Proceedings of the

XIV EUROPEAN ANNUAL CONFERENCE ON HUMAN DECISION MAKING AND MANUAL CONTROL

held at the
Delft University of Technology
Delft, The Netherlands
June 14 - 16, 1995

TUDelft

DELFT CONTROL

Proceedings of the

XIV EUROPEAN ANNUAL CONFERENCE ON HUMAN DECISION MAKING AND MANUAL CONTROL

held at the
Delft University of Technology
Delft, The Netherlands
June 14 - 16, 1995

Conference Chairman : Henk G. Stassen (Ed.)

Organisation : Peter A. Wieringa (Ed.)

Leo Brinkman

Mrs. Els Arkesteijn

Proceedings layout : Aad Gutteling
Man-Machine Systems Group
Laboratory for Measurement and Control
Faculty of Mechanical Engineering
and Marine Technology
Delft University of Technology
The Netherlands.

CIP-DATA KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Proceedings

Proceedings of the XIV European annual conference on human decision making and manual control : held at the Delft University of Technology Delft, The Netherlands, June 14-16, 1995 / Henk G. Stassen, Peter A. Wieringa (ed.). - Delft : Delft University of Technology, Faculty of Mechanical Engineering and Marine Technology
ISBN 90-370-0132-7
Subject headings: man-machine systems / human operators / process automation.

Publisher: Delft University of Technology
Faculty of Mechanical Engineering and Marine Technology,
Library WbMT
Mekelweg 2,
2628 CD Delft,
The Netherlands
Phone: + 31 15 2786765
Editors: H.G. Stassen, P.A. Wieringa.
Phone: + 31 15 2786400
e-mail: secrmr@wbmt.tudelft.nl.

FOREWORD

This volume contains the proceedings of the fourteenth European Annual Conference on Human Decision Making and Manual Control; for short the European Annual Manual or EAM. This time the EAM was held at the Delft University of Technology at Delft, The Netherlands, from July 14-16, 1995. This volume contains the full manuscripts of the papers as presented at the meeting.

In 1981 the first EAM was held. The idea was not new, at that moment the NASA -University Annual Conference on Manual Control, the Annual Manual, was already more than 15 years the yearly meeting for PhD students working in the field of Man-Machine Systems in the USA. The essence of the conference was to create an opportunity for young researchers to discuss their work in an early stage of their project. So the conference intends to be a platform to discuss in an informal way the projects rather than to present the results of an almost finished project. With this purpose in mind, it was appreciated to see that some of the leading staff in Man-Machine System research were also participating, and, in this way, stimulating and starting the discussions. The major difference between the conference in the USA and Europe is that the Annual Manual was just focussed on Manual Control, whereas the EAM also covers Supervisory Control research.

Since 1981 a series of conferences have been held in eight European countries, namely:

1981: The Netherlands, Delft, Delft University of Technology
1982: Federal Republic of Germany, Bonn, Forschungsinstitut für Anthropotechnik
1983: Denmark, Roskilde, Risø National Laboratories
1984: The Netherlands, Soesterberg, Institute for Perception TNO
1985: Federal Republic of Germany, Berlin, Technical University of Berlin
1986: United Kingdom, Wales, Cardiff, University of Wales, Institute of Science and Technology
1987: --
1988: France, Paris, Electricité de France
1989: Denmark, Lyngby, Technical University of Denmark
1990: Italy, Ispra, CEC Joint Research Centre
1991: Belgium, Lieges, University of Lieges
1992: France, Valenciennes, University of Valenciennes
1993: Germany, Kassel, University of Kassel
1994: Finland, Espoo, Technical Research Centre of Finland
1995: The Netherlands, Delft, Delft University of Technology

The intention is to continue the series; the next EAM is planned for 1996 in one of the southern countries of Europe.

Henk G. Stassen

Table of Contents

Session 1

Modelling Human Operator Behavior

Chairman: R. van Paassen, Germany

Simulation System for Behavior of an Operating Group (SYBORG)

-Development of an Individual Operator Behavior Model-

K.Takano*, K.Sasou & S.Yoshimura

Human Factors Research Center,
Central Research Institute of Electric Power Industry (CRIEPI)
2-11-1, Iwatokita, Komae-shi, Tokyo 201, Japan
takano@komae.denken.or.jp

Abstract

This paper describes the structure and mechanism of an individual operator behavior model developed, that is a major part of the Team behavior model, named SYBORG. This model has been developed on close contacts with experienced Nuclear Power Plant Operators, using the BWR training simulator. The results obtained from in-depth analysis of experimental chronology of operator's protocols and activities suggested that the existence of a mental model that becomes the kernel of operators' cognitive behavior in coping with an abnormal event. This mental model makes it possible for operators to help intentions formation of utterances and operations by means of advising possible causes, and further of envisioning the near-future scenario of plant dynamics. It was also suggested that both 12 kinds of knowledge categories in the long term memory and sensory information in the short term memory would be necessary to create the mental model in the working memory. Thus, a knowledge base, that was categorized into parameters KB, parameter parameters KB and event parameters KB etc., was prepared by experienced operators. For simulating operator's dynamic behaviors, it is also necessary to describe the operator's dynamic cognitive processes based on the created mental model used as a kernel. Therefore, the object modeling techniques (OMT) were adopted to translate the running processes of the mental model into computer programming. Simulated sequences of utterances and operations including thinking processes shown a reasonable agreement with events obtained in the experiment.

1.Introduction

In system that requires higher reliability and safety such as a nuclear power plant, it has been widely accepted that the basic principle for evaluating human-machine system is to consider that both humans (operators) and the machinery system (plant) should be regarded as a whole in evaluating the dynamic behavior (McRuer, Graham, Kredel and Reisener 1965). As a series of studies on the probabilistic risk analysis (PRA), it was implied that most important and difficult estimation was to identify human reliability and that the human reliability assessment must be necessary to estimate the reliability of an overall plant system. There are two ways how to evaluate system dynamics in which humans play a major role in control-loop: One is an approach to conduct experiments using plant simulator with cooperation of experienced operators, another is to develop a comprehensive simulator including both plant dynamics and human behavior. Because the experimental studies afford us only limited case-study results comparing to the resources required in the experiments, the Central Research Institute of Electric Power Industry (CRIEPI) has been making efforts to develop the Simulation System for the Behavior of an Operating Group, named SYBORG for several years (Sasou, et al 1993, Takano, et al 1994, Takano et al 1995).

---

*present address: c/o Professor J.T. Reason, Department of Psychology, University of Manchester Oxford Road, Manchester, M13 9PL UK. Email: takano@fs3.psy.man.ac.uk

The purposes of the applications of the SYBORG are:

1) To find possible paths to severe/large-scale accidents induced by a combination of machinery's functional failure and human inappropriate actions;
2) To design an adaptive interface for operators aiming at improving total reliabilities;
3) To develop an advanced intelligence operator support system on the way to the semi/full automatic control;
4) To propose more effective manners for upgrading cooperation within a team.

While many approaches, that are comparatively new originated at a few decades ago, are identified to simulate the human cognitive behavior, these are roughly classified into two categories (Cacciabue, 1994), these are micro cognition and macro cognition. Macro cognition modeling would be further divided into three categories, these are conceptual, deterministic and AI modeling.

Regarding micro cognitive modeling, Cacciabue (1994) mentioned that it is meant the detailed theoretical accounts of how cognition takes place in human mind. This modeling approach has been basically conducted in experimental psychology and medical research in brain mechanism. Basic and detailed theoretical modeling was studied individually such as on recognition, thinking and memories along the human information processing, however it would be difficult to connect these each other.

Macro modeling, in which human behavior in realistic tasks under actual working environments is tried to be explained and simulated (Cacciabue, 1994), made simplification on all cognitive process in order to describe a dynamic human behavior. The most simple and traditional theory was SR (Stimulus-Reaction) or SOR (Stimulus-Organization-Reaction) theory in the field of experimental psychology. Recently, in conceptual macro modeling, Card (1983) proposed a precursor model in which human information processing was divided into sensory processor, cognitive processor and motor processor. This model, that a required time consumption of each processor was determined based on experiments and on experiences, becomes a archetype at starting modeling. As next conceptual model, SL/SRK model, in which there are three level in human actions, skill, rule and knowledge, was proposed by Rasmussen (1983). This categorization also became a basis in understanding human behavior. As to the framework of the total information processing model, Baron's model (Baron, 1982) for nuclear power plant operator and Wickens model (Wickens, 1984) is nominated as important models because of introducing conceptions of the mental model and working memory, additionally the Wickens model suggested the existence of a limited attention resource. Each suggested the basic structure of human cognition and information processing mechanism interacting with external environments. Most recent conceptual model is COCOM (Contextual Control Model), proposed by Hollnagel (1993), implies that it is important to consider the effect of context, which should be influenced by external situations, for determining the control level of human actions. These conceptual and theoretical approaches cover a whole human mental activities, unfortunately it would be difficult to directly be extended to practical simulation models.

As to the deterministic models which was based on discrete simulation language such as SimScript and SAINT/micro SAINT, representative methodologies were the Human Operator Simulation, HOS, for radar monitoring tasks (Wherry, 1976), MAPPS for nuclear plant maintenance personnel (Siegel, Barttler, Wolf, Knee and Haars, 1984) and OPPS (Kozinskay, 1984) for nuclear plant operators. Huang (1991) developed a computer code using SimScript to simulate a team behavior in nuclear power plant based on plant simulator training experiments. This simulation suggested us the possibility of team behavior modeling. However, these models could only simulate the procedure based behaviors according to prescribed network, that is usually need a lot of time to describe by detail experimental observations. Knowledge is not explicitly defined but on the prescribed network as a task sequences, necessary time consumption and branch rule and furthermore these model cannot deal with unexpected knowledge base behavior.

At last, as the most recently developed model and principal approaches in simulating human cognitive behavior, the AI model would be referred. These approaches have been conducted in the field of nuclear power operators. The earliest approach in this area was conducted by Woods (1987), named Cognitive Environment Simulation, CES. This was the first attempt to realize the total human information processing in realistic environments by means of Artificial Intelligence, EAGOL/CADUCES for the use of medical diagnosis. The features are to divide cognitive behavior into three major activities, one is monitoring and analyzing the plant dynamics, second is building explanation for off-normal plant situation, and the last is managing the response for correcting abnormalities. Another feature is explicitly to define knowledge base that includes rules of identifying abnormalities, plant parameters vs. responses relationship, and parameters vs. parameters relationship. Cacciabue (1991) also developed the Cognitive Simulation Model, COSIMO, in which observed deviations in the plant transmitted to a blackboard as a cue and similarity matching is executed to identify abnormalities using fuzzy theory. If similarity matching is fail, frequency gambling would be applied to solve a conflict. Another model is the Cognitive and Action Modeling of an Erring Operator, CAMEO, developed by Fujita (1993). Common features to these models based on AI architecture is that basis is put on diagnosis, that is deduced from information about plant parameters acquired by monitoring, to identify what kind of event has happened. As results by diagnosis, it is possible to select a suitable set of operations. While these AI models are considered as up-to-date techniques, it is necessary to examine the similarities with the information processing process of real human operators because these has deficiencies in the respects of the contextual control and of behavior against unexpected/unfamiliar situations.

Reviews on recent approaches on human operator modeling suggested us following findings to be improved. Especially, because AI approaches has much possibilities and prosperity to simulate human cognitive behavior, it should be necessary to examine if it is true that the common mechanism emphasized on diagnosis would be robust in any situations observed at plant.

1) The previously proposed operator's model cannot deal with dynamic plant behavior, because researcher intended to make applications to static probabilistic risk assessment;.
2) Due to our observation on simulator experiments, diagnosis is only a part of human information processing strategies (Takano, et al 1994). Depending on the situations and context, the control level should be changed (Hollnagel, 1993);
3) Most of the previously developed simulation model only could deal with a few design based events because of less consideration with a universal mental model that real operators frame up in their mind based on their experience;
4) In spite of the importance of team behavior, every model excluded communication and group dynamics excepting Huang' model (1991).

Based on above consideration, this individual operator behavior model developed satisfies following necessities in order to simulate more realistically and universally. The development of the present individual model shall makes it possible to construct the team behavior model: SYBORG by connecting each individual models with HHI and also connecting them via HMI with "plant model".

1) To connect the operator model to dynamic plant simulator;
2) To adopt the universal mental model that was obtained by in depth analysis of simulator experiments with cooperation of experienced operators (Takano, Sasou and Yoshimura, 1994,1995);
3) To have a function of communicating with the other operators;
4) To have a function of action and movement for taking the available time into consideration.

2. Structure and functions of an individual operator model

2. 1 Structure

Because the final aim of this studies is to develop the team behavior model: SYBORG, this individual model should have functions in order to be able to behave as an individual in a team. Thus, in designing the structure of an individual model, interactions with human as well as with external environments should be taken into consideration. This concept, of course, leads to the necessities of model's capabilities of verbal communications and even necessities of introduction of the group dynamics as interactions between individual models. Our envisaged block diagram is shown in Fig.1, which consist of plant model (dynamic simulator), Human-Machine Interface model (HMI:

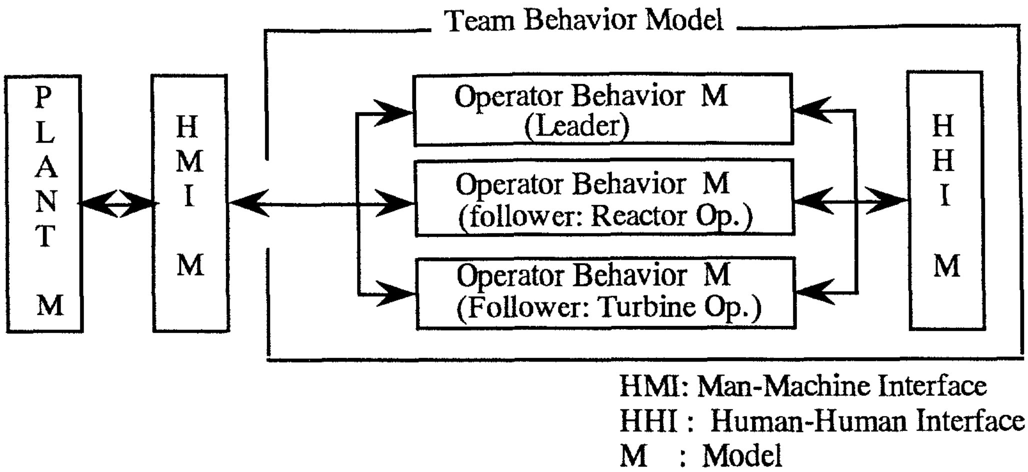
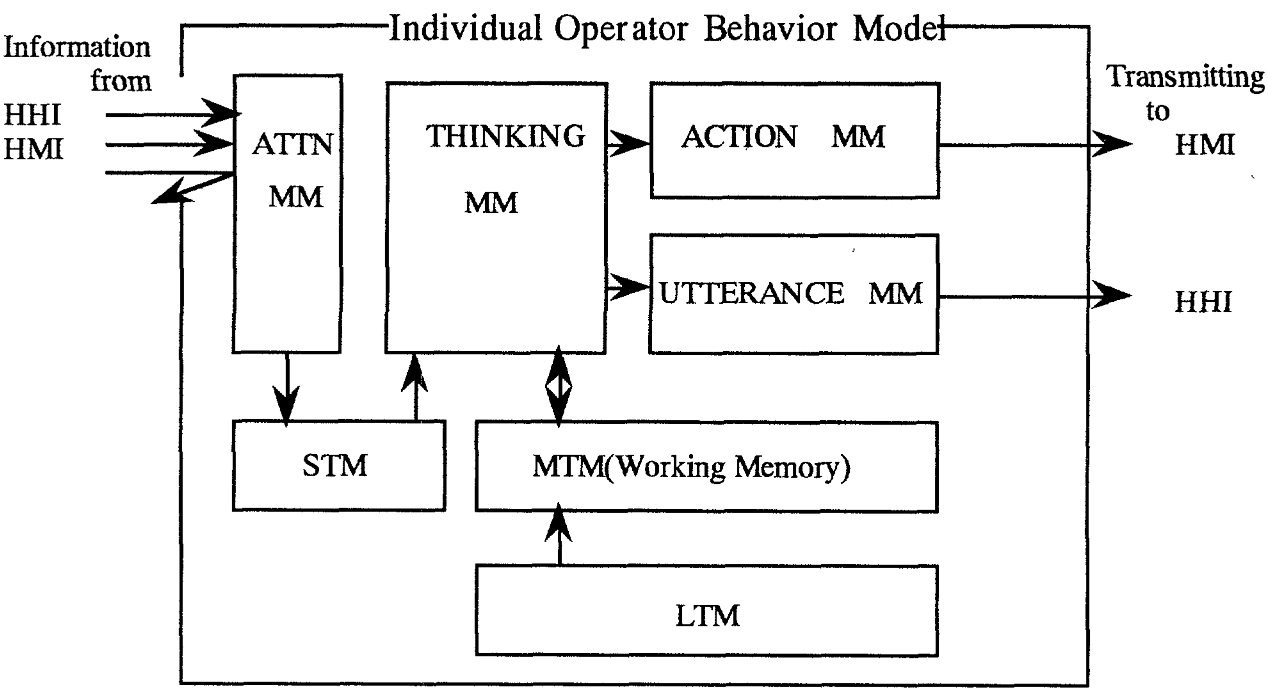


Fig. 1 Block diagram of the total system named SYBORG: Simulation System for Behavior of the Operating Group

control panel), three individual operator models and Human-Human Interface model (HHI). Here, the HHI plays a role of handling communications and group dynamics between each operator. In this paper, so we do not intend to mention details about the HHI as to treat it as a black box with defining explicit functions. Details were explained in the other author's paper (Sasou, et al 1994). The HHI model includes functions of the role assignment within a team, the management of communications, the disagreement solution between members.

Both on the above mentioned speculations and on conceptual models proposed by Wickens (1984) and Baron (1982), the block diagram presenting the structure of an individual model has driven as shown in Fig.2, which consisted of an attention micro model, a short term memory: STM, a thinking micro model, a working memory (mediate term memory: MTM), a long term memory: LTM, an action micro model and an utterance micro model.



Notes: HMI:Human-Machine Interface, HHI:Human-Human Interface
MM:Micro Model, ATTN:Attention, STM:Short Term Memory
MTM: Mediate Term Memory, LTM:Long Term Memory

Fig. 2 Block diagram of an Individual Operator Behavior Model and Information flow

2.2 Functions

The functions of each substitutes are described below including the fundamental consideration for getting the detailed design of this individual operator behavior model.

2.2.1 Attention Micro Model

This part is for introducing information inputted into the individual operator model. In the present work, two kinds of information were defined: one is information on HMI through visual organs (Panel Information: annunciator indications, indicator readings, and status of switches and ramps), second is information on utterance from other operators through auditory organs.(Utterance Information). This micro model executes screening whether information transmitted from external environments could successfully inputted into the model or not. The screening criteria was already reported in the previous paper (Takano, 1994). Criteria for utterance information were defined by the volume of utterance and sender's position and that for Panel Information is the kinds of indicator (each indicator has an affordance level for attracting attention), scope of visual field and operator's arousal and workload varying according to the circumstance. Actual screening can be realized using neural networks of back-propagation type calculation.

2.2.2 Short Term Memory

Information through the attention micro model is temporary accumulated here in order to smoothly transmit information between the attention micro model and thinking micro model, for instance, in cases when plural information would be inputted simultaneously or when the other part of the model would be working at the time. Accumulated information would be decayed with the time constant $T_{1/2}$=10 seconds and STM has a limited capacity up to seven records. These value are really preliminary and should be modified with reference to experimental studies.

2.2.3 Thinking Micro Model

This micro model is the most important part in the whole model. Through situation understanding and causal reasoning based on both information introduced through the attention micro model and knowledge base in LTM, this leads to the determination of how to encounter the situation for making plant status steady and finally controls executing procedure. Main frame concerning how to proceed thinking is closely related to the running process of the Mental Model explained simply below. The concept and creating process has already proposed by authors (Takano,et.al 1994 and 1995) including how to construct the mental model using Knowledge base in LTM. To recapitulate the concept and basic principle of this mental model:

1) An annunciator indication makes the operator aware of a deviation of the plant operating condition from normal range —— i.e. the occurrence of an abnormal event.
2) If a simple countermeasure occurs to the operator's mind, the countermeasure is applied reflexively by the operator as "Immediate Reaction" (e.g. starting up the standby pump in the case of a pump trip).
3) After the Immediate Reaction, he proceeds to envisioning the near-future scenario of plant evolution, and at the same time to seeking the cause of the abnormal event. The near-future scenario serves the operator in choosing the Key Parameter (KP) —— that which he considers to exert the strongest influence on the plant operating condition affected by the particular abnormal event, and to which he will concentrate his attention in monitoring the evolution of plant behavior.
4) The operator adopts Emergency Countermeasures to seek preventing the Key Parameter from further deterioration —— or if possible, to restore steady condition. The specific Emergency Countermeasure to be adopted would depend on the extent of deviation from normal range marked by the Key Parameter.
5) Upon identifying the cause of annunciator indication, the operator seeks and applies a Causal Remedy with the aim of eliminating the cause and restoring normal plant operating status.

The foregoing structure of a universal mental model is schematized in Fig. 3. The occurrence of an event, which would be noticed as a first annunciator indication to operators, imparts its effect on one

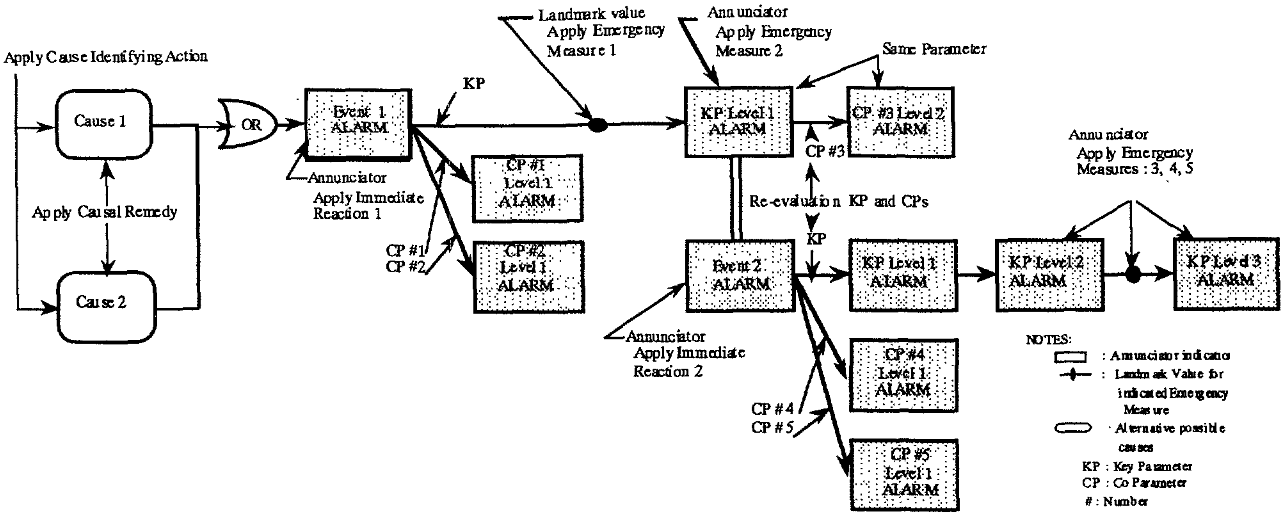


Fig.3 Pattern of universal mental model formed in an experienced operator's mind upon noting the first occurring annunciator indication

or more plant operating parameters. Such consequently affected operating parameters are termed Co-Parameters. Among the Co-parameters, the operator will choose a Key Parameter, which he considers to influence the most sensitively the plant operating condition affected by the particular abnormal event, and to which he will concentrate his attention in monitoring the evolution of plant behavior. The operator will next envision the near-future evolution scenario of the Key Parameter. The scenario would also incorporate the time factor and include foreseen annunciator indications and interlock actions. Considering this time factor, the operator will determine for the Key Parameter a Landmark Value at which he will apply Emergency Measures. Interlock actions —— set off by logical circuits to cause component tripping or reactor scramming —— will also influence one or more plant operating parameters.

The mental model, that is to be created in Working Memory (Mediate Term Memory: MTM) according to the situation, denotes the envision of near-future scenario of plant dynamic evolution and also points out the list of possible causes which could cause the present situation. The merits of this mental model proposed are (1) even applicable to events that have not been envisaged yet and (2) enable to combine dynamically elements of knowledge, thinking processes and behaviors. By means of this mental model, thinking process can be regulated and intentions of action and utterances can be also formed. Intention formation process will be mentioned hereafter.

2.2.4 Working Memory (Mediate Term Memory: MTM)

This is a part where information acquired through sensory organs and information retrieved from LTM is arranged in order to make operator's intentions formation. The Working Memory consists of largely 2 parts illustrated in Fig.4, one is a capacity for mental model and another is for pages. That page, which is created corresponding to each plant parameters defined in the mental model, stores the records of evolution of plant parameters, annunciator indications, utterances by other operators, and operation performed. The combination of the mental model and corresponding pages is defined as an Index. Thus, the single index would be formed in MTM when an encountering event involves single cause, however, plural indexes would be formed when the event induced by plural causes simultaneously. The mental model in MTM can be revised if important information would be introduced here.

2.2.5 Long Term Memory (LTM)

Authors already proposed the structure and contents of knowledge base involved in the Long Term Memory (Takano, Saso, Yoshimura 1995). In that paper, knowledge base contains 13 elements including plant static configurations, plant parameters specifications, plant dynamics and three-leveled countermeasures' procedure etc. Also LTM contains the HMI location data (x, y and z coordination

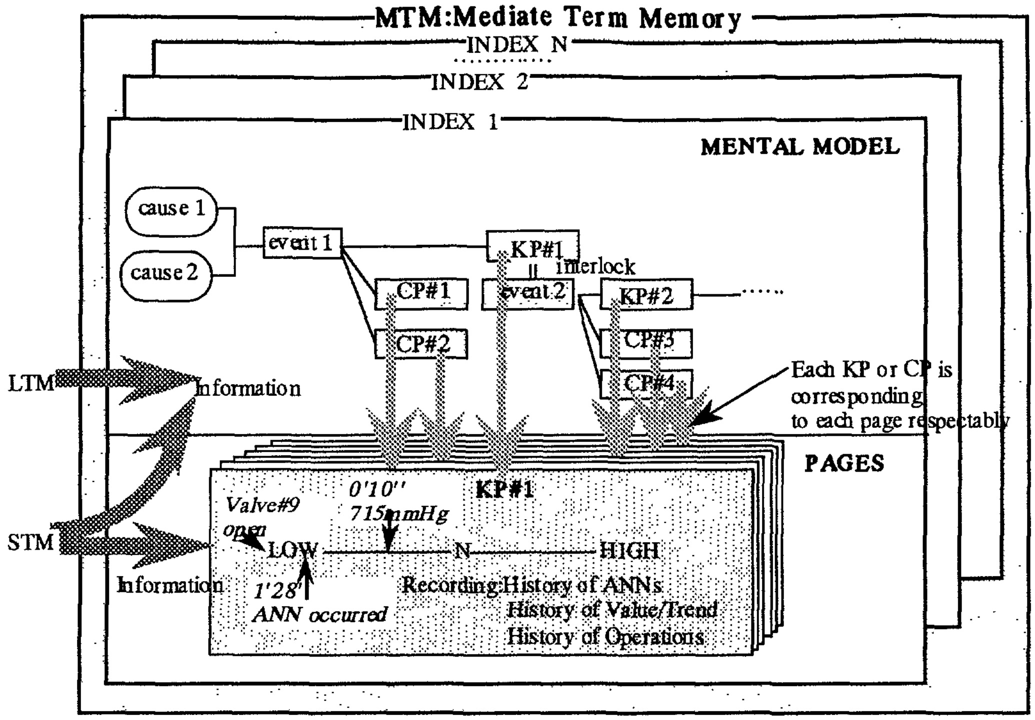


Fig.4 Schematic diagram of the mediate term memory involving the mental model envisaged and pages recorded

2.2.5 Long Term Memory (LTM)

Authors already proposed the structure and contents of knowledge base involved in the Long Term Memory (Takano, Saso, Yoshimura 1995) illustrated in Fig.5. In the Figure, the structure of the LTM is drawn as a union of the elements of KB, and each KB is corresponding to the substance of the object model expression of mental model that was shown in Fig.3, yet the object model originated

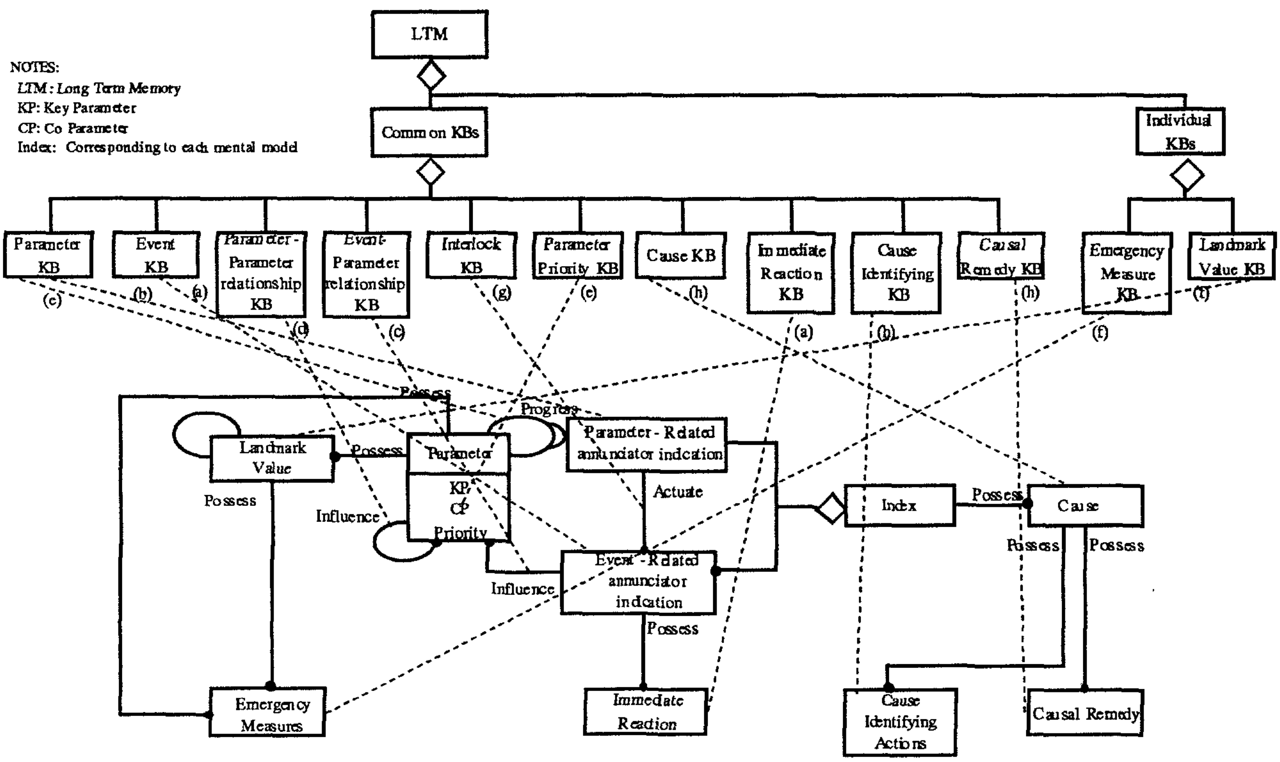


Fig.5 Structure of the operator's mental model and interrelationships existing between individual elements of knowledge base (KB) and the object model expression of the mental model

Table 1 Substances of individual elements of knowledge base involved in the formation of a mental model in the operator's mind

Knowledge Base (KB)	Substance of KB
Common KB	KBs commonly shared by all operators
Individual KB	KBs differing from operator to operator
Parameter KB	Details of each individual parameter, including annunciator settings and normal operating ranges
Event KB	All events that are subject to annunciator indication
Parameter - Parameter Relationship KB	Principal co-parameters influenced by a parameter that exceeds prescribed threshold
Event - Parameter Relationship KB	Principal parameters influenced by the occurrence of an event
Interlock KB	Event brought about by interlock action of logical circuit (component trip, reactor trip)
Landmark Value KB	Value of Key Parameter set by operator for applying Emergency Measure
Parameter priority KB	Graduated list of parameters, for selecting the Key Parameter (that which influences most sensitively the plant operating condition)
Cause KB	List of possible causes to be considered by operator upon noting annunciator indication
Immediate Reaction KB	Allopathic remedial measure to be envisaged by operator, based on his experience, immediately upon noting an annunciator indication
Emergency Measure KB	Measures to be envisaged by operator in order to prevent further deterioration of Key Parameter
Cause Identifying KB	Procedure for determining the cause as it was, by investigating relevant information or field observations
Causal Remedy KB	Procedure for eliminating the cause, or its effect on plant operating condition

at the Object Modeling Techniques will be explained later. As described in Table 1, the LTM involves 13 elements of knowledge base (KB); these are specifications of plant parameters (parameter KB and interlock KB), plant dynamics related (event KB, parameter-parameter relationship KB, event-parameters relationship KB and interlock KB) and three-leveled countermeasures' procedure etc. Beside this structure, LTM also contains the HMI location data (x, y and z coordination is represented for annunciators, indicators and operation knobs/lumps), available menu of actions and utterances.

2.2.6 Action Micro Model

Intentions formed in the thinking micro model can be classified into two categories, one is for actions (action intention), another is for utterances (utterance intention). This Action Micro Model realizes the action intention into execution. Actions involves, of course, a procedural operation and further walking and body movements that are reaching hand, reading indicators and switching on/off etc. It is also possible to calculate the standard time required for actions by using prescribed time table for various motions based on the modified Method Time Measurement (Nagasaka, 1994). Then, this estimation makes it enable to evaluate the workload due to operator's actions. At last, actions also includes "reading indicators" and "hearing utterances from others" in the simulation.

2.2.7 Utterance Micro Model

The utterance micro model have realized the intention formed by thinking micro model. The kinds of utterance is classified into 12 categories and defined based on the experiments that was conducted by authors using full-scope nuclear power plant simulator (Sasou, Nagasaka & Yukimachi, 1993). For each utterance category, some typical sentences were defined and stored in utterance data base. It is also possible to calculate the required time for speaking sentence retrieved from database.

3. Design of Simulation Model for an Individual Operator

3.1 Design Tool for Programming a Dynamic System

Flowchart as the traditional designing tool is not fully suitable for describing complicate time dependency between processes such as human information processing. Because both suspending information processing on the way of thinking and parallel processing will be anticipated in the simulation, it is difficult to validate whether developed flowcharts could satisfactory describe the dynamic functions involved and could represent contextual relationships or not.

Therefore, authors adopted the Object Modeling Techniques (OMT) for describing this dynamic simulation code. This method utilizes both a procedural hierarchy and a data hierarchy synthesized by introducing a conception of a class, as were used to be treated separately. Applying the OMT, three type of drawing should be prepared: one is the Object Model representing the static structure of a whole model describing interrelationship with among defined classes, second is Dynamic Model representing dynamic behavior in each class including interactions between classes using states transition chart and the last is the Functional Model representing data transfer and data processing within and between classes. As the OMT offers above mentioned three kinds of visual drawing results, not only understanding and modifying those but also transferring into program is comparably easy.

3.2 Object Model

As shown in Fig. 1, SYBORG contains three operators. To draw this object model for an individual operator, interface and interaction with HMI and HHI should be taken into consideration. Thus, instead of designing only a single operator, it had better to design a whole simulation model for SYBORG so as to grasp the whole frame to be developed even limiting to determine necessary classes in an individual operator model. In order to develop the object model, each model and micro models shown in Fig.2 were re-evaluated and divided into classes so that each class could have a simple function for realizing the running process of proposed mental model. Figure 6 shows developed object model, in which each rectangle indicates "class". In Table 2, brief explanation was

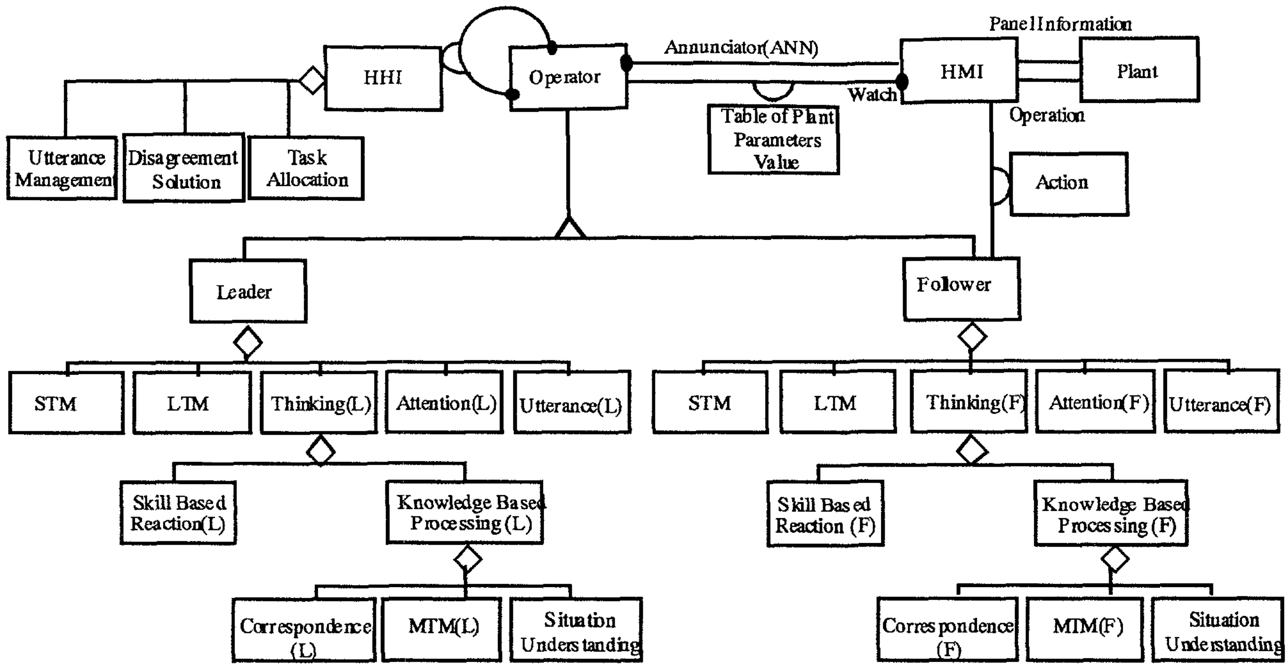


Fig.6 Described object model of overall structure of the SYBORG based on Object Modeling Techniques (OMT)

described for each class appeared in the object model. The class "operator" corresponds to operator team simulation model including three operators: a leader and two followers. Around operator class "HHI" and "HMI" is defined as a interface and connected to operator with a line representing "relationship". The class "plant" indicates the plant simulator connected to the HMI for displaying simulation results and reflecting operator's actions via a table of updating plant parameters indications and operation outputs. The HHI consists of three classes, that is really in operators or in among operators, those are the Utterance management, the Disagreement solution and Task allocation. These classes were defined for regulating team behavior, so called the group dynamics. As to the detail contents of these, Sasou, et.al (1994) mentioned in the other report, therefore herein we denote only functional aspect which closely relates to designing an individual operator model. The class operator is a "super-class" consisted of a leader "sub-class" and two follower "sub-classes"; one is a turbine operator and the other is a reactor operator. "Part of class" belonging to a class leader is the same as those of follower excepting class "action". Parts of class belonging to individuals are the Short Term Memory, Long Term Memory, Thinking, Attention, Utterance and Action. However, because the same label attached to a class does not always mean the same contents, classification is made between leader and follower with (L) for leader and (F) for follower. Each class is corresponding to already defined micro model in Fig. 2, for example a class "thinking" is coping with thinking micro model etc. Out of further analysis of the simulator experiments, the class "thinking" was divided into two "part of class"; one is Skill based reaction and another is Knowledge based processing. This skill based reaction plays a part in controlling prescribed cataloged behaviors such as distributing the inputted utterances to suitable class required or instantaneous reactions i.e.:. walking in front of indicator and reading it when annunciator flickering and making a reply to talker. The knowledge based processing, which is a running mechanism of author's proposing mental model, is further divided into three "parts of class". These are "Situation understanding", "Working memory (Mediate term memory: MTM)" and "Correspondence". Context made in the knowledge based processing is as following: (1) Creation of the mental model by information inputted and information retrieved from LTM; (2) Building a strategic goal based on the mental model; (3) Envisaging operation depending on the Key Parameter. Dynamic behavior of these are described later.

Table 2 Brief Explanation of each class defined in the object model shown in Fig.6

Name of class	Brief Explanation and Definition of the Class
Plant	Calculating plant dynamics including transient caused by malfunctions
MMI	On the board, displaying each value of plant parameters according to the calculation made by the plant and receiving/sending operators' manipulations
Operator	Super class defined as a union consists of one leader and two followers
HHI	Managing the human - human interactions between a leader and followers
Utterance Management	Indicating the person to be transmitted the utterance and calculating the loudness of the utterance made by an operator
Disagreement Solution	Determining/Coordinating each operator's opinion of correspondence if it being different between a leader and a follower in charge of the correspondence
Task Allocation	Distributing tasks to be done between operators by taking task priority, individual roles and the number of tasks to correspond to
Leader	The operator to supervise comprehensive correspondence and to work the HHI suitably
Follower	The operator to make actions based on his own basis from observations or decision made by the HHI
Action	Realizing an activated intention of making an action or a series of actions formed by his own thinking class to be implemented
STM	Temporarily storing information received via attention class
LTM	Necessary knowledge to create the mental model shown in Fig.3 including 13 categories of knowledge bases shown in Fig.5
Attention	Receiving auditory and visual information, including utterances, annunciator indications, panel indicators and switches/lamps within limited attention capacity varying according to workload and arousal level
Thinking	Forming intentions of utterances and actions for dynamic behavior
Utterance	Realizing an activated intention of making an utterance formed by his own thinking class to be spoken
Skill Based Reaction	Monitoring whether there is information in STM or not. If there is information, classifying it into defined categories and transmitting it to proper class and forming intentions of squared and prescribed reactions like as a confirmation of ANN occurred or Indicator ordered
Knowledge Based Processing	Creating a mental model in the MTM and forming intentions of utterances and actions according to the information processing route described in dynamic models (see Section 3.3) which was based on the mental model
MTM	As a working memory with the structure shown in Fig.4, noting the mental model envisaged and information relating to plant dynamics introduced
Situation Understanding	Understanding the situation by framing up the mental model that enable envisioning/reasoning plant dynamics, and managing the MTM refreshments, and defining the task priority using conceptions of Index and Key-parameter
Correspondence	Finding out suitable countermeasures and controlling a procedure until any effects will be seen

3.3 Dynamic Model

According to the instruction of the OMT, dynamic models should be developed for each ultimate underlying classes in the object model shown in Fig.6. The classes to be drawn the dynamic model are "Attention", "Skill based reaction", "Situation understanding", "Correspondence", "Utterance", "Action", "Short term memory" and "MTM". As to LTM, so it is static as not necessary to make a dynamic model. Dynamics and information transmission inside these class and between classes is not independent rather than closely related each other. Then, it is necessary to investigate the interaction and information flow between classes, even inside a class for a specific class like as "Correspondence". By revealing these interactions, dynamic models of each class could be detailed. Therefore, at first, overall information flow will be revealed. Following it, because information flow relating to the knowledge based processing are somewhat complicate, slightly detailed diagram will be shown by closing up inside a "Thinking class".

3.3.1 Overall information flow of SYBORG

At first stage of this designing, it is the most important to assign the roles of each class. To assign the roles and functions, the relationships and information flow between each class that builds up the individual operator behavior model. Figure 7 shows the overall information flow diagram including the other activities between each class defined in Fig.6. In the SYBORG, we assumed that there are no interactions between two followers: they are a reactor operator and turbine operator, for simplicity. In facts, the simulator experiments conducted by authors implied less interactions between followers than those between a leader and followers. As shown in Fig.7, only interactions between a leader and follower could be seen because of the above reason. This figure shows that all information introduced into an individual behavior model is via "Attention" and prompt and squared reaction will be made by "Skill based reaction (SBR)" which includes such actions as making a reply, distributing information introduced into the proper class that requested it, reading the indicator on request and also reading the annunciator indication activated etc. Following this, information that has possibilities to change plant situations will transfer to "Knowledge based processing (KBP)" in order to frame up and refresh the mental model related, and in succession, the Key-parameter to be concentrated should be chosen to notice it to "Task assignment". Information to be transmitted to the task assignment also accompanies the index and its priority because it is possible to define plural indexes of which a single mental model could be existed. Plural indexes mean the case that there occur multi-malfunctions in the plant. Priority means the highest emergency to be corresponded among indexes created (see Fig.4). The mechanism of the task assignment is involved in HHI, by which each operator would be noticed

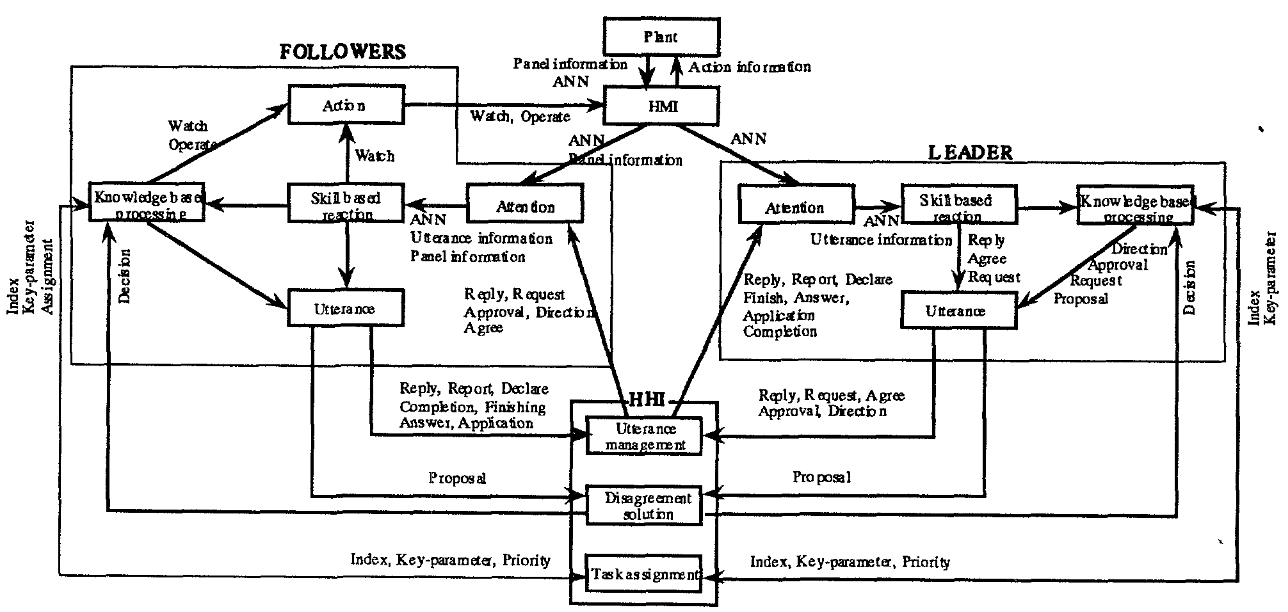


Fig.7 Diagram of the overall information flow between classes defined in SYBORG

the index to be implemented, the assignment rules is depending on the group characteristic, prescribed roles of each operator, the number of indexes created and its priorities. In this paper, the simplest pattern of task assignment was adopted for further studies. Therefore, SYBORG would correspond to the index of the highest priority.

Utterances are originated from both the skill based reaction (SBR) and the knowledge based processing (KBP), as usual, every utterance excepting a proposal shall be transmitted to the "Utterance management" to determine volume and the person to be transmitted. Regarding to proposal, it should be transferred to the "Disagreement solution" to determine suitable and coordinated correspondence among group. Definitions and contents of terms appeared in Fig.7 are denoted in Table 3 below. Because, in Fig.7, the class knowledge based processing was presented as a aggregation class for simplicity, further explanation will be made in following session concerning sub classes of knowledge based processing.

Here, we make some example of information flow for easy understanding; once ANN has occurred, it is to be transmitted via HMI to Attention for both leader and followers. The follower in charge of confirming ANN will form an intention to make an action in SBR. The intention will be realized by the Action to read the content of ANN. The content of ANN re-entered into Attention and transmit via SBR into KBP to form the mental model. After framing up the mental model in both leader and follower, correspondence will be continued by KBP based on the mental model

Table 3 Definition and contents of terms appeared in Fig.7

TERM	CONTENTS AND DEFINITION
Panel information	Status and reading of indicators, controls and lamps displayed on the HMI board and readings of ANN signboards
ANN	A flickered annunciator indication
Operate	Making an action to switches and controller on the HMI board
Watch	Reading indicators on the HMI board
Declare	Announcement of applying an immediate reaction to seek approval from the leader
Agree	Approval to the follower to the declare above
Completion	Announcement of the completion of the immediate reaction
Direction	Announcement to indicate implementation of each step of an Emergency measure procedure from Leader to Follower
Application	Announcement of a follower to seek approval from leader to implement a step of an Emergency measure procedure
Approval	Approval from leader to the application made by follower
Finish	Announcement of finishing a step of an Emergency measure procedure
Request	Asking a suitable follower to watch and tell me the indication reading
Answer	Answer against the request
Report	Announcement to notice indication reading and ANN occurrence
Proposal	Announcement of proposing the most suitable Emergency measure to seek decision from the disagreement solution in HHI
Reply	Announcement against every utterance heard or accepted
Decision	Assumed announcement originated by the disagreement solution as a result of being coordinated among the leader and the proposer to notice him the decision against the proposal made by the follower

3.3.2 Information flow inside a knowledge based processing

Overall information flow has been explained, however, it is not still clear how to run the knowledge based processing. Thus, we denote further about the information flow inside a class "Knowledge based processing". Figure 8 shows the processes after the skill based reaction; an important information, like as the panel data (ANN), the report from the other operator, or the answer against a query will be introduced to subclass "Situation understanding" to frame up or to revise the mental model in MTM. If the situation recognized previously has changed, a new or revised index, which

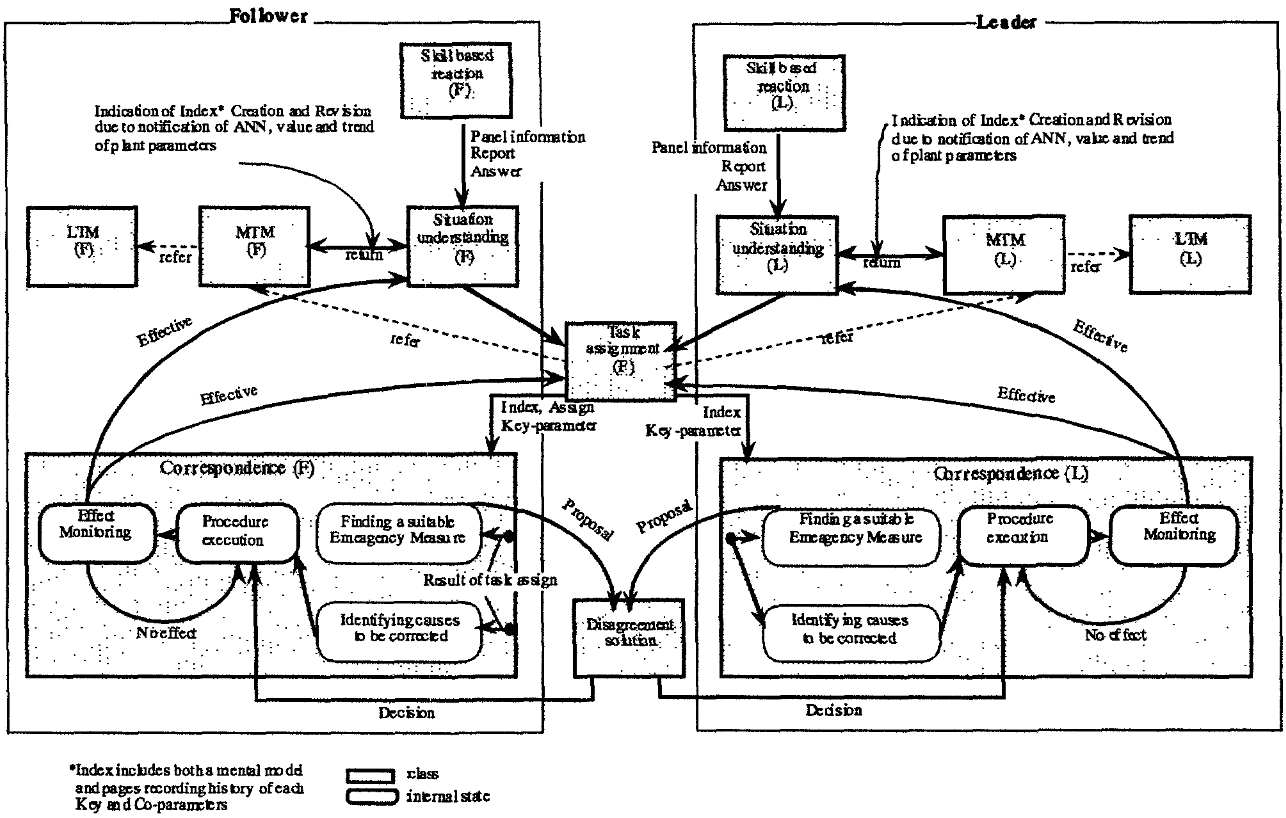


Fig.8 Schematic diagram of information flow inside the Knowledge Based Processing

includes both the mental model and pages, is transferred to "Task assignment" with index priority. In the task assignment, it is determined who is suitable to make a correspondence with it according to each operator's role sharing, busyness and task's priority. The result of assignment is noticed to the class "Correspondence" of a person who treats the subject directory and a leader at the same time. In the correspondence, as a first step, an emergency measure or cause identifying measure is to be proposed after necessary confirmations of the plant situations to the "Disagreement solution". Proposal made will be discussed in the disagreement solution by seeing whether proposal made by the leader was same as or not. Decision, which will be made in taking considerations of the group dynamics, is also noticed to the procedure implementation part. After implementation, monitoring should be continued to see whether it is effective or not. If no effect, bringing back to the previous part to find out an alternative measure, or if it was effective, the fact is transferred to the situation understanding and task assignment in order to delete the index.

3.3.3 An example of dynamic model "Situation understanding"

Dynamic models should be developed for all classes defined in Fig.6 excepting LTM. Because LTM has no dynamic part so that it is Read Only Memory (ROM). This dynamic model expresses the running process of the mental model including when and how to make actions and utterances. Of course, this model should describe the cognitive processes obtained as the results of simulator experimental protocol. Dynamic model, as shown in Fig.9, contains finite states inside, and it denotes transitions between states, branch criteria and contents of treatment under rules of the object

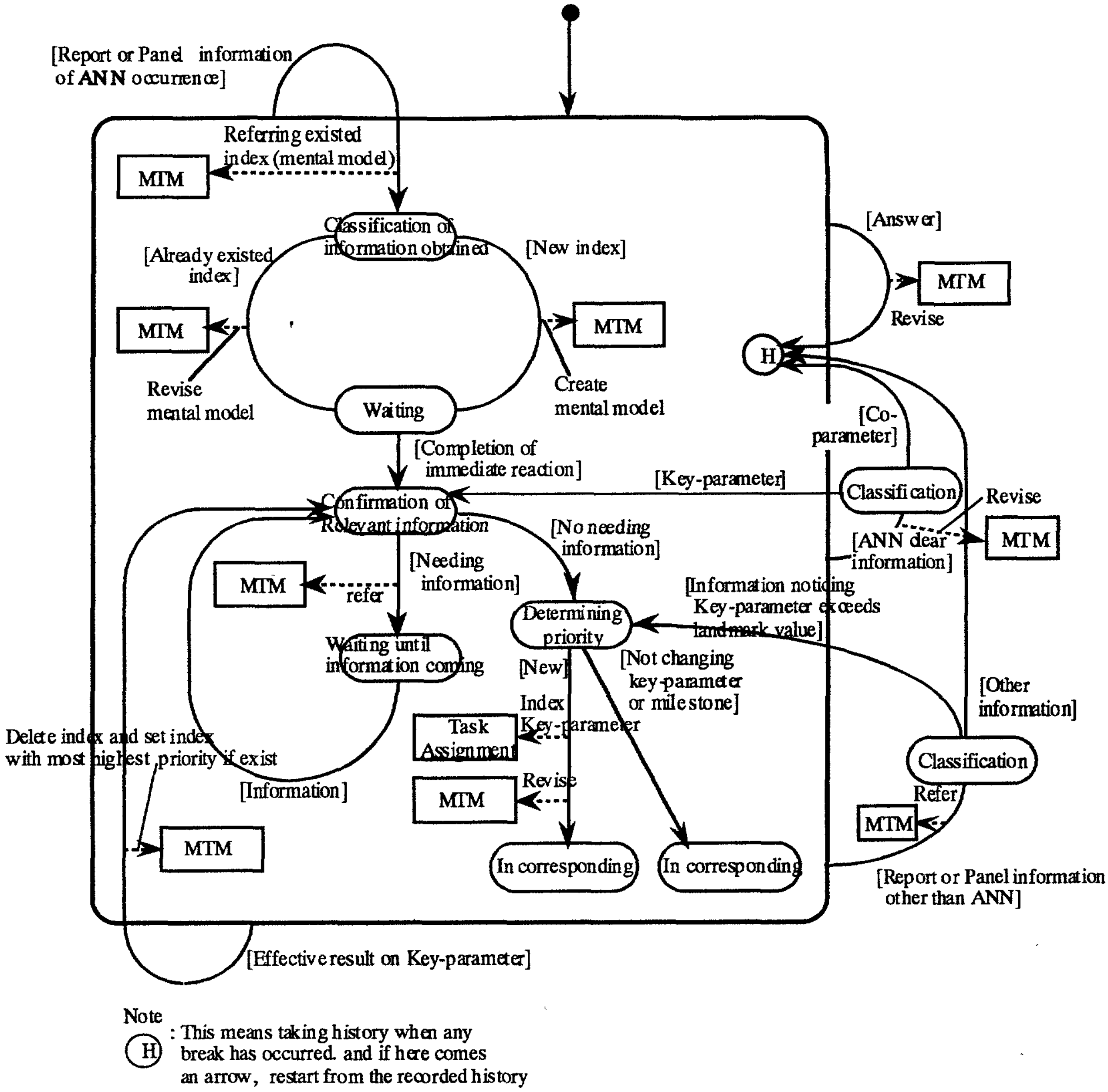


Fig.9 The dynamic model of the "Situation understanding" based on the object modeling techniques, (OMT)

techniques, that has proposed by Rumbaugh et.al (1991). In this paper, only a dynamic model of the class "Situation understanding" is to be shown in convenience. It is common for leader and followers. The processing route is as following:

1) After received ANN information, if it is a quite new ANN in no relation with the existed key parameter, a new index including a mental model is created in MTM. However If it is the ANN related to the already existed key parameter, revision of the mental model will be made.
2) Confirming the end of immediate reaction, observation of relevant parameters is conducted by himself or is obtained as an answer to his query. Those observation results shall be stored as historical information into pages defined in MTM.
3) The key-parameter shall be chosen among parameters in the mental model by taking importance / emergence into considerations. Importance and Emergence is regarded as a priority.
4) Information concerning an index which has the key-parameter of the highest priority is transferred to "Task assignment" in HHI. Assigned jobs per each operator is noticed to "Correspondence" of each operator respectably.
5) Each operator executes necessary cause identification procedures or emergency measures after

observing preliminary plant condition. Cause identification and emergency measures can be progressed in parallel.

6) Continuously monitoring the trend of the key-parameter, if operator identifies an effect anticipated, the effectiveness will be sent to MTM and task assignment to finish the job. If not effective, operator tries to find alternative measures and makes execution.

Above mentioned is an example of dynamic models already developed, each dynamic model should be developed for each class respectably to simulate operator's dynamic behavior even including mental processes.

3.4 Functional model

Above mentioned object model determined overall structure of our simulation model and clarified functional boundaries between each class. Next Dynamic model developed could describe the dynamic behavior of each class and could reveal the running process of the mental model that had been obtained in depth analysis of simulator experiment. However, there is indeed another important portion, that is how to create the mental model from knowledge base stored in LTM. The structure about knowledge base was already presented in author's previous paper in detail (Takano, et.al 1995). Thus, in this paper, a brief but essential explanation will be made.

Describing the information flow and concrete data processing within or between MTM and LTM could be realized by making a "Functional model". It can be also reveal the creating process of the mental model using knowledge base in LTM. Figure 5 shows the generalized format of the mental model described by OMT in relation with elements of knowledge base in LTM. We could identify 13 elements of knowledge base (KB) in LTM, that is explained in Table 1, in order to create the mental model. The mental model created in MTM concerning to a specified event, that is the "Condenser Hot Well water level LOW", is shown in Fig.10. And concrete description of the substances of major four representative elements of KB is also shown in Table 4. Detail descriptions of substances of KB could be obtained by close discussion with experienced operators.

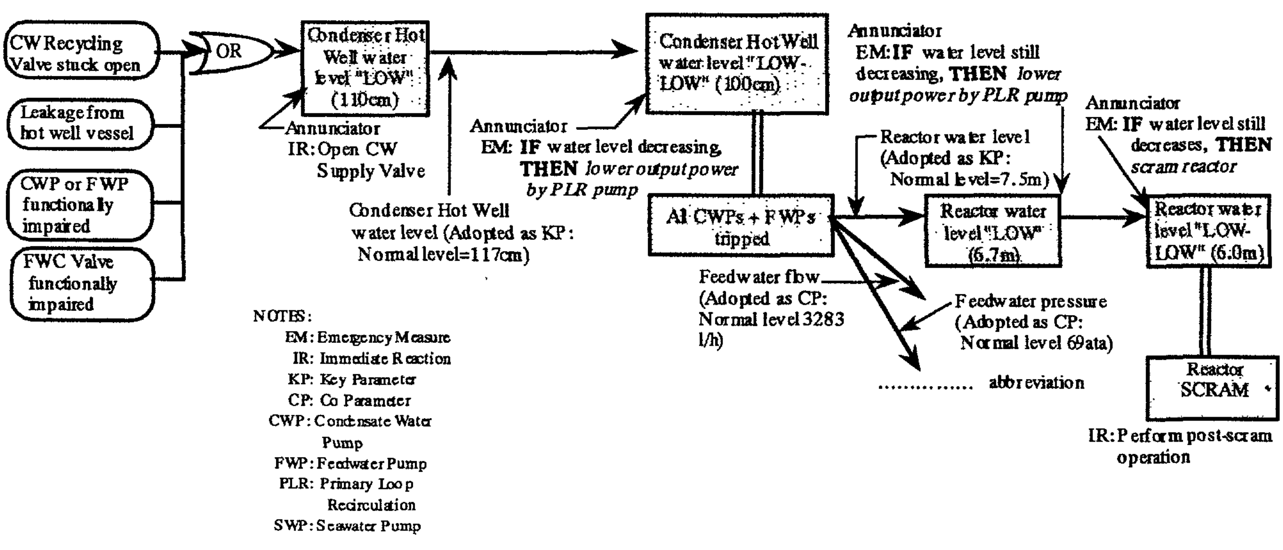


Fig.10 Mental model formed in operator's mind upon noting annunciator indication of "Condenser Hot Well water level LOW"

Necessary components to create the mental model shown in Fig.10 are both all elements of KB described in Table 1 and creation procedure itself. Then, we made the procedure to create the mental model of possible malfunctions occurred in the plant by the Functional model of OMT.

The creating process of the mental model is stored in MTM, then, every kinds of the mental model due to the malfunction could be formed instantaneously when the first ANN has occurred in the plant only by interacting with LTM. Creation process is shown in Fig.11, which was described under rule

Table 4 Substances of four representative elements of knowledge base relating parameter involved in the simplified simulator shown in Fig.1

Parameter KB						Parameter - Parameter RelationshipKB	Interlock KB	Cause KB	
	Annunciator setting								
Parameter	L-L	LOW	N	HIGH	H-H	Co-parameters		Excessive decrease	Excessive rise
SWP suction flow	NA	90%	100%	110%	NA	1.Condenser vacuum ↓ 2.Generator output ↓	None	1.SWP functionally impaired 2.SWP tripped 3.SWP entrance plugged	1.SWP over-speeding
Condenser hot well water level	100cm	110cm	117cm	125cm	NA	None	1.CWP/FEP tripped (at LOW-LOW)	1.CW recycling valve stuck open 2.Feedwater leaking from supply system 3.Malfunctioning CWP/FWP 4.Malfunctioning FWC valve	1.CW supply valve stuck open 2.Malfunctioning CWP/FWP 3.Malfunctioning FWC valve
Condenser vacuum level	660hPa	715hPa	723hPa	NA	NA	1.Generator output ↓	1.Reactor scrammed (at LOW-LOW)	1.Low SWP suction flow 2.High seawater temperature 3.Low exhaust gas flow	NA

Notes: CWP: Condensate water pump; FWC: Feedwater control; FWP: Feedwater pump; L-L: LOW-LOW; H-H: HIGH-HIGH; KB: Knowledge base; N: Normal; NA: Not applicable; SWP: Seawater pump of OMT. This model has indicated every processes from initial intake of the first indication of ANN until final creation of the specified mental model even up to the revision of it in relation with each element of KB.

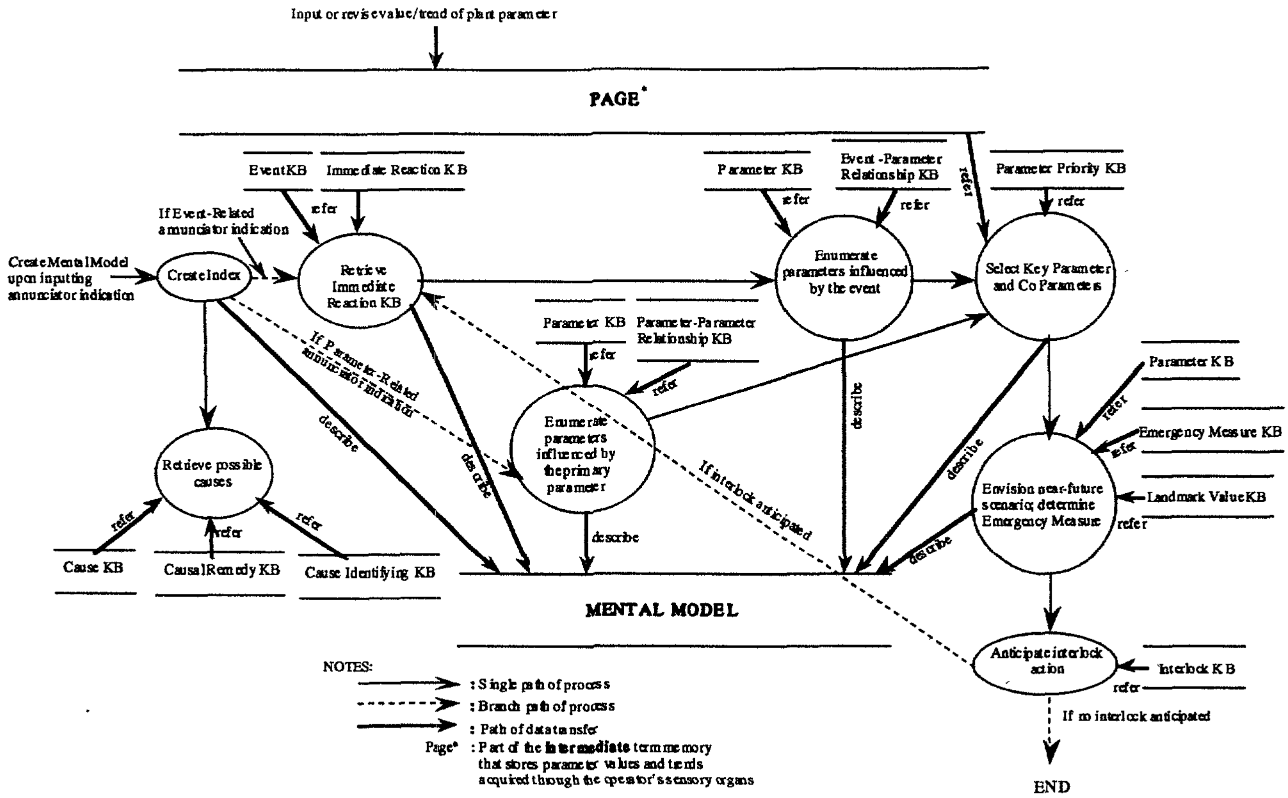


Fig.11 Functional model describing the process creating the specified mental model using each element of KB stored in LTM

4. Simulation results and comparison with experiments

4.1 Development of the simulation code

In this section, programming manner and accompanying condition is described according to the above mentioned design. At first, the hardware environments are to be explained, following to that, software environments will briefly commented.

4.1.1 Hardware Environments

Total system contained of two of EWSs (Sun spark 10 equipped with enriched main memory and hard disk) and three personal computers (PC-9801) corresponding to three operators each other, shown in Fig.12. All computers are connected to a common bath through a LAN. Left hand side EWSs is for calculating the plant transient phenomena and right hand side of EWSs is for simulating human behavior. It looks like to be able to simulate a team behavior, however, HHI built in the system was the simplest one that allow no consideration about group dynamics in this stage. Thus, exactly speaking, it is for simulating an individual operator behavior.Each PC, equipped with the voice composer, is available to announce utterances from each operator.

4.1.2 Software environments

Programming had been developed on C^{++} language for OMT based modeling, with using C language as a connection each other. For displaying of simulation results, the DATA VIEW package software was utilized. Furthermore, SYABELINBOO (speaking kids) was also adopted as a voice composer for three PCs. To develop the programming concerning the functional model, we adopted the finite state machine as a convenient tool. Additionally, we didn't adopt any knowledge engineering tool such as NexpertObject or GII because the inference engine of this simulation is the mental model mechanism and knowledge base was arranged specifically shown in Fig.5 and Table 4.

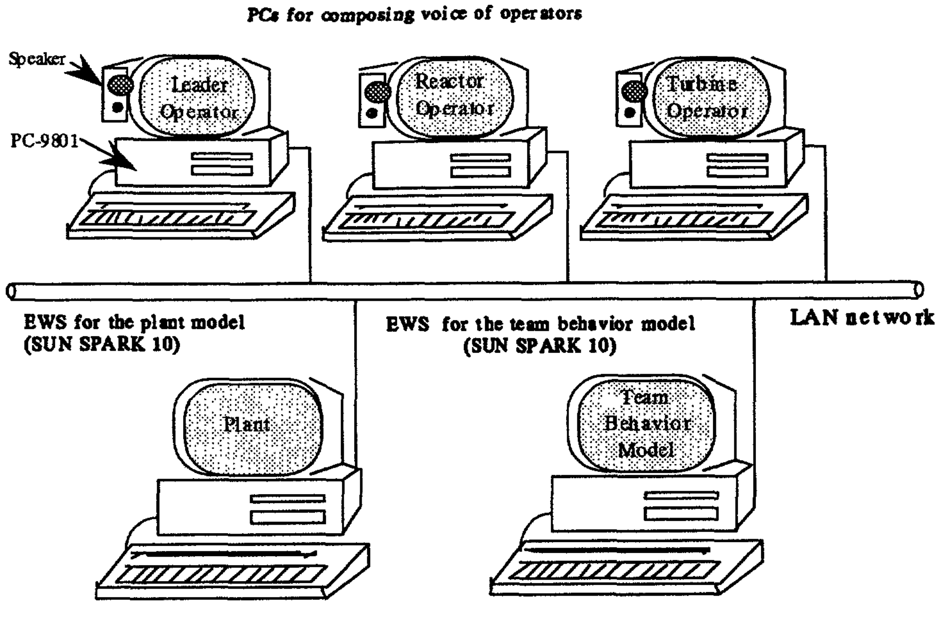


Fig.11 System configuration of SYBORG with two EWSs and three PCs

4.1.3 Simulation program developed

The program for the object model had been developed using C++ by directory converting the object model drawn in Fig.6. Recently, a convenient tool named OMTOOL was supplied commercially, this made it possible to translate the figure into concrete program. It implied that development of the program for the object model is quite easy and modification necessary in near future is also conveniently available.Next, the dynamic model was translated into program by using the finite state machine. For the describing program efficiently any library registered in C++ was applied. Also the functional model was translated to program using the same language, however, the description of KB was made by C language, connection was made also using C language. Event time sequenced was managed by originally developed code similar to the SimScript for operating this simulation code.

4.2 Simulation result

Simulation was conducted on the system shown in Fig.11. However, because it is necessary to validate this system whether it could present correct and similar human behavior or not, then the simplified plant system was introduced (Fig.12). This plant system was also used for manned experiment with the same malfunctions adopted in the simulation. The results of the manned experiment will be introduced later. This simplified plant simulator was developed by modifying a representative BWR plant (Saso, et.al 1994). This simulator made it possible to simulate the transient of 15 kinds of malfunctions and of their combination anyway. Simulation has been made for every malfunctions, then, we could got results of human behavior including utterances/actions against plant behavior. Fig.13 shows one of their result describing time-sequential protocols and events realized in the simulation due to a failure-open of the Condenser Water Recycling Valve. As shown in this figure, operators have noticed the occurrence of an off-normal condition at plant by the ANN indication. At the moment they have framed up their mind the mental model shown Fig.10, hereinafter the processes were progressed by the OMT model, that was described above in chapter 3, according to the mental model. Water Recycling Valve stuck open, in which case the initial indication on simulator display was "Condenser Hot Well Water Level LOW". They would envision the near-future scenario of plant dynamics evolution in terms of the Condenser Hot Well Water Level, which the operators would choose as a Key Parameter: If this Water Level falls to "LOW-LOW", the interlock logic would trip all CWPs and FWPs; this should rapidly lower the reactor water level, which upon marking "LOW-LOW", the interlock logic would scram the reactor. Based on the foregoing mental model, the operator would first open the CW Supply Valve as a Immediate reaction

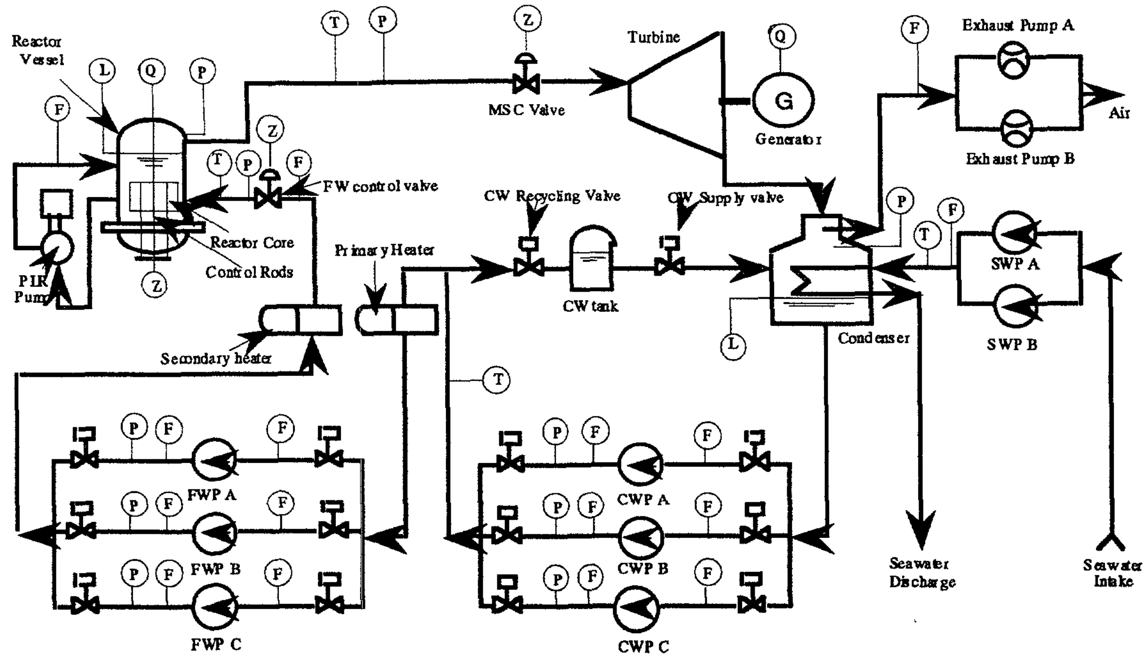


Fig.12 Schematic diagram of the simplified plant simulator developed

to maintain the hot well water level. Following this operation, the operators would envision proceeding first to monitor related plant parameters such as CWP/FWP flow rates, and then the statuses of FWP and CW recycling Valves, to seek the cause of the annunciator indication. Upon finding the CW Recycling Valve stuck open to be the cause, they would proceed immediately on the Causal Remedy of closing the valve by manual operation. If the Key Parameter happens to drop rapidly before the operator can identify the cause, the Emergency Measure to be adopted would be to decrease output power by lowering PLR pump speed. The simulator indicated that the foregoing Immediate Reaction and Causal Remedy would suffice to restore to normal state the Key parameter——i.e. Condensate Water Level——and then the Immediate Reaction, which was to open the CW Supply Valve, should be canceled, and the Valve should be reset.

Simulation result shows only interactions between the plant and operators and between operators. We could also obtain the internal processing trails in operator's mind as a typical output of the EWSs, but no introduction was made in this paper because it was quite too much to show them. However, it has been confirmed that data obtained concerning to the internal processes showed a reasonable agreement with our design mentioned above in chapter three, especially with the dynamic model on OMT.

4.3 Manned experiments for verification of this simulation model

2 teams of well trained students aged 19-25 years old, to whom our staffs made 100 hour training including the basics of plant dynamics and real operational practices using simulator, and one team of experts who had experiences as operators in real NPP. Each team contains of 3 operators. Under circumstances seen in Fig.14 with the control panel, experiments against each malfunction had been conducted for verifying this operator behavior model. In each unit of this experiments, the time sequential operators' protocol and actions were recorded to a VCR tape and an audio tape, of course, operations and plant behavior was accumulated in the hard disk of EWSs. After experiments, any utterances and actions made by operators and any ANNs appeared at the control panel in experiments were written down to papers by reviewing all records.

4.4 Comparison between the experiments and simulation

Comparison was made between experiments and simulations in two ways. One is the comparison of the events, that involved the records of ANNs and every levels of operations, Next is comparison of

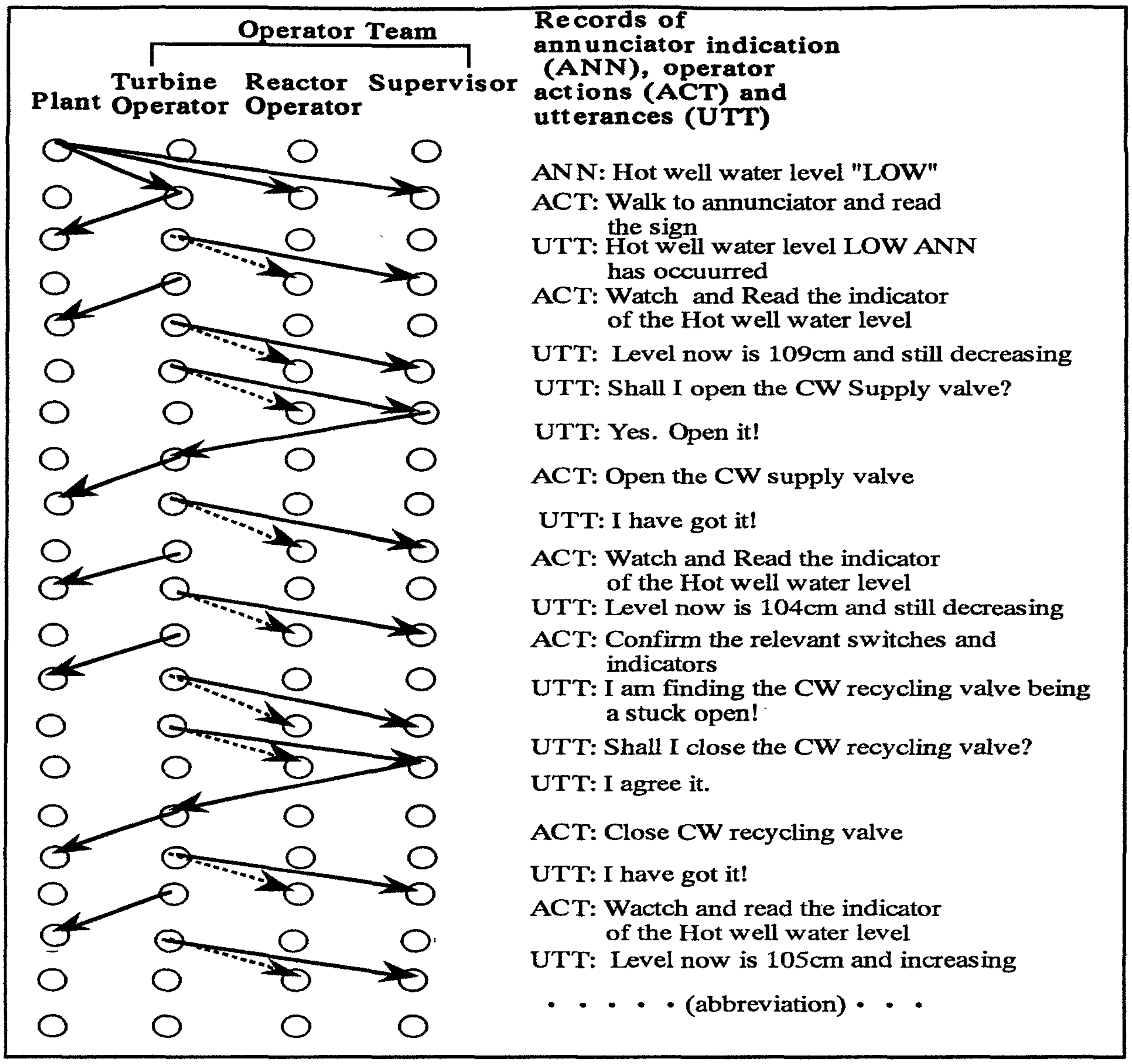


Fig.13 An example of the simulation result obtained applying to a failure-open of CW recycle valve

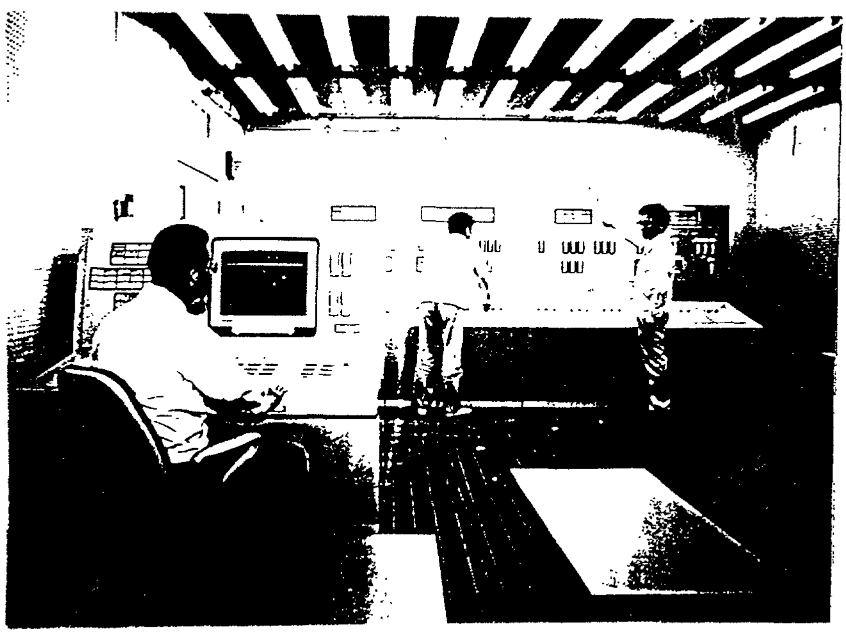


Fig.14 Typical scene of manned experiments

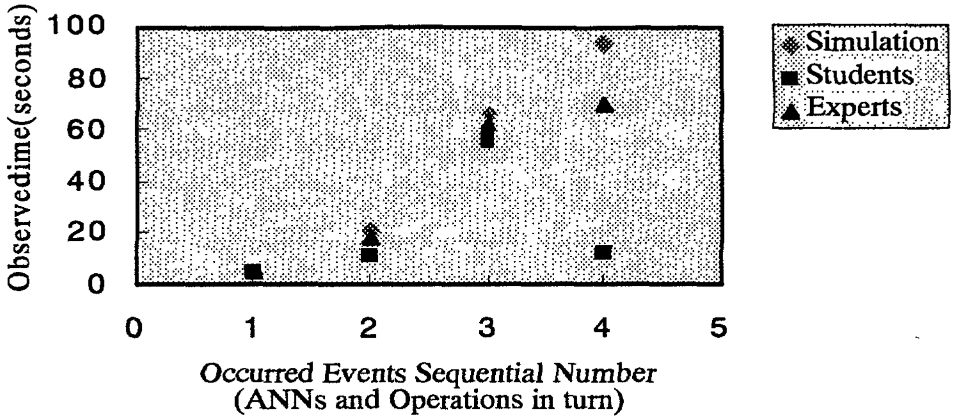
utterances obtained by both experiments and simulations. Objective malfunctions were; (1) a functional impairment of Seawater Pump (SWP) ; (2) a leakage at the pipe behind the Condensate Water Pump (CWP) ; (3) a sudden increasing of Seawater temperature. All conditions both in experiments and simulation were quite the same. And no information concerning which kind of malfunction would be subjected was presented to testees before starting experiments.

4.4.1 Comparison of events obtained

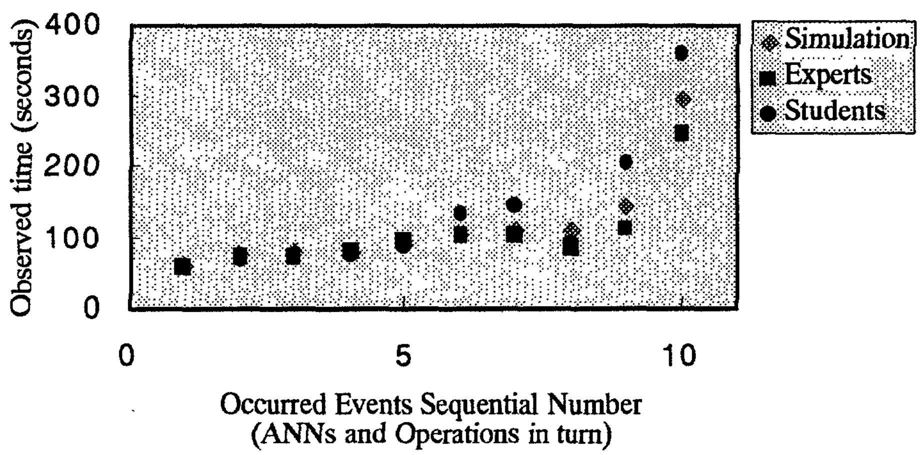
In this session, comparison was made of interactions between the plant and operators excluding utterances between operators. Figure 15(1)-(3) shows the results of this comparison regarding to interactions (events) occurred in both experiments and simulations. The horizontal axis means the sequential number of events from the beginning, as usual, first number represents the occurrence of ANN. This sequential number does not include utterances between operators. If there was the same ANN or operation observed both in the experiment and simulation, the same sequential number was afforded in turn. Nevertheless, there are no same ANN or operation respectably, then the corresponding mark was presented as a vacancy. Vertical line means the time recorded in seconds from the start point of experiments, but the time of first sequential number has no meaning because it was previously set by experimental staff at random. In figure 15 (1), No.4 operation by students, that was a operation to stop the CWP A, is quite isolated than the other data. It was a operation against the standard rule, a kind of human error of out of sequence. Also, in the figure (3), there were a lot of defects in the experts record. This is inching operations of degrading the reactor power by decreasing the speed of PLR pump. These operations were duplicated over and over. Then, it means that the experts team could control the plant not to decrease the reactor power than others. Excepting these point, simulation result located among experimental results made by students or experts. It was interesting that consequences to the plant at last were similar in every cases. This implied that simulation was going well regarding to the interactions between plant and operators.

4.4.2 Comparison of utterances obtained

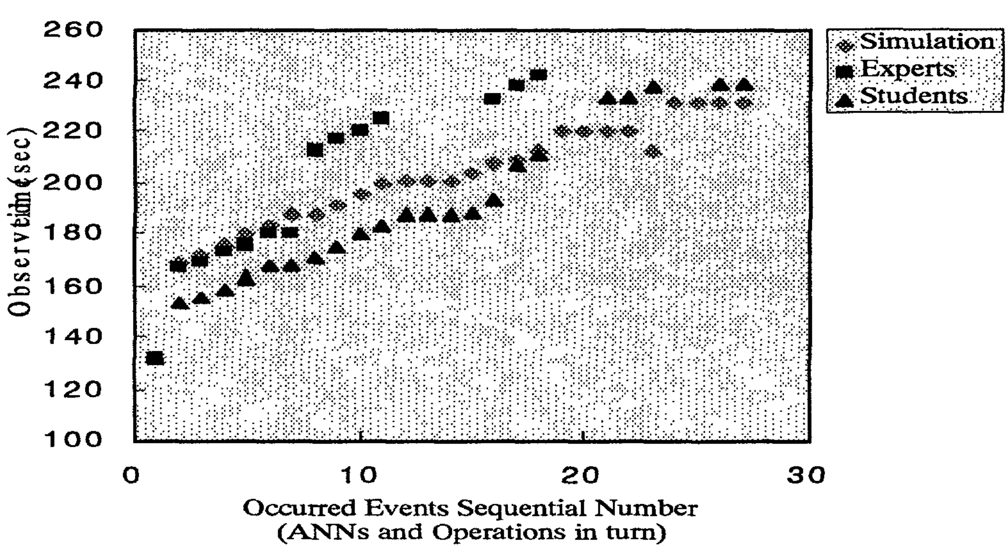
As one of the most unique features of our simulation, utterances between operators could be available. In this simulation, the simplest HHI mechanism was envisaged because of focusing on the individual operator behavior. The comparison can be made, however, between experiments and simulations. Fig.16 shows a typical example of this comparison applied to the above mentioned malfunction (2), a leakage at pipe behind a CWP. This figure shows the beginning part of the simulation. Classification of utterances should be made both in experiments and simulations and it was marked following to the sentence expressing content of an utterance. Every utterances could be classified into some categories defined in Table 3 even in the experiments. Each corresponding utterance was connected with a bar with arrow. Center column is the results of simulation, right hand side is by students team, and left hand side shows utterances made by expert, respectably. The utterances made by simulation was more than experiments because real human combine two or three utterances together at the same time. And at the last of Fig.16, the timing of an application "Shall I open a CW supply valve" is out of order. This is why the immediate action could be done after the ANN of Hot Well water level high in the simulation. However, human operator could do this before ANN because they have already detected the lowering of the level of the Hot Well. This point should be modified, but this modification will be more essential relating to the structure of the mental model. Then, it will be taken into accounts at the second stage in near future. Excepting these points, more or less, every important conversations were satisfactory simulated even compared to those of the experts team. To compare these mathematically each other will be severely difficult, it should depend on qualitative comparison. We cannot show the other case-studies in this paper, however, most of them had got the similar history of utterances. Generally speaking, the number of utterances observed in the experiments were less than those observed in simulation. It means that real humans made a rational conversation to minimize frequency by merging any kinds of utterances whenever it was possible. Furthermore, through the whole results, histories of utterances made by simulation were closer to those of experts team rather than students because students made abbreviation any necessary conversations when it had been already trivial. On the contrary, experts must have



(1) Functional impairment of a SWP



(2) Leakage at pipe behind a CWP



(3) Sudden increasing of Seawater Temperature

Fig.15 Comparison of events observed between experiments and simulation

Experts Team		Simulation		Students Team	
(T) Something has come! Leakage in turbine building has occurred!	Report	(T) ANN "Leakage in Turbine building" has occurred	Report	(T) There is leakage in turbine building	Report
(S) Oh! Leakage in turbine building	Reply	(S) Yes.	Reply	(S) Yes	Reply
(R) Ok! Leakage in turbine building!	Reply	(T) There is a leakage in CWP B	Report	(T) Hi? Pump B has a leakage	Report
(T) Exhaust flow of CWP B seems less than suction flow	Report	(S) Indeed!	Reply	(S) Yes. Start up the C just now!	Reply Direction
(T) So, you know, Shall I start up C	Declare	(T) Shall I change to alternative CWP	Declare	(T) Shall I start up C	Declare
(S) Yes! let's	Agree	(S) Please!	Agree		
		(T) Condenser Hot Well water level is now 112 and still decreasing!	Report		
		(S) Yes.	Reply		
		(S) Please correspond with monitoring Hot Well water level, Turbine operator	HHI		
		(T) Yes.	HHI		
		(T) I have got CWP started up!	Completion	(T) I have got C started up!	Completion
		(T) Yes.	Reply	(S) Oh! OK.	Reply
(T) Exhaust valve opened!	Completion	(T) I have got CWP's exhaust valve opened	Completion		
(T) OK!	Reply	(T) Yes.	Reply		
(T) I have got CWP started, then B stopped	Completion Declare	(T) I have CWP B stopped!	Completion	(T) B stopped now!	Declare Completion
(S) All right.	Reply	(S) Yes.	Reply	(S) Yes. Let's	Agree
(T) B stopped!	Completion	(T) Condenser Hot Well water level is now 111 and still decreasing!	Report	(T) Hot Well water level now, 111	Report
(S) Hi!	Reply	(S) Yes.	Reply	(S) Yes.	Reply
(T) Hot Well water level now 111	Report	(T) CWP A is normal	Report	(R) Reactor water level now is 7.5.	Report
(T) Shall I open CW supply valve. Now opened!	Declare Completion	(S) Indeed!	Reply	(S) Yes, 7.?	Reply Request
(S) OK!	Agree	(T) ANN "CW Hot Well water level High/Low" has occurred!	Report	(R) It's steady state	Report
		(S) Yes.	Reply	(S) Yes.	Reply
(S) Oh! Rapid decreasing! A large amount of leakage, it is!	Report	(T) Shall I correspond CW leakage quickly!	Proposal	(T) Let CW supply valve be opened. Hot Well "level LOW"	Report Declare
(T) A,C,A,C.....I see Hot Well 110!	Report	(S) Correspond CWP leakage, please.	Decision	(S) Hoi!	Reply Agree
(T) It's "Hot Well water level LOW"	Report	(T) Condenser Hot Well water level is now 109 and still decreasing!	Report	(R) Reactor water level and pressure is now quite steady	Report
(S) Hi!	Reply	(S) Yes.	Reply	(S) Yah!	Reply
(T) Well? Then, where leakage located?	Proposal	(T) Shall I open CW supply valve.	Application	(S) Woul you isolate CWP B? Is it OK!	Proposal
(S) Isolate B, B!, sure it is.	Direction	(S) Please!	Approval		

Fig.16 Comparison of utterances observed in both experiments and simulatioı

reported everything even if it is trivial for reconfirmation. Above mentioned simulation results looks like satisfactory for us even including utterances.

5. Conclusion

It has been 3 years from starting this project, at first, authors made wide spreading survey on human modeling. We found that there had been less information about human modeling in process control (Sasou et.al 1991); one of them were only on tracking performance, another on the inching operation expressing human behavior as a transformation function, and the others on the task networks.
However we could found that some comprehensive approaches were made by COSIMO (Cassibue, 1991) and CES (Woods,1987), and these were a basis for this study, furthermore the conceptions proposed by Hollnagel (1993), Wickens (1984) and Baron (1982) suggested us the new frame of human modeling. This is why we started this work to aim the modeling of a team behavior as the more general human modeling, which is able to deal with every kinds of malfunctions occurred in NPPs. Description in this paper is a part of this study as a first step because of less consideration about HHI mechanism and of simplified plant simulator. We have been conducting further research to develop more precise HHI mechanism by introducing the group dynamics and improving plant characteristic by introducing a full scope simulator. Even more, we have been making effort to introduce human error mechanism in this model in order to overcome prescribed purpose and to improve human characteristics by considering human emotion and reactions under stressful situation.
The SYBORG described here are as the first step of our continuous efforts, however, we enabled this individual to have a function of speaking and hearing for communications. And our model could treat with general problems, that is functional impairment occurred in plant as a malfunction, by discovering the mechanism of the mental model as a so called inference engine. The most remarkable advance is creation process of the mental model in relation with knowledge base and its running process designed by OMT. Relatively programming of OMT based design was said to be easy, so it is easy for not only designing but also modifying Above mentioned results obtained probably shows reasonable coincidence with experiments. It implies that, in near future, a full scope simulation including precise plant simulator and human behavior models would be available for various purpose.

References

Baron S. (1982) Framework for modeling supervisory control behavior of operators of NPPs, Proc. workshop on Cognitive modeling of Nuclear plant control room operators, NUREG/CR-3114

Cacciabue C. (1991) COSIMO, A cognitive model in blackboard architecture. Symp. AI. Pychol ., Reliability Engineering and System Safety, Special issue on cognitive science approach to process control.

Cacciabue C. (1994) Academic and practical needs for operator modeling and simulation. Second Workshop on SuperSimulators for NPP, Tokyo, November 2

Card SK. (1983) The psychology of human - computer interaction, Lawrence Erlbaum Associates, Hillsdale, NJ.

Fujita Y., Yanagisawa I., Itoh J., Yamane N., Nakata K., Kubota R. (1993) Modeling operator with task analysis in mind. ANS Topical Mtgs. on NPP Instrumentation, Control and Man Machine Interface Technologies, Illinois USA.

Hollnagel E. (1993) Human reliability analysis: Context and control. Academic Press, London UK

Huang YH. (1991) A system model for operating crew behavior during accident sequence. MIT., Ph.D in Nucl. Eng.

Kozinsky EJ., Gray LH., Beare AN., Barks DB., Gomer FE. (1984) safety related operator actions: Methodology for developing criteria. NUREG/CR-3515.

McRuer D., Graham D., Kredel E., Reisener W. (1965) Human pilot dynamics in compensatory systems - theory, models and experiments with controlled elements and forcing function variations. AFFDL-TR65-15, Wright Patterson AFB

Rasmussen J. (1983) Skills, rules and knowledge: Signals, signs and symbols, and other distinctions in human performance models. IEEE Trans. on Sys. Man, Cybernetics, SMC-13, 257

Sasou K., Takano K., Iwai S. (1991) A study on the operating team activity in NPPs - Present status and conceptional design of the SYBORG. CRIEPI Rep. No.S91001 (in Japanese)

Sasou K., Takano K., Yoshimura S. (1993) Development of a team behavior simulation model. Fourth European Conf. Cognitive Science Approaches to Process Control, Frederiksborg, Denmark

Sasou K., Takano K., Yoshimura S. (1994) Behavior simulation of team in NPP - Detail design of HHI (human human interface) model to simulate decision making and task assignment among operator models, CRIEPI Rep. S93002 (in Japanese)

Sasou K., Iwai S., Takano K., Yoshimura S. (1994) Behavior simulation of team in NPP - Development of control panels to verify team behavior simulation. CRIEPI Rep. S94002 (in Japanese)

Siegel AI., Barttler WD., Wolf JJ., Knee HE., Haars PM. (1984) Maintenance personnel performance simulation (MAPPS) model: Summary description. NUREG/CR-3626.

Takano K., Sasou K., Yoshimura S., Iwai S. (1994) Creation of operators' mental models in coping with abnormal event occurring in NPP. Int. J. Human Computer Studies. in submission

Takano K., Sasou K., Yoshimura S. (1995) Knowledge base structure to create the operator's mental model in coping with an abnormal event at NPP. Int J. Human Computer Studies. in submission

Wherry RJ. (1976) The human operator simulator HOS. In Sheridan TB. & Johansen G. (Eds.), Monitoring behavior and supervisory control. Plenum Press, New York.

Wickens CD. (1984) Engineering psychology and human performance. Charles Merrill, Columbus, OH

Woods DD. (1987) Cognitive environment simulation: An artificial intelligence system for human performance assessment: Modeling human intention formation. NUREG/CR-4682.

TOWARDS A CONTROL-THEORETIC MODEL OF PILOT MANUAL CONTROL BEHAVIOUR WITH A PERSPECTIVE FLIGHT-PATH DISPLAY

Max Mulder

Delft University of Technology, Faculty of Aerospace Engineering, P.O. Box 5058, 2600 GB Delft, The Netherlands

e-mail: m.mulder@lr.tudelft.nl

Abstract. In future aircraft, conventional displays might be replaced with perspective displays, which show the future planned trajectory in a synthetic three-dimensional world. The manner in which the flight information is presented strongly influences the pilot's ability to control the aircraft along the desired trajectory. A research project has been initiated to understand and mathematically model the manual control behaviour of a pilot who maintains the reference trajectory by using a perspective flight-path display. The primary issues of obtaining such a model will be discussed and, based on a theoretical analysis of the main characteristics of the pilot's guidance task, modelling solutions will be postulated in the paper.

Keywords. Manual control, perspective flight-path displays, human operator modelling.

1. INTRODUCTION

The use of fast and reliable graphics computers in the cockpit enables the designer to dramatically change existing display formats. One possibility is the introduction of the third dimension into the primary flight display. The resulting perspective display format would present a spatial analog of the out-of-the-windshield visual scene. The display can be further augmented with a number of synthetic display elements, which are designed to help the pilot in performing his control or monitoring task. For instance, it is possible to include the future desired trajectory in the perspective scene, resulting in a *tunnel-in-the-sky* display (Figure 1).

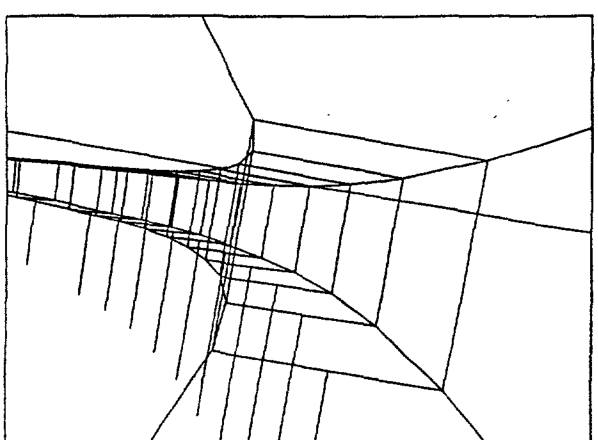


Fig. 1. Tunnel-in-the-sky display.

The application of a perspective display has considerable impact on the pilot's manual control task. In a conventional cockpit, the pilot mentally reconstructs the aircraft's spatial and temporal situation from a number of two-dimensional, planar displays. With a perspective flight-path display, the guidance and short-term navigation information is presented in a spatial, pictorial format, allowing the pilot to directly update his mental representation.

At the Delft University of Technology, a research project has been initiated to understand and, ultimately, mathematically model the manual control behaviour of a pilot controlling his aircraft along a reference trajectory using a three-dimensional, perspective flight-path display. Modelling the pilot's control behaviour involves analyzing his short-term planning behaviour, his utilization of future control-oriented information (*preview*) and his information processing characteristics. The model should be able to account for changes in pilot control behaviour and performance due to different display design parameters and additional display features, such as flight-path vectors or flight-path predictors.

Some important issues of obtaining such a model will be discussed in detail in this paper. Based on a literature survey as well as a theoretical analysis of the main characteristics of the pilot's guidance task, possible modelling solutions will be proposed.

2. PERSPECTIVE FLIGHT-PATH DISPLAYS

2.1. *The concept*

It has been hypothesized since the early 1950's that a pictorial, three-dimensional, true-perspective display, showing a direct contact analog of the through-the-windshield visual scene could be the optimal way of presentation. In the last decade a remarkable increase in research concerning the possible implementation of all kinds of three-dimensional, perspective displays. In many experimental studies it has been shown that presenting the guidance information by means of a perspective primary flight display can improve the information transfer from machine to man to a large extent (Wilckens 1971, Grunwald *et al.* 1981, Wickens *et al.* 1990, Theunissen and Mulder 1994, 1995).

Numerous experimental programs indicated that combining command guidance, by means of a spatial reference such as a tunnel, with a forward-looking true-perspective flight display can yield performance on MLS[1] (curved, steep) landing approaches superior to that achieved with conventional flight displays (Grunwald *et al.* 1981). The integration of command guidance in a perspective display shows the pilot the desired flight-path as a function of time and position. This enables him to anticipate to the changes in the reference trajectory, to conduct flexible manoeuvres, and to maintain his situational awareness. In the case of such a display, the task of the pilot consists of controlling the aircraft through the virtual three-dimensional tunnel, a typical aircraft guidance task (Figure 2).

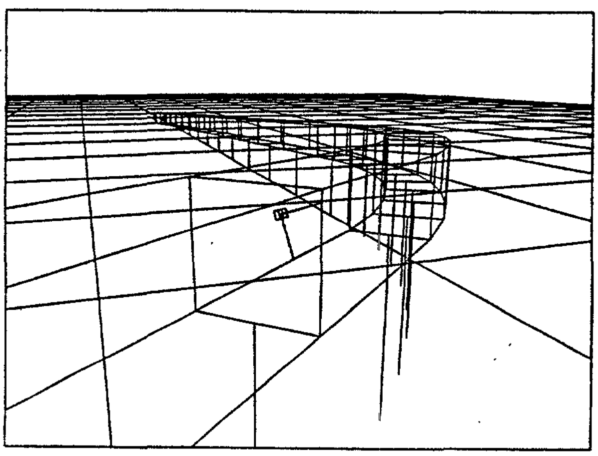


Fig. 2. Aircraft guidance through a virtual tunnel.

An important design advantage of synthetic, three-dimensional displays is that one may enhance pilot's control performance by including abstract, computer-generated display features, which are not available in the usual out-of-the-windshield visual scene. For instance, by incorporating flight-path command information together with flight-path prediction information, one can determine or manipulate the task to be executed to a considerable extent (Grunwald and Merhav 1978, Theunissen 1993).

[1] *MLS*: Microwave Landing System

3. MODELLING PILOT CONTROL BEHAVIOUR WITH SPATIAL INSTRUMENTS

3.1. *Literature survey*

A perspective flight-path display has considerable impact on the pilot's manual control task. Instead of mentally computing the relative flight-path position information from two separated, orthogonal, planar displays (Primary Flight Display (PFD) and Navigation Display (ND)), this information is presented in an integrated manner in a single three-dimensional display. Since the pilot's control behaviour stems from the mental construction and utilization of his internal representation of the control task, it can be expected that presenting the aircraft's guidance information differently will have a major effect on his control behaviour (Haskell and Wickens 1993).

In 1993 a research project was started to examine and mathematically model the manual control behaviour of a pilot maintaining the reference trajectory using a perspective display. Starting point in the project was a literature survey, published in (Mulder 1994), some of the results of which are discussed here.

3.2. *Consequences of a spatial display format*

3.2.1. *Effects on modelling.* First of all, presenting the guidance information by means of a perspective display has important consequences for the attempts to *model* the operator as a controller and information processor. Instead of reading off a number of discrete instruments, the operator is now confronted with a synthetic three-dimensional world. The different elements of information that are important for controlling the aircraft can now be perceived directly from the display screen. Further, the effects of his control actions are presented in the same, *spatial*, format, allowing the operator to close several, possibly spatial, feedback loops.

The functional information must be perceived from this display. By providing the operator with a pictorial analog of the outside visual world, enhanced with additional synthetic position information, the information is presented in a way humans are used to in daily life. Since man is very experienced and therefore highly proficient at processing these moving three-dimensional scenes, it

can be expected that the information necessary for guiding and manoeuvring the aircraft along the trajectory, can be perceived in a single observation. In other words, the information is presented in a manner that is highly compatible with the pilot's internal representation of his task.
Identifying the exact sources of information is a difficult problem. In a normal visual scene, many potential cues are available to the operator to construct a mental picture of his locomotion through the environment. According to Gibson (Gibson 1950), it is the optic array of light that contains important features or cues that are directly controlled during visually guided locomotion. In the ecological paradigm of perception as an active process, the observer controls some optical invariants in the optic array to perform a given task.

3.2.2. *Empirical approach.* A mathematical approach to describe the behaviour of an operator who observes a spatial display is difficult, if not impossible. Therefore, most of the research conducted to evaluate the pilot's performance with spatial instruments is of *empirical* nature (Haskell and Wickens 1993): The operator has to perform a simple, often one-dimensional, control task in which a virtual vehicle with highly simplified dynamics must be controlled by using an extremely simplified spatial display. The effects of varying some display characteristics on the operator's overt control behaviour is then examined. These empirical efforts can give meaningful insight into the relation between the presented cue(s) and the functionality or importance of this cue for the operator's control task. Nonetheless, these efforts often cause contradictory results and do not lead to an understanding of the relationships between what has been presented and the resulting control actions of the operator. Moreover, the results from such an empirical investigation cannot be extrapolated to a different display-control situation.

3.2.3. *Modelling approach.* A mathematical model can provide a *systematic* tool to investigate some of the information-processing characteristics of the operator. In mathematically modelling the operator using spatial instruments, the main question of interest is the way in which the operator extracts the control-oriented information from the visual scene. In other words, when confronted with a spatial display, what elements in this visual scene does the operator use in computing his control signal and what elements does he use to decide whether or not and when to initiate the control signal?
This is exactly the question which is addressed by Grunwald (Grunwald and Merhav 1976, 1978), possibly the first who attempted to model the phenomenon of what he called the *visual field control* task. According to Grunwald, for controlling a vehicle using a spatial display, it is essential that the control-oriented information to be perceived from the visual field is obtained from a considerable forward view of the control situation ahead. The operator mentally computes the future path of the vehicle, and by comparing the actual future path with the desired trajectory, a control action will be decided upon. Grunwald's work emphasizes the importance of the availability of trajectory preview information in a spatial display.

3.3. *Automobile driving?*

The pilot's task of controlling an aircraft along a space-constrained trajectory has several significant similarities with the more common control task of automobile driving (Mulder 1994, Theunissen and Mulder 1994, 1995). The primary resemblances that have considerable impact on the operator's control behaviour are:

- the *preview* of the trajectory ahead, allowing anticipatory control,
- the *boundary control* nature of the task, which can result in distinctly different control strategies from the compensatory error-minimization task,
- the estimation of the control-oriented information from the outside out-of-the-window visual scene.

Because of these striking resemblances between the car driving task and the aircraft flight-path control task with a perspective display, it is postulated that the pilot adopts a control behaviour that is comparable to that of an automobile driver. Since the pilot's control behaviour is linked to his information-processing characteristics, it is expected that a perspective flight-path display allows him to use, in the least, visual cues comparable to the cues in the automobile driving task.

3.4. *Automobile driver models*

In the last decades a truly remarkable amount of driver models has been developed. Validated driver models exist for various driving tasks such as the lane-change manoeuvre and the obstacle-avoidance manoeuvre. Here, our primary interest goes to the driver models for driving on straight roads and the curved lane tracking task.
As mentioned above, the main question in modelling the driver's control behaviour is the identification of the perceptual feedbacks that are established in a car driving task. Obviously, the perceptual cues are a function of the vehicle motions relative to the environment. The functionality and usefulness of visual cues are usually validated in the following ways:

- by considering the guidance and control requirements of the man-vehicle system (system *controllability*), e.g. (McRuer 1977),

- by determining the availability and functionality of the visual cues in the visual field (system *observability*), e.g. (Gordon 1966),
- by conducting eye-movement experimental studies, which could lead to an identification of the most utilized visual cues, e.g. (Gordon 1966, Kondo and Ajimine 1968).

The control-theoretical analysis is commonly conducted in the frequency domain using *multi-loop* extensions of the classical cross-over-model modelling methodology (McRuer *et al.* 1977). Throughout the years the investigations have pointed in the direction of *multi-level* models, in which the operator's control behaviour is described by several, more or less independent, and strongly task-related submodels (Reid 1983).

3.4.1. *Effect of trajectory preview.* Central in most driver-modelling attempts is the notion of *duality* of information presented to the driver by the forward view of the road (Wierwille 1967, Donges 1978). First of all, the visual field of the driver provides information on the instantaneous and future course of the road, so that the driver can extrapolate not only the present but also the future course of the driver-vehicle's forcing function. The information on the forcing function is utilized by the driver for vehicle *guidance* along the desired path. Secondly, static and dynamic visual cues in the visual field contain information on the instantaneous deviations between the vehicle's actual path and its desired path. The driver utilizes the corresponding instantaneous visual cues to *stabilize* vehicle motion with respect to the forcing function.

As a result, the driver is generally modelled on two levels of control. The higher level represents the utilization of the previewed future reference trajectory, resulting in a feedforward *anticipatory* control activity. The lower level is assumed to represent the continuous control of vehicle position and heading, resulting in a feedback *compensatory* control activity.

3.4.2. *Curve-negotiation.* An experienced car driver has a proficient internal representation with respect to the vehicle handling qualities and the disturbances acting on the vehicle, but also in terms of perceived changes in the future trajectory. This enables the driver to anticipate the oncoming changes by initiating an open-loop control action some time before the actual event occurs. This is reflected in the eye-movement studies of Kondo (Kondo and Ajimine 1968) and Shinar (Shinar *et al.* 1977), who found that a driver exhibits anticipatory lateral eye movements some time before the change in trajectory actually started. Consequently, Shinar postulated that, in what he called the *curve-negotiation process*, the driver estimates from the road ahead of the vehicle the trajectory changes in terms of curvature and roadway characteristics. Based on this information and the actual state of the vehicle, the operator uses his internal representation to compute an anticipatory control action. The quality of this control action is, besides the experience of the driver, dependent on the accuracy of the perceived out-of-the-window information and the transformation of this information into terms compatible to the driver's internal representation.

3.4.3. *Some modelling aspects.* The division of the driver control action in anticipatory and compensatory control actions has become a well-established theoretical hypothesis. The way in which the two control modes are modelled, however, shows many alternative viewpoints, especially for the anticipatory control. This is a result of the relatively many potential ways in which one can model the use of the future trajectory information. Does the driver fixate at a single point ahead of the road, does he look at a fixed angle into a roadway curve, or does the driver use a continuous span of future errors? Further, does he act on perceived changes in future position, or heading, or road curvature?

In general, using a one-point looking distance approximation of the effect of the previewed road ahead can lead to satisfactory results (McRuer *et al.* 1977, MacAdam 1981, Guo and Fancher 1983). The modelling attempts of Grunwald (Grunwald and Merhav 1976), however, showed that it takes at least a two-point looking distance model to model the operators' control behaviour properly. This is confirmed by the work of Hess (Hess 1981, Hess and Modjtahedzadeh 1989) and others, who attributed this fact to the estimation of the *trend* in the forcing function, which is impossible with a single-point looking model. Yet, it is the trend that makes the preview worthwhile, since this trend allows an inversion of the low-frequency system dynamics, resulting in a 'perfect' feedforward control (Hess 1981).

Concerning the perceptual cue which is used in anticipatory control in curved lane tracking, almost all modelling attempts consider the road curvature as the primary variable of interest (McRuer *et al.* 1977, Donges 1978, Reid 1983).

3.5. *The two-level driver model of Donges*

One of the automobile driver models will be described here to illustrate some of the modelling issues discussed above, but also because it will be used later to show the validity of the optimal control modelling methodology postulated in the next section.

Donges (Donges 1978) proposed a two-level model

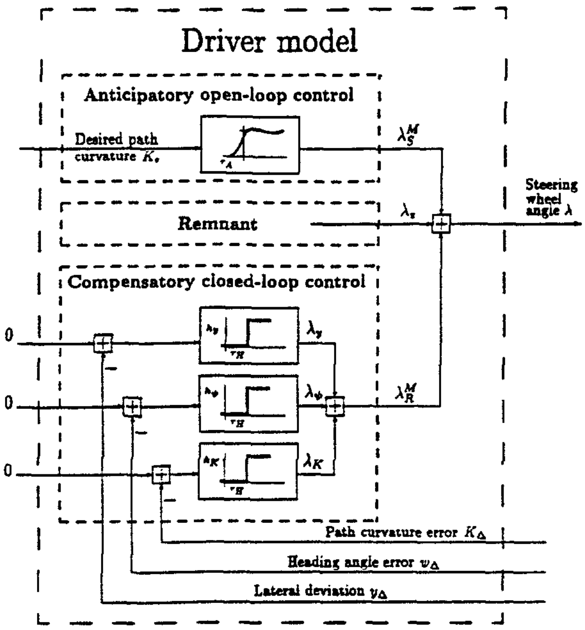


Fig. 3. A two-level model of driver steering behaviour (Donges 1978).

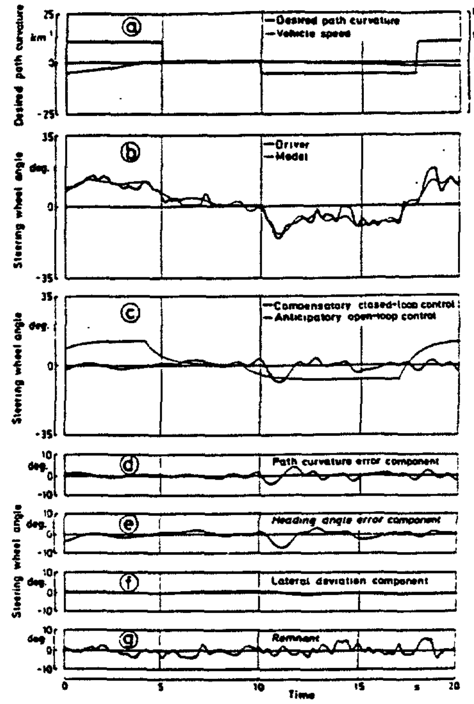


Fig. 4. Time histories of steering wheel angles of driver, driver model and driver model components (Donges 1978).

of driver steering behaviour. Although the multi-level modelling methodology was not new, Donges was the first who experimentally validated the multi-level model structure and identified (in the time domain) the principle driver model parameters.

As mentioned above, the driver's main task in steering a vehicle is to extrapolate from the visual field the vehicle's desired path and the vehicle actual motions relative to this desired path. Consequently, the driver model of Donges is defined to consist of two levels (Figure 3):

1. the *guidance level*, involving the perception of the instantaneous and future course of the driver-vehicle's forcing function, i.e. the roadway ahead. The curvature of the road is considered to be the most suitable quantity for the sensory expression of the forcing function. The driver responds to this change in curvature by an open-loop, anticipatory steering wheel action.
2. the *stabilization level*, in which all deviations from the forcing function are compensated for in a closed-loop control mode. Here, the essential information quantities are:
   - the curvature error, i.e. the discrepancy between the curvatures of the desired and actual path (proportional to the lateral acceleration),
   - the path angle error, i.e. the discrepancy between the tangents of the desired path and the actual path (proportional to the lateral velocity),
   - the lateral position error.

An important result of the experimental validation is that the driver maintains a constant curve-entrance anticipation time of approximately 1.1 second, independent of the vehicle's velocity. The anticipatory control acts on the oncoming curvature change by means of a first order lag response, with a negative dead time. The model time simulations show a clear depiction of the basic driver steering behaviour, as is illustrated in Figure 4. Especially the division of the control actions in the anticipatory (feedforward) and compensatory (feedback) level is very illustrative.

4. TOWARDS A CONTROL-THEORETIC MODEL

In this section, the goal of the present research is defined. A methodology for modelling pilot control behaviour with a perspective flight-path display will be chosen. Based on a theoretical analysis of the pilot's guidance task, modelling solutions are postulated.

4.1. *Research goal*

The research focuses on modelling the characteristic elements of pilot control behaviour for a controller who uses spatial instruments. To model these elements properly, a model structure should be chosen that has the potential of incorporating the main characteristics of the operator's control behaviour. Instead of developing a model that can account for the human factor in any manual control situation, it should be able to examine our particular research interests:

- the *information-processing* characteristics of the pilot with perspective displays,
- the utilization of *trajectory preview* information,
- the effects of changing primary *design variables* of the perspective flight-path display (e.g. field-of-view, tunnel size, amount of trajectory preview),
- the effects of *additional display symbology*, such as flight-path vectors and flight-path predictors.

A perspective flight-path display can lead to a number of different control strategies (Theunissen and Mulder 1994, 1995), dependent on the task description. Although this flexibility is often stated as one of the virtues of a perspective display, the task considered here will be the continuous tracking task: *Fly the reference trajectory as accurate as possible, in spite of atmospheric disturbances acting on the aircraft.*

4.2. *Choosing a modelling methodology*

The first step towards a mathematical formulation of the operator's behaviour with a spatial display is the development of a control-theoretical model structure. When the basic model structure has been developed, the different parts must be validated experimentally.

4.2.1. *The classical methods.* The attempt to use describing function models to describe the manner in which observers translate the visual information of the outside world to appropriate control actions, is called here the *classical* approach. Generally, this approach lacks the possibilities to deal with some man-machine interactions in a conceptual manner. Especially the perceptual and motor limitations, which are important in the manual control context discussed here, can be dealt with only in an indirect manner. Although it is obvious that humans can not make extremely accurate observations concerning the vehicle's state, the classical approach does not incorporate this fact within the model structure. Observations concerning the state of the controlled system are usually stated to be *directly observable* from the outside visual world, without making any assumptions concerning the identification of usable cues or perceptual strategies. The states are perceived, in one way or another. The effects of the perceptual limitations of the operator are generally accounted for by increasing his remnant spectrum, decreasing his gain, or increasing his effective time delay.
With our research goals in mind it is clear that the classic approach lacks a conceptual structure to deal with the perceptual characteristics of the operator. Classical control theory <u>can</u> be used, however, to obtain a basic understanding of the elementary characteristics of the closed-loop system, and provide a control-theoretical basis for further research.

4.2.2. *The optimal control methodology.* Mathematically modelling a pilot operating with a spatial display format is commonly conducted using the well-known Optimal Control Model (OCM, Kleinman *et al.* 1970) (e.g. Grunwald and Merhav 1976, 1978, Korn *et al.* 1981) or a generalized control-theoretic model based on modern optimal control theory (Wewerinke 1989). Main reason for this is the superiority of the optimal control approach in providing a *conceptual* basis for incorporating task-related variables and some of the perceptual characteristics of the operator, such as his observation inaccuracies and thresholds, within the manual control framework. Our attempts to obtain a mathematical pilot model strongly leans on this fact.

4.3. *Determining a modelling approach*

As mentioned above, a mathematical model that can deal with the entire range of pilot control behaviour with a perspective display, is not the primary goal. So let's consider the pilot's guidance task in somewhat more detail.
The reference trajectory to be flown will mainly consist of straight sections interrupted by lateral or vertical transitions. Lateral transitions consist of a change in aircraft heading, leading to a (lateral-horizontal) curved section in the approach path. Vertical transitions consist of a change in the aircraft's flight-path angle (and perhaps speed), leading to a (vertical) curved section or, most probably, to an instantaneous change (twist) in the trajectory.
From a system-theoretical point-of-view, the reference trajectory to be flown leads to a series of different system steady-states and transition periods between these steady-states. This has important consequences for (modelling) pilot control behaviour, as was also realized in the automobile driving research discussed above (McRuer *et al.* 1977, Reid 1983). Maintaining a certain state of a system against disturbances leads to a regulation task. The transition between steady-states of the system, however, is different, both from a modelling as a perceptual point of view. This distinction will be discussed below.

4.3.1. *Straight sections of the trajectory.* If the reference trajectory is straight, the task of the operator comes close to a conventional continuous error-minimization task. On a straight section, in case of no position and angular errors with respect to the trajectory, the tunnel image will be

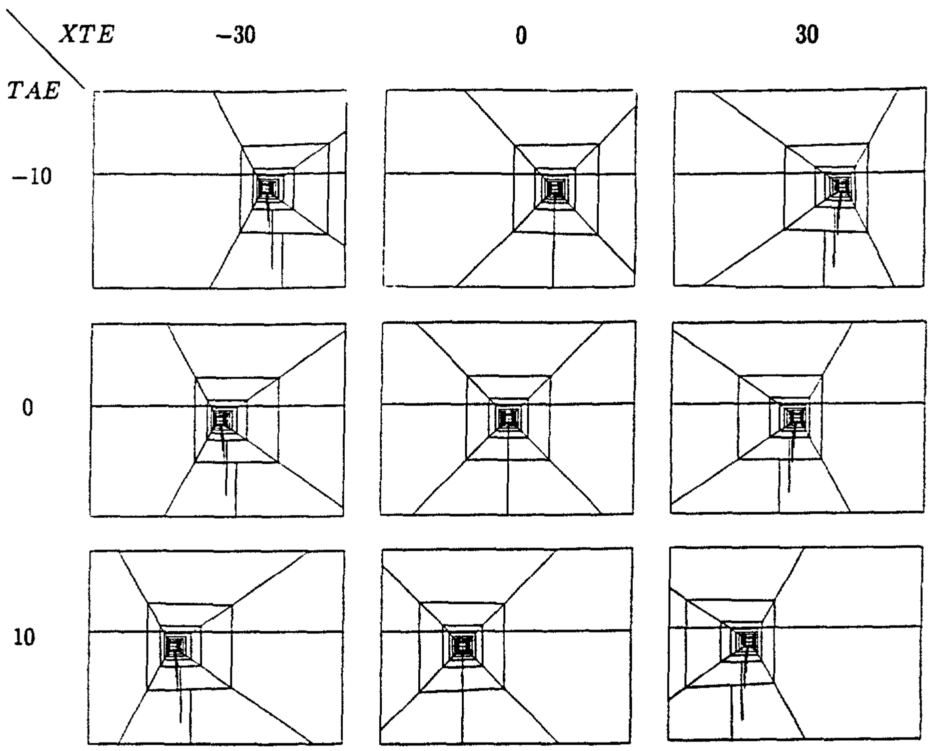


Fig. 5. Distortion of symmetric tunnel image due to position (cross-track-error XTE, in [m]) and/or angular errors (track-angle-error TAE, in [deg]).

symmetric[2] (Theunissen and Mulder 1994, 1995). Any deviation from the trajectory will lead to a distortion of this symmetric condition (Theunissen 1994) (Figure 5). In other words, to minimize the discrepancy between the actual and the desired trajectory, the operator must maintain a symmetric tunnel image, a regulating task which can be described mathematically as the minimization of some criterion.

Therefore, it is postulated here that the conventional Optimal Control Model is well-suited to model pilot's control behaviour on straight sections of the trajectory. The control-oriented information must then be modelled through the observation part of the OCM, by means of geometrical relations (the laws of linear perspective and motion perspective) between the visual cues of distorted symmetry and the aircraft's position and angular errors with respect to the reference trajectory.

Note that on straight sections the effect of preview of the trajectory ahead is twofold. First of all, a sufficient preview allows a better estimate of the track-angle-error (Theunissen 1994). Secondly, due to the increasing optic flow, an extended preview emphasizes the dynamic character of the task and will probably lead to a better estimate of higher derivatives of the vehicle's motion relative to the trajectory (Mulder 1994).

[2] Note that this symmetric condition is as well a *static* symmetry as a *dynamic* symmetrical optic flow.

The OCM approach was followed successfully by Wewerinke (Wewerinke 1989), who modelled the pilot conducting an approach-to-land task using an out-of-the-windshield perspective image of the runway. Another suitable example of this approach is the work of Blaauw (Blaauw 1984), who modelled the automobile driver using a comparable outside world observation model.

Experiments to validate the OCM approach on straight tunnel sections will be conducted at the end of this year. The experimental set-up will be simple: the pilot must maintain the trajectory using only a central tunnel-in-the-sky display and a side-stick. Thus, no vestibular or peripheral cues will be available.

4.3.2. *Curved sections of the trajectory.* A curved section of the trajectory can form a transition between two straight sections, for a small change in aircraft heading, or form a steady-state curve in itself, i.e. for large changes in aircraft heading. Tracking a curved trajectory is fundamentally different from the following of a straight section, which was also recognized in the automobile driving research (Gordon 1966, Reid 1983).

Generally, two stages can be identified in tracking a curve, the *curve-anticipation* and the *curvature-tracking* phase. In the curve-anticipation stage, the aircraft is still on a straight section immediately preceding the curve. The pilot estimates the oncoming change in trajectory curvature and,

based on his internal representation, computes and initiates an open-loop feedforward control action to anticipate the curve. As a result of this control action, the vehicle will start to develop the required lateral acceleration to follow the oncoming curvature.
In the curvature-tracking stage, the aircraft is in the curve itself and the pilot tries to minimize the perceived curvature error, heading error and position error. While in the curve, an estimate of the heading error and position error is difficult, due to the lack of a symmetrical condition (Theunissen and Mulder 1994, 1995). Further note that when the tunnel image has a steady-state *dynamic* condition, the curvature error and heading error are both zero. This steady-state dynamic condition is probably the most prominent cue in curvi-linear motion, a fact which was hypothesized first by Gordon (Gordon 1966).
The curve-anticipation phase involves the *timing* of the feedforward control action and the estimation of the required *amplitude* of this control action. From a information-theoretical point-of-view (Theunissen and Mulder 1995), an oncoming change in curvature can be perceived from the tunnel display, except from very close to that part of the tunnel that is immediately near the own position. Previous research (Theunissen and Mulder 1994) has indicated that the pilot is able to extract *temporal* cues from the display.

4.3.3. *Implications.* From the discussion above it is clear that modelling pilot control behaviour along, or transitioning to, curved sections is more difficult than for straight sections of the trajectory. In tracking a constant curve, however, a steady-state does exist, so, equivalent to the tracking of straight sections, it should be possible to describe the control of this steady-state using the same, OCM-based, methodology.
Curve-anticipation involves the timing and amplitude estimation of a feedforward control action, based on the perceived change in the future trajectory to be followed. The next section will discuss a model, an extension to the OCM, which could account for this stage of operator behaviour.

5. TOMIZUKA's PREVIEW MODEL

Of all the operator preview models available in literature, the preview modelling approach of Tomizuka (Tomizuka and Whitney 1975, 1976) is probably the most sophisticated one. Instead of a preview model that is especially developed for a specific situation and resulting in satisfactory results in only that particular application, the model of Tomizuka is derived in the strict mathematical framework of optimal control. The resulting operator model is similar to the Optimal Control Model, but extends its operation to the use of future control-oriented information.

5.1. *The optimal finite preview problem*

Common feedback control gives in many cases satisfactory performance in tracking command signals and for compensating disturbances. One could expect, however, to achieve an even better performance (in some sense) when the controller would have some knowledge of the future course of the command signal and/or disturbances compared to the case when one would have no idea about the future. This kind of problem, where information about the future is available, is called the *preview* problem (Sheridan 1966).
It is important to know how to utilize the future information in an *optimal* fashion. Depending on the amount of a priori information of the command signal, the problem can be classified as follows (Tomizuka and Whitney 1975):

- *nothing* is known of the future course of the system forcing function. This leads to the conventional regulator control problem, in which the controller acts solely on the instantaneous discrepancy between actual and desired system output,
- *statistical knowledge is known* a priori, which is the problem considered here,
- *complete knowledge* of the future course of the forcing function, which is in fact the conventional optimal tracking problem.

To obtain a solution of the optimal preview problem, the future course of the forcing function is divided in two segments:

1. the part that the controller can preview, a part which is *deterministically* given,
2. the part that the controller cannot preview. Here, the only available a priori information is *statistical* in nature.

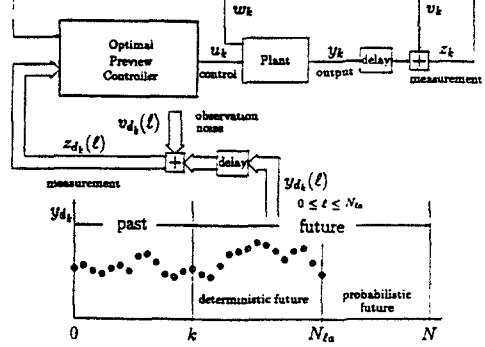


Fig. 6. Optimal finite preview problem (Tomizuka 1975).

Figure 6 illustrates the resulting optimal *finite* preview problem. The optimal preview controller is an optimal tracker which operates on the discrepancy between the actual and the desired plant output, but with the advantage of previewing a part of the future command signal trajectory.

5.2. *Mathematical model description*

In (Tomizuka and Whitney 1975, 1976), the finite preview control problem is formulated and solved within the framework of discrete-time stochastic optimal control theory. The solution, which will be discussed in very general terms only[3], follows closely the solution of the OCM.
The dynamics of plant $\mathcal{P}$ are given by:

$$x_{k+1} = Ax_k + Bu_k + Ew_k \quad (1)$$
$$y_k = Cx_k \quad (2)$$
$$z_k = y_{k-d} + v_k \quad (3)$$

Furthermore, the dynamics of the command generator $\mathcal{C}$ are given by:

$$x_{d_{k+1}} = A_d x_{d_k} + B_d w_{d_k} \quad (4)$$
$$y_{d_k} = C_d x_{d_k} \quad (5)$$
$$z_{d_k}(\ell) = y_{d_{k-d}}(\ell) + v_{d_k}(\ell) \quad (6)$$

where $y_{d_k}(\ell)$ represents the commanded output signal at ℓ time units in the future beyond time k $(0 \leq \ell \leq N_{\ell a})$.
As mentioned above, the optimal finite preview (OFP) controller is an optimal tracker which minimizes the error between the plant $\mathcal{P}$ output y_k and the output y_{d_k} of the command signal generator $\mathcal{C}$ over the control interval of interest $k \in [0, N]$. Following the OCM solution, a cost functional J is formulated, which combines the effect of the output error to be minimized as well as a control-rate input weighing to account for the operator's neuromotor dynamics:

$$J = E\left\{[y_k - y_{d_k}]^2 + g\left[\frac{u_k - u_{k-1}}{\Delta t}\right]^2\right\} \quad (7)$$

where Δt represents the discretization time.
The OFP problem can now be stated as to find an optimal control input u_k^{opt}, based on the available information (measurements z_k and $z_{d_k}(\ell)$), that minimizes the cost functional J.
The solution of the OFP problem is very similar to the solution of the OCM, as is illustrated in Figure 7. It involves the cascade combination of an optimal estimator (Kalman filter), an optimal least-mean-square predictor and a proportional state feedback. Due to the duality of the available information, the optimal controller can be described as the parallel combination of two independent channels:

- a *feedback* controller (L_x), which acts on the estimated system state. This channel is identical to the equalization network of the OCM.
- a *feedforward* controller (L_d and L_{x_d}), acting on the estimated previewed command generator output interval and command generator state respectively.

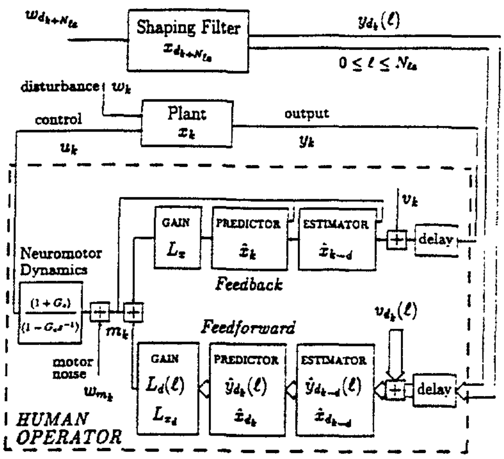


Fig. 7. Manual preview control model structure (Tomizuka 1976).

The OFP controller extends the conventional OCM with the availability to use previewed command trajectory information. Both models are able to deal with the main operator perceptual characteristics, such as observation noise, observation time delay, motor noise and neuromotor dynamics, in a conceptual manner.

5.3. *Application*

Tomizuka applied the optimal finite preview modelling theory to a conventional manual compensatory tracking task with preview (Tomizuka 1976). A one-axis compensatory tracking task was executed, in which the length of the preview interval was manipulated. As could have been expected, it was found that the preview drastically improved the tracking performance compared to zero-preview tracking.
Tomizuka's preview model was used to model the operator's control behaviour. The model parameters were adjusted to the values consistent to the published experimental data (Kleinman *et al.* 1970). Compared to the OCM, only the observation noise(s) $v_{d_k}(\ell)$ related to the previewed command signal interval $(0 \leq \ell \leq N_{\ell a})$ posed some difficulty, since no experimental data exists on how to choose these noise intensities. The performance of the model was reported to be in close agreement with experimental data.
In the next section the optimal finite preview model will be applied in the context of automobile driving.

[3] Note that for ease of reference the nomenclature of (Tomizuka and Whitney 1975, 1976) is maintained here.

6. APPLICATION OF TOMIZUKA's PREVIEW MODEL TO THE AUTOMOBILE DRIVING TASK

6.1. *Background*

It was discussed in section 3 that an automobile driver is generally modelled on two levels of control: The guidance level and the stabilization level. In the automobile driver model of Donges (Donges 1978), the guidance level consists of an anticipatory feedforward controller acting on the oncoming changes in the future trajectory curvature. The stabilization level consists of a compensatory feedback controller acting on the instantaneous deviations from the trajectory.
Considering the structures of Donges' two-level model of driver steering behaviour (Figure 3) and Tomizuka's preview model (Figure 7), the similarity is evident. In this section, the preview model will be used to reproduce some of the characteristics of the Donges model. The optimal control approach, however, requires the determination of the perceptual characteristics of the operator, in terms of observation and motor noise intensities and the neuromotor dynamics. These model parameters, however, are not relevant in the classical approach followed by Donges, so they are not available here. Therefore, the discussion will be restricted to the perfect measurement case.

6.2. *Model description*

The concept of Donges' model has been inserted in the OFP operator model structure. The curvature of the trajectory to be followed is taken as the output of the command signal generator. The feedforward channel of the OFP model therefore operates solely on the future trajectory curvature. The feedback channel operates on the instantaneous deviations from the reference trajectory, i.e. the curvature error, the heading error and the lateral position error.

6.2.1. *Vehicle dynamics.* The vehicle dynamics (Donges 1978) represent the highly simplified lateral dynamics of an automobile. The driver's steering wheel angle λ_{sw} (in [rad]) results in a change in the curvature K_i (in [m^{-1}]) of the vehicle:

$$K_i(t) = E_L \lambda_{sw}(t - \tau_F) \quad (8)$$

where $E_L = 0.049$ [$m^{-1}rad^{-1}$] depicts the steady-state steering sensitivity (independent of vehicle velocity V ([m/s]) and $\tau_F = 0.2$ [sec] is the vehicle's time delay.
The curvature error K_Δ is defined as the difference between the actual vehicle curvature K_i and the commanded vehicle curvature K_s:

$$K_\Delta(t) = K_i(t) - K_s(t) \quad (9)$$

The following kinematic equations describe the relations between the curvature error K_Δ and the heading angle error ψ_Δ ([rad]) and lateral deviation error y_Δ ([m]) respectively:

$$\psi_\Delta(t) = \psi_\Delta(t_0) + \int_{t_0}^{t} V(\theta)K_\Delta(\theta)d\theta \quad (10)$$

$$y_\Delta(t) = y_\Delta(t_0) + \int_{t_0}^{t} V(\theta)\psi_\Delta(\theta)d\theta \quad (11)$$

In other words, the vehicle dynamics can be classified as a double integrator with a time delay. In the following, the velocity of the vehicle V is fixed at 50 km/h (13.9 m/s).

6.2.2. *OFP model parameters.* Because the discussion will be restricted to the perfect measurement case, only four model parameters must be set. The driver's time delay τ_d is fixed at 0.25 seconds, while the neuromotor lag time constant τ_N is varied between 0.1 and 0.3 seconds. These values correspond well to the published OCM data (e.g. Kleinman *et al.* 1970) and also with the experimentally found time delay reported by Donges, which lied between 0.35 and 0.55 seconds[4]. The vehicle disturbance noise intensity was set to zero, while the command signal shaping filter[5] white noise input intensity was varied to obtain reasonable curvatures.

6.3. *Results*

6.3.1. *Effect of preview.* Figure 8 shows the feedforward gains $L_d(\ell)$, over a preview interval of 1.5 sec., as a function of the neuromotor lag constant τ_N. The shape of the feedforward gains appears very similar to a well-damped impulse response. In other words, the OFP feedforward channel can be seen as what Sheridan classified as an *extended convolution* preview controller (Sheridan 1966).

Several remarks can be made. First of all, the feedforward gains are zero for the time interval between 0.0 and 0.25 seconds, apparently due to the operator's dead time τ_d. The effect of the neuromotor lag time is obvious: When the driver is willing to initiate fast, high-amplitude control actions, only the near future is important. In other words, he does not need much anticipation time. These fast control actions, however, lead to high vehicle lateral accelerations which affects the comfort of the manoeuvre. To decrease these accelerations, the driver will smoothen his control actions (decreases the open-loop driver-vehicle cross-over frequency), and as a result he has to look further

[4] Note that the time delay of Donges can be seen as the sum of the OFP's time delay and neuromotor lag constant.

[5] The command shaping filter was a second order low-pass filter with adjustable bandwidth and a fixed damping of 1.0.

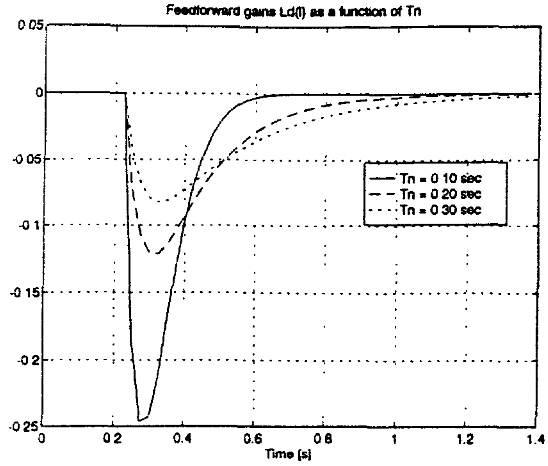


Fig. 8. Feedforward gains as a function of preview time.

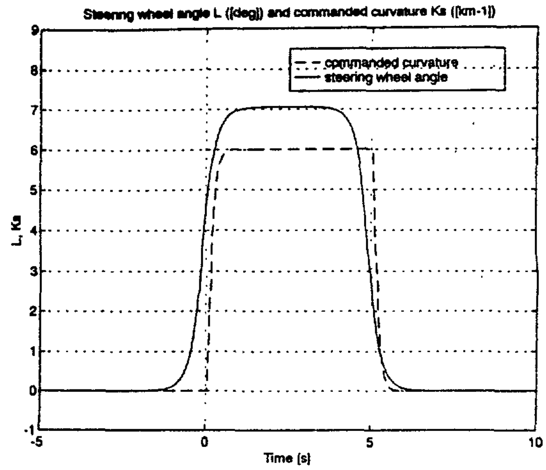


Fig. 10. Time histories of commanded curvature and control action of the OFP controller.

ahead to maintain the necessary feedforward control action. In the model this is reflected in an increasing τ_N[6].

Donges (Donges 1978) reports an experimentally found maximum lateral acceleration of approximately 0.3 g, which corresponds with a neuromotor lag constant between 0.2 and 0.3 seconds. As can be seen from Figure 8, this means a minimum anticipation time between 1.0 and 1.4 seconds, which corresponds well to the experimentally found value of 1.1 second. Note that this anticipation time of the OFP controller is indeed independent of the vehicle's velocity, which is a result of the simplified vehicle dynamics.

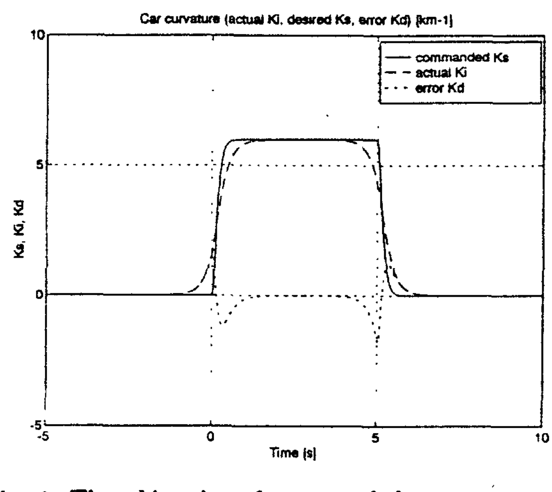


Fig. 9. Time histories of commanded curvature, actual curvature and curvature error for the OFP controller.

Figure 9 shows some simulation time histories. The commanded curvature is a block with amplitude of 6 km^{-1}, corresponding to the experiments of Donges. As can be seen from this figure, the OFP controller anticipates and follows the commanded change in curvature very well. Figure 10 shows the commanded curvature together with the control action of the OFP model. The figure clearly illustrates the anticipatory character of the control action, which is initiated approximately 1.0 to 1.5 seconds before the actual curvature change occurs.

[6] One should keep in mind that it is not the driver's neuromotor system that changes, but his bandwidth, a confusion common when using OCM-based pilot models

6.3.2. *Comparison to a no-preview controller.* Figures 11 and 12 show simulation time histories for the zero-preview controller together with the OFP controller. Obviously, the performance of the preview controller is superior. The zero-preview controller has no ability to anticipate the oncoming trajectory curvature and is therefore, contrary to the OFP controller, unable to mimic the experimentally found driver steering behaviour.

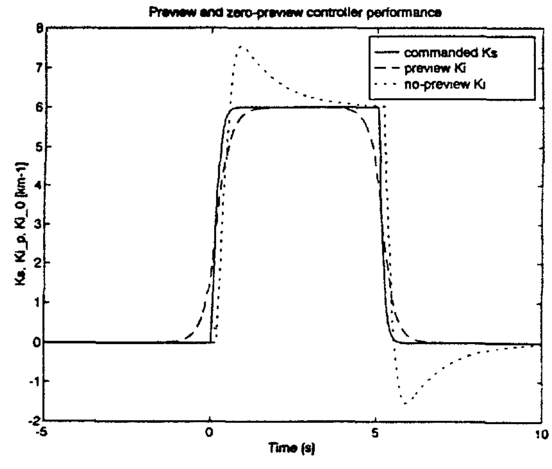


Fig. 11. Time histories of commanded curvature, actual curvature and curvature error for the OFP and *zero-preview controllers.*

6.4. *Conclusion*

The optimal finite preview operator model has been used to reproduce some of the characteristic elements of automobile driver behaviour as reported in (Donges 1978). Although the demon-

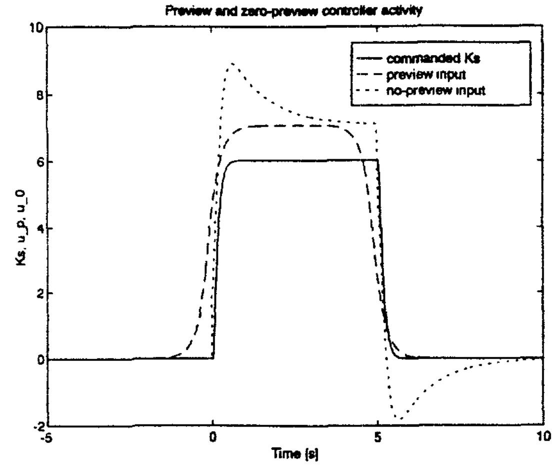


Fig. 12. Time histories of commanded curvature and control action of the OFP and zero-preview controllers.

stration was restricted to the perfect measurement case, the OFP operator model gave satisfactory results. Especially the parallel operation of the anticipatory feedforward and the compensatory feedback information-processing channels shows a strong resemblance to most of the lane-tracking automobile driver models. An advantage of the OFP modelling methodology is that it gives a conceptual insight into what part(s) of the future are most relevant to control. It is postulated that the OFP model could as well reproduce driver behaviour on lane-change or obstacle-avoidance tasks, in which the anticipatory channel operates directly on perceived future commanded vehicle position instead on trajectory curvature.

7. DISCUSSION/CONCLUSIONS

A research project has been initiated to understand and model the manual control behaviour of a pilot controlling the aircraft along the reference trajectory using a perspective flight-path display. This paper has discussed some important issues of obtaining such a model.

A literature survey revealed that there exist strong resemblances between the pilot's guidance task with a tunnel-in-the-sky display and the common task of automobile driving. Based on this fact and on a theoretical analysis of some elements of the pilot's guidance task, possible modelling solutions are suggested.

On straight sections of the trajectory it is postulated that the well-known Optimal Control Model, enhanced with an adequate observation model can provide a basis for the modelling efforts. The OCM approach can give meaningful insight into the main characteristics of pilot control behaviour with a perspective display.

Modelling the pilot on tracking curved trajectories poses more difficulties, since it requires a transition response instead of a steady-state regulator. In the automobile driving research this distinction was also recognized and lead to multi-level driver models, in which the driver is modelled as a combination of an anticipatory and a compensatory controller. The anticipatory part acts as a feedforward controller on some perceived element (usually the road curvature) of the future trajectory to be followed.

A model has been presented in which the effects of the previewed trajectory can be incorporated into the conventional OCM model structure: The so-called optimal finite preview (OFP) controller, postulated by Tomizuka (Tomizuka and Whitney 1975, 1976). Because at this time no experimental results are available, the OFP controller has been applied to reproduce some of the characteristic elements of driver steering behaviour reported by Donges (Donges 1978). The OFP model showed the same properties of the two-level model of Donges by incorporating the parallel operation of an anticipatory feedforward and a compensatory feedback information-processing channel. Moreover, one of the advantages of the OFP modelling approach is that it gives meaningful insight into the useability of the future control-oriented information. It is postulated here that the OFP controller can be used to model the pilot in following the transients in the reference trajectory with a perspective flight-path display.

In the near future, the OCM and OFP models will be validated in pilot-in-the-loop experiments. When the basic model structure has been validated, it should be able to predict the effects of changing primary display design parameters of the perspective display as well as the effects of additional display symbology.

8. ACKNOWLEDGEMENTS

The author would like to thank Erik Theunissen for his cooperation and support in the present research project[7].

9. REFERENCES

1. Blaauw, G.J., (1984), *Car Driving as a Supervisory Control Task*, PhD thesis, Institute for Perception TNO, Soesterberg, The Netherlands.
2. Donges, E., (1978), *A Two-Level Model of Driver Steering Behaviour*, Human Factors, Vol. 20, No. 6.
3. Gibson, J.J., (1950), *The Perception of the Visual World*, Houghton-Mifflin, Boston.

[7] Further, the author is grateful to *Sonic Youth* for helping him to stay awake at 3am while finishing this last-minute paper.

4. Gordon, D.A., (1966), *Perceptual Basis of Vehicular Guidance*, Public Roads, Vol. 34, No. 3.
5. Grunwald, A.J., Merhav, S.J., (1976), *Vehicular Control by Visual Field Cues -Analytical Model and Experimental Validation*, IEEE-SMC, Vol. SMC-6, No. 12.
6. Grunwald, A.J., Merhav, S.J., (1978), *Effectiveness of Basic Augmentation in Vehicular Control by Visual Field Cues*, IEEE-SMC, Vol. SMC-8, No. 9.
7. Grunwald, A.J., Robertson, J.B., Hatfield, J.J., (1981), *Evaluation of a Computer-Generated Perspective Tunnel Display for Flight-Path Following*, NASA TP 1736.
8. Guo, K., Fancher, P.S., (1983), *Preview-Follower Method for Modeling Closed-Loop Vehicle Directional Control*, Proc. of the 19th Annual Conference on Manual Control.
9. Haskell, I.D., Wickens, C.D., (1993), *Two- and Three-Dimensional Displays for Aviation: A Theoretical and Empirical Comparison*, The International Journal of Aviation Psychology, Vol. 3, No. 2.
10. Hess, R.A., (1981), *Pursuit Tracking and Higher Levels of Skill Development in the Human Pilot*, IEEE-SMC, Vol. SMC-11, No. 4.
11. Hess, R.A., Modjtahedzadeh, A., (1989), *A Preview Control Model of Driver Steering Behaviour*, Proc. of the IEEE Conference on Systems, Man and Cybernetics, 1989.
12. Kleinman, D.L., Baron, S., Levison, W.H., (1970), *An Optimal Model of Human Response -Part I: Theory and Validation, Part II: Prediction of Human Response in a Complex Task*, Automatica, Vol. 6.
13. Kondo, M., Ajimine, A., (1968), *Driver's Sightpoint and Dynamics of the Driver-Vehicle-System Related to it*, SAE 680104.
14. Korn, J., Gully, S.W., Kleinman, D.L., (1982), *Validation of an Advanced Cockpit Design Methodology via Workload/Monitoring Tradeoff Analysis*, Proc. of the 18th Annual Conference on Manual Control.
15. MacAdam, C.C., (1981), *Application of an Optimal Preview Control for Simulation of Closed-Loop Automobile Driving*, IEEE-SMC, Vol. SMC-11, No. 6.
16. McRuer, D.T., Allen, R.W., Weir, D.H., Klein, R.H., (1977), *New Results in Driver Steering Control Models*, Human Factors, Vol. 19, No. 4.
17. Mulder, M., (1994), *Displays, Perception and Aircraft Control - A literature survey of theory and modelling of pilot behaviour with spatial instruments*, Report LR-762, Faculty of Aerospace Engineering, Delft University of Technology.
18. Reid, L.D., (1983), *A Survey of Recent Driver Steering Behaviour Models Suited to Accident Studies*, Accident Analysis and Prevention, Vol. 15, No. 1.
19. Sheridan, T.B., (1966), *Three Models of Preview Control*, IEEE-HFE, Vol. HFE-7,No.2.
20. Shinar, D., McDowell, E.D., Rockwell, T.H., (1977), *Eye Movements in Curve Negotiation*, Human Factors, Vol. 19, No. 1.
21. Theunissen, E., (1993), *A Primary Flight Display for Four-Dimensional Navigation - Influence of Tunnel Size and Level of Additional Information on Pilot Performance and Control Behaviour*, Proc. of the AIAA '93 FST Conference, Monterey, CA.
22. Theunissen, E., (1994), *Factors Influencing the Design of Perspective Flight-Path Displays for Guidance and Navigation*, Displays, Vol. 15, No. 4.
23. Theunissen, E., Mulder, M., (1994), *Open and Closed Loop Control with a Perspective Tunnel-in-the-Sky Display*, Proc. of the AIAA '94 FST Conference, Scottsdale, AZ.
24. Theunissen, E., Mulder, M., (1995), *Availability and Use of Information in Perspective Flightpath Displays*, Proc. of the AIAA '95 FST Conference, Baltimore, MY.
25. Tomizuka, M., Whitney, D.E., (1975), *Optimal Discrete Finite Preview Problems (Why and How is Future Information Important?)*, Journal of Dynamic Systems, Measurement, and Control.
26. Tomizuka, M., Whitney, D.E., (1976), *The Human Operator in Manual Preview Tracking (an Experiment and its Modelling via Optimal Control)*, Journal of Dynamic Systems, Measurement, and Control.
27. Wewerinke, P.H., (1989), *Models of the Human Observer and Controller of a Dynamic System*, PhD thesis, University of Twente, Enschede.
28. Wickens, C.D., Todd, S., Seidler, K., (1990), *Three-Dimensional Displays: Perception, Implementation, and Application*, AMRL-TR-90.
29. Wierwille, W.W., Gagné, G.A., Knight, J.A., (1967), *An Experimental Study of Human Operator Models and Closed-Loop Analysis Methods for High-Speed Automobile Driving*, IEEE-HFE, Vol. HFE-8, No. 3.
30. Wilckens, V., (1971), *On the Dependence of Information Display Quality Requirements Upon Human Characteristics and Pilot/Automatics-Relations*, Proc. of the 7th Annual Conference on Manual Control, NASA SP-281.

PILOT-IN-THE-LOOP STUDIES INTO MANUAL CONTROL STRATEGIES WITH PERSPECTIVE FLIGHTPATH DISPLAYS

Erik Theunissen
Delft University of Technology, Faculty of Electrical Engineering
P.O. Box 5031, 2600 GA Delft, The Netherlands (e.theunissen@et.tudelft.nl)

and

Max Mulder
Delft University of Technology, Faculty of Aerospace Engineering
P.O. Box 5058, 2600 GB Delft, The Netherlands (m.mulder@lr.tudelft.nl)

Abstract. Conventional flight directors are typical command displays, requiring the pilot to apply a closed-loop compensatory control strategy. In contrast, perspective flightpath display present status information, and as a result allow a wider variety of control strategies to be applied. This paper addresses the different control strategies which are possible with perspective flightpath displays, and discusses two experiments which have been performed to gain more insight into compensatory and error-neglecting control with perspective flightpath displays.

Keywords. perspective flightpath displays, aircraft guidance, manual control.

1. INTRODUCTION

Navigation can be defined as "to direct the course of an aircraft". The guidance task comprises the control of the aircraft to keep position and velocity errors within the constraints specified by the navigation performance requirements. The conventional instrument for the guidance task is the flight director, presenting steering commands.

An alternative guidance display is the perspective flightpath display. A perspective flightpath display presents an integrated view of the desired trajectory specified in the three spatial dimensions on a two-dimensional display. Due to their spatial nature, perspective flightpath displays present navigation and guidance data in a way which is fundamentally different from conventional planar data formats used today, which in turn is likely to influence pilot control strategies. In a study into 2-D and 3-D displays for aviation, Haskell and Wickens (1993) report that the way in which a task was performed differed as a function of the displays employed. The importance of the fact that different displays can result in qualitative, strategic differences is stressed by pointing out that "when making empirical comparisons between different display types, researchers must evaluate measures other than performance on only one type of task; they must go beyond performance in any case and examine task performance strategies".

At Delft University of Technology, research into perspective flightpath displays for guidance and navigation is performed in the context of the Delft Program for Hybridized Instrumentation and Navigation Systems (DELPHINS). Figure 1 presents an example of the DELPHINS Tunnel-in-the-Sky display.

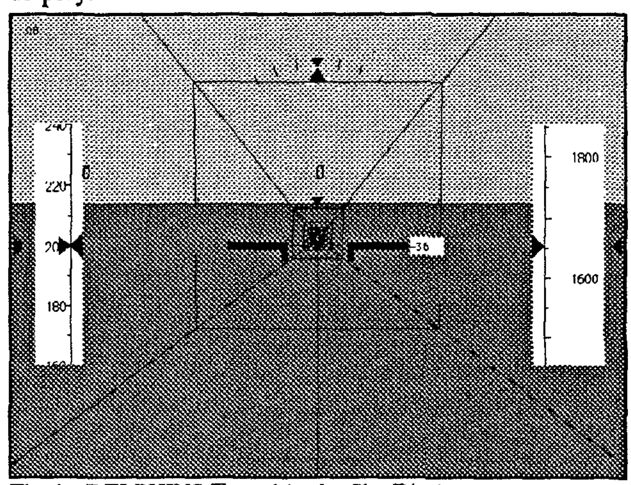


Fig. 1. DELPHINS Tunnel-in-the-Sky Display

To investigate the different control strategies and the influence of several display augmentation concepts which are possible with perspective flightpath displays, pilot-in-the-loop studies have been performed. This paper discusses the results obtained from two pilot-in-the-loop studies in the context of the specific aspects of the data presentation which allow the different control strategies to be applied.

2. GUIDANCE DISPLAYS

Flight director commands are based on a weighted combination of position and angular errors, presented in one dimension. As a result of the integration of multiple parameters into a single dimension, the pilot is unable to extract information about the specific errors from the flight director display. Furthermore, since the error-gains of the display are determined by the flight director algorithms, the possible bandwidth the pilot can apply for scanning and executing the flight director commands is very limited. Finally, the data which is required to maintain adequate spatial and navigational awareness requires the scanning of several other instruments, while the integration of this data has to be performed by the pilot. This process involves mental rotation and scaling operations, which costs time and may introduce errors.

With a perspective flightpath display, the spatial presentation of the imaginary flightpath in the three-dimensional environment can be used to combine guidance data with the data required for spatial and navigational awareness, thus alleviating the pilot from scanning several instruments and performing the mental integrations of the separately displayed position and orientation data into a spatially coherent picture.

With such a display, the pilot is required to fly through a synthetic tunnel which is a representation of his desired three-dimensional flight-path.

Perspective flightpath displays have been discussed since the early fifties, and various concepts for aircraft guidance have been evaluated in simulation (Wilckens and Schattenmann, 1968; Grunwald, 1984; Wickens et al., 1989; Theunissen, 1993), some even in actual flight (Filarsky and Hoover, 1983; Theunissen, 1995).

3. TUNNEL-IN-THE-SKY DISPLAYS

Figure 2 presents a line-drawing of the DELPHINS Tunnel-in-the-Sky display. In this display, the desired flightpath is indicated by the tunnel. In (Theunissen, 1994) it is illustrated how information about position and orientation errors can be extracted from the distortion of the symmetrical shape of the tunnel. The moving horizon presents attitude, while heading information is presented on the horizon line. Altitude, airspeed, and bank are displayed by means of separate indicators. To avoid distortions between the perspective presentation of the three-dimensional flightpath and the attitude presentation, the visible pitch attitude range

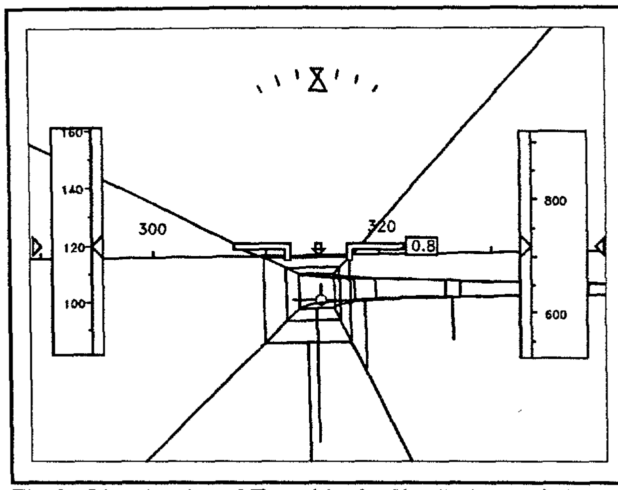


Fig. 2. Line drawing of Tunnel-in-the-Sky display symbology

corresponds to the geometric vertical field of view. To accommodate the fourth dimension, reference speed is presented by means of a bug on the speed-tape. The display also provides the possibility to present integrated speed information by means of a moving window in the tunnel.

Additional information can be presented to aid the pilot with the guidance task. This information has been divided into three levels: Unprocessed status information, processed status information, and command information.

3.1. *Unprocessed status information*

To aid the pilot in maintaining the correct flightpath angle, a flightpath vector can be presented (Figure 3).

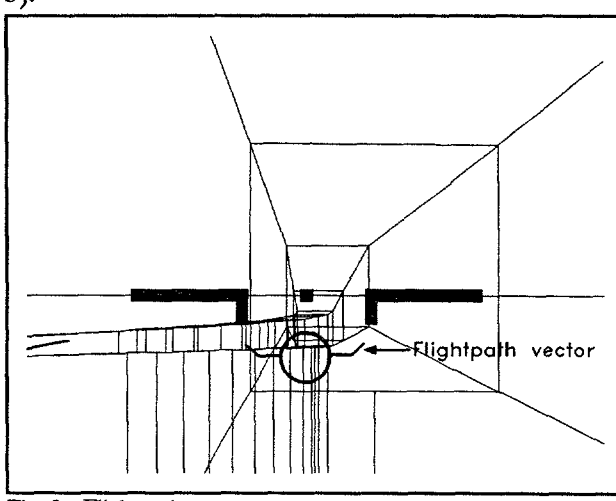


Fig. 3. Flightpath vector

This vector indicates the current direction of the velocity vector of the aircraft relative to the aircraft attitude symbol. Because a flightpath vector presents raw data, it is classified as unprocessed status information.

3.2. *Processed status information*

An airplane is a higher order dynamic system, and the pilot has to determine his control actions by predicting the future system state as a result of his actions. To aid the pilot, a predictor symbol indicating the future position and bank angle can be displayed. To resolve the position ambiguity of the predictor symbol, the cross-section of the tunnel at the position of the predictor is indicated by a transparent window (Figure 4). A predictor is classified as processed status information.

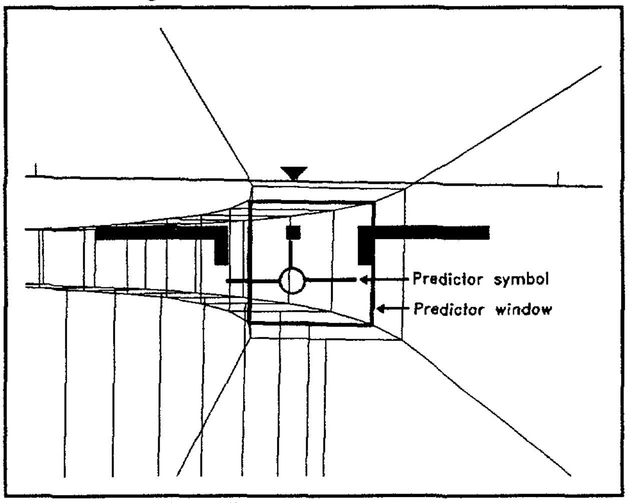


Fig. 4. Flightpath predictor

The size of the tunnel determines the size of the reference window. Thus, with decreasing tunnel size, the resolution of the reference increases. The gain of the position error data, however, is not affected. Both the predictor reference window and the predictor move. The task of keeping the predictor in the center of the reference window is a pursuit tracking task.

3.3. *Command information*

Instead of presenting status information to aid the pilot controlling the aircraft, command information can be presented, e.g. by means of a flight-director.

4. CONTROL STRATEGIES

Ample research has been performed on human control behaviour in compensatory tracking tasks (McRuer et al., 1965). Perspective flightpath displays however, present the pilot with integrated trajectory preview combined with an indication of the allowed deviations, and research into pilot control behaviour when presented with this kind of information is relatively scarce. In (Mulder, 1994) an extensive literature review about the modelling of pilot control behaviour with spatial displays is presented.

With car driving the situation is different. Various models have been proposed to describe driver control behaviour in relation with the visual environment. Since the nature of the control task (boundary control) and the visual cues are quite similar for the guidance task with a perspective flightpath display and car driving, it is expected that there also is a similarity in control strategies.

Concerning car driving, McRuer et al. (1977) present an approach in which they distinguish between compensatory, pursuit and dual mode control behaviour. With compensatory control, the driver uses lateral position and heading errors. With pursuit control the driver takes advantage of the trajectory preview to initiate an open-loop control action to follow the desired path, i.e. the driver applies feedforward control. With dual mode behaviour, the driver initiates an open-loop control action which is succeeded by closed-loop compensatory control.

Gordon (1966) states that "The behaviour involved in steering an automobile has usually been misunderstood. It is less a matter of aligning the car with the road than it is a matter of keeping the focus of expansion in the direction one must go". The velocity field provides information on the speed and direction of the vehicle's forward motion. The driver may become aware of the misalignment of the car by slewing shifts in direction, and by side-slipping sidewise movements which exceed the human visual position and movement thresholds. The driver's perceptual response is based upon an integration of these and other sources of information.

On the basis of human perception theory, it is difficult to determine which of the combinations of slew, sideslip, rate, and amplitude the driver perceives. The driver responds to a total situation, not to isolated or ranked cues. This indicates the necessity of determining a single parameter to describe and predict driver responses. Godthelp (1984) introduced the so-called Time-to-Line Crossing concept, which is based on the assumption that there is a relation between the remaining time the vehicle under control is within a certain boundary, and the moment a control action is initiated.

Most of the available vehicle control models are based on the fundamental assumption that drivers control their vehicle with permanent visual feedback. However, as it is commonly accepted, visual feedback is sometimes interrupted. Godthelp (1984) investigated the potential role of visually open-loop strategies and error-neglection in vehicle control. He assumed that the time available for a driver to

control his vehicle in an open-loop mode largely depends on the accuracy of the open-loop generated steering-wheel action and the time available for error-neglection.
The control activity indicates the amount of effort invested in the control task. For continuous closed-loop control tasks, frequency domain techniques are very useful for describing control behaviour. However, for non-continuous control behaviour encountered during error-neglection and open-loop control, time domain techniques may be more appropriate.

5. SIMULATOR EVALUATION

As indicated in the previous section, it is expected that a similarity in control strategies between car driving and flying a tunnel-in-the-sky display exists. To evaluate pilot performance and control behaviour and obtain suitable values for the design parameters of the perspective flightpath display, several experiments have been conducted in the moving-base flight simulator at Delft University of Technology. In an experiment performed in 1993, closed-loop compensatory control was investigated. In 1994, error-neglecting control strategies were examined. The simulated aircraft was a twin-engine business jet, which is also used for the in-flight experiments.

5.1. *Experiment I*

In 1993 pilot performance and control behaviour when flying a Tunnel-in-the-Sky with the addition of a flightpath vector (FPV) and with a flightpath position predictor (FPP) for different error gains was investigated (Theunissen, 1993).

5.1.1. Experimental setup

Five pilots, of whom two student pilots and one none-pilot, participated in the experiment. The study consisted of a 3x2 within subject design. Pilots flew three different tunnels (22.5, 45, and 90 m width), in two different configurations (FPV, FPP). Each condition was replicated five times, resulting in a total of 30 flights for each pilot. The order in which the tunnels were presented in a certain configuration was balanced to be able to compensate for possible learning effects.
Pilots started their flight at an altitude of 1200 ft about 4 miles away from the runway threshold. The task of the pilot was to fly the curved approach as accurate as possible using the Tunnel-in-the-Sky display, and land the aircraft. Pilots were required to maintain an airspeed of 120 knots. The airspeed was indicated by a green bug on the speed-tape. No additional speed cues were presented in the display. At the beginning of the flight, the aircraft was already in the landing configuration, so no aircraft configuration changes had to be made by the pilot.
Before the experiment started, pilots were briefed on the display and the approach. After the briefing, the training sessions started. To reduce the learning effect, pilots performed eight flights in each display configuration. The standard deviation of their horizontal and vertical path error was calculated for these flights and used as a measure of performance. If performance still appeared to improve after the first eight training flights, more training flights were issued.

5.1.2. Results

Results showed that both in the FPV and the FPP configuration, tracking accuracy increased linearly with decreasing tunnel size (Figure 5).

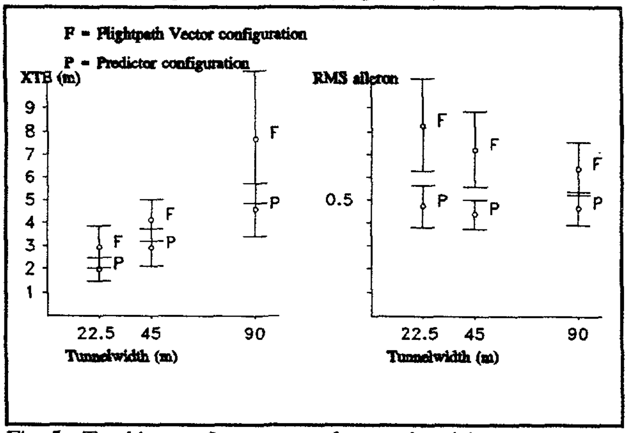


Fig. 5. Tracking performance and control activity

With the FPV, control activity was linearly related to error gain, whereas with the FPP no significant difference in control behaviour was found for the different tunnel dimensions. Figure 6 shows the average *XTE* for the five consecutive segments of the approach.

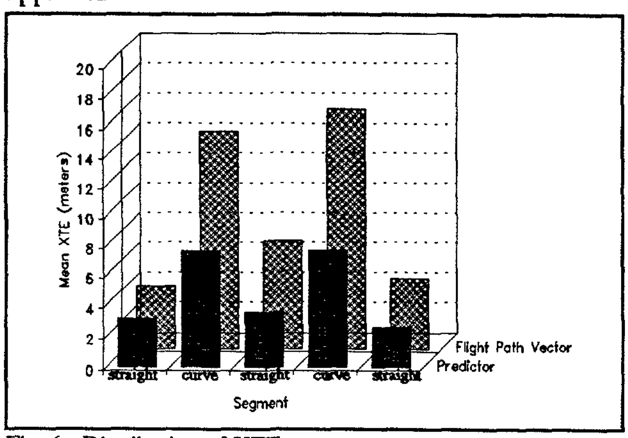


Fig. 6. Distribution of XTE

The second and fourth bar present the results for the curved sections, and show a decrease in tracking accuracy. This can partly be contributed to the increasing difficulty of the control task, and the fact that it is impossible to perfectly transition from a straight segment to a circular one. Since in a curve a FPV presents no adequate lateral guidance cues, this will also decrease tracking accuracy.
To illustrate the difference between control activity with the FPV and with the FPP, Figure 7 presents a cumulative distribution of the aileron deflections.

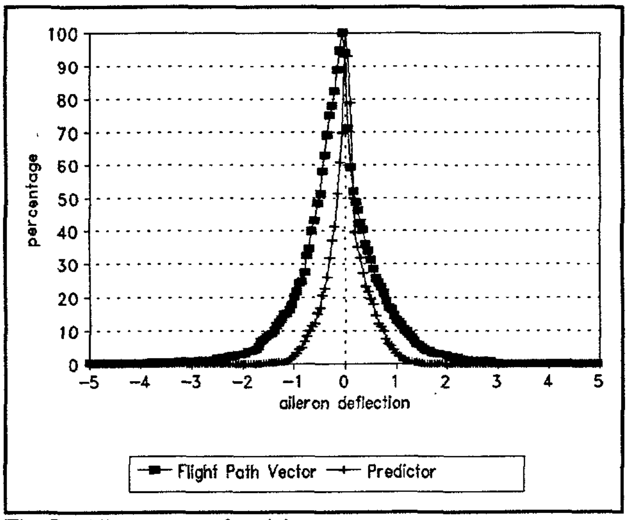


Fig. 7. Aileron control activity

As can be seen from this Figure, approximately 20% of the deflections made in the FPV configuration exceeds the maximum deflections made in the FPP configuration. The fact that in the FPP configuration control activity does not significantly increase with an increase in position error gain, suggests that in the presence of an adequate prediction of the future position and attitude, the pilot does not use the error information presented by the tunnel, but only the error presented by the predictor for the control task. Thus, when the pilot is told to fly as accurate as possible, he is likely to use the information with the highest error gain he can process to perform this task. In case of an additional flight director or predictor, the pilot will mainly concentrate on the data presented by this indicator, and control behaviour will be dominated by closed-loop compensatory and pursuit control respectively. The perspective presentation of the flightpath provides the information which allows the pilot to anticipate changes in the trajectory. When the task relaxes to maintaining the position error below the thresholds indicated by the walls of the tunnel, a shift towards more open-loop control strategies is possible.

5.2. *Experiment II*

In a follow-on study, error-neglecting control strategies were investigated (Theunissen and Mulder, 1994). The goal of this study was to determine what causes the pilot to initiate error-corrective actions. It was hypothesized that the moment an error-corrective action is initiated is strongly related to the time remaining before the aircraft crosses one of the boundaries indicated by the tunnel walls, and that the pilot uses temporal range information from the display for his decision to intervene.

5.2.1. Experimental setup

Five subjects, all airline pilots, were instructed to fly an approach to landing. After several training sessions, each pilot performed thirty approaches. To prevent the pilots from becoming accustomed to a particular approach, six different approaches were presented in a random order. To prevent them from applying a dominantly closed-loop compensatory control strategy, they were explicitly instructed that the goal was not to fly as accurate as possible, but to remain inside the tunnel using minimal control effort. A relatively low error gain was used by presenting tunnels with a width of 135m.

5.2.2. Data analysis

Data analysis was performed for data relevant to aircraft control in the lateral-horizontal plane, both for a first and a second order *TWC* model. The assumption for the first order model is that the pilot does not use a yaw component in his estimate, and consequently assumes a straight trajectory. This is comparable to the Time-To-Contact (Lee, 1976) and Time-To-Passage (Kaiser and Mowafy, 1993) models. Equation 1 presents the second order model:

$$TWC = \frac{\cos^{-1}\left(\cos(TAE) - \frac{r\left(\frac{width}{2} - XTE\right)}{V}\right)}{r} + \frac{TAE}{r} \quad (1)$$

In Equation 1 *width* represents the tunnel width [m], *XTE* the cross-track error [m], *TAE* the track-angle error, *V* the velocity [m/s] and *r* the yaw rate [rad/sec].
For both models, the consistency between the direction of the control actions and the prediction of the tunnel intersection (left or right) was analyzed. When the model predicts an intersection of the left tunnel wall, and the pilot initiates an error corrective action to the left, the outcome of the model is regarded as inconsistent with pilot control behaviour. At the time a control action was identified as an

error-corrective control action, all variables of interest (*XTE*, *TAE* and *TWC*) were recorded.

5.2.3. Results and discussion

Figure 8 and 9 present a distribution of the *XTE* and *TAE* variables respectively, at the moment an error-corrective control action was initiated.

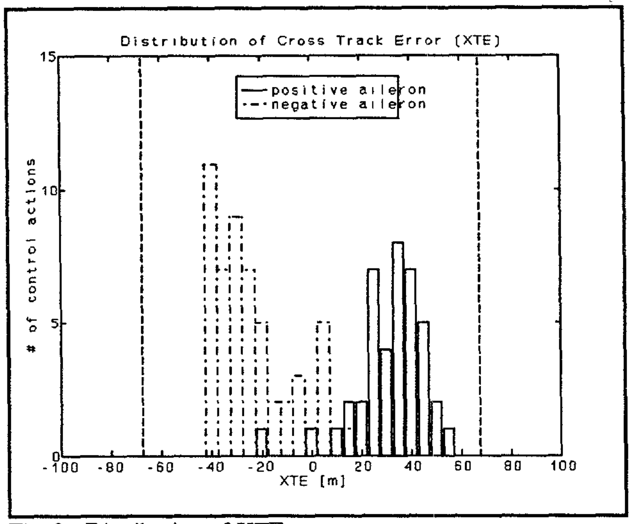


Fig. 8. Distribution of XTE

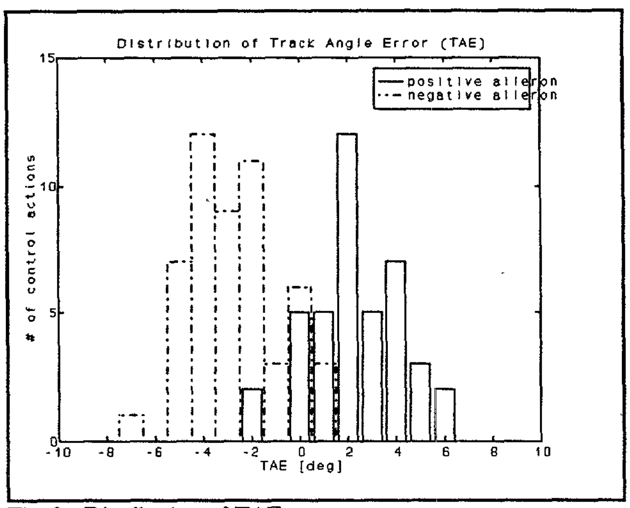


Fig. 9. Distribution of TAE

These figures indicate that there exists a large variation between the magnitudes of these variables and the number of initiated control actions. Furthermore, since no minimum threshold can be established in these distributions, it can be concluded that no individual guidance variable is solely responsible for switching from error-neglecting to error-correcting control, which strengthens the hypothesis of an integrated parameter.

In the curved segments the first order model produced completely inconsistent predictions, whereas the second order model was highly compatible with the direction of the control actions performed by the pilot. On the straight sections, both the first and the second order model predicted compatible control directions. The results showed that the *TWC* estimates of the second order model yielded a significantly smaller standard deviation as compared to the first order model. The first order model often (>50%) produced *TWC* estimates which exceeded 20 seconds, and it was concluded that the pilot does take yaw into account on the straight segments.

When examining the distribution of the *TWC* (Figure 10), it can be seen that no control actions were made for *TWC* values smaller than approximately 4 to 5 seconds.

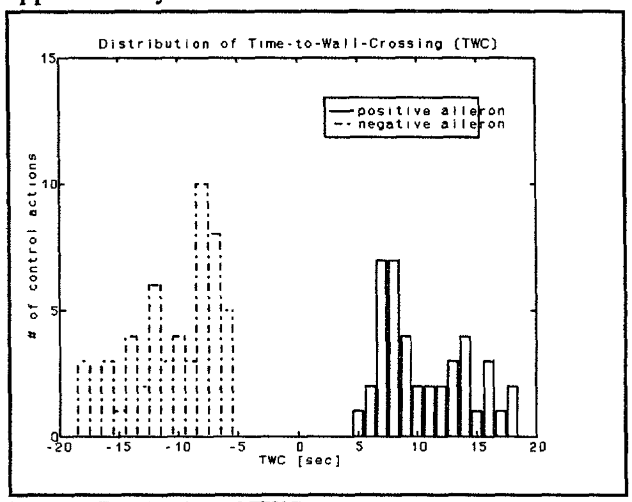


Fig. 10. Distribution of TWC

This strengthens the hypothesis that pilots maintain a certain temporal spacing from the boundaries represented by the tunnel walls, which they directly perceive from the display.

A statistical analysis (non-parametric Kolmogorov-Smirnov), showed that none of the distributions were from a normal population. Furthermore, no statistically significant differences were found between the distributions of control actions related to preventing crossing the right or left tunnel walls, allowing both distributions to be combined. As can be seen from Figure 11, showing box-plots of the combined distributions, the error-corrective control actions are initiated for a wide range of individual guidance variables.

For the *TWC* variable, however, the ratio between the standard deviation and the mean remains rather small, supporting the hypothesis that pilots maintain a certain temporal spacing from the boundaries represented by the tunnel walls, which they directly perceive from the display. The temporal spacing varies between pilots, and is believed to be determined by a self-chosen safety margin which, in turn, is largely determined by the familiarity the pilot has with the airplane and its handling qualities.

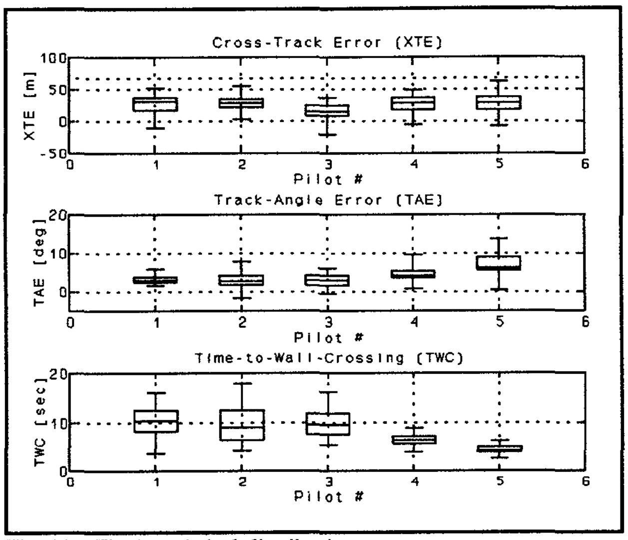


Fig. 11. Final statistical distributions

6. CONCLUSION

The first study illustrates that the size of the perspective flightpath can be used to help the pilot obtain a certain required performance. Introduction of the flightpath predictor can be used to increase the pilots' performance, while reducing control activity. Such a display combines the best properties of two concepts. It presents integrated trajectory, position and attitude information which contributes to the pilots spatial and navigational awareness and increases his level of confidence, and it presents processed status information which allows him to spend less effort to control the system.

The second study illustrates that one of the advantages of the perspective flightpath display is that, due to its integrated presentation, pilots do not have to mentally integrate the values of position and angular errors and error rates and verify whether the outcome exceeds a certain threshold, which would be required for error-neglecting control with non-integrated displays. Instead, the ego-referenced spatial presentation of guidance data allows pilots to extract temporal range information which enables them to apply an error-neglecting control strategy.

7. REFERENCES

Filarsky, S.M. and Hoover, S.W. (1983) *The command flight path display*, Naval Air Development Centre, Warminster, PA.

Godthelp, H. (1984) *Studies on Human Vehicle Control*, PhD Thesis, Institute for Perception TNO, Soesterberg, The Netherlands.

Gordon, D.A. (1966) 'Perceptual Basis of Vehicular Guidance' *Public Roads*, Vol. 34, No. 3, pp. 53-68.

Grunwald, A.J. (1984) 'Tunnel Display for Four-Dimensional Fixed-Wing Aircraft Approaches', *J. Guidance*, Vol. 7, No. 3, pp. 369-377.

Haskell, I.D. and Wickens, C.D. (1993) 'Two- and Three-Dimensional Displays for Aviation: A Theoretical and Empirical Comparison', *The International Journal of Aviation Psychology*, Vol. 3, No. 2, pp. 87-109.

Kaiser, M.K., and Mowafy, L. (1993) 'Visual Information for Judging Temporal Range', *Proc. of Piloting Vertical Flight Aircraft*, pp. 4.23-4.27, San Francisco, CA.

Lee, D.N. (1976) 'A Theory of Visual Control of Braking based on Information about Time-to-Collision', *Perception*, Vol. 5, pp. 437-459.

McRuer, D.T., Graham, D., Krendal, E., Reisener, W. (1965) *Human Pilot Dynamics in Compensatory Systems - Theory, Models, and Experiments with controlled element and Forcing Function Variations*, AFFDL-TR-65-15.

McRuer, D.T., Allen, R.W., Weir, D.H., Klein, R.H. (1977) 'New Results in Driver Steering Control', *Human Factors*, Vol. 19, No. 4, pp. 381-397.

Mulder, M. (1994) *Displays, Perception and Aircraft Control - A Survey of Theory and Modelling of Pilot Behaviour with Spatial Instruments*, Report LR-762, Delft University of Technology, The Netherlands.

Theunissen, E. (1993) 'A Primary Flight Display for Four-Dimensional Guidance and Navigation: Influence of Tunnel Size and Level of Additional Information on Pilot Performance and Control Behaviour', *Proceedings of the AIAA '93 FST Conference*, Monterey, CA.

Theunissen, E. and Mulder, M. (1994) 'Open and Closed Loop Control With a Perspective Tunnel-in-the-Sky Display', *Proceedings of the AIAA '94 FST Conference*, Scottsdale, AZ.

Theunissen, E. (1994) 'Factors influencing the design of perspective flight path displays for guidance and navigation', *Displays*, Vol. 15, No. 4, pp. 241-254.

Theunissen, E. and Mulder, M. (1995) 'Error-Neglecting Control with Perspective Flightpath Displays', *Proceedings of the Eight International Symposium on Aviation Psychology*, Columbus, OH.

Theunissen, E. (1995) 'In-flight Application of 3-D Guidance Displays: Problems and Solutions', *Proceedings of the '95 IFAC MMS Conference*, Cambridge, MA.

Wickens, C.D., Haskell, I., and Harte, K. (1989) 'Ergonomic Design for Perspective Flight Path Displays' *IEEE Control Systems Magazine*, pp. 3-8.

Wilckens, V. and Schattenmann, W. (1968) 'Test Results with New Analog Displays for All-Weather Landings' *AGARD Conference Proceedings No. 55*, pp. 10.1-10.31

Session 2

Human Reliability Models I

Chairman: P.A. Wieringa, The Netherlands

PRISMA - SAFETY

Prevention and Recovery Information System for Monitoring and Analysis of Safety Factors

T.W. van der Schaaf

Eindhoven Safety Management Group

Graduate School of Industrial Engineering and Management Science
Eindhoven University of Technology
P.O. Box 513
Pav. U8
5600 MB Eindhoven
The Netherlands

1. Introduction

PRISMA consists of set of tools to monitor and analyse process deviations and their consequences on safety, reliability, quality and the environment. Originally developed to manage human error in the chemical process industry, it is now being applied also in the steel industry, energy production and hospitals. The initial focus on safety consequences has been extended to provide an integral approach to manage all adverse consequences, based on the assumption that a common set of causal factors is responsible for these various effects.

The main goal is to build *quantitative database* of process deviations, from which conclusions may be drawn to suggest optimal countermeasures to the responsible managers. These countermeasures may be directed not only at *prevention of errors* and faults, but also at *promotion of recovery* factors to ensure timely corrective action. As such, PRISMA often uses not only actual but rare accidents as input, but also (the abundantly available) *near misses.*

This paper on PRISMA methods and techniques will serve as a common introduction to other papers by the Eindhoven Safety Management Group, by briefly describing:

- the Causal Tree incident description method;
- the Eindhoven Classification Model of System failure;
- The Classification/Action Matrix to suggest optimal countermeasures.

2. Causal Tree incident description method

(Van Vuuren and Van der Schaaf, 1995)

Causal trees, derived from fault trees, are very useful to present the critical activities and decisions during the development of an incident in a chronological order, and also show how all the different activities and decisions are logically related to each other. For near misses the causal trees are divided in two parts: the 'failure side' (left side in fig. 1), which gives an overview of all the activities that have lead to the failure, and the 'recovery side' (right side in fig. 1), which gives an overview of the activities that prevented it from developing into a real accident. The importance of recovery is discussed by Van der Schaaf (1995).

By using causal trees it becomes clear that there is never one single cause of a near miss, but that it is always a combination of many technical, organisational and human causes. The 'root causes' of the near miss, which are found at the bottom of the failure side of the causal tree, are the main products of this first phase of the PRISMA analysis.

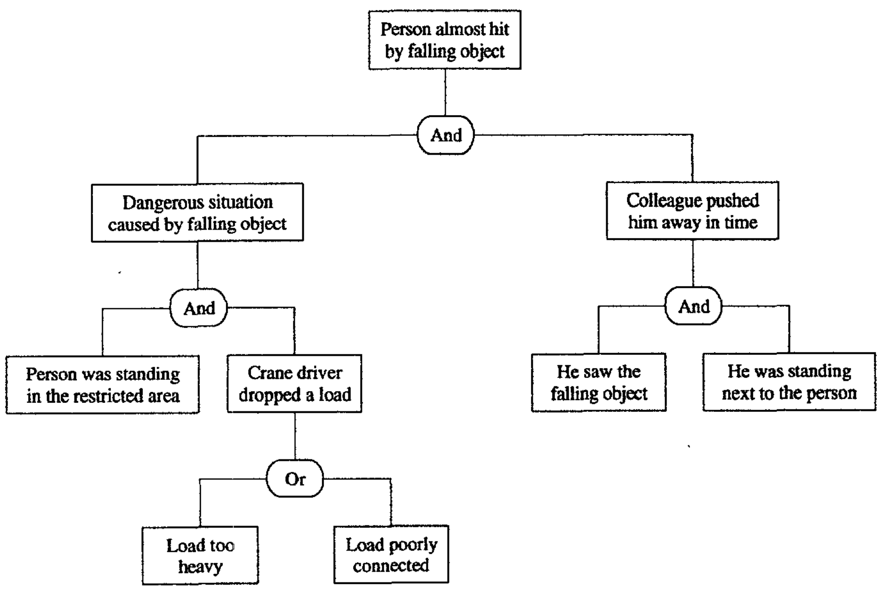


Fig. 1. Example of a causal tree describing a near miss.

3. Eindhoven Classification Model (ECM) of system failure

(Van der Schaaf, 1992)

3.1. General background, and a model of incident causation

In the past (and present!), discussions on the contributing factors of industrial safety have always had strong "fashion" aspects. It seems as if every new technology starts out by stressing *technical* factors like design and construction specifications, then moves on to "discovering" the importance of *errors by individual human operators* and finally switches to *organisational and management* aspects as the most important factors. Reason (1991) describes these historic transitions very clearly for *railroad transportation* systems by noting the succession of an "engineering age" to a "human error" period and finally to the "socio-technical" era. The remarkable thing about a more recent very safety-conscious technology like *nuclear power* generation is the fact that according to Reason these very same transitions seem to be fully recapitulated within its much shorter span: up to the mid-seventies safety measures in nuclear power plants were primarily directed at minimising the consequences of technical failures. This industry's human error concern (inspired by several very clear incidents and accidents) from 1975 on, first focussed on execution failures (like slips and lapses); after the Three Mile Island accident in 1979 one suddenly also realised the importance of cognitive failures or mistakes (like diagnostic errors and the selection of inappropriate recovery strategies).

In PRISMA an *integral approach* is followed by assuming that *all three types* of contributing factors must be investigated and acted upon. The most important question then becomes: what is the *relative* importance of each factor and of the interactions between them?

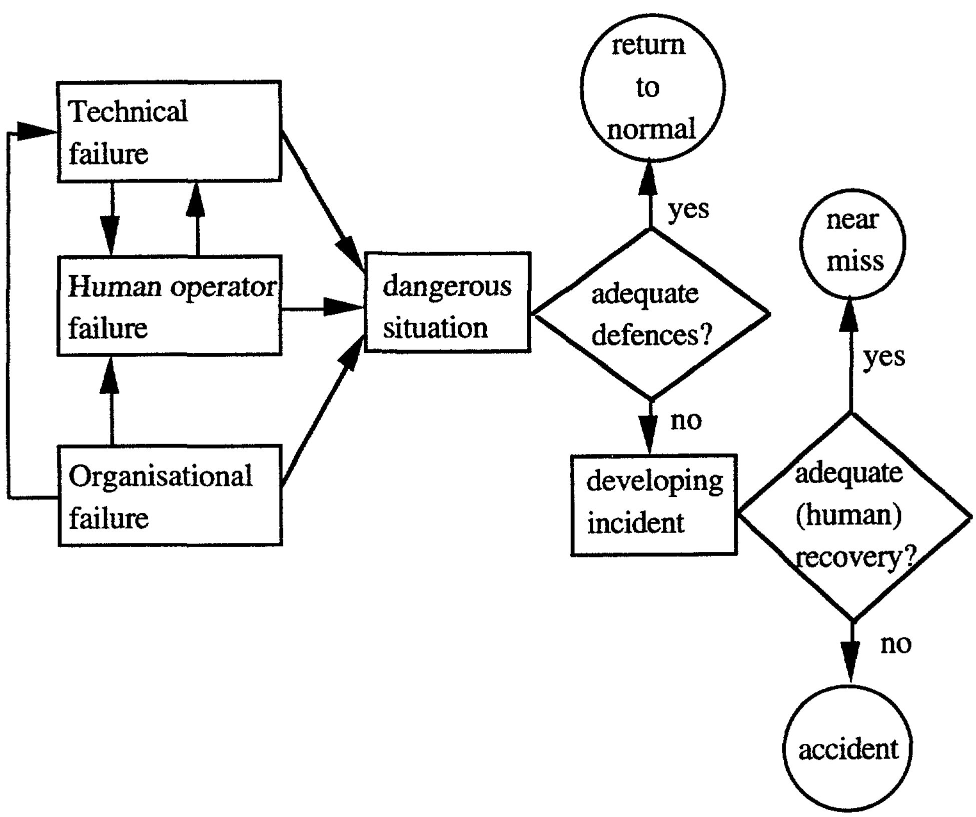


Fig. 2. A simple model of incident causation.

Figure 2 shows a simple model of the main components involved in incident causation, and also *defines* three basic terms: *accident*, *incident*, and *near miss*. Technical and organisational failures can initiate the chain of deviations leading to incident causation events both directly, and indirectly (e.g. by inducing operator failure). Only very seldom is an actually *dangerous situation* assumed to follow from such failures. Even if it does, the "built- in" defenses of the process to be controlled will *usually* be adequate (e.g. automatic safety systems; standard procedures for control of deviations, etc.) but *not always*. In the latter case the *potential incident* is allowed to develop further and it is usually up to the flexibility, experience, intuition, etc. of the human operator to try to *recover* from this undesired chain of events, and restore the original situation again, or at least to prevent major injuries, damages, etc. This *(human) recovery* phase will have been adequate if the developing potential incident is *detected, diagnosed and corrected accurately and in time;* in that case the potential incident is changed into a *"near miss"*, that is an occurrence with potentially important safety-related effects which in the end was prevented from developing into actual consequences. If human recovery was inadequate (e.g. too late; incorrect; or not even attempted) the potential incident will develop into an actual *"accident"* , that is an occurrence with actual adverse consequences (e.g. injuries (or worse), material damages, environmental pollution, etc.). The term *"incident"* refers to the combined set of occurrences of both accidents and near misses.

According to figure 2. the basic causes of incidents (also called *"root causes"*) are located at the very beginning of the chain of events, and may be Technical Failures, Organisational Failures, Human Operator Failures, or (most likely) a combination of these. On the basis of this assumption we have done several studies on the relative roles of these three groups of main factors in the last seven years. These studies looked at all kinds of system failures, with consequences varying from zero to (near-) disaster, in three different chemical plants in the Netherlands. Using a preliminary version of the PRISMA system failure model applied to the *original* in-house incident reports we could classify the main root cause in about 90% of all cases, with the following distribution:

– Technical Failure	30%
– Organisational Failure	10%
– Operator Failure	50%
– Unclassifiable	10%

(It is interesting to note that these figures were practically identical for all three companies, in spite of large (cultural) differences between them. These figures should however not be taken too literally; e.g. because of likely underreporting of organisational factors this figure will probably increase if investigated further.)

Nevertheless, the main conclusion must be that all three factors are of importance, with operator behaviour as the most dominant one. How then should we interpret the abundance of newspaper reports and other sources which proclaim that 90% to 99% of all such incidents are caused by "human error"?

Firstly, it may refer to the *use of "human" error in a useless way* when "human" refers not only to the person(s) directly involved in the incident but also to the designers, constructors and managers who gave the "operator" the tools to be used during task performance; in the end of course practically any design or management aspect of these "tools" and tasks may be traced to some human action or decision. At the same time such a definition then becomes useless because it does not give a single clue as to where and how improvements might be made.

Secondly a reason might be that the *incident investigation usually is rather superficial* in the sense that it often looks only at the events *directly* preceding the observable end components of a usually long and complex set of real root causes (see figure 1) and their interactions. These observable end components often consist mainly of human actions (or the lack of them).

Therefore, we propose to define "human" by *choosing a certain focus* or starting point, dictated by the goal of the investigation. If one is interested in the behaviour of control room operators then *their* task should dictate the classification: "human" error refers to *their* behaviour, "organisation" might refer to the procedures *they* have to follow for instance, and "technical design" refers to *their* workplace (e.g. control-room layout) and equipment,etc. One could equally be interested in the role of the engineers of course, but in that case the engineers become the human/"operator" component in the analysis; company guidelines for Engineering and Design Practices are then part of the "organisation" factor, etc.

In short, PRISMA uses a *goal-directed classification* because then we can fruitfully distinguish between the three main factors determining industrial safety.

3.2. Classification model of process supervision and control errors

3.2.1. Introduction

In section 3.1 it was argued that human behaviour is amongst the most dominant factors causing system failures in chemical process control. In this chapter we will present a detailed classification system of operator errors derived from the Rasmussen SRK model (Rasmussen, 1976). First the "human error" part will be operationalised, taking the process control operator task as the focal point. Then the other main categories of system failure, Technical and Organisational Failure, will be added in order to arrive at a complete model of system failure. An explicit link between classification results and proposed actions will then be presented in the form of a Classification/Action Matrix. In this way the combination of model + matrix may become an actual safety management tool.

3.2.2. Classification of human errors in process supervision and control

Rasmussen (1976) has provided the basic model of human error based on three levels of behaviour: skill-, rule- and knowledge-based (S-B, R-B, K-B). This SRK model has been operationalised to describe operator errors in process control tasks by combining it with *characteristic task elements*, which as a whole cover the entire spectrum of operator subtasks.

On the basis of our own informal observations and discussions in process control rooms, and the "classical" literature on operator tasks the following *set of elementary operator requirements* in order to carry out process control tasks correctly, has been identified:

a) firstly, the correct *status* and dynamics of the system to be controlled must be known to the operator;
b) also, the (main) *goal,* or priorities of goals, must be known and understood by the operator;
c) the operator in question must be *qualified* (on the basis of training) to do the job, and
d) if applicable, he must obtain a *temporary permit* for activities where extra risk is involved;
e) the preparation of the job itself starts by informing other operators, if necessary, of the work to be done (*coordination*), in view of the potential effects on *their* tasks;
f) when arriving at the job location the local system status should be *checked* to comply with the expected conditions in as far as these would be relevant for the job;
g) the job itself should be *planned* correctly, i.e. the correct methods should be chosen and carried out in the correct order;
h) the prescribed tools and information sources for a proper job performance should be present and used;
i) the execution of the required actions themselves implies successful correct movements; both *controlled, i.e.* intended, detailed, movements (e.g. to manipulate tools and request information), and maintaining the correct *body* position in order to make the controlled movements possible.

According to our interpretation of the SRK-model, *a* and *b* rely mainly on K-B behaviour; *c, d, e, f* and *g+h* all involve decisions to carry out the job in a certain way (i.e. according to the rules and procedures) and rely therefore on R-B behaviour; finally, the actual execution aspects *i* rely on S-B behaviour.

Figure 3. presents the classification codes, the labels of error categories and a typical example of each. These classification codes for Human operator behaviour *(H)* first distinguish between K-B, R-B and S-B behaviour *(K, R* and *S* respectively); at the most detailed level of classification these combined codes are further subdivided in the order of the elementary operator requirements mentioned above; the first subcategory of K-B error ("system status") is coded as HK1; the second type of K-B error ("goal") as HK2, etc.

	error code	descriptive label	example
K-B	HK1	system status	- not realising that part of the plant is inoperative because of maintenance
	HK2	goal	- aiming at "overspec" production instead of at "right-on-spec"
	HR1	license (permanent)	- not qualified for a certain task
R-B	HR2	permit (temporary)	- no permit obtained, although required
	HR3	coordination	- not informing control-room operator of one's actions outside in the plant
	HR4	checks	- not ensuring that system status is as expected
	HR5	planning	- choosing wrong method for correct goal
	HR6	equipment/information	- using wrong tools/process data
S-B	HS1	controlled movement	- making typing error on keyboard
	HS2	whole-body movement	- slipping, tripping, falling

Fig. 3. Subcategories of operator errors, with labels and brief examples.

The practical experiences in process control rooms mentioned earlier clearly point at *biases* in describing and classifying incidents: a tendency on the part of safety officers to concentrate on clearly visible S-B elements of task performance (e.g. pressing the wrong button) rather than on less obvious R-B errors (e.g. planning), let alone on the mainly cognitive, internal "activities" involved in K-B behaviour.

Therefore, in using the classification, the following *fixed order* is proposed (as indicated from top to bottom in figure 3.) to arrive at the best-fitting error category for causal factors of accidents and near misses: first K-B errors, then R-B and finally S-B errors. In this way *the above mentioned bias might be counteracted.*

3.2.3. Extension to classification of system failure

Human (operator) Error cannot be separated entirely from the Technical and Organisational context of task performance (see figure 1). At the very least one should know the importance of Human Behaviour relative to that of the Technical and Organisational factors in understanding the causes of accidents and near misses. On the basis of our own pilot CCR studies the following extensions are suggested:

The main category "Technical Factors" is subdivided into

- *Engineering* (T*E*): wrong design.
- *Construction* (T*C*): correct design which was not followed accurately during the construction phase.
- *Materials* (T*M*): rest category, for those material defects not classifiable under TE or TC.

The main category of Organisational Factors is subdivided into:

- *Operating Procedures* (*OP*): refers to the (inadequate) *quality* of procedures (completeness, accuracy, ergonomically correct presentation), not whether they are followed or not!
- *Management Priorities* (*OM*): refers to any *de facto* pressure by top- or middle management to let production prevail over safety.

Again, just as in the previous paragraph, a *fixed order of analysis* is advocated: firstly, one has to make absolutely certain that the technical design etc. of one's work environment is fully adequate; only after the technical design has been established as adequate, one may question whether the organisational context was a causal factor. Finally, when the Technical and Organisational aspects are found to be in perfect order, we should focus on human behaviour as a possible failure factor. In this way we hope to counteract the sometimes strong *bias* within a company's culture to start and stop the analysis at the level of the end-user and leave the technical and organisational context of any mishap unquestioned.

3.2.4. Eindhoven classification model of system failure

Integrating all the above, we arrive at the complete "Eindhoven Classification Model of System Failure in Process Control tasks" (see figure 4).
The classification model may also be used as a *checklist* to assist in the stage where *all relevant causal factors* should be collected for the investigation of a specific incident. In this respect it is not unlike *generic Fault Trees*, like the one used in the MORT system, which enable investigators not only to present their analysis results in a highly visible way, but also ensure that a multitude of non-technical factors ("specific job oversights and omissions", "management system factors") are carefully checked for their relevance.

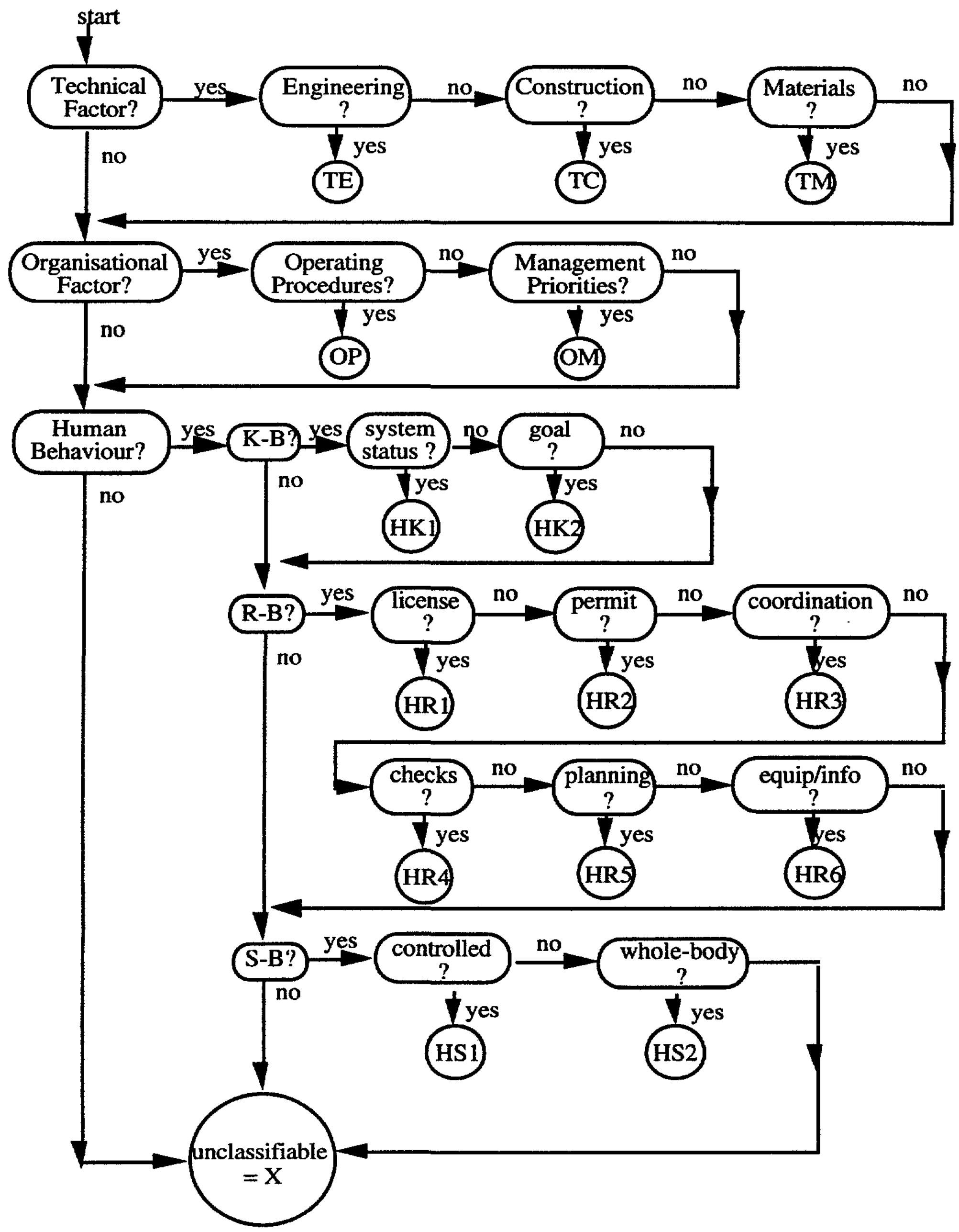


Fig. 4. The complete Eindhoven classification model of system failure.

4. Classification/action matrix

In order to develop an actual tool for safety management it does not suffice to stop at the analysis stage of failure classification mentioned above. These classification results have to be translated into proposals for *effective preventive and corrective action*. To fulfill this purpose a socalled Preliminary Classification/Action Matrix is proposed below (see figure 5.).

Its rows consist of the final classification codes as defined in figure 4., while its columns represent the following *five classes of actions*:

- *Equipment:* redesigning of hardware, software or interface parts of the man-machine system;

- *Procedures*: completing or improving formal and informal procedures for efficient *and* safe task performance;
- *Information & Communication* completing or improving available sources of information and of communication structures;
- *Training* improving (re)training programmes for skills needed;
- *Motivation* increasing the level of voluntary obedience to generally accepted rules by applying principles of positive behaviour modification.

	Equipment	Procedures	Information & Communication	Training	Motivation
TE	ⓧ				
TC	ⓧ				
(TM)					
OP		ⓧ			
(OM)					
HK1			ⓧ		no!
HK2			ⓧ		no!
HR1				ⓧ	
HR2				ⓧ	
HR3				ⓧ	
HR4				ⓧ	
HR5				ⓧ	
HR6				ⓧ	
HS1	ⓧ				no!
HS2	ⓧ				no!

Fig. 5. Classification/Action Matrix belonging to figure 4.

In the matrix the most preferred action in terms of expected effectiveness for each classification category is indicated by x . The last column's "*no*!'s" refer to particularly ineffective management actions, which are none the less often encountered in practice.

The entries in figure 5 were determined as follows:

- Technical Failures (TE, TC) point at technical improvements of the *equipment*; also S-B errors, because of their "automatic" nature, can best be prevented by changing the work environment, not by changing the person.
- Inadequate procedures (OP) imply improvements in the area of *procedures.*
- K-B errors (HK1, HK2) should be prevented by investing in operators' background

Decision Support for Human Decision Making in Production Scheduling

VINCENT C.S. WIERS

Eindhoven Univrsity of Technology
Graduate School of Industrial Engineering and Management Science
P.O.Box 513, Paviljoen
NL-5600 MB Eindhoven, The Netherlands
E-mail: v.c.s.wiers@bdk.tue.nl

Presented at the XIV European Annual Conference on Human Decision Making and Manual Control
June 14 – 16, Delft, The Netherlands

Abstract: Production scheduling is an important factor of the operational performance of production organisations. However, in practice, scheduling still is a typical human task. Attempts to apply decision support to the production scheduling task often have failed. In the article, two aspects of this problem will be dicsussed: the invalidity of the mental model of the designer regarding the production system to be controlled, and the gap between the mental model of the user and the mental model of the designer. To solve the first problem, two dimensions will be presented to assess the applicability of a technique in a certain situation: (1) the extent to which a scheduling technique can adequately model the structure of the shop floor, and (2) the extent to which the processing part of a technique matches the nervousness and the goals of the production unit. It will be argued that co-operative design and implementation can be used to solve the second problem.

1. Introduction

The planning and control of production departments, which is referred to as production scheduling, lies at the heart of the performance of a manufacturing organisation. Due to market demands regarding service, order flow times and flexibility, the need for adequate scheduling in practice has greatly increased during the last decades. In the same period, academic research in production scheduling has moved from purely theoretic exercises to techniques that are aimed at real-world problems. Some of these techniques have been implemented in standard software [Wortmann *et al.*, 1995]. However, successful implementations of scheduling techniques in practice are still very scarce [King, 1976; Graves, 1981; McKay *et al.*, 1988; Rodammer and White, 1988; Buxey, 1989; McKay *et al.*, 1989]. Moreover, most of the successful implementations have only been realised in highly controlled industries, i.e., process industry and mass-assembly. Therefore, in many companies, scheduling is still a typical human task. However, production scheduling is also a very complex task.

In this paper we will address the problem of applying decision support to production scheduling, while avoiding the pitfalls that have held back earlier efforts. Problems that cause many implementations of scheduling software to fail will be discussed. We will argue that these problems do not solely stem from problems with techniques themselves, but also have to be seen in a broader context, i.e., the organisation of production control decisions. We will argue that specific characteristics of humans are often underemphasised when designing and implementing scheduling techniques. Some parts of the scheduling tasks fit very well with the abilities of human cognition, while other sub-tasks pose difficulties. Therefore, a task-allocation decision has to be made based on these characteristics.

The paper is structured as follows: in Section 2 we will give a description of production scheduling. In Section 3 we will give an overview of available scheduling techniques. In Section 4 we will explain why the implementation of techniques in practice often has failed. In Section 5 we will outline human strengths, weaknesses, and preferences regarding the scheduling task. In Section 6, possible improvements for scheduling are given. Lastly, in Section 7, our conclusions are given.

2. The Production Scheduling Task

In this Section, scheduling, and its relation to other production control decisions, i.e., planning and sequencing are discussed. The relation between these functions is depicted in Figure 1. In production *planning* the amount of required production in a specified time horizon is determined by the amount of (expected) customers orders. The output of planning consists of material requirements in time. These requirements are passed to the production departments. A production unit is an outlined part of the production process, that on short term is autonomous regarding the use of resources [Bertrand *et al.*, 1990] (note: production unit and shop floor are used as equivalent terms in this paper). The production departments are controlled by *production scheduling.* Scheduling focuses on the allocation of finite resources within the production departments to fulfil the material requirements, and determines for each resource (e.g., machine–capacity), the points in time when operations are executed, under the following constraints: Finite capacity resources; precedence relations (i.e., routings); start–dates and due–dates of jobs. These constraints that are satisfied by scheduling are often *not* satisfied in the planning function.

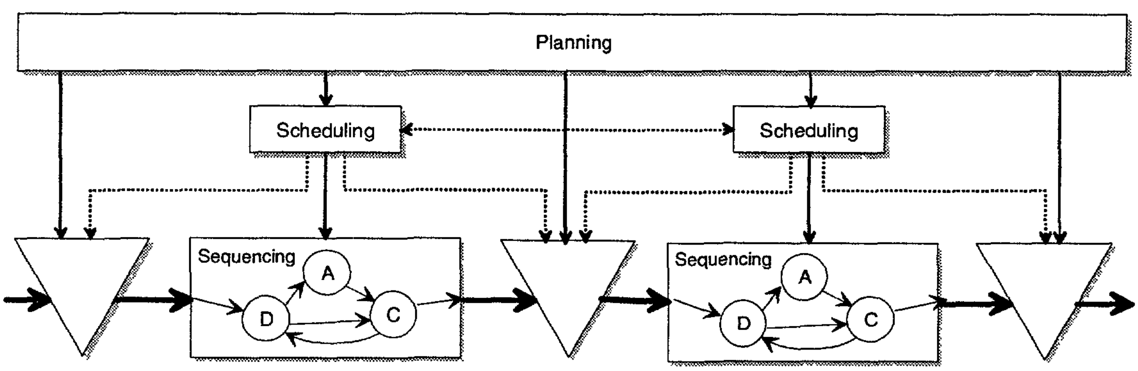


Figure 1: The relation between scheduling, planning and sequencing

Scheduling should optimise (or at least satisfy) certain *goals* that are deduced from organisational objectives, for example: service level; resources utilisation; set up costs; inventory costs; order flow times. After scheduling, the schedule is transferred to the production unit.

At each workcenter in a production unit, every time a job is completed, a decision has to be made which order will be processed next; this decision is referred to as *sequencing.* The existence of an explicit sequencing decision at the shop floor depends on how scheduling is carried out. Two possible aggregation levels of scheduling exist:

Resource level. Jobs are scheduled for each resource (workcenter). This type of scheduling leaves no sequencing decision freedom to the shop floor.

Production Unit level. Jobs are only scheduled for the total production unit. The operations of the job are not scheduled. This type of scheduling leaves the sequencing decision to the shop floor.

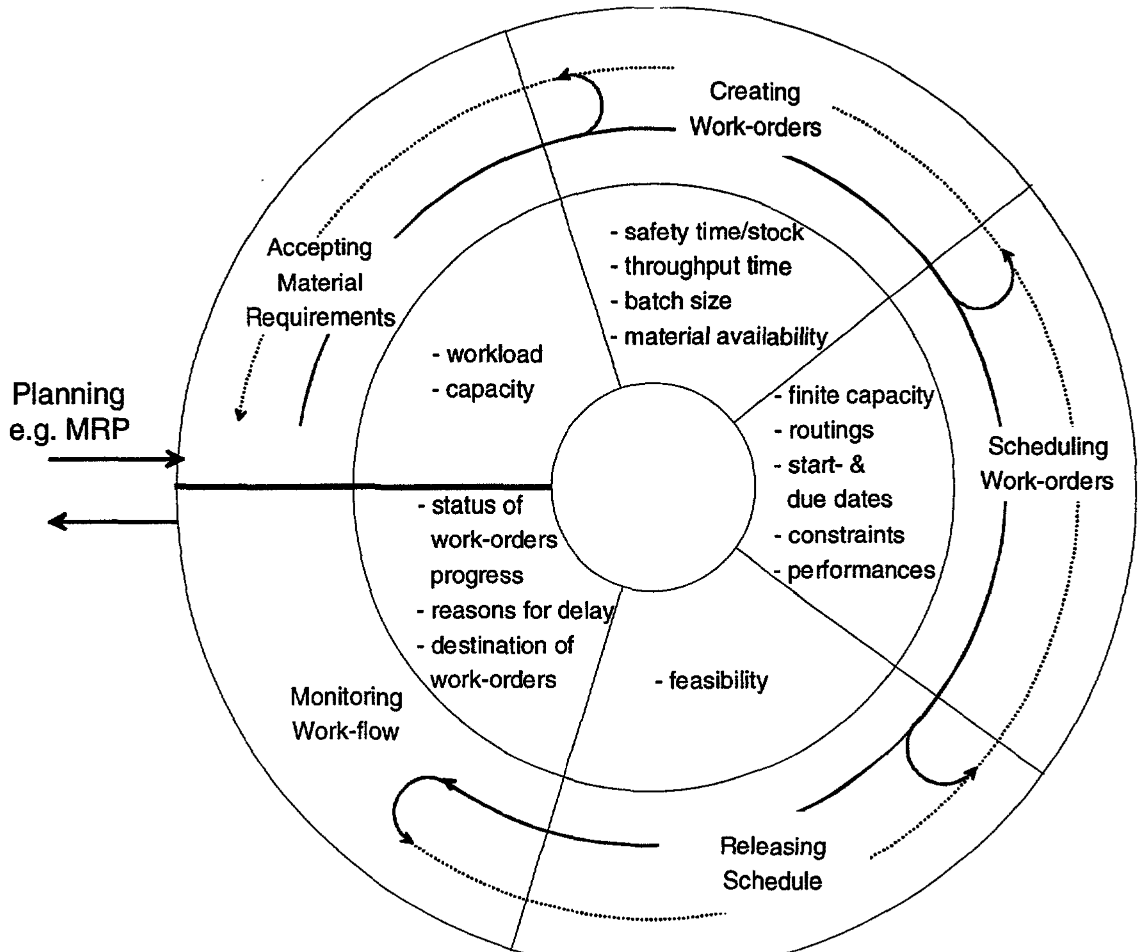


Figure 2: The scheduling task. The inner layer represents the information required for carrying out the activities in the outer layer.

In Figure 2, the activities required to carry out the scheduling task are depicted in the outer layer of the circle, and the information required to carry out these activities is depicted in the inner layer of the circle. The activities have to be carried out clockwise; if unforeseen events occur it may be necessary to carry out the activities counter-clockwise.

3. Techniques

A large amount of scheduling techniques, and information systems based on these techniques, is available. We will not give an extensive overview this vast field of research, instead, types of techniques that are relevant to our discussion will be mentioned here. A comprehensive overview of techniques can be found in Morton and Pentico (1993).

Priority dispatching rules. Priority dispatching rules are usually based on job characteristics, e.g., processing time or due–date. Some rules also take the status of the total shop into account, i.e., work-load. An example of a dispatching rule is the Shortest Processing Time (SPT) rule (i.e., if a job is completed at a work–center, the job with the shortest processing time in the queue will be processed next). Priority dispatching rules can be executed by the scheduler or at the shop floor. If the scheduler uses these techniques, the shop floor is simulated by a time-clock, and operations are scheduled according to the priority rule. If the operators on the shop floor use these techniques, the decision about which order to select next from the queue for the workcenter if an order has been completed is based on the priority rule.

Optimisation techniques. These techniques find an optimal schedule by enumerating many (possibly all feasible) schedules and choosing the best one according to a specific performance criteria, e.g., makespan, average order flow time. Examples are: Branch and bound techniques, mathematical programming.

Bottleneck methods. These techniques make a distinction between bottleneck and non–bottleneck resources. The bottle-neck resource is scheduled first to ensure maximal utilisation. Then the critical non-bottle-neck resources, i.e., resources preceding the bottle-neck, are scheduled to continuously provide the bottle-neck with work. A well-known example of a bottle-neck technique is Optimised Production Technology (OPT).

Knowledge Based Techniques. Knowledge based techniques model the shop floor by means of many hard and soft constraints. These rules are often obtained by eliciting knowledge from (experienced) schedulers. These techniques are usually aimed at generating feasible schedules [Randhawa and McDowell, 1990].

4. The Gap between Theory and Practice

There often are differences between the mental models of the designer and the user regarding the production unit to be controlled. This ultimately results in the human scheduler not using the technique. These differences are depicted in Figure 3. In this Section we will focus on the problem of matching the mental model of the designer with the production unit, i.e., match 2 in Figure 3. To achieve an adequate match we will present two dimensions of the scheduling task: the model of the production department, and the processes in the production department (see also the inner and outer layer of Figure 2).

4.1 Model

Techniques rely on a certain model of the production unit to generate a production schedule, and consists of the machines, tools, personnel, workorders, etc. that are present in the unit. However, because of the very complex and dynamic nature of shop floors, often many assumptions have to be made to enable economically feasible modelling. If techniques generate schedules that are based on an invalid model of the shop floor, these schedules have to be adjusted by humans. However, if the schedule is adjusted manually, the performance criterion the schedule was based on will suffer from these actions, and the reason for using such a technique becomes doubtful. Priority dispatching rules suffer less from this problem, because these techniques do not necessarily

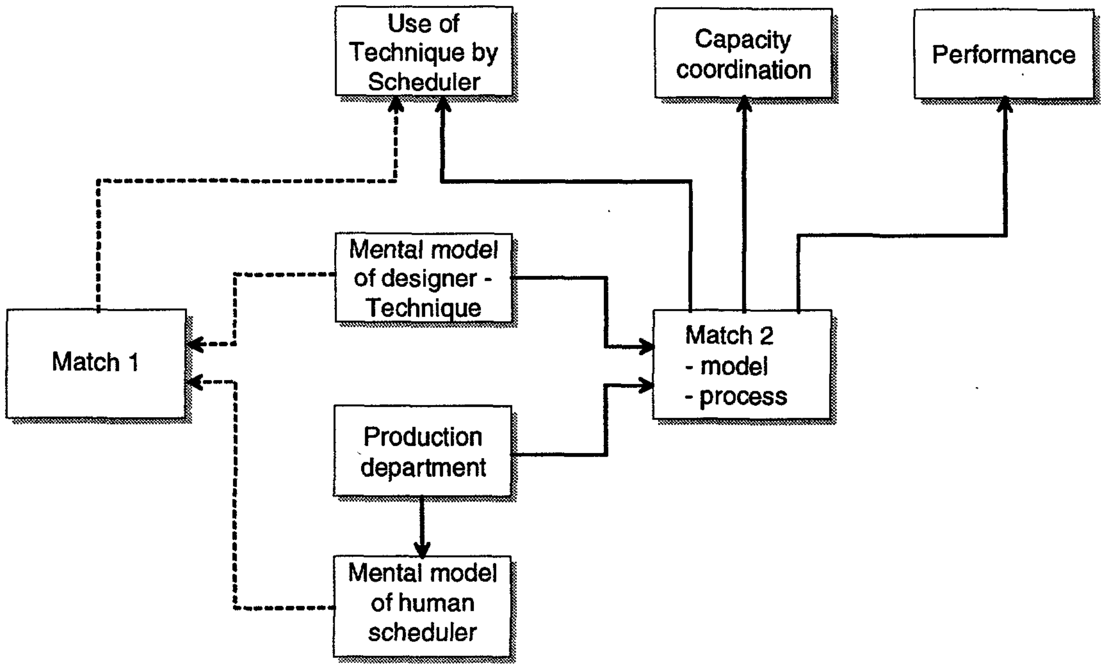


Figure 3: Conceptual model of the research

rely on a model of the shop floor. Instead, they can be executed by operators on the level of workcenters.

> *An example of an assumption regarding the structure of the shop is that jobs cannot be split. However, in practice, jobs often can be split and produced simultaneously on two or more machines (and possibly the second machine needs a longer processing time). In this case (and similar cases), scheduler or the shop floor workers will probably violate the schedule and split jobs if necessary.*

Problems related to the model of the shop floor occur especially with mathematical scheduling techniques. Although it is theoretically possible to model the shop floor in great detail, the level of detail has to be paid by the response time of the scheduling process and the costs of the modelling efforts. Artificial Intelligence (AI) approaches are only slightly more successful in modelling the shop floor. The strength of AI, i.e., to model many hard and soft constraints of the shop floor, also comes with weaknesses. Because these rules are constantly subject to change, the knowledge elicitation and modelling process will practically never end. Also, because AI systems usually are based on knowledge elicited from human schedulers, the AI system only duplicates human performance and cannot create a better schedule than a human scheduler. Moreover, AI techniques also suffer from long response times if a detailed model of the shop floor is used.

4.2 Process

4.2.1 Nervousness

Dependent on the type of production unit, many disturbances occur at the shop floor. Therefore, schedules have to be adjusted frequently. However, some techniques do not offer the possibility to manually adjust the schedule. Instead, each time something relatively small occurs, the technique generates a complete schedule. Techniques that are only able to completely generate a schedule usually suffer from the *final–state* problem [Wortmann *et al.*, 1995]. This means, that a small change in the status of the shop floor causes major changes in the production schedule.

For example: job X has been scheduled to start at Monday, 10.00 AM. However, the material required for job X is not present, so it cannot be started. To keep the schedule up-to-date, job X has to be rescheduled to a later start-time in the schedule. The schedule technique now calculates a new optimal schedule with job X rescheduled. This new optimal schedule is totally different from the "old" schedule.

In many production units, disturbances as in the example happen every few minutes. If a schedule generation technique would be used here, the system would generate a completely new schedule every few minutes. The scheduler and the personnel on the shop floor will not accept such a nervous behaviour and will eventually reject the constant load of schedules.

A priority dispatching rule can be executed by the humans at the shop floor level and is thereby much more robust than techniques of the second category, which are executed on a higher organisational level. If dispatching rules are used on the shop floor, no fixed sequence is present, and each time a job is completed at a machine, the shop floor worker will choose the next order according to the priority dispatching rule. Also, if humans decide to deviate from dispatching rules this does not have major consequences on other jobs in the shop. So, dispatching rules enable humans to use their own mental model of the shop floor.

4.2.2 Criteria

Of the techniques discussed in 3, priority dispatching rules are the most broadly used in practice. Also, these priority rules often belong to implicit scheduling guidelines that schedulers already use, unaware of their meaning in production planning and control theory. However, priority dispatching rules are not very 'smart' and they generate schedules that are far from optimal with respect to a specific performance criterion. Most of these rules use local information (at the workcenter level), which leads to local optimisation, i.e., insensitive to the overall state of the production unit. Mathematical techniques give the best results if we solely look at the ability to optimise schedules conform one or a combination of performance criteria. Bottle-neck techniques primarily focus on the utilisation of bottle-necks, and are not very well suited to optimise schedules conform other goals. Knowledge based techniques are not in the first place used to optimise schedules, instead they are able to generate a schedule that is feasible conform to the many constraints in the knowledge base.

5. Human Factors

In the previous Section, problems related to the match between scheduling techniques and the production unit to be controlled were discussed. We assumed that the mental model of the scheduler adequately matches the production unit. However, there are some aspects within production scheduling that pose cognitive problems to the human scheduler. These problems will be discussed below.

5.1 Limited cognitive abilities

Short-term Memory. The short term memory of humans applies to the amount of information that a human can pay attention to simultaneously. A human can have approximately seven 'chunks' of information in his or her short term memory [Anderson, 1990]. One information chunk can apply to one or a coherent set of task elements. To manage all task elements in the scheduling task, the task should be regarded on a certain level of aggregation to conform to the limited capacity of short term memory. A high aggregation level would be to look at all jobs simultaneously and not discriminate between groups or individual work orders. A low level of aggregation would be to look at each task element separately, for example each job. It will be clear that the first option is not suited for manipulating the jobs in the schedule. The last option violates the short-term memory constraint. However, no in-between levels of decomposition are available. When constructing a schedule, one has to take into account that an action regarding one job influences most other jobs. However, the human scheduler is forced to decompose the task anyway. This results in sub-optimising parts of the schedule subsequently, and the human is therefore not able to optimise the schedule towards some performance criterion.

Long-term Memory. The long term memory of humans is, contrarily to the short term memory, not limited regarding capacity. Limitations of long-term memory lie in our restricted ability to *recollect* information [Anderson, 1990]. These restrictions especially apply to abstract and numerical information. The limitations of the long-term memory of humans has effect on the ability to assess effects that occur as a result of actions with some delay. As a result, the human scheduler is not able to assess the effects of his or her actions, and to balance actions aimed at the short-term and actions aimed at the long-term. Also, the human is not able to learn how to improve his or her performance, because the schedule performance cannot be evaluated.

5.2 Advanced cognitive abilities

Humans are very well equipped to cope with many 'soft,' qualitative task elements. Humans are superior to existing scheduling techniques and information systems regarding the following characteristics [McKay *et al.*, 1989]:

Flexibility, adaptability and learning. Humans can cope with many stated, not–stated, incomplete, erroneous, and outdated goals and constraints. Furthermore, humans are able to deal with the fact that these goals and constraints are seldom more stable than a few hours.

Communication and negotiation. Humans are able to influence the variability and the constraints of the shop floor; they can communicate with the operators on the shop floor to influence job priorities or to influence processing times. Humans are able to communicate and negotiate with (internal) customers if jobs are delayed, or communicate with suppliers if materials are not available as planned.

Intuition. Humans are able to fill in the blanks of missing information required to schedule. This requires a great amount of 'tacit knowledge.' At the time of collecting this knowledge it is not always clear which goals are served by it.

5.3 Human Preferences

Apart from cognitive strengths and weaknesses, humans often prefer their own judgement to the application of techniques and information systems. This especially goes for the situation where they are confident about their expertise [Kleinmuntz, 1990], which is certainly the case in the scheduling task, which requires a high level of expertise. Schedulers communicate intensively with people on the shop floor, and many have advanced from a shop floor position to a scheduling position. However, decisions made by overruling formal techniques are not necessary superior to the decisions suggested by these techniques, although these decisions are hard to compare.

Humans also have less confidence in techniques if they are not able to understand how the technique works. Of the techniques discussed in Section 3, priority dispatching rules are probably the best understood by the planner. Priority rules often already belong to implicit scheduling guidelines that schedulers use. Also, bottle-neck techniques find their origin in simple principles. Contrarily, optimisation techniques are opaque to the scheduler, because it is not clear how such a technique finds a particular schedule. AI techniques suffer from the same drawback.

6. Improving Scheduling

In this Section, alternatives for improving scheduling in practice will be given. These alternatives focus on the match between the production unit and the designers' view of the production unit, i.e., the scheduling technique (match 2 in Figure 3). To avoid problems with problems regarding the mental model of the scheduler and the production unit (match 1 in Figure 3), the scheduler (and the operators on the shop floor) should be involved in the selection and implementation of a specific technique.

6.1 Applying Techniques

Techniques can be used to improve scheduling if an appropriate technique is used for a specific production unit. We have already argued in Section 4 that two dimensions should be used when selecting a technique for a specific production unit: the model of the shop floor, and the scheduling process. A technique should match a production unit regarding these dimensions. If techniques are not able to match the production unit, the scheduler and/or the operators on the shop floor will probably disregard the schedule generation technique.

Priority dispatching rules. Priority dispatching rules do not really have a shop floor model; they are applied to queues for individual workcenters (machines). These rules are used to decide which order in a queue should be handled next if a workcenter gets idle. Therefore, priority rules are only able to schedule individual queues, and do not or scarcely co-ordinate all workcenters in a coherent way. Routings of products, personnel capacity, special constraints, etc. have to be handled manually. The scheduling process is based on a heuristic and is not optimal regarding some or a combination of performance criteria. Because of the limited scope of the model of the shop, priority dispatching rules often lead to sub-optimisation. The main reason to use these rules is that they are relatively simple to use. They are especially suited for complex job-shops, where only one type of capacity has to be co-ordinated, and where much uncertainty exists. Because rules can also easily be carried out by the operators on the shop floor, the scheduler does not have to generate a schedule in advance.

Optimisation Techniques. Today, builders of optimisation techniques are able to model the shop floor in a sophisticated way, including multiple resources, complex constraints, etc. However, the level of detail of the shop floor model has to be compromised with the response time of the system, which usually is quite low. This compromise of course depends on the re-scheduling requirements of the production unit to be scheduled. The scheduling process is able to optimise the schedule conform one or a combination of performance criteria. However, as described in Section 3, these techniques suffer from the final state problem. Therefore, these techniques should only be used with production units with very few disturbances. Because these techniques are difficult to understand by the scheduler, they should be implemented carefully, and should only be used in situations where the role of the scheduler is limited e.g., continuous process industries or mass-assembly lines.

Bottleneck techniques. Bottle-neck techniques use a simplified but potentially effective model of the shop floor because only the bottle-neck and the processes feeding the bottle-neck have to be scheduled. However, not in all production systems it is possible to locate the bottle-neck. An advantage of bottle-neck scheduling is that the logic appeals to the scheduler. However, the operators on the shop floor loose all decision freedom, and it is often very difficult to persuade operators to follow the schedule exactly, as it involves many set-ups that sometimes are regarded as unnecessary. Moreover, these systems reschedule at each disturbance, which further harms their credibility towards the scheduler and operators. Bottle-neck techniques should therefore only be used in stable, flow-oriented production systems.

Knowledge Based Techniques. The large advantage of knowledge based techniques above 'conventional' techniques is their ability to model many hard an soft constraints of the shop floor. The limitations of knowledge based techniques regarding the model of the shop floor lie in the difficulties to build an maintain such a knowledge base. To build a knowledge base, an expert has to be present, which is not always the case. Furthermore, the amount of rules that apply to a shop floor can be very large, which hampers the modelling process. Furthermore, these rules are constantly subject to change. The scheduling process of such a technique delivers a 'feasible' schedule which is not necessarily optimal conform one or a combination of performance criteria. The more sophisticated the model of the shop floor is, the longer the response time will be. Therefore, knowledge based techniques should only be used in shops with a large amount of rules that can be elicited in an economically feasible way.

6.2 Human Factors

The reason for applying a technique in scheduling lies in the cognitive weaknesses of human schedulers. Sub-tasks within scheduling that require the handling of many jobs, procedures, capacity constraints, etc. to achieve some specified performance criteria are eligible be supported by techniques. The reason for giving decision responsibility to the scheduler or to the operators of the shop lies in the cognitive strengths of humans, which are especially employed in production units where scheduling tasks require flexibility, communication and intuition.

A possible solution to the problem of schedule fulfilment would be to force the shop floor to literally execute the schedule, and to penalise behaviour deviations. However, this solution ignores the complexity of scheduling in practice. It would make the job of schedulers and operators less interesting and sometimes even frustrating, for example when they see obvious solutions for problems in the schedule which they are not allowed to carry out. Problems regarding schedule fulfilment can be compensated by:

- *Putting more flexibility on the shop floor.* This can be achieved by increasing the number of tools, procedures for checking the availability of drawings and raw materials, and technically decreasing the amount of machine disturbances. Furthermore, the availability and quality of materials can be improved by better co-ordination between the (internal) supplier and the production unit. Multi-deployment of operators, machines, and materials will increase flexibility. Also, decreasing the complexity of the production unit by standardising, and thereby decreasing the number of: materials, processes, end-products, etc. can be used to improve flexibility.
- *Giving more decision freedom to the scheduler.* The reason for giving decision freedom to the scheduler lies in the cognitive strengths of humans, which are especially useful in production units where schedule fulfilment is a problem. In these situations, the schedule should not primarily be constructed and optimised, but continually re–constructed by reacting to short–term disturbances and events.
- *Creating possibilities for informal communication.* For example between schedulers or between a scheduler and the operators on the shop floor. Informal communication can for example easily be created by geographically placing people together.

7. Conclusions

In this paper we have addressed the problem of giving adequate decision support to human schedulers in practice. This problem has been attacked many times by academia and practitioners in the field of production scheduling, however, without much success. We have first outlined the scheduling task and its relations with other production planning and control components. We have given an overview of available techniques for production scheduling. We then presented a framework to match the mental model of the designer, i.e., the technique, with the production unit to be controlled. The gap between the mental model of the designer and the mental model of the production scheduler can be overcome by co-operative selection/design and implementation of a technique or software for production scheduling. In Section 6 we applied the framework to existing techniques, and we indicated when these techniques should be applied in practice.

References

ANDERSON, J.R. (1990). *Cognitive Psychology and its Implications.* Third edition. Freeman, New–York.

BERTRAND, J.W.M., WORTMANN, J.C., WIJNGAARD, J. (1990). *Production Control: A Structural and Design Oriented Approach.* Elsevier Science Publishers B.V., Amsterdam.

BUXEY, G. (1989). Production Scheduling: Practice and Theory. *European Journal of Operational Research*, 39, pp.17–31.

GRAVES, S.C. (1981). A Review of Production Scheduling. *Operations Research*, Vol. 29, pp. 646 675.

KING, J.R. (1976) The Theory Practice Gap in Job Shop Scheduling. *The Production Engineer*, Vol. 55, March, pp. 137 143.

KLEINMUNTZ, B. (1990). Why we Still Use our Heads Instead of Formulas: Toward an Integrative Approach. *Psychological Bulletin*, Vol.107, No.3.

MCKAY, K.N., SAFAYENI, F.R., BUZACOTT, J.A. (1988). Job Shop Sheduling Theory: What Is Relevant? *Interfaces*, Vol. 18, pp. 84 90.

MCKAY, K.N., BUZACOTT, J.A., SAFAYENI, F.R. (1989). The Scheduler's Knowledge of Uncertainty: The Missing Link. In: *Knowledge Based Production Management Systems*, J. Browne (Ed.). Elsevier Science Publishers B.V., North Holland, IFIP.

MORTON, T.E., PENTICO, D.W. (1993). *Heuristic Scheduling Systems (with applications to production systems and project management).* John Wiley & Sons, New York.

RANDHAWA, S.U., MCDOWELL, E.D. (1990). An Investigation of the Applicability of Expert Systems to Job Shop Scheduling. *International Journal of Man-Machine Studies*, Vol. 32, pp.203-213.

RODAMMER, F.A., WHITE, K.P. (1988). A Recent Survey of Production Scheduling. *IEEE Transactions on Systems, Man, and Cybernetics*, Vol.18, No.6.

THOMAS, L.J., MCCLAIN, J.O. (1993). An Overview of Production Planning. In: *Logistics of Production and Inventory.* S.C. Graves, A.H.G. Rinnooy Kan, P.H. Zipkin (Eds.). Elsevier Science Publishers B.V., Amsterdam.

WORTMANN, J.C., EUWE, M.J., TAAL, M., WIERS, V.C.S. A Review of Capacity Planning Techniques within Standard Software Packages. To appear in Production planning & Control.

Risk Management in Hospitals: Predicting versus Reporting Risks in a Surgical Department

N.W.S. van der Hoeff[1] and T.W. van der Schaaf

Safety Management Group
Eindhoven University of Technology
P.O. Box 513
Pav. U-8
5600 MB
Eindhoven
The Netherlands

1. Abstract

Medical errors are increasingly unacceptable in modern society; for ethical reasons, but lately also because of a rising trend in the number and amount of damages paid to ex-patients. As a result, a large training hospital and the Eindhoven Safety Management Group have started a joint project "Risk Management in Hospitals". In this project, the PRISMA safety management approach based (see PRISMA paper of Van der Schaaf) on a decade of work in the chemical process and steel industry was tested in the domain of medical treatment.

The project consists of five stages:
1. a prediction of the possible risks to patients.
2. an inventory of the real risks to patients.
3. the development and implementation of a voluntary incident reporting system.
4. the development of a set of minimum norms which a surgical department has to satisfy.
5. the development and implementation of a quality system based on the two methods of predicting and registering incident causes and the norms developed, not only on the Surgical Department (SD) but on every ward of the hospital.

In the first stage, the processes on the SD were described using a systems approach by In `t Veld. The predictions of the risks were made by using the Failure Mode and Effects Analysis (FMEA)-method.

In the second stage, an inventory of actual incidents and their causes was made by using Critical Incident Interviews as an input to the PRISMA incident analysis approach. A comparison was made between the results of the FMEA and the Critical Incident Interviews, thus comparing predicted vs. reported risks.

Now (June 1995) the project is halfway stages three and four. In stage three, a pilot has been started on the SD with an operational incident reporting system. In stage four, the set of minimum norms for every ward will be developed and tested: this will happen in the next few months.

In the future, the two methods will be used to predict and report also medical complications, accidents and near accidents with regard to negative physical consequences for nurses

---

[1]now working as risk manager at the Catharina Hospital in Eindhoven, The Netherlands

and doctors, events with negative effects on the environment, and events that cause higher costs for the hospital.

Important concepts are defined in the appendix.

2 Introduction

2.1 Motivation for the project

This risk management project is a joint project of the Eindhoven Safety Management Group and the Catharina Hospital in Eindhoven, The Netherlands. The Catharina Hospital is a large teaching hospital with a clinical Surgical Department (SD) consisting of ten operating theatres (about 11,500 clinical operations a year) where a wide range of operations is performed.

The goal of risk management is to control the processes patients are going through on a SD in order to minimize the risks for patients and to obtain better care.

Additional goals for risk management are:
- the trend of rising insurance premiums and
- future quality legislation.

2.2 The research design of the project

The project consists of five stages:
Stage 1: A prediction of the possible risks to patients. This stage is divided into two parts:
In the first part, a model was made of the processes that take place in the SD. The processes were described by the system approach (In 't Veld [10]). In the second part, a prediction of the possible risks to patients was made. The predictions were made by using FMEA (Failure Mode and Effects Analysis) [1][3].
Stage 2: An inventory of the real risks to patients. This second stage is divided into three parts. They provide a basis for the next stage: a voluntary incident reporting system which is a permanent risk management tool. In the first part, twenty Critical Incident Interviews were taken of which seventeen were useful and were described in detail by qualitative Causal Tree Analysis (see PRISMA paper of Van der Schaaf). In the second part of this stage, a comparison was made between the results of the FMEA and the Critical Incident Interviews. Finally in the last part, conclusions and recommendations were made on the basis of:
- the results and use of the process model;
- the results and use of the FMEA;
- the results and use of the Critical Incident Interviews;
- the results and use of the Eindhoven Classification Model;
- the results of the comparison between the FMEA and the Critical Incident Interviews.

Stage 3: The development and implementation of a voluntary incident reporting system [9]. This stage consists of three parts. The first part is a participant design with the SD staff of an incident reporting system based on the theoretical concept of a Near Miss Management System [8]. To make a participant design possible, several training sessions were held for nurses and doctors. In the second part a reporting system is set up. This implies a two month period in which the reporting system is actually operational under supervision of a member of the Safety Management Group. The goal is a fully autonomous reporting system after these two months. The last part of this stage is an evaluation, to improve the reporting system on the SD and decide on other future incident reporting systems on other wards in the hospital.

Stage 4: The development of a set of minimum norms which a ward has to satisfy. A surgical department must satisfy these norms to achieve process control and a maximum of safety for patients. This set describes what must be discussed, agreed and recorded between nurses and doctors at important policy making moments. It also forces the participants to decide and record not only who is authorized for certain tasks but also who is responsible.

Stage 5: The development and implementation of a quality system based on the two methods of predicting and registering incident causes and the norms developed, not only on the SD but on every ward of the hospital.

3 Stage one: Predicting possible risks to patients

3.1 The process model of the Surgical Department

Three possible methods of describing the processes undergone by patients were examined:

- "SADT" (Structured Analysis and Design Techniques) [5];
- "Data Flow Diagramming" [2][3];
- "steady state" systems (system approach described by In 't Veld [10]).

The processes were finally described by the system approach [7]. The value of the "In 't Veld" method has been demonstrated in the construction of a process model for the SD. Strong points of the "In 't Veld" method are:

- processes are modelled in natural sequence (according to the time aspect in a flow);
- the possibility of using different aggregation levels;
- the use of measurement and control loops with norms.

This model must meet the following requirements:

- it must give insight into the processes that take place in the department;
- it must contain all aspects that are necessary and relevant for process control. These aspects are:
  - patient;
  - personnel:
    - medical;
    - nursing;
    - paramedical;
  - equipment;
  - instruments;
  - information;
  - interfaces (with other departments);
  - materials:
    - use;
    - consumption.
- it must be able to be used as an input to an assessment method for possible risks (FMEA, HazOp (: see section 3.2));
- it must give insight to all users so it can be used as a communication means between members of the project and others.

The process model of the Surgical Department consists of four main parts:

1 preparation of the patient;
2 anaesthesia;
3 operating on the patient and maintaining the anaesthesia;
4 recovery.

Each part is described by using the system approach of In 't Veld (see figure 1).

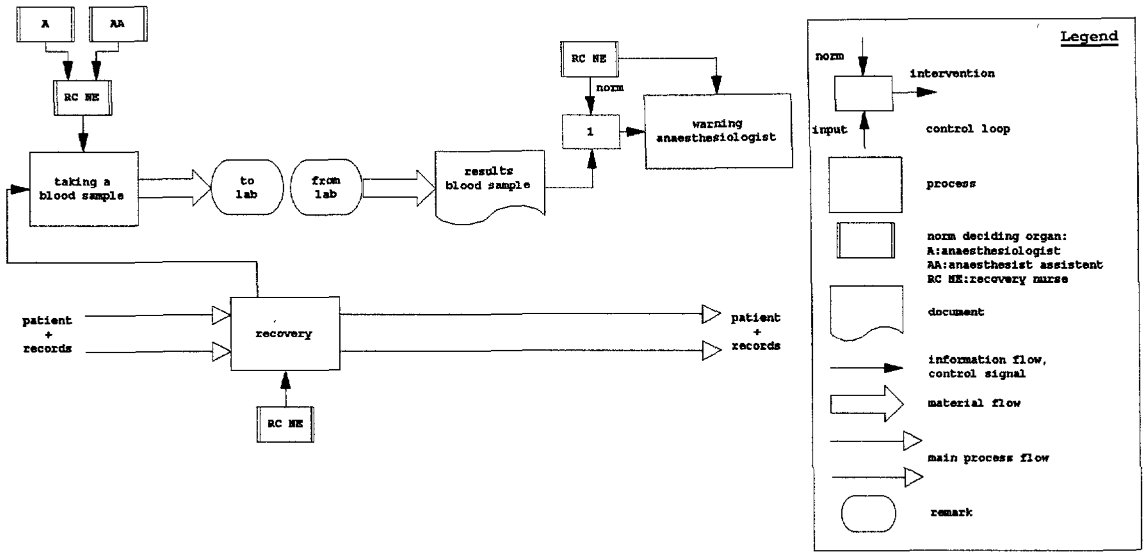


Figure 1 An example of the system approach of In 't Veld: a part from the process model of the SD (taken from the main part "recovery")

3.2 The assessment method for possible risks: the Failure Mode and Effects Analysis (FMEA)

Two possible methods for predicting the risks for patients were examined:

- FMEA [1][3];
- HazOp (Hazard and Operability) [2][3].

The predictions were finally made by using FMEA because of previous experience with the FMEA technique, and the absence of an experienced HazOp leader [7]. The main goal of the FMEA is to predict the nature and frequency of possible risks for patients on a SD.

FMEA is a technique which can be used on a system that can be separated into individual components. These components can be hardware blocks or functional blocks. The person who does the assessment should have a clear understanding of the functions of every component with all inputs and outputs. The failure modes of every component can be examined in a systematic way to establish the effects and causes of the failure modes. This information will be registered on a special form which contains the following columns:

- the failure mode(s) of every component;
- the effects of a certain failure mode;
- the seriousness of these effects;
- the causes of a failure mode;
- the frequency of occurrence of a certain cause;
- the extent to which a certain cause can be corrected.

The original FMEA concept works with one group which consists of members of all the relevant functions. Because of the extent of the group, however two groups were formed. The first group consisted of SD staff members holding functions belonging to the first, second and fourth part of the process model; they did the FMEA for the parts "preparation" and "recovery". The second group held functions belonging to the second and third part of the process model; they did the FMEA for the parts "anaesthesia" and "operating".

Every group did four two-hour sessions:
- the first session was held for the two groups together and consisted of a briefing about what would happen, when and how;
- the second session generated the failure modes, the possible effects and the seriousness of these effects;
- the third session generated the causes of the failure modes, the frequency of occurrence of these causes and the extent to which a certain cause can be corrected.
- the fourth session was used to generate counter measures for the causes with the highest RPN (Risk Priority Number = seriousness • frequency • extent of correction).

The results contain more than 800 causes. These causes were used as input to PRISMA, and classified according to the Eindhoven Classification Model so they can be compared with the results of the Critical Incident Interviews. The classification of the causes was done by three persons. They tried to reach consensus about the classification of a cause. The results were calculated in two forms:
- with a weight-factor for the estimated frequency (see figure 2). This weight-factor compensates the frequency of occurrence for a certain cause;
- without this weight-factor (see figure 3).

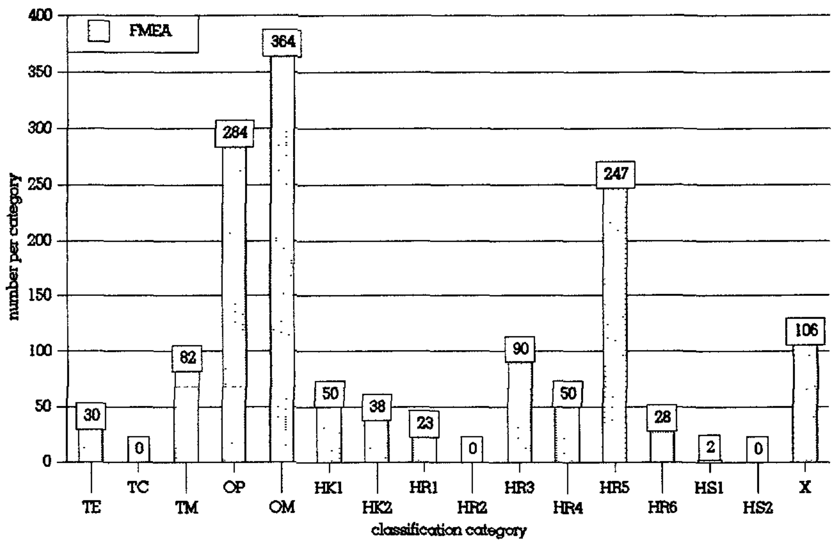


Figure 2 PRISMA profile of classified causes per classification category for both FMEA groups, combined with weight-factors: n=1394 (n=number of classified root causes)

As an important by-product, participating in the FMEA has raised the enthusiasm of the workers in the SD for the risk management project.

4 Stage two: Making an inventory of the real risks for patients

4.1 Critical Incident Interviews

Twenty Critical Incident Interviews (Flanagan [4]) were used to assess the real risks. This technique was used to collect confidential data about incidents, the Interview itself is also confidential.

The Interviews were intended to elicit information about accidents and near accidents which:
- happened recently;

- were closely experienced by the interviewee.

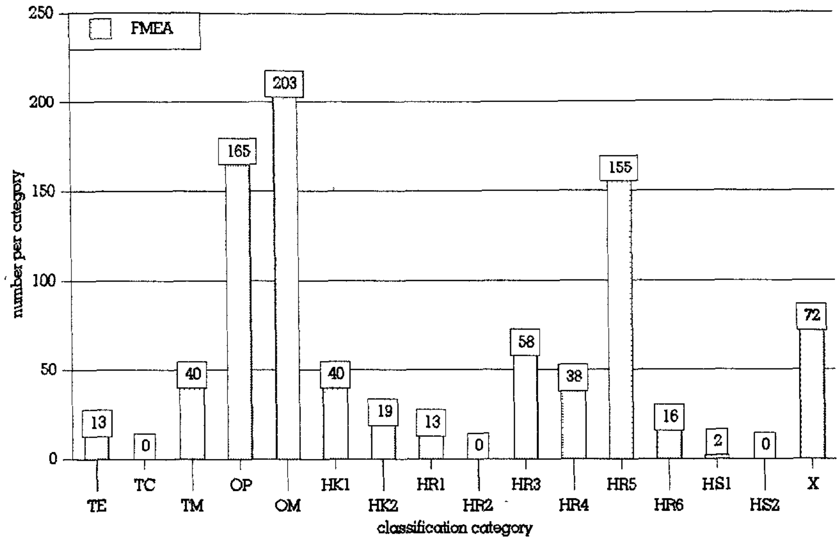


Figure 3 PRISMA profile of classified causes per classification category for both FMEA groups, combined without weight-factors: n=834

The main goal of the Interviews was to make an assessment of the nature and frequency of the risks for patients on the SD. In this way:
- a file was created which is used to get a provisional insight in the nature and frequency of the root causes of incidents on the SD;
- the value of the Eindhoven Classification Model for classifying the root causes of system failure on the SD was checked;
- a reference data base was created whereby:
  - the validity of the predictive FMEA method can be determined;
  - the quality (that is a check on reporting bias) of the future voluntary incident reports can be determined;
- an example could be given to the workers on the department to show them where a voluntary incident reporting system could lead to.

The Critical Incident Interviews were held in two sessions to check the consistency of the results. The first session yielded eleven usable Interviews and the second session six. At least one Interview was held for every kind of function in the SD.

The participants of the FMEA-sessions and the Critical Incident Interviews consisted of two separate groups, except only one person which took part of both the FMEA-sessions and a Critical Incident Interview because this was impossible to avoid.

The 17 usable Critical Incident Interviews were used as input to PRISMA:
1. A Causal Tree Analysis was made from all the Interviews and they were reviewed by the interviewer and the person who was interviewed. When events and/or relations were wrong or forgotten, the Causal Tree was altered.
2. The results of the Critical Incident Interviews yielded 95 root causes (see figure 4: n_{cii}=number of Critical Incident Interviews). The classification of these root causes was done by the same three persons who classified the causes found by the FMEA. The classification was done in the same way as for the FMEA: the classification team members tried to reach consensus about the classification of a root cause.

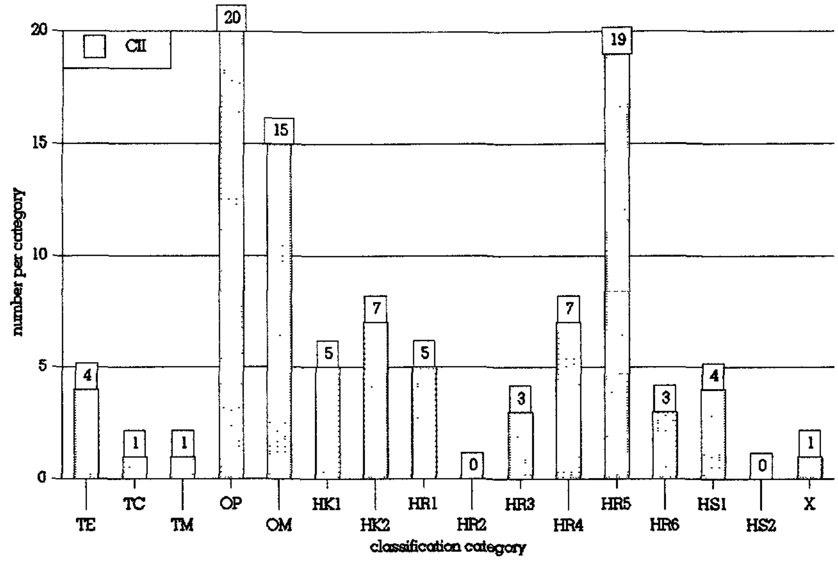


Figure 4 PRISMA profile of classified CII root causes per classification category: $n_{cii}=17$ and $n=95$

The most important results of the Critical Incident Interviews:

- the major part of the incidents is not reported to the FONA commission (FONA: Faults, Accidents and Near Accidents) and therefore unknown to the hospital's management;
- every incident has one or more recovery factors. More than ⅓ of these factors did not lead to successful and complete recovery. The main reason for this is that medical doctors are not listening to recovery oriented remarks of nursing staff;
- a list of systematic failures was made during the Critical Incident Interviews so measures can be taken to eliminate these systematic failures;
- PRISMA is applicable to medical incidents in a SD [6].

Further research on how to integrate and standardise the legally required FONA reports and the voluntary incident reporting system is recommended.

4.2 Comparison between the PRISMA profiles of FMEA and Critical Incident Interviews

Four comparisons were made:

1. between the results of the complete FMEA with weight-factors and the results of all Critical Incident Interviews;
2. between the results of the complete FMEA without weight-factors and the results of all Critical Incident Interviews;
3. between the results of the "anaesthesia" and "operating" sessions of the FMEA without weight-factors and the "anaesthesia" and "operating" results of the Critical Incident Interviews;
4. between the results of the "anaesthesia" and "operating" sessions of the FMEA with weight-factors and the "anaesthesia" and "operating" results of the Critical Incident Interviews.

Comparison 3 and 4 were made for two reasons: first to see if it was possible to make a comparison between the results of the two methods on sub department level (that is the sections "anaesthesia" and "operating") and second, because there were enough Critical Incident Interviews ($n_{cii}=13$) available to make such a comparison possible.

The result of the statistical comparisons is that Spearman's rank correlation coefficients are almost 5% significant (comparison 1 and 2), 5% significant (comparison 3) or even 2,5% significant (comparison 4).

As can be seen in figures 2, 3 and 4, the three classification categories with the most causes are the same for both methods: the organisational factors "operating procedures" (ECM: OP) and "management priorities" (ECM: OM), and the human behaviour factor "planning" (ECM: HR 5).

This stage has demonstrated the potential of a voluntary incident reporting system.

5 Stage 3: The voluntary incident reporting system

The first part is the designing of an incident reporting system based on the theoretical concept of a Near Miss Management System (see also the PRISMA paper) whereby a participant approach is used. Because of this approach, the preferences of the SD can, if possible, be included in the system. In this way the acceptation and success of the system are enlarged. A commission has been appointed which designs the incident reporting system and controls the system when it is operational (this commission is further called the incident reporting commission). It consists of four nurses (which together cover all four main parts of the processes on the SD) and an anaesthesiologist and surgeon as representation of the doctors on the SD. The incident reporting commission was trained in several sessions how to analyse incidents using the PRISMA tools. In this way, they were able to participate in the design of the incident reporting system.

The second part is the setup of an operational reporting system. During the first two months, incidents will be handled by the incident reporting commission under supervision of a member of the Safety Management Group, not only to help the commission but also to alter the design of the reporting system when this is necessary after the evaluation. This evaluation after the first two months is not only to improve the reporting system on the SD but also to get a better starting position for future incident reporting systems on other wards in the hospital. The second and the last part of this stage are still in progress.

6 Stage 4: The development of a set of internal certification norms

The question behind the development of a set of minimum norms is wether it is possible to design a structure and norms which must be satisfied to assure process control for the medical and nursing processes on the ward.

A theoretical framework for a clinical treatment and nursing flow is given in figure 5 [11]. Every policy making moment has a medical and nursing component. A policy making moment must lead to an activity plan. This activity plan consists of a list of activities for doctors and nurses. The relevant doctors and nurses agree and record for every action who is authorized and who is responsible.

This stage of the research is also still in progress but it seems feasible to design a structure like the one in figure 5, as well as establish "quick-scan" norms to achieve a list of uncontrolled processes which can be altered to enlarge the reliability.

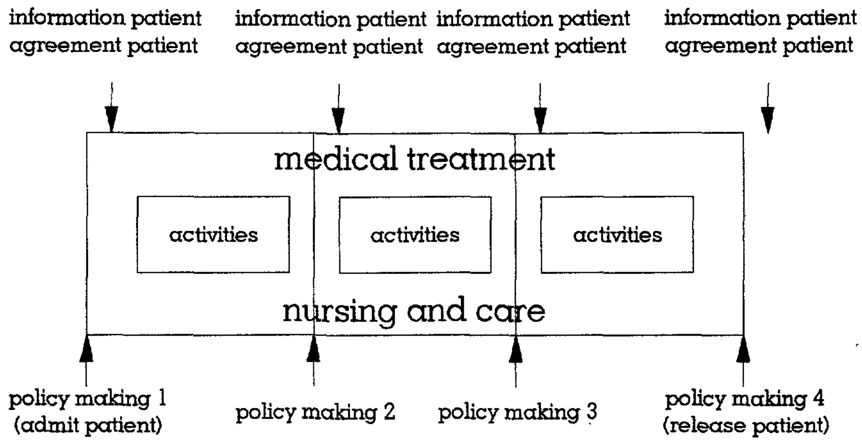


Figure 5 Clinical treatment and nursing flow [11]

7 Stage 5: The development and implementation of a quality system

It is possible to create a quality system with the two risk management instruments (FMEA and incident reporting system) and the quick scan list of norms. The quality system will not be only developed and implemented on the SD but on every ward of the hospital.

Eventually, the process deviation consequences on which the quality system will focus, shall be broadened to include not only unwanted physical consequences for a patiënt but also incidents with regard to negative physical consequences for nurses and doctors, events with negative effects on the environment, and events that cause higher costs for the hospital.

8 Literature

[1] Bij, J.D. van der; Monhemius, L.; Mulder, F.A.; Tuijl, H.F.J.M. van; Wortman, J.C.
Inleiding kwaliteitsmanagement
Eindhoven University of Technology, 1992

[2] Chudleigh, M.F.; Clare, J.N.
The Benefits of SUSI: Safety Analysis of User System Interaction
Cambridge Consultants Limited, Camebridge, England

[3] Dahnjal, S.S.; Fink, R.
Safety Analysis of Clinical Laboratory Systems

[4] Flanagan, J.C.
The Critical Incident Technique
Psychological Bulletin, Vol. 51, No. 4, 1954

[5] Heins, W.
Structured Analysis and Design Technique [SADT]: Application on Safety Systems
MoSHE 1993: chapter 14, Delft University of Technology, 1993

[6] Hoeff, N.W.S. van der
Modellen voor de classificatie van de oorzaken van medische incidenten
Eindhoven University of Technology, Internal Report, May 1994

[7] Hoeff, N.W.S. van der
Risicomanagement op een OK-afdeling; voorspellen vs. registreren van incidentoorzaken
Eindhoven University of Technology, MSc thesis, January 1995

[8] Schaaf, T.W. van der
Near Miss Reporting in the Chemical Process Industry
Eindhoven University of Technology, PhD thesis, 1992

[9] Timmermans, P.P.M.
Risicomanagement op een OK-afdeling; het ontwikkelen van een vrijwillig incidentmeldingssysteem
Eindhoven University of Technology, interim report, April 1995

[10] Veld, J. in 't
Analyse van organisatieproblemen
H.E. Stenfert Kroese bv, Leiden-Antwerpen, 1987

[11] Weegh, G.J. Op de
Klinisch behandel- en verpleegtraject
Catharina Hospital Eindhoven, Internal Report, 1995

Appendix: Definitions of important concepts

The following definitions are used:

- Risk
  The possible events leading to process deviations which can have unwanted physical effects for the patient.
- Risk management
  Striving after process control in a systematic way to minimize the risks for patients.
- Process deviation
  Every deviation from normal in the process which can have unwanted physical effects for the patient; this consists of:
  - all accidents;
  - all near accidents;
  - all complications.
- Incident
  This consists of:
  - all near accidents;
  - all accidents.
- Complication
  A process deviation with temporary or permanent negative physical consequences for the patient of which the cause is unknown.

Because complications need another management system (based on modelling: causes unknown (also see Van der Schaaf [8])) than incidents (based on monitoring), they are not included in this project.

- Accident
  A process deviation that is not a complication and that leads to temporary or permanent negative physical consequences for the patient.
- Near accident / near miss
  A potential accident that
  - either by active (human) intervention
  - or by chance

  turns out well.

- Recovery
  The active (human) intervention and/or the chance whereby a potential accident results in:
  - a near accident or
  - a much smaller accident (where the temporary and/or permanent effects were limited).
- Voluntary incident reporting system
  A Near Miss Management System (Van der Schaaf [8]), based on monitoring. In such a system the root causes of near accidents and of actual accidents are:
  - classified,
  - registered and
  - analysed

  with the purpose of learning from these incidents to prevent or limit the effects of future incidents.
- PRISMA (Prevention and Recovery Information System for Monitoring and Analysis)
  Safety management and incident analysis tool-kit, consisting of:
  - Causal Tree Analysis to obtain the root causes of an incident;
  - the Eindhoven Classification Model to classify root causes in order to create structure in the enormous variety of root causes;
  - the Classification/Action Matrix which suggests effective countermeasures (see also PRISMA paper of Van der Schaaf).

Session 3

Human Reliability Models II

Chairman: T.W. van der Schaaf, The Netherlands

NEAR MISSES AS ACCIDENT PRECURSORS.

A.M.Mulder and T.W. van der Schaaf

Safety Management Group
Graduate school of Industrial Engineering and Management Science
Eindhoven University of Technology
P.O. Box 513, Pav. U8
5600 MB Eindhoven, The Netherlands

Abstract

The topic of this paper is whether, and under which conditions, the causes of near misses can be regarded as identical to accident causes: near misses as precursors to accidents. Usually the assumption is made that the pattern of 'root causes' for accidents and near-misses are the same. However this assumption has never been tested properly. Based on pilot data, empirical research questions are formulated and an experimental design is presented to test the assumption. Two databases are being compared: one containing accident descriptions and causes, the other consisting of near misses. Both sets of data were/are generated by the same production task situations and workforce. This comparison will be made in two ways: 1.<u>quantitative</u>, at the level of relative frequencies of types of root causes; 2.<u>qualitative</u>, comparing the process of incident development from root causes via AND- and OR-gates to the top-event. In this paper, pilot data of a safety project in the steel industry and an experimental design needed for testing the overlap of accident- and near-miss causal factors will be discussed.

1. Introduction

In January 1993 a three-year safety project supported by the European Coal and Steel Community was started at Hoogovens IJmuiden. The main purpose of the project is a structural improvement of the safety level by the implementation of a 'Near Miss Management System (NMMS)', called "SAFER" (Systematic Analysis of Faults, Errors and Recoveries). With this system near misses are registrated, described, analysed and evaluated. The system focuses on prevention of human, organizational and technical failures and on promotion of timely, corrective (human) recovery. There is a special interest for the low-probability, high-consequence risks in High Tech-installations, where no accidents have occurred yet.

An important issue in the project is the investigation of the possibility to get insight in fundamental <u>accident</u> causes by registration and analysation of <u>near misses</u>. It is necessary to make sure that measures based on near misses are effective and contribute to the prevention of actual accidents.

2. Theoretical approaches

In many cases the assumption is made that the pattern of root causes for accidents and near-misses is the same. Actual accidents are often seen as the top of the so-called iceberg. Under the top a lot of near misses happen, whenever luck or recovery stopped the causation proces just in time. In figure 1 a simple incident causation model (Mulder, 1994) is presented near the iceberg-model. The iceberg represents the frequency of occurence of the different sub-parts in the causation-proces. In this figure it is assumed that the causation proces for accidents and for near misses is the same.

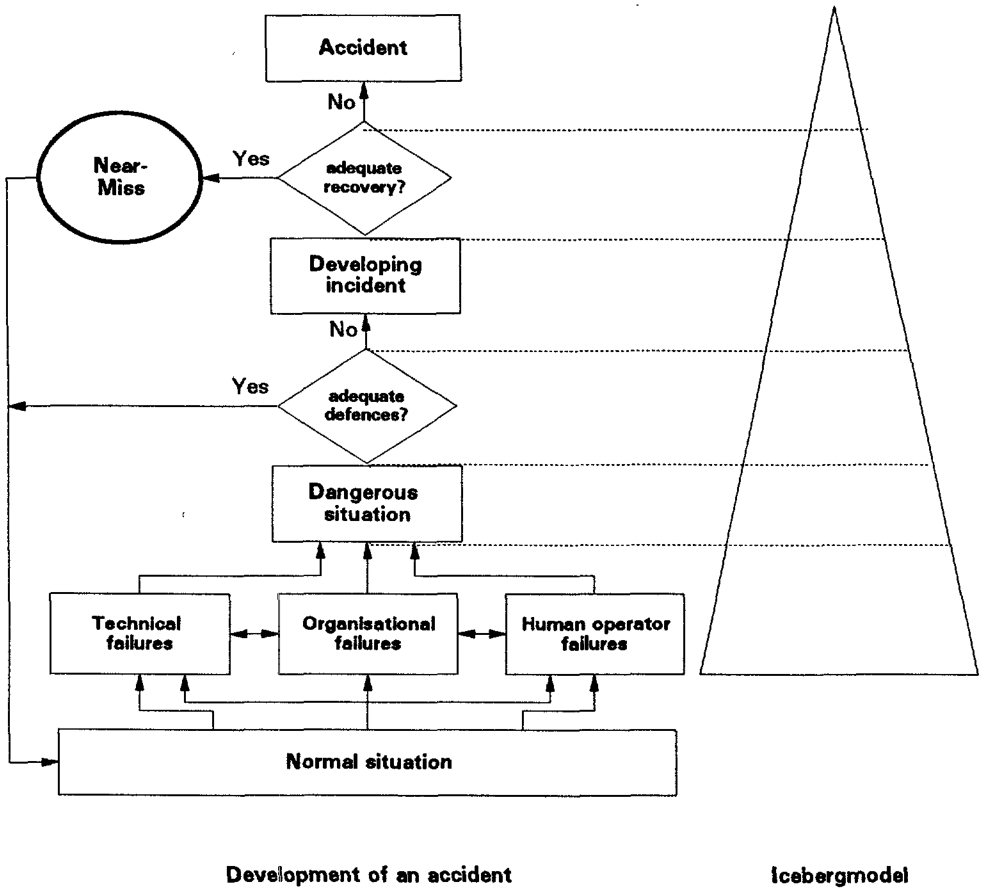


Figure 1: Accidents versus Near misses

This assumption has never been tested properly. In literature the following remarks are made about this issue:

- " One influential tradition in safety (Heinrich 1931) places emphasis on the 'near miss' as a member of the same family as the accident, having the same causes and lacking only the physical damage." (Hale, Glendon (1987) (p.132))
- "The idea behind the pyramids is valid: fatalities are built on the same foundations as Lost Time Injuries and unsafe acts and should not be treated seperately." (Groeneweg, 1994). However no proof or reference is given.
- "Review of comparisons between situations where conflicts and emergency manouevres (but no harm) occur and situations where actual harm occurs show them to be similar." (Hale, Glendon (1987) (p.17))
- "The much more numerous unsafe situations (both chronic and sudden) and even more abundant human errors not resulting in serious consequences are assumed to have the same psychological root causes as the tiny subset that actually develops into an accident." (Van der Schaaf, 1992)
- "Another vital assumption is that the three levels (accidents, near misses, dangerous situations) of the iceberg are directly related in the sense that they show largely overlapping sets of 'root causes'." (Van der Schaaf, 1992)
- Different studies have proven that information gathered by investigation of near

misses can be used for prevention and reduction of actual accidents. (Bird, 1974)

- "Near Misses are not equal to accidents because with a near miss there has been a recovery. This recovery is not a coincidence but a consequence of a to the proces inherent boarding propability." (Hale, 1995)

3. Pilot studies

The PRISMA (Prevention and Recovery Information System for Monitoring and Analysis of Safety Factors) approach to safety management (van der Schaaf, 1995b) developed in and for the chemical process industry had to be redesigned for implementation in the steel industry. Adjustment of the organisational part of the classificationmodel (Van Vuuren, 1995) resulted from analysis of 'Critical Incident Interviews (CII)' (in two plants, 25 CII/plant). In figure 2 the results of the CII's for plant A are presented. Besides the CII's also usable accident-reports of the same plant were described in causal trees. These results are presented in figure 3. More detailed results are shown in Appendix 1.

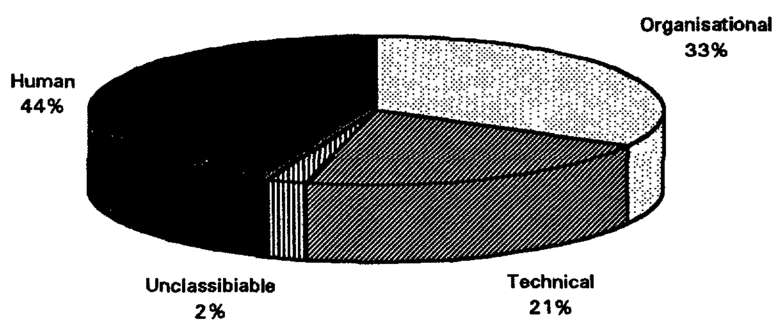


Figure 2: Distribution of 154 failure factors in 25 near misses in plant A.

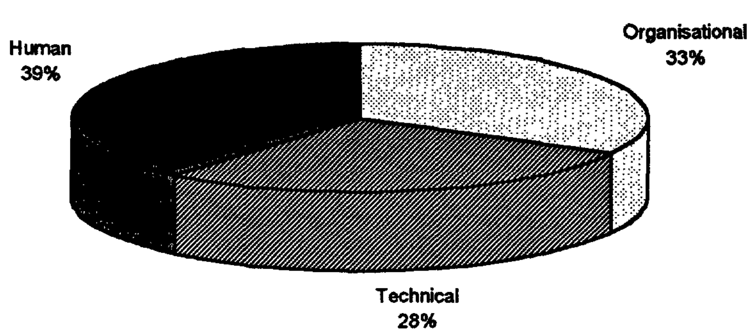


Figure 3: Distribution of 83 failure factors in 14 accidents in plant A.

The results point in the direction of a large overlap in frequencies of root causes of near misses and accidents as far as main categories are concerned. However, an experimental design is needed to give grounded answers to the empirical research question.

Concerning the pattern of causes the following question can be asked: Is it true that only when the difference between accidents and near misses is luck or coincidence their causal patterns are the same? This can be separated in <u>two questions</u>:

1. Is it more than only luck or coincidence what makes a situation a near miss instead of a real accident, what kind of recovery finds place?
2. And if so, in what way does this influence the causal-pattern of accidents and near misses?

<u>Question 1:</u>
A positive answer to the first question can be given based on pilot data. Pilot data from the steel industry (Mulder, 1994) showed that recovery is not always a matter of luck, but that in a lot of near misses different recovery-aspects could be distinguished.
Recovery can be classified according to (Van der Schaaf, 1995a):

- type of recovery factor: technical design of process and of man-machine interface; organizatinal and management factors or human operator factors.
- operator reaction after symptom detection. In this way recovery can consist of three phases: detection of symptoms, localisation of their cause(s) and correction to return the system to its normal status.

From the same set of the 25 near misses presented in figure 2, the recovery factors were classified. The results of the classification according to type of recovery factor are presented in figure 4. Figure 5 presents the classification results according to operator reaction after symptom detection.

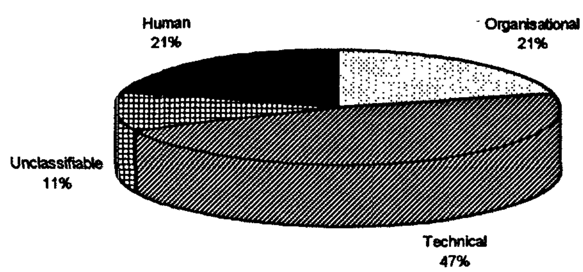


Figure 4: Distribution of 34 recovery factors in 25 near misses in plant A.

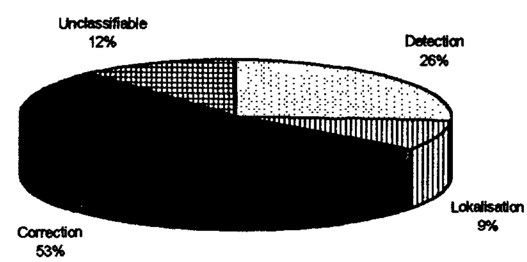


Figure 5: Distribution of 34 recovery factors according to recovery phase in 25 near misses in plant A.

Also in a number of other case studies (an energy production plant and a surgical ward) recovery seems to be more than only luck. For more detailed information on human recovery of errors see Van der Schaaf (1995a).

<u>**Question 2**</u>

The figures presented above show that not only luck or coincidence makes the difference between accidents and near misses. Now it becomes necessary to get more insight in the way recovery influences the causal pattern of near misses and accidents. Therefor, the following examinations:

1. Investigation of near misses to see why and exactly when recoveries took place.
2. Investigation of accidents to see why no recovery did not took place, and if there was any possibility to prevent or recover at all.

4. Empirical research questions

Main research question: whether, and under which conditions, the causes of near misses can be regarded as identical to accident causes: near misses as precursors to accidents.

Research on the main empirical question should produce two different products: insight in the overlap in sets of root causes and a method to come to qualitative analysis.

4. 1. <u>Insight:</u> the extent to which the overlap of causal-patterns of near misses and accidents prove to be true. This insight refers to two different levels of the causal tree. In figure 6 a fictive example of a causal tree is presented.

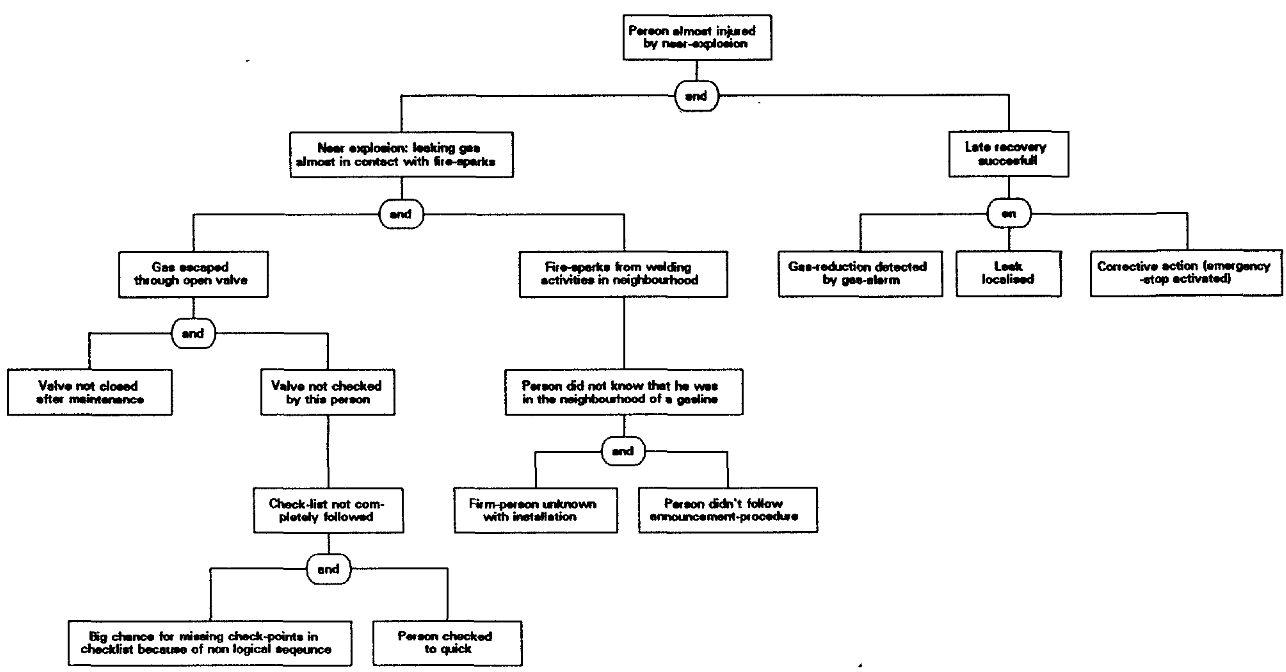


Figure 6: Causal tree example

The first level regarding overlap of frequencies of root causes is concerned with the bottom of the causal tree (see figures 2,3 and 4). The second level refers to the overlap in causal-*pattern* of the total causal tree.
For this second level questions like the following are important:

- In what stage of the developing process of an accident have recoveries started and succeeded?
- How did certain organisational root causes develop over time? Which other organisational, human and technical factors where caused by these organisational root causes?
- Which are the frequencies of combinations of factors in causal trees? For instance combinations of factors connected by an AND-gate horizontally or vertically.

2. Method: For the insight regarding to causal-*pattern*, a method needs to be developed for quantitative analysis of qualitatively complex information presented in the causal tree. This method shall be too comprehensive to use for every selected near miss, but can be very usefull in cases of complex, unique (often low probability, high-consequence) near misses.

5. Experimental design

Two databases are being compared: one containing accident descriptions and causes, the other consisting of near misses. Both sets of data were generated by the same production task situations and workforce. This comparison will be made in two ways:

1. quantitative, at the level of relative frequencies of types of root causes;

2. qualitative, comparing the process of incident development from root causes via AND- and OR-gates to the top-event.

An important issue is on which aggregation-level the comparison should be done:

- Comparison based on matched pairs, that means a single near misses is compared with a single similar accident.
- Comparison based on samples, that means comparing a large group of near miss data with a large group of accident data.

In this case comparison will be based on the second aggregation-level, because of the impossibility to find selection criteria for comparising near misses and accidents based on pairs. Selection based on same consequences is not possible. The NMMS assumes that consequences happen very late in the developing proces and can be very different (damage to production/environment or human). The overlap should consist in the different steps of the developing-proces itself, therefore consequences can not be the determining part in this case.

6. Methodological aspects

The empirical research questions bring forward two methodological questions:

1. Reliability and validity of information gathered in the form of causal trees.
   At least the following types of reliability will be tested:
   - Same people building causal trees of the same near miss/accident.
   - People at different organisational levels building causal trees of the same

near miss/accident.

- Comparison of accident reports with causal trees of the accident.
- Comparison of building causal trees after filling in an accident report with building trees immediately after it happened.

2. Minimum number of near misses/accidents for sensible statistical analaysis.
An important issue here is that a single near miss or accident consists of more than one root cause. So besides the total number of near misses there is also the total number of root causes that plays an important role.

Pilot studies in the steel industry as well as in the chemical process industry have shown that 25 near misses, with an average total of 160 root causes, were enough to obtain a stable pattern of relative frequencies.

7. Conclusion and future research

Pilot data showed that there is more than only luck that makes the difference between accidents and near misses. A disctinction in sorts of recoveries can be made. A classification-model for recovery factors is needed to get insight in the recovery-proces. Now it is time to examine in which way recovery is influencing the causal pattern of near misses and accidents. The empirical research to test the assumption of near misses as precursors of accidents in the way it is described in this paper will be done during the rest of 1995. The software necessary to build a database is already developed and in place. The database is feasible, and seems to be a usefull management-tool.

Future topics:

1 Recovery modelling and the integration of classification into PRISMA.
2 Importance of safety culture on the mechanism and the effects of a NMMS.
In what way is a NMMS working? Is there a difference between short term and long term effects, on which goal (modeling, monitoring, alertness) should be emphasised? After receiving the answers, the question can be asked how safety performance of an organisation is influenced by a) an improved safety consciousness of the employees and management and how by b) structural measures based on statistical patterns of structural root causes.

Acknowledgement

The SAFER research project is supported by a grant from the European Coal and Steel Community.

References

Bird, F.E. (1974). *Loss Control Management.* International Loss Control Insitute, Logan ville.

Groeneweg, J. (1992). *Controlling the controllable.* Leiden: DSWO press.

Hale, A.R. and Glendon, A.I. (1987).*Individual behaviour in the control of danger.* Elsevier, Amsterdam.

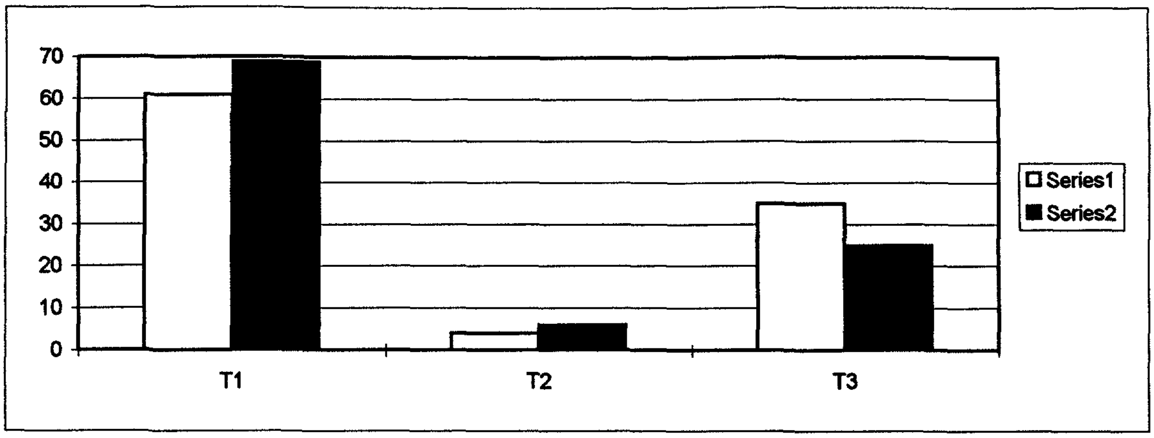
Heinrich, H.W. (1931). *Industrial accident prevention.* MCGraw-Hill, NewYork.

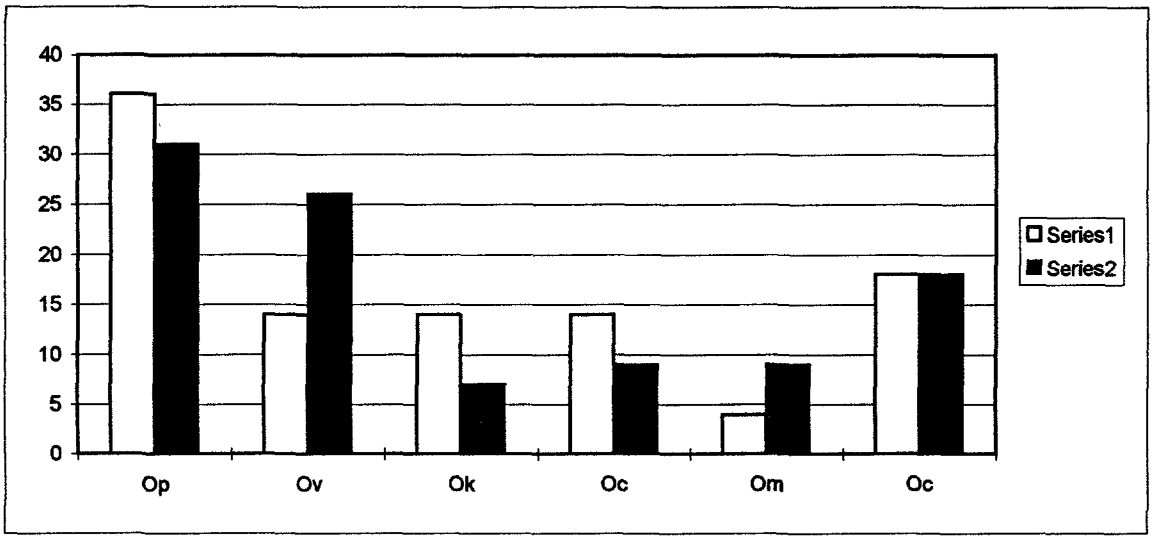
Mulder, A.M. (1994). *Progress report 1993 of the SAFER project.* Report Hoogovens steel industry (in Dutch), IJmuiden.

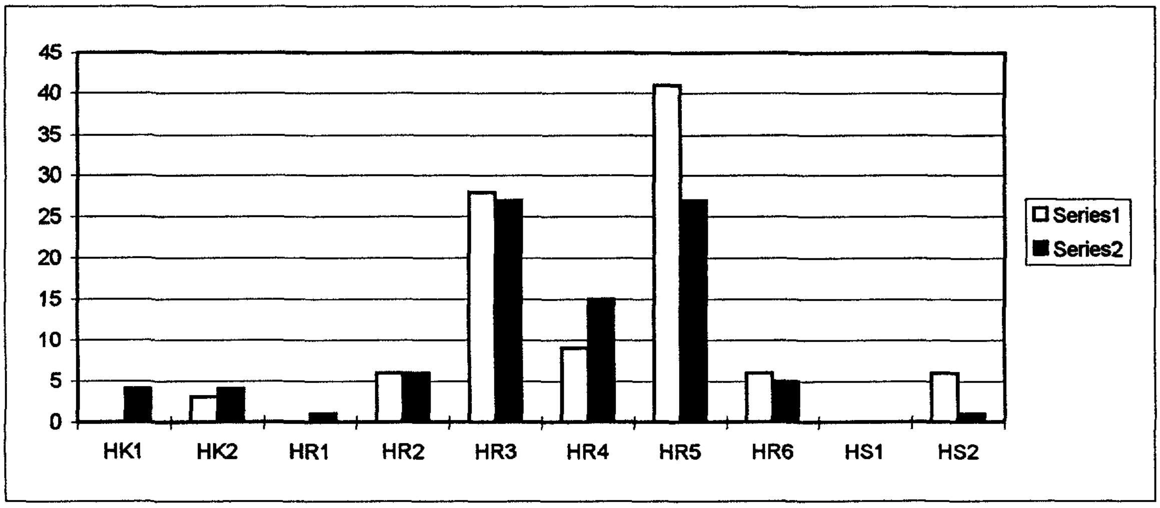
Schaaf, T.W. van der (1992). *Near miss reporting in the chemical proces industry.* Ph.D. Thesis, Eindhoven University of Technology.

Schaaf, T.W. van der (1995a).*Human recovery of errors in man-machine systems* . MIT.
Schaaf, T.W. van der (1995b). *PRISMA-SAFETY*. This volume.
Vuuren, W. van (1995). *Modelling organisational incident factors in complex technical systems.* This volume.

Appendix 1







Detailed comparison of PRISMA-profiles of failure root causes for 14 accidents (series 1) and 25 near misses (series 2) in a steelplant.

MODELLING ORGANISATIONAL INCIDENT FACTORS IN COMPLEX TECHNICAL SYSTEMS

W. van Vuuren

Safety Management Group
Graduate School of Industrial Engineering and Management Science
Eindhoven University of Technology
PO Box 513, Pav. U-42
5600 MB Eindhoven, The Netherlands

Abstract

To show the importance of modelling organisational incident factors in complex technical systems, two pilot projects were carried out within the Dutch chemical process and steel industries. Too often only human and technical factors are seen as the main contributors to industrial safety. The pilot projects have shown that organisational factors cannot be neglected, and that insufficient analysis of incidents (both accidents and near misses) is a prime reason for not recognising such factors. Based on these results, a PhD project has been started to develop an explanatory model for the organisational causes of safety-related incidents in complex technical systems.

1. INTRODUCTION

Accident investigation has always played an important role in the improvement of complex technical systems. If the causes of accidents are subdivided into three groups (technical, human and organisational), it appears that changing *technology* was the classic approach to prevent accidents. Based on the assumption that man will always make mistakes, engineers tried to make technical systems 'foolproof'. When the accident rates did not decline to acceptable levels, researchers started to focus on *human behaviour* as a cause of failure. Only during the last decade have the importance of *organisational and management factors* as causes of incidents been accepted [Reason, 1991]. However, this growing interest in the organisational causes of system failure still has not led to any widely accepted explanatory theoretical model. Neither are there any practical tools for finding, describing, classifying and correcting organisational factors. The few real attempts were made by researchers with a psychology background [Wagenaar, *et al.*, 1995], or came from technical risk analysis projects. Many questions remain unanswered.

The first part of this paper will focus on the results of two pilot projects. These projects were started to seek answers to two research questions.

1. How significant are organisational factors, relative to human and technical factors, as causes of safety related incidents?

This significance was expected to be high. However, in an overview given by Wagenaar [1983], the predominance of 'the human factor' (80 to 100%) in accident research is clearly shown. This leads to the second research question.

2. Why does almost all in-company accident research suggest that only human and technical factors are important?

It was predicted that it takes more than just a rough investigation to come up with organisational causes. Technical and human types of failures are easy to detect, because the time between such failures and the real incident is usually limited. Using the distinction made by Reason [1990] these failures would be termed: 'active errors', because the effects of active errors are felt almost immediately. Organisational failures, on the other hand, normally occur much earlier in time and the effects lie dormant within the system, waiting for the 'right' circumstances to manifest themselves. Reason [1990] would term such failures: 'latent errors'. Because of this time delay, organisational failures are much harder to detect. Therefore, in one project current methods of investigating incidents were also examined.

The first pilot project, which was carried out in the Dutch steel industry [Van Vuuren, 1993], concentrated on the first research question. The second pilot project, which was carried out in the Dutch chemical process industry [Van Vuuren, 1994], focused on the second research question. Besides answering the two research questions, the data collected in the two projects was also used to develop tools to describe and classify the organisational factors. For both companies, the practical aim was to set up a 'near miss' reporting system to improve the safety levels in each by gaining greater insight into the causes of incidents. However, this practical objective will not be discussed in this paper.

Based on the results of the two pilot projects, a PhD project has been started to model the organisational causes of safety-related incidents in complex technical systems. This PhD project will be described in the second section of this paper. The paper will end by drawing some general conclusions.

2. STEEL INDUSTRY PROJECT

2.1 Plant profile

For practical reasons, such as the considerable size of the steel plant, only the coking plant was asked to participate in the project. The coking plant produces coke for the steel-making process. The plant also purifies the gas created in the coking process, to keep the air emissions well below the levels permitted by law, and to extract valuable by-products from the gas. There are approximately 300 employees working at this coking plant, divided into a production group and a maintenance group.

2.2 Aims of the research project

As stated earlier, the data collected in the coking plant was used to:

1. answer the first research question: how significant are organisational factors, relative to human and technical factors, as causes of safety related incidents?
2. develop tools to describe and classify the organisational factors.

2.3 Method

Since setting up a near miss reporting system was the practical aim, near misses were used as input for answering the research questions. Near misses are usually estimated to occur one or two orders of magnitude more frequently than actual accidents. So by using near misses, it is possible to collect a lot of information within a short period. The assumption was made that near misses have the same causes as real accidents.

The near misses were reported voluntarily in confidential 'critical incident interviews'. This technique is based on the 'critical incident technique' [Flanagan, 1954], a technique originally developed to trace the 'critical job requirements' for such purposes as training and selection of personnel. In a similar way, it is possible to look at the critical activities or decisions in the development of a near miss. The interviews were carried out on a confidential basis, minimising the reluctance to report and later to talk about the near misses. A selection was made to get an equal number of 'maintenance' and 'production near misses'.

After interviewing the people involved in each near miss, the information was used to build a 'causal tree' (see fig. 1). Causal trees, derived from fault trees, are very useful to present the critical activities and decisions in a chronological order. They also show how all different activities and decisions are logically related to each other. For near misses, causal trees are divided in two parts: the 'failure side', which gives an overview of all the activities that have led to the failure, and the 'recovery side', which gives an overview of the activities that prevented it from developing into a real accident.
By using causal trees it becomes clear that there is never one single cause of a near miss. It is always a combination of many technical, organisational and human causes. The 'root causes' of the near miss, which are found at the bottom on the failure side of the causal tree, are the main products of the analysis. These root causes should be used to decide which preventive measures to take.

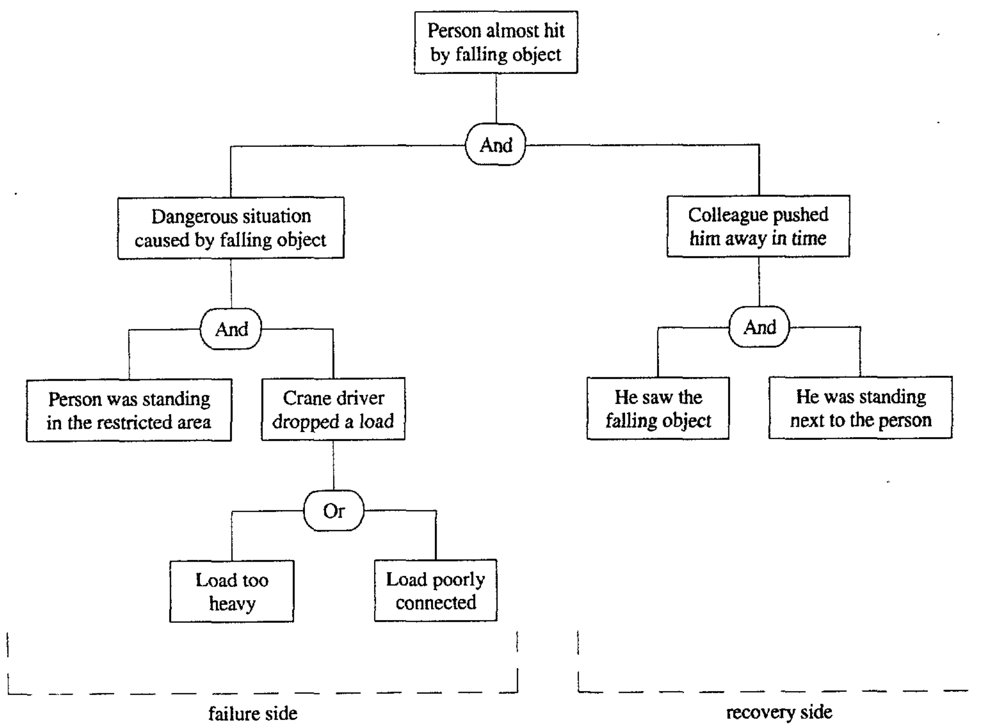


Fig. 1. Example of a causal tree describing a near miss.

To be able to take these preventive measures, it is first necessary to classify the root causes. Without classifying the root causes, the information will be too descriptive and therefore too unclear to be of use. By classifying the root causes, only a limited number of technical, organisational and human categories remain. As a starting point, the 'Eindhoven classification model' was used (see fig. 2). This model was originally developed for the chemical process industry, with the main focus on human failure. Therefore, the model had to be adjusted for the steel industry. An extension of the organisational categories was also needed.

To adjust and extend the classification model the technique of analytic generalisation was used [Yin, 1991]. Analytic generalisation uses the following iterative nature of 'explanation building':

* make an initial statement or an initial proposition
* compare the findings of an initial case against such a statement or proposition
* revise the statement or proposition
* compare the revision to the facts of the second, third or more cases *and*
* repeat this process as often as needed.

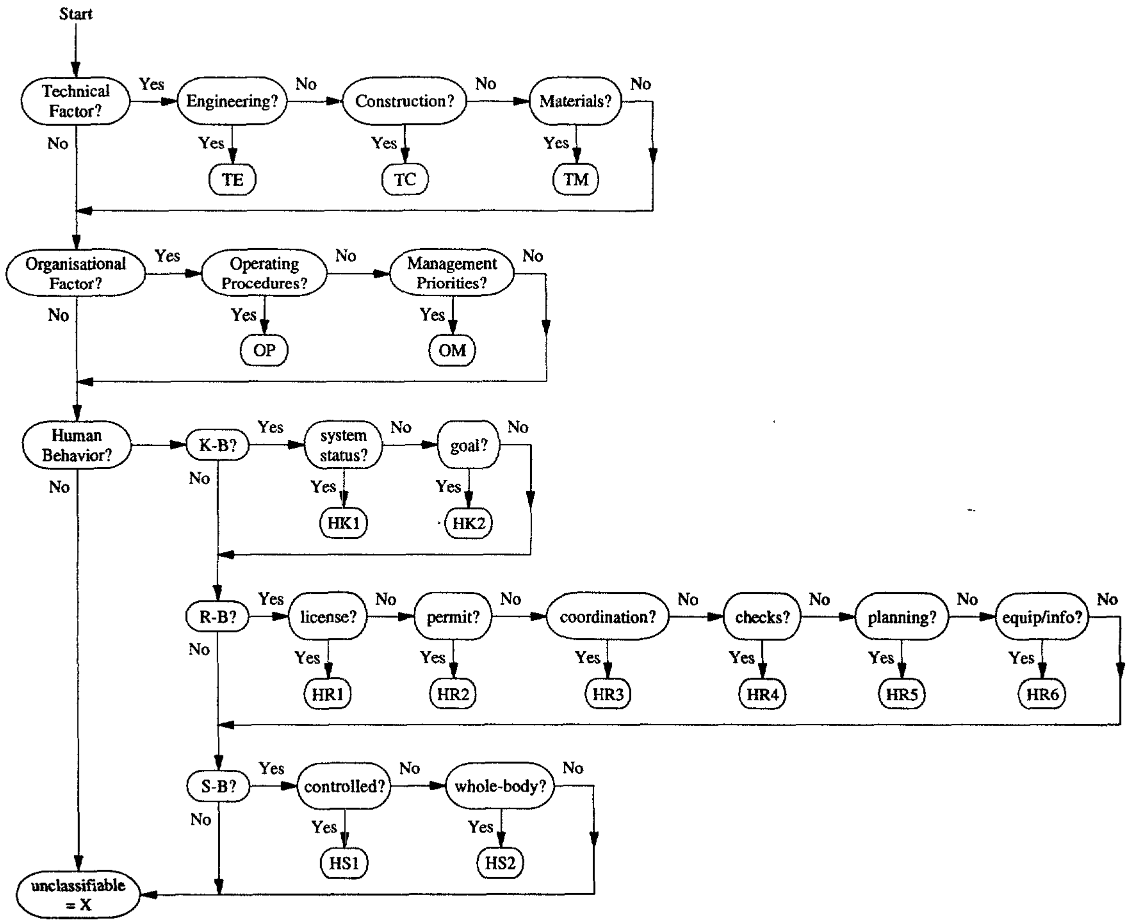


Fig. 2. The Eindhoven classification model [Van der Schaaf, 1992].

In this research project, the Eindhoven classification model was used as the initial 'statement'. Instead of using an initial case, 15 near misses were investigated and used as input to the first revision process. After the first revision of the model, five new near misses were investigated, and the model was revised again. This process was continued until the model did not change any more.

2.4 Results and discussion

After 25 near misses, the classification results stabilised and enough data was collected to answer the first research question. By describing 25 near misses, 164 root causes were identified and classified according to the adjusted classification model. The distribution of the root causes is presented in fig. 3. On the basis of these results it can be concluded that the organisational causes are important and should not be neglected. In this study they even appeared more frequent than the technical causes.

The second aim was to develop tools to describe and classify organisational factors. Critical incident interviews and causal trees were used for the description of the near misses. The company policy had been to use standard forms to analyse its incidents. However, this means that only predefined questions are asked. By using critical incident interviews and causal trees, a reconstruction of the incident takes place, and the incident investigator is free to ask any question that can help him to make this reconstruction. In this way, the description will be much better, and more comprehensive information is gained about the causes of the incident.

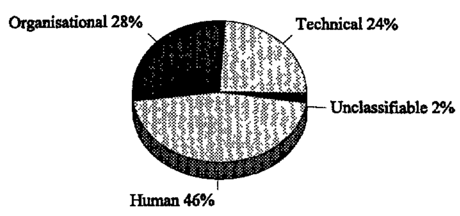


Fig. 3. Distribution of 164 root causes in the steel industry project.

The second tool that was developed was the adjusted and extended classification model. To adjust the existing model to the steel industry, only a limited redefinition of the categories was needed. It was necessary to extend the Eindhoven classification model on the organisational part. With the two original categories 'operating procedures' and 'management priorities' it was not possible to classify all organisational root causes. After the process of analytic generalisation, four organisational categories were added to the two existing ones. Within the following six categories it was possible to classify all the organisational root causes.

* *Management priorities*: refers to any kind of pressure by management to let production prevail over safety. For example, maintenance personnel often have to work during short planned production stops.
* *Responsibilities*: refers to the inadequate allocation of responsibilities to persons. Especially in emergencies, this can lead to the omission of vital actions.
* *Operating procedures*: refers to the inadequate quality of procedures. (not whether they are followed or not) Critical factors are completeness, accuracy and ergonomically correct presentation of operating procedures.
* *Transfer of knowledge*: refers to the lack of education about situational knowledge, which is not included in the operating procedures. Especially new personnel have to be informed about things like system characteristics, known dangers and the existing operating procedures.
* *Bottom-up communication*: refers to all the bottom-up signals that do not result in any reaction. For safety and motivational reasons it is very important to listen to ideas or complaints from the shop floor.
* *Culture*: refers to the inappropriate way of acting by a group. Groups within a company are used to do certain tasks 'their way', and not the prescribed way.

3. CHEMICAL PROCESS INDUSTRY PROJECT

3.1 Plant profile

The main products of this plant are Propylene Oxide (PO) and Tertiary Butyl Alcohol. Hundreds of everyday commodities, ranging from cosmetics to antifreeze and furniture cushions to car bumpers, contain PO or PO derivatives. Tertiary Butyl Alcohol is used primarily to produce the gasoline component Methyl Tertiary Butyl Ether, which has good octane-improving properties and facilitates the production of unleaded petrol. Since both raw products are extremely flammable, a high safety level is a top priority of management. All employees (about 350) were asked to participate in the project, including all employees of the maintenance contractors.

3.2 Aims of the research project

Just as in the first project, the practical aim was to set up a near miss reporting system. The collected data was used to:

1. answer the first research question in a different domain.
2. answer the second research question: why does almost all in-company accident research suggest that only human and technical factors are important?

Also, the refinement of the tools described in the previous chapter was considered as an important aim of this project.

3.3 Method

For this project, the same approach as in the steel plant was used. Critical incident interviews, causal trees and classification of root causes were used to collect the data and to answer the research questions. New to this project was explicit attention to the methods in which the company used to analyse its incidents, and the results of these analyses. This attention was necessary to answer the second research question. As stated before, an insufficient level of analysis of incidents was expected to be a prime reason for the low percentage of organisational failure found in most of the companies. To check this expectation, two accidents that had already been investigated by the company were used for the project. For these accidents, the investigation was conducted all over again, in the way described in the previous chapter. After the new investigation, a comparison of the results was made.

3.4 Results and discussion

During this project, 24 near misses were investigated, and showed a total number of 138 root causes. The distribution of the classified root causes is presented in fig. 4. This leads to the same conclusion as in the first project, namely that the contribution of organisational factors is significant and deserves attention during incident analysis.

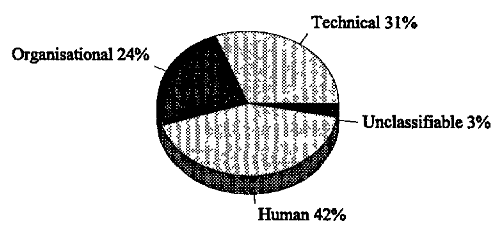


Fig. 4. Distribution of 138 root causes in the chemical process industry project.

To answer the second research question, two serious accidents already investigated by the company were investigated all over again. All the results for the two accidents are summarised in table 1.

In the original investigation, eight root causes were identified, six of which were classified as human causes. Only one root cause was classified as organisational. During the new investigation, 21 root causes were identified, of which five were classified as technical, ten as human and six as organisational. This distribution is comparable to that shown in fig. 4.

	Original results (number of root causes)			New results (number of root causes)		
Types of root causes	Techn.	Org.	Human	Techn.	Org.	Human
Accident 1	1	0	2	4	4	4
Accident 2	0	1	4	1	2	6
Total	1	1	6	5	6	10

Table 1 Results of the comparison.

It was not possible to study more than two accidents to help in answering the second research question. To obtain valid information, very recent accidents were needed. Only two accidents could satisfy this criterion. In spite of this limited number, the comparison between the original and new results convinced management to start using the new approach for its incident investigation. The original investigation method was shown to be inadequate. Results were incomplete and only covered the top level of the causal tree, which mainly addressed human and technical failures.

4. FOLLOW-UP RESEARCH

The two pilot projects clearly showed the importance of organisational failure. Also, the need for an extension of the existing classification model was shown. By using analytical generalisation, four organisational categories were added to the classification model. However, the new classification model only uses *'descriptive'* categories for organisational root causes. Using these categories, it is only possible to add a classification to the root causes, and this way simplify reality. Ideally, the categories should serve as a link between root causes and preventive counter measures. For this an 'explanatory' model has to be developed. Therefore, a PhD research project has been started with the following aims:

1. to develop and test an *explanatory model* for the organisational causes of safety-related incidents in complex technical systems, in order to be able to propose more effective and more efficient measures to increase safety performance
2. to find out which difficulties organisations have with detecting, analysing and interpreting 'organisational failure'.

The second aim partly arose from the two pilot projects. In second pilot project, it was clearly shown that an incomplete description of an incident will lead to problems in detecting organisational failure. It is also expected that unfamiliarity with organisational failure will lead to detection problems. If you are only used to looking for human and technical failures, it is very easy to overlook organisational failure. This is also true for the interpretation phase.

5. CONCLUSIONS

At the outset, two research questions were posed. After carrying out the two research projects described in this paper, it can be concluded that the results met the expectations. Based on the first research question, both projects were started with the expectation that the contribution of the organisational factors would be significant. The percentages found (28% in the steel industry and 24% in the chemical process industry) confirm this expectation. To make the best use of this information, the classification model was successfully extended to six descriptive organisational categories. Within these categories it was possible to classify all organisational categories.

For answering the second research question, the method of analysing incidents was investigated. It was expected that companies do not investigate deeply enough to identify organisational factors. The results of the second project did confirm this expectation. Only the top of the causal tree was considered during an original investigation. This top mainly consists of non-organisational factors. During a new investigation not only more root causes, but also organisational root causes were found. Causal trees were shown to be very valuable in describing incidents in an orderly, comprehensive way.

Incident analysis as described in this paper is very useful for risk management in complex technical systems. By only using predictive models like Fault Tree Analysis, a unique opportunity to learn from accidents or near misses is missed. Especially for low-probability, high-consequence risks it is crucial to learn as much as possible from near misses to prevent real accidents.

REFERENCES

Flanagan, J.C. (1954). The Critical Incident Technique. *Psychological Bulletin,* **51**, 327-358.

Reason, J.T. (1990). *Human Error.* Cambridge University Press, Cambridge.

Reason, J.T. (1991). Too little and too late: a commentary on accident and incident reporting systems. In: *Near miss reporting as a safety tool.* (Schaaf, T.W. van der, D.A. Lucas and A.R. Hale, (Ed.)), 9-26. Butterworth Heinemann, Oxford.

Schaaf, T.W. van der (1992). *Near miss reporting in the chemical process industry.* PhD thesis, Eindhoven University of Technology.

Vuuren, W. van (1993). *Near miss reporting in the Dutch steel industry (in Dutch).* Master thesis, Eindhoven University of Technology.

Vuuren, W. van (1994). *Near miss management in the chemical process industry (in Dutch).* Internal report, Eindhoven University of Technology.

Wagenaar, W.A. (1983). *Human error (in Dutch).* Inaugural lecture. University of Leiden.

Wagenaar, W.A., J. Groeneweg, P.T.W. Hudson and J.T. Reason (1995). Promoting safety in the oil industry. *To appear in Ergonomics.*

Yin, R.K. (1991). *Case study research: Design and methods.* Applied Social Research Method Series, Sage Publications.

REDUCING THE CONTROLLED FLIGHT INTO TERRAIN RISK BY HUMAN ERROR MANAGEMENT

Frank C. Dijkgraaf

Delft University of Technology, Faculty of Aerospace Engineering, P.O. Box 5058, 2600 GB, The Netherlands

Abstract. In the light of the latest airline scheduled airline traffic growth forecasts, significant improvements will have to be introduced to maintain the public's confidence in the safety of air travel. So-called "Controlled Flight Into Terrain", in which airworthy aircraft are inadvertently flown into the terrain (or water) with little or no awareness by the pilots, cause the major part of airline passenger and crew fatalities today. This paper categorises the causes leading to this specific type of aircraft accident and suggests improvements in the area of cockpit man-machine-interface and warning system improvements capable of reducing the CFIT-rate in future.

Keywords. Controlled flight into terrain, Human error management, Ground proximity warning systems, Synthetic/Enhanced terrain displays.

1. INTRODUCTION

Air travel is generally accepted as a safe means of transportation. According to the latest traffic growth forecasts however, the civil aviation safety record, in terms of yearly passenger fatalities or aircraft hull-losses, could deteriorate. Until the year 2003, the International Civil Aviation Organisation (ICAO) assumes a moderate 5% long-term traffic growth for world scheduled passenger traffic. As the averaged annual passenger fatalities and fatal accident rates in the 1984 trough 1993 period did not change substantially, extrapolation, based on the trend in traffic growth and accident statistics over this period of time, presents reasons to be concerned about the future flight safety record. According to this extrapolation, ICAO (Corrie, [4]) expects the number of passenger fatalities to increase with 74% to 1,200 per year, and the number of fatal accidents to increase with 77% to 40 per year by the year 2003. In order to maintain the public's confidence in the safety of air travel, significant safety improvements will have to be introduced.

So-called "Controlled Flight Into Terrain" (CFIT) accidents remain the number one cause of passenger and crew member fatalities today. This paper categorises the causes leading to this specific type of aircraft accident and suggests improvements in the area of cockpit man-machine-interface and warning system improvements capable of reducing the CFIT-rate in future.

2. CONTROLLED FLIGHT INTO TERRAIN

Definition and statistics

Adopting the definition as compiled by the Flight Safety Foundation's CFIT task force, a controlled flight into terrain or CFIT accident occurs when "an airworthy aircraft is inadvertently flown into the terrain (or water) with little or no awareness by the pilots". Since the advent of the jet transport in 1958, well over 9000 lives have been lost in CFIT accidents world-wide. Although reduced in number by the implementation of the ground proximity warning system (GPWS), CFIT accidents continue to occur. GPWS is a terrain warning system that alerts the crew whenever their aircraft's terrain clearance becomes endangered. Recent data (1988-1993 period) show that, in world-wide airline operations, 54% of all passenger and crew fatalities were a result of controlled flight into terrain. These fatalities were caused by 28 CFITs out of a total number of 76 fatal accidents (figure 1). The number of CFIT accidents per million flights by region is depicted in figure 2.

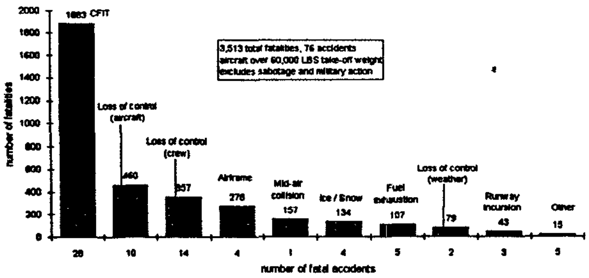


figure 1, World-wide airline fatalities, classified by type of accident (1988-1993), source: Boeing, in Hughes, [7].

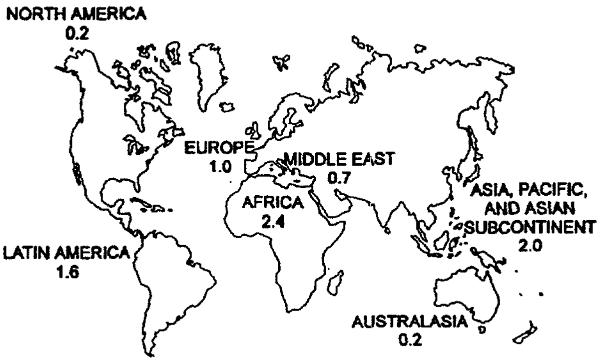


figure 2, CFIT Accident Hull Losses per Million Flights by Region (1959-1992 data), Source: Boeing Commercial Airplane Group in: Weener [10].

Over the period 1959-1992, the number of hull-losses due to controlled flight into terrain in Europe was five times higher than that in the North America region. This might be a result of earlier implementation of GPWS in the USA and the availability of a terrain-warning system for Air Traffic Controllers, the Minimum Safe Altitude Warning System (MSAWS). MSAWS is a software package for use at Automated Radar Terminal System (ARTS) Air Traffic Control facilities. The MSAWS provides the air traffic controller automatically with aural warnings when aircraft under his or her control penetrate the safety altitude for the region they are flying in. The MSAWS program started in 1977 but still only 50% of the project has been completed in the United States.

63% of all CFIT accidents occurred during the initial approach or final approach and landing flightphase figure 3. The relatively high (67) average number of casualties per CFIT accident can be explained by the high airspeed averaged in CFIT accidents (220 knots, 407 km/h).

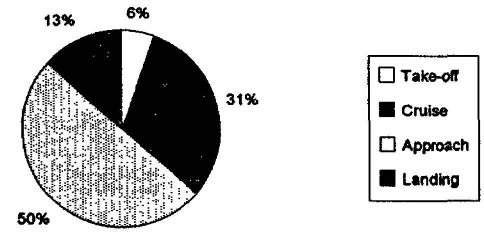


figure 3, CFIT Accidents by Flightphase, Source: Corrie [4]

Human error

"Experience indicates that most CFIT accidents can be related to poor visibility, navigation error, instrument reading error, visual misconception, vertigo, distraction, confusion and/or inattention. Thus, the expression "controlled flight into terrain" is one applied to those accidents that are normally attributed to "pilot error" as opposed to mechanical failure"

The essence of this quotation by Peter Penny, published in the ICAO BULLETIN of March 1975, still holds for present day aviation operations. It is important however to stress that the causes for aircraft accidents cannot be attributed to one action only. Even when a crew-error directly caused an accident, there usually are several other events which preceded it, and also attributed. According to Professor James Reason (Reason, [9]), aircraft accidents occur as a result of complex interactions between many causal factors. The causal factors may be categorised into three groups: active failures committed by those operating at the "sharp end", which are *necessary but insufficient causes for aircraft accidents*, local triggering factors such as weather conditions, and latent organisational failures. The total sequence of events leading to Reason's *organisational accident* is depicted in figure 4.

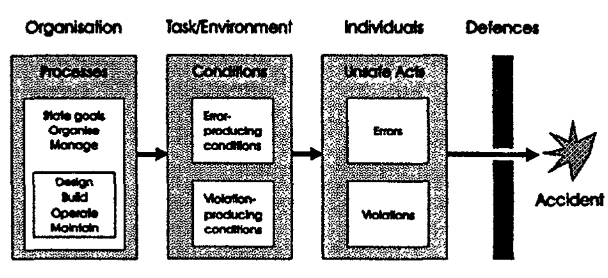


figure 4, The Elements of an Organisational Accident. (Redrawn from Reason [9])

This paper, however, would like to focus on the human-error side of the CFIT problem. In Reason's diagram this would concern the "Individuals" and "Defences" sections.

CFIT causal conditions

The factors that contribute to CFIT accidents may be categorised into two groups, identified by two "CFIT causal conditions". The first condition focuses on the aircraft's flightpath:

> *The aircraft's flight path will cause it to collide with terrain or water.*

The second condition concentrates on the crew's terrain awareness:

> *The crew has no awareness concerning the flightpath condition, or develops this awareness too late to avoid collision.*

Only when *both* of these conditions have been satisfied, a CFIT accident will occur. In order to construct a fail-proof CFIT protection, however, both of these problem areas will have to be addressed. According to this classification, two classes of CFIT prevention strategies may be distinguished: those which prevent aircraft from flying a flightpath towards terrain and those which make it obvious to the crew that they are flying towards terrain. First we have to understand under what circumstances and due to which error-chains any of these conditions may occur. From accident and incident reports, a number of causes for the two conditions mentioned above can be extracted.

3. ASSESSING THE LINKS IN THE CFIT ACCIDENT CHAIN

Flightpath Condition

As most flightpath deviations leading to CFIT accidents have occurred on the extended centreline of the destination runway (Bateman, [3]), flightpath deviations in the vertical plane seem to pose a greater threat than lateral deviations. In the July '88 through July '93 period, 13 out of 25 CFIT commercial jet aircraft hull losses occurred while executing non-precision[1], step-down, approach procedures. When observing the flight path profiles of these accidents it appears that in several cases the crew failed to level off after performing an altitude step (Bateman, [2]). The increased workload induced by performing a large number of step-down altitude changes in a short period of time, as required by some approach profiles, may cause the crew to lose track of their position on the profile. In September 1992 an Airbus A-300 crashed on approach for Kathmandu's Tribhuvan International airport (figure 5). The step-down approach for this airport comprises no less than 8 altitude steps along a 16 nautical miles long approach path.

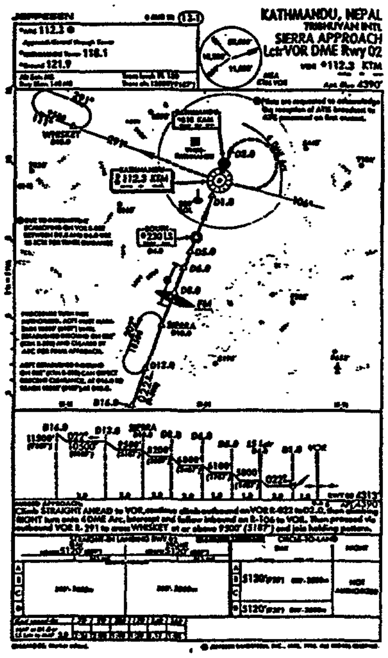


figure 5, Instrument Approach Plate for Kathmandu's Tribhuvan "Sierra" Approach. source:[8].

Undetected descents, resulting from erroneous autoflight mode selection have also been a factor in several incidents and accidents.

Misinterpretation of a departure procedure is illustrated by the 1989 accident with a Boeing 737-200. The aircraft crashed into a mountain in Taiwan after misinterpreting the Standard Instrument Departure (SID) chart. The SID for the flown departure called for a *right* turn after take of, instead the crew flew a *left* turn and, after being discovered by the copilot, corrected to the right too late.

Premature descent clearances or late heading changes, issued by air traffic control, may also result in flights towards terrain, and may remain undetected by the crew. Although monitoring the terrain separation of aircraft is not a primary responsibility of air traffic controllers, accident and incident reports not seldom cite the crew's confidence in terrain-free vectors issued by ATC personnel. It may well be that the high workload during the approach flightphase makes it tempting

[1] i.e. Approaches without using precision approach and landing guidance systems like the Instrument Landing System (ILS). Non-precision approaches do not offer vertical approach path guidance.

for the crew to trust controllers rather than to check any issued ATC vector for terrain before accepting it. High confidence in terrain-free vectors on the part of the crew reduces the redundancy, and thus the safety, of the aviation system.

As the altimeters used in aviation are barometric, common approach procedures require the crew to enter the local barometric pressure for the destination airport before commencing the approach. The approach controller of the destination airport provides the crew with current air-pressure data by radio communication. Several CFIT accidents are known to have occurred due to altimeter setting errors. Pilots have entered erroneous digits causing altitude indications with errors of more than 1000 feet (305 meters)! Misunderstanding of the correct data during the radio contact with ATC is also known to have been a factor in several accidents.

Aviation is highly dependent on human-to-human voice communication. This is also a leading source of error in the system, and one that is difficult to combat. Communication in aviation operations can be divided into intra-cockpit and extra-cockpit or radio communication. Intra-crew communications provide the only way for captain, first officer, and flight engineer (in three-crew cockpits) of working as a team, they include requests by the pilot flying for specific action by the pilot not flying, and acknowledgements to these requests. Typical extra-cockpit communications are requests to and from air traffic control, and confirmation of reception of the message (readback). Radio communication between air traffic control and crew provides clearance, weather, and traffic information that is not available by other means. Two examples of error in communication:

En route from Colombia to Seattle, the crew of a Metro III aircraft received this descent clearance:
"Nectar one six nine three Metro, you are cleared to cross Hobart at 8,000, Seattle at or above 4,000. Maintain 4,000. No delay expected. Contact Seattle Approach Control over Hobart for further clearance, over"
The captain, who was experienced on the route replied:
"Roger, this is uh nine three Metro is cleared to... uh... Hobart... to cross there 4,000 or above, the range station at 4,000, and we report to you at uh Hobart, over"
Control replied: *"Negative. Report Hobart to Seattle Approach Control"*, thus correcting the last and least important of the two mistakes in the repeat-back. The aircraft descended to 4,000 feet and crashed into a mountain.

A Boeing 747 approaching Nairobi in the middle of the night was cleared by the controller to *"seven zero zero zero"* feet. The first officer repeated back *"five zero zero zero"*. The controller should have corrected the mistake, but it was allowed to continue. Fortunately the captain saw the ground through intermittent cloud and carried out an overshoot.

In general, errors made in communication arise from non-standard or ambiguous phraseology, lack of communication between crew members or mishearing words with similar pronunciation.
Lack of information exchange between crew members has also led to several accidents in the history of aviation. Cockpit voice recorders (CVRs), taping the last seconds before the accident often show the pilot's chilling inability to say the words that might save them.

Terrain Awareness Condition
The primary factor leading to the loss of the crew's terrain situational awareness is of course the outside visibility. No flight towards terrain will remain undetected for long during operation in daylight Visual Meteorological Conditions (VMC). Under Instrument Meteorological Conditions (IMC) pilots have to determine their position by relating their current position and flightpath vector with the approach plate's[2] terrain information. This process involves mental conversion of the north-up oriented approach plates towards a track-up "mental terrain picture". Under high workload situations this may involve too much time to be correctly performed, mental terrain pictures constructed prior to the initial approach may fade during the relatively high number of actions needed to complete this flight-phase. The CVRs of crashed aircraft often expose expressions of uncertainty by crew-members concerning the whereabouts of the surrounding terrain.
When flying under Visual Flight Rules (VFR) in deteriorating weather conditions, pilots may have the tendency to try and maintain visual contact with the ground, instead of cancelling VFR and continuing under an Instrument Flight Rules (IFR) flight plan. Several accidents have been attributed to

---

[2] Instrument Approach Plates are paper maps that contain all information for successfully performing an approach to one specific airport runway.

this phenomenon, most of which occurred to aircraft of small regional operators, whose pilots were used to fly under VFR during almost all flights. An example:

In December 1991 a Beechjet Be 400 collided with a mountain summit near Rome, Georgia, US, shortly after take-off. The aircraft was told to by ATC to *"remain VFR because we have traffic four and five right now south-east of Rome"*. The crew had trouble maintaining VFR in the fog, but did not inform the controller. The cockpit voice recorder installed in the aircraft indicated that the pilots recognised that the aircraft was close to obscured terrain. At 9.39:39 (nine minutes past half ten and 39 seconds) the Captain told the First Officer: *"We're gonna have to get away from that mountain down there pretty soon"*, and at 9.39:52: *"You're getting close. You're gonna (have to) go to the right"* The First Officer answered that he could not *"see over there"*. The captain then stated that if they maintained their present course, they could run into an airplane on approach to Rome and pointed out there was a mountain in one direction and an antenna in another that would be hidden by fog. At 9.40:07 the captain directed the first officer to fly *"back to the right"* and the first officer stated *"I can't see over there that's why I wanted to go the other way"*. The CVR recording stopped at 9.40:55. The accident report stated that the probable cause of the accident was:

"the captain's decision to initiate visual flight into an area of known mountainous terrain and low ceilings and the failure of the flight crew to maintain awareness of their proximity to the terrain".

4. GROUND PROXIMITY WARNING SYSTEMS

To alert the crew whenever their loss of terrain-awareness has developed into a hazardous situation, the ground proximity warning system (GPWS) has been developed. Development of GPWS started in 1967, when it was recognised that the radio altimeter[3], a requirement for the Category II "All Weather Landing" instrument package, did not encompass the maximum alerting for exposure to collision with the ground under all type of flying conditions. The GPWS concept was unique in a number of respects. It was the first warning system in general use to combine information of a number of hitherto unrelated aircraft sensors to produce a *single warning output. It was also the first cockpit* warning system to use a synthesised human voice to provide the primary warning to the crew. The ground proximity warning system computer unit accepts inputs of radio altitude, barometric altitude rate, ILS glide slope deviation and landing gear and flap discretes, which it manipulates mathematically to determine the onset of terrain proximity. Audible warnings or alerts are operator specified synthesised voice commands such as "WHOOP WHOOP, PULL UP!" or "TERRAIN! TERRAIN!". Red "PULL UP", "GND PROX" and/or "BELOW G/S" lights, flashing in the glareshield, form the visual output of the GPWS. The GPWS functions only when the radio altitude is less than 2,500 feet above ground level. Power to the system is controlled only by a circuit breaker, pilot inputs are not required and, when the aircraft is flown in normal profiles, the GPWS warning should never be heard.

Modern ground proximity warning systems offer four warnings for a (predicted) dangerous situation with respect to the terrain: modes 1 (excessive rate of descent), mode 2 (excessive terrain closure rate), mode 3 (altitude loss following take off), mode 4 (insufficient terrain clearance). Another two warning modes include a warning for excessive deviation below the glideslope of an ILS approach (mode 5) and for a descent below a minimum radio altitude selected by the crew (mode 6). Although its introduction has significantly reduced the CFIT rate, incident and accident reports, indicate that GPWS is not capable of providing a fail-proof safety-net against CFIT. More than half of the aircraft lost in CFIT accidents during the July '88-July '93 period had been equipped with GPWS (figure 6).

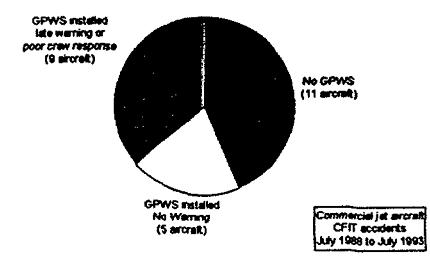


figure 6, GPWS effectivity, Large commercial Jet Aircraft, 25 CFIT accidents, July 1988-July1993 (redrawn from Bateman [3])

The ground proximity warning system has been designed as a CFIT safety net and therefore should comply with two characteristic requirements. First of all the system should not intervene until normal

[3] The radio altimeter uses continuous or pulsed radar signals to measure the distance to terrain directly below the aircraft.

procedures and crew vigilance have failed to assure a safe terrain clearance. Secondly, when safety is likely to be endangered, the system should be capable of alerting the crew under all circumstances. Present GPWS models are known to be subject to both primary and secondary error: there are circumstances under which the system issues unnecessary warnings, as well as situations in which the system does not alert the crew for a potentially dangerous situation with respect to terrain. These two areas do not include all deficiencies of current GPWS equipment however. Accidents have occurred in which the GPWS did issue terrain warnings, but too late for the crew to complete a successful recovery. Simulated GPWS warning time prior to projected impact has been depicted in figure 7.

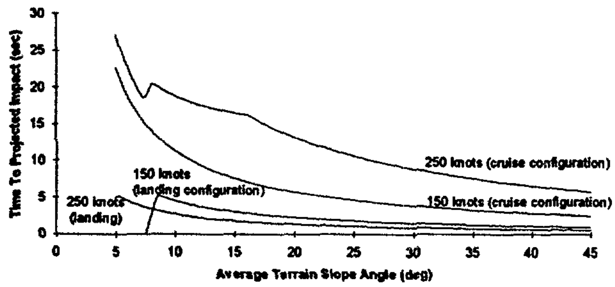


figure 7, GPWS warning time prior to projected impact. Data from computer simulation with "GPWS SIM" (Dijkgraaf, [6]).

For the flightpath of an aircraft flying horizontally towards a single mountain with a constant slope angle the GPWS warning time was measured. When the aircraft was configured for landing (landing gear down and locked, landing flaps selected) warning times did not exceed 5 seconds in advance of a possible impact!

In other cases a timely GPWS warning was issued, but the crew responded too late (or even not at all). The most severe limitation of the GPWS however, is its inability to guarantee a safe recovery from any warning. In other words: even when a procedural evasive manoeuvre will be initiated immediately after the first warning, GPWS does not guarantee sufficient terrain clearance throughout the recovery flight path. Due to the lack of a "forward looking" sensor input, present ground proximity warning systems have to determine terrain hazard by extrapolating the slope angle of the terrain directly below the aircraft. Warnings will then be issued some time prior to a projected collision with an extrapolated mountain slope. It is this technique, in use since the introduction of GPWS in the late 1960s, that lies at the basis of most of the GPWS drawbacks. Whenever terrain steepness increases along the trajectory the aircraft is flying, the system will alert the crew relatively late, on the other hand when the aircraft flies at a safe altitude over sheer cliffs, the system will issue unwanted alerts.

5. MANAGING THE ERRORS LEADING TO CFIT

Having summarised the numerous human-error related causal factors which form the CFIT chain-of-events, and having observed that the ground proximity warning system is not a fail-proof CFIT safety-net, what can be done do to reduce the CFIT-rate in future?

In his report "Intervention Strategies for the Management of Human Error" (Wiener [11]), Earl Wiener of the University of Miami hands several lines of defence against human error. In his terms, "error management" must be distinguished from "error reduction" or "elimination". "Management" in this sense means that one strives to build into systems and operator methods by which one can either eliminate or reduce human error, or if this is not possible, to minimise its consequences. According to Wiener:

"The Human remains a vital component in complex systems found in aviation and elsewhere because he/she possesses remarkable perceptual capabilities, among them the ability to detect subtle deviations from normal. This capability should be assigned to the front end of the lines of defence against human error. Human error is the price we pay for the flexibility of the human brain. It is a price that must be minimised by effective intervention strategies and lines of defence."

Concluding his report Wiener presents 5 levels at which technology and humans may combine to manage rather than necessarily prevent error. These lines global of defence are:

1. Prevent the error in the first place, or make it as unlikely as possible. This is done by training, procedures, management, and quality assurance.
2. If an error is introduced into the system, make it as conspicuous as possible through display design and traditional human factors ("error-evident displays").
3. If the first two methods fail to block or remove the error, design the system, probably through software, to trap the error and prevent it from affecting the system. This level of defence may or may not require further developments of artificial intelligence.
4. Provide sophisticated warning and alerting systems.
5. Make certain that there is a recovery path from any error.

Returning to the controlled flight into terrain problem, Wiener's global lines of defence (see also figure 8) indicate several shortcomings with respect to CFIT avoidance in the present aviation system.

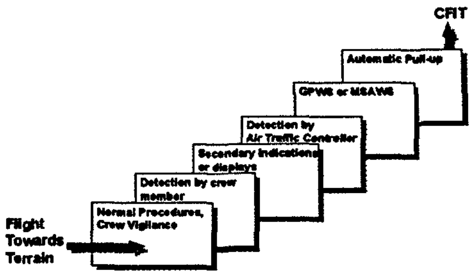


figure 8, Lines of defence against CFIT.

The first defence line (training, management etc.) is currently being promoted by the ICAO and the Flight Safety Foundation's CFIT Task Force in the form of CFIT awareness programmes and GPWS training aids. It is of extreme importance that pilots learn to understand CFIT "traps" and learn to perform a successful recovery from any GPWS warning. "Error-evident" displays, the second line of defence Wiener proposed, may be introduced in the form of some sort of terrain display that could improve the crew's terrain awareness. Such a display could increase the probability of detection by the crew of a inadvertent flight towards terrain. It may also take over the crew's task of mental conversion of the north-up oriented paper approach plates towards a track-up oriented picture. The error-evident display does not have the intelligence to detect errors.

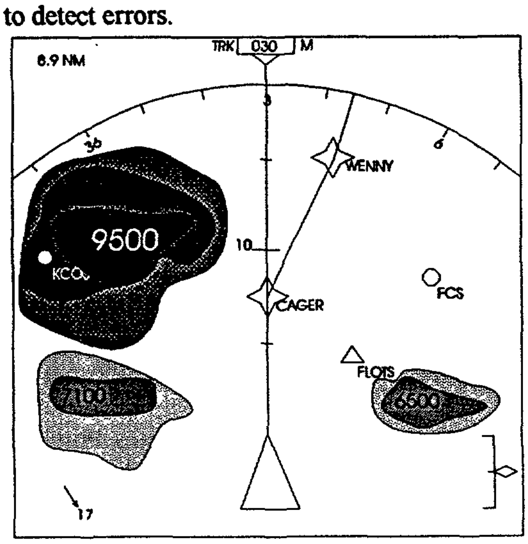


figure 9, Example of terrain elevation information on navigation display. In that cases it is not likely these pilots would ignore the terrain display.

It is merely a display system, and the management of human error ultimately depends on human intelligence to detect the error. As stated earlier in this paper, in many CFIT accidents from the past, pilots expressed their doubts concerning the position of high terrain relative to the aircraft. An example of terrain display depicted on the aircraft's navigation display is presented in figure 9. The terrain data required for terrain displays could be "synthetic", in the form of a terrain elevation database stored in a computer's memory, or obtained via various "enhanced-vision" sensors that depict *real* terrain. Using a terrain elevation database would require a reliable and accurate navigation system to align the synthetic terrain and the real world, but might well be more cost-effective than installing expensive infra red sensors, millimeter-wave radar and low-light-level television equipment. Enhanced vision equipment is not dependent of navigation accuracy and does not hold the dangers hidden in database-errors, but is hindered by weather-conditions. Combining the two systems, enhancing *real* sensor data with synthetic terrain information could be a solution but will get even more expensive.

Future ground proximity warning systems should assure the success of procedural escape manoeuvres, eliminate unwanted warnings and no-warning situations and improve crew response to warnings. In order to achieve these requirements, the use of a coarse digital terrain elevation database is expected to offer the best cost/performance ratio. By using terrain elevation data, aircraft performance and pilot response time as GPWS input, it should be possible to continuously compute escape flightpaths, both in the vertical and horizontal plane. Postponing terrain warnings until the last of these flightpaths intersects a safety margin above terrain, will considerately reduce the number of unwanted warnings while assuring a recovery from every warning. By using approach recognition logic, the warning system should be able to distinguish between stabilised approaches towards an airport and stabilised approaches towards terrain. Besides these system improvements, it will still be necessary to assure timely and efficient crew response to terrain warnings. To achieve this, the terrain preview capability as used for the ground proximity warning computer should also be offered to the crew by means of a terrain display. It is assumed that the preview capability offered by this display will cause the crew to discover developing terrain hazards long before the GPWS warning is issued and, in case the GPWS still catches the crew by surprise, reduces pilot response time by indicating the position of the terrain causing the alert. MSAWS software for air traffic control or enhanced vision equipment could in that case be used as an independent monitor.

6. CONCLUSIONS

Rather than to try and eliminate human error as a controlled flight into terrain cause, efforts should be directed towards human error management, thus avoiding the error where possible while reducing the severity of the consequences of any unavoidable errors. Well-designed procedures, man-machine-interfaces, training programmes and CFIT awareness programmes should avoid human error where possible. In addition, pilots should be provided real-time terrain information to make flightpath deviations towards terrain obvious. Acting as a final line of defence, improved terrain-warning systems should incorporate "forward looking" capability and take into account the pilot's inherent response delay to unexpected warnings, thus offering a fail-proof controlled flight into terrain safety net.

7. LIST OF ABBREVIATIONS

ATC	Air Traffic Control
CFIT	Controlled Flight Into Terrain
GPWS	Ground Proximity Warning System
ICAO	International Civil Aviation Organisation
IFR	Instrument Flight Rules
ILS	Instrument Landing System
IMC	Instrument Meteorological Conditions
MSAWS	Minimum Safe Altitude Warning System
VFR	Visual Flight Rules
VMC	Visual Meteorological Conditions

8. LITERATURE

1. Bateman, C.D., *Past, Present and Future Efforts to Reduce Controlled Flight into Terrain (CFIT) Accidents,* Proceedings of Flight Safety Foundation's 43rd International Air Safety Seminar, Rome, 1990.
2. Bateman, C.D., *Flight Into Terrain and the Ground Proximity Warning System*, Engineering report 070-4251, Sundstrand Data Control, Inc., February 1991
3. Bateman, C.D., *International CFIT Task Force Data and Dissemination Group,* paper presented at the 46th International FSF Air Safety Seminar, Kuala Lumpur, Malaysia, 1993.
4. Corrie, S.J., *Programme to address CFIT problems includes plan to improve GPWS provisions,* in: ICAO Journal, November 1993.
5. Corrie, S.J., *Potential growth in air travel demands renewed effort to improve safety record,* in: ICAO Journal, December 1994.
6. Dijkgraaf, F.C., *Controlled Flight Into Terrain, causes, solutions, future research,* Graduation Report, Delft University of Technology, Faculty of Aerospace Engineering, Delft, The Netherlands, 1994.
7. Hughes, D., *Safety Group Highlights CFIT Risk For Regionals,* in: Aviation Week & Space Technology, May 9, 1994.
8. McCarthy, P., *A Tale of Two Cities,* in: IFALPA Quarterly Review, December 1993.
9. Reason, J., *Identifying the Latent Causes of Aircraft Accidents Before and After the Event,* in: Proceedings of the International Society of Air Safety Investigators 22nd Annual Seminar, Canberra, Australia, November 4-7, 1991.
10. Weener, E.F., *Presentation for IATA Safety Subcommittee (SAFAC)* Montreal Canada, August 1993.
11. Wiener, E.F., *Intervention Methods for the Management of Human Error,* NASA Contractor Report 4547, NASA Ames Research Center, August 1993.

About Faults, Errors, and other Dangerous Things

Matthias Rauterberg

Work and Organizational Psychology Unit
Swiss Federal Institute of Technology (ETH)
Nelkenstrasse 11, CH-8092 ZUERICH
+41-1-63-27082, rauterberg@ifap.bepr.ethz.ch

Abstract

In this paper the traditional paradigm for learning and training of operators in complex systems is discussed and criticised. There is a strong influence (the doctrine of 'mental logic') coming from research carried out in artificial intelligence (AI). The most well known arguments against the AI-approach are presented and discussed in relation to expertise, intuition and implicit knowledge. The importance of faults and errors are discussed in the context of a new metaphor for cognitive structures to describe expertise.

Keywords: fault, error, learning, training, cognitive structure, expertise, intuition

1 Introduction

"I learned more from my defeats than from my victories" (Napoleon, ca. 1800)

Why is this statement sometimes (or always) true? To answer this question we need a new understanding of human errors, inefficient behaviour, and expertise. In this paper we will discuss the importance of learning from unsuccessful behaviour. What percentage of unanticipated events (e.g., accidents) is caused by human error? This is a question that vexed researchers for years in the context of human interaction with complex systems.

The classical understanding of human errors is characterized by a *negative* valuation of erroneous behaviour, something that must be avoided. The Western Culture is constrained by this *taboo*: Not to talk about faults, errors and other dangerous behaviour! This taboo keeps us to present our self as successful as possible. We are–normally–not allowed to discuss in public how and what we could learn from our faults and errors.

Rasmussen defines human errors as follows [20]: "if a system performs less satisfactorily than it normally does–because of a human act–the cause will very likely be identified as a human error". Human errors are the most important cause of accidents or near-miss accidents. Heinrich [9] analysed insurance company records and got the result that approximately 85 percent of accidents are due to human error. Nagel [15] presents the results of his analysis: approximately 70 percent of accidents in aviation operations are classified as human errors. Accidents are categorised as caused by either unsafe acts of persons (e.g., operator error) or by unsafe conditions (cf. [9] and [23]). One consequence of using this dichotomy is often to blame the individual who was injured or who was in charge of the machine that was involved in the accident.

In fact, it is probably meaningless even to ask what proportions of accidents were due to human error. The more important question is what can one learn from his or her errors, and how are these insights and the derived knowledge embedded in the individual cognitive structure.

2 The Artificial Intelligence (AI)-Approaches

The AI-approaches can be distinguished in two different research tracks: (1) the traditional rule-based approach, and (2) the connectionistic approach. In the following we concentrate us on the rule-based approach.

One of the most elaborated modelling approach is Soar [13] [16]. Newell [17] describes Soar as follows: "Soar is ... a *symbolic computational system*. ... Soar is organised around *problem spaces,* that is, tasks are formulated as search in a space of states by means of operators that produce new states, where operators may be applied repeatedly, to find a desired state that signifies the accomplishment of the task. ... Soar is organised entirely as a *production system,* that is, its long-term memory for both program and data consists of parallel-acting condition-action rules. ... Soar incorporates a *goal hierarchy.* ... Soar learns continuously from its experience by *chunking,* which constructs new productions (chunks) to capture the new knowledge that Soar developed (in working memory) to resolve its difficulties" ([17] pp. 30-32).

Soar is based on impasse-driven learning. "While Soar is performing a task by using the behaviour model in working memory, it is also learning. It is building chunks every time it impasses from one problem space to another, ... These chunks constitute the acquisition of knowledge for doing the task" ([17] pp. 62-62).

The knowledge generated by chunking and stored in the long-term memory represents only successful trials (i.e., solving an impasse). Knowledge of unsuccessful attempts (i.e., not solving an impasse) is not in memory. Learning in Soar means that long-term memory contains evidence only of the sequence of *effective actions.*

But, how would it be if the majority of the knowledge of the long-term memory of humans consists only of *unsuccessful trials?* Soar seems to be a typical representative of a theory driven approach for *error-free skilled behaviour* (cf. for other modelling approaches [4] pp. 80ff and [19] pp. 14ff). Why do we believe that an empirical driven approach – looking to the concrete task solving behaviour of people – is better than a theory driven approach? The answer refers to the following assumption.

2.1 Implicit assumption

Most of the known modelling approaches is based on the implicit assumption that the "mental model maps completely to the relevant part of the conceptual model, e.g. the user virtual machine. Unexpected effects and errors point to inconsistency between the mental model and the conceptual model" ([27] p. 258). This one-to-one mapping between the mental model and the conceptual model of the interactive system implies a *positive* correlation between the complexity of the observable behaviour and the complexity of the assumed mental model. But this assumption seems to be – in this generality – wrong.

Based on the empirical result in [21] (see chapter 3.2), that the complexity of the observable behaviour of novices is larger than the complexity of experts' behaviour, we must conclude that the behavioural complexity is *negatively* correlated with the complexity of the mental model. If the cognitive structure is too simple, then the concrete task solving process must be filled up with many heuristics or trial and error behaviour [21].

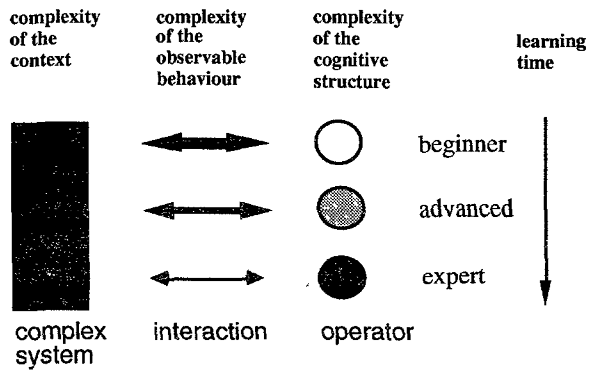


Fig. 1. The relationship between the complexity of the human interaction and of the cognitive structure.

Learning how to solve a specific task with a given system means that the behavioural complexity decreases and the cognitive complexity increases (cf. Fig. 1). Now, one of the central question is: What kind of knowledge is stored in the cognitive structure? Before we are able to give a preliminary answer to this question, we have to discuss the consequences of the traditional AI-paradigm.

2.2 Undesirable consequences

The famous classification of Rasmussen [20] in skill-based, rule-based, and knowledge-based behavior is one consequence of taking the AI-approach seriously. Human behaviour can be therefor classified (1) as error-free skilled behaviour (cf. [13]), (2) as inefficient behaviour (cf. [29]), and (3) as erroneous behaviour (cf. [28]).

It is very important to notice that in a rule-based expert system all unsuccessful trials during a task solving process (e.g., heuristic search) are thrown away after finding the correct solution path. A traditional expert system has no power of recollection of all these unsuccessful trials. Only correct solution paths are stored and could be retrieved later. This is an undesirable consequence of the classical AI-approach.

Newell [16] describes the chunking in Soar as to be not sensitive to success and failure, only to whatever resolves an impasse. "An impasse can be resolved as much by failure–the decision to leave a space as unprofitable–as by success" ([16], p. 313). It would be interesting to see whether this type of failure is the only one in Soar. Although Rumelhart and Norman [24] claimed to investigate and to model specific errors of highly skilled typists, they exclude explicitly all "mechanism involved in learning".

Soar shows the direction of an alternative modelling approach for human learning processes. Each human has at least one recollection of an unsuccessful problem solving strategy in his or her individual learning history. This argument has a strong empirical evidence (cf. also [21]).

A second consequence of the traditional AI-approach is the fact that all inferences of a heuristic problem or task solving process must have a 'mental logic'. The most glaring problem is that people make mistakes. They draw invalid conclusions, which should not occur if deduction is guided by a 'mental logic'.

2.3 Critical statements

The doctrine of 'mental logic' can certainly be formulated in a way that meets the methodological criterion of *effectiveness*. The trouble with mental logic is thus empirical. Johnson-Laird [12] describes six main problems: (1) People make fallacious inferences. (2) There is no definitive answer to the question: Which logic, or logic's, are to be found in the mind? (3) How is logic formulated in the mind? (4) How does a system of logic arise in the mind? (5) What evidence there is about the psychology of reasoning suggests that deductions are not immune to the content of the premises. (6) People follow extra-logical heuristics when they make spontaneous inferences. Why does cognitive psychology constrain the modern research to the doctrine of 'mental logic'? To come up with an answer, we have to look on the discussion and review coming from the Dreyfuses.

To Dreyfus [6], the world of the subjective is more important than that of the objective; reality is defined from within–in terms of the individual and his power to perceive and act, to know truths that are unutterable. Dreyfus concludes that some of the things' people do are intrinsically human and cannot be mechanised. To Dreyfus they are *intuition*, *insight*, and *comprehension*–the ability to immediately grasp complex situations, resolving ambiguities, weeding the relevant from the irrelevant (cf. the "exformation" described by Nørretranders [18]).

According to Dreyfus, the conviction that we can formalise reality, explaining everything with rules, began–as far back as the days of ancient Greece–and has become so dominant in the twentieth century that few people question it. This is one explanation for the doctrine of 'mental logic'.

Mary Henle [10] declares: "I have never found errors which could unambiguously be attributed to faulty reasoning." She suggests that mistakes arise because people misunderstand or forget premises, and because they import additional and unwarranted factual assumptions into their reasoning.

The Dreyfuses [5] argument that only novices use facts and rules. But as we become expert, we forget the rules and act intuitively, automatically adjusting our behaviour to the perceived constraints. Most scientists assume that these kinds of abilities are based on the unconscious and simultaneous processing of signals coming from the eyes, the ears, and the hands. But the Dreyfuses [5] believe that intuition defies rational powers of description, that it can't be computerised. Like judgement and wisdom it is one of the atomic element of our world (i.e. *irreducible*).

We share the critique of the Dreyfuses, but we do not follow their conclusions. To overcome the deadlock and 'mystical' situation following the Dreyfuses we need a new understanding of knowledge that gives an expert the ability to act intuitively.

3 Empirical Studies of 'Erroneous' Behaviour

"There are no perfect humans, there are only perfect intentions" (Buddha)

Our basic assumption is that human behaviour cannot be erroneous. Of course, human decisions and the behavioural consequences of these decisions can be classified as erroneous and faulty, but from a pure introspective standpoint – from the internal *psycho*-logic of the subject – each decision is the best solution fulfilling all actual constrains and restrictions: lack of information and/or motivation, lack of knowledge and/or qualification, over or under estimation of the task and/or context complexity etc. In this sense we share the position of the Dreyfuses.

3.1 The 'law of requisite variety'

Humans need variety to behave and to adapt. A total static environment is insufferable. Ashby ([2] p. 90) summarises his analysis of regulation and adaptation of biological systems as follows: "The concept of regulation is applicable when there is a set D of disturbances, to which the organism has a set R of responses, of which on any occasion it produces some one, r_j say. The physico-chemical or other nature of the whole system then determines the outcome. This will have some value for the organism, either Good or Bad say. If the organism is well adapted, or has the know-how, its response r_j as a variable, will be such a function of the disturbance d_i that the outcome will always lie in the subset marked Good. The law of requisite variety then says that such regulation cannot be achieved unless the regulator R, as a channel of communication, has more than a certain capacity. Thus, if D threatens to introduce a variety of 10 bits into the outcomes, and if survival demands that the outcomes be restricted to 2 bits, then at each action R must provide variety of at least 8 bits."

If we try to translate this 'law of requisite variety' to normal human behaviour then we can describe it as follows: All human behaviour is characterized by a specific extent of variety (see Fig. 2). If the system – in which the human has to behave – constrains this normal variety then we can observe 'errors'. In this sense an error is the necessary violation of system's restrictions.

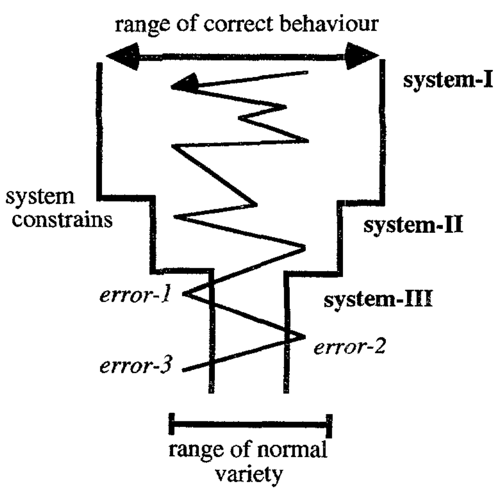


Fig. 2. The correlation between 'erroneous' behaviour and constraining behavioural variety.

If a system constrains human behaviour to only one possible 'correct solution path' then we can observe a maximum of violations, say errors. Husseiny and Sabri [11] counted 644 "critical incidents" in a representative study analysing complex systems (this is equivalent to an error rate of 16%); they noted that in "non nuclear complex systems" the rate of slips lies only between 3% and 7%. Most complex systems are designed to constrain the operator's behaviour to a minimum of variety. Ulich [25] arguments against this 'one best way' doctrine of system design because users differ inter- and intra-individually. A system must have a minimum of flexibility to give all users the opportunity to behave in an error-free way.

To investigate the relationship between behavioural and cognitive complexity we try to observe individual behaviour in its 'natural sense'. All deviations of the correct solution path are only interpreted as exploratory behaviour caused by the need for variety.

3.2 About the relationship between behavioural and cognitive complexity

In one of our experiments we compared the task solving behaviour of novices (subjects without experiences of electronic data processing EDP) with the behaviour of experts (subjects with a lot of EDP experience). We could show, that the complexity of the observable task solving process (the 'behavioural complexity') of novices is significantly larger than the complexity of the observable behaviour of experts (see Fig. 3). All twelve novices and all twelve experts solved exactly the same four different tasks (see [21]).

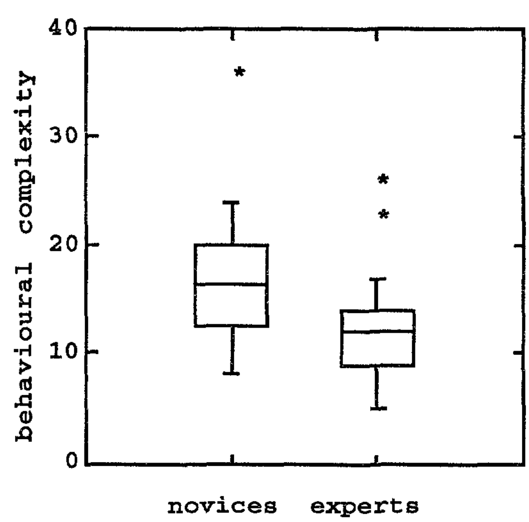


Fig. 3. The Box-and-Whisker Plot shows the 'behavioural complexity' for 'novices-experts' (N=24).

This important result seems to be – at the first glance – trivial. But what does it really mean? If we assume – and this is highly plausible – that the complexity of the mental model of the experts is significantly larger than the cognitive complexity of the novices, then we must conclude that the correlation between behavioural and cognitive complexity is negative! And, this result is not trivial. Note in Fig. 3, that the minimal task complexity of all four tasks was only reached by one expert (task one: behavioural complexity = 6, the 'one best way'). We do not argument that this minimum cannot be reached, but to constrain human behaviour only to this minimum leads directly to the paradox of 'errors' of high skilled and over-trained experts.

3.3 What do we learn from errors?

Arnold and Roe assume ([1] p. 205), "that errors may have great functionality for the user, especially during learning. When the user is able to find out what has caused the error and how to correct it, errors may be highly informative. This implies that one should not try to prevent all errors." This hypothesis was tested later in an empirical investigation by Frese et al [7].

Frese et al [7] describe the following four reasons for the positive role of errors in training: (1) "the mental model of a system is enhanced when a person makes an error ... (2) mental models are better when they also encompass potential pitfalls and error prone problem areas ... (3) when error-free training is aspired, the trainer will restrict the kind of strategies used by the trainees, because unrestricted strategies increase the chance to error ... (4) errors not only appear in training but also in the actual work situation." They compared two groups: one group with an error training (N=15), and a second group with an error-avoidant training (N=8). In a speed test the error-training subjects produced significant fewer errors than the error-avoidant group.

Gürtler ([8] p. 95) got the same results in the context of sport: "there, where more accidents were counted in the training phase, appeared less – above all of less grave consequences – accidents during the match. Few accidents during the training correlate with accidents of grave consequences during the match."

Van der Schaaf ([26] p. 85) concludes, that "every time an operator, manager, procedure, or piece of equipment 'behaves' in an unexpected way and thereby prevents a likely breakdown of the production system ... or restores the required levels of safety and reliability, these *positive deviations* could be detected, reported and analysed in order to improve the qualitative insight into system functioning on the whole." This conclusion is not only valid for the global 'accident driven' design process "on the whole", this statement is also valid on the individual level of operating a complex system.

4 An Alternative View on the Knowledge of Mental Models

First, let us shortly summarise the traditional approach for learning based on training. To avoid unnecessary knowledge about unsafe acts beyond stable system's reaction operators are only trained on key emergency procedures. The beneficial effects of *extensive* training of these key emergency procedures are that they become the dominant and easily retrieved habits from long-term memory when stress imposes that bias. Sometimes emergency procedures are inconsistent with normal operations. To minimise the uncertainty coming from these inconsistencies Wickens demands the following design: "Clearly, where possible, systems should be designed so that procedures followed under emergencies are as consistent as possible with those followed under normal operations" ([28] p. 422).

We try to argument against this position. But, what is wrong with this traditional position? Nothing, of course not! Except the assumption that "knowledge about 'unsafe acts beyond stable system reactions' is *unnecessary* or *dangerous*". If our experimental results (the *negative* correlation between behavioural and cognitive complexity, see [21]) are correct (and there is no evidence that they are not correct), then we must conclude that the cognitive structure of experts contains knowledge about unsuccessful trials (see [22]). What does this result mean for the cognitive structure of mental models about complex systems? Our conclusion is that humans need for effective and correct behaviour in critical situations a huge amount of knowledge 'about unsafe acts beyond stable system reactions'.

4.1 A metaphor for traditional learning and knowledge acquisition

To describe the traditional training procedure and the intended effects on the cognitive structure of operators, we introduce the following metaphor: The cognitive structure is a 'landscape' (cf. 'Tabula rasa'). This landscape–without knowledge–has a flat structure. Learning and training of correct behaviour means 'to run a ditch' (cf. Fig. 4). The 'flow' of the actual behaviour can be described with 'a rolling ball' on this landscape. The 'course of the ball' describes exactly the observable behaviour.

Intensive training results in 'deepening the ditch' (cf. Fig. 5) to make sure that the operator behaves correctly especially in critical situations. But what happens if there is no ditch for the actual and–in the worst case–dangerous situation?

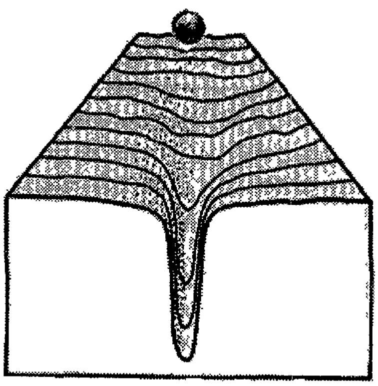


Fig. 4. The 'landscape'-metaphor for the cognitive structure with a 'rolling ball' to symbolise the 'flow' of actions at the *beginning* of a training.

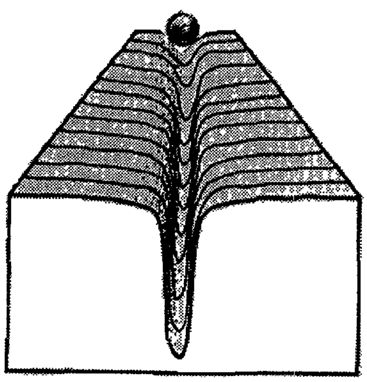


Fig. 5. The 'landscape'-metaphor for the cognitive structure with a 'rolling ball' to symbolise the 'flow' of actions at the *end* of a training.

Bainbridge [3] describes very clearly the problems arising when an operator has to take over a complex process during a monitoring task. To take-over process control is especially problematic when the system runs into an unknown state.

In this situation there is no 'ditch' to guide the 'ball'. Training in a simulator is one possible consequence, better is permanent on-line control in the real process. Operators should have on-line control over the real process (cf. [3]). High skilled operators tend to lose the potential to be aware of the whole process. They need a special qualification to get *open minded* (to increase their perceptual range; cf. [14]).

4.2 An alternative metaphor for learning and knowledge acquisition

What would happen, if operators are trained in simulators to get experience 'about unsafe acts beyond stable system reactions'? Following our metaphor the cognitive structure could be described as filled up with knowledge about unsuccessful behaviour: the 'hills' and 'mountains' (see Fig. 6).

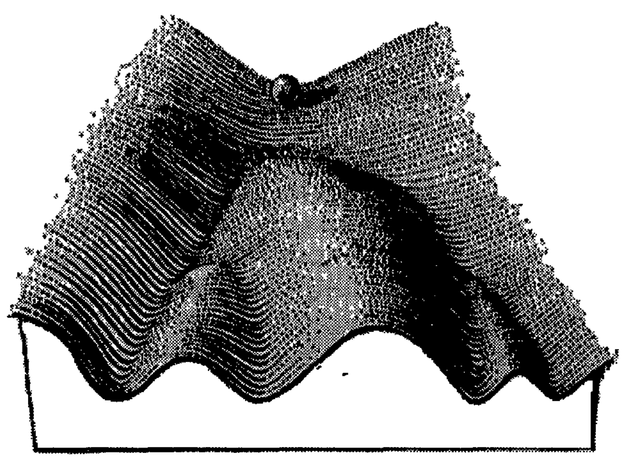


Fig. 6. The 'landscape'-metaphor for the cognitive structure with a 'rolling ball' to symbolise the 'flow' of actions and 'hills' as knowledge about unsuccessful behaviour.

The decisions for the actual behaviour are carried out to avoid *known* errors and faults! The 'ball' is guide by the 'valleys' and 'dales'. To minimise the effort is the basic principle for each actual decision. The implicit knowledge about effort comes from previous faults, errors and other dangerous behaviour. This is a possible explanation of *real* expertise.

Of course, this alternative metaphor for learning and knowledge acquisition can not replace the traditional view. This alternative is a completion, and probably a valid and very helpful one.

References

[1] Arnold, B. & Roe, R. (1987) User errors in Human-Computer Interaction. In: M. Frese, E. Ulich & W. Dzida (Eds.) Human Computer Interaction in the Work Place. Amsterdam: Elsevier, pp. 203-220.

[2] Ashby, W. R. (1958) Requisite variety and its implications for the control of complex systems. *Cybernetica* 1(2):83-99.

[3] Bainbridge, L. (1982) Ironies of Automation. In Analysis, Design, and Evaluation of Man-Machine Systems (G. Johannsen & J.E. Rijnsdorp, Eds.) Düsseldorf: VDI/ VDE, pp. 151-157.

[4] Booth, P. A. (1991) Errors and theory in human–computer interaction. *Acta Psychologica* 78: 69-96.

[5] Dreyfus, H. & Dreyfus, S.(1979) Mind over machine–the power of human intuition and expertise in the era of the computer. New York: The Free Press.

[6] Dreyfus, H. (1979) What computers can't do–the limits of artificial intelligence. New York: The Free Press.

[7] Frese, M., Brodbeck, F., Heinbokel, T., Mooser, C., Schleiffenbaum, E. & Thiemann, P. (1991) Errors in training computer skills: on the positive function of errors. *Human-Computer Interaction* 6:77-93.

[8] Gürtler, H. (1988) Unfallschwerpunktanalyse des Sportspiels. In: E. Rümmele (Ed.) Sicherheit im Sport – eine Herausforderung für die Sportwissenschaft. Köln: Strauss, pp. 91-100.

[9] Heinrich, H. (1959) Industrial accident prevention (4th edition). New York: McGraw-Hill.

[10] Henle, M. (1978) Foreword to R. Revlin & R. E. Mayer (eds.) Human Reasoning. Washington: Winston.

[11] Husseiny, A. & Sabri, Z. (1980) Analysis of human factor in operation of nuclear plants. *Atomenergie, Kerntechnik* Vol. 2.

[12] Johnson-Laird, P. (1983) Mental models. Cambridge (UK): Cambridge University Press.

[13] Laird, J., Newell, A. & Rosenbloom, P. (1987) SOAR: An architecture for general intelligence. *Artificial Intelligence* 33:1-64.

[14] Moray, N. & Rotenberg, I. (1989) Fault management in process control: eye movements and action. *Ergonomics* 32:1319-1342.

[15] Nagel, D. (1988) Human error in aviation operations. In: E. Wiener and D. Nagel (Eds.) Human Factors in Aviation. San Diego: Academic Press, pp. 263-303.

[16] Newell, A. (1990) Unified Theories of Cognition. Cambridge: Harvard University Press.

[17] Newell, A. (1992) Unified theories of cognition and the role of SOAR. In: J. Michon and A. Akyürek (Eds.) SOAR: A Cognitive Architecture in Perspective. New York: Kluwer, pp. 25-79.

[18] Nørretranders, T. (1991) *Mærk verden*. Kopenhagen: Gyldendal.

[19] Oostendorp, H. van & Walbeehm, B. J. (1995) Towards modelling exploratory learning in the context of direct manipulation interfaces. *Interacting with Computers* 7(1): 3-24.

[20] Rasmussen, J. (1986) Information Processing and Human-Machine Interaction. (System Science and Engineering Vol 12, A. Sage, ed.) New York: North-Holland.

[21] Rauterberg, M. (1993) AMME: an automatic mental model evaluation to analyze user behaviour traced in a finite, discrete state space. *Ergonomics* 36: 1369-1380.

[22] Rauterberg, M. (1995, in press) From Novice to Expert Decision Behaviour: a Qualitative Modelling Approach with Petri Nets. In: Proceedings of 6th International Conference on 'Human-Computer Interaction', Tokyo/Yokohama, July 9-14, 1995.

[23] Reason, J. (1990) Human Error. New York: Cambridge University Press.

[24] Rumelhart, D. & Norman, D. (1982) Simulating a skilled typist: a study of skilled cognitive-motor performance. *Cognitive Science* 6:1-36.

[25] Ulich, E. (1994, 3rd edition) Arbeitspsychologie. Stuttgart: Poeschel.

[26] Van der Schaaf, T. (1992) Near miss reporting in the chemical process industry. Proefschrift, TU Eindhoven.

[27] Van der Veer, G., Guest, S., Haselager, P., Innocent, P., McDaid, E., Oesterreicher, L., Tauber, M., Vos, U. & Waern, Y. (1990) Desiging for the mental model: an interdisziplinary approach to the definition of a user interface for electronic mail systems. In: D. Ackermann and M. Tauber (Eds.) Mental Models and Human-Computer Interaction 1. Amsterdam: North-Holland, pp. 253-288.

[28] Wickens, C. (1992) Engineering Psychology and Human Performance (2nd edition). New York: HarperCollins.

[29] Zapf, D., Brodbeck, F., Frese, M., Peters, H. & Prümpers, J. (1992) Errors in working with office computers: a first validation of a taxonomy for observed errors in a field setting. *International Journal of Human-Computer Interaction* 4:311-339.

Session 4

Mental Load and Trust

Chairman: N. Moray, France

Development of a navigator model by the use of mental workload measures

Fulko van Westrenen

Safety Science Group
Department of Technology and Society
Delft University of Technology

Abstract

This paper describes a navigator model clearing the approach into a harbour. The model is a three stage decision model, containing tracking, short-term planning and long-term planning behaviour. This model is verified by measuring the mental workload of pilots whose task is to bring a ship to berth. The mental workload is obtained by ECG recording, from which heart rate and heart rate variability (the 0.1 Hz power density) were calculated. It is shown that the workload varies with the difficulty of manoeuvring the ship, and the situation ahead.

Introduction

Due to technological innovations more information has become available for the navigator. This is especially important during the approach and departure of a harbour, when the risks of grounding and collision are highest. In most harbours around the world a pilot comes on board to assist the crew with the navigation task. Normally the pilot performs the entire navigation task. If better information is to be provided, we need to know what information is required at what time and place. Very few attempts have been made to do a thorough task analysis (Mara, 1968). Navigation is a difficult process to describe. Most information needed is obtained from the visual scene, few orders are given, and performance is hard to quantify. One of the few variables that can be quantified is mental workload. This provides an opportunity to measure the effort during decision making.

In process control all relevant parameters must be known. But when a pilot comes on board he has very little knowledge about the specific characteristics of the ship. A number of characteristics is available on a manoeuvring sheet, but since this information is only valid at full sea, the pilot will not even look at it. Personal observations and interviews have learned that the pilot only knows the basic dimensions of the ship, and he estimates characteristics by simple rules of the thumb and experience. Yet he will bring the ship to berth with a very high success rate.

Task analysis

In task analysis decomposition methods are best known. They start from the goals and narrows down towards detailed operations. From this an analysis can be done of all necessary information, controls, training, workload, allocation, etc. (Kirwan, 1992). A key factor in this approach is the set of operations. Unfortunately in shipping, and especially in piloting ships, there are very few operations. For operations all the system components must be known beforehand. In piloting this information is not available: all ships and crews are different and the environment is highly variable, making a unique operation for each instance necessary. All other decomposition methods suffer from the same drawback. A different approach is needed.

What is known, are the goals. Most other elements are not known beforehand, with the exception of the topography of the harbour with its infrastructure. To describe the task of the pilot, a goal directed approach is chosen to construct a pilot model. The centre of focus is what is what must be done, not how it is done.

Navigator Model

Many navigator models have been made over the years. Most of them are based on open-sea manoeuvring. Few of them are actually focused on inland navigation, and these often take a control theoretic approach (Papenhuijzen, 1994). The navigator model presented here is a decision model with three levels of control (see figure 1).

On the top level the overall planning of the voyage, long-term planning (LTP). Decisions on this level range from very simple decisions such as a free berth and the availability of tugs, to very complicated decisions about an overall plan to sail the ship to its berth. Most planning is done before the pilot boards the ship, and the pilot will check on board if everything is in order. Characteristic of long-term planning is that procedural and it is done once.

The middle level is short-term planning (STP). When a pilot brings in a ship, he will not plan the track as a whole in advance: he will sail the ship from waypoint to waypoint. Each waypoint will have its own objective: a new course, a curve, a change in current, an intersection. In STP the decision is made about a situation at hand. The controls which the pilot has available are very limited: rudder and engine. The decision horizon depends largely on the manoeuvrability of the ship in combination with the environment. Due to the inherent inertia of the ship, decisions often has the form of a point of no return, like in during the take-off sequence of an airplane. The choice made on this level is which track to follow. Short-term planning is based on knowledge about the forthcoming local situation, traffic, the training of the pilot, standard manoeuvres, and his personal style. This, and the manoeuvring characteristics of the ship, will be applied to plan a track. The manoeuvring characteristics and the required accuracy determine what type of control he will apply: speed-course (V,φ) or thrust-rudderangle (F,δ). Ships react slow to new settings, depending on the ship's size and type of machine control. Therefore, setpoints must be planned well in advance. Short-term planning is a discrete process, it has the form of methods, it is proactive, occurring at specific points, depending on the on-coming situation.

The lowest level of control is tracking behaviour. The track that has been chosen on the STP-level must be followed as well as possible. For this the navigator must observe the position and movements of the ship relative to the environment. When deviations from the desired track occur corrections must be made. Important factors are wind, current, ground-effect, and various elements in ship control, and the accuracy with which the position fixing can be done. Tracking is a continuous process, it is based on techniques, and it is reactive by nature. The most important characteristics are summarised in table 1.

Table 1. Most characteristic elements of the three levels of control.

level	character	type
long-term planning	procedures	once
short-term planning	methods	discrete
tracking	techniques	continuous

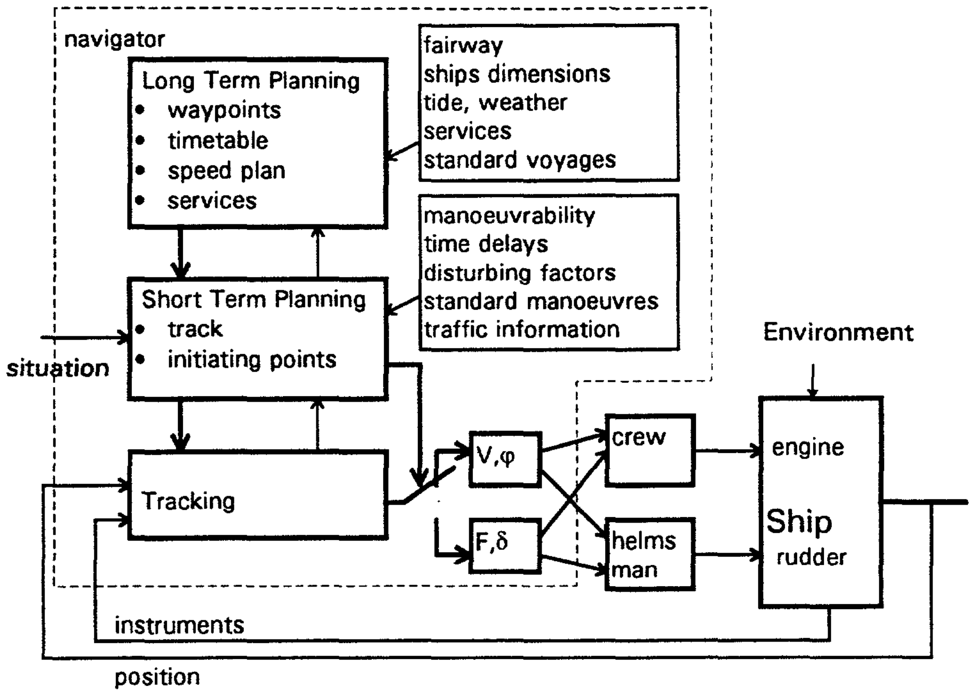


Figure 1: A navigator model.

Workload

In this paper, the term 'workload' always refers to mental workload. Physical workload resulting from boarding and leaving the ship, or irregular working hours is not considered. Performance shaping factors are also not considered.
Not all factors that may add to the workload are present at all times, and even if some factors are present, they may be unimportant. A voyage can be cut down into a number of sections, each characterised by a few elements. These elements determine the track to be chosen. From this track it follows what manoeuvres must be made, and how they are to be initiated.

Two elements are needed for tracking: observation and control. After a track is chosen it must be possible to observe the deviation from the desired track. The uncertainty of the position, related to the required accuracy, will add to the workload. During tracking different types of disturbances act of the ship. It is assumed that they are not always of importance. When the effect of the disturbance is below a certain threshold it has no effect on the task. Over the threshold it does have effect. One could model the influence of such disturbances as a linear combination adding to the workload. The use of speed-course control is considered to be easier than power-rudder control. This would also linearly add to the workload.. All these workload components are directly related to the situation present. Additionally, changes in the disturbances will add temporary to the workload.

Short-term planning will result in a workload present well before a situation occurs. It will be related to a specific situation and occur at a time that is related to the manoeuvrability of the ship. It will depend either on the fairway with its local problems, in which case it will be related to a specific location/situation, or it will depend on other ships, in which case it will be related to the traffic situation.

Long-term planning will lead to an increase of workload for a short time when arriving on board, and play a role later only if the planning has to be changed for reasons unknown at first.

Measurement

To verify the model a method must be applied which will provide workload information about the transitions. Generally, there are four ways to measure mental workload: task demand load in relation to performance, primary task measures, secondary task measures, and physiological measures. Much has been written about each of them and about their possibilities and limitations. (see for instance: Hancock & Meshkate, 1988; Sheridan & Stassen, 1979; Johanssen et al, 1979)
Performance measures are very hard to obtain in navigation tasks. The number of unknown factors is high, and personal preferences will lead to different outcomes.
Subjective workload estimates are not very good for measuring transients in workload. They provide an overall rating of the workload. Due to the specific working conditions, one may doubt the possibilities for self-rating by pilots. The TLX-ratings obtained provide information for wild speculation. Our idea is that pilots have learned to ignore their perception of workload.
To apply a secondary task, the primary task must be known exactly, which is impossible due to the very nature of this project. Additionally, a secondary task will increase the workload and is intrusive by its very nature, which is considered impermissible (Damos, 1991).
Despite the fact that the relation between mental workload and various physiological measures is not clear at all, some measures have shown high correlation between workload and the physiological reaction. Blood pressure, hormonal concentrations, and muscle tension are well known examples. Heart rate and heart rate variability (the power in the 0.1 Hz-band) were chosen. It is easy to record, sensitive, especially for central processing and visual perception, and non-intrusive. Although its reliability and validity have often been questioned, there is evidence that this measure is a reliable measure of workload. A few of these authors are Vincente et al (1987), Wastell (1989), Mulder, G (1983), and Mulder, L.J.M. (1988). The major disadvantage is its problematic interpretation, a drawback for all physiological techniques. It was assumed that the effects caused by a little physical exercise would be small compared to the differences caused by mental workload.

In a series of recordings, four experienced Rotterdam pilots participated. Each pilot did at least two inward bound and two outward bound voyages. The ships were selected for maximum size from the ships available. Ships for the river were preferred to ships for Europort. All pilots were videotaped, and marker points were plotted on a map for specific locations and traffic.

The ECG recordings were corrected for recording errors, and transformed into power density signals using Carspan 1.99 (Mulder, 1988). The power density was calculated using a Direct Fourier Transform based on the Integral Pulse Frequency Modulator model (Rompelman, 1986). The time window of 50(s) was found to be acceptable for both a sufficient spectral resolution and stationary signal. As a smoothing filter, a modified discounted-least square filter was applied with τ=30(s) (SWOV, 1994).

As was argued when you need to start, you need to know these goals. But goals, unlike actions, cannot be observed. Goals can only be obtained by interviewing, and sometimes by interpretation of the actions by an expert. Interviewing during the voyage may be a serious disturbance of the ongoing process, and was rejected for that reason. A compromise was found by asking the pilot afterwards what had happened during the voyage, what the plans were, and what the problems were in executing these plans. Over the time, some knowledge was obtained about the pilot's task, and this could be used in asking specific questions about the events that had taken place. Additionally, some knowledge about standard voyages was obtained by interviewing outside the recorded voyages.

Results

Because of the many differences between the voyages, no statistical analysis has been done over different pilots, ships, and situations. Voyages were analysed by comparing the voyage with the heart rate and power density. One voyage has been chosen as an example. In figure 2 one can see the track, and in figure 3 the resulting ECG-data. Apart from climbing on board, there is no physical exercise by the pilot. The heart rate of all pilots on board stabilised around 90 beats/min, about 10 beats higher than on the pilot-ship. Climbing on and off board can be seen very clear in all heart rate plots. Beforehand we distinguished between inward-bound and outward-bound voyages. In the heart rate plots this difference shows very clearly when entering the harbour. The heart rate increased significantly, to reach its peak at the start of the turn. This effect could not be found with outward-bound ships. The same could be found during critical manoeuvres and passing bridges. This effect was attributed to arousal, that is directly related to hormonal effects (Gellatly & Meyer, 1992).

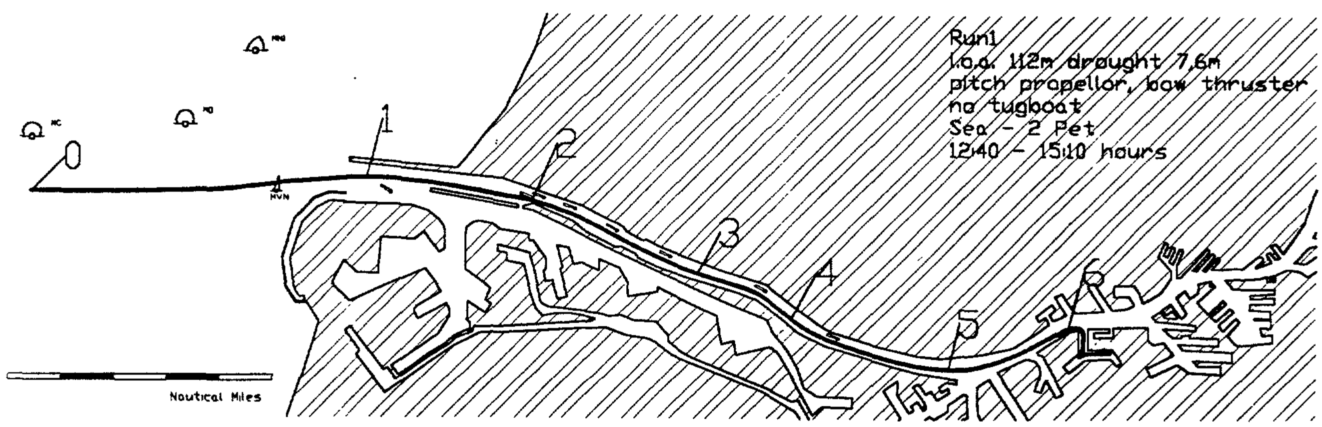


Figure 2. A voyage from sea to a Rotterdam harbour. The nmubers correspond with figure 3.

In computing the power density it was found that in the harbour many artefacts occurred. This is very troublesome because they act like Dirac-functions, masking all other data in that time segment. It is known that stress-like situations often result in extra-systolic contractions of the first ventricle, recording as an extra heartbeat, and disturbing the Fourier transforming process. A second outcome was that the power density is highly variable, which made it difficult to interpret. Moving average filters and frequency based filter techniques were found unsatisfying. The Modified Discounted Least Square filter was found to be most useful. This filter is an a-symmetrical exponential time-weighted filter.

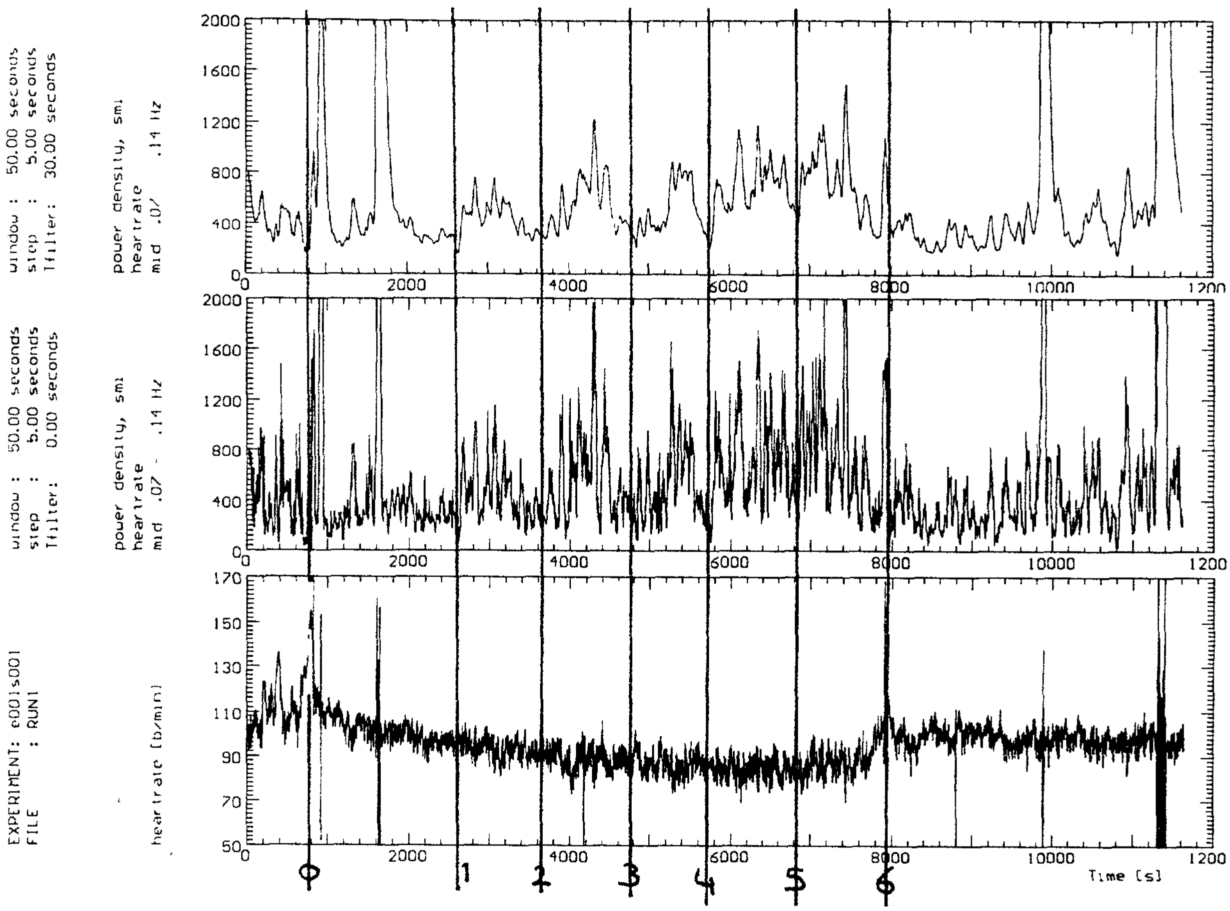


Figure 3: Physiological data from the voyage shown in figure 2. From bottom to top: the heart rate, the power density (unfiltered), the power density (filtered). The numbers on top of the vertical lines correspond with numbers in figure 2. Note that the mental workload is reversely related to the power density of the 0.1 Hz-band. The peaks in the power density result from various artefact in the heartrate.

The increase in heartrate is clearly visable during boarding at sea (0), and when entering the harbour (6). At marker 2 there is an increased workload due to traffic and construction work (between 2 and 3). Marker 3 is the result of a ship that came to the wrong side of the fairway. Marker 4 is located at a place where difficult currents exist. Marker 6 placed just before a location of usually intense traffic.

The overall profile of the power density gave a good indication of the workload (face value). The peaks in workload as indicated by the power density occurred at pre-defined places, such as before a significant change in the current, or when decisions had to be made about passing a ship coming from another fairway. Orders, passing of other ships and disturbances could often be well identified. As is usual with these techniques, lot of variation is left unaccounted for. In the first series, peaks in workload were found which later, during interviews with other pilots, were identified as specific problems.

Conclusions

The use of ECG-recording seems a valuable tool for developing a navigator model. Data obtained from interviews and the ECG-recordings provided valuable information about the pilot's task. For a more sophisticated interpretation a thorough talk-through afterwards is required, making the interpretation much easier. In field experiments such as these, this is not always easy to achieve. Often there is very little time between ships, due to high time-pressure, and the work is done around the clock. Only with full co-operation from the entire organisation time is made available.
The data obtained seem to confirm the model presented here. All major workload peaks which were found to be related to locations or situations, were confirmed in interviews. For a more detailed model, more detailed analysis is needed. We think it will be possible to identify most of the dominant factors in workload using this technique.
For the interpretation of the heartbeat data, the pilot actions, the goals, the characteristics of the ship, the location of the ship, and the situation it is in must be known. To do this, a lot of information must be recorded during and after the voyage. Unfortunately, due to many unknown factors, not all changes in power density may be explained.
For a good analysis of the heartbeat data a good reconstruction of that data is most valuable. Disturbances make the time-window in which it occurs useless, and when disturbances occur regularly it can make a large portion of the recording useless. It seems to us that a better reconstruction may be achieved than present available with Carspan 1.99.

Acknowledgements

We would like to thank the Rotterdam Pilots, the participating pilots, and the captains and crews of the ships for their co-operation and hospitality.

Literature

Damos, D.L., 1991, Dual task methodology: some common problems. In: D.L. Damos (ed), *Multiple task performance*. Taylor and Francis, London.

Gellatly, I.R. & J.P. Meyer, 1992, The effects of goal difficulty on physiological arausal, cognition, and task performance. *Journal of Applied Psychology*, vol 77, no 5, pp 694-704.

Hangcock P.A. & N. Meshkati (eds), 1988, *Human mental workload*. Advances in psychology 52, North Holland, Amsterdam.

Johanssen N.; N. Moray; R. Pew; J. Rasmussen; A. Sanders & C. Wickens, 1979, Final report of the experimental psychology group. In: N. Moray (ed), *Mental workload: it's theory and measurement*, p 101-114, Plenum Press, New York.

Kirwan B. & L.K. Ainsworth (eds), 1992, *A guide to task analysis*. Taylor & Francis, London.

Mara, T.D., 1968, *Human factors in ship control, Vol III: Function and information requirements of deck watch officers and other deck personnel*, Clearinghouse, Springfield, U.S.A..

Mulder, G., 1983, Attention, effort and sinusarrythmia: how far are we? In: J.F. Orlebeke, G. Mulder & L.J.P. van Doornen (eds), *Psychophysiology of cardiovascular control.* p 407-424, Plenum Press, New York.

Mulder, L.J.M.; H.J. Van Dellen; P. Van Der Meulen & B. Opheikens, 1988, Carspan: a spectral analysis program for cardiovascular time series. In: F.J. Maarse; L.J.M. Mulder; W. Sjouw & A. Akkerman (eds) *Computers in psychology: methods, instrumentation & psychodiagnostics.* pp 39-47. Swets & Zeitlinger, Lisse.

Mulder, L.J.M., 1988, *Assessment of cardio vascular reactivity by means of spectral analysis.* State University of Groningen, the Netherlands, Thesis.

Sheridan T.B. & H.G. Stassen, 1979, Definitions, models and measures of human workload. In: N. Moray (ed), *Mental workload, it's theory and measurement.* Plenum Press, New York, 1979.

Papenhuijzen, B., 1994, *Towards a human operator model of the navigator.* Delft University of Technology, the Netherlands, Thesis.

Rompelman, O., 1986, Tutorial review on processing the cardiac event series: a signal analysis approach. *Automedica*, vol 7, pp 191-212.

SWOV, 1994, Control strategies for a highway network; A joint research project of the SWOV, the Technical University Delft and the Institute for Perception TNO, sponsored by the Dutch Ministery of Transport and Watermanagement. Part I, II and III. R-94-34 I/II/III. SWOV, Leidschendam, the Netherlands.

Vincente K.J.; D.C. Thornton & N. Moray, 1987, Spectral analysis of sinus arrhythmenia: a measure of mental effort. *Human Factors*, vol 29, no 2, pp 171-182.

Wastell, D., 1989, Mental effort and task performance: towards a psychophysiology of human computer interaction. In: D. Diaper et al (eds), *Human-Computer Interaction - INTERACT '90.* North Holland: Elsevier Science Publishers B.V., Amsterdam.

The Dynamics of Mental Effort

F.R.H. Zijlstra

Work & Organizational Research Centre
P.O. Box 90153 Tilburg University
e-mail: f.r.h.zijlstra@kub.nl

1. Introduction

The difficulty with respect to research on (mental) workload is perhaps most clearly illustrated by the diversity in concepts that have been proposed and the delicate distinctions between those concepts. Some examples are: objective workload, functional workload, perceived workload, experienced workload, subjective workload, *relative* subjective workload. These concepts are needed to explain the inter-individual differences with respect to the results in workload studies (Moray, 1982; Vidulich & Tsang, 1986; Hart & Staveland, 1988; Tsang, 1994).
Implicit in this line of reasoning is the expectation that when tasks are identical, the demands of these tasks imposed upon the operators are similar and therefor should lead to identical results in workload studies. Although there is also some awareness that individuals differ from each other to some extent which may account for the fact that various individuals may perceive a particular load differently, this is usually regarded as an artefact.
However, what we actually are interested in is the *costs* at which human operators execute their tasks in order to predict (and prevent) situations of potential breakdown. In order to get an adequate estimation of these costs the (differences in) capacities of the human operator should explicitly be taken into account in any model or study concerning workload. In this paper it will be argued that the concept of (mental) effort may be very fruitful in tackling these problems. We will in fact restrict ourselves to task with a predominantly cognitive nature, and therefore refer to *mental* workload and *mental* effort when we speak of workload and effort.
After discussing some conceptual issues, some experimental results will be presented to illustrate and support the conceptual remarks.

2. Some conceptual remarks

The concept of workload refers to the fact that the demands of a task impose a certain load on the human operator. With respect to *mental* workload it is referred to the degree to which the information processing system is occupied, or loaded. This concept became popular because in certain situations it seemed that operators were 'over loaded', or 'over charged'. In these situations operators started to make errors, or developed symptoms of acute, or chronic stress. These are the situations in which the demands of the task (or entire work situation) exceed the available capacities of the human operators. Inherent in this notion is the assumption that the capacity of the information processing system is limited. The nature of these limitations is therefore something which is of interest to engineers and designers of man-machine systems.
On the other hand the focus of research quite often points to the demands of the task. It is assumed that by analyzing these demands one may get a good understanding of the *impact* of those demands on the human operator, and thus the workload.

However, there are some theoretical and empirical problems with this approach. First of all the concept of workload is a rather *static* concept. Since this concept is directly related with the demands of the task it is implicit that as long the task demands do not change the workload does not change either. This is in contradiction with empirical results which show that one and the same task may have a different impact on various operators, but also on the same operator at various times of the working day (Zijlstra, 1993).
A second problem is that from a work psychological perspective this approach is not plausible, since this approach implies that people are just passively exposed to task demands. It can be quite easily demonstrated that people actively intervene in their work situation. Actually that is what the word 'operator' stands for. For example, if someone is experiencing noise hindrances from outside, or a cold coming in, the first thing to be observed is the window or door being closed in order to deal with these inconveniences. And if one expects some kind of inconvenience it can be observed that people may anticipate by making sure that doors or windows are closed. This example illustrates that people are actively engaged and dealing with the situations they are in. This of course is also true for work situations. By choosing various work strategies operators can actually influence the impact of the demands that are imposed on them. Again there is empirical evidence that workers, by using various strategies, manipulate their own task load (Sperandio, 1978: Teiger, 1978; van Aalst, et al., 1986). These strategies ensure that their work capacity is evenly distributed throughout the day. Furthermore workers actively monitor and supervise their own activities in order to see whether their actions have brought them any closer to the goals they have set themselves. Consequently they may decide to change their action pattern, or strategy (i.e. try a different approach, try harder, etc.). This means that workers also may anticipate upon future actions.

The conclusion of the arguments presented above should be that, in order to estimate the costs at which operators execute their tasks, the focus should not be exclusively on the demands of the task, but we should focus more on the capacities of the human operator. Or rather to say the *potential* the operator has to perform (i.e. 'performance potential', Meijman & O'Hanlon, 1984). A person's performance potential refers to all the capacities and resources (i.e. skills, knowledge, experience, etc.) that one person may potentially be able to mobilize to perform a particular task, or to meet particular demands. In line with Kahneman's ideas (1973) we assume that the capacities and resources of an individual are influenced by his momentanuous psycho-physiological state. Loss of sleep, fatigue as a result of prolonged work, circadian effects, intake of alcohol or medicins, and lack of external stimuli may deteriorate the worker's psycho-physiological state (see also O'Hanlon, 1981; Hockey & Hamilton, 1983). On the other hand worker's are continuously acquiring new experiences which may lead to new skills. This influences the performance potential in a positive way. From this point of view it is evident that there are inter-individual differences with respect to the potential of individuals. But there are also intra-individual differences. This means that the potential that an operator is able to mobilize may vary from day to day, or perhaps even from moment to moment. Therefor it should be stressed that the operator's potential to perform is not steady, but is continuously changing.

The conclusion of this line of reasoning is that the effects of task demands are determined by the *interaction* between these demands and the operator's performance potential. To be more precise the costs of carrying out a particular tasks depend on the degree in which an operator is prepared to mobilize his potential. This implies an additional remark to be made. Although a task may impose certain demands upon the operator, or one could say charges the performance potential, this does not automatically 'trigger' an operator to mobilize his

potential. This requires an active and conscious decision by the operator to execute that particular task. Consequently a motivational aspect is involved as well.
From a work psychological perspective it is assumed that a person has to 'accept' the task. Therefor a distinction is made between the 'objective' task and the 'subjective' task (Hacker, 1978, 1986). The objective task is the formal prescription of what the organization expects a worker to do, this may refer to all practical aspects as 'what', 'when' and 'who'. These information can usually be found in the task description, and is usually published in announcements of particular vacancies.
The subjective task refers to the worker's perception of what the task is, or in other words, what he thinks the organization expects him to do, and how, etc. So, the subjective task means that the worker has reformulated the task into a 'personal goal'. And in fact this personal goal should constitute the motivation to act. However, it should be noted that sufficient degrees of freedom are a prerequisite. Due to organizational and technical constraints, or safety procedures, the degrees of freedom may be reduced. This also reduces the possibilities for individual 'redefinition', which may cause additional problems that can lead to deteriorated performance of a Man-Machine system.

Fluctuations in the performance potential may account for fluctuations in workload estimations. Therefore a focus on demands is not sufficient. The focus should be on the degree to which a worker is prepared to meet those demands. In fact what counts is the degree to which a worker mobilizes his available potential. Or one could say the amount of effort a worker is prepared to invest. The effort investment may be considered as an adequate indication of the level of costs at which an operator is performing his task (Zijlstra, 1993). Apart from fluctuations in the performance potential (acquiring new skills, etc.) over a relatively lage frame of time, the potential an operator has available is also influenced by changes in the psycho-physiological state (fatigue, motivation, etc.) that may fluctuate within a rather short time frame. This may result in the fact that one particular task may constitute a different *functional* load for an individual at different moments, even within one day. And as a consequence that person has to mobilize more effort to execute the same task at one moment than at another moment. When the actual psycho-physiological state of an operator does not match with the state that is, taking the task demands into account, required at that moment, the operator has to mobilize additional effort to keep his performance level within acceptable limits. As a consequence he may be able to perform at an acceptable level, but at a higher level of costs (Zijlstra, 1993).
Therefor the amount of (mental) effort that an operator has to invest is an adequate indicator of the level of costs at which that operator is performing. Whenever an operator performs a task, even the most simple task, he has to invest some effort in order to execute this task successfully. In vigilance tasks the demands of the work situation (staying alert) are not extremely high when the operator is fresh and fit to work after a good night sleep. However, the same demands may be very high when the operator is weary. At that moment his actual state does not match the required state. Such a dismatch can be overcome by additional effort investment of the operator to adjust his state.

In conclusion we may state that effort refers to the degree to which an operator is mobilizing his performance potential to meet the demands that are imposed by the work situation. Metaphorically speaking effort can be symbolized by fuel, but we should keep in mind that the size of the barrel containing teh fuel may fluctuate. Effort measurements may provide us with an adequate estimation of the level of costs that are involved in executing tasks, or perhaps better to say the efficiency level at which an operator is operating (Zijlstra, 1993).

This statement will be illustrated with some results of experimental and field studies in which the performance potential of human operators have been manipulated.

3. Why subjective measures?

The discussion on pro's and con's of subjective measures actually has never stopped. There are still researchers who are sceptical with respect to the use of subjective measures in general and in research on workload and effort in particular. The arguments against subjective measures not only refer to the aspect of reliability, but also concentrate on the aspect of validity. (cf. O'Donnell & Eggemeier, 1986). On the other hand, empirical evidence makes a plea for subjective measurements (cf. Gopher & Braune, 1984).

However, I would like to make a point that goes beyond the issue of whether subjective ratings are reliable or not. The experiences of an operator affect his subsequent behaviour, and thus also their performance and physiological responses. When an operator is feeling tired, or when he is less motivated he will probably act accordingly. This means that he may decide to avoid high demanding situations and may choose a strategy that is in accordance with his feelings.
Although subjective measures have proven to be quite sensitive (Casali & Wierwille, 1983), the question remains whether methods based on human experience of effort can indeed provide the information they claim to provide. In other words the construct validity of instruments in this domain should be of special interest.

In the next section some empirical results will be presented that demonstrate the fact that the performance potential is an important factor that has to be taken into account. Furthermore the results demonstrate that a simple uni-dimensional rating scale, The Rating Scale Mental Effort - RSME, is a sensitive and useful instrument for measuring effort. These results are part of an extensive study on the validity of this scale (Zijlstra, 1993).

4. The dynamics of effort

The first indication that the performance potential indeed is a very relevant aspect to take into account comes from a study among bus drivers. As part of a large project in which the workload of bus drivers was studied, 27 bus drivers agreed to participate in an extensive research program. The drivers were examined on several different days. At least two of these days were working days and one was a duty-free day. In this study the drivers were required to rate their effort expenditure at various moments during their working day.
Furthermore they had to perform specially developed laboratory tasks that were designed to measure workload effects on a routine working day. These laboratory tasks had to be performed three times: at the start of the working-day (at about 9.00 hours), during the lunch-break (13.00 hours) and at the end of the working day (17.00 hours). This procedure was repeated on a second working day. First the results of this laboratory part of the study will be described.

These laboratory tasks are so-called visual memory-search tasks (Sternberg, 1969; Massaro, 1975; Mulder, 1980). During such a task a set of stimuli is presented on a video screen for a short period. This set of stimuli consists of a variable number of letters (1 - 4), the so-called 'display set'. The subjects are asked to indicate whether the (or one of the) letter(s) presented belong(s) to a previously shown set of letters, which should be memorized (the so-called 'memory-set'). The subjects have to respond by pushing one button for a 'yes' response and another button for a 'no' response to show that one or more of the letters of the presented display-set does or does not belong to the memory-set. The memory-load of the task can be manipulated by varying the number of letters in the memory-set. This means that the difficulty of the task or, one should say, the task load, is varied. (For a detailed description of this task see: van Dellen et al., 1985; Aasman et al., 1988).

In the experiment to be described here memory-sets of two letters and of four letters were used. Furthermore a situation was created in which a subject was requested to count how often certain letters from the memory-set were presented. This resulted in four different task load levels with increasing information loads:

1. 2 letters in memory and pushing a button in response;
2. 4 letters in memory and pushing a button in response;
3. 4 letters in memory and counting in response;
4. 4 letters in memory and counting and pushing a button in response; this task may be regarded as a dual-task.

The task load levels, each consisting of a high number of trials, were presented in a random order. After each task load level the subjects had to rate their effort expenditure by means of the Rating Scale Mental Effort (RSME). Other measurements were also made, like reaction times and heart rate variability. A task like the one described above takes about 20 minutes to complete.
The results of the ratings on the first working-day are presented in Figure 1^a and the results gathered on the second working-day are presented in Figure 1^b.

here Figures 1 (a and b)

According to both figures the RSME is able to differentiate systematically between various information load levels. The mentally more demanding tasks are given higher scores than the mentally less demanding tasks. Analysis of Variance showed significant effects for the factors 'task level' ($F(3,24) = 31.2$; $p < .001$, and task level 1 < task level 2 < task level 3 < task level 4) and 'time of the day' ($F(2,25) = 7.4$; $p = .003$, and score on time 9.00 < score time 13.00 < score time 17.00). No significant effect for the factor 'working day' was found. This is important, because comparing the effort-scores on both working days (Fig. 1^a vs. 1^b) gives an indication about the reproducebility of this rating scale as the second working day is a replication of the first one. Since there are no significant differences we can conclude that the Rating Scale Mental Effort provides us with reliable measurements. The RSME scores of working day 1 correlate (product moment correlation) $r_{pm} = .81$ ($p=.001$) with those of working day 2.
The increase of the RSME score accompanying the increase in information load is an indication that the RSME provides information on mental effort as 'controlled information processing': the higher the information load, the more mental effort is needed to process the information.

The differences between the various levels of information load were also reflected in the reaction times because a considerable increase in reaction time was seen between the simplest condition (level 1) and the dual task condition (level 4). In addition some effects of the

increase of information load could be traced in the physiological parameters. Heart rate was found to be significantly faster in the dual task condition than in the simplest condition while heart rate variability was lowest in the dual task condition (see for an extensive description of reaction time and heart rate results: Aasman et al., 1987, 1988). These physiological reactions are presumed to accompany mental effort exertion (Mulder, 1980).
The findings with respect to the reaction times can be regarded as an indication of 'converging' evidence of the validity of the RSME while the physiological results can be regarded as an indication of the 'concurrent validity' of the RSME.

Furthermore it appeared that the RSME score for each task level was systematically highest at the end of the working day (17.00 hours), while the information load of the experimental task was the same as at the start of the working day. By 17.00 hours the subjects had done a day's work which brought about a change in their psycho-physiological state. The bus drivers had been taxing their resources and so their performance potential was lower: they were tired. The increase in the RSME score can be regarded as an indication that the RSME also provides information on effort like 'executive resource control'. Since a bus driver's psycho-physiological condition was sub-optimal he had to compensate by investing extra effort in order to perform at an acceptable level.

RSME scores were also gathered while the drivers were performing their real task of driving a bus. Several times during their working day the bus drivers had to rate how much effort they were exerting. The numbers of passengers transported and the time left to spare upon arrival at the end of his route was also noted. These parameters were used as indices for the task load. Figure 2 presents the results of this effort rating.

here Figure 2 (a en b)

As can be seen in Figure 2 the bus drivers reported that they have to invest more effort to carry out their task at the end of the day. Analysis of variance teaches us that there is no significant effect of the factor 'working day' (w1 vs w2) on the one hand but that on the other hand it appears that the factor 'time' (end of the day versus early morning) shows a significant effect ($F(5,23) = 11.6$; $p = .001$). No significant interaction effects were found. It turned out that an increase in number of passengers during the morning period correlated ($r_{pm} = +.30$) with an increase in effort ratings and a decrease in spare time at the end of the bus route correlated $r_{pm} = -.33$ with an increase in effort ratings (These task load indices correlated $r = -.31$ when scores of differences were calculated). During the afternoon period the correlations became respectively greater: $r = +.46$ and $r = -.64$ for passenger numbers and amount of time to spare. The correlation between both task load indices also increased ($r = .-53$) (see also Meijman, 1991 page 180).
The dip in the curve of the effort ratings in the mid morning period (the differences between times 1, 2, and 3 was found to be significant) and to a lesser extent (not to a significant extent) in the afternoon period. This reflects the fact that around those periods (after the morning rush-hour and before the evening rush-hour) there are fewer passengers and there is less traffic. This makes the bus driver's task easier. Moreover we saw the same pattern on both working days. The scores of working day 1 correlate highly ($r= .71$; $p= .001$) with those of working day 2. This again is an indication that subjects give reliable estimates.

Additional information can be gathered from other experiments where the RSME has been applied. One such experiment had to do with the tasks of nautical officers responsible for navigating a ship. Their tasks were simulated on a ship's-simulator (see also Perdok, 1984). There are usually several persons present on the bridge of a sea-going vessel at any one time and each of these people has his own specific task be it connected with navigation, communication or controlling the ship's engine-room (speed regulation). When a ship is equipped with advanced technological instruments like sophisticated automatic pilot systems it is possible to make do with fewer crew members on the bridge. This simulation study was about 'one-man manning' situations on the bridges of sea-going vessels. The subjects were given various course permutations along which they were expected to sail while remaining responsible for the navigation, communication (coast-ship) and engine-room (speed control). These appeared to be the most relevant task-aspects. The scenarios variations were: differences in weather conditions and speed and in the courses of oncoming shipping.

The subjects of this study were eight nautical officers (with a mean age of about 26 years) and each had at least three years of nautical experience. Half of this group had recently finished a training with the new 'one-man manning' equipment, a so-called 'integrated training', hereafter referred to as the Seaman Integrated Training (SIT). The others had been trained in the traditional way, that is to say, in the navigational aspects of their work. Their training is hereafter referred to as the Seaman Not-adequate Training (SNT) because it did not adequately prepare them for one-man-manning situations.

While the subjects executed their task they filled in the RSME and after completing the task they filled in the 'Schaal Ervaren Belasting' (Scale Experienced Load - SEB; Meijman, 1991). This SEB consists of a questionnaire with items pertaining to symptoms which indicate various degrees of fatigue. With the RSME the subjects had to rate the set of task elements as a whole and individually as well. Their performance was also registered. Several methods were used for determining the quality of the performance (Perdok, 1984). Amongst others things the passing distance between the ship being monitored and other ships served as a measure of performance quality. In navigation circles a passing distance of less than one nautical mile is regarded as dangerous.

During the experiment the researchers discovered that the subjects had different strategies for avoiding collisions than they had expected them to have. So it was that the theoretical gradations of difficulty that had been manipulated in the various scenarios lost their validity. The average scores for the various scenarios are presented in Table 1.

Table 1.: Comparison of RSME scores of two groups of nautical officers with respect to the various tasks and the set of tasks as a whole.

VARIABLES	Average score SIT-group	Average score SNT-group	F	sign. (α=.05)
$RSME_{Tasks}$	31	48	5.2	.03
$RSME_{Navigation}$	28	44	4.8	.04
$RSME_{Engine\text{-}room}$	27	39	2.9	.09
$RSME_{Communication}$	28	36	1.9	.17
Miles (passing by)	2.3	1.8	4.4	.05
SEB	8	11	5.5	.03

The SNT group of subjects reported having put significantly more effort into the tasks as a whole and into the task of Navigation. They had higher SEB scores indicating that they had endured heavier workloads and their performance was worse. This means that their qualification discrepancy was not compensated even though they exerted more effort. The difference with respect to the Communication and Engine-room tasks was not significant although the SNT group had higher RSME scores. It may be noted that it is not remarkable that there is no significant difference with respect to Communication as this can generally easily be time-shared with other tasks.
Moreover it appeared that the RSME score correlated with the performance-indicator 'passing-distance' ($r = -.57$; $p<.05$). This means that higher effort-scores are reported when other ships are passed at shorter distances.

This study demonstrates that where one group is more qualified for a task than another group the difference in effort investment between the two groups can be measured with the RSME.

In previous study we have mainly dealt with situations where the psycho-physiological state of subjects was negatively affected because subjects were more or less fatigued. This has lead to a decrease of performance potential. The changes in performance potential that these situations created were of a temporary nature. Consequently, by investing more effort subjects could see to it that their performance did not worsen.
In this study we have been dealing with a more stable aspect of performance potential: the level of qualification. Moreover the performance potential of one group was 'up-graded' by means of training so that these individuals became better suited to the task. This resulted in a lower level of effort exertion, in better performance and in a lighter workload as indicated by the SEB.
This conformed with the theoretical concept of effort.

Another aspect that is of relevance to the validity of the RSME concerns the relation between the RSME and the SEB. In the study described above we have seen that one group of subjects reported that they felt more fatigued after completing the task (as expressed by means of the SEB) than the other group whilst they reported having invested more effort in the execution of the task. The SEB-scores correlate $r = +.55$ with the RSME-scores as reported by Meijman (1991, page 144).
Similar results are reported by Meijman for other studies. One of these studies was a replication of the earlier mentioned bus driver's study. In this replication study the bus drivers were examined in three different conditions. They did an early morning shift, a normal shift and a late shift.
The RSME scores for respectively morning, normal and late shifts correlated very positively with the SEB-scores: .71; .56; .74 after 4 hours of work, and .70; .77; .66 at the end of the working day (see Meijman, 1991; page 175).

The findings of Wiethoff et al. (1988) published in a study into the effects of sleep deprivation are also interesting with respect to the RSME. They report that the RSME adequately discriminates between various levels of task load. They also found that subjects who had been deprived of sleep for one night had to put more effort into the task than control group subjects, especially when the task was very simple. When the task was rather difficult it appeared that the rise in RSME scores was less. According to Wiethoff et al. subjects found it harder to stay awake when they were working on the easy task than when they were working with a difficult task and therefore they had to put in more effort to compensate for

their sub-optimal state.
These findings can be added to results of the other studies that have been mentioned.

5. Conclusion

The first question to be answered, of course, is whether this scale really measures mental effort investment. In section 3, on the theoretical aspects of mental effort, it has already been stated that the level of effort expenditure is determined by the objective task demands in relation to the 'performance potential' of the individual. Practically speaking this means that if it is assumed that an instrument measures mental effort expenditure, this instrument should reflect differences between various task load levels as well as changes in the performance potential of the individual.
In the first study in which this rating scale was used it appeared to differentiate systematically between various levels of information load and there are indications that changes in the performance potential of individuals are also reflected in the RSME score. The other studies support the conclusion that changes in the performance potential affect the level of effort expenditure. This suggests that the Rating Scale Mental Effort does indeed measure effort. Furthermore it has been demonstrated that the RSME is easy to apply both in laboratory settings and in real work situations. Subjects give reliable estimates of the various task load levels when using the RSME and the instrument proved to be non-obtrusive.

Additional information on the psychometrics of the RSME have been obtained in an experimental study (Zijlstra, 1993). These results support the conclusions thusfar that the RSME is a useful instrument.

From a theoretical perspective these results support the conclusion that variations in the performance potential may account for fluctuations in costs of task performance. This means that the focus in work load studies should be shifted from the task to the costs of the human operator.

6. References

Aalst, R. van, Kemper, H.C.G., Diederen, A.L.J. & Pijpers, V.T.J., (1986). Arbeidsbelasting van huisvuilbeladers na een reorganisatie van de huisvuildienst. *Tijdschrift Sociale Gezondheidszorg*, 64, 619-623.

Aasman, J., Wijers, A.A., Mulder, G., and Mulder, L.J.M., (1988), Measuring mental fatigue in normal daily working routines. In: P.A. Hancock and N., Meshkati (eds.), *Human Mental Workload.* Amsterdam: North-Holland publishers. p. 117-138.

Aasman, J., Mulder, G., and Mulder, L.J.M., (1987). Operator effort and the measurement of heart rate variability. In: *Human Factors*, vol. 29(2), 161-170.

Casali, J.G., and Wierwille, W.W., (1983). A Comparison of rating scale, secondary-task, physiological, and primary-task workload estimation techniques in a simulated flight task emphasizing communications load. In: *Human Factors*, vol. 25 (6), 623-641.

Dellen, H.J. van, Aasman, J., Mulder, L.J.M., and Mulder, G., (1985). Time domain versus frequency domain measures of heart rate variability. In: J.F. Orlebeke, G. Mulder, and L.J.P. van Doornen (eds.), *The Psychophysiology of Cardiovascular Control.* New York: Plenum Press. p. 353-374.

Gopher, D., and Braune, R., (1984). On the psychophysics of workload: why bother with subjective measures? *Human Factors*, vol. 26(5), 519-532.

Hacker, W., (1978). *Allgemeine Arbeits- und Ingenieurspsychologie, Psychische Struktur und Regulation von Arbeitstätigkeiten.* Bern: Huber Verlag.

Hacker, W., (1986). *Arbeitspsychologie. Psychische Regulation von Arbeitstätigkeiten.* Bern: Huber Verlag.

Hart, S.G., and Staveland, L.E., (1988). Development of NASA-TLX (Task Load indeX): Results of empirical and theoretical research. In: P.A. Hancock, and N., Meshkati (eds.), *Human Mental Workload.* Amsterdam: North-Holland publishers. p. 139-184.

Hockey, G.R.J., and Hamilton, P., (1983). The cognitive patterning of stress states. In: G.R.J. Hockey (ed.), *Stress and Fatigue in Human Performance.* New York: Wiley. p.331-362.

Kahneman, D., (1973). *Attention and Effort.* Englewood Cliffs (N.J.): Prentice Hall.

Massaro, D.W., (1975). *Experimental Psychology and Information Processing.* Chicago: Rand-McNally.

Meijman, Th.F., en O'Hanlon J.F. (1984). Workload, an introduction to psychological theories and measurement methods. In: P.J.D. Drenth, Hk. Thierry, en Ch.J. de Wolff (eds.), *Handbook of Work and Organizational Psychology.* New York: Wiley and Sons. p. 257-288.

Meijman, Th.F., (1991). *Over Vermoeidheid.* (Acad. Dis. R.U. Groningen). Amsterdam: Studie Centrum Arbeid & Gezondheid.

Moray, N., (1982). Subjective mental workload. In: *Human Factors*, vol. 24, no. 1, 25-40.

Mulder, G., (1980). *The Heart of Mental Effort.* Acad. Thesis. University of Groningen (The Netherlands).

O'Donnell, R.D., and Eggemeier, F.T. (1986). Workload assessment methodology. In: K.R. Boff, L. Kaufman, J.P. Thomas (Eds.), *Handbook of Perception and Human Performance, Vol. II: Cognitive Processes and Performance.* New York: Wiley & Sons. p. 42-1 - 42-49.

O'Hanlon, J.F., (1981), Boredom: practical consequences and a theory. In: *Acta Psychologica*, vol. 49, 53-82.

Perdok, J. (1984). *Methods to study mental workload in one-man-manning situations.* Paper presented at the 'Second International Bridge Design and Operation Forum, Hamburg august 1984. MARIN Wageningen.

Sperandio, J.C., (1978). The regulation of working methods as a function of workload among air traffic controllers. *Ergonomics*, 21, 193-202.

Sternberg, S., (1969). Memory scanning: mental processes revealed by reaction time experiments. In: *American Scientist*, vol. 57, 421-457.

Teiger, C., (1978). Regulation of activity: an analytical tool for studying work load in perceptual motor tasks. *Ergonomics*, 21, 203-213.

Tsang, P.S., (1994). The roles of immediacy and redundancy in relative subjective workload assessment. *Human Factors*, 1994, 36(3), 503-513.

Vidulich, M.A., and Tsang, P.S., (1986). Techniques of subjective workload assessment: a comparison of SWAT and the NASA-Bipolar methods. *Ergonomics*, 29(11), 1385-1398.

Wiethoff, M., Gunther, Th., Mulder, G. en Mulder, L., (1988). Mentale inspanning bij slaapdeprivatie. In: B.G. Deelman en G. Mulder (red.), *Nederlands Psychologisch Onderzoek 3: Bijdragen aan de Experimentele Gerontologie en het Onderzoek naar Mentale Belasting en Stress.* Lisse: Swets & Zeitlinger. p. 159-172.

Zijlstra, F.R.H., (1993). *Efficiency in Work Behaviour; A Design Approach for Modern Tools.* (Acad. Dissertation Delft University of Technology). Delft: Delft University Press.

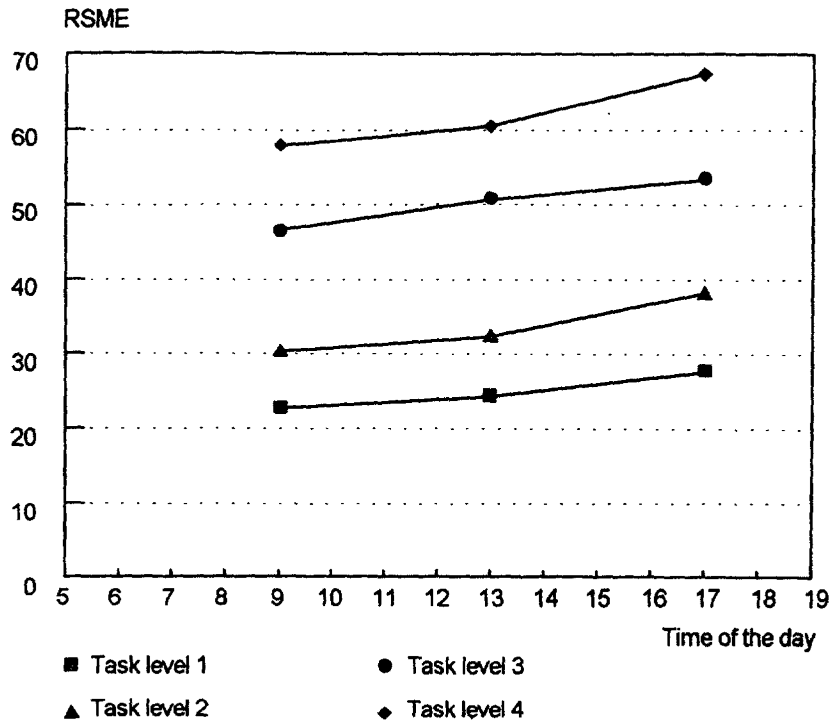


Figure 1:a: *Laboratory task, working-day 1*

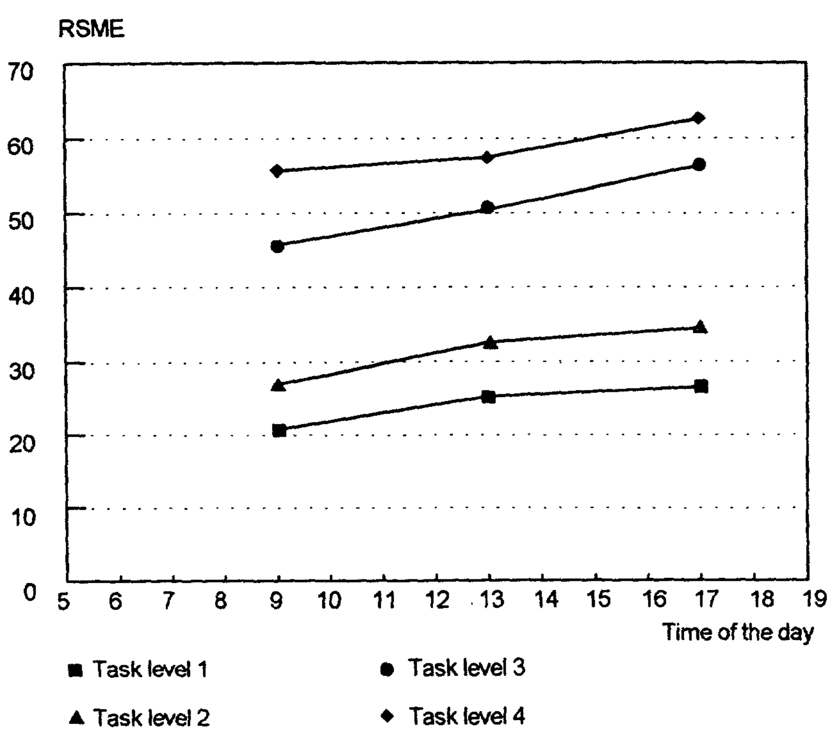


Figure 1.b: *Laboratory task, working-day 2*

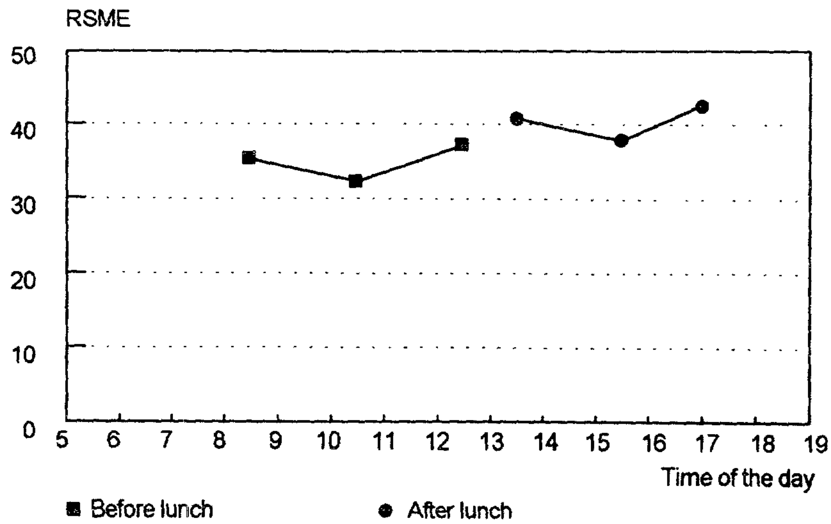


Figure 2 a: *RSME during working-day 1*

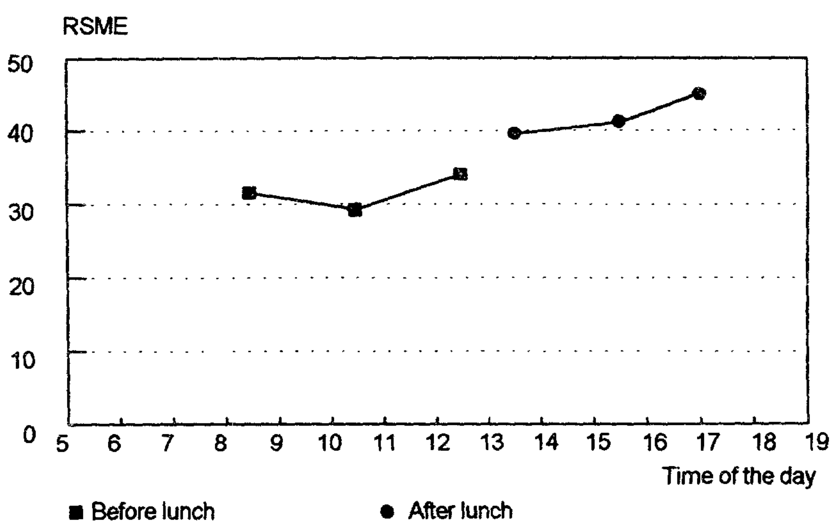


Figure 2 b: *RSME during working-day 2*

MENTAL MODELS, STRATEGIES, AND OPERATOR INTERVENTION IN SUPERVISORY CONTROL

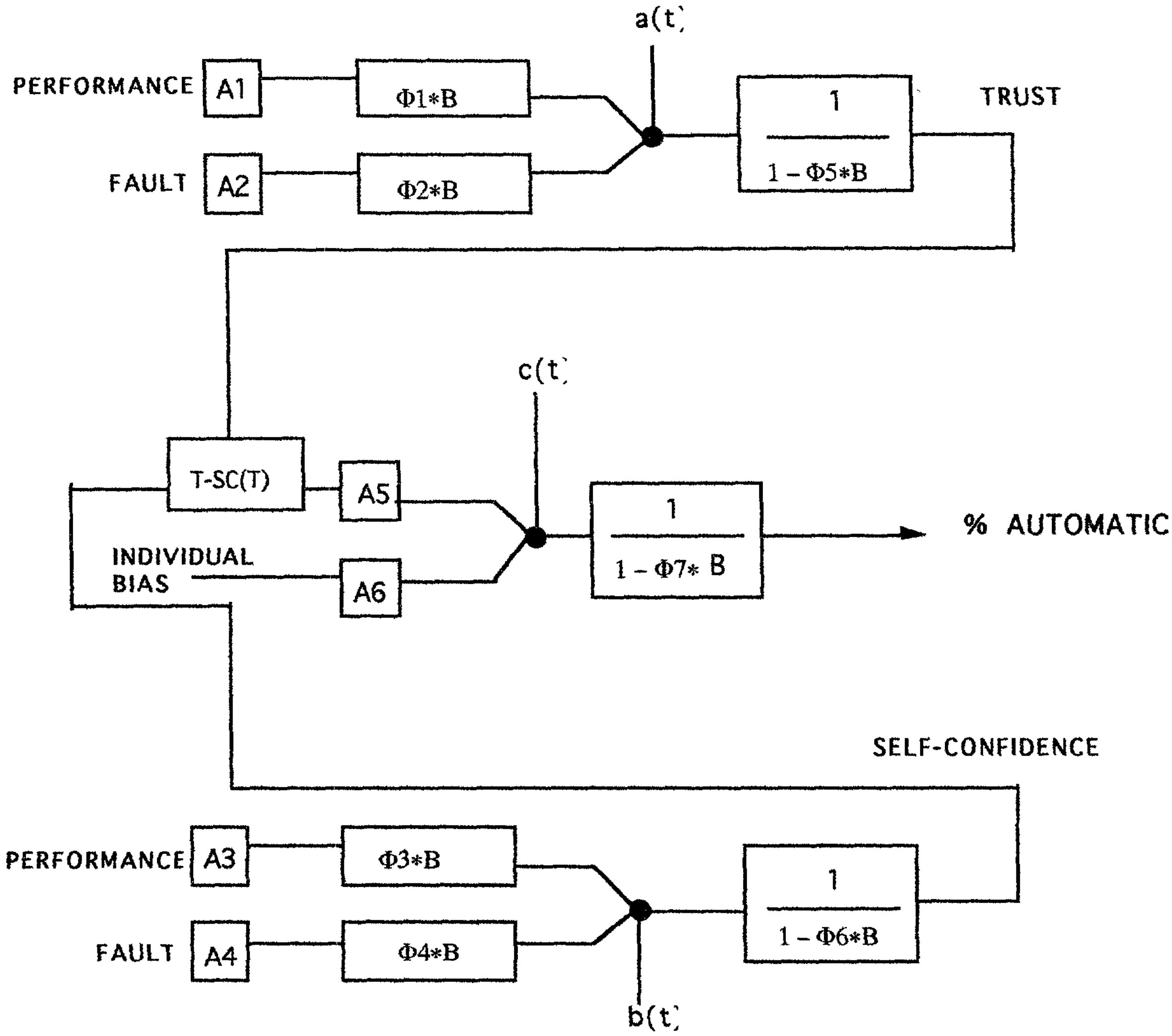
Neville Moray
LAMIH-PERCOTEC, Université de Valenciennes, 59304 France
G. Jamieson
University of Illinois at Urbana-Champaign, Illinois, USA

Abstract

Recent work has led to qualitative and quantitative models of operator intervention in supervisory control of continuous and discrete human machine systems> These models however apply to data pooled over operators and task conditions . Close examination of data show large individual differences in strategies of monitoring and control which are correlated with quality of performance. These strategies seem to be related to operators' mental models of the plant which they are controlling. We propose a new approach to identifying those aspects of a system which are likely to be incorporated into mental models of novice operators, based on the work of Conant, and describe some preliminary results of applying this method.

Introduction

A fundamental characteristic of supervisory control of industrial systems is the decision by operators to intervene in the operation of an automated system in order to exercise manual control of the process. In recent years work from several laboratories (Lee and Moray, Muir, Parasuraman, Riley,) has led to quantitative and qualitative models of operator intervention. Typical of these models are those developed by Lee (op. cit.) in which performance characteristics (such as the efficiency of production) and task characteristics (such as the magnitude and frequency of faults) were found to affect the trust and self-confidence of operators, which in turn affected the probability that they would switch between manual and automatic control. If we speculate slightly beyond the data analysis performed by Lee we arrive at the model shown in Figure 1 which summarises the characteristics of intervention in continuous process control (a simulated pasteurisation plant).



FLOW CHART MODEL OF OPERATOR INTERVENTION IN SUPERVISORY CONTROL OF CONTINUOUS PROCESSES (PASTEURISATION PLANT)

(AFTER LEE AND MORAY)

Figure 1

This model can be also represented as time series equations allowing us to predict the proportion of time spent in each mode:

Trust(t)

$$= \Phi5(\text{Trust}(t-1) + A1\text{Performance}(t) + A1\Phi1(\text{performance}(t-1) + A2(\text{Fault}(t) + A2\Phi2\text{Fault}(t-1)+a(t) \quad (1)$$

SC(t)

$$= \Phi6(\text{SC}(t-1) + A1\text{Performance}(t) + A1\Phi3(\text{performance}(t-1) + A2(\text{Fault}(t) + A2\Phi4\text{Fault}(t-1)+b(t) \quad (2)$$

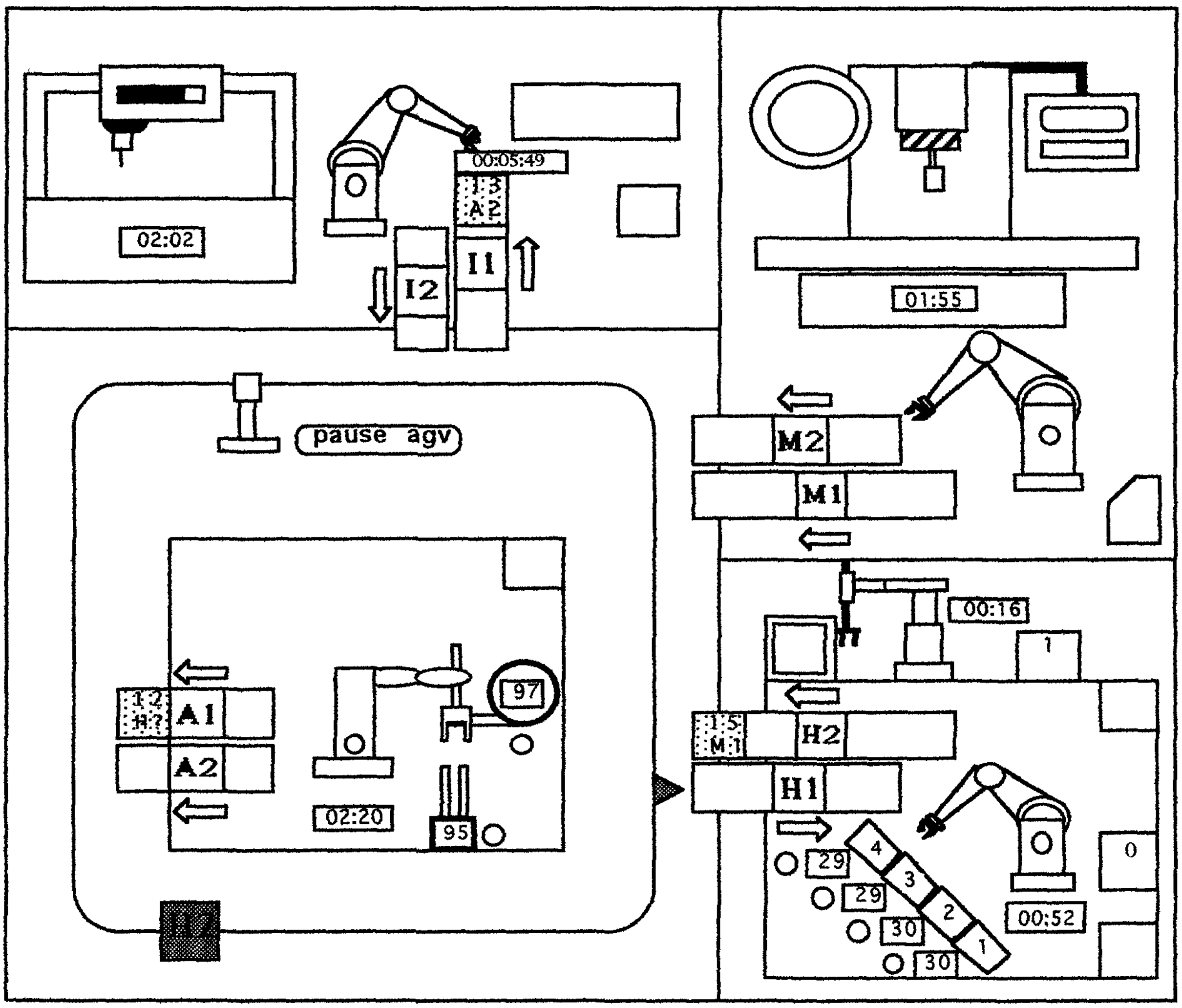
%Auto

$$= \Phi7\%\text{Auto}(t-1) + A5*((\text{T-SC})(t)) + A6*(\text{Individual Bias}) + c(t) \quad (3)$$

which can be combined to give:

$$\%Auto = \Phi7\%Auto(t-1) + A5*((\Phi5(Trust(t-1) + A1Performance(t) + A1\Phi1Performance(t-1) + A2Fault(t) + A2\Phi2Fault(t-1)+a(t) - \Phi6SC(t-1) + A1Performance(t) + A1\Phi3Performance(t-1) + A2Fault(t) + A2\Phi4Fault(t-1)+b(t)) + A6*(Individual\ Bias) + c(t) \quad (4)$$

For an interpretation of the coefficients see the original papers by Lee and Moray.



UIUC-CIM: A SIMULATED DISCRETE MANUFACTURING PLANT USED BY HISKES (1994)

Similar results were reported by Hiskes (1994) using a simulation of a computer integrated manufacturing discrete process, although it was not possible to develop a predictive model at the level of detail reached by Lee because of the nature of the process. His system is shown in Figure 2.

Among Hiske's results were the following:

1. Effect of faults on intervention:
   14 out of 14 operators used automatic control less on the second day, (mean decline = 16.5%,) when faults occurred. ($P < 0.001$)

2. Effect of faults on trust:
   12/14 operators showed a decrease in trust on the second day when faults occurred. ($P \approx 0.01$)

3. Effect of faults on (T - SC):
   12/14 operators showed a decline on in the value of (T-SC) with no change in self-confidence.on the second day. ($P \approx 0.01$). Note that faults only occurred when automatic scheduling occurred, and hence there was nothing in the expeeriment to cause SC to vary. Once manual scheduling occurred there were almostn no faults seen.

Individual Differences in Strategy: the Role of Mental Models

The above results, while being very satisfactory as far as they go, rely on the pooling of data over many subjects and many trials or, where the notion of trial is not applicable, over many events during operations. However, if we examine more closely the details of moment to moment behaviour, it is apparent that there is great variability in how operators perform these tasks. (It should be emphasiseed that while many of the operators appear to reach asymptotic performance, others do not, and there is a need for research both into the acquisition of these skills and into the stability of strategies.)

In an unpublished pilot experiment for those of Hiskes, Moray, Phillips and Quaid found the results shown in Table 1.

What is noteworthy here is that when faults were rare or very rare, nonetheless some operators chose to use manual control of the discrete manufacturing process even though the scheduling algorithms were effective. In this experiment and those of Hiskes,, faults only occur during automatic operation, and hence operators can preempt the possibility of faults at the cost of considerable workload. This kind of strategy has appeared in some of our later work. It appears that some operators decide, once the system shows any faults at all, that is worth their while to undertake very considerable manual intervention, probably in order to prevent the possibility of faults occurring. This seems to be true even when the probability of a fault is very low indeed, less than one fault per production run of 35 minutes. But not all operators do this. There is great variation between individuals.

PALLETS SCHEDULED MANUALLY

P(fault):	33%	20%	14%	11%
SUBJECT				
1	4	2	2	0
2	1	0	0	0
3	2	0	0	1
4	3	1	2	0
5	3	0	0	0
6	0	0	0	1
7	0	1	0	1
8	2	1	0	1
TOTAL	15	5	4	4

TABLE 1

Table 2 shows related results found by Hiskes. In his experiment, the probability of faults might either be constantly high, constantly low, progressively more frequent or progressively less frequent. In all cases there were very many redundant scheduling activities performed manually. That is, even when there was no immediate evidence that faults were occurring, and even when the probability of faults was quite low or very low, some operators began to take charge of all the scheduling actions so as to prevent possible faults from occurring in the first place.

These kind of results suggest that the operators have a well formed mental model of what will happen when faults occur, and use that model for tactical or strategic decisions for improving, even if not optimising, the system performance. But again, there is a very great variation in individual performance, strategies and tactics.

Such individual differences were also found by Lee in his continuous process control task. Here it was very clear that these differences were related to differences in performance. In particular he analysed the monitoring strategies used by several individuals, including one who was very effective when the plant was normal, and was also effective when faults occurred, and another who was poor even with the normal plant and very poor when confronted by faults. graphical representations of the monitoring tactics used by these two operators are shown in Figure 3.

Table 2 Summary Statistics On Redundant Scheduling In Simulated Discrete Manufacturing

CONDITION	FAULTS SEEN	ReS
Overall		
High Mean Fault Rate	31	25.8
Low Mean Fault Rate	96	1.2
High-High		
High Mean Fault Rate	7	23.1
Low Mean Fault Rate	25	1.6

The Column marked ReS shows the number of times the operator rescheduled a pallet which had already been (apparently correctly) scheduled by the computer, in order to ensure that no errors would occur. In the condition where there were many faults which could occur (High-High) operators only saw 7, whereas where fewer faults occurred they saw more (25). This is because when faults occurred frequently the operators realised that they only occurred when scheduling was done automatically. Hence they scheduled many pallets (23.1) redundantly to the destination already chosen by the computer. Those who saw relatively few faults did not form a mental model of how they arose, and so did not reschedule so many pallets. Hence paradoxically fewer faults were actually seen in the condition where more occurred.

Lee found marked indivudal diferences in the pasteurizere environment. He recorded the sequence of demands for information as an index of monitoring strategies, and found that good operators tended to look at many more parts of the system than poor operators during fault management, although there was little difference during the operation of the normal system. Poor operators showed inappropriate cognitive tunnel vision. We assume that the reason for these differences is that the operators have different mental models of the causality which is present in the plant, since there is no constraint on the order in which monitoring observations are made, and it is up to the operators to interpret the task demands in such a way as to monitor the process as efficiently as possible. Given that all operators saw the same system, why should they acquire such different mental models, and can we proceed further to describe the nature of these models?

Identifying "Empirically Normative" Mental Models

If there are many alternatives that operators "choose" for their mental models, why do we so often talk about "the" mental model which operators should have of a system which they supervise? We have argued elsewhere (Moray, 1989) that we should think of a mental model as a homomorphic mapping of the properties of a physical system into a representation in the mind. To say this is to say that the contents of a mental model are a partial representation of the original physical system. What are the features which one might expect to be carried over from the physical system into the mental model? Does it make sense to talk of a "normative" mental model which the operator "should" possess?

There is an obvious sense in which one might define such a normative model. For example, in the Optimal Control Model (OCM) approach to human performance modelling to which Henk Stassen has contributed so much, one speaks of the need to have a well defined model of the system as part of the Kalman filter. And while there are often several possible such models depending on the choice of the variables to be measured, the notion of the normative complete model is clear. If operators are to be optimal controllers, we expect them to possess one of these models. In particular, the dimensionality of such models is well defined - they require the minimal degrees of freedom which are necessary and sufficient to represent the dynamics of the system.

But when we observe operators monitoring and controlling very large systems, is it reasonable or even possible to expect them to have such complete mental models? There are very tight constraints on the rate at which they can sample state variables, due to the limitations of eye-movements and the time required to access computer displays, and dynamic short term working memory is very limited. If the bandwidth of the system is moderately high and the degrees of freedom large, then it is probably unreasonable to expect operators to develop complete, that is isomorphic, mappings into their mental models; and even the homomorphic models may be too large for them to construct. As noted elsewhere (Moray, 1976) that there is empirical evidence, for example from the work of Iosif (1968,1969a,b) that operators seem to regard a large system as composed of subsystems within which variables are so tightly coupled that it suffices to sample only some of them to determine the state of the subsystem, and that in effect the number of subsystems, not the number of variables, is what determines the structure of the mental model. If then many systems are too complex to be completely mapped into the mental model, can we determine what is the most likely subset of variables to be modelled, so that we can still speak of a "good" and a "poor" mental model, relative to such a subset? Such a subset of information about a system is not normative, because it is acknowledged to be imperfect. But one might speak of the "empirically normative" model, and ENM, meaning that given the limitations of human information processing, this is what one would expect a good operator to be able to discover, even if, from a truly normative viewpoint such as OCM, it is acknowledged to be suboptimal.

We suggest that the work of Conant (1976) provides a way to approach this problem. Conant has shown, following the work of Ashby (1956), how one can use Information Theory to identify the natural decomposition of a complex system. By considering the flow of information between variables one can identify subsystems, defined as sets of variables between which there is tight coupling, while there is loose coupling between one subsystem and another. The tight coupling means that the value of one variable can stand, to some extent, proxy for the values of other variables within the subset, so as effectively to reduce the dimensionality of the system. this is of course what Iosif's operators seem to have discovered. One might proceed as follows. Take a record of the values of all the displayed variables in a system at a sampling rate sufficient to satisfy the sampling theorem, and perform a Conant analysis on the resulting history of the system. The result should be a "best empirical decomposition" of the system as a whole. Essentially this amounts to performing cross correlation among all the variables, and Conant's method allows one to compute cross-correlation and auto-correlation *functions*, not just point estimates. Moreover, one can make use of qualitative variables (such as colours) as well as quantitative variables. Given our knowledge of the rate at which operators can switch among displays, the size of working memory, etc., we can estimate which relationships are likely to be detectable by an operator. For example, even if there is a strong coupling between a change in one variable, perhaps under my control, and another variable, but with a lag of, say, several minutes, it will be extremely difficult for me to detect it.

Because of the ability of Conant's method to handle qualitative data, it is practical to use it to analyse systems such as those used by Hiskes. The analysis and modelling of discrete manufacturing systems is extremely difficult, since the causality is unlike that of continuous systems. For example, an intermittent fault in a discrete manufacturing system, such as the failure of an automated guided vehicle to stop at the correct place for some but not all pallets, is only visible from time to time. Its effect does not propagate through the system in the same way that a change in pump rate or heat supply propagate continuously in a petrochemical plant. Nonetheless, using variables such as the name of the p[lace where a product is to found at a particular time, Conant's method can in principle analyse the structure of such a system. Furthermore, in principle, the effects of the coupling of the operator to the system when he or she intervenes can also be detected.

One might then arrive at the representation of system causality which is most likely to be detected (as correlations among the values of displayed variables), and hence most likely to be mapped from the physical system into the mental model - the "empirically normative" mental model. in so far as operators do not even manage to acquire such a mental model, they are falling short even of what we can reasonably expect from them psychologically. Obviously, the discovery of such an ENM would be valuable for many purposes, such as guiding display design, assessing training, etc..

First Steps Towards ENM Methodology

In recent months we have been working with Conant on the first application of this approach. We have used the PASTEURISER simulation to trace the pattern of operator monitoring and intervention, and have tried to force operators' attention onto different parts of the system by the appropriate use of various faults. In addition we have used Conant's method to analyse the ENM of the PASTEURISER plant. Our hypothesis is that the best and most effective operator mental models approximate to the ENM of the system revealed by Conant's analysis, and that operators whose sampling pattern seems to be consonant with the aim of pasteuriser will be the most effective operators. If this is successful, this approach may offer new insights into the detection and analysis of mental models of supervisory control of complex systems. In effect, such an approach allows us to take seriously the notion of the "human-machine *system*" since Conant's method allows us to see how the structure of the total system seems to change as a result of the role played by the operator. The operator re-organises the system when he or she becomes coupled to it, and one must then assume that the mental model is not a model of the plant, but of the system of which the operator is a dynamic component. The operators in a sense model themselves, not just the plant. Examples of such an ENM for two operators will be shown in in the presentation, in each case both when the plant is normal and when a fault has developed. The "natural decomposition" of the system is different for the two normal cases although the physical plant is the same in each. The way in which the operator is coupled to the plant through observation and intervention however makes the two *systems* different. This is even more apparent in the natural decomposition of the faulty plants, where the responses of the two operators clearly differ, and as a result one can see that they are attempting to regulate the fault in quite different ways. We are currently developing this method further in collaboration with Conant.

Acknowledgements

Our thanks are due to Professor Roger Conant for his assistance in adapting his method to our analysis, and for the generous giving of his time and effort to run the analyses.

Bibliography

Ashby W.R., 1956. *An introduction to cybernetics.* London: Chapman and Hall

Conant, R.C. 1976. Laws of information that govern systems, *IEEE Transactions on Systems, Science and Cybernetics, SMC-6,* 240-255.

Hiskes, D. 1994. Unpublished M.Sc. Thesis, Department of Mechanical and Industrial Engineering, University of Illinois at Urbana-Champaign; Engineering Psychology Research Laboratory.

Iosif, G. 1968. La stratégie dans la surveillance des tableaux de commande. I. Quelques facteurs déterminants de caractère objectif. *Revue Roumanien de Science Social-Psychologique, 12,* 147-161.

Iosif, G. 1969a. La stratégie dans la surveillance des tableaux de commande. I. Quelques facteurs déterminants de caractère subjectif. *Revue Roumanien de Science Social-Psychologique, 13,* 29-41.

Iosif, G. 1969b. Influence de la correlation fonctionelle sur parametres technologiques. *Revue Roumanien de Science Social-Psychologique, 13,* 105-110.

Lee, J.D. 1991. *Trust, self-confidence and operators' adaptation to automation.* Unpublished Ph.D. Thesis, University of Illinois at Urbana-Champaign. Engineering Psychology Research Laboratory Technical Report EPRL-92-01.

Lee, J.D. and Moray, N. 1992. Trust, control strategies and allocation of function in human-machine systems. *Ergonomics, 35,* 1243-1270.

Lee, JD and Moray, N. 1994 Trust, self confidence and operators' adaptation to automation. International Journal of Man-Machine studies. in press.

Moray, N., "Attention, Control, and Sampling Behavior," *Monitoring Behavior and Supervisory Control,"* (T.B. Sheridan and G. Johannsen, eds.) Plenum Press, 221-245, 1976.

Moray, N. 1989. A lattice theory approach to the structure of mental models. *Philosophical Transactions of the Royal Society of London, series B, 327,* 447-593.

Session 5

Decision Support Systems I

Chairman: E.A. Averbukh, Germany

Although the security class of a network is defined rather precisely, a lot of ambiguity exists in determining it, because the net state cannot be determined exactly, the severity of the violations is not defined uniquely, the effects of a violation are open to different interpretations, and the goals and the constraints of the system are usually expressed using linguistic terms. Hence, a fuzzy logic based system presents advantages, as it can deal with this type of uncertainty, while keeping the computational load low. This is important as the extensive simulations of the network take a long time and require powerful, usually specialized hardware.

This paper describes a fuzzy logic rule-based system (INSANE— INtelligent Security Analysis for power NEtworks) for determining the security class of a power network. The security analysis of a power network is primarily concerned with the "normal" and the "alert" states in which the network is operating within the specified limits. These classes are further divided into four security classes in this study, in order to assess the severity of the violations in case of a contingency. An index of "optimism" or risk awareness is defined which allows the company to adjust its risk awareness for particular situations.

The rest of this paper is organized as follows. Section II describes the notion of network security. The security classes considered in this paper and the way they are assessed at present are explained in this section. This section also contains a brief description of the 380kV Dutch transmission network which is considered in this study. Section III describes the fuzzy logic based INSANE system for determining the security class of the network. Various aspects of the system such as data acquisition, knowledge base, inference mechanism and the user interface are explained. Section IV reports some results obtained using INSANE. Section V presents the conclusions.

II. Security Assessment of Power Networks

Reliable operation of power transmission networks requires the assessment of the network security so that correct control actions may be taken for maintaining continuous supply of power. In addition to the reliability, other constraints must also be considered such as minimizing the operation costs. For this purpose, the operators first assess the security of the network. In this context, security assessment means that the system state must be classified into one of the security classes that are defined in literature.

A. Security classes

In literature, the network state is usually divided into three classes [3]:

1. the normal state,
2. the emergency state,
3. the restorative state.

The distribution network is said to be in the normal state when the power system is operating within all limits, all the potential users are supplied with power and the system remains within limits after an occurrence of a single contingency. The network is said to be in the emergency state when one or more violations have taken place. The network is in the restorative state if the network is operating within the limits but not all potential users can be supplied with power due to a violation beforehand. For purposes of steady state security, the normal state is augmented with the "alert state" which is the state of a network operating within limits, but one or more violations will occur should a single contingency take place. From a security analysis point of view, the interesting security classes are the normal state and the alert state, since the control actions that are taken when the network is within these classes should prevent the occurrence of the emergency state. It is usually assumed that when the system is operating in the alert state, the operator should try to bring the system into the normal state. In practice, however, the operators evaluate the severity of the violations in case of a contingency and adjust their actions accordingly. It is therefore possible that the system is within the alert state but the operators do not take action because the alert state is not very severe or because it is expected that the network will leave the alert state within a short period by itself. In order to model the operators' concept of the severity of the violations, the normal state is divided in this study into "Very Secure State (VSS)" and "Normal Secure State (NSS)", while the alert state is divided into "Slightly Insecure State (SIS)" and "Very Insecure State (VIS)". The security of the network decreases from VSS to VIS. The fuzzy rule-based system INSANE classifies the network security state into these four classes.

B. The Dutch 380kV Network

During this study, the Dutch 380kV system has been studied. Fig.1 depicts a simplified diagram of this system. It consists of a number of nodes connected by lines in a ring structure. The potential users are provided with energy through the 220kV and 150kV subnetworks that are connected to the 380kV network by

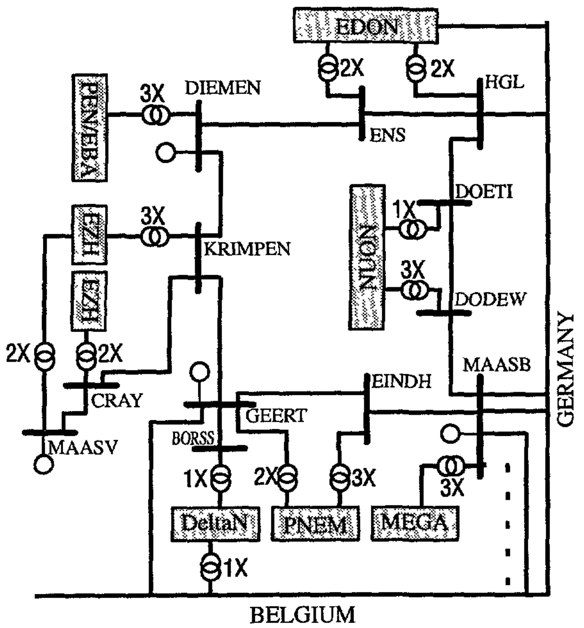


Fig. 1. Simplified diagram of the 380kV Dutch transmission system.

transformers. In Fig.1 the names of the nodes denote various towns, while the rectangular blocks denote the 220kV and 150kV subnetworks of the regional energy suppliers. These low voltage terminals also belong to the system under security analysis. The system is connected to the power systems of Germany and Belgium by a number of international tie-lines. Energy can be imported through these lines or generated in one of the main generators connected to the system.

C. Security assessment

At the control centre of the Dutch power network, the system is continuously supervised by system operators. These operators have access to the data about the system through a SCADA (Supervisory Control And Data Acquisition) system. About every 15 minutes an on–line security assessment program is run. This program uses a power flow algorithm for solving the system equations under given operating conditions and all possible single contingencies (the so–called $n-1$ criterion). The worst violations that can occur in this analysis are presented to the operators, who must then interpret the results for assessing the security class of the network. The data presented to the operators shows which violations occur as a result of which contingency together with some coded information about the location and the size of the violation. For assessing the network security, the operators seek answers to the following questions:

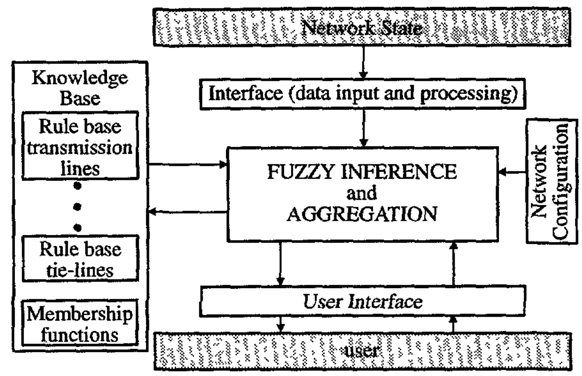


Fig. 2. General structure of INSANE.

- Where are the weak points in the network?
- Do the violations have a local character only?
- How should various violations be classified?
- What is the degree of security?
- How large are the security margins?

The present security analysis programs do not answer these questions directly. Instead, they present only numerical information. Thus, a system that (partly) answers some of these questions is desirable for supporting the decision of the operators and increasing the quality of the system operation.

III. Fuzzy Security Assessment

This section describes *INSANE*, a *fuzzy rule-based* system for determining the security class of a power transmission network with a ring structure. Fig.2 shows the general structure of INSANE. The heart of the system is the fuzzy inference and aggregation mechanism. The actual information about the network state (percentage of loads and voltages) is fed in from the EMS. Information about the network configuration is provided externally. The knowledge about the physical principles and the heuristics is stored in various rule bases. Each rule base contains rules for determining the security class of the system, based on the information from a particular type of components. The overall security class is determined by combining the fuzzy outputs of each rule base. The output of the fuzzy inference system is a fuzzy security class. The result is defuzzified by using a method which allows the selection of different modes of operation corresponding to various degrees of caution. A risk index is introduced for quantifying the risk awareness of the user. Communication with the operators is achieved via a graphical user interface. The following paragraphs explain in more detail the assumptions made for the design of the fuzzy security analysis program and the individual components of the system.

A. Assumptions

INSANE has been designed for networks with a ring structure such as the Dutch 380kV transmission network. Many of the assumptions are related to this particular structure. These assumptions are summarized below.

- The network that is considered has a ring structure.
- The occurrence of a contingency does not change the topology of the network. The ring structure is protected under all contingencies.
- The fuzzy rules describe the *approximate* behaviour of the network. This approximate behaviour is sufficient for the purpose of security analysis.
- Impedence of the parallel branch of the ring is very large compared to the impedence of the parallel components. In practice, this implies that when a contingency occurs, the load of the malfunctioning component can effectively be taken over completely by the parallel components. For example, when one of the two parallel circuits, each carrying 15% of their nominal load is disabled, the other circuit starts transmitting the load of the broken circuit, i.e. it starts transmitting 30% of its nominal load. The other components in the network are not affected.
- Many generators are directly feeding the network so that the outage of a generator has negligible effect on the power that is being transmitted through the network. The outage of a generator influences mainly the voltage drops.
- Only a single contingency can occur at a time (the so called $n-1$ criterion).
- Only the steady state violations are considered for the purpose of security analysis.

Experts from the field of electrical power systems have confirmed that these assumptions are valid during the normal operation of the 380kV Dutch transmission network.

B. Data acquisition

Measurements about the state of the transmission network are available from a SCADA system. These measurements indicate the power loads through various components (circuits, transformers etc.) and the voltages at various stations. INSANE does not use any other related data about the network such as the setpoints of the generators and the positions of the power switches. In principle, a direct connection between the SCADA system and INSANE can be made. However, such a connection has not been made in this study. Instead, the information about the network state is presented to INSANE in a file in which data is written according to certain conventions.

In addition to the actual loads and voltages in the network, INSANE uses information about the configuration of the network, provided to it externally. This information is necessary, amongst others, for localizing the violations that may occur. This information is also used for determining which rule bases should be used with which components in the network.

C. Knowledge base

INSANE determines the security class of the network, based on a contingency that may occur at one component of the network. The components for which a contingency may occur include the circuits connecting the nodes, the transformers connecting the network to the subnetworks, the tie-lines and the generators that are connected directly to the 380kV network. Consequently, knowledge must be available to the system about how each of these components contribute to the overall security of the network. This knowledge is provided in a number of rule bases, one for each type of component. Each rule base describes how a particular type of component affects the network security for different degrees of loads and voltage drops. The rules are given by experts from the field of electrical power systems based on their knowledge of the network, physical principles and experience obtained through simulations of the network.

A. Transmission lines

Transmission lines connect two nodes of the ring. A transmission line may consist of two circuits (double-circuit line) or three circuits (triple-circuit line). Based on the experience about the network, the degree to which a line is loaded is described by four linguistic labels which are represented by membership functions:

1. slightly loaded (SL),
2. normally loaded (NL),
3. slightly overloaded (SO),
4. heavily overloaded (HO).

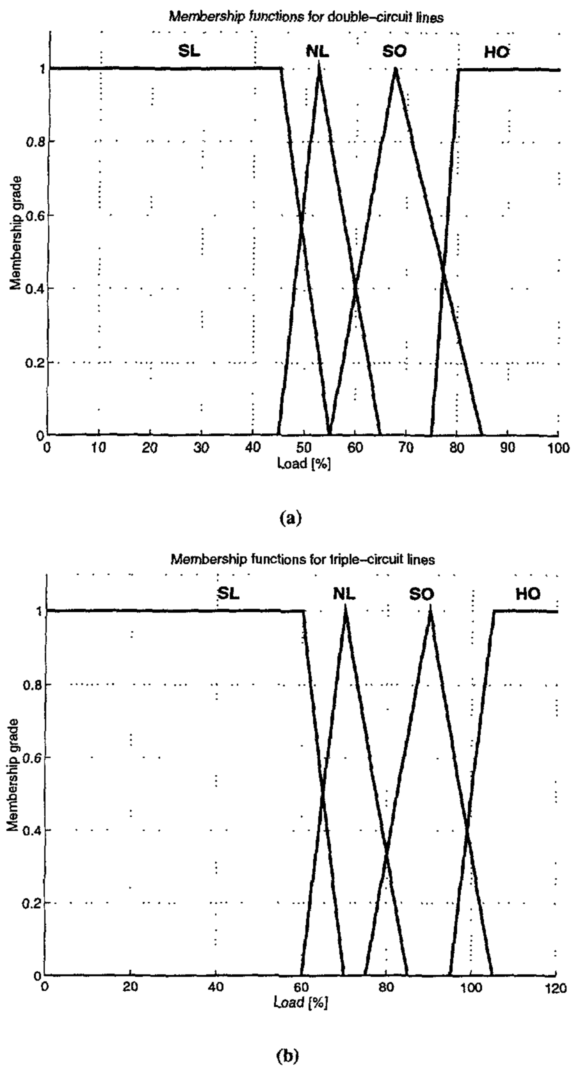


Fig. 3. Membership functions for the linguistic labels SL, NL, SO and HO defined for (a) double-circuit lines and (b) triple-circuit lines.

It is assumed that the circuits belonging to the same line transmit equal amounts of power. Then, fuzzy rules may be provided which map the degree of load of the transmission line to the security class of the network. The rules have the following form:

If the current load is SL then security class is VSS

The same rules can be used both for double-circuit lines and triple-circuit lines. However, the definition of the linguistic variables SL, NL, SO and HO are different for the two cases because a circuit in a double-circuit line has to transmit more power in case of a contingency, than a circuit in a triple-circuit line. Fig.3 shows the membership functions defined for double- and triple-circuit lines.

B. Transformers

The transformers connect the subnetworks to the main network. It is seen, from Fig.1, that a subnetwork is connected in two different ways to the ring network:

1. By a single cluster of transformer, such as the connection of PEN/EBA. The transformers in this type of clusters will be called 'type A' transformers.
2. By a double cluster of transformers, such as the connection of EDON. The transformers in this type of clusters will be called 'type B' transformers. The two clusters of transformers connecting a subnetwork to the ring are named as $T1$ and $T2$ respectively.

The same rules that apply to transmission lines also apply to 'type A' transformers. 'Type B' transformers are influenced by the transformers in their companion clusters and thus the rule base must consider the effects of the transformers in $T1$ and $T2$ simultaneously. The rules now look like:

If the transformer in $T1$ is SL
and the transformer in $T2$ is SL
then security class is VSS

The membership functions SL, NL, SO and VO are defined in a similar way as with the transmission lines.

C. Generators

A contingency at a generator can result in a voltage drop of more than 10% of the nominal voltage. This has consequences for the security class of the network. It is assumed that there are enough generators feeding the net so that the transmitted power is not changed significantly. The voltage drop is also divided into four classes of increasing severity and represented by membership functions:

1. very small (VS),
2. small (SM),
3. big (BI),
4. very big (VB).

A rule base now maps each linguistic label for the voltage drop to the corresponding security class. The rules are again of the following form:

If voltage drop is VS, then security class is VSS

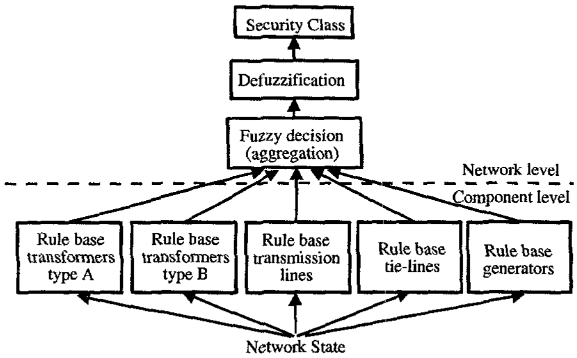


Fig. 4. Determination of the security class of the network. The information about the network state is first combined at the component level and then aggregated and defuzzified for obtaining the final classification.

D. Tie-lines

The tie-lines with the neighbouring countries are typically used for importing electricity. When a contingency occurs at a tie-line, it is known from experience that the power that the line has been transmitting is rescheduled over other tie-lines. A rule base can now be defined, based on the experience with the network, for performing this rescheduling. This rule base can be implemented as a table which indicates how much the power transmitted through a particular tie-line increases approximately as another falls out. The rules for the transmission lines may then be applied to the tie-lines as well. For analyzing the voltage drops because of a contingency at a tie-line, the tie-lines are considered as generators.

D. Inference mechanism

The inference mechanism combines the information about the current network state and the rules for security analysis in order to arrive at a classification of the security of the network. The classification is done first at the component level and then at the network level. Fig.4 shows how INSANE arrives at the final classification of the network security. The inference at the component level determines the security class of the network when considering only one type of component in the network. Each rule base thus determines a seperate outcome for the security class. At this level, Max–min composition with the Mamdani minimum operator [1] is used for inference. The output of the inference is a fuzzy set with four elements, where the membership value for each element denotes the membership of the network to a particular security class. The resulting fuzzy classifications (outputs of each rule base) must now be combined for arriving at an overall classification of the network security. In general, this aggregation can be done by using one of the decision functions available from the fuzzy decision theory. INSANE uses the maximum operator for this aggregation, because there is not a specific preference for one of the rule bases and this operator allows the analysis of a variety of security classifications including the extreme considerations such as the worst-case analysis.

After the aggregation of the intermediate results using the maximum operator, INSANE presents the membership degrees with which the network belongs to each security class to the human operators. From a classification point of view, however, the operators need to know which security class the network belongs to, so that they can take the required actions. Hence, the fuzzy classification should be defuzzified. INSANE uses a modified version of the generalized averaging operator [2] for the defuzzification purpose. The defuzzification is achieved by

$$g = \left\{ \frac{\sum_{j=1}^{4} w_j c_j x_j^s}{\sum_{j=1}^{4} w_j c_j} \right\}^{1/s} \qquad (1)$$

where $g \in [1,4]$ is a class indicator related to the security class of the network and $x_j \in \{1,2,3,4\}$ correspond to the security classes VSS, NSS, SIS and VIS respectively. c_j indicate the membership degree that the security of the network is classified as class x_j and w_j are weight factors related to the importance of the particular class for the defuzzification purposes. In general, the weight of the secure states are lower since more attention needs to be paid to the components that result in an insecure class. The security class of the network is found by determining the class that is closest to the calculated value of g and presented to the operators with the help of the graphical user interface.

The parameter $s \in \mathbb{R}$ in (1) can be interpreted as an index of risk awareness [6]. For negative values of s, the influence of the secure states on the defuzzification increases. In general, this corresponds to an increased awareness for cost–efficient operation, since one need not take costly actions when the network in the secure state. For positive values of s, the influence of the insecure states on the defuzzification increases. This corresponds to an increased awareness for risk aversion and very secure operation. Therefore, by changing the value of s, a particular strategy such as cost–effective operation or risk aversion may be implemented, depending on a production company's goals and priorities.

When the operation is not based on complete risk-aversion, some risk is being tolerated as the network

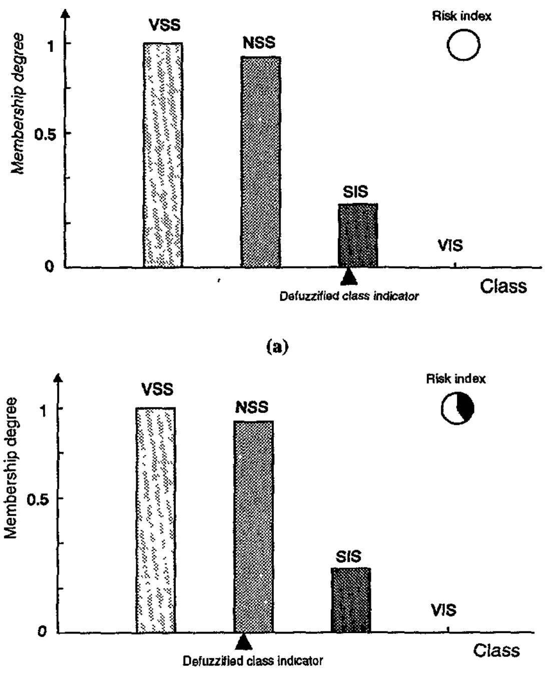


Fig. 5. Fuzzy security classification with (a) extreme risk awareness and (b) reduced risk awareness.

may be classified unjustly as more secure. It is important to have an indication of the risk level so that possible consequences and costs can be evaluted. For this purpose, a risk index is defined as follows

$$\text{risk index} = \frac{\sum_{\text{class}+1}^{4} c_j p_j}{\sum_{j=1}^{4} c_j p_j} \quad \text{if class} < \text{VIS} \quad (2)$$

$$\text{risk index} = 0 \quad \text{if class} = \text{VIS}, \quad (3)$$

where p_j are again weight factors that indicate the importance of a particular class for the calculation of the risk index. The less secure classes contribute more towards an increased value of the risk index.

Fig.5(a) shows the security class for a particular state of the network when the operation is completely risk aware. The final classification corresponds to SIS. There are no components in the VIS and the risk index is zero. Fig.5(b) shows the same network in the same state when the risk awareness of the operation is reduced. The final security class is now NSS and some risk is being taken as the risk index indicates.

E. User interface

A prototype of INSANE has been written in MATLAB, a well-known software package for scientific computation. An easy to use program with a suitable user interface is targeted, keeping in mind that INSANE is a decision making support system for the operators. MATLAB (version 4.2c) offers a reasonable set of tools for developing graphical user interfaces with sliders, push-buttons, menus and other standard controls. Making a full use of these features, INSANE has a mouse-driven graphical user interface (GUI) that consists of the following items (see also Fig.6).

Control panel in the upper left corner of the screen contains the menus and push buttons for controlling the program functions. It also displays information about the "case" (data file) loaded and the security classification.

Two *text information screens* below the control panel present standard information about the component loads to the operator. Only the components belonging to the classified security class are shown.

In the upper right part of the screen, the results of the *fuzzy inference are displayed in the class membership window* in the form of membership degrees of the individual classes. The defuzzified value (shown as the class pointer) actually determines (after rounding off to the nearest class) which class is displayed in the control panel. The dark (on screen red) area of the pie-chart in the upper right corner of the window indicates what risk is taken by accepting the INSANE's decision (based on the risk strategy setting in the control panel). Additional information can be asked from INSANE by clicking the mouse button on the bars corresponding to the individual classes. Then all the components belonging to that particular class are displayed in the text information windows.

The *bar chart* in the lower right part of the screen presents the loads of the transformers and lines in a graphical form.

IV. RESULTS

The fuzzy security assessment program INSANE has been tested for some network states and configurations which were designated as being typical or relevant by the experts. Even though the fuzzy rules give an approximate description of the network and no detailed simulations are performed, it was found that the classification of INSANE agreed with the classification obtained from the experts, in many cases. Hence, INSANE is able to determine the security class of the network rather accurately. Several test cases were generated and the classification of INSANE was compared with the classification of the operators who supervise the network. In all these cases the security

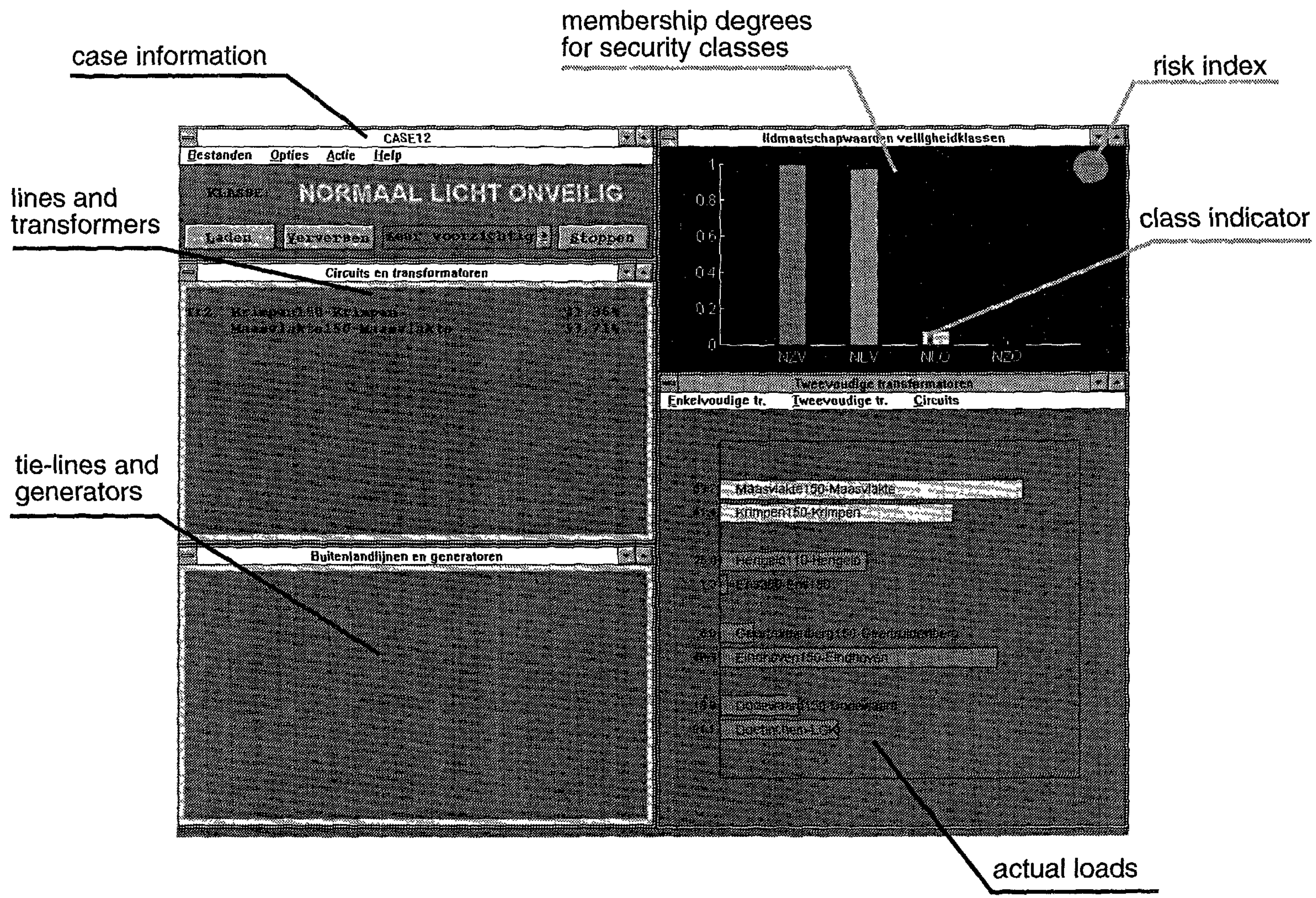


Fig. 6. User interface screen.

class determined by INSANE agreed with the security class determined by the operators. The operators have confirmed that the classification of INSANE coincides with their knowledge of the network. There were a few cases in which the security class determined by INSANE did not agree with the security class determined by the experts from the field of electrical power systems. However, the discrepancy between INSANE and the experts was found to be minimal in these cases.

V. Conclusions

A fuzzy rule-based system for the assessment of the security class of a power transmission network has been described in this paper. The system (called INSANE) uses the measurements of the network state, the network configuration, physical knowledge about the system and heuristics for determining the security class. The transmission system must have ring structure. The information about the network is utilized in a hierarchical manner, combining it first at the component level and then at the network level. The fuzzy logic system can take into account different strategies such as increased or reduced levels of risk awareness, which is related to the operation costs. A graphical user interface has been built for realizing effective interaction with the operator. The developed system has been tested on some real-life cases obtained from the 380kV Dutch transmission network. The operators have confirmed that the classification of the developed system corresponds to their knowledge of the network.

The division of the "normal" and the "alert" states into four different classes of increasing insecurity has proven to be useful in practice. In this way, the concept of the "severity of a violation" which the operators use in practice can be modelled. In addition to the determination of the network security, a decision support system for security analysis should also be able to suggest control actions for realizing the goals of the system. The extension of INSANE for suggesting proper control actions is the scope of future research.

Acknowledgements

This project has been done in cooperation with the Laboratory for Electrical Power Systems of the Delft University of Technology and has been sponsored by SEP and KEMA. The authors thank H. Braam for his contribution to the project and to the software development in particular.

References

[1] D. Driankov, H. Hellendoorn, and M. Reinfrank. *An Introduction to Fuzzy Control.* Springer Verlag, Berlin, 1993.

[2] U. Kaymak and H. R. van Nauta Lemke. A parametric generalized goal function for fuzzy decision making with unequally weighted objectives. In *Proceedings of the Second IEEE International Conference on Fuzzy Systems*, volume 2, pages 1156–1160, Mar. 1993.

[3] T. E. D. Liacco. The adaptive reliability control system. *IEEE Transactions on Power Apparatus ans Systems*, PAS–81(5), May 1967.

[4] T. A. Mikolinnas and B. F. Wollenberg. An advanced contingency selection algorithm. *IEEE Transactions on Power Apparatus ans Systems*, PAS–100(2), Feb. 1981.

[5] K. Tomsovic and J. Ling. A proposed fuzzy information approach to power system security. In *Proceedings of The Third Symposium on Expert Systems Application to Power Systems*, pages 427–432, Tokyo–Kobe,Japan, Apr. 1991.

[6] H. R. van Nauta Lemke, J. G. Dijkman, H. van Haeringen, and M. Pleeging. A characteristic optimism factor in fuzzy decision-making. In *Proc. IFAC Symp. on Fuzzy Information, Knowledge Representation and Decision Analysis*, pages 283–288, Marseille, France, 1983.

KNOWLEDGE-BASED MULTIMODEL AS A BASIS FOR THE DESIGN OF HUMAN-MACHINE SYSTEMS IN PROCESS CONTROL

S. Ali and J. Heuer

University of Kassel, Laboratory for Human-Machine Systems, D-34109 Kassel, Germany

Abstract: This paper presents an integrated design approach for the construction of Human-Machine Interfaces. This approach uses knowledge elicitation in the form of natural language as its design basis. Fuzzy techniques are used to translate the knowledge into different types of models which subsequently merge into a multimodel - based user Interface. Each model in itself fosters the operator's understanding of the process on specific skill levels. The design stages, fuzzy mechanisms and visual representations will be presented in detail.

Key Words: process control; human-machine interface; qualitative model; fuzzy logic

1. INTRODUCTION

Due to increasing product quality requirements and economic efficiency needs, technical processes are becoming more complex. This complexity is aggravated by the use of complex plant concatenations, difficult process control strategies and extensive automating functions. In order to meet the crucial demands of the technical process, the human-machine interface also becomes more intricate and complex. Additionally operators are placed under stress by their task of eliminating plant failures. They have to repair malfunctions quickly and efficiently, an activity that is not supported sufficiently by conventional user displays (Bainbridge, 1987). Therefore there is a need to develop novel user displays that are adapted to the requirements of operators and give inexperienced operators the opportunity to gain a coherent understanding of the technical process (Heuer et al., 1995). One approach for the improvement of this kind of human-machine interfaces consists in the participation of operators during the design process (Ali et al., 1993; 1994a).

This paper presents an integrated approach with new design methods for human-machine systems in process control. All models and examples mentioned are based on the simulation of a chemical destillation column (Gilles et al., 1990). This simulation represents a typical, highly complex system and is thus very suitable for demonstrating the methodology.

In conventional man/machine interfaces, the visualisation of dynamic processes is mostly realised by the use of topological representations (VDI/VDE 3695, 1986). This kind of user interfaces becomes very intricate when complex plants have to be visualised. A solution to this problem can be found in considering a set of different models that is suitable for representing the process on different levels of abstraction. This set forms the so-called „multimodel". A multimodel is acquired through task and process analysis and includes a qualitative states-causal model and a goals-means model. All these models are based on the experience and knowledge of operators and are located as objects in the knowledge-base of the human-machine system. Great emphasis is put on the integration of the different models within a multimodel to achieve a consistent visualisation of the dynamic process on the user display (multimodel - based user display). The knowledge of a multimodel is acquired in collaboration with experienced operators on the basis of natural language. It will be shown that the use of fuzzy logic is suited for the translation of the natural language procedures into the multimodels as well as for the visualisation of approximate (non-crisp) values.

2. APPLICATION FIELD

The example used here to demonstrate the development of the different models and their integration in the human-machine system, as well as their implementation in the design of user displays, is a process simulator of a distillation column. The mathematical kernel of the process simulator was developed at the University of Stuttgart. The process simulator runs on a *VAX station 4000* with the operating system *VMS*. The chemical process separates an ideal mixture of Benzol and Toluol. The distillation system contains several components such as a column, a reboiler at the bottom of the column, a condenser, reflux drum, pumps and control units managed by the operators. Fig. 2 presents a schematic flow diagram of the distillation column.

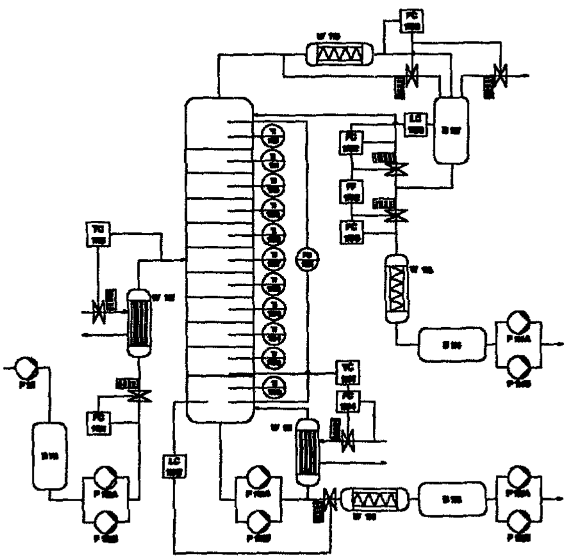


Fig. 1: Flow diagram of the distillation column

3. DIFFERENT MODELS OF THE TECHNICAL PROCESS

In the human-machine system, different models of the technical process are used to represent the process on different levels of abstraction. These models are allocated and integrated within the so called „multimodel". The multimodel is considered both in the design of the user display and in the development of the information management system of the human-machine system. A multimodel is acquired through task and process analysis and includes a qualitative states-causal model and a goals-means model. The qualitative states-causal model is based on the knowledge of qualitative states, qualitative transitions, and a qualitative causal model of the dynamic process. The goals-means model is based on the Multilevel Flow Model (MFM) theory; see Lind (1982; 1988) . Both these models are based on the experience and knowledge of operators and are located as objects in the knowledge-base of the human-machine system.

Great emphasis is put on the integration of the different models within a multimodel to achieve a consistent visualisation of the dynamic process on the user display. The different models mentioned are used as a basis for the design of human-machine interfaces (HMI), therefore the level of abstraction required for these models depends on that application, in this case the supervisory and control tasks (Shen et al., 1992). The qualitative modelling of physical systems is not an inaccurate or approximate inference; rather more the essential features and distinctions must be determined and combined (Struss, 1992).

3.1. Qualitative States-Causal Model

Experiences with user interfaces in process control have shown that the visualisation of process states can be substantial aid to the operator (Larsson, 1992). The operators need to know the current state of the process is and towards which state the process moves as a consequence of the operator's actions or plant malfunctions. Moreover, it is necessary to mediate for which reason the dynamic process is located in a particular state or why the process leaves this state. The combination of states, transitions and causal coherence makes it possible to present this information to the operator. Qualitative states and transitions are modelled by a *Qualitative States-Transitions Model*, causal coherence is described by *a Qualitative Causal Model*. The integration of both models results the so called *Qualitative States-Causal Model*.

Qualitative states-transitions models are acquired with the help of task and process analysis through which essential states and transitions of the dynamic process are identified. The calculation of qualitative states is based on measured process state variables and their tendencies, and describes the actual state of a dynamic process. On the other hand the calculation of qualitative transitions is based both on influence and state variables, and mediate mainly information about the next possible states.

Causal networks are designed on the basis of the knowledge and experience of operators, that may be obtained from a causal diagram analysis (Funke, 1992). The description of a technical system acquired from operators will be transformed into a causal network which contains qualitative process variables and their connections (Fishwick et al., 1992). These qualitative variables and change of rates are located as objects in the knowledge-base of the human-machine system. There are two kinds of qualitative variables within the causal model, *state-* and *input variables*. State or storage variables can be influenced only from input or other state variables. Input or influence variables can be changed by the operators.

Fig. 2 gives an overview of the qualitative causal model of the destillation column. It was possible to reduce the number of the process variables (about 120 process variables) to 12 qualitative variables suitable for the causal model. The causal model is used in the design of the human-machine interface, therefore great care is taken to select the variables of the qualitative causal model so that all components and process variables needed in the process control are properly represented.

The values of the objects *qualitative variables* and *qualitative connections* of the causal model as well as of the objects *qualitative states* and *qualitative transitions* are described with non-crisp values. A qualitative causal connection contains the functionality of a non-linear transfer function describing a change-of-rate of a state variable. The non-linearity is modelled with fuzzy rules. For example the connection c1 (see Fig. 2) which describes the influence of the heating *RB* on the bottom filling level *BL* can be modelled by the following rules:

if RB=off then c1=zero
if RB=weak and BL=empty then c1=-fast
if RB=weak and BL=little then c1=-middle
...

The modelling of states and transitions is based on fuzzy rules including qualitative variables and connections of the qualitative causal model. Fig. 3 shows a part of the qualitative states-transitions model. The value strength of process situation in a particular state or a transition visualised on the user interface are approximate (non-crisp) and are managed from an object with a local fuzzy-logic module located in the multimodel component of the management system. The fuzzy logic implementation of the states-causal model makes it possible that the process can be in more than one fuzzy state at the same time. This corresponds with the operator's perception of the process state, e.g. a process can be "almost stationary", and it is also shown in the display; see also Fig. 9. If, e.g., the dynamic process is located in the initial state *COLD* where the bottom temperature *BT* is less than 85 °C, the temperature in the middle of the column *MT* is less than 80 °C, and the head temperature *HT* is less than 77 °C. At the same time the state BOTTOM_HEATING is also active if the bottom temperature *BT* equals ca. 85 °C (see Fig. 5) and the bottom temperature tendency *dBT* is positive. The transition between the states *COLD* and *BOTTOM HEATING* is based on the previous state *COLD* and the operator activity *HEAT_BOTTOM* described by the change-of-rate *C7*. These states are described by the following rules:

if BT=less85 and MT=less80 and HT=less77 then state=COLD
if BT=ca85 and dBT=pos then state=bottom_heating
...
if state=COLD and C7=/zero then transition=COLD_BOTTOM_HEAT

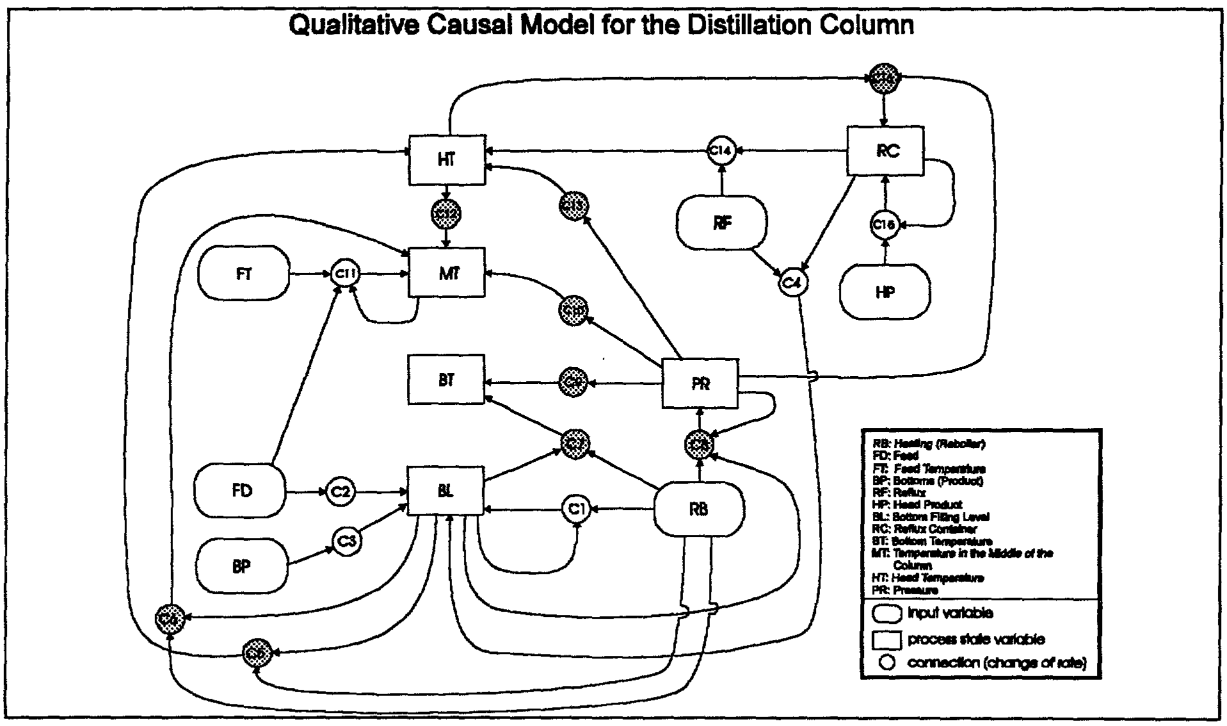


Fig. 2: The qualitative, causal network of the destillation column

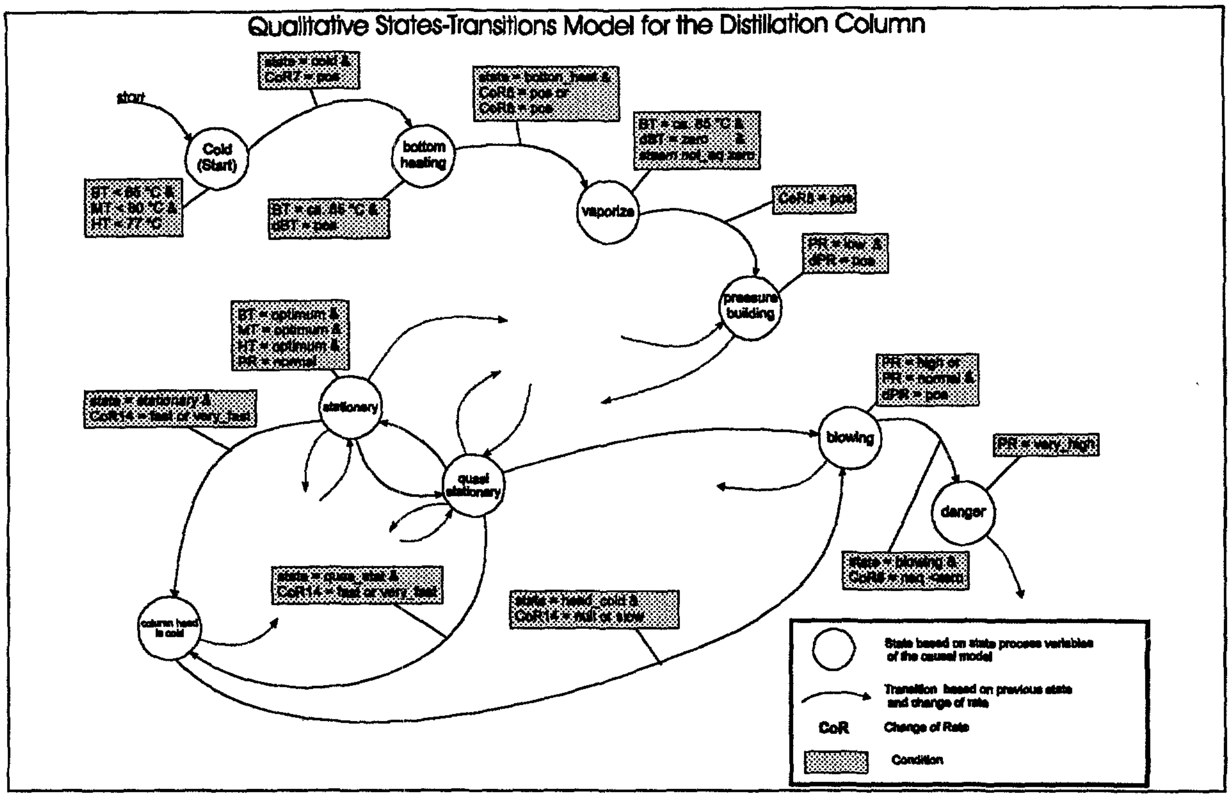


Fig. 3: Part of the states-transitions model of the destillation column

3.2. Goals-Means Model

Lind (1981; 1993) has suggested multilevel flow models for the presentation of process information at different levels of abstraction. Interfaces based on MFM stress the presentation of functional information rather than topological information. Six types of flow functions are needed to represent the mass, energy and information flows that govern the process. The two main ones are storage and transport processes; others are distribution, barrier, etc. For all types of processes, graphical flow model symbols have been suggested as elements of a graphical representation language for functional relationships. Extensions to this graphical representation language include also symbols for sensors and controllers (Johannsen, 1992). Such models are also suitable to provide support in fault diagnosis tasks. In Larsson (1992) and Sassen (1993) the MFM - method is implemented for fault diagnosis purposes in the application domains of bio-chemistry and nuclear power plants. The fault diagnosis is an essential sub-task of the supervisory control (Sheridan and Johannsen, 1976). Presentation of the information structure based on functional and goal-oriented aspects facilitates the problem solving tasks of the human users. In the Laboratory of Human-Machine Systems at the University of Kassel, the MFM - method is used to visualise the process information on user displays in goal- and functional-oriented way (Rauh, 1994).

The functional information may be shown on the display at three different levels of detail. Global information is presented by a goal tree that summarises which goals and sub-goals in the human-machine system are met. More detailed information is represented by flow structures, that present the flow functions that support one or a few goals in the tree. The highest level of detail, but on a local scale, is provided by a topological presentation of the components implementing the flow functions.

The fulfilment of the goals in the means-ends tree is again mediated by fuzzy logic. This means that a goal can be “half fulfilled”. The goal state is represented on the display by a small pie-chart, displayed in the (circular) goal icons; see also Fig. 8. For example the sub-goal *Bottom_Fill_Level_Goal* *„Bottom Filling Level BL is ca. 60 %“* may be described with the following fuzzy rules:

if BL=optimal ***then*** *Bottom_Fill_Level_Goal=fulfil*
if BL=/optimal ***then*** *Bottom_Fill_Level_Goal=not_fulfil*

4. TOOL FOR FLEXIBLE DESIGN OF MULTIMODEL-BASED USER DISPLAYS

Most parts of the human-machine system *INFO system (Information, Navigation, Failure handling and Organisation - Management System)* have been implemented during the project “Participative design of user displays in the process control” (Ali et al., 1994b). The main goal of this project is the involvement of the end users, in this case the operators, in the design process. Therefore, the *INFO* system follows the rapid prototyping philosophy and supports the iterative design of user displays. Fig. 4 shows the principle architecture of the *INFO* system. The main components of the *INFO* system are the graphical tool *Dynavis-X*, the *INFO* management system *(INFO-MS)* containing the process control management system *PC-MS*, the qualitative multimodel management system *QuMultModel-MS* and the Hypermedia system *HypMedSys*. All of these components are implemented in *C++* and run with the graphical tool on a *SUN SPARC* Workstation with the operating system *UNIX*. The graphical tool is responsible for the representation of the user display, the design and configuration of graphical objects and last but not least the interaction with operators. Whereas the *PC-MS* co-ordinates the tasks: the process data acquisition, the data processing, the management of local and global data as well as the actualisation of process data. Moreover, the *PC-MS* system logs the operator actions during experiments for evaluation purposes. The results of the protocol data are considered in subsequent design stages.

The *PC-MS* is implemented in a shell structure. The basic elements of the kernel are *module-* and *process variable* objects. The element organisation is managed with the help of dynamically linked lists. A coupling program connects the *PC-MS* with the process simulation (destillation column) running on a *VAX station 4000* with the operating system *VMS*. The data transfer between the simulation and the *PC-MS* based on the *DECnet* protocol.

The *INFO* HMI visualises the process in several representations, such as topology, causal nets or goal-oriented models, and retrieves the animation data from the *PC-MS*. The *PC-MS* gets the actualisation data either from the technical process, from animation objects (simple simulations located in *PC-MS*) or *QuMultModel-MS*. The structure of the database of the qualitative multimodel within the *QuMultModel-MS* is based on the qualitative states-causal model, process states, process transitions and goal-oriented hierarchies. The qualitative model contains objects such as process variables, connections, goals, and states. Every object contains different attributes such as a fuzzy variable, fuzzy sets, a local fuzzy rule base and fuzzy inference machine. The values of attributes are used to animate graphical objects on the user displays (multimodel - based user displays).

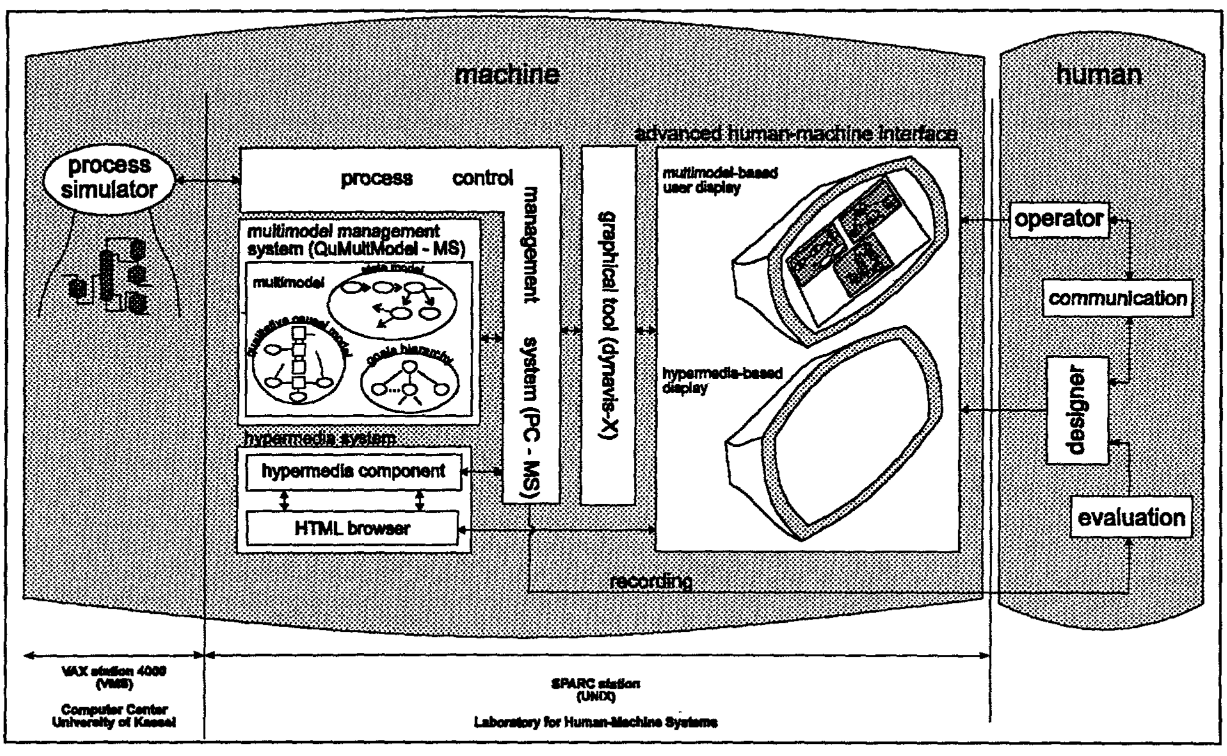


Fig. 4: Principle architecture of the *INFO* system

5. MODEL DESCRIPTION WITH FUZZY LOGIC

The operator's knowledge about the technical process is mostly available in natural language (Zimmermann, 1991). For the construction of qualitative models, this knowledge must be recorded and converted into a computer representation. Fuzzy logic is especially suited for this step. The linguistic values used by the operators to describe the properties of a variable x are combined into a fuzzy set A. For each term A_i links a linguistic variable, e.g. "large", with a membership function of base variable x.

Trapezoid, and triangular membership functions are used in the multimodel to describe the values of process control variables. Fig. 5 shows the membership function of a fuzzy term with a trapezoid shape.

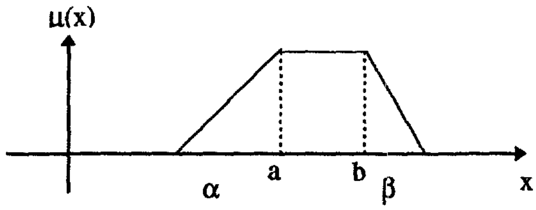


Fig. 5: membership function of a fuzzy term with a trapezoid shape

The input variables used for a fuzzy rule within the objects of the multimodel are qualitative variables and connections, change-of-rates of the qualitative causal model and tendency variables. All these input variables get their values from the dynamic process managed from the information management system. The output variables of the multimodel are change-of-rates, fulfilment - grade values for goals as well as strength values for states and transitions calculated from local fuzzy inference machines. The values of these variables are used to animate graphical objects on multimodel - based user displays. Fig. 6 shows two fuzzy sets for the input variable *BT „temperature in the column base"* and the output variable *fulfilment - grade*. Every qualitative input variable used is realised within the multimodel component of the information management system as an object-oriented class with a local fuzzy object, the name of the input process variable, a fuzzy set etc. An output variable is also realised as an object and contains a local fuzzy object, a fuzzy set, a name list of the input variables, the name of the output process control variable or variables visualised later on the multimodel - based user display, and last but not least a local fuzzy rule base with fuzzy inference machine. As an example, the syntax of the qualitative state variable (output variable) located in the initialisation file for the qualitative multimodel looks as the following:

<QualitativeModel>
.
.
<QualitativeVariable>
.
.
</QualitativeVariable>
.
.
<QualitativeState>
<Param>
QuStateName: *QuStateKol*
ProcessVarName: *QuModelModule.QuStateKol*
Min: *-100*
Max: *200*
Relation: *State = f{BT, MT, HT, dBT, dMT, dHT, PR, dPR};*
</Param>
<FuzzySet>
Subset: *{unknown, -100, -75, -100, -100 }*
Subset: *{danger, -100, -50, -75, -75 }*
Subset: *{blowing, -75, -25, -50, -50 }*
Subset: *{quasi_stat, -50, 0, -25, -25 }*
Subset: *{cold, -25, 25, 0, 0 }*
Subset: *{base_heat, 0, 50, 25, 25 }*
Subset: *{vaporise, 25, 75, 50, 50 }*
Subset: *{pressure_build, 50, 100, 75, 75 }*
Subset: *{stationary, 75, 125, 100, 100 }*
</FuzzySet>
<Rule>
Rule: *{BT=<85 & MT=<80 & HT=<77*
-> State=cold}
Rule: *{BT=ca85 & dBT=pos*
-> State=base_heat}
Rule: *{BT=ca85 & dBT=zero & steam=/zero*
-> State=vaporise}
Rule: *{PR=low & dPR=pos*
-> State=pressure_build}
Rule: *{PR=normal & dPR=pos*
-> State=blowing}
Rule: *{PR=high & dPR=pos*
-> State=blowing}
Rule: *{PR=very_high*
-> State=danger}
.
.
</Rule>
</QualitativeState>
</QualitativeModel>

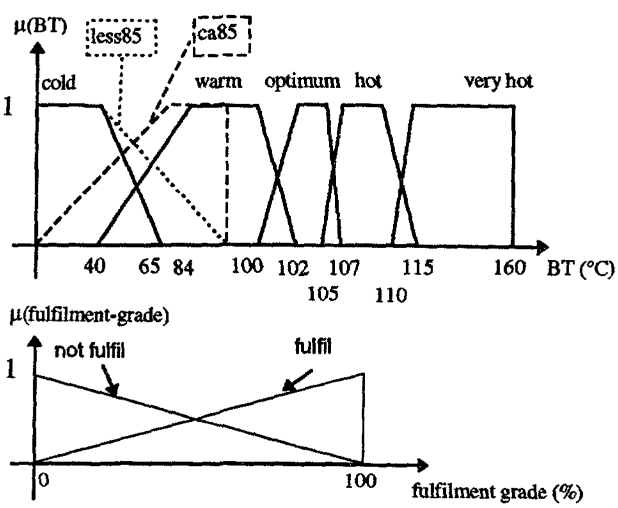


Fig. 6: Fuzzy Set for the input variable *BT „temperature in the column base"* and the output variable *fulfilment - grade*

The inference process for output variables of the classes *connections* and *fulfilment - grades* proceeds in four steps; after Kantrowitz, M. et al. (1995).

1. Fuzzification. The membership functions defined on the input variables are combined with the actual values that are measured or produced by model calculation to determine the degree of truth for each premise.

2. Inference. The truth value for the premise of each rule is computed, and applied to the conclusion part of each rule. This results in one fuzzy subset to be assigned to each output variable for each rule. The local inference machines currently used use the MIN inferencing. In MIN inferencing, the output membership function is clipped off at a height corresponding to the rule premise's computed degree of truth (fuzzy logic AND).

3. Defuzzification, which is used when it is useful to convert the fuzzy output set to a crisp number. There is more than one defuzzification method. Two of the more common techniques are the CENTROID and Given a certain process state, the PC-MS retrieves its data from the dynamic process. The data is stored in the appropriate data objects. The goal-object (or any other object inside the multimodel) retrieves a value from the PC-MS and uses this value as an input for the local fuzzy inference machine. This inference calculates a resulting value using the CENTROID method. This value is then used to update the corresponding variable object in the PC-MS. Finally, the value is used to visualise the degree of fulfilment (of the given goal) on the user display.

MAXIMUM methods. In the CENTROID method, which is used in this system, the crisp value of the output variable is computed by finding the variable value of the centre of gravity of the membership function for the fuzzy value. In the MAXIMUM method, one of the variable values at which the fuzzy subset has its maximum truth value is chosen as the crisp value for the output variable.

The inference process for output variables of the classes *states* and *transitions* contains only the steps 1 and 2, because on the multimodel - based user display sub-sets' values (sympathy vector) of a qualitative state or transition describing the different sub-states or sub-transitions will be visualised. However, for the output variables of the classes *connections* and *fulfilment* defuzzificated values are used for visualisation.

Fig. 7 illustrates the basic functionality and integration of the different components of the system. The example shows the calculation of the degree by which a certain process goal is met.

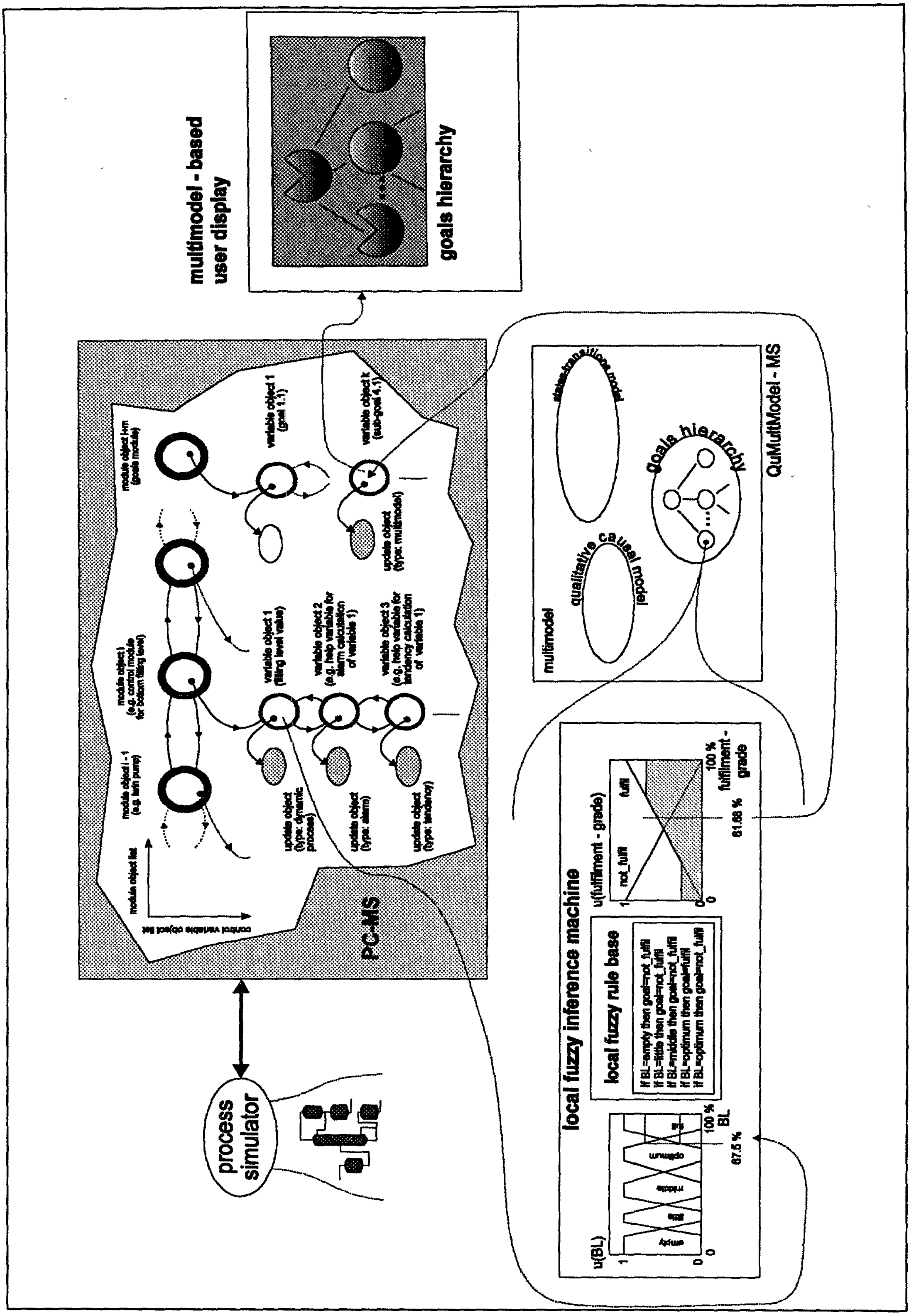


Fig. 7: Basic functionality and integration of the different components of the system

6. A MULTIMODEL-BASED USER DISPLAY

Fig. 8 presents the different visualisations of the multimodel-based display. Different representations of the technical process such as topology, causal coherence, goals-means hierarchy, states and transitions are visualised. On the lower left side of the screen four dynamically minimised windows for the different models are located. These windows cannot be covered by other displays, and thus the operators are always informed about the current state. In Fig. 8 the goals-means hierarchy display is activated. The minimised window of the goal hierarchy mediates whether the goals are fulfilled or not. In the detailed visualisation of the goal hierarchy a goal fulfilment is visualised by a pie-chart symbol. The calculation of every pie-chart value is based on fuzzy rules and is executed by a local fuzzy inference-machine. On the lower right side, the multilevel flow modelling representation of the highest hierarchical level is visualised. The visual momentum (Woods, 1984) between the different windows, with different graphical representations on different abstraction levels and different degrees of detail, is guaranteed by text labels and also dependent on the skills and knowledge levels of the users.

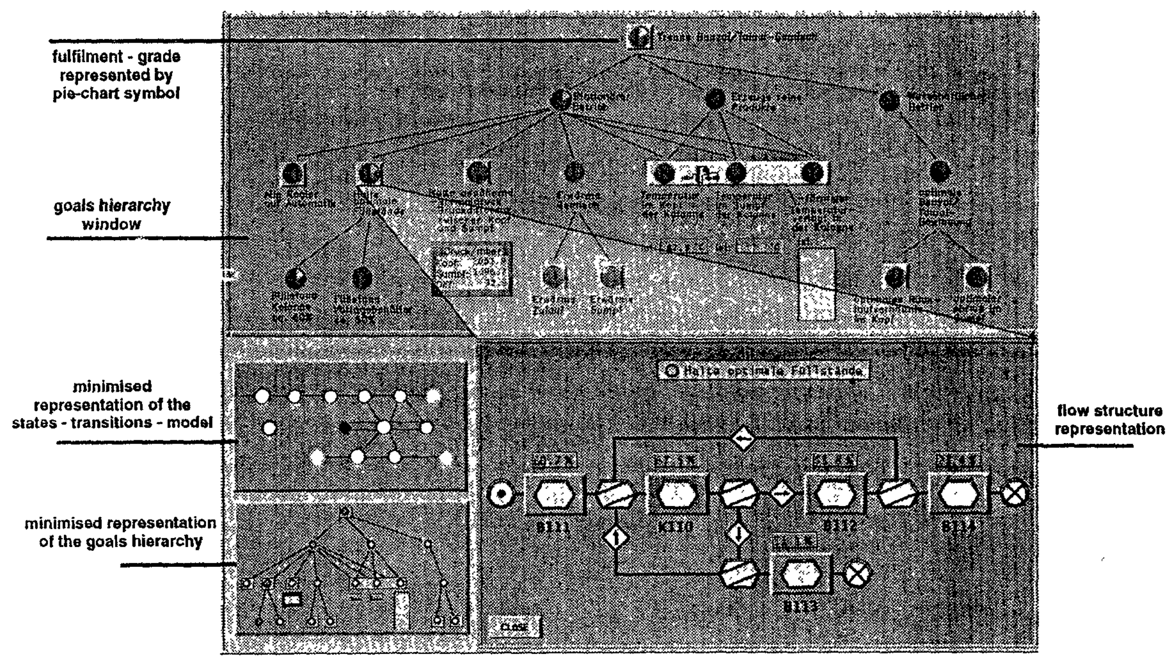


Fig. 8: Multimodel-based Display with goals-means-oriented representation

Fig. 9 illustrates the general entities of a states-causal based display. The minimised representation of the states-transitions model is shown in the lower left-hand corner of display. Details of this representation can be enlarged to get more information a bout the qualitative states-transitions. The operator can select any given state or transition to gather information about the rules underlying this specific state/transition. These rules are presented in the windows shown on the lower right-hand side of the picture. The model representations can also be used to activate the corresponding topological views of the process.

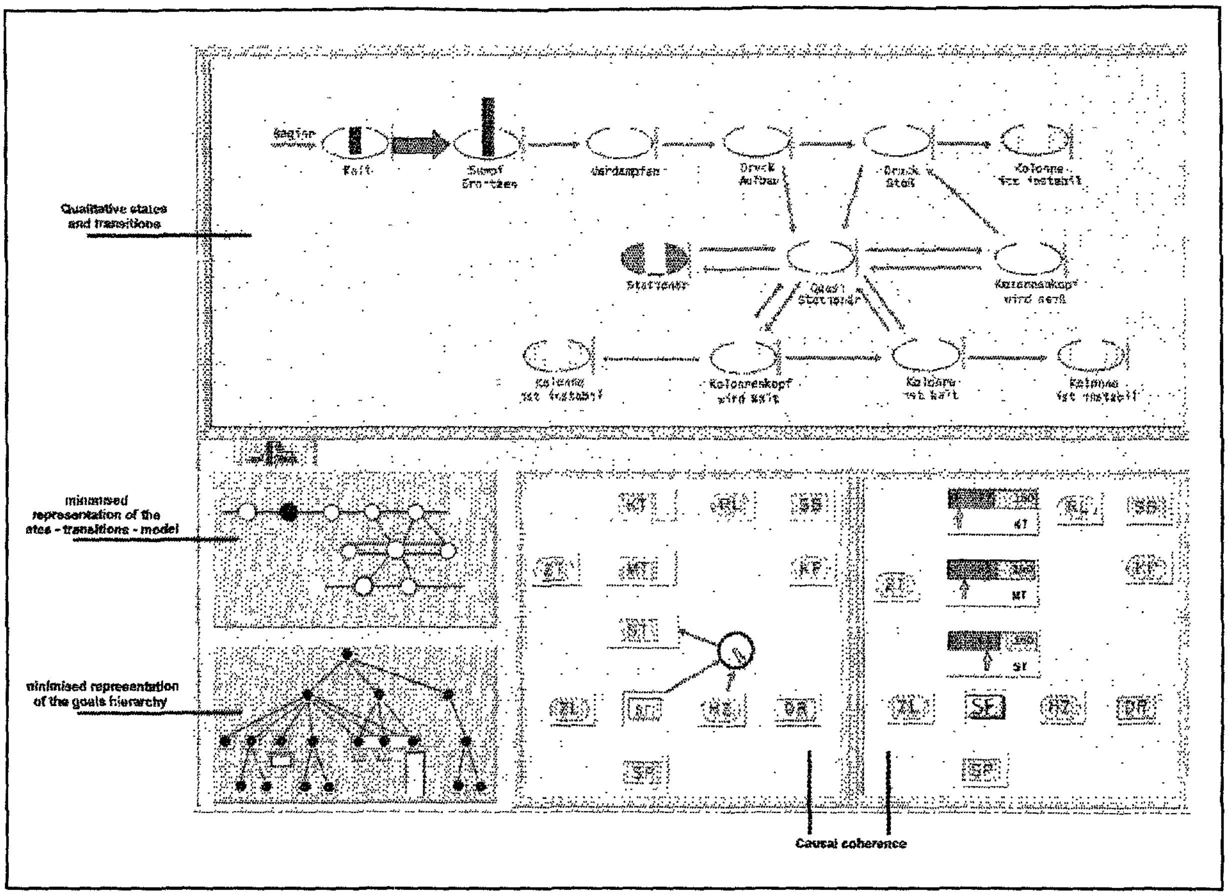


Fig. 9: Multimodel-based Display with states-causal-oriented representation

7. CONCLUSION

In this paper it was shown how different models of the dynamic process on different levels of abstraction can be used in the design of human - machine systems. These models will be allocated and integrated within the so-called „multimodel", and are considered both in the design of the HMI as well as of the knowledge-base. This kind of representation provides the process control operators with a solid based for their supervisory and control tasks (Sheridan, 1993).

Fuzzy logic techniques are used to translate operator's knowledge that is available in the form of natural language into knowledge-base as well as to calculate animation values visualised on user displays.

The assumed positive effect of the proposed methodology on the operators daily work has to be examined be intensive experimental studies. Several different studies are planned, each of which focuses on a different aspect of the operator's overall work context. A first series of experiments is currently conducted. They investigate the effects of different ways to mediate knowledge about the process. First results can be expected soon.

REFERENCES

Ali, S., Heuer, J., Hollender, M. and Johannsen, G. (April 1993). Participative design of human-machine interfaces for process control systems, INTERCHI'93, Amsterdam.

Ali, S., Heuer, J., Hollender, M. and Johannsen, G. (May 1994). Partizipative Bediengestaltung von Oberflächen für Prozeßleitsysteme, in Proc. Leitwarten-Kolloquium, Köln.

Ali, S., Heuer, J., Hollender, M. (1994). Partizipative Erstellung von Bedienoberflächen für Prozeßleitsysteme durch interaktive Gestaltung und Bewertung. Zwischenbericht zum DFG-Forschungsvorhaben JOH-139/6-1. IMAT-MMS, Universität Kassel.

Bainbridge, L. (1987). The ironies of automation. In: J. Rasmussen, K. Duncan and J. Leplat: New Technology and Human Error, 111-120. Chichester: Wiley.

Fishwick, P. A. and Luker, P. A. (Eds.) (1991). Qualitative Simulation, Modeling and Analysis. New York: Springer.

Funke, J. (1992). Wissen über dynamische Systeme: Erwerb, Repräsentation und Anwendungen. Berlin: Springer.

Gilles, E.D., Holl, P., Marquardt, W., Schneider, H., Mahler, R., Brinkmann, K. & Will, K.-H.. Ein Trainingssimulator zur Ausbildung von Betriebspersonal in der Chemischen Industrie (1990). *Automatisierungstechnische Praxis, 7*, 344-350.

Heuer, J., Ali, S. and Hollender, M. (June 1995). Mediation of mental models in process control through a hypermedia man-machine interface, in IFAC-MMS'95, Cambridge, MA.

Johannsen, G. (1992). Towards a new quality of automation in complex man-machine systems, Automatica, 28, 355/373.

Johannsen, G. (1993). *Mensch-Maschine-Systeme*. Berlin: Springer.

Kantrowitz, M., Horskotte, E. and Joslyn, C. (1995). Answers to Qusestions about Fuzzy Logic and Fuzzy Experts Systems, WWW-address http://www.cs.cmu.edu/Web/Groups/AI/html/faqs/ai/fuzzy/part1/faq.html

Larsson, J. E. (1992). Knowledge-based methodes for control systems. ISRN LUTFD2/TFRT--1040--SE. PhD Thesis from Lund Institute of Technology, Department of Automatic Control.

Lind, M. (1981). The use of flow models for automated plant diagnosis. In Rasmussen, J. und Rouse W. B. (eds.) Human Detection and Diagnosis of System Failures. New York: Plenum Press.

Lind, M.. (1982). Multilevel flow modelling of process plants for diagnosis and control. *Int. Meeting Thermal Nuclear Power Safety*, Chigaco, IL.

Lind, M. (1988). System concepts and the design of man-machine interfaces for supervisory control. In: L.P. Goodstein, H.B. Andersen, S.E. Olsen (Eds.), *Tasks, Errors and Mental Models*. Taylor and Francis, London, pp. 269-277.

Lind, M. (1993). Multilevel flow modeling. *AAAI'93 Workshop on Functional Reasoning*. Washington, July 9-15 1993.

Rauh, J. (1994). *Erstellung einer MFM-Benutzeroberfläche für eine Destillationskolonne*. Diplomarbeit Labor für Mensch-Maschine-Systeme, Universität-Gesamthochschule Kassel.

Sheridan, T. (1993). Telerobotics, Automation and Supervisory Control. Cambridge, MA: MIT Press.

Shen, Q. & Leitch, R. (1992). Application studies of fuzzy qualitative simulation. In P. Borne and S. G. Tzafestas (eds.): Mathematics of the Analysis and Design of Process Control. New York: Elsevier.

Sassen, A.-M. (1993) Design issues of human operator support systems. Ph.D. thesis TU Delft.

Struss, P. (1992). Qualitative Modellierung physikalischer Systeme auf dem Weg zu Anwendungen. KI, 4, 50-53.

VDI/VDE Richtlinie 3695 (1986). *Preformatted displays on video display units for the control of process plants.* VDI-Verlag GmbH, Düsseldorf.

Woods, D. D. (1984). Visual momentum: A concept to improve the cognitive coupling of person and computer. *Int. J. Man-Machine Studies*, Vol. 21, pp. 229-244.

Zimmermann, H.-J. (1991). Fuzzy Set Theory and its Applications. Boston, MA.: Kluwer Acadimic Publishers.

Fuzzy predictive control based on human reasoning

R. Babuška J. Sousa H.B. Verbruggen

Delft University of Technology, Department of Electrical Engineering, Control Laboratory
Mekelweg 4, P.O. Box 5031, 2600 GA Delft, The Netherlands, tel: +31 15 783371
fax: +31 15 626738, e-mail: r.babuska@et.tudelft.nl

Abstract

Human knowledge is an important source of information for modeling and control of complex dynamic processes. Fuzzy sets proved to be suitable for dealing with subjective uncertainty encountered when incorporating human knowledge in the design of automatic control systems. Besides direct fuzzy control, in which the control law is explicitly described by If–Then rules, the knowledge-based approach can be applied at a higher level for formulating the control objectives and constraints. Appropriate control actions are then found by means of a multistage fuzzy decision making algorithm, using optimization over a finite horizon as in conventional predictive control. Compared to the standard quadratic objective function, the knowledge-based approach gives the designer more freedom in specifying the desired process behavior. By using fuzzy models, the uncertainty arising from the modeling of complex and partially unknown systems can be represented at the same conceptual level as is the uncertainty in the goals and constraints. Finally, a model-based search for an optimal control strategy can be combined with model-free reinforcement techniques inspired by human learning.

Keywords: Predictive control, fuzzy decision making, optimization, learning.

1 Introduction

Complex, nonlinear and partially unknown systems, encountered for instance in chemical process industry, biotechnology or climate control, present big challenges for automatic control. While the conventional linear control techniques often fail or can be applied only locally, human operators are able to control these systems across a wide range of operating conditions. *Knowledge-based* control tries to integrate the knowledge of human operators or process engineers into the controller design. *Fuzzy control*, one of the most popular techniques, has been successfully applied to a large number of consumer products and industrial processes [10, 12]. Most of the applications of fuzzy control use a *descriptive* approach introduced in the seventies by Mamdani [4]. The operator's knowledge is verbalized as a collection of If–Then control rules, that are directly translated into a control algorithm, as schematically depicted in Fig. 1.

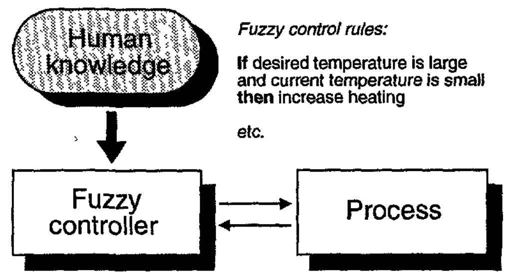


Figure 1: In conventional fuzzy control, operator's knowledge is verbalized as a collection of If–Then control rules.

With this methodology, no explicit model of the process is required, which can significantly reduce the development time if sufficient knowledge is available. If this is not the case, the design must rely on the tuning phase, which may be a tedious and time-consuming trial-and-error procedure. In industrial environments, an on-line experimental tuning is often not acceptable for e.g. safety, economical and environmental reasons. Moreover, it has been observed, that human control skills are sometimes difficult to verbalize since the operator's control strategy can be based on various control principles simultaneously, combining feedforward, feedback, predictive strategies in a complex, time-varying fashion. In that case, an operator may not be able to explain why he or she chooses a particular control action. Experience from knowledge acquisition also shows that opinions of different operators may be very different or even contradicting [5]. Being based on a human control strategy, the descriptive approach is also not suitable for control problems that go beyond the capabilities of the human operator, such as optimization. Process operators usually tend to react quite cautiously and do not want to force the system to the limits of the allowable regions.

In this paper we discuss an alternative approach, where human knowledge is used to specify the control objectives and constraints, not the control protocol itself. A decision making algorithm selects the control actions that best meet the specified criteria, see Fig. 2.

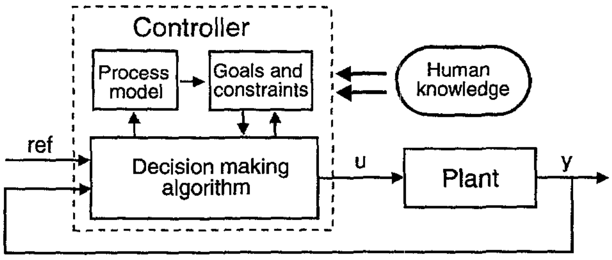


Figure 2: Controller based on objective evaluation and fuzzy decision making.

Since this *prescriptive* approach is closely related to predictive control, we first review the basic concepts of conventional predictive control, than we motivate the approach using fuzzy decision making. Finally, we discuss practical issues related to optimization, its computational complexity and we briefly describe a model-free optimization scheme using reinforcement learning.

2 Predictive control

Model-based predictive control has become an important research area of automatic control theory and it also has been widely applied in industry [7]. Reasons for this success are the ability of MBPC to control multivariable, nonlinear systems under various constraints in an optimal way (with respect to the specified objective function). The working of a MBPC is as follows. A model of the process predicts the process behavior over a specified (finite) prediction horizon, as shown in Fig. 3.

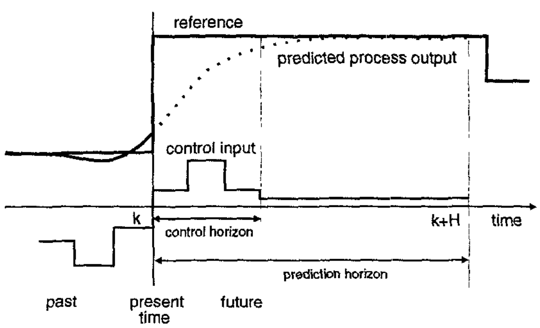


Figure 3: The principle of model based predictive control.

A predictive controller uses an optimization algorithm to calculate a sequence of future controller outputs over a control horizon, such that a specified objective (cost) function is minimized. Most of the objective functions are some modifications of the quadratic form [3]:

$$J=\sum_{i=1}^{H}\alpha_i(w(k+i)-\hat{y}(k+i))^2+\sum_{i=1}^{H_c}\beta_i\Delta u(k+i-1))^2 \quad (1)$$

where $\hat{y}$ denotes the predicted process output, w the desired process behavior (reference trajectory) and u the future control signal. H is the prediction horizon and H_c is the control horizon. Vectors α and β determine the weighting of the output error and the control effort with respect to each other and with respect to time. It is important to realize that the cost function is only a suitable mathematical approximation of the control objectives. While its quadratic nature is convenient for finding analytical solutions for linear models, it may be less suitable for achieving the "real" control goals, such as fast rise time, small overshoot, good damping, etc. Though many authors provide tuning rules that attempt to relate the desired performance to the setting of the individual parameters in (1), see e.g. [9], in practice, extra constraint (such as overshoot constraints, etc.) often must be imposed in order to meet the prescribed goals. The quadratic cost (1) minimizes the variance of the process output, which might not be always desirable. In many processes, it is sufficient to keep the controlled variables within certain limits and more accurate control is not desired since it makes the production more expensive. Too a tight control also reduces the information contents of the data that otherwise may be used for adapting the process model. Though optimal w.r.t. (1), the control response may have some undesirable features such as non-minimum phase closed-loop behavior for a minimum phase plant.

From practical reasons, it is desirable to have a direct control over the influence of the individual components of the objective function on the controller performance. It is advantageous if the degree of compensation among the partial goals and among the goals and constraints can be specified by the designer. This additional freedom can be achieved by choosing a different representation of the objective function, e.g. as a combination of fuzzy goals and constraints, as shown in the following section.

3 Objective function with fuzzy goals and fuzzy constraints

The idea of decision making in fuzzy environment was introduced in the beginning of seventies by Bellman and Zadeh [1]. In fuzzy decision making, the goals, constraints and also the systems under control can be fuzzy. An example of a fuzzy goal is "the product concentration should be *about* 80%", where concentration is the controlled variable and the vague expression *about 80%* is represented by a subjectively defined fuzzy set, for instance as shown in Fig. 4. For a crisp measurement x, the degree of satisfaction of a fuzzy goal G is determined by the membership degree of the measurement in the fuzzy set G, $\mu_G(x)$. For a value expressed as a fuzzy set [1] F, the degree of satisfaction of the goal G is computed as a degree of similarity between the two corresponding fuzzy

[1] Also the process values can be fuzzy, consider for instance notions based on human perception that can be expressed as rule-based combinations of measured variables (e.g. comfort may be defined using rules combining temperature and humidity).

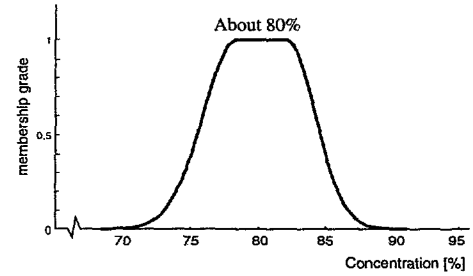


Figure 4: A membership function for the fuzzy goal about 80%.

sets, e.g. as $\max(\mu_F(x) \wedge \mu_G(x))$.

Simple fuzzy goals can be combined in more complex goals or can be refined by adding other conditions, e.g. "the product concentration should be *about* 80% but (and) not *substantially higher* than 82%". Here the goal "not substantially higher than 82%" can represent for instance a temporary restriction that can be added without removing or modifying the original goal. The logical connective *and* and the negation operator *not* are represented as intersection and complement of the fuzzy sets respectively, $G = G_1 \cap \bar{G}_2$, or in terms of membership degrees, $\mu_G(x) = \mu_{G_1}(x) \wedge (1 - \mu_{G_2}(x))$. Fuzzy constraints can be represented in a similar way as fuzzy sets C_i.

Fuzzy decision D for n fuzzy goals $G_1, G_2, \ldots, G_n$ and m fuzzy constraints $C_1, C_2, \ldots, C_m$ is a confluence of these goals and constraints. If we require simultaneous satisfaction of the goals and constraints, we may define D as intersection of the corresponding fuzzy sets:

$$D = G_1 \cap G_2 \cap \ldots \cap G_n \cap C_1 \cap C_2 \cap \ldots \cap C_m$$

or in terms of membership degrees

$$\mu_D(x) = \mu_{G_1}(x) \wedge \mu_{G_2}(x) \wedge \ldots \wedge \mu_{G_n}(x) \wedge \\ \wedge \mu_{C_1}(x) \wedge \mu_{C_2}(x) \wedge \ldots \wedge \mu_{C_m}(x)$$

The *maximizing decision* x^m is any $x \in X$ that maximizes $\mu_D(x)$, i.e.

$$\mu_D(x^m) = \bigvee_{x \in X} \mu_D(x)$$

Optimizing the system's performance over a finite horizon, as in predictive control, corresponds to finding an optimal sequence of decisions in a multistage decision making process. Assume that the system under control is described by a state transition equation

$$x(k+1) = f(x(k), u(k)) \qquad (2)$$

Given the current state $x(k)$ we are interested in finding a sequence of actions $u_k, \ldots, u_{H-1}$, corresponding to the maximizing decision. This is a nonlinear optimization problem that can be solved e.g. by dynamic programming.

Example: Predictive control of a container crane. We give a simple example of predictive control with fuzzy goals and constraints of a container crane shown schematically in Fig. 5. A simulation model of a real container

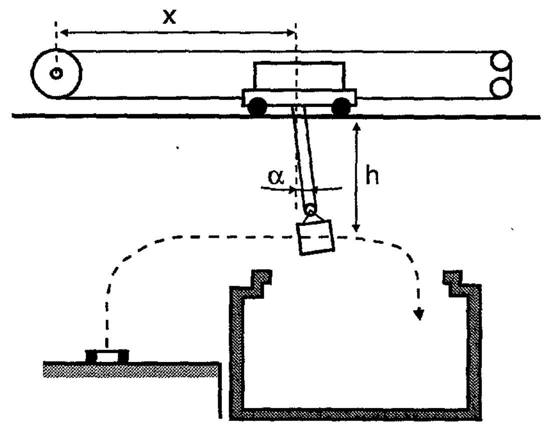


Figure 5: Schematic drawing of a container crane

crane of the port of Kobe was taken over from [8]. The system output variables are the trolley position x, the length of the rope h and its angle α. The torque of the trolley drive and the torque of the hoist motor are the manipulated inputs. In our example we only consider setpoint change of the trolley position from $x = 35$ m to $x = 45$ m on which we compare three different objective functions:

1. Sum of squared errors between the reference and the actual trolley position of x ($J = \sum_{i=k+1}^{k+H}(r(i) - x(i))^2$.
2. Minimization of the overshoot in the trolley position x, using a fuzzy goal with a membership function shown in Fig. 6 a.
3. A combination of 2) with a criterion for minimizing the variance of x around the setpoint, using membership function shown in Fig. 6 b.

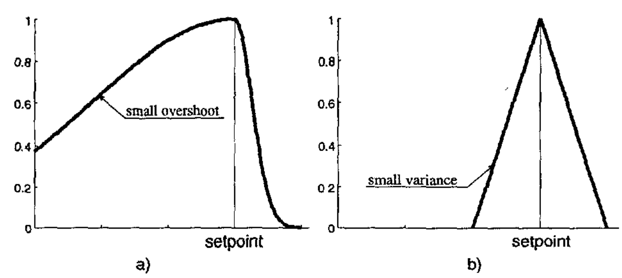


Figure 6: Membership functions for small overshoot (a) and small variance (b).

Note that the variance reduction term in 3) aims at a similar goal as 1). The two goals in 3) are combined using the logical *and* operator, i.e. the minimum of the membership degrees, to represent that both the goals should be satisfied simultaneously. Note that there is no compensation between the criteria involved, as opposed to (1). The prediction and control horizon were both 8 steps.

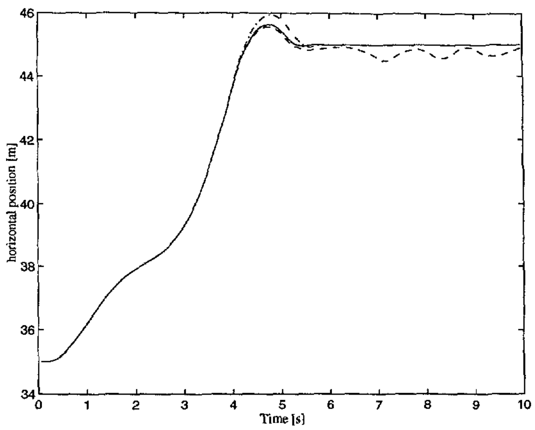


Figure 7: Simulation results for three different objective functions: minimum variance (dash-dotted line), overshoot (dashed line) and combination of the two using the minimum operator (solid line).

The simulation results shown in Fig. 7 clearly show that the oscillation obtained with the overshoot criterion alone can be eliminated by simply adding another criterion for variance minimization.

4 Approximating processes using fuzzy models

For engineering purposes, *mathematical models* are often constructed, based on, for instance, differential or difference equations, derived from physical laws and with parameters estimated via experimental identification. For well-defined systems, these standard mathematical tools lead to good models, even though the modeling process is often very tedious. There are, however, many systems where the underlying physical mechanisms are not known or are so complex that a mathematical model is difficult to obtain and to use. On the other hand, such systems can be described quite simply and with a sufficient accuracy in a verbal way, using fuzzy If–Then rules. Fuzzy sets are used for *partitioning* the continuous domains of the system input and output variables into a small number of overlapping regions labeled with linguistic terms such as LOW, HIGH, etc. A fuzzy model describes the system by establishing relations between the input and output labels. These relations can be expressed in the form of If–Then rules, mapping fuzzy regions from the premise space to other fuzzy regions in the consequent space. For instance, the following rule

If $voltage$ is $HIGH$ then $speed$ is $HIGH$

relates HIGH voltage on an electric motor to HIGH speed of its rotor. Fuzzy sets for linguistic terms HIGH are defined in their respective domains of voltage and rpm. Fuzzy inference mechanism ensures interpolation between the rules, providing answers to inputs that are not defined in the rule premises. This idea of linguistic fuzzy modeling was introduced in the pioneering papers of Zadeh and applied later on by Mamdani to fuzzy control of dynamical plants, see e.g. [4]. Instead of an explicit description by rules, the mapping can be defined via a *fuzzy relation*. The construction of this so-called fuzzy relational model is based on the theory of fuzzy relations and relational equations, see e.g. [6]. The output fuzzy set Y is computed from the input fuzzy set X via relational composition

$$Y = X \circ R$$

A dynamic system, such as (2) can be described as a composition of the input fuzzy set $U(k)$, state fuzzy set $X(k)$ and a relation R describing the system

$$X(k+1) = U(k) \circ X(k) \circ R$$

Fuzzy models have several useful properties. First, their belong to the class of general function approximators, e.g. a fuzzy model can approximate a smooth function to any degree of accuracy, as shown by Wang [13] (among others). Secondly, different kinds of information can be integrated for building fuzzy models, such as knowledge expressed as If–Then rules and numerical data (process measurements). Finally, the mathematical framework for representation of fuzzy models is convenient for analytical manipulations, such as analysis of the model, its inversion, etc. Fuzzy models also can be seamlessly integrated in a predictive control system based on the fuzzy decision making approach, as shown in the previous section.

5 Optimization based on reinforcement learning

The previously described approach requires a reliable model of the plant and if a significant model–plant mismatch appears, the controller performance rapidly degrades. Modeling and identification of complex systems is a difficult, time-consuming and expensive task, resulting in the fact that most of the design effort (sometimes as much as 80%) is spent on developing a good process model [7]. In many cases, an accurate model cannot be obtained which places the use of MBPC out of question. Therefore, techniques for optimizing the control policy without an explicit model of the plant are desirable.

Here again, it is useful to take inspiration from the way humans adapt their behavior in a particular environment without knowing an accurate model of that environment. Many learning tasks (think for instance of learning to play tennis) consist of repeated trials (attempts to hit the ball) followed by a reward (a nice shot) or punishment (picking the ball from the ground). Each trial can be a dynamic sequence of actions (run, taking a stand, hitting the ball) while the feedback (reinforcement) comes only at the end. Therefore, a large number of trials may be needed to figure out which particular actions were correct and which must be adapted.

A family of algorithms inspired by human and animal learning is known as *reinforcement learning*. Reinforcement learning assumes that there is no direct evaluation of the quality of the selected control action. Instead, an indirect evaluation is received, possibly after a sequence of control actions, in terms of (dis)satisfaction of the control objectives and/or violations of constraints. The reinforcement system learns how to predict the outcome of each particular action, using techniques like *temporal difference* [11] or *Q-learning* [14]. This prediction is used to adapt parameters of a suitable general function approximator (e.g. a neural network) to iteratively approximate the optimal control policy. This approach differs from other optimization techniques in two respects: i) no model of the system is needed, ii) the optimization is not done at once, it is distributed over a large number of small steps that gradually approximate the optimal policy.

Most of the approaches based on neural networks do not employ any prior information and learn the control policy from scratch. It is well known that *prior knowledge* can speed up the learning process (in our example, it might be useful to know how to grip the racket and how to stand for backhand and forehand). In more serious control tasks, prior knowledge is essential for a correct functioning of the system at the very start of the learning process (think of driving a car without knowing the function of the pedals, or stochastically exploring the effect of individual control variables in an unstable chemical process). The ability of fuzzy rule-based system to use prior knowledge on one hand and to approximate nonlinear functions on the other hand is advantageous for combining fuzzy models with reinforcement learning [2]. By using reinforcement learning, possibly approximate and imprecise prior knowledge expressed in terms of fuzzy rules can be refined on line, during the control process.

In order to briefly explain the principle of reinforcement learning, let us assume (without loss of generality) that the system to be controlled is represented by a state transition function (2). When the system is fuzzy, f is a rule base. The goal is to learn (or adapt) an associative mapping $\pi : X \to U$, by maximizing a (scalar) evaluation of the performance (reinforcement). Here π is so called *policy function* and is equivalent to a controller, that for a particular state $x \in X$ computes a control action $u \in U$, see Fig. 8. When controlling dynamic systems, the reinforcement evaluation is usually available only after a sequence of state–action pairs $\{(x(k), u(k))\}$.
For discrete control actions $U = \{u_1, u_2, \dots, u_m\}$, the policy π can be formed as a composition of two functions: a function approximator $g : X \to R^m$ that assigns

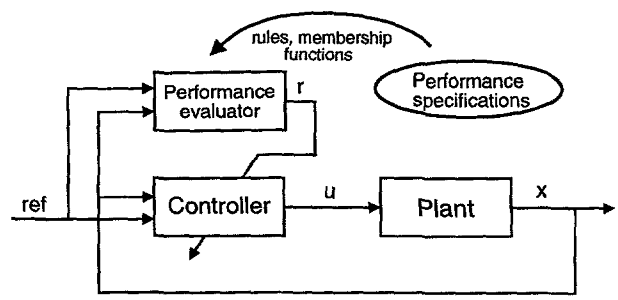


Figure 8: Control scheme with reinforcement learning.

each u_i a real value representing its merit (cf. degrees of fulfillment in fuzzy inference) and a fixed function $M : R^m \to A$ that selects a particular control action (cf. the aggregation/defuzzification block). Often, M is a maximum selector combined with a stochastic action modifier that ensures proper exploration of the search space.

The basic principle of reinforcement learning can be explained as follows. The goal is to maximize the total reinforcement over time, which can be expressed as a discounted sum of the immediate payoffs $r(x, u)$ computed by the performance evaluator (see Fig. 8). The sum, so called *value function* $V : X \to R$, is defined as:

$$V(x) = \lim_{N\to\infty} E\{\sum_{k=0}^{N-1} \gamma^k r(x(k), \pi(x_k))|x_0 = x\} \quad (3)$$

where the constant $\gamma \in [0, 1)$ is a discount rate. $V(x)$ can be approximated, using the *temporal difference* operator

$$\Delta(x) = r(x(k), u(k)) + \gamma V(x(k+1)) - V(x(k))$$

which is a difference of predictions of $V(x)$ at two consecutive time steps. The estimate $\hat{V}(x)$ is updated, using $\Delta(x)$

$$\hat{V}^{k+1}(x) = \hat{V}^k(x) + \beta\Delta(x)$$

where β is a small positive constant. Finally, the learning rule for the merit function $g(x)$ is:

$$g^{k+1}(x) = g^k(x) + \alpha(\rho(x(k), u(k)) - V(x(k))) \quad (4)$$

where α is a small positive constant and $\rho(x(k), u(k))$ is the expected total reward obtained if u is applied to the system at state $x(k)$ and then policy π is followed. This estimate is not available, but it can be approximated as:

$$\rho(x(k), u(k)) \approx r(x(k), u(k)) + \gamma\hat{V}(x(k+1)) \quad (5)$$

Using (5) in (4) gives the update law for $g(x)$:

$$g^{k+1}(x) = g^k(x) + \alpha\Delta(x)$$

Obviously, the reinforcement learning (RL) methods solve the same optimization problem as the dynamic programming (DP) methods. The difference is that the DP methods are off-line, while RL techniques try to learn the optimal policy on-line, concurrently with the system operation. The DP methods search the entire space $X \times U$, while the RL methods operate on the a subset of states $\tilde{X}$ that occur during the system operation. Since $\tilde{X}$ may be significantly smaller than X, RL methods do not suffer from the curse of dimensionality as much as the DP methods. RL methods also do not require a system model and can be extended to continuous spaces of u.

6 Concluding remarks

In this paper we discussed some links between conventional predictive control, multistage fuzzy decision making and reinforcement learning. These three approaches, originating from different roots, nicely complement each other and in combination may be applied to control of complex dynamic systems that are difficult to deal with using standard methods. Let us summarize the main features of the proposed approach:

- The control objective function can be specified as any suitable combination of the terms representing degree of satisfaction of the goals and constraints, ranging from a simple conjunction to a problem-specific rule base. Such a rule base, for instance, can capture context-dependent importance of the control goals and constraints or different ways of their aggregation [15].
- With such an objective function, each control action can be explained in terms of partial decisions and degrees of satisfaction of the individual criteria, which is suitable for tuning and monitoring purposes.
- Subjective uncertainty can be easily incorporated in the specification of goals and constraints.
- The process itself can be represented as a fuzzy rule-based or a relational model. In this way, also ill-defined, partially unknown or highly nonlinear systems can be modelled in a transparent way.

For the higher flexibility one has to pay by increased computational costs, since the optimization problem is in general nonlinear and often non-convex. Reinforcement learning is considered as an on-line optimization method which can simultaneously add adaptivity features to the controller. Potential applications of the proposed methodology include such systems where humans have been or are part of the control structure or the environment, such as climate control, telemanipulation, crane control, etc.

References

[1] R. E. Bellman and Lotfi A. Zadeh. Decision making in a fuzzy environment. *Management Science*, 17(4):141–164, 1970.

[2] H.R. Berenji and P. Khedar. Learning and tuning fuzzy logic controllers through reinforcements. *IEEE Transactions on Neural Networks*, 3:724–740, 1992.

[3] D. Clarke. Advances in model-based predictive control. In D. Clarke, editor, *Advances in Models-Based Predictive Control*, pages 3–21. Oxford University Press, Oxford, 1994.

[4] E.H. Mamdani. Applications of fuzzy algorithms for control of simple dynamic plant. In *Proceedings IEE*, number 121, pages 1585–1588, 1974.

[5] Sander F.J. Ouwerkerk. *Development and Implementation of a Fuzzy Logic Controller for a Chemical Process*. Master's dissertation, Shell Nederland B.V., Delft University of Technology, Delft, The Netherlands, February 1995. (in Dutch).

[6] W. Pedrycz. An identification algorithm in fuzzy relational systems. *Fuzzy Sets and Systems*, 13:153–167, 1984.

[7] J. Richalet. Industrial applications of model based predictive control. *Automatica*, 29:1251–1274, 1993.

[8] Y. Sakawa and Y. Shindo. Optimal control of container cranes. *Automatica*, 18(3):257–266, 1982.

[9] A.R.M. Soeterboek. *Predictive Control - A Unified Approach*. PhD dissertation, Delft University of Technology, Delft, The Netherlands, November 1990.

[10] M. Sugeno, editor. *Industrial applications of fuzzy control*. Elsevier Science Publishers B.V. (North-Holland), Amsterdam, The Netherlands, 1985.

[11] R.S. Sutton. Learning to predict by the method of temporal differences. *Machine learning*, 3:9–44, 1988.

[12] T. Terano, K. Asai, and M. Sugeno. *Applied Fuzzy Systems*. Academic Press, Inc., Boston, 1994.

[13] L.-X. Wang. Fuzzy systems are universal approximators. In *Proc. IEEE Int. Conf. on Fuzzy Systems 1992*, pages 1163–1170, San Diego, USA, March 1992.

[14] C.J.C.H. Watkins and P. Dayan. Technical note: Q-learning. *Machine learning*, 8:279–292, 1992.

[15] S. Yasonobu and S. Miymoto. Automatic train operation system by predictive fuzzy control. In M. Sugeno, editor, *Industrial Applications of Fuzzy Control*, pages 1–18. North-Holland, 1985.

Decision support in anaesthesia monitoring

Philip M.A. de Graaf

Delft University of Technology, Department of Technology and Society, Safety Science Group,
Kanaalweg 2B, 2628 EB Delft, The Netherlands
Phone +31 15 78 62 64, Fax +31 15 78 18 43, E-mail philip@it.et.tudelft.nl

Abstract

Ever since the first anaesthesia was administered, a very important subject of research in anaesthesia has been the improvement of patient safety. An important result of this research has been the advent of new monitoring devices. Partly because of these new devices the anaesthetist has to manage large amounts of data during his task. This is one of the reasons why most incidents in anaesthesia involve human error. The research we will present in this paper, is part of a project that was started to develop methods for supporting the anaesthetist in his task. In this paper we will propose a research method that should lead to a thorough understanding of the anaesthesia process. A number of approaches are presented to model the anaesthesia process, along with a number of methods that can be used to test the validity of the model. Finally a method is discussed to identify where things go wrong and what kind of support the anaesthetist needs to prevent this.

Introduction

Ever since the first reported anaesthesia procedure in 1847 and the first reported death during anaesthesia in 1848, improving patient safety has been an important research topic. Both anaesthesia techniques as well as patient monitoring have been subject to a large number of improvements. In our research we concentrate on the monitoring task of the anaesthetist.

Monitoring patient condition serves two purposes in anaesthesia. If we look at the normal anaesthesia process as a manual feedback control loop, monitoring the patient is the anaesthetist's feedback in this process. The other purpose of monitoring is to detect process deviations. A deviation process or incident is any event causing a change in patient physiology or equipment behaviour that exceeds the limits of normal process variability.

Observation of the patient and the surgical procedure he is going through, was and is an essential part of the monitoring process. However, during the last two decades more and more monitoring devices have become available to continuously monitor a large number of patient functions (parameters). In the early 80's every monitoring device was in a separate unit with its own display and user interface. This resulted in a monitoring setup which was very difficult to survey. In the late 80's and early 90's, monitoring equipment has become available that integrates all monitoring modalities in one device. These devices have one integrated display for all measured variables and one user interface to control alarm and display settings. Although data-acquisition and data-display are integrated now, data processing is still done separately for every parameter. The anaesthetist must therefore continuously monitor between 2-10 measured physiological signals and between 10-30 features extracted from these signals. The only support he gets in this task are a number of high and low limit alarms on the feature values, of which he should set the limits himself. Clinical research shows that as little as 3% of all alarms triggered represent a problem with the patient (Kestin et al, 1988). Therefore most anaesthetists ignore or disable most alarms. This means that the anaesthetist has to work with large amounts of data offered to him, without any support to manage them.

The anaesthetist uses the data offered to him to obtain information about the status of the patient. This information is used to decide whether intervention is necessary, and if so what kind of intervention. Means of intervention are: administration of drugs in different dosages, infusion of different kinds of fluids and blood with different infusion rates, and changing administration of anaesthetic gasses and oxygen. All together there are about 50 different controls which can be adjusted or applied individually as well as combined with each other.

Considering the complexity of the anaesthetists task described above, it is not surprising that recent research shows that 75 % of all incidents involve human error (Chopra et al, 1992). Chopra's study was based on voluntary significant observation reports. A significant observation was defined as "any deviation, however minor, from acceptable safe practice or working condition". For patients in good condition administering anaesthesia is a routine task. Human error probability for routine tasks for a trained operator is in the order of 10^{-3} (NRC, 1975). Reliable reports on incident rates in anaesthesia are not available. There is however some research based on voluntary incident reports. The problem of underreporting that is inevitable in this kind of research can be accounted for by using expert judgement. Research based on this method shows an overall incident rate of $2 \cdot 10^{-3}$ for healthy patients (Paté-Cornell, 1994). From this we may conclude that incident rate is in the same order of magnitude as human error probability.

One approach to improve patient safety is to improve anaesthetist performance by training and organisation of the entire workspace. This will however hardly affect the 10^{-3} error probability for routine tasks. To improve performance in this respect it is better to support the anaesthetist in his task in order to minimise the effects of the errors he is destined to make.

Support can be offered in a number of different ways ranging from automation of subtasks (like automatic control of blood pressure) to complete support of the anaesthesia task. The research being described in this paper is part of a project that is started to find ways to support the anaesthetist in his monitoring task with the goal to improve patient safety. If we want to be able to justify the means of support we want to offer, a thorough analysis of the anaesthesia monitoring process is necessary. The project goals can be formulated as follows:

- Develop methods to automate a number of basic data and information processing tasks, using advanced automatic signal processing and pattern recognition algorithms.
- Try to acquire a thorough understanding of the anaesthesia monitoring task in terms of data and information processing and the process of decision making, in order to identify how the anaesthetist should be supported.

In this paper we will concentrate on the second project goal. If we analyse this goal we can identify three different issues that should be tackled. To acquire understanding of the anaesthesia monitoring task we will have to build some kind of model to describe it. If we have build a theoretical framework, which we want to use to identify and solve practical problems, it will have to be validated. Only if we have such a valid model of the anaesthesia monitoring process it will be possible to identify where things can go wrong. In short this comes down to the following research issues:

- Construction of a model that describes data, information and activity flow in the anaesthesia process for both normal and deviation processes.
- Development of a method to test the validity of the model that is build.
- Identification of the monitoring tasks for which support is needed, based on the validated process model.

In the following sections we will discuss possible approaches to tackle each of these issues.

Analysis

In the introduction we briefly touched on terms like data, information and decisions. Since these are the key subjects of investigation we will briefly illustrate how these concepts are interrelated. In general we can draw the chart in figure 1.

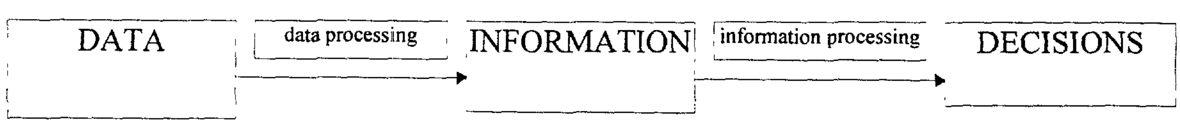


Figure 1 Relations between data, information and decisions

The relations illustrated in figure 1 are almost trivial but form the basis for any model of information and data flow, and decision making in the anaesthesia monitoring process. The anaesthetist has access to a lot of data which are only part of which is relevant in a given situation. One could say that he is continuously processing data to obtain information about relevant aspects of the condition of the patient. Based on this information he decides what should be done to change or maintain the current patient state.

We already stated that the anaesthesia process can be divided into two important states. The normal process and the deviation process. We will start with a description of both states.

Normal process

One of the main causes of the complexity of the anaesthesia control process is that the variables to be controlled cannot be measured directly. The most important variables to be controlled during anaesthesia are tissue oxygenation and depth of anaesthesia (sufficient reduction of pain sensation and consciousness). There are no means of measuring these parameters directly in a way that is practically feasible. This means that controlling patient state, even under normal conditions, requires a lot of data and information processing by the anaesthetist. We could say that the anaesthetist builds an internal model of the patient and the anaesthetic delivery system based on knowledge, experience and a large amount of data about the patient. This model is used to continuously estimate the state of the patient in terms of adequacy of tissue oxygenation and anaesthesia. Most of the anaesthetist's activities are directed to the optimisation of these two aspects of patient state. This is what we will call the normal process flow.

Deviation process

Events causing large variations in the two control variables mentioned can occur during normal process flow. Only if an incident occurs a deviation process is initiated. In this context an incident is an event that causes a substantial change in the physiology of the patient or a problem occurring somewhere in the delivery of anaesthetic gasses or medication. This results in a mismatch between the data coming from the patient and the internal model of the anaesthetist. This is what we will call a deviation process. After detection of the mismatch or the event itself, the anaesthetist will try to diagnose the problem. If he can identify the problem, the anaesthetist will try to correct it. During this process the internal model will be used to diagnose and solve the problem and it will be continuously adjusted. If the problem cannot be corrected he will try to bring the patient in the best possible state given the problem that occurred. In this case he will also use and adjust the internal model of the patient during the process. If the new internal model is build, the process switches back to the normal state. If either the deviation is not detected in time, the wrong diagnosis is made, or the new internal model is wrong, the mismatch between the internal model and the real process will continue to exist. This can eventually lead to patient injury. If this happens the incident has evolved into a critical incident.

In the introduction we argued that it is impossible to prevent incidents from occurring. We could however try to prevent an incident from developing into a critical incident. We therefore want to support the anaesthetist in detecting a possible mismatch between his internal model and data coming from the patient.

In the next section we will describe a number of different process models and discuss their value for our research based an a number of requirements we will place upon them. We will also present a number of different methods to test a theoretical model. Finally we will give an indication of how we could identify steps in the process where things can go wrong.

Methods

The methods we will discuss in this section are not yet in a definite form. We only want to describe a number of possible approaches to acquire some insight in the anaesthesia process.

Model

If we want to offer support to the anaesthetist by automatic data and information processing, we should have detailed understanding of how the monitoring task is carried out by the anaesthetist himself. We can try to model the anaesthesia process in terms of information and data flow and processing. If this (set of) model(s) can be tested and validated, it can be used to identify which task can and should be done by the computer and which tasks are better left to the anaesthetist. Based on the foregoing we can formulate a number of requirements the model(s) should meet. These requirements are listed below:

- The model(s) should cover all data and information processing aspects of the anaesthesia monitoring and control process.
- The model(s) should cover the normal process, the detection of deviations, as well as the deviation process.
- The model(s) should be consistent with the most recent views on human cognitive processes.
- Data and information flows should be modelled in such a way that all transformations performed on them are explicitly described.
- The model(s) should define the sequence in which the information and data processing tasks are carried out.

Dynamic decision making model

The first model we will discuss is the model developed by Gaba (1994). He models the thought processes of anaesthetists as they administer anaesthesia and respond to perio-operative problems. This model is illustrated in figure 2. It was developed for both normal and deviation process flow. It also incorporates current views on human cognitive modelling. However, data and information flows are not explicitly modelled. Furthermore there are a lot of actions described in this model without a clear place in the sequence of events. Especially the supervisory control level and the resource management level contain a lot of actions without a clear place in the execution flow. Another problem of the Gaba model is that there is no explicit transition from the normal process to the deviation process. This model can be very useful to describe the process in a comprehensive way. However, if we are interested in specific information and data processing tasks a more specific model is needed. We therefore believe that for our purpose this model is not appropriate. If we would like to use it significant changes to it would be necessary.

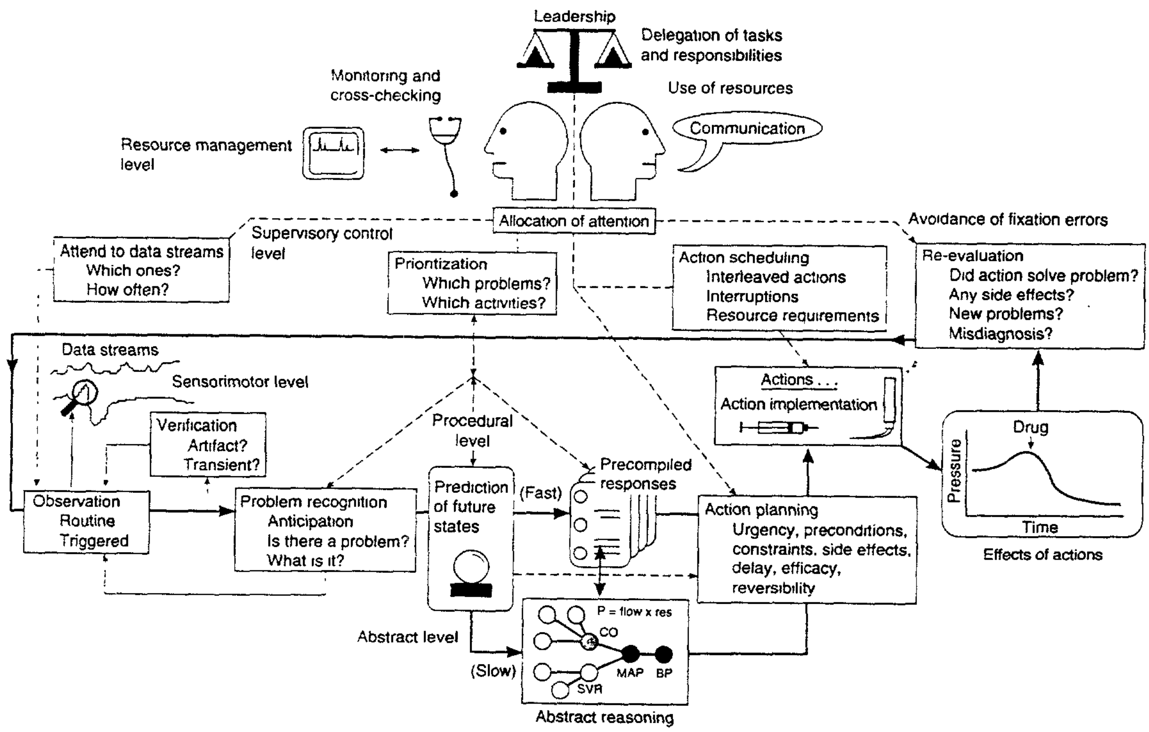


Figure 2 Complex model of dynamic decision making and crisis management in anaesthesiology (Gaba, 1994)

Deviation process model

A totally different approach is to model the behaviour of the anaesthetist only for detection and occurrence of deviation processes. There are various models developed for this purpose (Gerdes, 1994; Rasmussen, 1987; Hale, 1987). We will discuss only the Glendon and Hale model which is illustrated in figure 3. The model gives a good description of the cognitive processes involved in detecting and controlling problems. Data and information flows however, are not incorporated in the model. Another limitation of the Glendon and Hale model is that the normal process flow is not modelled. For the development of methods to support the anaesthetist, it is not possible to separate the detection of deviations from the normal process flow. This model could be useful only to identify where things can go wrong once a deviation is detected.

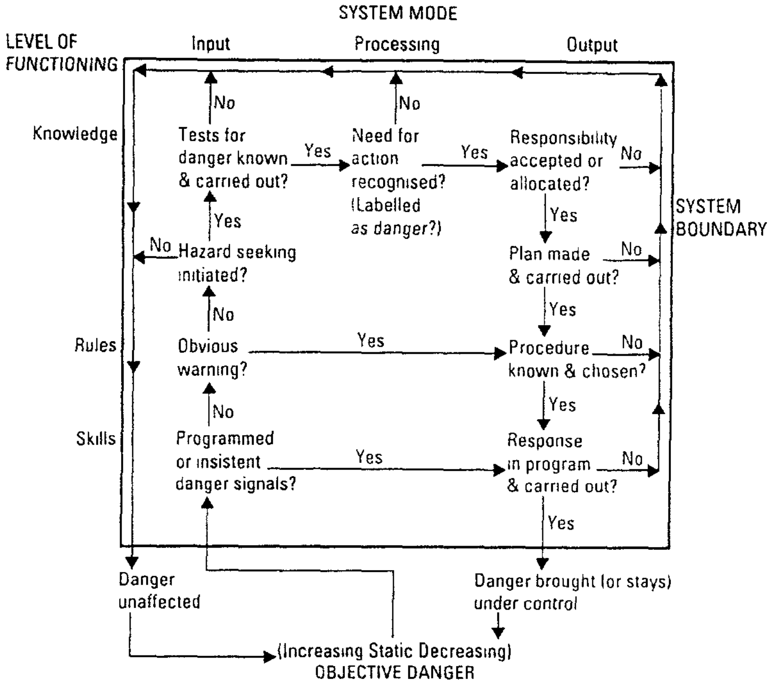


Figure 3 Behaviour in the face of danger (Hale, 1987)

Feedback control model

Another approach is to try to model the entire anaesthesia monitoring and control process as a feedback control system. In figure 4 a possible realisation of such a model is given. This is only a very elementary model to demonstrate the concept of such a model. One of the biggest shortcomings of this model is that it is very difficult to map the different elements in the model to human cognitive processes. Cognitive processes are often distributed over a number of elements or are only a small part of an element. This means that it will be difficult to identify where things can go wrong based on knowledge about human cognitive errors. Furthermore the model does not cover deviation processes. This could be incorporated in the model but problem solving is not really a control task. The model is suited to identify the kind of information used for each element, because information flows can be made very explicit. It is also possible to define a clear sequence of events. For detailed analysis of data and information processing a more detailed model will be necessary.

Considering the above it is clear that none of the approaches covers the entire anaesthesia process in sufficient detail. Therefore we will either have to take a deviation process model for the deviation process and a feedback control model for the normal process. We then also need a clear criterion for switching between these two models . We can also build a hybrid model combining the two models into one.

We have not discussed the possibility of a mathematical model of the process yet. In the past, attempts have been made to build a complete patient model. It seems however, that it is impossible to build a general mathematical model describing the entire process. We therefore choose a more qualitative approach that does offer the possibility to build a complete process model.

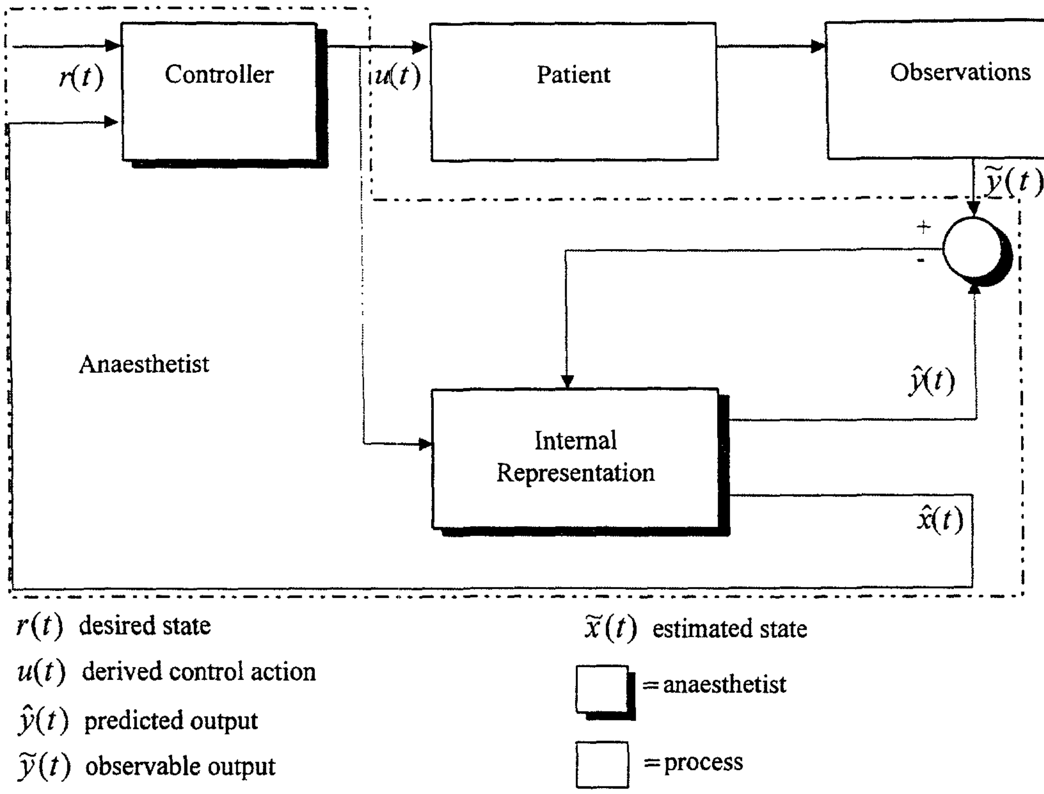


Figure 4 Feedback control model of the anaesthesia process.

Model validation

If we have build a model or combination of models we can use a number of different methods to validate it. Each technique has its own advantages and disadvantages. We will discuss a number of methods with their pros and cons.

Observation

The first possible method is observation of the anaesthetist. This technique is very useful to obtain a preliminary impression of the validity of the model. The results of these observations can be used to adjust the model where necessary. A disadvantage of this technique is that it is time consuming and subjective. It is therefore not suitable for drawing general conclusions about the validity of the model. We can however use this method to get the final version of our model to which a more thorough method of validation will be applied.

Expert opinions

Another approach to validate the model are expert opinions. By confronting anaesthetists with either imaginary or practical cases, a detailed analysis of the monitoring process is possible. We can also get some insight in the relevance of different types of data and information in different situations. Finally, we can get a basic understanding of the cognitive processes involved in decision making. Since this is a time consuming form of data collection we can only interview a small number of anaesthetists most of which will be working in teaching hospitals. This method can be very useful to formulate hypotheses which can be tested on a large population of anaesthetists using questionnaires.

Questionnaires

The only possible method for larger scale data collection are questionnaires. This method is only useful if we have a number of hypothesis we would like to test. We can ask a large number of anaesthetists to judge a relatively large number of cases with different amounts of information presented within each case. The information to be presented and the cases to be judged will be based upon the knowledge from expert opinions research. This method can therefore not be done without the previously mentioned methods. If it is practicable to do this kind of research, we can reliably draw more general conclusions about the validity of our model.

The three methods together cover the entire validation process. It is not yet clear whether these methods are all feasible and practicable. If it is not possible to analyse the entire anaesthesia process in this way we can try to limit our research to very specific cases. This might be necessary to prevent that we drown in the data we will collect.

Identifying sources of human error

If we have a valid model with detailed information about the process of data and information usage and processing, we can start to investigate where things can go wrong.

Incident analysis

A possible source of data for human error investigations are incident reports collected for previous research. If we can map human errors in these reports to elements in our model, we can demonstrate where support is most needed and in what way. This will only work if the model we build is detailed enough to pinpoint places where errors can be made.

Cognitive error identification

Another source of human error identification is a more theoretical approach. We can use knowledge about possible errors in human cognitive processes to identify which processes in the anaesthesia monitoring task are likely to go wrong. This means that the model we build must be suitable for identification of errors in cognitive processes.

Discussion

In this paper we have tried to give a brief description of the problems we want to address in our project. We have presented a number of methods we could use to model the process we are investigating. We concluded that it is difficult to capture everything in one model. Although we will still try to find such a model, it might be sufficient to use two models next to each other to be able to model the different aspects of the process each in its own way. We also presented a number of different methods to test the validity of the model(s) used. Although it is not yet clear whether all of these methods will be feasible, we think that a hierarchical approach as presented above is optimal. The success of the incident analysis depends largely on the success of the other parts of the research. Even if a model can be build and validated, it is still not clear whether the available incident reports will contain the proper information for identification of elements in the model where things go wrong. For cognitive error identification the problem lies in the compatibility of the model and cognitive model theory. By taking into account these demands on the process model(s) it should be possible to achieve the goal mentioned in the introduction. We hope we can thus derive the specifications for a anaesthesia monitoring support system.

References

Chopra V, Bovill JG, Spierdijk J. Koornneef F, Reported significant observations during anaesthesia: a prospective analysis over an 18 month period. *British Journal of Anaesthesia*, 1992. **68**: 13-17.

Gaba DM, Fish KJ. Howard SK. *Crisis management in anaesthesiology*. Churchill livingstone, New York, 1994.

Gerdes V, Method to identify cognitive erros during accident management tasks. To be published in: *Proceedings of the IFAC Symposium MMS '95*, Cambridge, 1995.

Hale AR, Glendon AI, *Individual Behaviour in the Control of Danger*. Elsevier Science Publishers BV, Amesterdam 1987.

Kestin IG, Miller BR, Lockart CH, Auditory Alarms during Anaesthesia Monitoring. *Anaesthesiology*, 1988; **69**: 106-109.

NRC, *Reactor Safety Study* WASH 1400. NUREG 751014. 1975.

Paté-Cornell M, Murphy DM, Lakats LM, *Anaesthesia patient safety: Probabilistic risk analysis and benefits of organizational improvements*. Report to the Anaesthesia Patient Safety Society, Department of Industrial Engineering and Engineering Management, Stanford University, 1994.

Rasmussen JT, A framework for classifying errors. In: Rasmussen et al; *New technology and Human Errors*, Wile, Chichester, 1987.

Short TG, O'Regan A, Lew J, Oh TE, Critical incident reporting in an anaesthetic department quality assurance programme. *Anaesthesia*, 1992, **47**: 3-7.

Session 6

Decision Support Systems II

Chairman: H.G. Stassen, The Netherlands

Automatic Prevention of Routine Errors in the Rouse Fault - Tracking Task

M. Masson

Eindhoven University of Technology
Safety Management Group
Graduate School of Industrial Engineering and Management Science
P.O. Box 513, Pav. U-5, 5600 MB, Eindhoven, The Netherlands
Email: M.Masson @BDK.TUE.NL

Abstract

Developing a prototype of a computerised system aimed at preventing routine errors in tasks favouring routinisation, called CESS (Computerised Execution Support System), is a project started at the University of Liège and continued at the Eindhoven University of Technology. While the original system uses a *"cognitive"* model of the user, theoretically and historically related to COSIMO, the version currently under development uses a less complex but computationally more efficient *"engineering"* approach to error prevention. The paper addresses the research context that has motivated the development of these prototypes. It then summarises the original CESS structure and functions and briefly reports the results of a first experiment using a simple laboratory task. It then focuses on the objectives and the current development stage of a second research phase, in which the new system will be tested on the Rouse 1 fault diagnosis task.

1. Introduction

Routine errors are committed in domestic as well as in work environments. The nature of the risks they bring about does however not depend first and foremost on their nature. Indeed, as any kind of predictable errors, they all present similarities on the psychological level and share common psychological and, more particularly, *cognitive* mechanisms. The main difference lies indeed in the degree of tolerance of the environment, which can, depending on latent risks and available safety barriers, make of a routine error a pure harmless act or the triggering condition of a catastrophy.

From an ergonomical point of view, the ideal would be to preserve or indeed encourage the conditions that favour skill acquisition, all the while preventing routine errors. Errors are thus dealt with neither by suppressing nor by modifying the operator but rather by acting on the characteristics of the environment. This philosophy of assistance recommends to maintain opportunities of exploration and learning, all the while developing safety envelopes (e.g. Wiener, 1985), designed to prevent, filter or manage errors. Such an orientation requires a technological intelligence capable of adapting itself to man or the situation, which implies resorting, in some stage of the design, to models of the user and/or of the situation. Systems that possess such capacities belong to the class of *adaptive systems*, most of which demonstrating also capabilities of anticipation (see Masson, 1995).

Developing a prototype of an adaptive system aimed at preventing routine errors in tasks favouring routinisation is a project started at the University of Liege and continued as a post-doctoral research at the Eindhoven University of Technology. While the original system called CESS (Masson, 1992, 1994) uses a "*cognitive*" model of the user, theoretically and historically related to COSIMO (e.g., Masson, 1991, 1994; Cacciabue et al., 1992; Decortis, 1992), the version currently under development uses an "*engineering*" approach, which constitutes a less complex but computationally more efficient version of the system.

The paper addresses the theoretical research context that has motivated the development of these prototypes. It then summarises the original CESS structure and functions and briefly reports the main results of a first experiment. It then focuses on the objectives and current development stage of the research project running at Eindhoven.

2. Human failures and routine errors

In the areas of high-risk activities, such as aeronautics, electrical power, steel industry, chemistry and petrol industry, errors can lead to human, ecological and economic catastrophes. This means material damages, environmental degradations and economic disasters, as well as the loss of human lives - all consequences that are simply unacceptable. Such catastrophes as Three Mile Island, Bhopal, Challenger, etc., invite the scientific community and civil authorities to invest more effort in undestanding the very nature of human error and to develop innovative error prevention and/or management and recovery measures (e.g., Frese, 1991; Rasmussen, 1994; van der Schaaf, 1995) aimed at improving safety.

Psychologists indeed emphasise the need for maintaining human operators in complex systems because of their combined capacity to deal with unpredictable situations, to manage crises and conflicts, to compensate for the deficiencies of information systems and procedures, and to *acquire expertise* through interaction with the systems they control or supervise (Masson & De Keyser, 1993). For a human operator, acquiring expertise involves not only the acquisition of knowledge, models, theories or methods regarding his work environment, but also hierarchies of interpretation and action schemes, know-how and physical or cognitive *automatisms* that allow him to adapt himself at a low cost to the environment, and to reduce its complexity. Skill acquisition is a highly adaptive mechanism. It provides the operator with cognitive and behavioural structures that guide interpretation, prepare action, serve as a support to anticipation and retrospection, and reduce workload (e.g., Bainbridge, 1989). There is, however, a "price to pay" for these advantages (Reason & Embrey, 1980; Reason, 1990) to the extent that skilled activity induces a risk of capture by habit, that is to say of *routine error*.

Routine errors, as well as fixation errors they sometimes induce (e.g., De Keyser & Woods, 1990; Masson, 1991, 1994), are among the errors that industry meets the most frequently. They also belong to the most difficult to erradicate. They have indeed sometimes been qualified as "diabolic errors", because they resist to most of the given advices and classical training sessions, their actors knowing in principle what to do and *not* to do. This research tries to come up with a reply to the former paradox: promoting skill acquisition while preventing routine errors. The kind of counter-measure proposed here, namely automatic prevention through a situation-sensitive computerised system, is surely not the only possible one, but it might be promising for industrial applications. That is the reason why we think it deserves further research.

3. A theory of routine errors

More than one hundred years ago, William James (1890) already noticed that the successful practice of any type of activity results in the gradual delegation of control from a high-level, closed-loop and attention-driven control mode, to a low-level, open-loop, and schema-driven processing mode, characterising *routine* behaviour. As illustrated by Reason in a video tape on human error (Reason and Embrey, 1980), devolution of control results in the setting up of semi-autonomous cognitive processors that progressively acquire a part of autonomy in the release and control of activity. The more we get skilled by the repetition of the same piece of activity, the more this activity brings into play such semi-autonomous knowledge and control structures. With acquiring skill, a substantial part of the motor as well as of the cognitive activities thus become governed by low-level processors, which are hierarchically organised in long-term memory. They operate mostly locally, in environments triggering them through local calling conditions (Norman, 1981).

Conscious attention normally controls the selection of these low-level processors, together with the order in which they are brought into play. Conscious attention is also used to check the quality of activity, by comparing it to assigned set-points. High-level control takes place in working memory and involves attentional checks upon activity progress, that is upon the running of the lower-order schemes involved. However, this control function can sometimes be inappropriately exercised, being either ommitted or performed too early or too late (Reason, 1987, 1990).

The control of activity thus implies two control modes: the attentional or consciously directed mode, and the schematic or automatic mode (e.g., Schneider and Shiffrin, 1977; Norman and Shallice, 1980). These two modes operate in very contrasting ways: the former demands attention, is resource-consuming and is performed in a sequential way, while the latter requires nearly no effort, saves resources and is mainly parallel.

The process of routinisation thus presents the double advantage of reducing mental workload (Bainbridge, 1989) and of freeing attentional resources from the situation at hand. Following Reason (1987, 1990), there is, however, *a drawback* to this mostly adaptive mechanism, as routines are sometimes automatically activated in an inappropriate way or on an inappropriate object, simply because they were *frequently* or recently used in the past, or because they are triggered by an environment *similar* to environmental descriptions held in memory. Erroneous activation of normally highly adapted and efficient activity segments is the mechanism that leads to routine errors or, more precisely, to capture errors (Norman, 1981, 1990; Reason, 1990; Hollnagel, 1991).

4. Computer support to routine error prevention

Most of the support systems developed in the domain of man-machine interactions (see Woods, Johannesen, and Potter, 1990) tend to support activities such as diagnosing, planning, estimation and regulation, that mainly resort to problem solving. But as pointed out in particular by Rasmussen (1987), a substantial part of operators' activities *do not resort to problem solving but to know-how and skill.* Problem-solving oriented support systems would thus ideally need to be associated with other types of assistance, dealing with more routinised activities, as can be found in aeronautics, power plants (Masson, Malaise, Housiaux, and De Keyser, 1993) or train and car driving (Kruysse, 1992).

Designing such a system aimed at providing assistance in routinised activities is the purpose of this research. Its methodology of development is summarised as follows. Some of the human errors are best understood as a consequence of the dual nature of cognition. They thus can be both *explained* and *predicted* in the light of that theoretical cognitive framework. Among all human errors, routine errors are probably the most common. They also are the best *understood* and thus also the most *predictable*. That is the reason why they are called “systematic errors”. Like any highly predictable phenomenon, routine errors can thus be *modelled* and eventually *simulated*, on the basis of such a model.

Two main-error prevention designs are investigated in this project.

1. The first design consists of integrating a routine error simulation tool in the support system, in order to anticipate the occurrence of errors and remedy them in a preventive way (Masson, 1992, 1994). Because the routine error simulation system tends to duplicate *some* of the

properties of human cognition - in the way COSIMO simulates the process of knowledge extraction in operators, we call it a *"cognitive system"*.

2. The second design consists of approximating the performance of the former cognitive system, using a simpler and less computer power demanding algorithm. Because such a design does not claim to duplicate any property of human cognition, but only to take advantage, in an indirect way, of what is known about routine errors, we call it an *"engineering system"*.

5. Routine error prevention using cognitive simulation

Developing and testing an original support system dedicated at preventing and/or correcting routine errors using cognitive simulation was the objective of a first research phase (Masson, 1992; 1994). A software prototype called CESS (*Cognitive Execution Support System),* was programed in LISP on a SUN computer and tested in the laboratory with a very simple task favouring routinisation: the transcription of series of letters into corresponding series of digits, using an alpha-numerical code.

5.1. Functioning of the system

CESS works by anticipating in a sketchy manner the way a user would act or react to a familiar task environment, by simulating in real time some basics of human routinised behaviour. CESS anticipates such a reaction on the basis of a combination of *similarity* and *frequency* criteria (Reason, 1987, 1990), in the way COSIMO simulates the process of knowledge extraction by the operator, in the first stage of complex accident management (Caaciabue et al., 1992)

In such circumstances, the actions a subject is most likely to perform, or the segment of activity which is more likely to be activated, are determined both by the resemblance between the situation and similar situation exemplars framed in memory, what Reason (1987, 1990) calls the *similarity* criteria, and by the *frequency* distribution of these similar situations in the past. The cognitive system used in CESS accounts for the similarity factor by computing and updating, in the course of the activity, a similarity score between the current situation and situations records available in memory. This similarity score is then combined with a frequency score for each situation. The calculation and combination of these similarity and frequency scores - which, as an artefact, are carried out in a sequential way - gives each situation matched its final *activation level.*

When the resulting activation value is considered sufficient (by comparing it to a threshold used as a control parameter), the candidate obtaining the highest activation score is qualified for guessing the subject's answer. According to the "winner takes it all" principle, only the candidate with the highest activation score is taken into account for making this prediction.

This candidate is then evaluated against the current situation, for which an answer is known (or can be derived or approximated from another source). If this candidate appears to be erroneous, warning messages are sent by the preventive assistance module, *before* the user has carried out any action on the system. Such warnings tend either to modify an inappropriate perception or to correct, in a preventive way, the suspected erroneous answer.

Errors that would still occur in spite of this prevention system are processed by a corrective module, which gives the user an opportunity to modify his action before to be released in the task environment and bring about negative consequences. The corrective module thus operates like a filter, blocking any erroneous action and giving the subject some chances to rectify it [1].

5.2. Experimental testing

5.2.1. *Experimental task and procedure*

130 subjects (mostly students in psychology) aged between 18 and 30 were asked to transcribe 240 series of 9 or 10 letters into corresponding series of digits, using an alpha-numerical code. The letter series were presented one by one on the screen of a SUN computer. The answers were entered through the keybord and validated by the "return" key, which had also the effect of displaying the next letter series. The speed of progression was thus under the subjects' control, as no emphasis was put on time. The subjects were indeed asked to find a good personal balance between accuracy and speed, while performing the task within a "reasonable" time limit (about half an hour).

Routine errors were induced by controlling the frequency of occurrence and the similarity between a) a normal series, b) 10 single-change series, differing from the former only by 1 digit, and c) 2 double-change series in which an additional modification was introduced.

- the normal series "A B C D E F G H I" had a frequency of .95 (= 228 / 240);
- the 10 single-change series "A B C D E F G G H I" had a frequency of about 0.04 (=10/240), and
- the 2 double-change series "A B C D F G G H I" had a frequency of about 0.01 (=2/240).

---

[1] In a routinised activity, the subject very often knows the correct answer to the situation. The problem thus does not consist of building or discovering an initially unknown answer - which would resort to problem solving - *but not to use a strong answer, which turned to be wrong because of some unnoticed change in the situation.*

First, the normal series were presented 40 times, in order to make it "dominant" both in the user's and in the CESS' "mind", and to create some anchor effect. The single-change series were then inserted approximately every 20 presentations. The double-change series were finally presented as the two last modified series, at the end of the task.

5.2.2. Hypotheses

Four basic hypotheses were considered.

1. Series featuring changes are more propitious to errors, called here "routine"errors, than normal series.
2. Preventive and/or corrective assistance reduce the rate of such errors.
3. Double-change series are more error-prone than single-change series, and single-change series more than normal series.
4. The error reduction produced by the preventive and/or the corrective assistance is higher in double- than in single-change situations.

5.2.3. Results

These four hypotheses were globally supported by the data.

Presence and number of changes. Series featuring a double-change entailed a higher error rate than single-change series, and single-change series a higher error rate than normal series ($Pr > F = 0.0001$, $df = 2$).

Preventive assistance. The preventive assistance provided by CESS decreased the use and need for correction in the course of activity, but was *not* sufficient to eliminate all errors when used alone. Preventive assistance reduced in *absolute* terms the error rate in each situation type ($Pr > F = 0.0023$, $df = 1$), the reduction being higher for double-change series than for single-change series. Considered in *relative* terms, the reduction was, this time, higher for single- than for double-change series, demonstrating *the particular resistance of double-change series to error prevention.* However, as the interaction between preventive assistance, performance and type of series (i.e. number of changes) was not significant ($Pr > F = 0.1782$, $df = 2$), these last two results are only indicative.

Corrective assistance. The corrective assistance provided by CESS reduced in *absolute* terms the error rate in each situation type ($Pr > F = 0.0001$, $df = 1$), the reduction being higher for double-change series than for single-change series. Considered in *relative* terms, the reduction profile was once again inverted, that is higher for *single-* than for double-change series. This demonstrates *the particular resistance of double-change series to error correction.* Moreover, the interaction between corrective assistance, performance and type of series (i.e., number of changes) was here significant (Pr > F = 0.0016, df = 2). We also would like to notice that, with such a simple task, corrective assistance could also virtually have eliminated any error by forcing automatic correction - an option which was, for an obvious reason, not tested in the experiment.

In summary, these results confirm the particular error inducing character of situations featuring changes, when normal situations are massively encountered during the task. The error risk was also lower for single- than for double-change situations, *the latter being particularly resistant both to prevention and correction approaches. The number of errors committed in all error-prone situations was, however, significantly reduced by CESS, a result which is considered promising.*

Chronology of performance. The task was chronologically divided in 24 series of 10 presentations. Differences obtained when comparing the error rates between them were very significant (Pr > F = 0.0090, df = 23). When considering only the 12 situations featuring changes, the differences were also highly significant (Pr > F = 0.0001, df = 11).

Interpretation. These results are interpreted by a difference in the nature of the control processes brought into play during task performance. By reference to the theoretical framework presented in the beginning of the paper, the task can be considered as a succession of phases, which vary from a psychological viewpoint.

- The beginning of the task is a kind of *learning* phase, where most of the attentional resources are allocated to the transcription process. Some errors appear in the first problem presentations then disappear rapidly, because of the task's simplicity and its very low variety.

- The first presentations of the single-change series then entail an increase of the error rate, because they match or activate the strong-but-locally-wrong normal answer, thus triggering *routine errors*.

- In the following phase, the single-change series progressively acquire a status of *exception*. They are progressively identified as such and processed carefully, on a semi-attentional mode. Better mastered, these situations thus produce less errors, and even their detection seems also to routinise.

- Because they are erronously processed as something between "normal modified series", normal series and real two change series, the last two double-change series produce a final increase of the error rate.

Qualitative approach. The analysis of the nature of errors committed in each situation supports the former interpretation in terms of error mechanisms. In single-change situations, the success rate is approximately 90%, 70% of the errors being the answer given to the normal series. In double-change situations, the success rate drops to 64 % at the first presentation, 55 % of the errors (i.e., 20 % of the answers) being the answer prescribed in the single-change situations and 22 % the one given in the the normal situation. The success rate then increases up to 74 % at the second presentation, with 42 % and 12 % corresponding respectively to the answer prescribed in the single-change and normal situation; the other less frequent errors being a mix of those prescribed in the three situation types.

6. Routine error prevention using an "engineering" approach

6.1. Rationale for a new design

The main drawback of the former "cognitive" version of CESS was that it used a limitless long-term memory, in which all situations, answers and outcomes where extensively recorded. Such a requirement is unacceptable for a whole range of industrial applications, where computer power and memory are limited. That is the first reason why it was decided to redesign the system, while trying to keep most of its original efficacy.

The second reason was that modified situations, and particularly double-change situations, induced an increase in the variety of errors. Despite the extreme simplicity of the task used in the first experiment, *this variety highlights how difficult it is to predict with certainty the precise form of errors*, an objective which was mainly responsible for the complexity of the first design. *If only the error risk, that is the occurrence of the error-prone situations, has to be foercasted, then the system can be simplified* [2], without losing much of the initial efficacy.

Another type of modification was also introduced, as double-change situations were both very error-prone and particularly resistant to error prevention and error correction. Double-change situations probably thus require some dedicated type of support.

6.2. The new system

Processing in specific ways single- and double-change situations, or not, can itself becomes an experimental variable. To compare the two solutions, two different systems area being developed.

6.2.1. Error prevention not specific to the number of changes

This system processes without any distinction between single- and double-change situations.

---

2 This can be argued because the process of situation recognition or situation assessment is very simple with the kind of tasks used in this research. In other cases, it would on the contrary be necessary to increase the algorithm complexity, using, for example, *connectionist* techniques.

The system has *one* limited memory of the task situations (the "symptom sets" in the Rouse task, see below), the size of this memory being a parameter. The system warns the user when a) this memory is filled, b) only with the same record, and c) the current situation differs from it.

The policy for memory updating is "First In First Out"(FIFO), which by definition preserves the presentation order of inputs. Two phases can then be distinguished: initial *memory loading*, in which situations exemplars are recorded up to reach the memory capacity, and *memory updating*, which allows to acquire new situations exemplars while forgetting old ones.

6.2.2. *Error prevention specific to the number of changes*

This system distinguishes single- and double-change situations.

The system has *two* limited memories of the task situations, the size of which being parameters. The first one, called the Situation Memory (SM), records each and every new situation, while the second, called the Change Memory (CM) records modified situations only. The conditions for accessing the CM are the following ones: a) the SM has to be filled to capacity, b) it contains only records of the same situation, and c) the current situation differs from these SM records.

The system then specifically warns the user about the presence of a *double*-change situation when a) the CM it is filled, b) it contains only records of the same situation, and c) the current situation differs from but is *similar* to these CM records.

The policy for memory updating is in both cases "First In First Out" (FIFO), which preserves the presentation order of the encountered situations. Memory loading and updating once again keep the system aware of the situations' history.

6.3. Routine error prevention in the Rouse task

The task used used in the first experiment was over-*simplified* when compared to the tasks operators perform in real-work environments, even these that involve large segments of routinised activity. It was thus decided to further investigate the potentialities of CESS in a more valid task, the Rouse 1 fault finding task (Rouse, 1978).

6.3.1. *Objectives of this research phase*

The main objectives of this second research phase are:

- to gain insight into the factors determining routine errors during performance in complex tasks, such as fault diagnosis;
- to develop and test a prototype of a computerised support system aimed at preventing routine errors in such tasks;
- and to evaluate the possibilities for extending the above results, obtained in laboratory, towards pratical applications in the process industry.

6.3.2. *An adapted version of the Rouse task*

The Rouse 1 fault diagnostic task requires fault finding in graphically displayed networks of interconnected AND-gate components. The logics of a network and the main task characteristics are the following ones:

- all components are AND-gates; for a component's output to be 1, both the component itself must be working and all its input connections must be 1;
- the signals are transmitted from left to right and there is no feed-back loop;
- connections are restricted to components situated in adjacent columns;
- only one component fails in each network;
- there is no failure propagation nor possible degradation or repair during the task;
- the task is static;
- the diagnosis activity does not combine or compete with any other kind of task;
- the task is usually performed by a single subject.

The subject has to identify the faulty component, using different sources of evidence:

- the so-called "symptom set", a vector of binary 1/0 values available at the output of the network from the beginning of the task;
- and the results of tests made either on connections between components or on components themselves.

During task performance, a chronological list of the tests made and results obtained so far are displayed at the right-hand side of the network, until the subject has identified the faulty component, in which case the problem has been solved. The subject has to solve a series of problems presented on screen. The normal experimental procedure is the following one:

- each problem involves a new network;
- in each network, the connections between components are determined randomly;
- the faulty component varies also randomly in each network.

For a more extensive description of the task, refer for example to (Rouse, 1978).

These conditions are rather different from those process operators meet in real work settings, as they usually work with the *same* environment (e.g., in a control room) which give them opportunities to develop control skill and specialised knowledge about the system's structure, functioning, reliability and typical temporal evolutions (e.g., Masson and De Keyser, 1993; Masson et al., 1993; see also Patrick, 1993). These are indeed the conditions where large segments of semi-automated activity are developed and *routine errors* commited, either by misunderstanding or by inadequately mastering "nearly normal" situations.

To remedy this inconvenience and create conditions likely to induce more pattern-matching based problem solving in the Rouse task, the following modifications are introduced:

- all problems involve the *same* network;
- in this network, the connections are thus never modified;
- the distribution of faulty components and the frequency of failure are controlled.

In the first experiment, three types of situations were used: normal, single- and double-change situations. The same conditions are introduced in the Rouse task. Modifications always concern the *symptom set.*

Routine errors are expected to be induced by controlling the frequency of occurrence and the similarity between a) a normal symptom set, b) a single-change symptom set, and c) a double-change symptom set. However, as there is usually no one-to-one relation between symptom sets and failed components, there is thus one more factor to be controlled, the location of the faulty *components* in the network. Indeed to any symptom set generally corresponds a whole set of components which explain the signals got at the network output and which might thus have failed. This set of possibly faulty components is called the "Consistent Fault Set" or CFS (e.g., Toms and Patrick, 1987; Patrick, 1993; Brinkman and van der Schaaf, 1993).

6.4. Experimental testing

6.4.1. *Experimental task and procedure*

The distribution of the faulty components over the task is controlled. A problem is considered as finished when and only when the subject has proposed the correct diagnosis. In case an erroneous diagnosis is proposed, the subject has to go back to the task, until the right diagnosis is made. Such a task setting has the advantage to equalise the component failure rate for all the subjects.

The frequency of single- and double-change symptom sets and the component failure rate remain to be fixed during a pilot study. The number of problems and eventually the number of sessions needed to induce routinisation remains also to be fixed during this pre-session. The subjects of the experiment will be students in Industrial Engineering from the University of Eindhoven. In a later stage, the same experiment might, however, be proposed to process operators, in order to evaluate the robustness and generality of the results formerly obtained with students.

6.4.2. *Task profile, single- and double-change situations*

The temporal profile of the task is similar to the one used in the experiment on the letter series. Three task phases are distinguished.

1. The learning phase

The normal symptom set is presented a certain number of times, in order to allow skill aquisition and to induce routinisation. The CFS corresponding to this normal symptom set is called CFS0 (for 0 change). In the network presented in figure 1, the components that belong to CFS0 are components 1, 3, 4, 5, 7, 10, 11, 12, 13, 18, 19 and 26. The faulty component, called FC0, varies inside CFS0. It is expected that the less this variety, the easier will be the process of routinisation and the larger the risk of routine errors in the rest of the task.

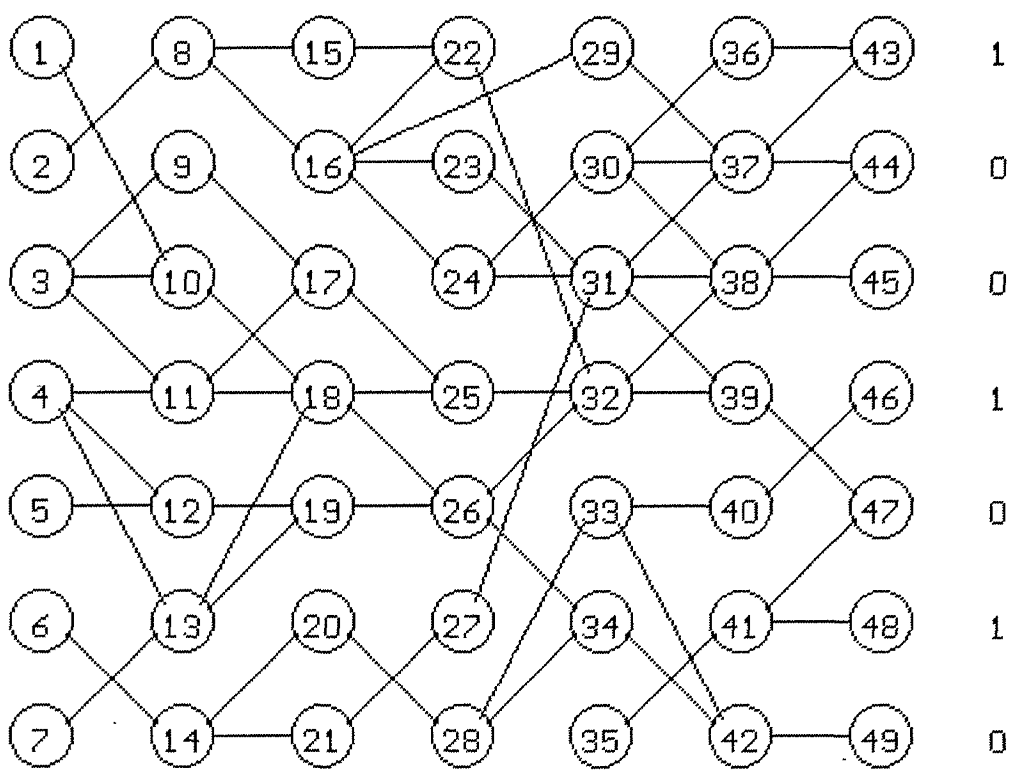


Figure 1. The normal situation. The symptom set is 1 0 0 1 0 1 0. The CFS is 1, 3, 4, 5, 7, 10, 11, 12, 13, 18, 19, and 26, and the faulty component varies inside the CFS.

2. Regular rare injections of a single-change symptom set

One single-change symptom set is inserted in the task in a nearly periodic way. To this single-change symptom set corresponds a consistent fault set, called CFS1. FC1 is kept always the same inside CFS1, in order to facilitate routinisation and favour routine errors.

In order not to allow the subjects to find it too easily by chance, FC1, component 15 in figure 2, will be located not too close to CFS0.

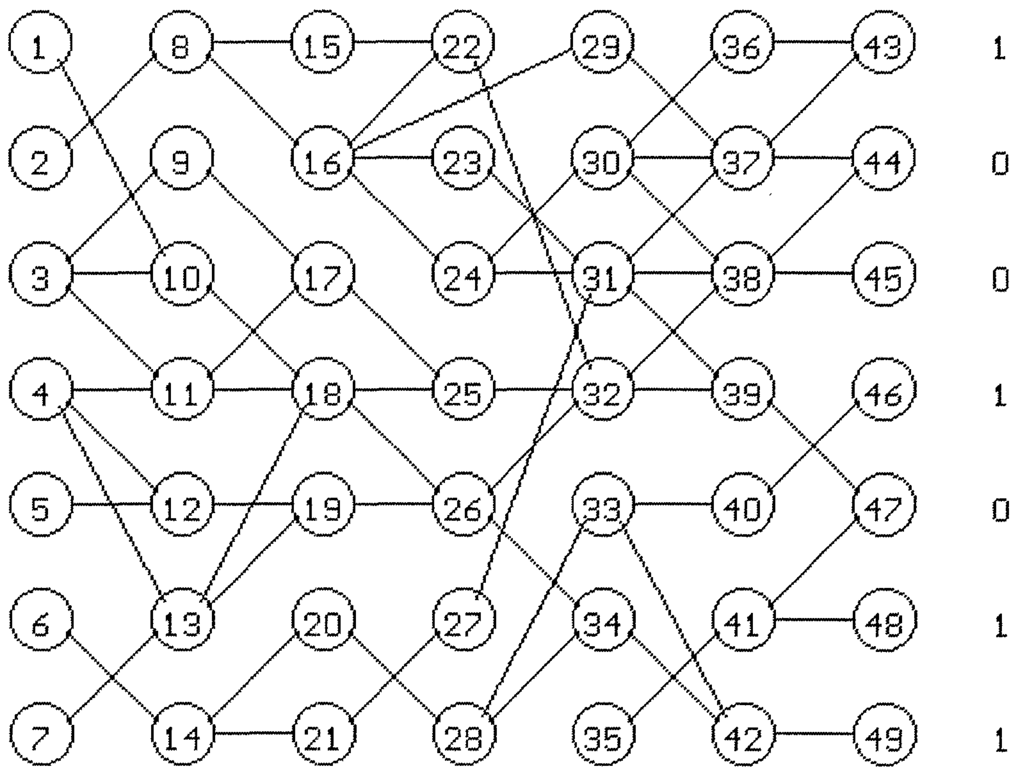


Figure 2. The single-change situation: the last digit of the symptom is now 1. The CFS is 9, 15, 17, 22, 25, and 32, and the faulty component is always component 15.

3. Two insertions of a double-change symptom set

One double-change symptom set will be inserted at the end of the task, instead of the (expected) single-change one. The consistent fault set corresponding to this double-change symptom is called CFS2. The faulty component, called FC2, will always be the same inside CFS2, in order to facilitate the analysis of performance in this particular hard to manage situation. To ensure that the subject won't find it too easily by chance, FC2, that is component

21 in figure 3, is located in a different network area than CFS0, and in a position opposite to component 15 (FC1).

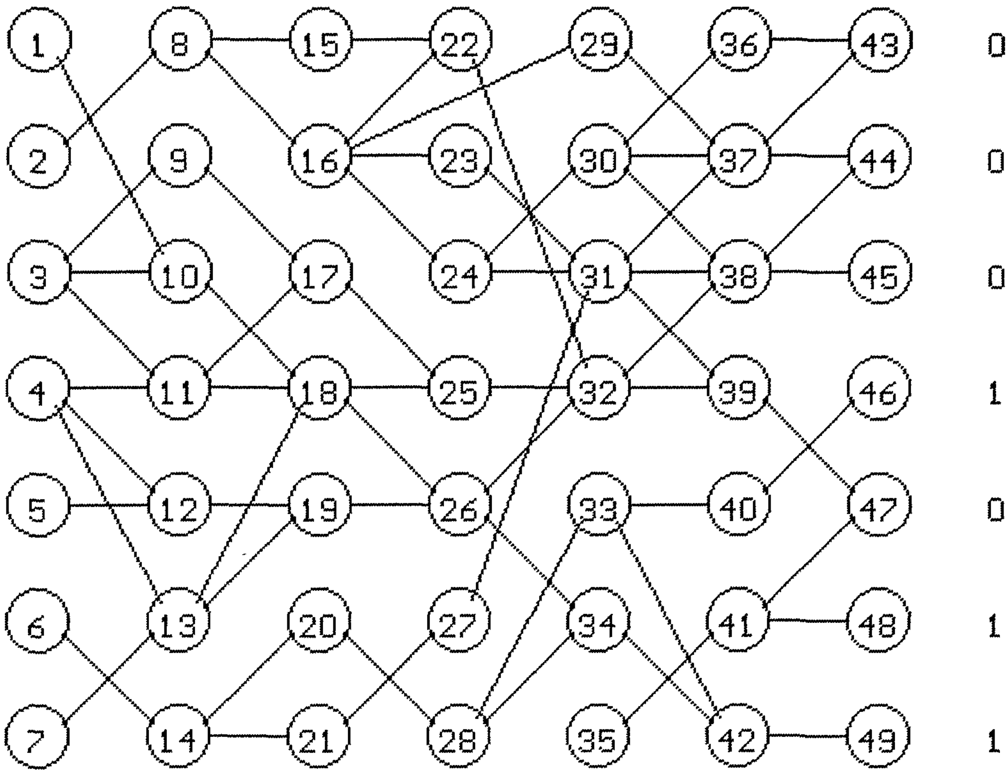


Figure 3. The double-change situation: the first digit of the symptom is now 0 and the last one remains 1. The CFS is 2, 8, 16, 21, 23, 24, 27, and 31, and the faulty component is always 21.

6.4.3. *Performance measurements*

By definition, as a good diagnosis is required in order to stop a problem, if the quality of the final diagnosis would be used as a measurement of performance, no error might ever be possible. *Sensitive performance measurements* are thus needed, like time to solve the task, number of tests made, number of premature and of incorrect diagnoses, number of redundant tests, and some kind of "deviation from optimality" variable (e.g., Brinkman, 1993). This particular performance measurement remains to be determined, but it is thought that information theory might provide an appropriate candidate.

6.4.4. *Additional experimental condition: visualising the CFS on screen*

According to Rouse and Hunt (1984), the main benefit of computer aiding lies in its ability to make full use of the available evidence and in particular of the 1 outputs (the 1 values in the symptom set). It is thus foreseen to provide the subjects with a system allowing to visualise the CFS on screen, as used for example by Brinkman and van der Schaaf (1993).

6.4.5. *Main hypotheses*

The basic hypotheses are similar to those of the first experiment:

1. Situations featuring changes are more propitious to errors, called "routine"errors, and to subsequent performance degradation, than the normal situations.
2. Preventive assistance provided by CESS reduce the rate of routine errors or decrease performance degradation.
3. Problems featuring double-change symptom sets are more error-prone than those featuring single-changes, and these more than normal situations.
4. The efficacy of the preventive assistance not specific to the number of changes is higher in single- than in double-change situations.

More particularly, in line with the results of the first experiment:

4.1. Efficacy is greater in *double-* than in single-change situations, when the reduction of performance degradation is expressed in an *absolute* way [i.e., performance with assistance - performance without].

4.1. Efficacy is greater in *single-* than in double-change situations, when the reduction of performance degradation is expressed in a *relative* way [i.e., (performance with assistance - performance without) / performance without assistance].

5. The efficacy of the preventive assistance specific to the number of changes should be the same in single- and in double-change situations. We, however, expect that this won't be the case, because of the very difficulty of double-change situations. An interaction is thus expected between performance, type of situation and type of assistance, despite the specific design effort dedicated to cope with double-change situations.
6. Visualising the CFS will improve performance in all types of situation.
8. Visualising the CFS will equalise the efficacy of *each* assistance type regarding single- and double-change situations.
8. Visualising the CFS will equalise the efficacy of *both* assistance types.

7. Conclusion

Routine errors are among the most common form of slips, both in everyday life and in work environments. They are also the most difficult to erradicate, their actors knowing usually what to do before the error is committed. This research is an attempt to solve this problem using an adaptive computer support, in tasks where such kind of assistance is possible.

A good ergonomical approach requires indeed to preserve or even to encourage the conditions that favour skill acquisition, all the while preventing routine errors. Errors are thus dealt with neither by suppressing nor by modifying the subject, but by acting on the environment, *by taking advantage of what is known at a theoretical level about errors.*

The task used in a first experiment, that is the transcription of a long sequence of short letter series into digits, produced promising results. This task is, however, over-*simplified* when compared to real operator tasks. It was thus decided to investigate the potentialities of the error prevention system in a more valid task, namely the Rouse 1 fault tracking task.

While the original system uses a very simple "*cognitive*" model of routinised activity, the version currently under development has adopted a less complex but computationally more efficient "*engineering*" approach to error prevention. One of the main results of the initial experiment was indeed that double-change situations were very error-prone and did also increase the variety of error forms. Consequently, and despite the extreme simplicity of the task used, this variety of answers demonstrates the difficulty to predict with certainty the precise form of errors, an objective which was mainly responsible for the "cognitive" flavour of the original design. If only the risk of routine errors has to be predicted without trying to forecast the error form, then a totally new design can be adopted. It is hoped that this new design will meet or even exceed the original system's efficacy, all the while surpassing it in terms of cost / efficiency balance.

Acknowledgements

This research initiated at the University of Liège, Belgium, is continuing thanks to a grant from the Eindhoven University of Technology, The Netherlands. The author wants to thank J.A. Brinkman and K. Kuipers for their the essential contribution to the second research phase. The presentation of this paper in Delft was supported by a CEC Human Capital and Mobility Network on Human Error Prevention.

References

Amalberti, R. (1992). Safety in process control: an operator-centred point of view, *Reliability Engineering and System Safety*, 38, 313, 99-108.

Bainbridge, L. (1989). Development of skill, reduction of workload, in L. Bainbridge and S. A. R. Quintanilla (Eds) *Developing Skills with Information*, Wiley, London.

Brinkman, J.A. (1993), Verbal protocol accuracy in fault diagnosis, *Ergonomics*, 36, 11, 1381-1397.

Brinkman, J.A. & Schaaf, T.W. van der (1993). Analysis and support of fault diagnosis strategies, in M.J. Smith & G. Salvendy (Eds) *Human-Computer Interaction*, 115-120, Elsevier, Amsterdam.

Decortis, F. (1992). *Procesus Cognitifs de Résolution d'Incidents Spécifiés et Peu Spécifiés en Relation avec un Modèle Théorique*, Thèse de doctorat en Psychologie, Université de Liège.

De Keyser, V. & Woods, D.D. (1990). Fixation errors: failures to revise situation assessment in dynamic and risky systems, in A.G. Colombo & A. Saiz de Bustamente (Eds) *Systems Reliability* Assessment, 231-251. ECSC, EEC, EAEC, Brussels and Luxembourg.

Cacciabue, P.C., Decortis F., Drozdowicz, B., Masson, M. & Nordvik, J.P. (1992). COSIMO: a COgnitive SImulation MOdel of human decision making and behaviour in accident management of complex plants, *IEEE-SMC*, 22 (5), 1058-1074.

Frese, M. (1991). Error management or error prevention: two strategies to deal with errors in software design, in H.-J. Bullinger (Ed.) *Human Aspects in Computing: Design and Use of Interactive Systems and Work with Terminals*, 776-782, Elsevier, Amsterdam.

Hollnagel, E. (1991). The phenotype of erroneous actions, in G. R. S. Weir and J. L. Alty, (Eds) *Human-Computer Interaction and Complex Systems*, Academic Press, London.

James, W. (1890). *The Principles of Psychology*, Henry Hott, New-York.

Kruysse, H.W. (1992). How slips result in traffic conflicts and accidents. *Applied Cognitive Psychology*, 6, 607-618.

Masson M. (1991). Understanding, reporting and preventing human fixation errors, in T.W. van der Schaaf, D.A.Lucas & A. Hale (Eds) *Near Miss Reporting as a Safety Tool*, 35-40, Butterworth-Heinemann, Oxford.

Masson, M. (1994). *Prévention Automatique des Erreurs de Routine*. Thèse de Doctorat en Psychologie, University of Liège.

Masson, M. (1995). Réflexions sur les systèmes adaptatifs d'assistance aux opérateurs, *Le Travail Humain*, (accepté pour publication).

Masson, M. & De Keyser, V. (1993). Human error: lesson learned from a field study for the specification of an intelligent error prevention system, *Advances in Industrial Ergonomics and Safety*, 1085-1092, Proceedings of the Annual International Ergonomics and Safety Conference, June, Denver, CO.

Masson, M., Malaise, N, Housiaux, A. & De Keyser, V. (1993). Organisational change and human expertise in nuclear power plants: some implications for training and error prevention, *Nuclear Engineering and Design*, 144, 361-373.

Norman, D.A. (1981). Categorization of Action Slips, *Psychological Review*, 88, 1-15.

Norman, D.A. (1990). *The Design of Everyday Things* (2nd ed.), Double Day Currency, New York.

Norman, D.A. & Shallice, T. (1980). *Attention to actio : willed and automatic control of* behaviour, CHIP 99, University of California San Diego, La Lolla.

Patrick, J. (1993). Cognitive aspects of fault-finding training and transfer, *Le Travail Humain*, 56, 2-3, 187-209.

Rasmussen, J. (1987). Cognitive control and human error mechanisms, in J. Rasmussen, K. Duncan and J. Leplat (Eds) *New Technology and Human Error*, Wiley, London.

Rasmussen, J. (1994). *Cognitive Engineering*, Wiley, London.

Rouse,W.B. (1978). Human problem solving performance in a fault diagnostic task, *IEEE-SMC*, 8, 258-271.

Rouse, W.B. & Hunt, R.M. (1984). Human problem solving in fault diagnosis tasks, *Advances in Man-Machine System Research*, 1, 195-222.

Reason, J.T. (1987). Generic Error Modelling System (GEMS): a cognitive framework for locating common human error forms, in J. Rasmussen, K. Duncan and J. Leplat (Eds) *New Technology and Human Error*, Wiley, London.

Reason, J.T. (1990). *Human Error*, Cambridge University Press, Cambridge.

Reason, J.T. and Embrey, D. (1980). *Human Error. A Fall from Grace*, BBC video tape, Reliability Associates (Ed.), Signal Vision, Knutsford, Cheshire.

Schaaf, T.W. (1995) van der. Human recovery of errors in man-machine systems, *Proceedings of the IFAC MMS'95 Conference*, June, MIT, Cambridge, MA, Elsevier, Oxford (in press).

Schneider, W. & Shiffrin, R.M. (1977). Controlled and automatic human information processing: I. Detection, search and attention, *Psychological Review*, 84, 1-66.

Toms, M, & Patrick, J. (1987). Some components of fault-finding, *Human Factors*, 29(5), 587-597.

Wiener, E. (1985). *Cockpit automation: in need of a philosophy*, SAE Technical Paper Series, 851956, NASA Ames Research Center, Moffett Field, CA.

Woods, D.D., L. Johannesen & S.S. Potter (1990). *Structured Bibliography on Human Interaction with Intelligent Systems*, CSEL Tech. Report 1990-009, The Ohio State University, Columbus.

Session 7

Robotics and Prostheses

Chairman: P. Breedveld, The Netherlands

The Development of a Predictive Display for Space Manipulator Positioning Tasks

Paul Breedveld

Delft University of Technology, Faculty of Mechanical Engineering & Marine Technology, Laboratory for Measurement & Control, Man-Machine Systems Group, Mekelweg 2, 2628 CD Delft, the Netherlands.

Abstract. Currently, the space manipulator ERA (European Robot Arm) is being developed in Europe. It will be used on the International Space Station Alpha. When the ERA is manually controlled, its movements are monitored with camera's, like the end-effector camera, which is used for positioning the 'hand' of the robot. When the operator controls the ERA, he/she has to cope with slow and flexible dynamics. At the Delft University of Technology, research is carried out on the development of a man-machine interface for the ERA. The movements of the ERA are animated on a graphical workstation, and the input device is a Spaceball controller. This paper briefly discusses a novel graphical display that uses intersections between 3-dimensional objects to improve the spatial observability of the end-effector camera picture, and a real-time simulation model, that accurately simulates ERA's movements. Next, the paper describes the extension of the display to a novel *predictive display*, that compensates for the effects of the slow dynamics and the flexibility. Finally, the paper discusses the suitability of the predictive display to compensate the time delays when a space manipulator like the ERA is controlled from Earth. Some try-outs with the predictive display have shown, that it is very suitable for manually positioning the end-effector.

Keywords. Man/machine systems, Man/machine interfaces, Manual control, Telemanipulation, Telerobotics, Graphic displays, Predictive displays

1. INTRODUCTION

Around the end of this century, the first segments of the new International Space Station Alpha will be put into orbit around Earth. On the Russian part MIR-2 of the Alpha, a European *space manipulator* will be used. The manipulator is called ERA, which is short for 'European Robot Arm', and shown in Fig. 1. It is a 10 m long anthropomorphic manipulator with 6 degrees of freedom (DOF's), that is being developed and assembled at Fokker Space and Systems B.V. in Leiden, the Netherlands. The joints of the robot are driven by electric brushless DC motors, and its links are made of lightweight carbon fibre. Even though this is a very stiff material, the long and slender upper- and forearm of the ERA are flexible. Since the power of the electric motors is limited, however, the arm can only move slowly, and the vibrations will be small. The 'hand' of the ERA is called *end-effector*. On the end-effector front, a gripper is placed, which is used for grasping an object.

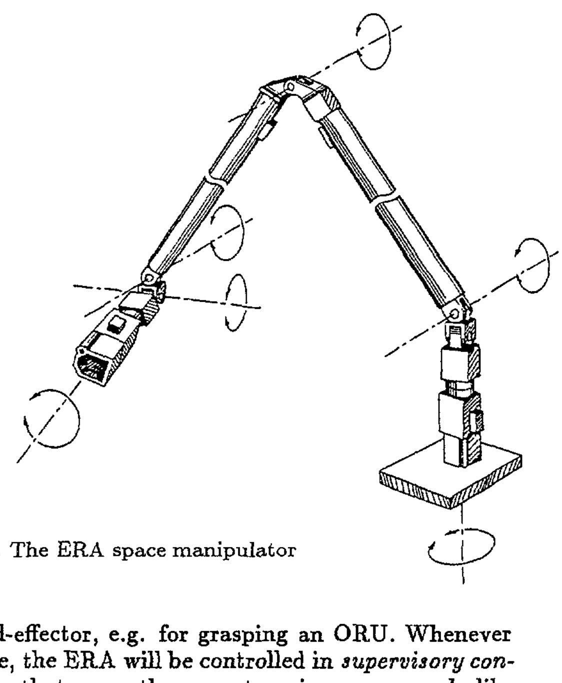


Fig. 1. The ERA space manipulator

On the MIR-2, the ERA will be used for servicing and inspection tasks and for the displacement of Orbital Replaceable Units or ORU's. These are containers with objects that must be moved from one part of the space station to another. During these activities, the robot will be controlled from inside the MIR-2 by an astronaut. Due to a lack of direct vision on the MIR-2, *camera pictures* will be used to monitor the movements of the ERA. The camera's are located on the manipulator itself and on the space station. An example of a camera, located on the manipulator itself, is the *end-effector camera*, that will be used for fine positioning the end-effector, e.g. for grasping an ORU. Whenever possible, the ERA will be controlled in *supervisory control*. In that case, the operator gives commands like 'put that there' to the manipulator, and the robot performs these tasks automatically. The operator uses the camera pictures to *monitor* the movements of the manipulator. If anything goes wrong, the operator intervenes, and switches to *manual control*. In that case, he/she uses a *control device* to move the manipulator into the desired direction.

In the case of manual control, the human operator is

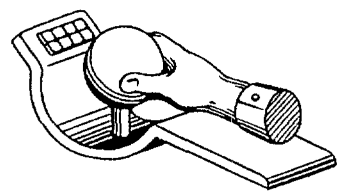


Fig. 2. The Spaceball controller

faced with three problems, that make controlling the ERA difficult:

1. The ERA is a multivariable system with a large freedom of movement. It is difficult for a human to control such a system manually;
2. It is difficult for a human to estimate the 3-dimensional movements of the manipulator from flat, 2-dimensional camera pictures;
3. The slow dynamics and the flexibility make controlling even more difficult.

At the Delft University of Technology, research is carried out on the development of a man-machine interface for manual control of a space manipulator like the ERA, and to find solutions for these problems. Besides space manipulators, the results of the project are also suitable for other applications, like endoscopic surgery, undersea robotics, and robotics in the nuclear industry [Sheridan, 1992].

The hardware facility consists of a fast Silicon Graphics Indigo II Extreme *graphical workstation* and a *Spaceball controller*. The graphical workstation is used to simulate the movements of the manipulator, and to animate the camera pictures; both in real-time. The Spaceball controller, Fig 2, is a 6 DOF force operated control device that consists of a sphere on a base. The sphere, or Spaceball, can be slightly translated and rotated in 3 perpendicular directions. The forces and torques the operator imposes on the Spaceball are transformed into translational and rotational velocities of the end-effector. The choice for end-effector velocity control and the selection of the Spaceball controller as control device for this project is motivated in [Breedveld, 1994]. Instead of velocity control, also end-effector *position control* could have been used. In [Kim, 1987], however, it is shown that position control is more suitable for *fast* manipulators with a *small* workspace, while velocity control is more suitable for *slow* space manipulators with a *large* workspace like the ERA. *Joint control* didn't seem to be interesting to investigate, since, for a human operator, this way of controlling is much more difficult to do.

The research of the project is phased as follows:

A. Activities in an operating point:
   1. Manual control of an ideal manipulator;
   2. Manual control of a flexible manipulator.
B. Activities along a track

In stage A, which is focused on positioning the end-effector, e.g. for grasping an ORU, the demanded accuracy is high, but the risk of an unexpected collision between the manipulator and the environment is small, since the movements of the manipulator are small. In stage B, which is focused on moving the end-effector along a track, e.g. for displacing an ORU, the demanded accuracy is small, but the risk of a collision

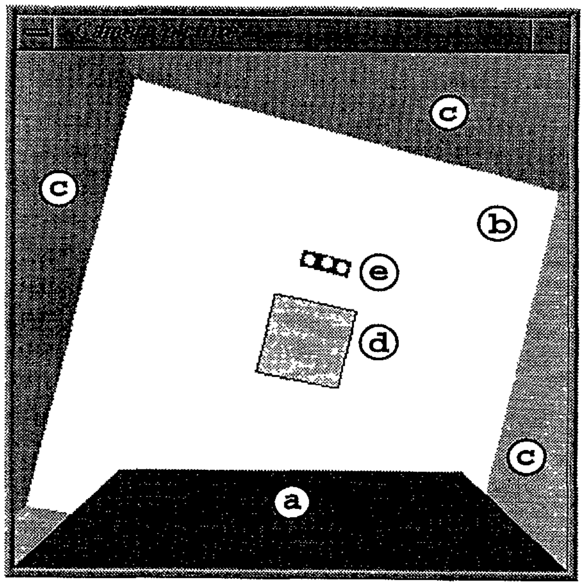


Fig. 3. The end-effector camera picture. *(a) End-effector (b) backplane (c) sideplanes (d) target (e) vision target.*

is large. In phase A1, the first and second problem mentioned above are investigated, and the manipulator dynamics are not implemented yet. In phase A2, the third problem is investigated, and the dynamics of the manipulator are modelled in an accurate real-time simulation model.

The developments in phase A1 of the research are described in [Breedveld, 1995 & Buiël, 1995a,b]. The solutions for the second problem mentioned above in this phase of the research are briefly discussed in section 2. The solutions for the third problem in phase A2 of the research, that have resulted in a *predictive display*, are described in sections 3 and 4. In section 5, the suitability of the predictive display to compensate the time delays when a space manipulator like the ERA is controlled from Earth is discussed. The paper ends with some concluding remarks in section 6.

2. DISPLAYS FOR POSITIONING THE END-EFFECTOR

The end-effector camera picture

When an object must be grasped, the end-effector camera will be used for fine positioning the end-effector. Fig. 3 shows a stylized impression of the end-effector camera picture, when the distance between the end-effector front and the object to be grasped is decreased to 1 m. The end-effector camera is placed *on top* of the end-effector and shows a view of the environment. The end-effector front is visible at the bottom of the picture. The environment is simplified to five planes. The *backplane* represents the object to be grasped. The four *sideplanes* represent the environment around the object. On the backplane, a *target* is present. This is a connection point on which the end-effector front must be accurately positioned with a desired accuracy of 3 mm and 1.5 $^{\circ}$, to grasp the object. Above each connection point on the MIR-2, a *vision target* is present. This object consists of a black base plate with three white discs, from which the middle one is placed on an elevation. It is used by a *proximity sensor*: a computer program that uses image processing techniques to calculate the distance and orientation of the end-

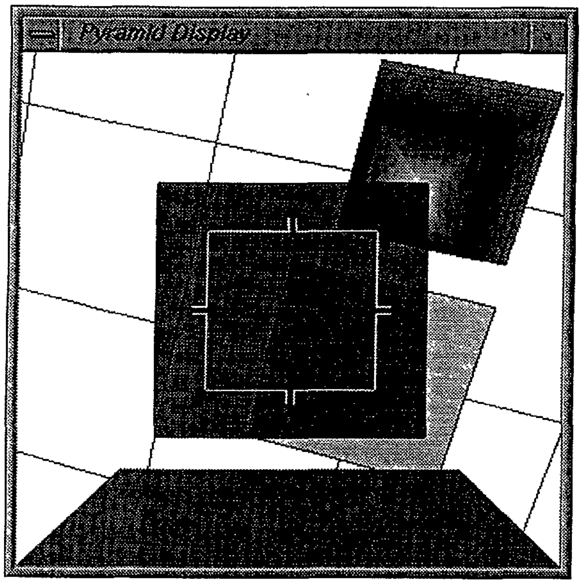


Fig. 4. Pyramid Display, at a distance of 50 cm from the target.

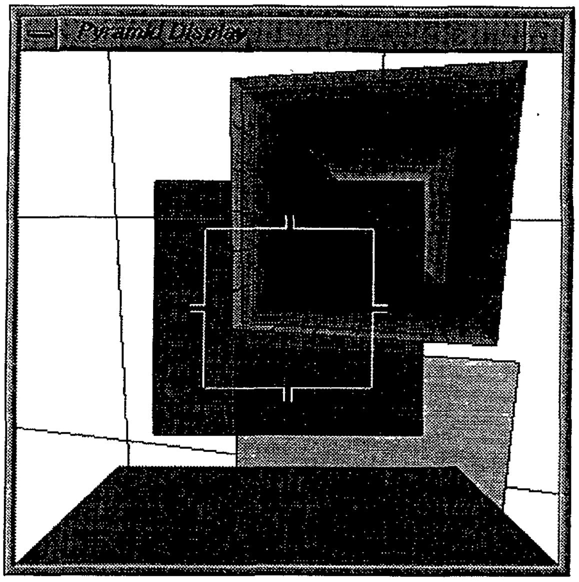


Fig. 5. Pyramid Display, at a distance of 5 cm from the target.

effector with respect to the target from the camera picture. This distance and orientation is described with six numbers: 3 translations and 3 rotations along the principal axes. These numbers can be used to position the end-effector on the target automatically. The slow dynamics and the flexibility of the manipulator, however, make automatic positioning difficult. In some cases, e.g. if an object is present between the end-effector and the target, or if the automatic controller is damaged, automatic positioning might even be impossible. Therefore, it is also possible to position the end-effector by hand.

Graphical overlays on the end-effector camera picture
In the last stage of the positioning task, the distance between the end-effector and the backplane becomes so small, that the four sideplanes get out of sight. Then, only the target and the vision target remain to give spatial information to the operator. Unfortunately, for a human, it is very difficult to estimate the position and orientation of the end-effector from the size and the locations of the three white discs. To assist the operator, the proximity information can be used to superimpose a *graphical overlay* on the picture that makes it more easy to position the end-effector.

In [Bos, 1995], a graphical overlay on the end-effector camera picture is described, in which the proximity information is used to animate six analogue dial indicators, that display the six translational and rotational misfits *individually.* In this project, the misfits are not displayed individually, but *integrated* in a novel overlay that consists of 3-dimensional graphical objects. Due to the integration, the operator can estimate the translational and rotational misfits *at a single glance.* The overlay is constructed in such a way, that the animated objects seem to be *part of the environment.* Due to this, it *increases* the operators involvement with the real situation. The working principle of the overlay, called the *Pyramid Display,* is briefly described below. In [Breedveld, 1995], the development of this display is discussed in detail, and some further improvements of the display are described.

The Pyramid Display
Figs. 4, 5 and 7 show the Pyramid Display, at a distance of 50 cm, 5 cm and 0.5 cm from the target respectively. To improve the sensation of depth, the proximity sensor information is used to project rectangular grids on large unicoloured planes in the environment. Besides the grids, the proximity sensor information is also used to animate three 3-dimensional graphical objects: an *insert-box*, a *frosted-glass* and a *pyramid.* The insert-box and the pyramid are fixed to the backplane, at the location of the vision target, which is therefore not visible anymore. The frosted-glass is a transparent square plane, fixed to the end-effector front, like the sight of a gun. In the first phase of the task, Fig. 4, the operator uses the frosted-glass as a viewfinder and moves it towards the pyramid. The rotational misfit of the end-effector can then easily be detected from the orientation of the box. In the second phase of the task, Fig. 5 and Fig. 6 from aside, the frosted-glass *intersects* the insert-box. Then, the rectangular grid in the box is used as a distance indicator, perpendicular to the backplane, and the remaining distance can easily be estimated. In the third phase of the task, Fig. 7, when the distance left to the backplane is decreased to only a few millimeters, the frosted-glass intersects the pyramid. The *size* of this intersection then visualizes the distance left to the backplane. The larger the intersection, the smaller this distance. The *shape* of the intersection visualizes the rotational misfit perpendicular to the backplane. If a

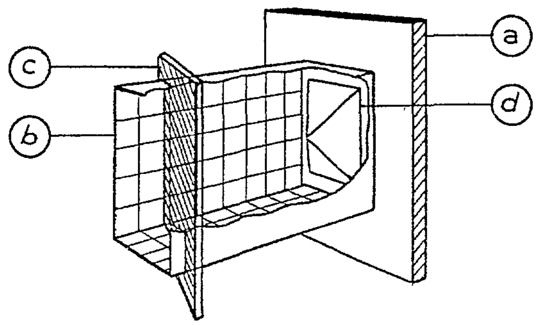


Fig. 6. Side view of the insert-box, the frosted-glass and the pyramid. *(a) Backplane, (b) insert-box, partially cut away, (c) frosted-glass, and (d) pyramid.*

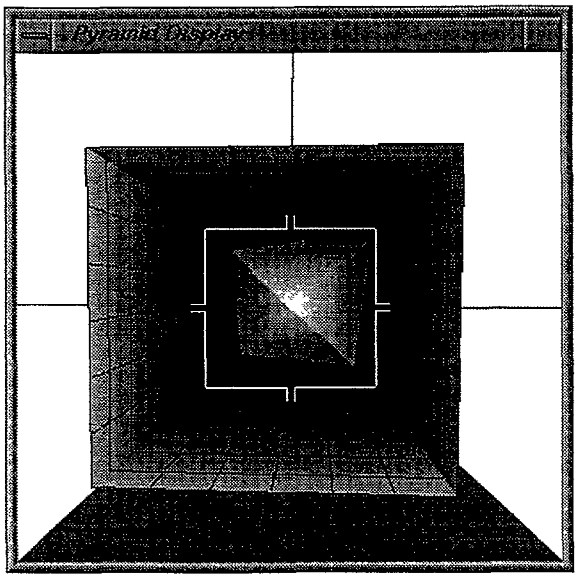


Fig. 7. Pyramid Display, at a distance of 0.5 cm from the target.

rotational misfit is present, the intersection is a trapezium or a kite, like in Fig. 7. If all rotational misfits are zero, the intersection is a square. When the intersection finally coincides with the square on the frosted-glass, the front of the end-effector is positioned against the backplane and the positioning task is completed. The intersection with the pyramid is also used to avoid collisions against the backplane: a collision occurs, when one of the four vertices of the intersection touches the base of the pyramid, like in Fig. 7.

The graphical objects of the Pyramid Display have contrasting colours that make them clearly distinguishable. The grids are coloured black, the insert-box blue, the frosted-glass transparent black, and the square on the frosted-glass bright yellow. The colour of the pyramid is given the function of *accuracy indicator.* This was done like a traffic light: In general, the pyramid is drawn bright red; an alarming colour, that enforces the operator to decrease speed. When all translational and rotational misfits, except for the distance left to the backplane, have reached their desired accuracies of 3 mm and 1.5 °, the colour of the pyramid changes to bright yellow. When the misfits are made even smaller, its colour changes to bright green, and the end-effector can be positioned safely. The yellow colour functions as a kind of security margin.

In phase A1 of the project, a number of man-machine experiments have been performed with the Pyramid Display. The experiments proved the display to be very handy for positioning the end-effector. After a relatively short learning phase of about 3.5 hours, almost all subjects turn out to be able to position the end-effector with surprisingly fast completion times between 17 and 34 s [Buiël, 1995b].

3. THE EFFECTS OF THE DYNAMICS

The real-time simulation model of the ERA

If the ERA is manually controlled, the operator uses the Spaceball controller to generate the desired velocity of the end-effector. ERA's *automatic velocity controller* transforms his/her control command into control signals for the six electric brushless DC motors,

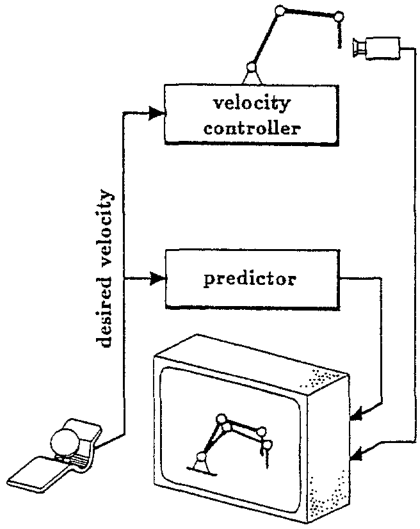


Fig. 8. The principle of a predictive display.

that make the end-effector move into the desired direction. In phase A1 of the research, the manipulator dynamics were not implemented yet. It was assumed that the ERA behaves *perfectly*, and that it responds immediately and exactly to a given control command. In reality, however, even with an optimized automatic controller, the behaviour of the manipulator is far from perfect.

In order to investigate the effects of the slow dynamics, the flexibility and the time delays in phase A2 of the research, the dynamics of the ERA were modelled in a real-time simulation model. The model is described in detail in [Diepenbroek, 1994]. The flexible dynamics of the ERA were modelled by using software of Fokker Space and Systems B.V. [Woerkom, 1994]. The electric motors and the automatic velocity controller were modelled in this project. The simulation model also includes a novel friction model, that simulates the effects of the friction in ERA's joints very accurately, but takes less computing time.

Some try-outs with the simulation model showed, that ERA's slow, flexible dynamics make manually controlling the manipulator difficult. This is mainly caused by the following two effects:

- Since it takes some time for the robot to accelerate, it is not immediately clear *in which direction the end-effector will move* when the operator moves the Spaceball;
- Since it takes some time for the robot to slow down, it is not immediately clear *at what position the end-effector will come to a standstill* when the Spaceball is released.

In order to reduce the risk of instability, the operator is forced to use a *move and wait* strategy; a cautious way of controlling, in which he/she generates a small incremental control command, then stops and waits for a few seconds to see what happens, then gives another small incremental control command, etc.

The principle of a predictive display

The two annoying effects above can be reduced by using a predictive display. The principle of such a display is shown in Fig. 8. The control commands are sent to the manipulator, and the movements of the manipulator are recorded with a camera and shown on a television screen in front of the operator. In order to compensate the two effects mentioned above, the control commands

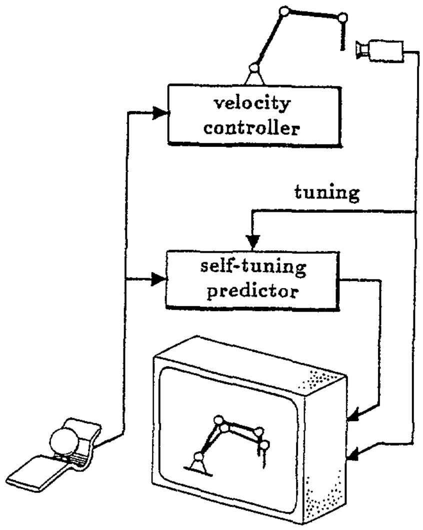


Fig. 9. A predictor with tuning.

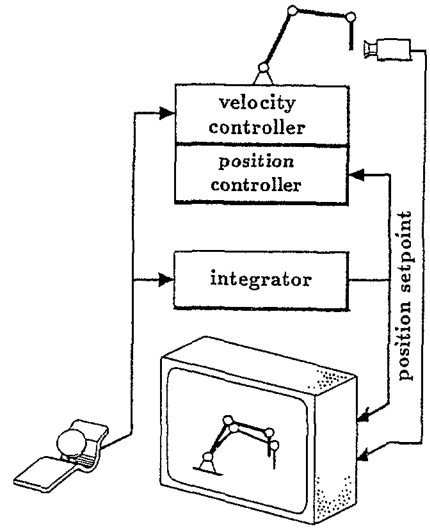


Fig. 10. A predictor that generates a position setpoint

are also used in a computer program, or *predictor*, that serves two purposes:

(a) it calculates in which direction the end-effector will move when the operator moves the Spaceball, and
(b) it calculates at what position the end-effector will come to a standstill when the Spaceball is released (this position will further be referred to as the *stopping position*).

The output of the predictor is used to overlay an image of a *ghost manipulator* on the camera picture, that initially coincides with the real one. When the operator moves the Spaceball, the ghost manipulator will detach from the real manipulator, which, in turn, will accelerate and follow the prediction. When the Spaceball is released again, the ghost manipulator stops moving, and the real manipulator will slow down, till finally, both pictures coincide again. If the prediction is accurate, and if the image is spatially easy to interpret, a predictive display can be very helpful for controlling the manipulator.

In phase A2 of the research, the principle above has been used to extend the Pyramid Display to a Pyramid Predictive Display. The development of this display will be described in the next section.

4. DEVELOPMENT OF THE PYRAMID PREDICTIVE DISPLAY

The structure of the predictor

If the operator controls the manipulator with a predictive display, he uses the stopping position to avoid collisions between the manipulator and the environment. In the literature, different ways have been used to calculate the stopping position in an accurate way.

In the predictor described in [Bos, 1991, pp. 88-121], the stopping position is calculated by using a simulation model with the dynamics of the manipulator. A simulation model, however, is always a simplification of reality. If such a predictor is used in practice, the predicted stopping position will always differ from the real one, and a *statical deviation* will be present between the final position of the real- and the ghost manipulator on the screen. This deviation can be reduced if the simulation model is made more accurately. However, a complex model takes much computing time. This is not desired, since the prediction must be calculated in real-time.

In the predictor described in [Breedveld, 1992], measured information from the camera picture is used to *tune* the predictor, to eliminate the statical deviation automatically (Fig. 9). If such a *self-tuning predictor* is used, however, the ghost manipulator will not immediately come to a standstill, when the operator releases the Spaceball, but it will slowly drift towards the real stopping position. In order to compensate this drift, the operator will have to generate more control commands to make the manipulator do what he/she wants. This is quite irritating, and a drawback of a self-tuning predictor.

Due to the problems with the two predictors above, in this project, it was not decided to use a predictor that describes the behaviour of the robot as accurately as possible, but to use the predictor as *setpoint generator* for the robot (Fig. 10). For the predictor, a simple *integrator* is taken, that integrates the velocity commands of the operator, and calculates a position setpoint. This setpoint is used to animate the ghost manipulator on the screen, and used in the automatic controller of the manipulator, that has been extended with an *automatic position controller*. This position controller eliminates the deviation between the setpoint and the real position of the end-effector automatically and moves the real manipulator towards the ghost.

Since the ghost manipulator now responds immediately and exactly to a given control command, the predictor serves both purposes (a) and (b) mentioned above. Since the predictor is a very simple integrator, that doesn't use a simulation model of the robot, no problems with the computing time are present anymore. Furthermore, a statical deviation in the final position of the real manipulator and the ghost is compensated by drifting the real manipulator towards the ghost, instead of the opposite, and the drawback of the self-tuning predictor is not present anymore.

A disadvantage of the strategy is, that the automatic controller of the manipulator must be extended with a position controller, which makes the system more complicated. However, the requirements on the position controller are not very large: if it eliminates the statical deviation within a reasonably short time, and if the overshoot is not too large, the operator can rely

Fig. 11. The pyramid (left) and the star (right).

almost completely on the position setpoint, and a simple automatic controller will meet the demands. Only when the manipulator moves very close to an object, the overshoot may cause a collision. In such a case, the operator cannot rely completely on the position setpoint, and he/she must also observe the movements of the real manipulator, and smoothen his/her control commands if it vibrates too much.

So, instead of using a very complicated simulation model, or a clumsy self-tuning predictor, a simple extension was made to the automatic controller of the robot, and the predictor is used as setpoint generator. With this strategy, the *total system*, which consists of both the manipulator ánd the predictor, has a simple structure and is easy to control. In spite of the extra automation, the movements of the manipulator are still completely manually controlled by the operator.

The visualization of the prediction

The predictive display in Fig. 10 shows a *2-dimensional side view* of a manipulator. The end-effector camera picture, however, shows a *3-dimensional front view* of the end-effector and the environment. In this case, the position setpoint must be visualized *spatially*, which is much more difficult.

In the predictive display of Fig. 10, the predictor integrates the velocity commands of the operator to a position setpoint for the end-effector *relative to the environment.* Since the camera picture in Fig. 10 only shows a manipulator, an image of a ghost manipulator was used to visualize the position setpoint. The end-effector camera picture, however, both shows the end-effector front ánd the environment. In this case, either an image of a *ghost end-effector front*, or an image of a *ghost environment*, can be superimposed on the picture to visualize the position setpoint. When the operator controls the end-effector camera picture, he/she focuses on the environment, and not on the end-effector front. Therefore, the ghost end-effector front seems to be less suitable to visualize the position setpoint, and the ghost environment is selected.

Note, that, instead of an image of a really existing object in the end-effector camera picture, also another, more abstract presentation could have been used to visualize the position setpoint. In the end-effector camera overlay described in [Ferro, 1994], for example, three analogue distance scales and a figure that looks like the artificial horizon of a plane, are used to visualize the real and the predicted misfits. In this project, however, it was decided to visualize the position setpoint in such a way, that the operator stays involved with the real situation as much as possible. Therefore, in the end-effector camera picture, a ghost environment is used to visualize the position setpoint, instead of a more abstract presentation, that would divert the attention of the operator from the real situation.

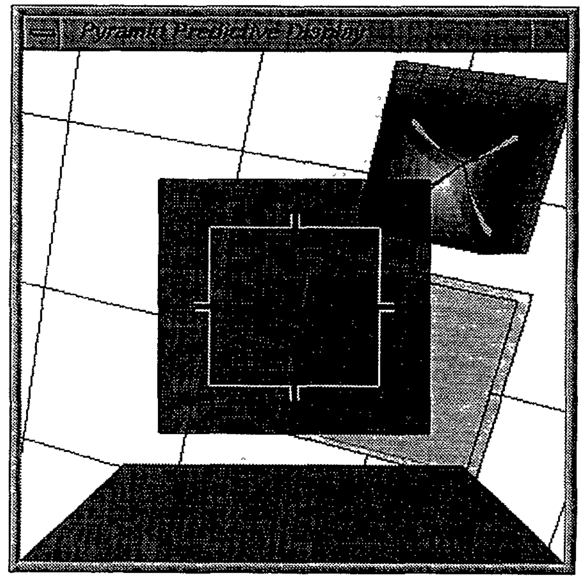


Fig. 12. Pyramid Predictive Display, at a distance of 50 cm from the target.

So, the predictive display shows both a ghost- and a real environment. In the predictive display, initially, the ghost environment coincides with the real one. When the operator moves the Spaceball, the ghost environment will detach from the real environment, which, in turn, will accelerate and follow the position setpoint. When the Spaceball is released again, the ghost environment stops moving and the real environment will slow down, till, finally, both environments coincide again. From the display, it must be easy to estimate the translational and rotational misfits of both environments. Due to the good results with the Pyramid Display in phase A1 of the research, the principle of this display is also used in the predictive display. In the Pyramid Display, an insert-box and a pyramid are projected in the environment, to visualize its position and orientation. If these two graphical objects are projected on both the ghost- and the real environment, however, the predictive display would contain 2 insert-boxes and 2 pyramids. This mixture of objects would look very crowded and confusing.

With the predictor of the preceding paragraph, the operator can rely almost completely on the position setpoint. The movements of the real environment are only important *at the end* of the positioning task, when the end-effector is very close to the backplane and the overshoot may cause a collision. In the Pyramid Display, the operator uses the intersection of the frosted-glass with the pyramid to avoid collisions against the backplane. The insert-box is not used for this purpose. This means, that, in the real environment, the insert-box can be omitted. This simplifies the predictive display to a ghost environment with an insert-box and a pyramid, and a real environment with only a pyramid. At the end of the positioning task, when both pyramids intersect the frosted-glass, the ghost pyramid is used to fine position the end-effector, and the real pyramid is used to avoid collisions. For the operator, however, the mixture of both intersections would still be difficult to interpret. Therefore, the display is simplified further.

In the Pyramid Display, a collision occurs, when one of the four vertices of the intersection on the frosted-glass touches the base of the pyramid (Fig. 7). So, the

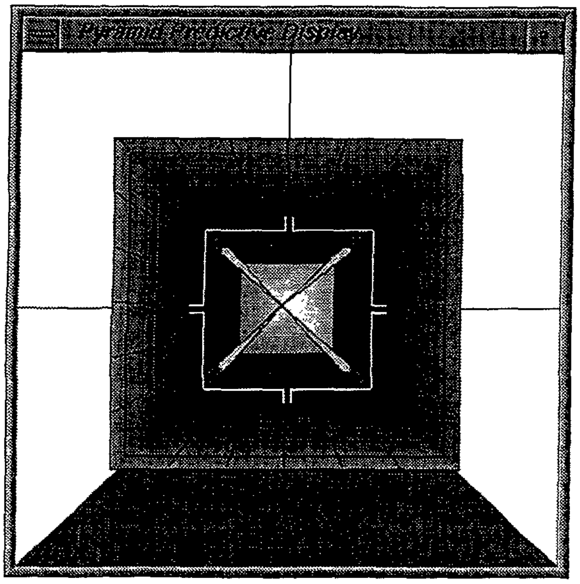


Fig. 13. Pyramid Predictive Display, at a distance of 0.5 cm from the target.

operator uses the *vertices* of the intersection to avoid collisions. This means, that, in the real environment, it is sufficient to draw only the *edges* of the pyramid. The object then simplifies to a *star* with four tips, as shown in Fig. 11.

The Pyramid Predictive Display

Figs. 12 and 13 show the finally resulting *Pyramid Predictive Display*, at a distance of 50 cm and 0.5 cm from the target respectively. The ghost environment is equal to the environment in the Pyramid Display and consists of the grids, a ghost target, the insert-box and the pyramid. The real environment consists of the back- and sideplanes, the real target and the star. The ghost target is drawn transparent, and the star is drawn with a *preference*: it is always visible, whether or not it lies behind the pyramid. The colours of the graphical objects are equal to the colours of the Pyramid Display. The accuracy indicator has been extended: the colour of the pyramid now changes from red via yellow to bright green, if the *position setpoint* reaches its desired accuracy, and the colour of the star changes from red via yellow to bright green, if the *real position* reaches its desired accuracy.

In the beginning of the positioning task, Fig 12, the operator moves the frosted-glass towards the insert-box of the ghost environment and uses the orientation of the box and its intersection with the frosted-glass to position the end-effector in front of the target. At the end of the task, Fig 13, both the pyramid and the star intersect the frosted-glass. Then, the operator uses the pyramid for fine positioning the end-effector on the ghost environment, and the star, to get an impression of the overshoot. If the intersection with the star vibrates a lot, the operator must be careful and slow down, till the vibrations of the manipulator are damped sufficiently. When both the pyramid and the star are coloured green, the end-effector can be positioned on the target safely.

Note, that in the last stage of the positioning task, the movements of the real environment, the predicted environment, and the accuracy indicator, are all visualized *in the middle* of the display, which is the location where the operator focuses on. Due to this, in the last stage of the task, when the end-effector must be fine positioned, he/she doesn't have to spread the attention to different parts of the display, which would be quite irritating.

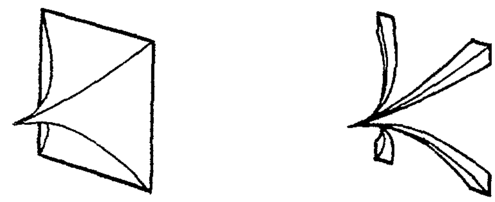


Fig. 14. The stretched pyramid (left) and the stretched star (right).

Improved shape of the pyramid and the star

In the display, the pyramid and the star have two purposes: they are used to visualize the resulting translational and rotational misfits, and they are used to avoid collisions. The first purpose requires, that the objects are *low*: the lower they are, the stronger they amplify the misfits when they intersect the frosted-glass. The second purpose, however, requires that the objects are *high*: the higher they are, the earlier they intersect the frosted-glass, and the earlier the operator can slow down the end-effector to avoid a collision.

In order to meet both conflicting requirements above better, in the Pyramid Predictive Display, the pyramid and the cross are *stretched* a little bit, as shown in Fig. 14. The low slope at the base of the stretched objects makes them very suitable to visualize the misfits, but their height still makes them very suitable to avoid collisions. However, the objects should not be stretched too much, since then, the size of the intersection changes strongly non-linear with the distance to the backplane, which is very confusing.

5. DISCUSSION

Prediction in the case of time delays

In future applications, space manipulators like the ERA will also be used on unmanned space vehicles, and manually controlled from Earth. Then, the control commands of the operator are transmitted into space and arrive at the manipulator with a time delay Δt up to about 3 s [Sheridan, 1992, p. 212]. The delay is *variable*, due to changes in the distance and in the number of satellites (1 or 2) between the manipulator and Earth, and due to the processing time within the satellites. The automatic controller of the manipulator transforms the delayed control signals into movements of the arm that are monitored with a camera. The video signal is transmitted back to Earth, and the camera picture appears on the television screen, again with a time delay Δt. In the control loop, the overall time delay of $2\Delta t$ up to about 6 s, makes controlling the robot very difficult.

In the case of the positioning task, the Pyramid Predictive Display can also be used to compensate the time delays. Since, in that case, the computer with the predictor is located on Earth, no delays will be present when the operator controls the predictor. The control commands are transmitted into space, and the real manipulator on the screen will follow the position setpoint after $2\Delta t$ s.

Compensation of the variations in the time delays

If a manipulator like the ERA is controlled from Earth, the control- and video signals will be transmitted digitally, as a discrete series of samples with a constant time interval, for example of 0.05 s. Due to the variable time delay $\triangle t$, the time interval at the arrival will not be constant anymore, and the signals will be deformed. In [Breedveld, 1992], a method is described by which the deformation in the signals can be compensated. This method uses two *buffers*; the first one near the manipulator and the second one near the television screen. At the arrival, the samples are put into the buffer, which is emptied on-line, again with the constant time interval of 0.05 s. Since the time interval between the numbers is constant again, the signals are not deformed anymore. The time interval at the input of the buffer varies, while the interval at the output is constant. Therefore, the size of the buffer is variable. Since the buffer should never become completely empty, the constant time delay between the moment of transmitting and the moment of emptying the buffer, must be somewhat larger than the maximum value of $\triangle t$. Due to this, the time delay increases a little bit. If a well working predictive display is used, however, it is not expected that a small increase in the time delay will make controlling the manipulator more difficult.

6. CONCLUDING REMARKS

At the moment of writing this paper, no man-machine experiments have been carried out with the Pyramid Predictive Display yet. Some try-outs, however, have shown that the display simplifies the positioning task strongly. In the near future of phase A2 of the research, a number of man-machine experiments will be carried out with the display, to investigate the effects of a predictive display on fine positioning the end-effector in the presence of slow, flexible dynamics and time delays. The Pyramid Predictive Display will also be compared with other kinds of predictive displays for end-effector positioning tasks, like the display described in [Ferro, 1992].

The Pyramid Display and the Pyramid Predictive Display, both developed in phase A of the research, are developed for fine positioning tasks, and help the operator in positioning the end-effector with a high accuracy. In phase B of the research, other kinds of displays are being developed for track following tasks, that help the operator in avoiding collisions between the manipulator and parts of space station. In the future of this phase of the research, a number of man-machine experiments will be carried out with different types of such displays, to compare them with each other, and to investigate their suitability for track following tasks.

7. ACKNOWLEDGEMENTS

This research is supported by the Dutch Technology Foundation STW and takes place in cooperation with Fokker Space & Systems B.V., the European Space Agency ESA, and the National Aerospace Laboratory NLR.

8. REFERENCES

Bos J F T (1991), Man-Machine Aspects of Remotely Controlled Space Manipulators, *PhD Thesis, Delft University of Technology, Fac. of Mechanical Engineering and Marine Technology, Lab. for Measurement and Control, Delft, the Netherlands, ISBN 90-370-0056-8, 184 p.*

Bos J F T (1995), Aiding the Operator in the Manual Control of a Space Manipulator. *Control Eng. Practice, Vol 3, No 2, pp. 223-230; and also in Proc. 5th IFAC Symp. on Analysis, Design and Evaluation of Man-Machine Systems, June 9-11, 1992, the Hague, the Netherlands, paper 1.3.3.*

Breedveld P (1992), Controlling a Space Manipulator from Earth with Predictive Display Techniques. *Proc. 11th European Annual Conf. on Human Decision Making and Manual Control, Nov. 17-19, Valenciennes, France, paper 27.*

Breedveld P (1994), The Design of an Optimal Man-Machine Interface for a Manually Controlled Space Manipulator. *Report N-469, Delft University of Technology, Fac. of Mechanical Engineering & Marine Technology, Lab. for Measurement & Control, 75 p.*

Breedveld P (1995), The Development of a Man-Machine Interface for Telemanipulator Positioning Tasks. *To be published in Prepr. 6th IFAC Symp. on Analysis, Design and Evaluation of Man-Machine Systems, June 27-29, MIT, Cambridge, USA.*

Buiël E F T, Breedveld P (1995a), A Laboratory Evaluation of four Control Methods for Space Manipulator Positioning Tasks. *To be published in Prepr. 6th IFAC Symp. on Analysis, Design and Evaluation of Man-Machine Systems, June 27-29, MIT, Cambridge, USA.*

Buiël E F T, Breedveld P (1995b), A Laboratory Evaluation of two Graphical Displays for Space Manipulator Positioning Tasks. *To be published in Proc. 14th European Annual Conf. on Human Decision Making and Manual Control, June 14-16, Delft University of Technology, Delft, the Netherlands.*

Diepenbroek A Y (1994) (*in Dutch*), Ontwikkeling van een real-time simulatiemodel van de ERA ruimtemanipulator met daarin wrijving. *Report A-704, Delft University of Technology, Fac. of Mechanical Engineering and Marine Technology, Lab. for Measurement and Control, Delft, the Netherlands, 98 p.*

Ferro D V (1992), Process and System for Remotely Controlling an Assembly of a First and a Second Object, *US Patent no. 5,119,305, 1992, Assignee: Aerospatiale Societe Nationale Indust., Paris, France.*

Kim W S, Tendick F, Ellis S R, Stark L W (1987), A Comparison of Position and Rate Control for Telemanipulations with Consideration of Manipulator System Dynamics, *IEEE Journal of Robotics and Automation, vol. RA-3, no. 5, pp. 426-436.*

Sheridan T B (1992), Telerobotics, Automation and Human Supervisory Control, *MIT Press, USA, ISBN 0-262-19316-7, 413 p.*

Woerkom P T L M van, et al. (1994), Developing Algorithms for Efficient Simulation of Flexible Space Manipulator Operations, *Proc. 45th Congress of the International Astronautical Federation, Oct. 9-14, Jerusalem, Israel, pp. 1-13.*

A LABORATORY EVALUATION OF TWO GRAPHICAL DISPLAYS FOR SPACE MANIPULATOR POSITIONING TASKS

Eric F.T. Buiël and Paul Breedveld

Delft University of Technology, Department of Mechanical Engineering and Marine Technology, Laboratory for Measurement and Control (Man-Machine Systems Group), Mekelweg 2, 2628 CD Delft, The Netherlands

Abstract: A space manipulator is a lightweight robotic arm mounted on a space station or a spacecraft. If the manipulator is manually controlled from a remote site (teleoperation), the pictures from cameras in the neighbourhood of the manipulator are the only aids the human operator can use. Besides, manipulator dynamics and the possible presence of time delays in the control loop also complicate manual control.
In space manipulator positioning tasks, the manipulator hand (the end-effector) has to be positioned accurately in front of a known (physical) target. Here, the human operator can only use the camera picture from a camera attached to the end-effector. In that picture, hardly any spatial cues are present. Therefore, it is difficult to assess the attitude of the end-effector relative to the target. This paper evaluates two graphical camera overlays that visualise the actual end-effector attitude. The first overlay, the *Indicator Display,* shows the exact position and orientation of the end-effector with a set of smart two-dimensional display indicators. In the second overlay, the spatial *Pyramid Display,* spatial information is presented by means of the intersections of three spatial graphical objects. The results of man-machine experiments show that the positioning task can be performed remarkably fast with *both* displays. Also, it appears that at the end of the task, the 'Pyramid operator' is more aware of the actual manipulator environment than the 'Indicator operator'.

Keywords: Teleoperation, Manual Control, Evaluation, Graphical Displays

1. INTRODUCTION

1.1 The manual control of a space manipulator

At Delft University of Technology (DUT), Department of Mechanical Engineering and Marine Technology, manual control of a space manipulator is an object of study (Bos, 1991). A space manipulator is a lightweight robotic arm mounted on a space station or a spacecraft. There, it performs inspection and maintenance tasks, e.g. the repair of a damaged satellite. Figure 1 shows an example of such a manipulator: the new European Robot Arm ERA[1] (van Woerkom *et al*, 1994). The ERA is ten metres long and has six Degrees-of-freedom (Dof's). At the end of the 20th century, it might become operational on the new global space station Alpha.

Recurrent tasks of a space manipulator (e.g. the replacement of Orbital Replaceable Units containing scientific experiments) may well be automated and performed under supervisory control. This does not seem plausible for tasks that are not well defined in advance (e.g. repair tasks). Here, the inventiveness of the human operator is required more often. Then, *teleoperation* (Sheridan, 1992) seems a suitable control method. With this method, the human operator controls the manipulator by hand

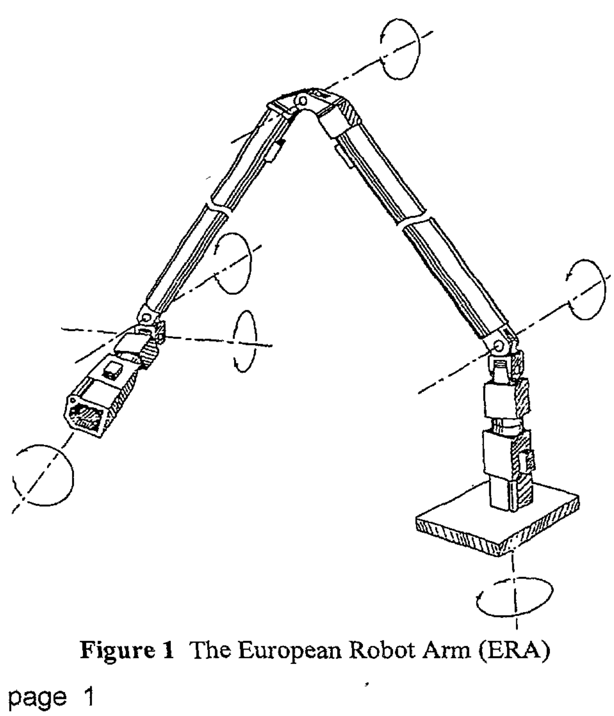


Figure 1 The European Robot Arm (ERA)

[1] The European Robot Arm is developed at Fokker Space and Systems B.V. in Leiden, The Netherlands.

from a remote location (e.g. a space station's manned module or a ground station on earth). Astronauts do not have to go outside their spacecraft to control the manipulator; the manipulator movements are controlled with the help of the pictures from cameras installed in the neighbourhood of the manipulator.

The mentioned teleoperation task is a hard job for the human operator. First, task execution suffers from the manipulator dynamics: because of the lightly constructed limbs, the manipulator will be flexible. Besides, when the operator controls the manipulator on earth, time delays are introduced in the control loop. These delays are caused by the transmission of control signals from earth to space, and back again. Finally, the lack of spatial information in the camera pictures complicates manual control. This paper focuses on that specific problem: because of the loss of the third dimension, it can be difficult to detect whether the moving manipulator may soon collide with one of the objects in its environment. Also, it can be difficult to assess the orientation of the manipulator-hand (the *end-effector*) relative to an object that has to be grasped and displaced.

A solution for the mentioned problem is available when position sensors are installed that measure the position of the manipulator relative to the objects in its environment. Then, the measured data can be presented in a graphical display shown on the man-machine interface console at the remote site. At DUT, new concepts for such displays are being developed. Current research is aimed at the development of a graphical display for an elementary subtask of a space manipulator: the *positioning task.*

1.2 The space manipulator positioning task

Before a space manipulator is able to grasp and displace an object, the end-effector must be positioned accurately in front of the object without damaging it. If this positioning task is performed in teleoperation, all end-effector Dof's (three translations and three rotations) have to be controlled manually. When the distance between the end-effector and the object has decreased to less than a few centimetres, it is difficult to detect whether the end-effector and the object collide.

Figure 2 shows an outline of the generalised positioning task currently investigated at DUT. In this task, the end-effector has to be accurately positioned upon a generic object. Here, the human operator can only use the information in the picture from the camera attached to the end-effector. If the end-effector is approaching the object that has to be grasped, this picture contains three characteristic elements: the top-side of the end-effector, the target and the vision-target. The target marks the area in which the front side of the end-effector must be placed to be able to grasp the object safely. The vision target consists of three cylinders in one line (the cylinder in the middle is taller than the others). From the position and size of each cylinder in the camera picture, a computer calculates the actual position and orientation of the target relative to the end-effector. These measures are used in the graphical display for the positioning task.

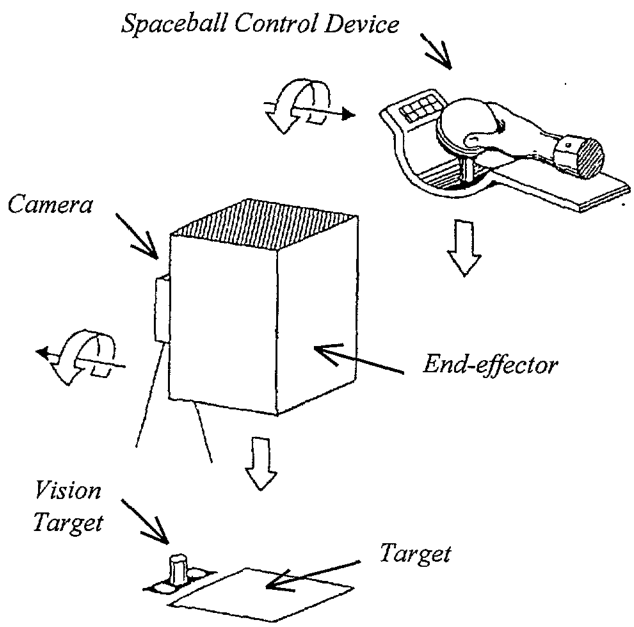


Figure 2 The space manipulator positioning task

In the investigated positioning task, the human operator controls the translational and angular velocity of the end-effector with a commercially available, force activated input device with six Dof's: the Spaceball®. If the operator grasps the Spaceball as if he grasps a car's gear lever, he might feel it as if he grasps the end-effector (*virtual grasping*; see also (Buiël and Breedveld, 1995)).

1.3 The graphical display for the positioning task

In fact, a conventional graphical display that simply shows the actual values of the position and orientation misfits might well suffice for fulfilling the task considered here. Then, to finish the task, the human operator only needs to drag six position and orientation indicators to zero one by one. However, the mentioned 'one-sensor-one-indicator interface' (Goodstein, 1981) enlarges the distance between the operator and the system to be controlled figuratively. With the introduction of the display, the operator acts like an automatic controller in a sense. If the end-effector positioning task is considered as a subtask in a much more comprehensive task, it seems wiser to develop a display that assists the operator in maintaining his spatial awareness of the manipulator environment. This *spatial display* can either be a spatial graphical camera overlay that visualises spatial cues that are missing in the end-effector camera picture, or a synthetic computer-generated image of the actual manipulator environment (an artificial camera picture). If it is well designed, the operator might feel it as if he is actually present in a three-dimensional environment (telepresence, (Sheridan, 1992)).

Currently, the pros and cons of the usage of a spatial display for the end-effector positioning task are investigated at DUT. This paper presents the outcomes of a laboratory man-machine experiment with two graphical displays. The first one is a typical example of a non-spatial, two-dimensional graphical camera overlay: the

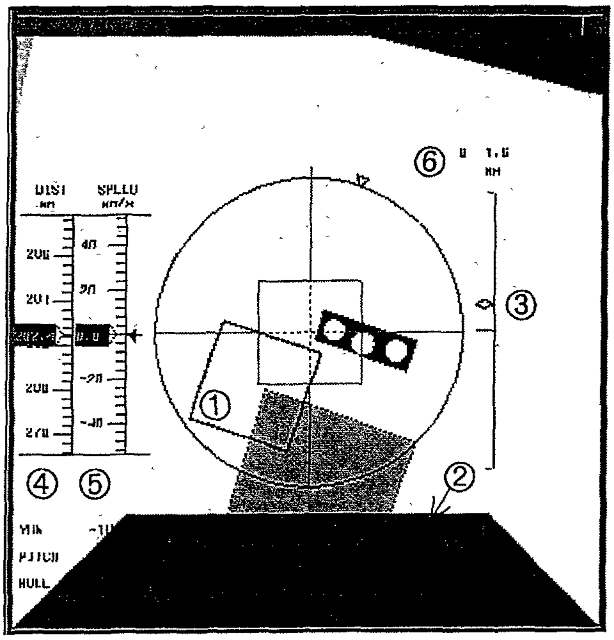


Figure 3 The Indicator Display
(① moving square, ② lateral position misfit, ③ vertical position misfit, ④ distance indicator, ⑤ speed indicator, ⑥ example of a digital misfit meter)

Indicator Display. The second display is a spatial display developed at DUT: the *Pyramid Display*. Detailed information about the work presented here can be found in Buiël (1995).

2. TWO GRAPHICAL DISPLAYS FOR SPACE MANIPULATOR POSITIONING TASKS

2.1 The Indicator Display

Figure 3 shows the Indicator Display in the end-effector camera picture. For this display, display technology that is quite conventional in current Aerospace Engineering has been applied to visualise the actual position and orientation of the end-effector.[2] Just like the artificial horizon in an aeroplane's primary flight display, a moving square near the middle of the Indicator Display (①: the 'horizon') represents the target's orientation relative to the end-effector. For the lateral and vertical position misfit resp., two moving pointers are displayed along the sides of the display (② and ③). A fixed pointer or moving scale altimeter visualises the distance between the end-effector and the plane the target is attached to (④). Next to that indicator, the end-effector velocity can be read from a second moving scale indicator (⑤). Finally, the exact value of each position- and rotation misfit can be read from digital meters (⑥).

All moving elements in the various misfit indicators of the Indicator Display move 'inside-out' (Wickens, 1992). This means that, the movements of those elements in the camera overlay correspond to the movements of the target in the end-effector camera picture. Each moving element changes colour at the moment it moves into (or out of) the safe region for the indicated misfit. At the moment all elements are coloured green (the 'safe' colour), the middle cylinder of the vision target is exactly in the middle of the display. Then, the operator can reduce the remaining distance between the end-effector and the target cautiously, and finally he can place the end-effector upon the target.

2.2 The Pyramid Display

The Pyramid Display (Breedveld, 1994 and 1995) emphasises the spatial cues that were hardly perceivable in the original end-effector camera picture (see Figure 4). First, to improve the sensation of depth, rectangular grids are projected on objects in the manipulator environment (①). Second, three spatial graphical objects are introduced to simplify the perception of the actual end-effector attitude: the frosted glass (②), the insert-box (③) and the pyramid (④). To the operator, each of these objects seems to be present in the manipulator environment. The insert-box and the pyramid seem to be attached to the vision-target; the frosted glass seems to be attached to the end-effector, just like the sight of a gun. To finish the positioning task, the human operator has to move the frosted glass towards the insert-box and the pyramid. When the end-effector approaches the target closely, the frosted glass intersects the insert-box and the pyramid successively. Then, the end-effector position and orientation can be perceived from these intersections.

Figure 5 shows the Pyramid Display at the moment the frosted glass intersects the pyramid. The size of this intersection is an accurate measure for the remaining distance between the end-effector and the target. The shape of the intersection visualises the orientation of the end-effector relative to the target. If a small orientation-misfit is present, the shape will be an irregular quadran-

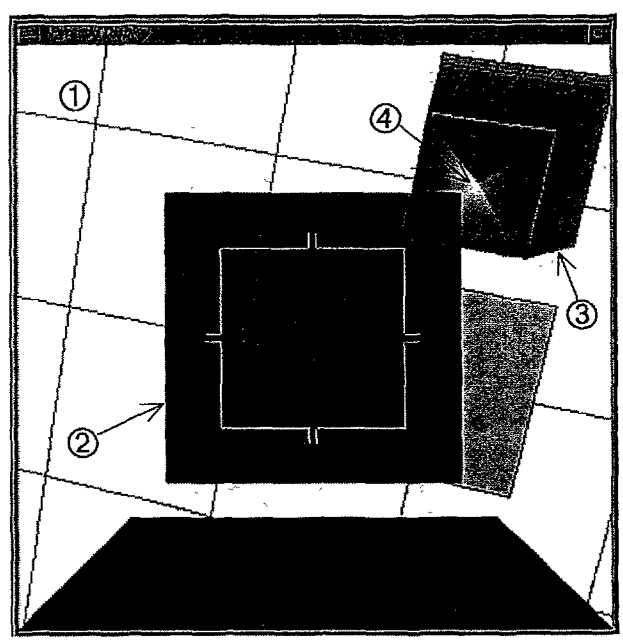


Figure 4 The Pyramid Display (① rectangular grid, ② frosted glass, ③ insert-box, ④ pyramid)

[2] The general design ideas for the Indicator Display came from the U.S. patented graphical display for remotely controlling an assembly of two objects, as developed at Aerospatiale Societe Nationale Indust., Paris, France (U.S. Patent Number 5.119.305, dated June 2, 1992).

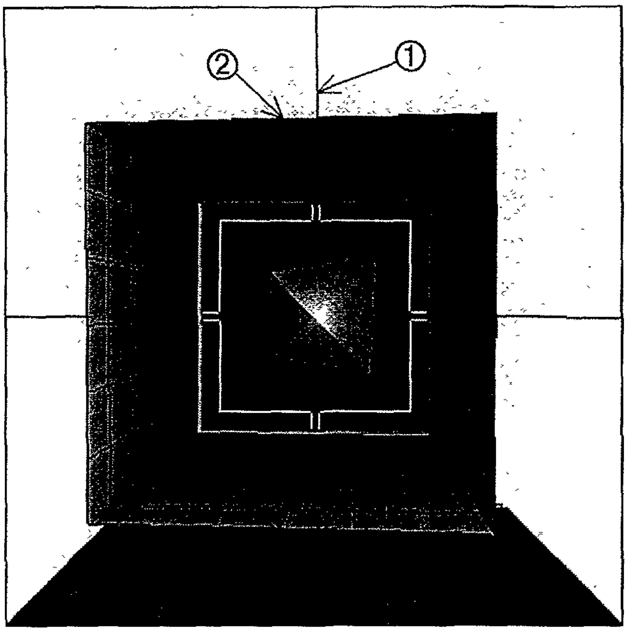


Figure 5 The frosted glass intersects the pyramid (① vertical gridline in the environment, ② longitudinal gridline upon the insert-box; see 4.1)

gle. When the intersection is almost square, orientation and position misfits are all within the desired accuracy. Then, the pyramid changes colour and the operator can place the end-effector upon the target.

2.3 Discussion

Without doubt, an operator using the Pyramid Display will be more aware of the three-dimensional environment the space manipulator is moving in than an operator using the Indicator Display. According to Gibson (1979), the optic flow patterns in the retinal picture of the human eye form an essential cue for spatial perception in daily life. To obtain a spatial impression of the space manipulator environment, the operator should utilise the flow patterns in the end-effector camera picture (i.e. the movements of the target and the other objects visible in that picture) in a similar way. This is more difficult for the operator using the Indicator Display, than for the operator using the Pyramid Display. In the display mentioned last, the presence of textured spatial objects results in a clearer motion perspective. Consider the insert-box to explain this. This box is much longer and bigger than the vision-target. Also, rectangular grids are projected on its sides. Because of this, it is easier to estimate changes in the attitude and position of the insert-box, than to do this for the vision-target. In this way, the insert-box acts more or less like an enhanced vision-target.

Despite the increased spatial awareness of the operator, the usefulness of the Pyramid Display for the end-effector positioning task is not known beforehand. To finish the task with this display, an operator can rely on spatial information only. No additional exact information about the actual end-effector velocity or the translation and rotation misfits in specific directions is provided; the 'Pyramid operator' is forced to *estimate* these quantities from the end-effector camera picture. Here, the 'Indicator operator' is in favour. This operator can always use the exact information in the Indicator Display. All the time, he can find the current velocity of the end-effector, and see which misfits are too large.

However, the question is whether the Pyramid operator really needs additional exact information like that provided by the Indicator Display. By nature, humans are very well capable to perceive spatial information from the outside world. The Pyramid Display capitalises on these capabilities. Here, the intersections of the frosted glass with the insert-box and the pyramid resp. may well provide sufficient information to be able to estimate the actual position and rotation misfits enough accurately. Also, the actual end-effector velocity may well be estimated from the optic flow in the camera picture. To verify this, in a laboratory experiment, the task performance with the Indicator Display has been compared to the task performance with the Pyramid Display.

3. A LABORATORY EXPERIMENT WITH THE TWO GRAPHICAL DISPLAYS

To measure the task performance with both displays, the members of two groups of six subjects have all been trained for a *simulated* space manipulator positioning task. The first group has been trained to perform this task with the Indicator Display; the second group has been trained to do this with the Pyramid Display. All subjects were students in Mechanical Engineering, without any experience in tasks like the simulated positioning task.

3.1 The simulation facility

Figure 6 shows the simulator that was used in the experiment. In this simulator, the movements of the European Robot Arm ERA (see the introduction to this paper) are simulated by means of a Silicon Graphics graphical workstation (①). This computer animates a simplified end-effector camera picture in real time (②); Figure 3 through Figure 5 show examples of that picture). At the start of the simulated positioning task, the end-effector is at one metre distance from the target. To finish the task correctly, the subject has to position the front side of the end-effector straight in front of this target, without coming into collision with the target itself. Here, position-misfits have to be reduced to less than 3.0 millimetres; orientation-misfits have to be reduced to less than 1.5°. The end-effector movements are controlled with a Spaceball (③).

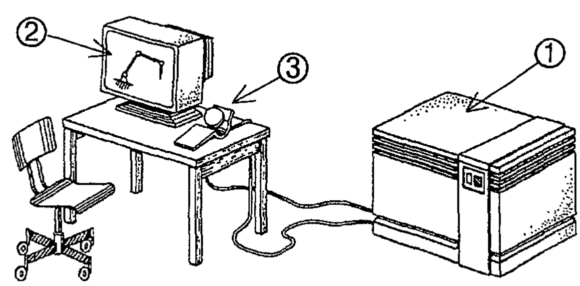


Figure 6 The simulation facility (① graphical workstation, ② end-effector camera picture, ③ Spaceball control device)

In the experiment discussed here, a simple mathematical model was used for the computation of the ERA movements. The model describes the ERA end-effector as an object with six Dof's, of which the movements are constrained by the upper limits for its translational and angular velocities only (0.1 m/s and 10°/s resp.). The end-effector movements are *not* hindered by the manipulator dynamics. In this way, the outcomes of the experiments do not only demonstrate the usefulness of both display concepts for ERA; they are also valid for other situations in which an object has to be positioned accurately in front of another object in teleoperation (e.g. in dredge technology (Jonkhof, 1995), or in laparoscopic surgery). Of course, for the ERA case, after this experiment more experiments will be needed to determine the influence of the slow ERA dynamics on operator performance with both displays.

3.2 The task training

At the start of the training for the simulated positioning task, each subject read the fundamentals of one of the displays in a manual first. After that, a large number of task runs had to be performed with the help of that display (at least 25 runs). Here, the subject was asked to search for a control strategy that results in *fast* task execution, but that guarantees *safe* task execution above all. Here, 'safe' means that no collisions between the end-effector and the object to be grasped were allowed to take place, except for the final intended contact between end-effector and target at the moment the end-effector is correctly positioned. The task had to be practised until the subject was able to perform it with sufficient and almost constant performance. This performance level had been attained at the moment the subject was able to perform eight successive task runs without any hazardous collisions between the end-effector and the target taking place. Also, the standard deviation of the completion times of these eight task runs had to be less than 3.5 seconds. To prevent the subject from becoming tired, a break was taken after each cluster of 25 practice runs.

3.3 Measurements

The task performance finally attained by every subject has been measured in a concluding experiment. In this experiment, each subject was asked to perform 24 task runs with the graphical display he or she had been trained for. Task demands were exactly the same as those in the training sessions.

After the experiment, the first part of a questionnaire was handed out. In this part, the subject was asked to express his or her opinion about the graphical display he or she had been trained for. Other questions dealt with the subject's control strategy for the end-effector positioning task and the mental load the subject experienced while erforming the task. At last, the subject was confronted with the display he had not been trained for. After he had erformed a large number of test runs with that display, he second part of the questionnaire was handed out. In his part, the subject was asked to draw a comparison etween both display concepts.

4. THE RESULTS OF THE EXPERIMENT

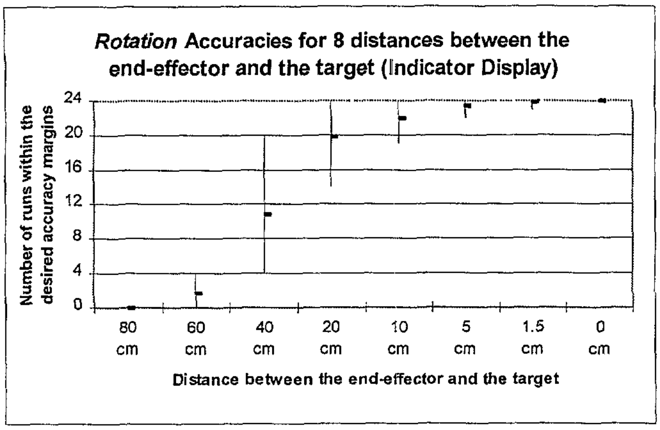
4.1 Control strategies

The control strategies with both graphical displays are visualised in Figure 7 and Figure 8 resp. Both Figures show two high-low graphs; one for the rotation misfits and one for the lateral and vertical translation misfits. Each graph shows the number of runs in a single experiment in which the considered misfits are within the requested accuracy at the moment the end-effector is at a specific distance from the target. In each graph, eight values of the distance between end-effector and target are considered. The horizontal dash in each of the eight high-low bars indicates the mean value for the mentioned number of runs (i.e. the mean of the six experiments); the top and bottom of each bar show the extreme values that were found for this number.

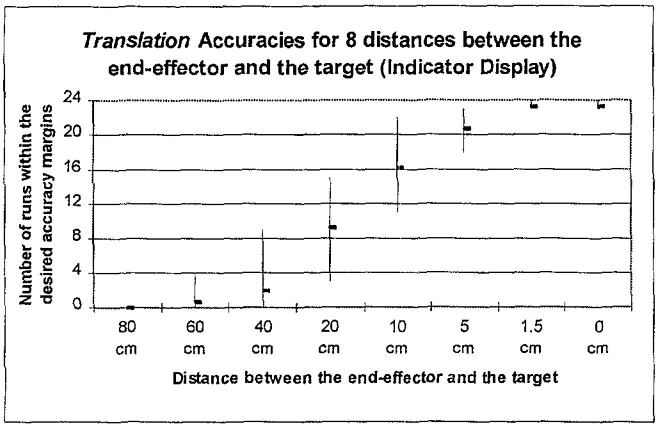
Control strategy for the Indicator Display

Generally, task execution with the Indicator Display can be split up in four stages. In the first stage, the subjects moved the end-effector towards the target at maximum speed. While the end-effector moved towards the target, the end-effector orientation was corrected gradually. At the moment the end-effector had approached the target up to about 40 to 60 cm, the subjects corrected the remaining orientation misfits with the help of the moving square in the middle of the display (2nd stage). After the rotation misfits had been reduced to acceptable values, the lateral and vertical position misfits were controlled with the help of the two moving pointers displayed along the sides of the display (3rd stage). Figure 7 confirms that the lateral and vertical position misfits were controlled after the rotation misfits had been corrected: when the distance between the end-effector and the target is in between 40 and 10 cm, the mean number of runs with 'accurate' rotation misfits is always (much) larger than the mean number of runs with 'accurate' lateral and vertical position misfits. When the rotation misfits and the lateral and vertical position misfits were small enough (distance < 5 cm in Figure 7), the end-effector was placed upon the target (4th stage).

Most subjects said that they had found some problems with the arrangement of the indicators in the Indicator Display in the last task stage. Here, a large number of indicators have to be checked sequentially. Of course, the subject has to take care that each misfit indicator indicates a value that is within the desired accuracy margin for that misfit. Also, the distance indicator has to be checked from time to time. Almost all subjects mentioned that it is almost impossible to read out these indicators at a single glance. Here, especially the distance indicator caused problems. If their eyes focused on the centre of the large circle in the middle of the display, subjects were just able to read out the orientation indicator and the lateral and vertical translation misfit indicators. The distance indicator was not clearly visible if the eyes focused on the middle of the display. Since that indicator is placed on the very left side of the display, the subject's eyes had to turn to the left to be

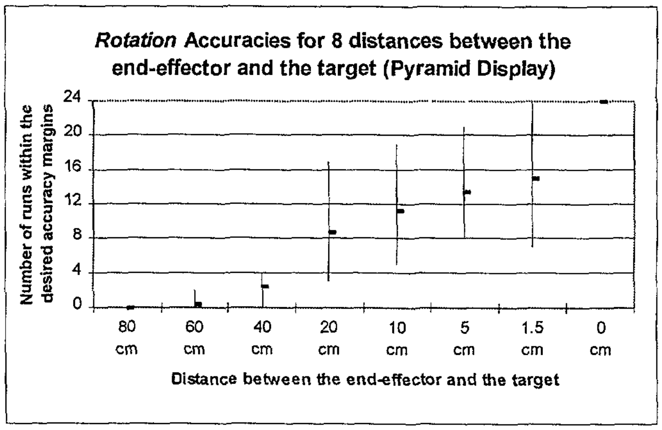


a) Control of rotation misfits

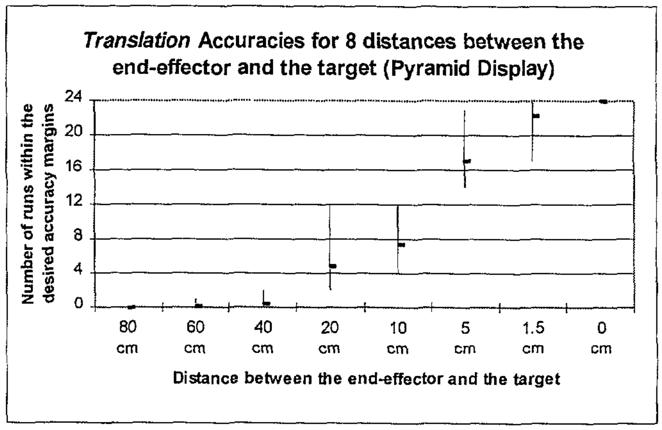


b) Control of translation misfits

Figure 7 Indicator Display control strategy



a) Control of rotation misfits



b) Control of translation misfits

Figure 8 Pyramid Display control strategy
(insert-box height = 10 cm; pyramid height = 1.5 cm)

able to read out the remaining distance between the end-effector and the target.

The above problem clearly shows that subjects really needed the distance indicator in the final stage of the positioning task performed with the Indicator Display; in this task stage, they were not able to assess the remaining distance between the end-effector and the target from the actual end-effector camera picture only. This indicates that with the Indicator Display, subjects are not enough aware of the spatial properties of the environment of the end-effector at the end of the positioning task. Because of this, most subjects assessed the probability of a collision with the Indicator Display to be larger than the probability of a collision with the Pyramid Display.

Control strategy for the Pyramid Display

The subjects that were trained with the Pyramid Display also guided the end-effector towards the target at maximum speed first. At the same time, they roughly corrected the alignment of the insert-box with the frosted glass. At the moment the frosted glass was about to intersect the insert-box for the first time (distance = 10 cm in Figure 8), the subjects considered the remaining position and orientation misfits. Here, two distinct strategies were found. Three subjects applied the 'pyramid strategy'. They moved the frosted-glass slowly towards the pyramid (distance = 1.5 cm in Figure 8), and corrected the remaining position and orientation misfits with the help of the intersection between the frosted glass and the pyramid. The other subjects applied the 'insert-box strategy'. They corrected the remaining misfits by looking consciously at the rectangular grids projected upon the insert-box and the environment of the target. Consider Figure 5 to explain this. Because the intersection of frosted glass and pyramid is not square, the end-effector is not perfectly aligned with the target in this figure. This can also be concluded from the fact that the vertical grid line marked ① (a part of the rectangular grid projected on the target object) is not in one line with the longitudinal grid line marked ② (a part of the rectangular grid projected on the insert-box). At the moment the end-effector is perfectly aligned with the target, both gridlines will be in one line.

The observation of two distinct control strategies for the Pyramid Display influenced the length of the high-low bars in Figure 8 (especially the bars in Figure 8a). At the moment the frosted glass has not intersected the pyramid yet (distance > 1.5 cm), the mean number of runs with 'accurate' misfits is (much) larger for the 'insert-box subjects' than for the 'pyramid subjects'.

4.2 Task performance and mental load

Task performance has been analysed by means of two performance measures: the number of collisions occurring in a single experiment and the mean completion time of the 24 task runs in a single experiment. At the end of this analysis, the attained level of task performance has been related to the mental load the subjects experienced while performing the positioning task. In this way, the suitability of both displays for the end-effector positioning task was investigated.

Number of collisions

The number of collisions occurring in the twelve concluding experiments was very small. In all Pyramid Display experiments, only three collisions occurred (each one was caused by a different subject). In the experiments with the Indicator Display, four collisions occurred (two of these collisions were caused by a single subject; two other subjects caused one collision only).

Task completion time

Figure 9 shows 95% confidence intervals for the mean completion time of the 24 task runs performed by each subject (each set of 24 completion times is assumed to be a sample from a normal distribution). Relatively large performance differences are present between the subjects trained for the Pyramid Display. In (Buiël, 1995) it is shown that these differences may well be explained from the operator's capability to perceive the spatial information in the Pyramid Display. For the Indicator Display, the performance differences between the six subjects are small. Apparently, the control strategy for this display caused almost equal problems to all subjects.

From Figure 9, it can already be concluded that with both graphical displays the positioning task can be performed quite fast. The grand mean run completion time for all subjects trained for the Pyramid Display is 23.9 sec.; for the subjects trained for the Indicator Display this is 24.7 sec. (the minimal time necessary to complete a task run was 10 sec.; this value can be computed from the upper limits for the translational and angular end-effector velocities). A two-sample student's t-test with df = 10 indicates that this difference is far from significant on the 5% significance level (t_{10} = 0.38, two-tailed p = 0.71).

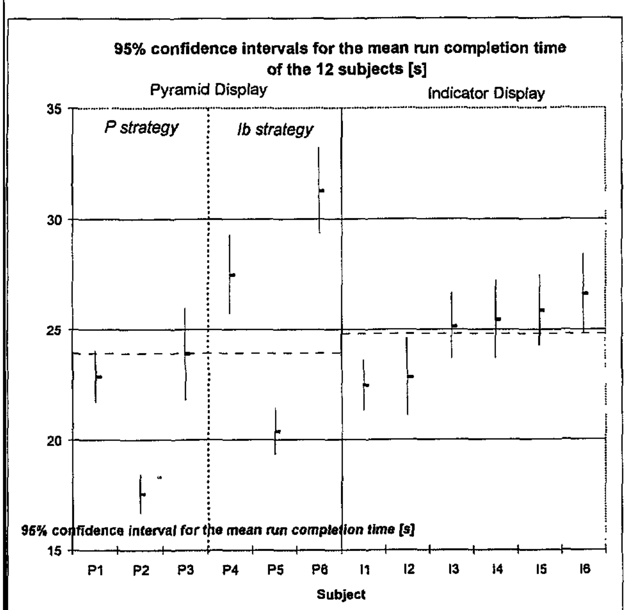


Figure 9 95% Confidence intervals for the mean run completion time of the 12 subjects

Table 1
Mental load experienced by the 12 subjects

Display	Mental load (Large values indicate large mental load)						
	1	2	3	4	5	6	7
Pyramid Display	1	3			2		
Indicator Display		3	1	1	1		

The table shows the number of subjects that experienced a specific level of mental load. Level 1 indicates that the positioning task is *not* mentally fatiguing; level 7 indicates that the task is mentally *very* fatiguing.

Therefore, it has been concluded that a possible performance difference between both displays is so small, that it is not worth mentioning. With the Indicator Display, as well as with the Pyramid Display, the end-effector positioning task can be performed very fast.

Mental load

Immediately after the experiment, each subject was queried for the mental load that he had experienced while performing the positioning task. The mental load was assessed by means of a discrete rating scale with seven levels at equal distances. Table 1 shows the mental load ratings of the twelve subjects. Again, no substantial differences are found between both graphical displays. In both groups, four subjects selected an option that indicates a low level of mental load (i.e. options 1, 2. and 3); the mental load experienced by the others was larger (options 4 or 5).

If the outcomes of the analysis of task performance and mental load are combined, it can be concluded that both graphical displays are suitable for the end-effector positioning task. With both displays, the task can be performed quickly and with small mental effort.

4.3 Subject opinions

After the training with the second display, each subject was asked to choose his favourite graphical display. Finally, eight subjects preferred the Pyramid Display, and three subjects preferred the Indicator Display. One subject liked the Pyramid Display just as much as the Indicator Display. The opinions of the subjects have been influenced by the order in which both displays were presented to them. Five of the six subjects that were first trained with the Pyramid Display preferred that same display in the end. The six subjects that were first trained with the Indicator Display were more divided about the choice of the best display.

Most subjects did not experience large difficulties with *both* graphical displays. This is illustrated in Table 2. This table shows the means and standard deviations resp. of the marks that each subject has assigned to both displays. In the same way as a Dutch school teacher judges his pupils, each mark had to be a whole number between 1 and 10. A number between 1 and 5 indicates that the display is unsatisfactory; other values indicate that it is satisfactory (if a display is marked with a 10, it is perfect). The table shows that both displays are marked

Table 2
The assessment of both displays by the 12 subjects

Graphical display	Training order		Grand
	PD/ID	ID/PD	mean
Pyramid Display	8.3 (sd 0.5)	7.7 (sd 0.7)	8.0 (sd 0.7)
Indicator Display	6.5 (sd 1.5)	7.5 (sd 0.5)	7.0 (sd 1.2)

The table shows the mean and standard deviation resp. of the marks that each subject has assigned to both displays. Each mark had to be a whole number between 1 and 10. A large value indicates that the display was highly appreciated.

as satisfactory. The appreciation for the Pyramid Display is a little larger than the appreciation for the Indicator Display. Again, the final assessments of the displays have been influenced by the order in which both displays were presented to the subjects.

In general, the subjects who preferred the Pyramid Display liked the amount of spatial information in this display. They said that with the help of these spatial cues, the actual attitude of the end-effector and the actual distance between end-effector and target can roughly be estimated in short time. With the Indicator Display, this takes more time, because here, especially at the moment the end-effector is very close to the target, hardly any useful spatial cues are present in the end-effector camera picture. Next to this benefit, those who preferred the Pyramid Display also mentioned that they found the Indicator Display much more crowded than the Pyramid Display. One of them said: "The Pyramid Display is simple, but it is also comprehensible; the Indicator Display is full of different indicators and numbers."

The subjects who preferred the Indicator Display liked the amount of exact information in this display. They said that with the Pyramid Display, it takes some time to find those misfits that are not within the desired accuracy margins at the very end of the positioning task. This is no problem in the Indicator Display, because each misfit value is displayed separately in that display.

5. CONCLUSIONS

The outcomes of the experiment show that with the spatial Pyramid Display, as well as with the two-dimensional Indicator Display, a non-flexible tele-manipulator can be positioned accurately in very short time, and with small mental effort. With the experiment, the pros and cons of the usage of a spatial display in teleoperation tasks have become more clear. With the Pyramid Display, the human operator should be able to perceive spatial information in the same way as he does in daily life. Indeed, most of the subjects that participated in this experiment liked the amount of spatial information in the Pyramid Display. Different strategies were applied to utilise this information. The experiment also shows that to some extent, the task performance with this display depends on the operator's capability to perceive that information. This was indicated by the relatively large performance differences found among the subjects that were first trained with the Pyramid Display.

The experiment also demonstrates the usefulness and vulnerability of conventional display technology for teleoperation tasks. The major benefit of the Indicator Display is that the human operator can always *exactly* see what is wrong (i.e. which position and rotation misfits are not within the desired accuracy margins). The display does not capitalise on the human's natural capabilities to perceive spatial information. Especially at the end of the positioning task, hardly any spatial cues are present in the end-effector camera picture. In the experiments, this resulted in a diminishing spatial awareness: the subject's eyes had to turn to the left of the display from time to time, to find the remaining distance between the end-effector and the target.

ACKNOWLEDGEMENTS

This research is supported by the Dutch Technology Foundation STW (Utrecht, The Netherlands) and Fokker Space and Systems B.V. (Leiden, The Netherlands).

REFERENCES

Bos, J.F.T. (1991). *Man-Machine Aspects of Remotely Controlled Space Manipulators.* PhD Thesis. Delft UT, Dept. of Mechanical Engineering and Marine Technology, Delft. ISBN 90-370-0056-8. 177 p.

Breedveld, P. (1994). *The Design of an Optimal Man-Machine Interface for a Manually Controlled Space Manipulator.* Delft UT, Dept. of Mechanical Engineering and Marine Technology, Lab. for Measurement and Control, Internal report N469, Delft. 75 p.

Breedveld, P. (1995). *The Development of a Man-Machine Interface for Telemanipulator Positioning Tasks.* Proc. 6th IFAC Symp. on Analysis, Design and Evaluation of Man-Machine Systems, Cambridge, Massachusetts. 6 p.

Buiël, E.F.T. (1995). *A Man Machine Experiment with two graphical displays for Space Manipulator Positioning Tasks.* Delft UT, Dept. of Mechanical Engineering and Marine Technology, Lab. for Measurement and Control, Internal report N485, Delft.

Buiël, E.F.T. and P. Breedveld (1995). *A Laboratory Evaluation of four Control Methods for Space Manipulator Positioning Tasks.* Proc. 6th IFAC Symp. on Analysis, Design and Evaluation of Man-Machine Systems, Cambridge, Massachusetts. 6 p.

Goodstein, L.P. (1981). Discriminative display support for process supervisors. *Human detection and diagnosis of system failures,*a NATO symposium, Roskilde, Denmark (eds. J. Rasmussen and W.B. Rouse), Vol. 3, pp. 21-36. Plenum Press, New York.

Gibson, J.J. (1979). *The ecological approach to visual perception.* Houghton Mifflin, Boston. ISBN 0-395-27049-9.

Jonkhof, M.N. (1995). *Design of a graphical display with spatial information for the manual control of a dredge crane.* Delft UT, Dept. of Mechanical Engineering and Marine Technology, Lab. for Measurement and Control, M.Sc. Thesis A687, Delft (in Dutch). 82 p.

Sheridan, T.B. (1992). *Telerobotics, Automation, and Human Supervisory Control.* The MIT Press, Cambridge, Massachusetts. ISBN 0-262-19316-7. 415 p.

Wickens, C.D. (1992). *Engineering Psychology and Human Performance.* 2nd Edition. HarperCollins Publishers, New York.

Woerkom, P.Th.L.M. van, A. de Boer, M.H.M. Ellenbroek and J.J. Wijker (1994). *Developing algorithms for efficient simulation of flexible space manipulator operations.* 45th Congress of the Int. Astronautical Federation, Jerusalem. 13 p.

Teleoperation with a dexterous robot arm

André van der Ham, Sandor den Braven, Ger Honderd, Wim Jongkind
Control Laboratory
Department of Electrical Engineering
Delft University of Technology
PO Box 5031
2600 GA Delft
The Netherlands

E-mail: a.c.vanderham@et.tudelft.nl

Abstract

In nuclear power plants snake-alike manipulators are used for telemanipulation. Thus, the telemanipulator arm is highly kinematically redundant. This paper discusses an experiment using two dissimilar robot arms configured to be used for teleoperation. The telemanipulator (remote) arm is kinematically redundant. Control schemes for 1:1 position and force bilateral control and for solving the control of the kinematically redundant manipulator are presented and have been verified in practice.

1. Introduction

Through the use of teleoperation systems in a nuclear environment, a human operator will not have to be exposed to high levels of radiation. The operator will control a slave manipulator, which is situated in the hostile environment, by generating position/force commands with a master system.

This paper presents a teleoperation system where the operator controls the slave manipulator by controlling a master manipulator which can be kinematically dissimilar to the slave. In order to control all the cartesian degrees of freedom of the slave, the master must at least have the same number of cartesian degrees of freedom.

The two manipulators will be controlled with a bilateral control system. This means that the operator can not only move the slave manipulator to a desired position by controlling the master manipulator, but he can also feel and control the forces exerted by the slave on the environment. This additional force reflection will improve the performance of the teleoperation system.

The basis of the bilateral control scheme is formed by position control of the manipulators. The snake-alike manipulators used in nuclear power plants are highly kinematically redundant. A control scheme is presented which, unlike most control schemes for such robots, makes full use of the kinematical redundancy.

The bilateral control system has been implemented and tested with a kinematically redundant robot as slave and a Scara robot as master.

2. Position control of a kinematically redundant robot

The position of a robot manipulator is defined as the position and orientation of the end-effector relative to the base frame of the robot in cartesian space. Given a certain set of joint positions of the robot (Θ), the forward kinematics are used to calculate the position in cartesian space (X):

$$X = Kin(\theta). \tag{1}$$

A robot is controlled by controlling the separate axes of the manipulator. So in order to control the robot in cartesian space, a conversion from a cartesian space description into a joint space description is needed. A straightforward method is to use the inverse kinematics to calculate the desired joint positions, given the desired cartesian position.

For a kinematically redundant robot the inverse kinematics problem has no closed form solution. Because a kinematically redundant robot has more joint variables than its number of cartesian degrees of freedom, there are an infinite number of ways to reach a certain cartesian position.

There are however numerical solutions to solve the inverse kinematics problem of a kinematically redundant robot. The solution that is presented here uses the transpose Jacobian for the conversion from cartesian space into joint space. When a manipulator is in contact with the environment, the mapping between the joint torques (τ) and the exerted forces (F) on the environment is given by:

$$\tau = J^T(\theta)F, \tag{2}$$

where $J(\Theta)$ represents the Jacobian of the manipulator. Figure 1 shows how this relationship is used to accomplish a cartesian control scheme.

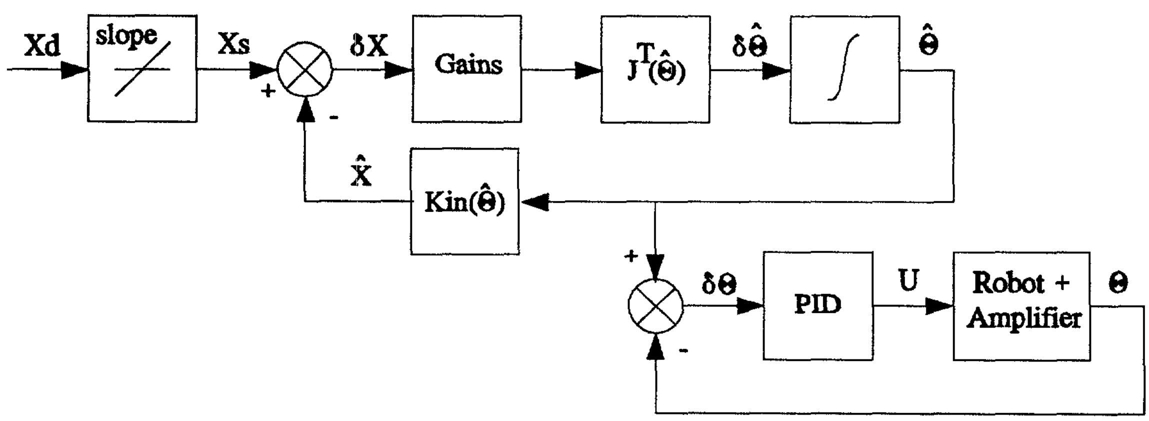


Figure 1: Transpose Jacobian inverse kinematics cartesian control scheme

The cartesian set point is constantly compared with the estimated cartesian position. The resulting errors are multiplied with gains and can be regarded as the cartesian forces that when applied to the end-effector of the manipulator, will reduce the cartesian error. With the transpose Jacobian the cartesian error is transformed into joint displacements. These joint displacements are added to the previous estimation of the desired joint positions as indicated by the integration term. The actual control is performed on joint level. The estimated joint positions are used as reference for the joint position controllers which use a simple PID algorithm.

The control scheme is based on a form of stiffness control of a robot which can be controlled in torque mode (figure 2).

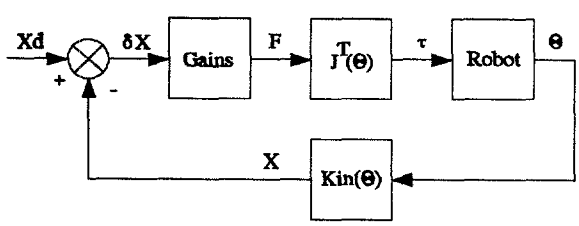


Figure 2: Stiffness control of a torque controlled robot

By substituting the standard model for each joint, one requires a model based inverse kinematics estimator. For convergence of the estimator the dynamics of the motor are disadvantageous and therefore may be omitted to allow faster convergence. In essence the motor model is simplified by a pure integrator with a certain gain. By sweeping the motor model gain together with the stiffness term, the estimator becomes as described in figure 1.

The inverse kinematics has been solved with an iterative algorithm. This is usually not beneficial to the control speed of the system. The estimation of the joint positions should not take too long. For small changes in the cartesian set point the estimation is rather fast. Therefore a slope filter is applied to limit the set point changes to a maximum cartesian velocity in order to achieve a good performance of the system.

The convergence speed of the numerical algorithm depends on the gains with which the cartesian errors are multiplied. These gains should be set for a sufficient high convergence speed. It is difficult to set the gains for the maximum possible convergence speed because this maximum will differ with different configurations of the robot manipulator.

3. Bilateral control of two robot manipulators

In order to control the slave manipulator, the operator needs a master system to generate position and force commands. In our case this interface is provided by a master manipulator with a force/torque sensor placed on the end-effector of the robot. The forces generated by the operator will be converted into position commands for a cartesian position controller of the robot as shown in figure 3.

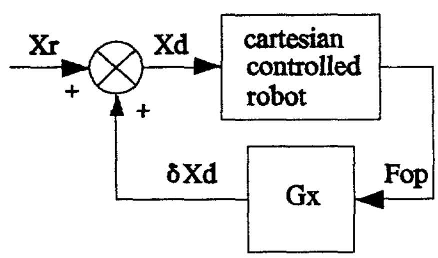


Figure 3: Cartesian stiffness control

The desired position displacement (δX) is calculated as:

$$\delta Xd = GxFop, \tag{3}$$

where Fop is the vector with forces and torques generated by the operator and Gx is a diagonal matrix. The desired displacement is added to a reference position (Xr). In this way the end-effector of the robot will appear to act like a spring with a certain stiffness along the cartesian degrees of freedom of the manipulator. As the action of a general spring with the same cartesian degrees of freedom as the manipulator is described by:

$$F = Kx\delta X, \tag{4}$$

where Kx is a diagonal matrix with the stiffness coefficients on its diagonal, the end-effector of the manipulator will have the same stiffness characteristics as this spring if Gx equals Kx^{-1}.

The bilateral control system is accomplished by using the slave position as reference for the position controller of the master manipulator and the master position (Xm) as reference for the position controller of the slave manipulator. This is shown in figure 4 where the subscripts 'm' and 's' represent respectively the master and slave variables.

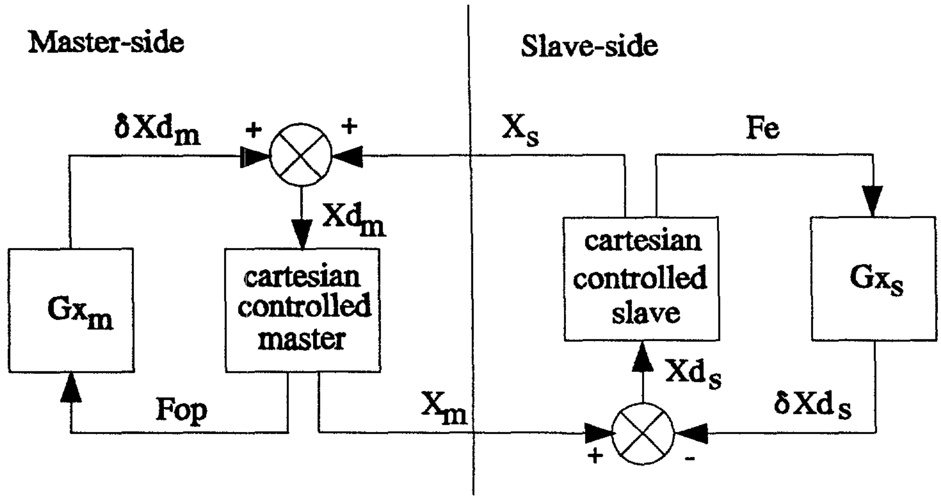


Figure 4: Bilateral control system for two robot manipulators

Due to the limited bandwidth of the position controller of the slave, the slave position will deviate a little from the master position during free motion (Fe=0). This position error is felt by the operator and he will always have to exert a certain force on the master manipulator to move the slave manipulator. With this force he is able to control the speed of the movement. The operator will also feel a position error that is caused when the slave operator makes contact with the environment.

The force reflection can be made clear with a few equations that can be derived from figure 4. Through the position controllers of both the master and the slave, the following equations will hold:

$$X_s = X_m - Gx_s Fe, \tag{5}$$

$$X_m = X_s + Gx_m Fop. \tag{6}$$

Substitution of equation 5 in equation 6 yields the next relationship:

$$Gx_s Fe = Gx_m Fop \tag{7}$$

Hence, if Gx_s is chosen equal to Gx_m, the forces exerted by the slave manipulator on the environment are equal to the forces exerted by the operator on the master manipulator.

4. Implementation and results of the bilateral control system

The bilateral control system has been implemented for the Octovera robot as the kinematically redundant slave manipulator. This is a manipulator which has six joints to control four cartesian degrees of freedom, the position (represented by three variables) and the orientation of the end-effector in the horizontal plane (one variable). The master manipulator is a Bosch Scara robot which has four joints to control the same four cartesian degrees of freedom as the Octovera robot. Both robots are equipped with a six degrees of freedom force/torque sensor at the wrist of the manipulator.

A transputer based system has been used to implement the bilateral control algorithms. The main feature of this transputer system is that it provides a multitasking environment in which several processes can run in parallel. This property is especially used to implement separate control processes for the two different robots. The two manipulators are thus really controlled in parallel.

The bilateral control system that has been presented can be divided into four different control levels as shown in figure 5. Up to and including the stiffness control level the control levels have been implemented separately for the two robots. The connection between the two robots is established in the last level, the bilateral control level.

bilateral control	
stiffness control	stiffness control
cartesian control	cartesian control
joint control	joint control
Master	Slave

Figure 5: Bilateral control structure

The control method that has been presented to control a kinematically redundant robot has been implemented and tested for the Octovera robot. Figure 6 shows the resulting cartesian position responses where the manipulator is moved from the (X,Y) position (1.5,0.0) to (1.4,0.1). The solid lines indicated by Xs and Ys represent the cartesian set point after the slope filter which has been set to a maximum cartesian velocity of 0.03 m/s. The dashed lines indicate the estimated cartesian position. The actual cartesian position is represented by the solid lines indicated by X and Y.

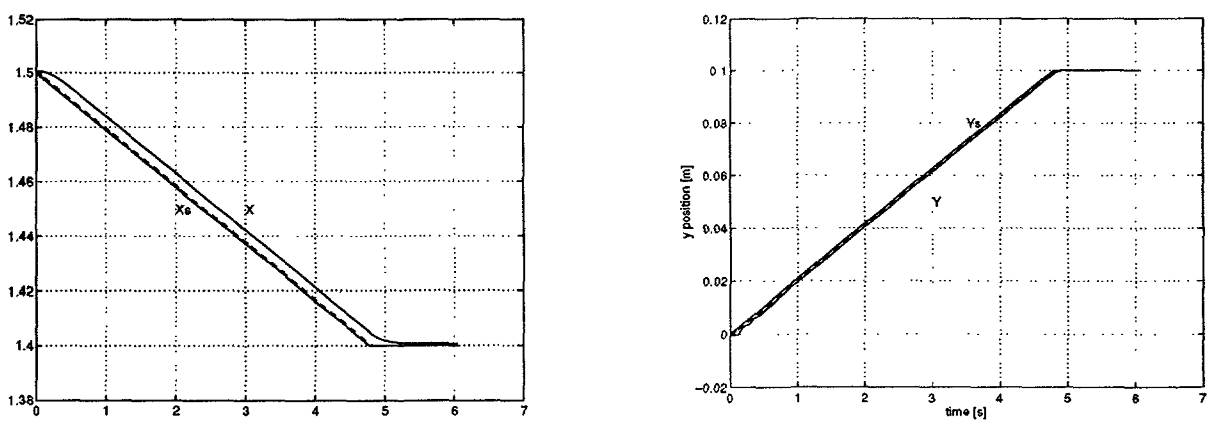


Figure 6: Cartesian responses for the kinematically redundant robot

The estimated cartesian position follows the cartesian set point very fast. The actual position in the Y direction follows the estimated position faster than in the X direction. This difference results in a deviation from the desired path and is due to different responses of the joint controllers of the robot.

The bilateral control system has been implemented with a linear stiffness of 5 KN/m for both robots. The stiffness of the two robots has to be the same for a 1:1 force reflection. Figure 7 shows the force and position responses of an experiment where the slave manipulator approaches an object in the X direction. The object has an estimated stiffness of 50 KN/m.

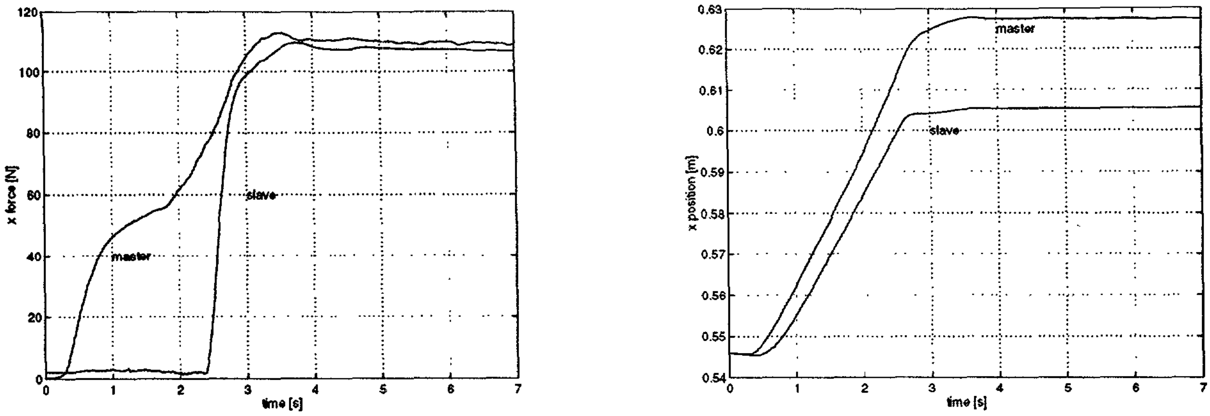


Figure 7: Force (left) and position (right) responses of the bilateral control system

During free motion the slave manipulator follows the master manipulator. The operator has to exert a force on the master in order to achieve this movement. When the slave makes contact with the environment this is felt by the operator. During contact the forces exerted by the slave are equal to the forces exerted by the operator on the master. So both 1:1 position and 1:1 force bilateral control have been realized.

At present the stiffness of the master is set to a constant value. Previous research (Ham, 1994) has pointed out that adaptation of the stiffness greatly improves the performance. The next step is to implement the adaptive laws for this system.

5. Conclusions

In this paper a teleoperation system for two dissimilar robot manipulators has been presented. This system is based on cartesian position control of the manipulators and establishes both 1:1 position and 1:1 force bilateral control.

A control method has been presented which uses all the joints of a kinematically redundant robot to control the cartesian position of this robot.

The bilateral control system has been implemented for a kinematically redundant slave manipulator and a Scara robot as master manipulator. The system works as intended, the operator can move the slave manipulator to a desired position and when the slave makes contact with the environment he can control the forces exerted by the slave on the environment.

References

Berger, J.S., *Design and realisation of an extension of the operating system of transputers and an application of this with the implementation of a complex robot control system.* Masters thesis (in Dutch), Delft University of Technology, Control Laboratory, 1994.

Craig, J.J., *Introduction to Robotics; mechanics and control.* Addison-Wesly publishing company, 1986.

Ham, A.C. van der, *A bilateral position-force controller with gain scheduling for telemanipulation.* IMACS international symposium on signal processing, robotics and neural networks, Villeneuve d'Ascq, France, April 1994.

Klomp, C., *Sensor based fine motion control.* Ph.D. thesis, Delft University of Technology, Control Laboratory, 1994.

Pieterse, H.L., *Kinematics of the Octovera.* Internal report (in Dutch), Delft University of Technology, Control Laboratory, 1994.

Improving feedback in body powered protheses

J. Herder[1], M. Munneke[2]
[1] Delft University of Technology, Faculty of Mechanical Engineering, Measurement and Control
[2] Free University of Amsterdam, Faculty of Human Movement Sciences, Currently at Leiden Academic Hospital AZL

Abstract
This paper concentrates on the control of hand prostheses. It is stated that body powered voluntary closing devices possess the best control potential. Feedback paths are identified and the observation is made that feedback is strongly influenced by the design of the prosthesis. Finally, the operating force feedback path is investigated in more detail. Psychophysical measurement methods point out (1) that operating forces should not be lower than 2.5N, and (2) that perforated self-adjusting shells should be used for good sensitivity in the information transmission through the interface.

Introduction

The Wilmer group is dedicated to applying principles of control engineering in mechanical design. Functional structures are conceived of from a system theory perspective, an approach that facilitates theoretical achievements such as feedback and feedforward to be implemented in a mechanical configuration.

The design of hand prostheses is one of the main application fields. Over the years, the Wilmer group assessed a set of basic demands for rehabilitation aids, that may be summarised as the triple-C criteria: cosmetics, comfort and control [1]. In this paper, the control aspect of hand prostheses will be emphasised.

Body power

Worldwide, most research aims to implement servo systems (myo-electric control) in order to restore motoric function [2]. However, servo systems require the availability of energy sources (external energy), and in spite of much research [3], no satisfactory artificial information feedback system has been realised. Feedback, essential for controllability of any system, is intrinsically present in body powered devices, which use muscle groups of the prosthesis user for control and energy supply.

Commonly, the shoulder girdle is used to supply operating forces. However, the burden of the shoulder harness makes people prefer the concept of elbow control [4], where elbow moment drives the prosthesis [5].

Most current body powered prostheses are voluntary opening: muscle action is required to open the prosthesis while a spring delivers the closing force. Disadvantages are that voluntary opening control works opposite to human physiology, and the closing spring must be stretched each motion cycle [6]. Voluntary closing hand prostheses are designed such that they close on muscle action. In figure 1, its principle is sketched. As the user's muscles generate pinching force, their propriocepsis supports prosthesis control [7], physiologically appropriateness is guaranteed [6,8], and pinching force is easier dosable compared to voluntary opening prostheses [9].

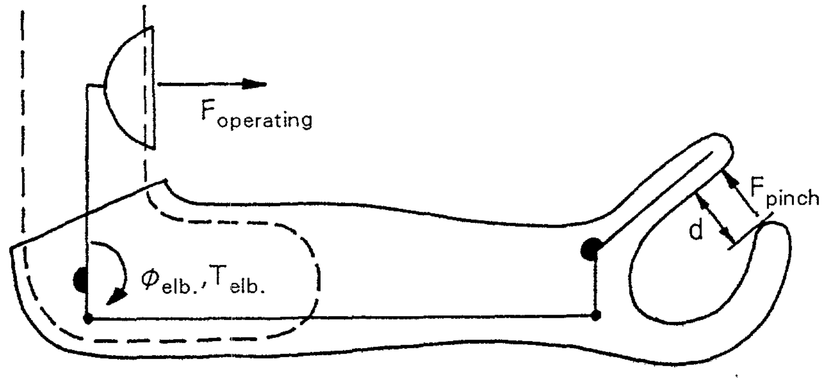


Figure 1 *Principle of an elbow controlled voluntary closing hand prosthesis*

Voluntary closing control

In goal-directed movements, the central nervous system conducts muscle action and reads the current status of the motoric system from sensors within muscles, tendons, around joints, and in tissue (propriocepsis) [10]. The prosthesis is driven by muscle action, which is accompanied by the generation of sensory information in the driving musculature, and at the locations of the human-prosthesis interface. Figure 2 shows the block diagram of this process. Except for vision of the hand opening, there is no direct feedback of the state of the variables hand opening and pinching force. However, if the prosthesis is designed such that a clear relationship exists between hand opening and elbow flexion, and between pinching force and operating and/or muscle force, then also the other feedback paths in figure 2 support prosthesis control. This phenomenon is called extended physiological proprioception [11]. Extended physiological proprioception is obstructed when the relationship between input and output of the prosthetic hand is disturbed by friction and by parasite spring forces from the cosmetic covering. Research projects were started to eliminate these influences [12,13].

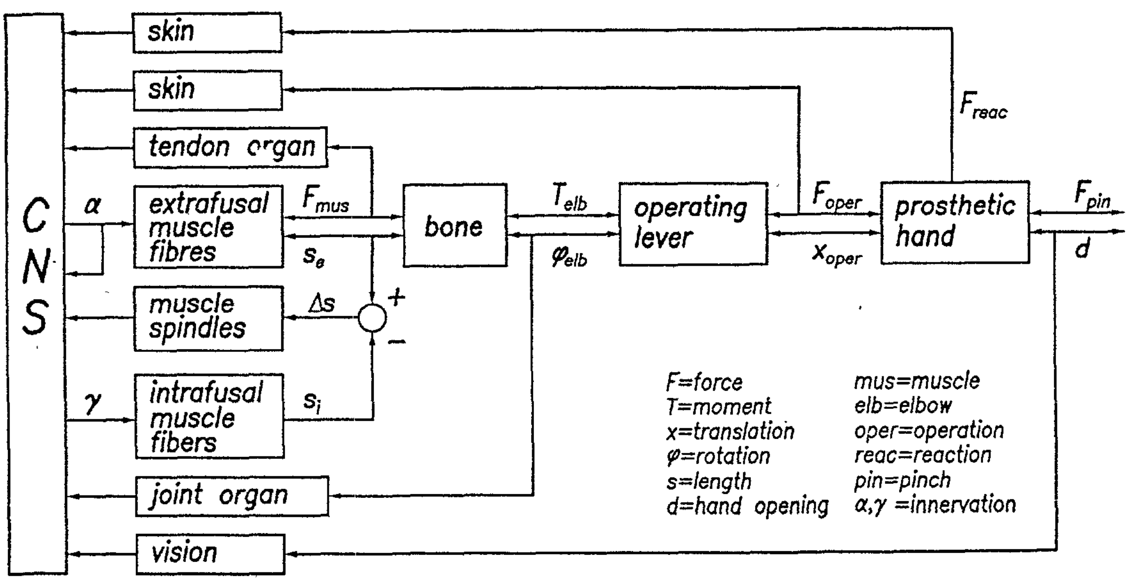


Figure 2 *Block diagram of elbow controlled hand prosthesis control. Signals correspond to figure 1.*

Methods

Being convinced from the benefit patients would have of a properly functioning voluntary closing hand, the Wilmer group started a research project to solve several practical problems [14], and to determine prosthesis geometry for optimal feedback.

Explorative research has been executed on optimising the relationship between hand opening and elbow flexion angle [15], but no reports were found on the prosthesis geometry for optimal force feedback in voluntary closing devices.

In order to make full use of proprioceptive capacity, the forces presented to the prosthesis user must match the highest sensitivity range of the sensors in the human body. In a well designed prosthesis, the muscle force, the operating force on the upper arm, and the reaction forces at the stump all contain information on the exerted pinching force (see figure 2). Of these, the operating force, exerted on the upper arm of the user, is not influenced by the weight of the prosthesis. For that reason, as a first step, the prosthesis design parameters were assessed for optimal feedback of the operation force.

Psychophysical measurement methods were used to assess (1) the sensitivity of the upper arm as a function of operating force level, and (2) the difference in sensitivity function when operating levers are used with or without perforated shells at the contact with the user's upper arm.

Psychophysical measurements

The science of psychophysics investigates the relationship between the perception of stimuli and their physical characteristics. This relation is called the psychometric function [16], which is dependent on the kind of stimulus, its location, its intensity, and its duration [10 p331/3].

Methods for quantifying the psychometric function are based on the comparison of two stimuli, occurring sequentially or at different locations. Indirect methods use one stimulus as a standard, while the other must be reproduced, or judged qualitatively with respect to the standard [16]. In direct methods, one stimulus is given a value, while the other stimulus must be rated with respect to this value, or a stimulus of another value must be reproduced [17].

For the purpose of this study, ingredients of several indirect methods were composed to a simple measurement procedure. The experimental is furnished with two operating levers, one at each upper arm (figure 3a). Operating forces are applied left and right simultaneously. Several dozens of stimuli are applied in pairs: the operating force on the right side is constant each time, while the left side is varied in arbitrary order. The experimental is to say whether the variable force (left) feels larger, equal, or

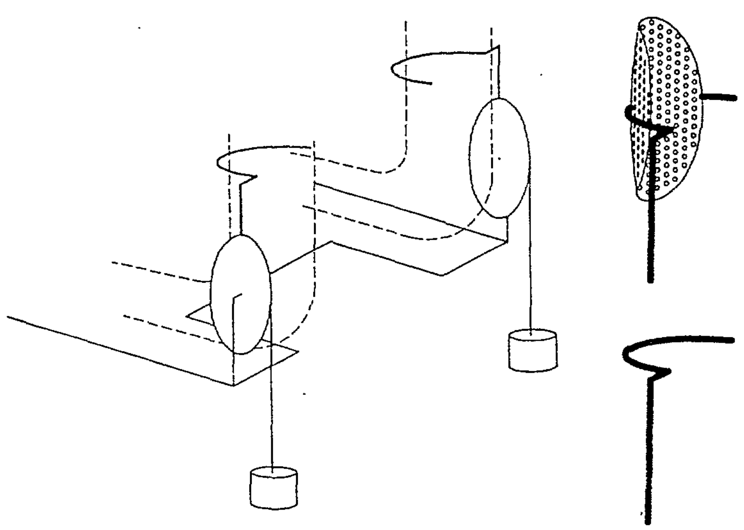


Figure 3 *(a) Experimental set-up. (b) operating lever with and without perforated shell.*

smaller than the constant reference (right). This procedure is repeated for four different reference force levels: 2.5N, 5N, 10N and 15N. Additionally, the same experiment was executed to assess the influence of a perforated shell that can be mounted on the operating lever (figure 3b).
Seven healthy young men were subject to the experiments. Of each reference force level, the lower and upper transition points were assessed (see figure 4). The force range between the transition points was defined as the uncertainty range: the experimental gave different judgements to the same variable stimulus.

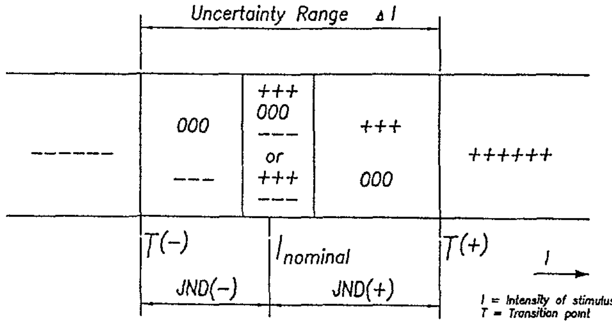


Figure 4 *Perception of stimulus I; + = higher, o = equal, - = lower than reference.*

Results

Many of the experimentals showed a systematic error: forces on one arm were judged differently than the same forces on the other arm. As variations of operating force are more important to note than the absolute force level, the systematic error is not taken into account. Consequently, the relative sensitivity is defined as the uncertainty range over the average of the upper and lower transition point: $\Delta I/(T^{(+)}-T^{(-)})$ (instead of the uncertainty range over the reference force level: $\Delta I/I_{nominal}$).
Figure 5a shows the relationship between reference intensity and relative sensitivity for the seven experimentals. It was found that an operating force level of 2.5N possesses a significantly lower relative sensitivity than operating forces of 5N and larger. For the latter range, the relative sensitivity amounts roughly 0.4 or 40%. Secondly, furnishing the operating levers with perforated self-adjusting shells improves relative sensitivity considerably at low operating forces (figure 5b).

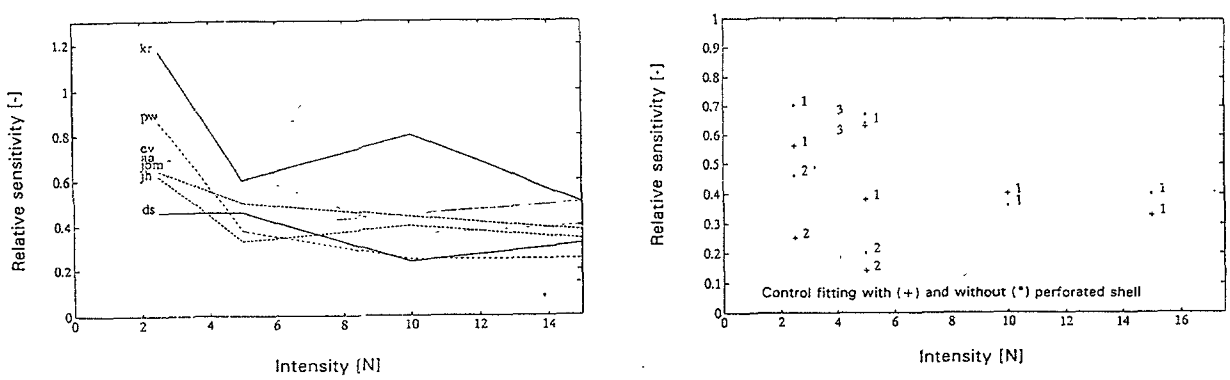


Figure 5 *(a) Relative sensitivity as a function of reference stimulus intensity for seven healthy young men, operating lever without perforated shell. (b) Influence of perforated shell.*

Conclusion

From a control engineering perspective, hand prostheses should be voluntary closing because that control concept possesses most feedback paths. Prostheses must be designed such that full use is made of this control potential. In this study, the operating force feedback was investigated, and it was found that operating forces should be 5N or higher to match the sensitivity of the human upper arm. If the contact area is enlarged by a self-adjusting perforated shell, sensitivity for low operating forces improves.

Future research will include feedback through reaction forces, and combination of force feedback and movement feedback. When feedback is optimised and practical problems are solved, a hand prosthesis with superior control qualities is expected to result.

References

1. Plettenburg DH, Cool JC (1994) Basic requirements for hand prosthesis and orthoses, *Congress Summaries 11th International Congress of the World Federation of Occupational Therapists*, London, p781/3.
2. Kritter E, Waukesha MD (1985) Myoelectric prostheses, *The journal of Bone and Joint Surgery,* 67A(4)654/7.
3. Körner L (1979) Afferent electrical nerve stimulation for sensory feedback in hand prostheses, *Acta Orthopaedica Scandinavica,* supplementum 178, Munskgaard, Copenhagen.
4. Lunteren A van, Lunteren-Gerritsen GHM van, Stassen HG, Zuithof MJ (1983) A field evaluation of arm prostheses for unilateral amputees, *Prosthetics and Orthotics International*, Vol 7, p141/51.
5. Kruit J, Cool JC (1991) Body powered hand prosthesis with low operating power for children, *Journal of Medical Engineering and Technology*, 13(1/2)129/133.
6. Klopsteg PE, Wilson PD (1968) *Human limbs and their substitutes*, Hafner Publishing Company, NY.
7. Radocy B (1986) Voluntary closing control: a succesful new design approach to an old concept, *Clinical Prosthetics and Orthotics*, The American Academy of Orthotists and Prosthetists, 10(2)82/6.
8. Kühlmann P (1982) Brunnstrom (In Dutch), In: Sneep R (ed), *Hemiplegie,* Stafleu ISBN 90-6016-516-0.
9. pinching force is easier dosable
10. Kandel ER, Schwartz JH (1991) *Principles of neural sciences*, third edition, Elsevier, NY.
11. Simpson DC (1974) The choice of control system for the multimovement prosthesis: extended physiological proprioception (e.p.p.), In: Herberts P et al, *The control of upper extremity prostheses and orthoses*, Thomas, Springfield, III p146/50.
12. Herder JL, Kuntz JP, Cool JC, Plettenburg DH (1994) Design of 'energy-free' mechanical structures, theoretical background and application in medical and rehabilitation technology, In: Alsté JA van, *Abstract Book Biomedical Science & Engineering Meeting*, 18-19 may, Garderen, The Netherlands, ISBN 90-365-0671-9, p59/62.
13. Kuntz JP (1995) *Rolling Link Mechanisms*, PhD Thesis Delft University of Technology.
14. Herder JL, Cool JC, Plettenburg DH (1994) The Wilmer voluntary closing hand prosthesis, *Congress Summaries 11th International Congress of the World Federation of Occupational Therapists*, London, p1046/8.
15. Arif B (1990) *Experiments with a prototype of a voluntary closing hand prosthesis* (In Dutch), Report s603, Delft University of Technology.
16. Gescheider GA (1976) *Psychophysics, method and theory*, John Wiley, NY.
17. Stevens SS (1975) *Psychophysics, introduction to its perceptual, neural, and social prospects*, John Wiley and Sons, NY.

Session 8

Manual Control

Chairman: J.C. van der Vaart, The Netherlands

THE ENVIRONMENT PROVIDES THE REFERENCE FRAME FOR SELF-MOTION PERCEPTION

Han F.A.M. van der Steen and Henk H. Kamphuis

Delft University of Technology, Faculty of Aerospace Engineering
P.O. Box 5058, 2600 GB Delft, The Netherlands

e-mail: han.steen@lr.tudelft.nl

Abstract. The perception of self-motion is needed to successfully control body motion in the environment. This perception is originated by senses that register inertial body motions and senses that determine the environment in which the body moves. The perceived environment strongly influences self-motion perception. We argue that this occurs since the motion perceptual system searches for a reference frame in the environment, in order to place inertial body motion signals. We apply this idea to formulate a self-motion perception model. In this model, optimal estimators are applied to tune the inertial and environmental sensory signals and to model attention on specific motion frequency ranges. A quantitative model for perceived self-motion about the vertical axis is described, which is based on literature data.

Keywords: perception, visual, vestibular, visual-vestibular interaction, psychophysics, self-motion, perception models, optimal estimators

1. Introduction

There are different ways to achieve a displacement of the body in the environment. Whatever the means of motion (walking, horse ride, driving a car), a specific set of body actions is required to move. These actions range from moving our limbs as we walk, or the pushing of a button in an elevator, to handling the controls of an aircraft.

In order to successfully control body motion, the perception of the body motion state in the environment is needed: perceiving self-motion. Virtually every sense contributes to the perception of self-motion. Each sensory system has, however, its specific domain for registering motion stimulations. The visual system, for example, provides information on the position and velocity of the environment with respect to the eyes. Inertial sensory systems, like the vestibular system, register forces and accelerations on the body. The inertial systems are therefore insensitive to constant linear body velocities and to rotational velocities about the vertical axis in the horizontal plane [Guedry, 1974; Howard, 1982].

Psychophysical studies have shown that the perceived environment strongly influences the perception of self-motion. The visual environment in particular has a strong effect on perceived self-motion [Wong and Frost, 1978; Büttner and Henn, 1981]. However, it is still unclear how the central nervous system fuses body signals and environmental signals to give a perception of motion of the body in the environment.

We propose a self-motion perception model based on a search for a reference frame in which inertial body motion signals are placed. We argue that the perceptual system seeks for the reference frame in the environment. As a result, environmental motion influences self-motion perception. First, we will consider and discuss perceptual phenomena from psychophysical studies on self-motion. Then we will present a self-motion perception model and show the results of the model for motions about the vertical axis, the yaw motion. This yaw model is based on literature data.

2. Self-motion perception phenomena

Many experiments have been conducted to determine the influence of inertial body stimulation and environmental presentation on perceived self-motion and orientation. Some important results are stated below.

2.1. *Inertial stimulation*

From perception threshold experiments on roll and heave motions [Hosman and Van der Vaart, 1978] and yaw motions [Benson et al., 1989], the sensitivity of self-motion perception in the frequency domain was found similar to that of the primary vestibular afferents in the squirrel monkey [Fernández and Goldberg, 1971; 1976].

From experiments with exclusive inertial yaw motion, it was shown that subjects perceive accelerations quite accurately, but report a fainting motion sensation when the motion is at constant velocity [Parsons, 1970; Guedry, 1974]. The decay of motion sensation is, however, slightly less than the vestibular afferents dynamics suggest. Equivalent dynamics for inertial stimulation is found in cells of the vestibular nuclei in the brain stem [Waespe and Henn, 1977] and in reflexive eye movements [Raphan et al., 1979].

An attentional effect on motion perception thresholds was elegantly demonstrated by Hosman and Van der Vaart [1978]. They determined self-motion perception thresholds by slowly increasing the amplitude of a sinusoidal roll or vertical motion. The subject did not, however, know the motion type nor its frequency. When the subject detected the motion, the amplitude was slowly decreased and subjects had to report when the experienced motion had vanished. A substantial difference between the detection threshold and vanishing threshold was found, indicating an attentional effect.

2.2. *Environmental presentation*

A moving visual perceived environment [Wong and Frost, 1978; Büttner and Henn, 1981] as well as a moving somatosensory [Lackner and Dizio, 1984], auditory [Lackner, 1977], or even tactile or proprioceptive perceived environment [Brandt et al., 1977] can evoke a strong self-motion sensation. This (illusory) self-motion sensation is usually referred to as 'vection'. In neurophysiological experiments on animals, cells in the vestibular nuclei have been found to be activated both during vection and during inertial stimulation of the whole body [Allum et al., 1976; Waespe and Henn, 1977].

Vection does not instantaneously occur but gradually builds up. From experiments on perceived yaw motion in a rotating striped drum, it was found that the time to full self-motion sensation depends on the visual acceleration applied [Melcher and Henn, 1981]. Furthermore, it was found that the layout of the environment is crucial for the generation of vection. When a 'natural' visual environment is at constant roll or pitch velocity around the subject, a complete head-over-heels motion sensation can be experienced [Howard and Childerson, 1994]. A perceived body tilt illusion occurs if a random dot pattern is used instead.

Perhaps introspectively, another phenomenon of environmental influence on self-motion perception can be illustrated. When a train on the opposite track starts to move, a passenger in the stationary train can perceive self-motion due to the relative visual motion in the peripheral visual field. When looking through the opposite window, where the waiting passengers walk on the platform, the perceived self-motion is recognised as an illusion and the self-motion sensation vanishes (almost?) instantaneously. This swift loss of self-motion sensation is not reported, however, from yaw experiments. The after effect of experienced self-motion persists longer when presenting a stationary environment which follows yaw velocity [Howard, 1982].

3. Phenomena discussed

3.1. *Inertial stimulation*

The entire inertial modality is often referred to as vestibular, suggesting that only the vestibular system is stimulated. Individual inertial senses are, however, hard to isolate. Other mechanoreceptors, such as somatosensory receptors, are stimulated as well in most experiments concerning vestibular psychophysics.

An enormous number of neuronal events takes place before the transduction of motion stimuli by the receptors leads to perceived self-motion. It is therefore striking that the magnitude of human self-motion perception from the perception threshold experiments [Hosman and Van der Vaart, 1978; Benson et al., 1989] resembles the squirrel monkey primary vestibular afferents so well.

From the similar dynamical behaviour for yaw motion of cells in the brainstem [Waespe and Henn, 1977], reflexive eye movements [Raphan et al., 1979], and self-motion perception [Parsons, 1970], it can be hypothesised that there is a signal in the brain that registers the motion of the body in the environment; a neuronal correlate for self-motion. This signal gives rise to reflexes of the body and evokes perceived self-motion. A time constant in these phenomena can be considered to be somewhat larger than the vestibular afferent time constant. This indicates a non-direct processing between vestibular afferent and registered self-motion. Raphan et al. [1979] state a storage of neuronal activity related to reflexive eye velocity which lengthens the decay of the vestibular afferent signal: 'velocity storage'.

3.2. *Environmental presentation*

Body motion with a certain magnitude in a specific direction through an inertially stable environment will result in an environmental motion of opposite magnitude and direction with respect to the body. If inertial body motion is absent, however, perceived self-motion appears to rely on environmental information. The reciprocity of body and environmental motion can therefore be considered to be represented in the brain. One way or the other, our perceptual system exploits this reciprocity [Wertheim, 1994].

From the experiments with rotating drums [Büttner and Henn, 1981; Melcher and Henn, 1981] it was shown that motion of the visual environment cannot account completely for self-motion sensation. Above certain acceleration amplitudes of the visual environment, the drum is perceived to move in the laboratory as well.

From the animal neurophysiological experiments [Allum et al. 1976; Waespe and Henn, 1977], it can be argued that there actually is some kind of neuronal correlate in the brain for self-motion in the environment, located in the vestibular nuclei of the brainstem. This signal could be responsible for reflexes that stabilise the body and can be considered to evoke experienced self-motion.

4. The subjective reference frame

From the influence of the environment on perceived self-motion, we argue that the self-motion perceptual system seeks for a reference frame to place inertial self-motion signals. The body needs this inertial reference frame to define its self-motion and to generate reflexes to support the control of body motion in the environment.

We argue that this reference frame is constructed from the perceived environment. The environment only provides, however, a subjective estimation of the inertial reference frame. When this subjective reference frame does not coincide with the inertial reference frame, self-motion illusions, such as vection, can occur. It should be noted that a self-motion illusion only exists when regarding a reference frame that is different from the subjective reference frame, such as the inertial reference frame.

When our body moves, the perceived environment continuously updates the subjective reference frame. The quickly vanishing motion sensation in the train illusion shows that the update of the subjective reference frame can be very sudden: the frame is placed in the less ambiguous platform frame. The latencies of several seconds that are found in the drum experiments, on the other hand, suggest a slower update.

5. Model for self-motion perception

In our model for self-motion perception, we describe perceived body motion which is evoked by inertial stimulation of the body and by environmental presentation. The head is assumed to be fixed to the trunk. The trunk is assumed to be rigid. An inertial stimulation of the whole body consequently is applied to the head. We consider the environment to be perceived only visually.

5.1. *Motion variables*

In the model, we have separated motion of the body (H) and motion of the environment (W).

The input of the inertial systems is body motion in the inertial reference frame. The visual input (V) is composed of a combination of two signals. First, motion of the environment can result in visual motion across the retina. This retinal motion will be referred to as *retinal slip*. We do, however, continuously fixate our eyes on locations in the environment. When being stationary in the environment and following a moving object, for example, the retinal slip will be non-zero, although the environment is perceived to be stationary. The retinal slip is assumed to be compensated by a registration of the eye motion in the socket. This registration is usually referred to as the *efference copy* [Von Holst and Mittelstaedt, 1950]. The efference copy is the second input of the visual system.

To incorporate retinal slip (R_s) and efference copy (EC), eye motion in the socket (E) is included in the model.

5.2. *Inertial and visual model inputs*

The inertial sensory systems are sensitive to forces on, and accelerations of the body in the inertial reference frame.

The motion of the eyes is described by the gaze (G); the sum of eye and head motion in the inertial reference frame:

$$G = E + H \tag{1}$$

When gaze and visual environmental motion are unequal, a flow of the visual environment across the retina will occur: retinal slip. This slip is defined as:

$$R_s = W - G \tag{2}$$

The input of the visual system is the subtraction of retinal slip and efference copy:

$$V = R_s - EC \tag{3}$$

5.3. *Inertial and visual dynamics*

The inertial dynamics (T_{ine}) are specific for the motion type, and are considered to resemble receptor physiology. For inertial yaw motion, there hardly is any other sensory system stimulated than the vestibular system. Consequently, the physiology found by Fernández and Goldberg [1971] can be applied for the inertial yaw dynamics. The visual dynamics (T_{vis}) is modelled as a time delay. The magnitude of this delay represents the slower perception of visual motion when compared to the processing of inertial motion [Hosman and Van der Vaart, 1988].

5.4. *Neural filters*

Neural filters are applied to transform inertial and visual afferent signals to an estimation of the stimulation that led to the signal. This transformation is considered as neuronal processing of the afferent signal in a motion estimate of the original stimulation. The visual and inertial neural filters are referred to as NF_{vis} and NF_{ine}, respectively, the output of the filters are H' and V'; the estimates of inertial and environmental motion.

To construct the original stimulation from an afferent signal, the sensory process that leads to the afferent signal has to be inverted. The inverse process therefore needs the characteristics of the sensory process. In the central nervous system a sensory signal will, however, always be corrupted by signals that do not correlate to the motion stimulation (noise). Consequently, no inverse process can be defined. In the neural filters, an optimal estimator (Kalman filter) is applied to handle the uncorrelated signals and achieve a (pseudo-)inverse process. Each neural filters contains an internal representation (IM) of the dynamics of the inertial and visual transformation of stimulus to afferent signal, and the intensities of the uncorrelated signals. Appendix A contains a more detailed description of the neural filters.

The neural filters are constructed in such a way that an attentional influence on the estimation can be modelled. This is achieved by shaping the input noise of the internal model to a specific frequency domain; see Appendix A.

5.5. *Visual attraction*

The heart of the model is the updating process by the subjective reference frame which is provided by the perceived environment. The visual environment is modelled to 'attract' the inertial motion estimate towards the subjective reference frame. This attraction mechanism is achieved by an interaction of the inertial and visual motion estimates, H' and V' respectively; see Figure 1.

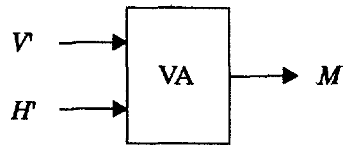


Fig. 1. The visual attractor combines the inertial and visual motion estimates and generates the motion signal.

5.6. *Self-motion signal*

The motion signal M is supposed to give rise to perceived self-motion in the subjective reference frame. The motion signal also directs eye motion in the socket (E) through the oculomotor system (T_{oms}).

The complete model is depicted in Figure 2.

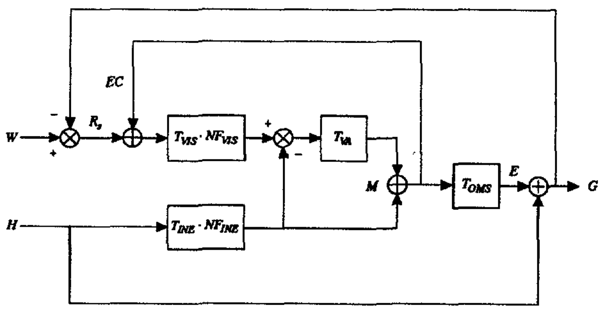


Fig. 2. The self-motion perception model.

6. Model for yaw motion

Results from experiments on neurophysiology, eye movements and perception of body motion are well documented for yaw motion. For this motion type, the model is described quantitatively and is based on literature data.

In the yaw model, the visual dynamics (T_{vis}) is modelled as a time delay of 150 ms, see Appendix B. This time delay incorporates both the delay of the retinal slip and the efference copy. The inertial dynamics (T_{ine}) relates input acceleration and afferent output and is modelled as the dynamics found by Fernández and Goldberg [1971]. This dynamics resembles an overdamped torsion pendulum with an additional lead component; see Appendix B.

The visual attractor compares estimated environmental yaw velocity and estimated inertial yaw velocity. The difference between these estimates is first-order low-pass filtered. The filtered difference and the inertial estimate are added and generate the motion signal M.

The oculomotor dynamics, generating compensatory horizontal eye movements from the motion signal (M), is considered to be unity. This contrasts to, for example, roll motions where compensatory eye movement can only reach a small torsion angle.

When the gains of both the efference copy and the retinal slip are taken unity and both signals have a same time delay in the central nervous system, the signals are synchronised, the visual input can be described by:

$$\begin{aligned} V &= W - G + EC \\ &= W - (E + H) + M \\ &= W - (M + H) + M \\ &= W - H \end{aligned} \qquad (4)$$

With these assumptions, the feedback loops of the efference copy and the retinal slip are avoided. Then, the model only has W and $(W - H)$ as the inputs. $W - H$ is the yaw motion of the visual environment with respect to the head. The yaw model is depicted in Figure 3.

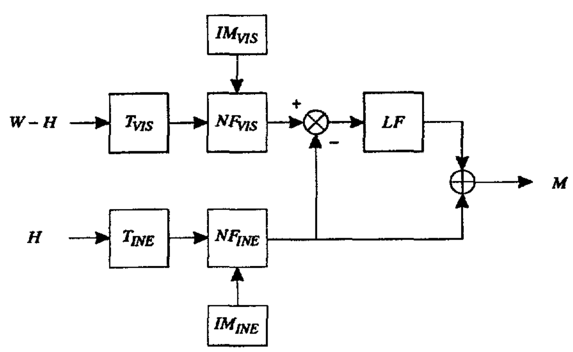


Fig. 3. The yaw model with H and $(W - H)$ as the input and the motion signal M as the output, containing the sensory dynamics (T), the neural filters (NF) with internal models (IM), and the low pass filter dynamics of the visual attractor (LF).

The settings in the neural filters are chosen for the specific stimulus condition tested. These settings are described next.

7. Settings of the neural filter parameters

The neural filters play an essential role in the model. The settings of the neural filters are the frequency domain of the (expected) motion stimulation and the noise intensities of the afferent signals (see Appendix A). The influence of these settings are demonstrated for both the inertial and the visual neural filter.

7.1. *Inertial neural filter*

The inertial neural filter provides the estimate of inertial stimulation of the body. Therefore, it contains the inertial sensory dynamics. The afferent signal is provided by T_{ine}, see Figure 3. A Bode plot of the inertial sensory dynamics [Fernández and Goldberg, 1971] is depicted in Figure 4. It should be noted that the input is a sinusoidal inertial angular velocity and the output is afferent firing rate.

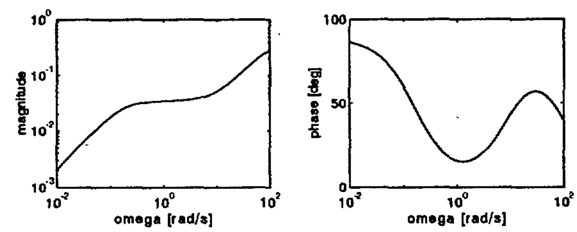


Fig. 4. Bode plot of the inertial sensory dynamics for yaw, according to Fernández and Goldberg [1971].

The neural filter is applied to estimate the angular velocity in the frequency range of interest. This range is set by the shaping filter (see Appendix A). The effect of the filter on the afferent signal for a frequency range of 0.1 to 10 rad/s is demonstrated in Figure 5.

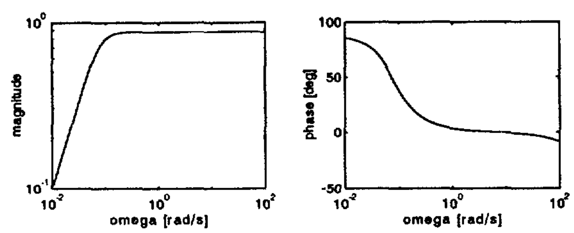


Fig. 5. Bode plot of the cascade of the inertial dynamics and its neural filter for a frequency range of 0.1-10 rad/s.

When comparing Figures 4 and 5, it can be seen that the transfer function of the cascaded systems is close to unity in the frequency range selected by the shaping filter, resulting in a unity gain and zero phase.

7.2. *Visual neural filter*

The dynamics of the cascaded visual dynamics and its neural filter is depicted in Figure 6. The selected frequency range is 0-0.1 rad/s.

The improvement of the visual transfer function

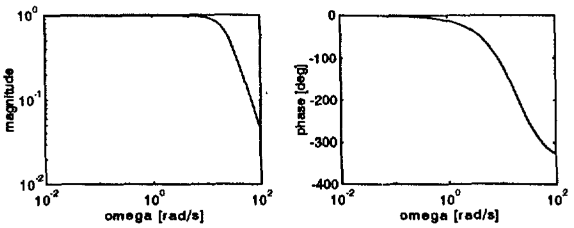


Fig. 6. Bode plot of the cascade of the visual dynamics and its neural filter.

is small. Due to large phase lags in the visual dynamics, a higher frequency range selected results in an insignificant improvement of phase lag while the gain drops far below unity.

8. Model results

The results of the model for self-motion perception are shown for three conditions. First, the results for an inertial yaw motion with a visible environment are presented: a rotation in the light (RL). Secondly, the results are shown for an inertial yaw motion without visual information: a rotation in the dark (RD). Finally, the results for a yaw motion with only visual input are shown. The latter type of stimulation is referred to as circular vection (CV).

8.1. *Rotation in the light*

In Figure 7 can be seen that the model correctly duplicates the sensation of yaw motion in a visual environment. The time constant of the visual attractor, τ_{va}, is chosen 1 s. The effect of the visual attractor reduces as τ_{va} increases. When increasing the time constant, perceived angular velocity sticks to the gain that is provided by the inertial estimate. The gain of the velocity estimate during rotation in the light can be improved by selecting a smaller value for τ_{va}. Ultimately, the gain converges to unity when τ_{va} is set to zero. The perceived self-motion will then be reciprocal to the motion of the visual environment. Independent of the time constant of the visual attractor, however, the response converges to the body velocity, as defined by the visual environment as the subjective reference frame.

8.2. *Rotation in the dark*

Model parameter settings for rotation in the dark are different from the settings needed for model simulations in the light. The noise-signal ratio of the visual system in the neural filter (see Appendix A) has to be set much higher because of the lacking input of the retinal receptors in the dark. Secondly, because of lacking visual information,

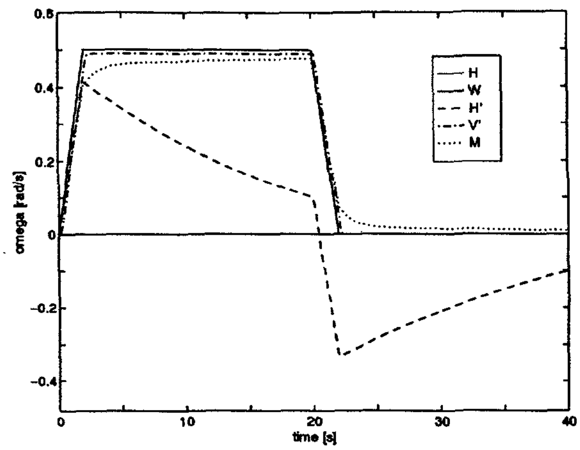


Fig. 7. Simulation of perceived rotation in the light with $\tau_{va} = 1s$ ($\beta_{ine} = 0.1$-$10\ rad/s$, $\rho_{ine} = -60\ dB$, $\beta_{vis} = 0.1\ rad/s$, $\rho_{vis} = -80\ dB$).

the visual dominance is much weaker. As a consequence, τ_{va} is taken larger. In the simulation, τ_{va} is 80 s.

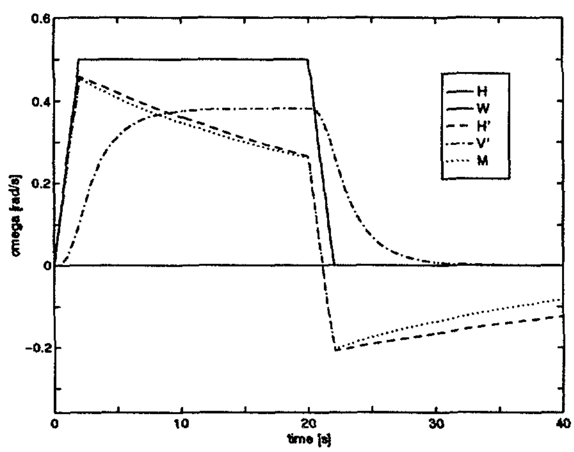


Fig. 8. Simulation of rotation in the dark with $\tau_{va} = 80\ s$ ($\beta_{ine} = 0.01$-$5\ rad/s$, $\rho_{ine} = -60\ dB$, $\beta_{vis} = 0.1\ rad/s$, $\rho_{vis} = 0\ dB$).

Figure 8 shows the decreasing yaw sensation when accelerations are absent. This decrease is slower than the decrease of the afferent output of the inertial receptors.

8.3. *Circular vection*

The responses to exclusive visual stimulation are solely dependent on the parameters of the visual system (time delay and visual neural filter settings). The most important parameter is the strength of the environmental updating, which is expressed in the time constant of the visual attractor.

Figure 9 shows that the model mimics the influence of environmental motion on perceived self-motion. The sensitivity of the model responses

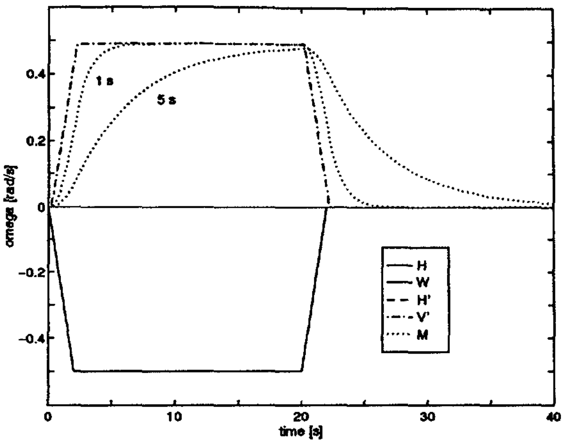


Fig. 9. Simulation of circular vection with $\tau_{va} = 1\ s$ and $5\ s$ ($\beta_{ine} = 0.1$ - $10\ rad/s$, $\rho_{ine} =$ -60 dB, $\beta_{vis} = 0.1\ rad/s$, $\rho_{vis} =$ -80 dB).

to changes of the visual attractor time constant is also shown in Figure 9; time constants of 1 s and 5 s are tested. The results from the time constant of 5 s resemble the circular vection experiments with rotating drums [Büttner and Henn, 1981].

9. Discussion

9.1. *Model outline and performance*

In the literature, different models have been suggested to account for the interaction of visual and inertial signals [Robinson, 1977; Raphan et al., 1979; Zacharias and Young, 1981; Borah et al.; 1988]. Those approaches concentrated on the notion that motion of the visual environment unquestioningly drives the self-motion signal; both to evoke compensatory eye movement and to perceive self-motion. Our proposal that the perceptual system pursues a reference frame in the environment for perceived self-motion is of a greater generality. As we showed for the perception of yaw motions, this concept appears to form an excellent basis for describing self-motion phenomena.

The optimal estimation method in a model for human spatial orientation was first applied by Borah, Young and Curry [1988]. The purpose of the Kalman filter in their model is, however, different from our purpose. They applied the filter to describe self-motion perception from visual-vestibular interaction. We have applied optimal estimators in the visual and inertial neural filters to tune the afferent signal from a single sensory modality. The filters straighten the magnitude and improve the phase of the afferent output.

With the shaping filter in the internal model of the neural filter, the bandwidth of the expected stimulus frequency range can be selected. One of the consequences is that a longer persisting self-motion sensation in the dark can be achieved, usually referred to as 'velocity storage' [Raphan et al., 1979]. In our model, this storage is not explicitly modelled. It is the result of an 'attempt' of the neural filter to estimate body velocity in the low frequency domain of inertial velocities.

Long latencies in the build up of vection emerge from rotating drum experiments. We have modelled this by selecting a visual attractor time constant of about 5 seconds. It should be noted, however, that a 'natural' environment lacks in the drum experiments. We think that the visual attraction can be described by a much smaller time constant of about 1 s, or even less, when a natural visual environment is applied [Howard and Childerson, 1994].

9.2. *Model extension*

As a good approximation, only the semicircular canals are stimulated when applying an inertial yaw motion about the gravitational vector. The dynamics of this stimulus transduction can therefore be described by the well documented vestibular physiology. When extending the model to more motion types, the question emerges whether or not a single (dominant) sensory system can be taken as the generator of afferent signals.

A reference frame that is in uniform linear motion with respect to the earth inertial frame, is inertial by itself. Therefore, we think that the reference frame updating by the visual attractor has similar dynamics for linear motions as it has for pure yaw motions. The updating dynamics could, however, be different for motions in which the orientation with respect to the gravity vertical changes.

The extension of the model for a non-rigid body will result in an extended definition of the efference copy. The efference copy has to be taken as a combination of eye motion in the socket and head motion on the trunk.

Attentional effects on perceived self-motion can be described when applying the concept of a shaping filter in the internal model. This enables future modelling of an 'active observer' in this model.

Acknowledgment

The Netherlands organisation for scientific research (NWO) is gratefully acknowledged for sponsoring this research.

10. REFERENCES

Allum JHJ, Graf W, Dichgans J and Schmidt CL (1976). *Visual-vestibular interactions in the* *vestibular nuclei of the goldfish.* Experimental Brain Research, vol. 26; pp. 463-485.

Benson AJ, Hutt ECB and Brown SF (1989). *Thresholds for the perception of whole body angular movement about a vertical axis.* Aviation, Space, and Environmental medicine, vol. 60; pp. 205-213.

Borah J, Young LR and Curry RE (1988). *Optimal estimator model for human spatial orientation.* In: 'Representation of three-dimensional space in the vestibular, oculomotor, and visual systems', by Cohen and Henn (Eds.). Annals of the New York Academy of Sciences, vol. 545; pp. 51-73.

Brandt Th, Büchele W and Arnold F (1977). *Arthrokinetic nystagmus and ego-motion sensation.* Experimental Brain Research, vol. 30; pp. 331-338.

Büttner U and Henn V (1981). *Circularvection: psychophysics and single-unit recordings in the monkey.* Conference proceedings of the Bárány Society; pp. 274-283.

Fernández C and Goldberg JM (1971). *Physiology of peripheral neurons innervating semicircular canals of the squirrel monkey. II. Response to sinusoidal stimulation and dynamics of the peripheral vestibular system* Journal of Neurophysiology, vol. 34(4); pp. 661-675.

Fernández C and Goldberg JM (1976). *Physiology of peripheral neurons innervating otolith organs of the squirrel monkey. III. Response dynamics.* Journal of Neurophysiology, vol. 39(5); pp. 996-1008.

Guedry FE (1974). *Psychophysics of vestibular sensation.* In H.H. Kornhuber (Ed.): 'Handbook of sensory physiology', vol. 6(2). Berlin: Springer-Verlag.

Holst E von and Mittelstaedt H (1950). *Das reafferenz Prinzip (The reafference principle).* Naturwissenschaften, vol. 37; pp. 464-476.

Hosman RJAW and Vaart JC van der (1978). *Vestibular models and thresholds of motion perception. Results of tests in a flight simulator.* Report LR-265, Faculty of Aerospace Engineering, Delft University of Technology, The Netherlands.

Hosman RJAW and JC Van der Vaart (1988). *Visual-vestibular interaction in pilot's perception of aircraft or simulator motion.* Proceedings of the AIAA Flight Simulation Technology Conference, AIAA paper 88-4622-CP

Howard IP (1982). *Human spatial orientation.* New York: Wiley.

Howard IP and Childerson (1994). *The contribution of motion, the visual frame, and visual polarity to sensations of body tilt.* Perception, vol. 23; pp. 745-751.

Lackner JR (1977). *Induction of illusory self-rotation and nystagmus by a rotating sound-field.* Aviation, Space, and Environmental Medicine, vol. 48(2); pp. 129-131.

Lackner JR and DiZio P (1984). *Some efferent and somatosensory influences on body orientation and oculomotor control.* In: Spillmann & Wooten (Eds.); Sensory experience, adaptation, and perception. Hillsdale: Erlbaum.

Melcher GA and Henn V (1981). *The latency of circular vection during different accelerations of the optokinetic stimulus.* Perception and Psychophysics, vol. 30; pp. 552 - 556.

Parsons RD (1970). *Magnitude estimates of the oculogyral illusion during and following angular rotation.* Journal of Experimental Psychology, Vol. 84; pp. 230-238.

Raphan Th, Matsuo V and Cohen B (1979). *Velocity storage in the vestibulo-ocular reflex arc (VOR).* Experimental Brain Research, vol. 35; pp. 229-248.

Robinson DA (1977). *Linear addition of optokinetic and vestibular signals in the vestibular nucleus.* Experimental Brain Research, vol. 30; pp. 447-457.

Waespe W and Henn V (1977). *Neuronal activity in the vestibular nuclei of the alert monkey during vestibular and optokinetic stimulation.* Experimental Brain Research, vol. 27; pp. 523-538.

Wertheim AH (1994). *Motion perception during self-motion: the direct versus inferential controversy revisited.* Behavioral and Brain Sciences, vol. 17(2); pp. 293-355.

Wong SLP and Frost BJ (1978). *Subjective motion and acceleration induced by the movement of the observer's entire visual field.* Perception and Psychophysics, vol. 24(2); pp. 115-120.

Zacharias GL and Young LR (1981). *Influence of combined visual and vestibular cues on human perception and control of horizontal rotation.* Experimental Brain Research, vol. 41; pp. 159-171.

A. Mathematical outline of neural filters

The neural filters represent low-level signal processing. They transform afferent output in an optimal estimate of sensory stimulations under a priori assumptions about:

1. the sensory system dynamics
2. the expected stimulus spectrum
3. the noise in each sensory measurement

These assumptions are referred to as the internal model. If the internal model satisfies the following conditions:

1. linear system representations
2. Gaussian white processes
3. stationary

the neural filter can be modelled as an optimal linear estimator or the well-known steady state

Kalman filter. Below, the complete set of equations are summarised to calculate a neural filter.

The transfer functions of the sensors are formulated in state-space:

$$\begin{aligned} \dot{\bar{x}}_1 &= A_1\bar{x}_1 + B_1\bar{u}_1 \\ \bar{y}_1 &= C_1\bar{x}_1 + D_1\bar{u}_1 + \bar{v} \end{aligned} \tag{5}$$

The state-space notation of the shaping filter:

$$\begin{aligned} \dot{\bar{x}}_2 &= A_2\bar{x}_2 + E_2\bar{w} \\ \bar{y}_2 &= C_2\bar{x}_2 \end{aligned} \tag{6}$$

If $\bar{u}_1 = \bar{y}_2$, then the shaping filter and the sensor are connected in series:

$$\begin{bmatrix} \dot{\bar{x}}_2 \\ \dot{\bar{x}}_1 \end{bmatrix} = \begin{bmatrix} A_2 & 0 \\ B_1C_2 & A_1 \end{bmatrix} \begin{bmatrix} \bar{x}_2 \\ \bar{x}_1 \end{bmatrix} + \begin{bmatrix} E_2 \\ 0 \end{bmatrix} \bar{w} \tag{7}$$

or:

$$\begin{aligned} \dot{\bar{x}} &= A\bar{x} + E\bar{w} \\ \bar{y} &= C\bar{x} \end{aligned} \tag{8}$$

The steady state covariance of the state is the solution of the Lyapunov equation:

$$\begin{aligned} \dot{\Phi}_{\bar{x}\bar{x},s} &= A\Phi_{\bar{x}\bar{x},s} + \Phi_{\bar{x}\bar{x},s}A^T + EQE^T \\ &= 0 \end{aligned} \tag{9}$$

Where Q is the system noise power spectral density.

The steady state covariance of the observation is:

$$\Phi_{\bar{y}\bar{y},s} = C\Phi_{\bar{x}\bar{x},s}C^T \tag{10}$$

The observation noise power spectral density is defined as a fraction of the covariance of the observation:

$$R(i,j) = \begin{cases} 0 & \text{if } i \neq j \\ \rho(i)\Phi_{\bar{y}\bar{y},s}(i,j) & \text{otherwise} \end{cases} \tag{11}$$

The steady state error covariance of the Kalman filter is the solution of the Ricatti matrix equation:

$$\begin{aligned} \dot{\Sigma}_s &= A\Sigma_s + \Sigma_sA^T + EQE^T - \Sigma_sC^TR^{-1}C\Sigma_s \\ &= 0 \end{aligned} \tag{12}$$

The steady state Kalman filter gain matrix:

$$G_{kf} = \Sigma_sC^TR^{-1} \tag{13}$$

The differential equation for the optimal observer is formulated as follows:

$$\dot{\hat{x}} = A\hat{x} + G_{kf}(\bar{y}_1 - C\hat{x}) \tag{14}$$

or

$$\dot{\hat{x}} = A_{kf}\hat{x} + G_{kf}\bar{y}_1 \tag{15}$$

where

$$A_{kf} = A - G_{kf}C \tag{16}$$

An observation equation can be defined by only measuring the sensory stimulation, $\hat{u}_1$:

$$\bar{y}_{kf} = \hat{u}_1 = \begin{bmatrix} C_2 & 0 \end{bmatrix} \begin{bmatrix} \hat{x}_2 \\ \hat{x}_1 \end{bmatrix} \tag{17}$$

or,

$$\hat{u}_1 = C_{kf}\hat{x} \tag{18}$$

where

$$C_{kf} = \begin{bmatrix} C_2 & 0 \end{bmatrix} \tag{19}$$

The state-space relating the afferent signal, $\bar{y}_1$, and the estimated sensory stimulation, $\hat{u}_1$, is summarised as:

$$\dot{\hat{x}} = A_{kf}\hat{x} + G_{kf}\bar{y}_1 \tag{20}$$

$$\hat{u}_1 = C_{kf}\hat{x} \tag{21}$$

B. Transfer functions of the sensory systems

The transfer function of the visual system in Laplace domain is modelled as a time delay. The time delay is approximated by a Padé filter to maintain model linearity. The Padé filter is complemented with a low-pass filter to account for physiological constraints on receptor transduction:

$$H_{vis}(s) = \frac{1}{1+\tau_v s}\frac{1-\frac{\tau_d}{2}s}{1+\frac{\tau_d}{2}s} \tag{22}$$

where τ_v is the time constant of the low-pass filter, and τ_d is the effective time delay of the visual system. The following values are implemented in the yaw model:

$$\tau_v = 0.1s \tag{23}$$

$$\tau_d = 0.15s \tag{24}$$

The transfer function of the whole body inertial motion sensors is according to Hosman and Van der Vaart [1978]. Relating input angular velocity and sensory afferent output, this function is described as:

$$H_{ine}(s) = \frac{s(1+0.11s)}{(1+5.9s)(1+0.005s)} \tag{25}$$

To transform input angular acceleration to input angular velocity, in which the visual output is defined, the numerator must be multiplied by s.

A Comparison Between System Theoretic Models and Neural Networks applied to Car Driving

P.H. Wewerinke* and H. Hilberink*

Abstract

This paper deals with two modeling appraoches to car driving. The first one is a system theoretic approach to describe adaptive human driving behavior. The second approach utilizes neural networks (NN).

The two model structures are compared based on criteria such as a priori knowledge of the process, optimality criteria, convergence speed, etc. Both a linear system and a nonlinear system is investigated.

In a simulation study both the lane keeping task and the overtaking task are considered. Both the model results and the NN results are compared and discussed in the paper.
The following step is to use an extended Kalman filter to estimate assumed unknown model parameters.
The next step is to simulate the car driving tasks in an experiment with real human operators, with key variable the level of experience. The objective is to assess the relative merit of both approaches to describe human learning behavior in car driving specifically and in operating dynamic systems in general.

1 Introduction

Man-machine systems often exhibit time-varying and adaptive characteristics. This can be caused by system- and environmental-related changes and/or by time-varying human operator behavior, adapting to the varying task or learning to optimize the performance on a given task. Globally this will be indicated in this paper with adaptive human operator (HO) behavior.

One popular approach to analyse and evaluate adapting man-machine systems is based on mathematical models. The subject of this paper is to compare two model approaches; the first one is rooted in system theory, the second one is based on the use of neural networks.
Generally system theoretic models have a normative structure in terms of goals, tasks, functions and actions, etc. System theoretic models can describe this hierarchy of interrelated aspects quantitatively.
A neural network (NN) is an input-output model of the HO. Based on data (experience) a desired (or given) input-output relationship can be learned.

Although the aim of this research is to assess the relative merit of both model approaches to describe adaptive human behavior in operating dynamic systems in general, car driving is considered as a specific application. This allows a concrete task analysis, HO functions and simulation

*Department of Applied Mathematics, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands

results.
The overall goal of car driving is to go from A to B in a certain way (safely, in a given time, etc.). The principal tasks derived from this are lane keeping, car following and overtaking slower vehicles, avoiding a collision with other cars. The system model of these tasks is discussed in the next chapter.

Chapter 3 deals with the neural network approach. Chapter 4 contains an analysis of both model structures. In Chapter 5 the simulation results of adaptive driver behavior are presented. Chapter 6 contains the concluding remarks.

2 System modeling of car driving

Car driving consists of a number of primary and secundary tasks, which result from the overall goal to go from A to B, such as lane keeping, car following, overtaking, navigating, etc. In the following lane keeping and overtaking will be considered as two representative examples.

2.1 Lane keeping

It can be assumed (Ref. 1) that lane keeping is based on aiming at a distance d ahead. The direction of this aimpoint ψ_a is a result of the car heading ψ and the lateral position y (Fig. 1), according to (assuming small angles)

$$\psi_a = \psi + y/d \qquad (1)$$

Equation (1) shows that for large d $\psi_a \approx \psi$ and for small d $\psi_a \approx y/d$; in other words, depending on the 'looking' distance ahead d, the driving task resembles more a relatively easy heading control task or a relatively difficult position control task.
Assuming that the driver is generating a steering wheel deflection δ proportional to the aimpoint direction ψ_a (gain K_h), with a first order lag with a time constant T_h, the closed loop system can be represented by the block diagram of Fig. 2.

A simple root locus analysis (Fig. 3) shows that aimpoint control resembles more heading control or position control depending on the additional lead (zero at $-u/d$, with u the forward driving speed) corresponding to the implicit heading feedback. The value of d determines the trade-off between system stability and the bandwidth of the path mode. So d can be optimized and can be related to the driver's experience level.

2.2 Overtaking

Overtaking is the most complex driving subtask including observing, information processing, decision making, planning, maneuvering and other traffic. The overtaking situation is shown in Fig. 4 for a two-lane road. The analysis for freeways is similar.

Assume the situation that car i intends to pass the preceding car j at distance X_j ahead. The possibility to do so depends on the oncoming cars k and ℓ. At time $t_0 = 0$ car i decides at distance X_j from car j to accelerate from speed u_j to speed u_m, with time constant T_u. The distances to cars k and ℓ (with speed u_ℓ) are X_k and X_ℓ, respectively. At time t_1 car i changes to the left lane. At time t_2 car j has been overtaken and car i returns to the right lane again.

The two conditions which must be satisfied are that at t_1 car k must be passed and car ℓ must be away far enough to allow the overtaking maneuver. This implies the two inequalities

$$X_k < X_j + (u_\ell + u_j)t_1 - S_k \tag{2}$$

$$X_\ell > X_j + S_j + S_\ell + (u_\ell + u_j)t_2 \tag{3}$$

with S_j the distance which car i is overtaking in the left lane with respect to car j and S_k and S_ℓ the safety distances to car k and ℓ, respectively.

Useful approximations can be derived (Ref. 2) for t_1 and t_2 (for t_1 and t_2 smaller than T_u). The result is

$$t_1 \approx \sqrt{\frac{2.7X_jT_u}{u_m - u_j}} \quad \text{and} \quad t_2 \approx \sqrt{\frac{2.7(X_j + S_j)T_u}{u_m - u_j}} \tag{4}$$

Combining eq. (4) with (2) and (3) shows that the decision criteria (2) and (3) depend on X_j (which can be identified with the overtaking strategy), car dynamics (T_u), speeds and safety margins.
The required gap between car k and car ℓ $(X_\ell - X_k)$ can be determined as a function of X_j. It turns out that the required gap can be substantially reduced already for a small X_j. So the possibility to trade-off $X_\ell - X_k$ and X_j allows car i to optimize its overtaking strategy depending on the momentaneous traffic situation.

The speed of car i is represented by a first order process driven by a commanded speed (u_c) and a first order disturbance input (w_u). Other traffic is modeled with a constant average speed. The specific speed of each car is unknown to car i and has to be estimated based on the observed (derivative of) distances to the other cars.
The system model of the total process can be expressed in the general form

$$X(k+1) = f(X(k), U(k), W(k)) \tag{5}$$

with $X = col(w_{u_i}, u_i, X_j, X_k, X_\ell, X_{ik}, X_{i\ell})$ with X_{ik} and $X_{i\ell}$ the adjoint states representing the inequalities (2) and (3), combined with (4).

The distances to other cars can be estimated based on the outside visual cues. The nonlinear relationships can be expressed in the general form

$$y_p(k) = g(X(k-i)) + v(k-i) \tag{6}$$

with v the observation noise and i the time delay. By means of an extended Kalman filter the states can be estimated. The estimated variables X_{ik} and $X_{i\ell}$ are used to decide whether (D_1) or not (D_0) overtaking is possible, in case the distance to the slower preceding car becomes too small. This involves a sequential decision process to determine continuously the mode of car i: 1. car following with speed u_j, 2. driving with speed u_c and 3. overtaking, which is basically a pre-program maneuver (accelerating at t_0 till u_m and lane changes at t_1 and t_2).

2.3 Adaptive driver behavior

Adaptive driver behavior can be related to the level of experience and to the adaptive human capability to traffic system developments. This can be related to

- the knowledge of the car response to control inputs (internal model of the vehicle dynamics);
- the use of the visual cues (accuracy, attention allocation);
- control strategy;
- the relationships between task variables, on which overtaking decisions and strategy, and time coordination are based.

In the system theoretic approach adaptive human behavior is modeled as an adaptive estimation process of unknown model parameters based on new data (experience). The question is what one's initial knowledge is of the process in a concrete situation, but conceptually it can be assumed that a given set of parameters θ of the system model are unknown. So learning is modeled as a parameter estimation problem.
The procedure to solve this is to add the parameters to the state vector (using the parameter model $\theta_{k+1} = \theta_k$) yielding an augmented nonlinear system with state $\bar{x} = col(x, \theta)$ of the same form as equations (5) and (6). This can be solved by means of an extended Kalman filter to estimate $\bar{x}$ and thus x and θ.

3 Neural Networks

Human operator behavior can be described as the relationship between task inputs y and control outputs u (inputs to the system). Learning this functional relationship between y and u can be described by a neural network (NN).

Basically a NN consists of a number of processing elements (or neurons) with weighted connections (Ref. 3). The weights represent the memory of the network and determine the input-output relationship. The NN has a given structure, determined by the number of neuron layers, the number of neurons per layer and the connections between the neurons. In this study a commonly used structure is assumed with one input layer, determined by the inputs y, one output layer, determined by the outputs u and one hidden layer. Only feedforward connections are assumed. Such a structure is known (Ref. 4) to be able to approximate any arbitrary input-output relationship (function).

Learning the NN, i.e. the scheme to adjust the weights, can be based on an explicit performance measure, e.g. to obtain a desired output. This is called supervised learning. A common learning strategy is the so-called backpropagation. This is also assumed in this study. Another learning strategy is called unsupervised learning, in which the adaptation scheme is not depending on the NN output.
The backpropagation algorithm can be considered as a first order gradient (steepest descent) method. It has been shown to provide good optimization results in many applications with favourable computational simplicity. More specific issues like convergence rate, final weights corresponding to a local minimum will have to be considered, but related to the concrete question how well the NN results agree with the pertinent human learning behavior.

For the car driving task the HO inputs y consist, for the lane keeping task, of heading and lateral deviation, and, for the overtaking task, of speed and relative distances to preceding and oncoming cars. The outputs u consist of steering wheel deflection, gas and brakes.

4 An analysis of both model structures

The system theoretic structure and the NN structure will be compared with respect to the following criteria

- a priori knowledge of the process;
- what is estimated;
- estimation criteria;
- iteration schemes;
- convergence;
- linear versus nonlinear systems.

Any a priori knowledge of the process can be incorporated in the system theoretic structure in terms of f and g, in equations (5) and (6), respectively. The quantitative relationship between system inputs and outputs can be expressed in terms of the parameter values and form of nonlinear difference equations.
In the NN structure no a priori knowledge is taken into account. This knowledge is, in principle, related to the initial weights, but this relationship is unclear. So, in practice, no a priori knowledge can be taken into account and generally arbitrary initial weights are used in the iteration procedure. However, in case it is known (or assumed) that an output is depending on certain variables, these variables can be assumed as NN inputs possibly with appropriate time delays similar to the linear ARMA model (Ref. 4). This is similar to the system theoretic approach with a given system theoretic structure but with unknown model parameters.

The system theoretic structure intends to estimate the model parameters and the system state based on the system input and (noisy) output. Based on the estimated state, also the output can be estimated. Based on the estimated system model new outputs can be generated for new inputs.
The forgoing parameter estimation scheme, on the basis of an extended Kalman filter, involves a nonlinear system model. This implies that it is unclear how good the estimate is (i.e. in a minimum variance sense), depending on the correctness of the system model. Only for a linear system the Kalman filter generates a minimum variance estimate.

In case the backpropagation scheme is used for the NN approach, also a mean-square estimation error criterion (E) is utilized, so the NN tries to estimate a desired output. The NN weigts are adjusted on the basis of a steepest descent criterion

$$W_{i+i} = W_i - \alpha \frac{\partial E}{\partial w}. \qquad (7)$$

The resulting weights correspond with a local extremum. Therefore, it is unclear how good the NN estimation results will be, although conceptually the iteration scheme is aimed at a minimum error variance solution.

The extended Kalman filter updates the estimated output of the previous time step by including the difference between the expected output and the measured output optimally, i.e. yielding a minimum variance estimation error in case the system model is correct. Convergence speed depends on this correctness of the system model and on the filter dynamics.

In the NN approach the convergence is determined by the first order method of equation (7), including the learning rate α. The value of α is a trade-off between convergence speed and overshooting. The simplest procedure involves a constant value of α. There are methods with variable α to improve the convergence speed (Ref. 2).

5 Simulation of adaptive driver behavior

As a first step the lane keeping task was considered. A linear system model was assumed (Ref. 2) to relate heading and lateral position to the steering angle, with a driving speed of 30 m/s. Human control response was modelled as a first order lag with a time constant of 0.5 second and an optimal gain.
The task considered was a commanded lane change of 1 meter. After 10 seconds an additional change of 1 meter was commanded. Ten seconds later a reduction of the lateral position to, again, 1 meter was commanded and, finally, after 10 seconds a reduction to zero was commanded.

The system model response is shown in Fig. 5. The model exhibits a stable response but with a delay of about 5 seconds (mainly determined by the aimpoint distance of 100 m).
The NN was trained at the lane keeping task by presenting the system model results of 1000 lane changes between (randomly) 0 and 2 meter at intervals between 0 and 10 seconds. After that, the same commanded lane changes as discussed above were presented to the NN. The resulted response is also given in Fig. 5, showing a rather close agreement with the system model results.

The overtaking task including the system model as described in Chapter 2 was simulated. It was assumed that all cars have a constant speed of 20 m/s and that car i tries to drive 30 m/s (u_m). The actual average speed, in the following expressed as the velocity v_r above the 20 m/s, is the key overtaking performance measure. The experimental (independent) variables considered are the traffic densities in the right and left lane (d_r and d_ℓ, in average number of cars per meter with a Poisson distribution), type of car, in terms of the time constant T_u and overtaking strategy, in terms of the parameter X_d (distance to the preceding car at the moment of accelerating).

The overtaking process is described in terms of the following phases (modes): mode 1: car i has to wait behind car j before overtaking is possible; mode 2: car i can accelerate because the preceding car j is away far enough; mode 3: car i can overtake and accelerate in the right lane; mode 4: car i overtakes and accelerates in the left lane; mode 5: car i decelerates nominally because of preceding car j; mode 6: car i breaks to avoid getting too close to car j.
The model results are shown in Fig. 6 for a simulation run of 6 minutes for the nominal condition ($d = 0.01/s$, $d_\ell = 0.005/s$, $T_u = 10\ s$ and $X_d = 20\ m$, in addition to 'reasonable' values for safety margins and decelerating). Because of the low traffic density in the left lane 12 overtakings were obtained (mode 3 = mode 4 = 12). The NN results will be discussed in the following. To simplify the picture, modes 1, 5 and 6 were pooled in one 'waiting' mode 1. Also mode 3 and mode 4 were pooled in one 'overtaking' mode 3. The result is shown in Fig. 7.

Next the NN was trained to learn the overtaking task. For this purpose the NN utilized a training data set of 360 s. The inputs were the distances to the other cars (X_j, X_k and X_ℓ) and the driving speed of car i (u_i). The outputs were the driving modes (6 or 3). About 7500 iterations were made to obtain the NN results.
These results are shown in Fig. 6 for the six modes and in Fig. 7 for the simplified case of three

modes. Fig. 6 shows that the NN has a difficulty to mimic precisely mode 1, 5 and 6 and to make a distinction between them. This makes sense because these modes differ in terms of slight differences in distance and speed to the preceding car. This was the reason to pool these modes into one waiting mode 1 to see if the NN could learn the distinction between this mode and the other two modes (mode 2: accelerating; mode 3: overtaking). The result in Fig. 7 shows that the waiting mode 1 exhibits quite a scatter but is generally clearly distinct from mode 2 and very clearly distinct from mode 3.

Especially the last result shows that the NN is able to learn the overtaking mode. The overtaking maneuver is a pre-programmed maneuver determined (e.g.) by corresponding initial and final time. As far as the initial time is concerned the NN is lagging somewhat, but the final time is duplicated precisely. This is important because it determines the end of the overtaking maneuver in the left lane before running into oncoming traffic.
After training the NN on the training data a test-set of data was used to check how well the NN has learned the task. The result is given in Fig. 8, showing the same good result as before, especially, again, with respect to the overtaking maneuver.

In order to get a better understanding of the NN problem to mimic the model results, the NN was trained (with 4000 iterations) based on the same inputs as before but as output the velocity u_i at the next time step ($u_i(k+1)$). The result shown in Fig. 9 indicates that the NN duplicates the speed rather closely, but the detailed low speed history (corresponding with braking and waiting (defined for a speed below 20.8 m/s), i.e., modes 6 and 1, respectively) is clearly deviating from the model at least for the number of iterations involved. This 'explaines' the foregoing difficulty to distinguish between modes 1, 5 and 6.

6 Concluding remarks

In this paper two modeling approaches are compared and applied to car driving. The first one is a system theoretic approach; the second one is based on neural networks. Both lane keeping and overtaking were modeled.

As a first step both tasks were modeled in terms of the system model of chapter 2. These model results were used to generate input-output data for the neural network, showing that the neural network was able to learn both the lane keeping and overtaking task.
For the lane keeping task an extended Kalman filter approach was used to learn unknown model parameters. The results are reported in Ref. 2. The next step is to do the same for the overtaking task.

The following step is to simulate the car driving tasks in an experiment with real human operators, with key variable the level of experience. The objective is to assess the relative merit of both approaches to describe human learning behavior in car driving specifically and in operating dynamic systems in general.

References

[1] Wewerinke, P.H., "Modeling car driving and road traffic", *Proc. IFAC MMS '95*, Cambridge, U.S.A., 1995.

[2] Hilberink, H., "A comparison of system theoretic models and neural networks applied to human learning involved in car driving", *Thesis (in Dutch),*, University of Twente, 1994.

[3] Simpson, P.K., "Artificial neural systems", Pergamon Press, 1990.

[4] Luenen, W.T.C. van, "Neural networks for control on knowledge representation and learning", *Ph.D. thesis*, University of Twente, 1993.

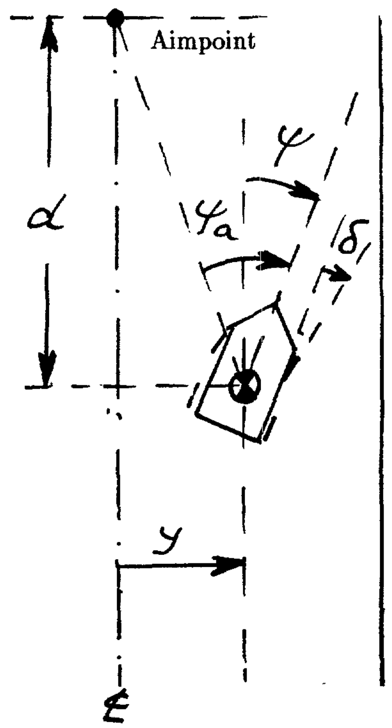


Fig. 1. Lane keeping task situation

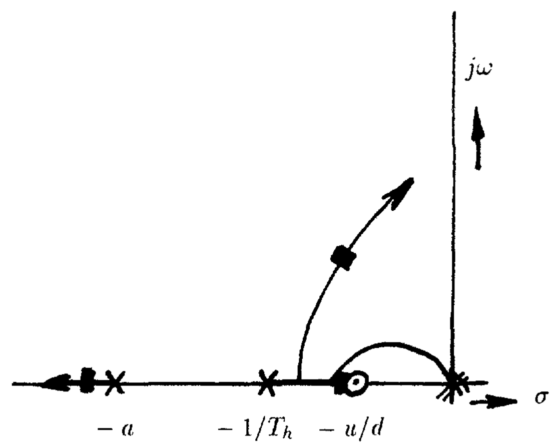


Fig. 3. Root locus lane keeping task

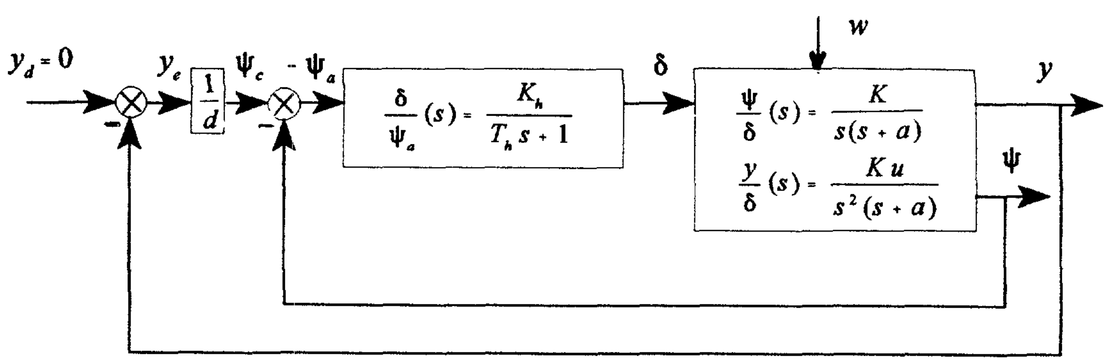


Fig. 2. Block diagram servo-system model of the lane keeping task

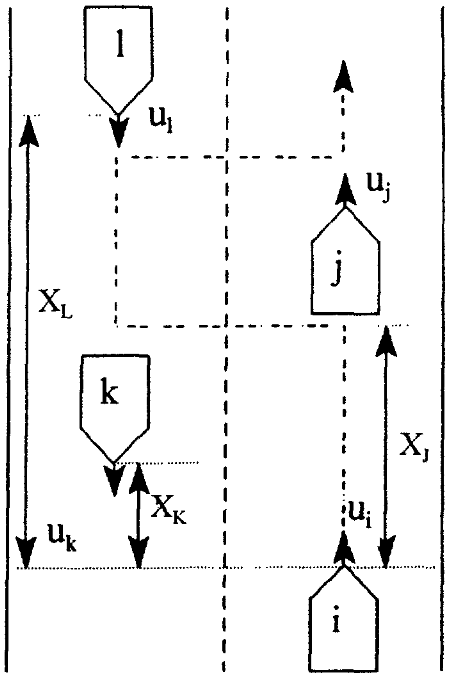


Fig. 4. Overtaking task situation

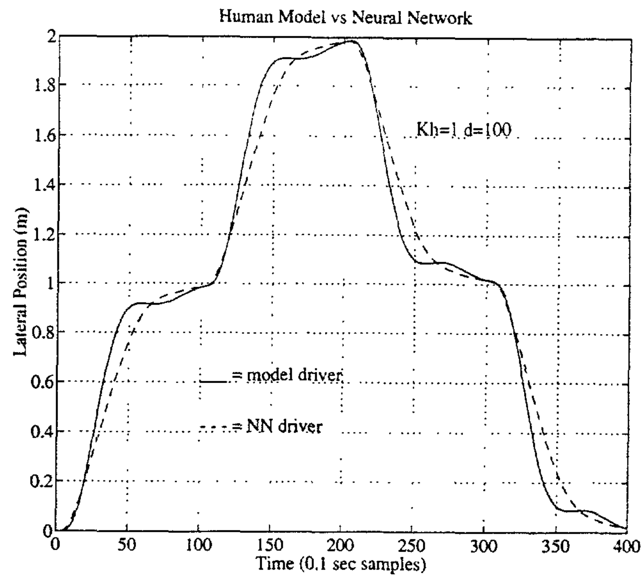


Fig. 5. Model results versus neural network results for the lane changing task

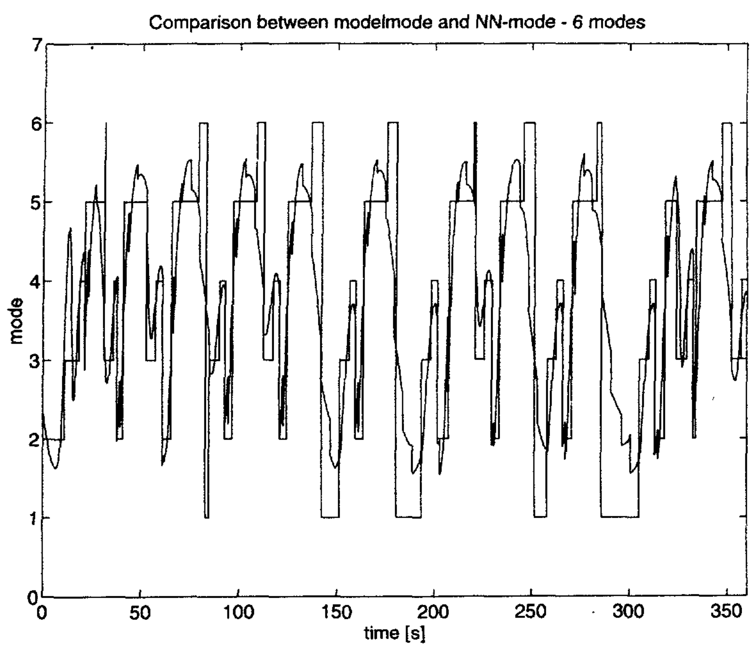


Fig. 6. Simulation results of the overtaking task in terms of 6 driving modes

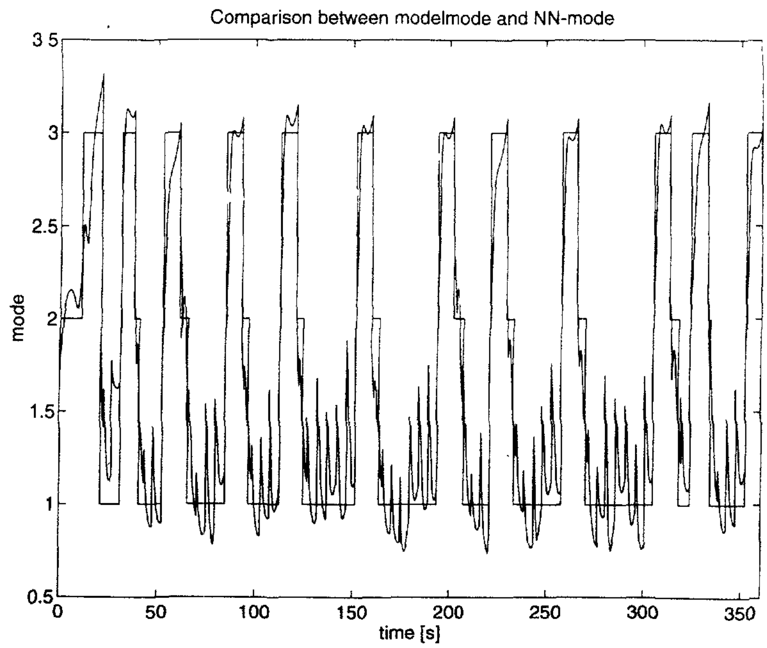


Fig. 7. Simulation results of the overtaking task in terms of 3 driving modes

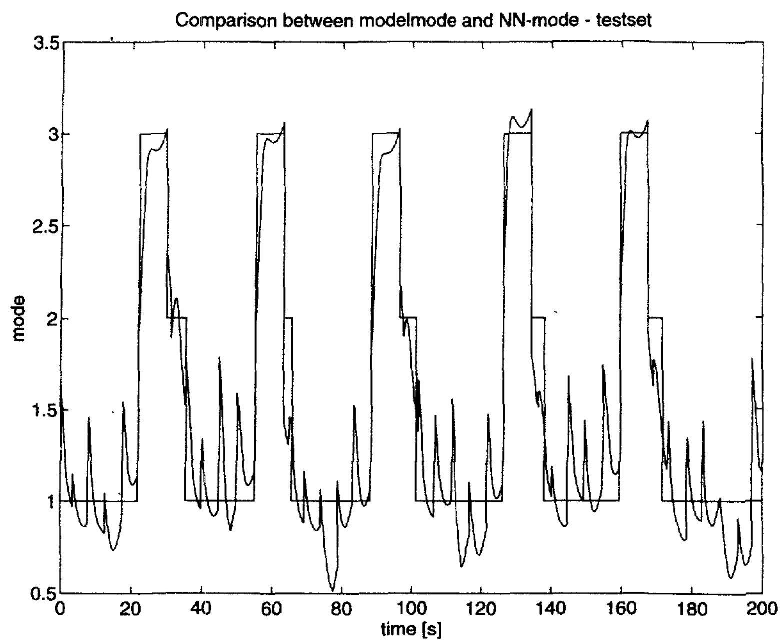


Fig. 8. Testing the neural network results for the overtaking task

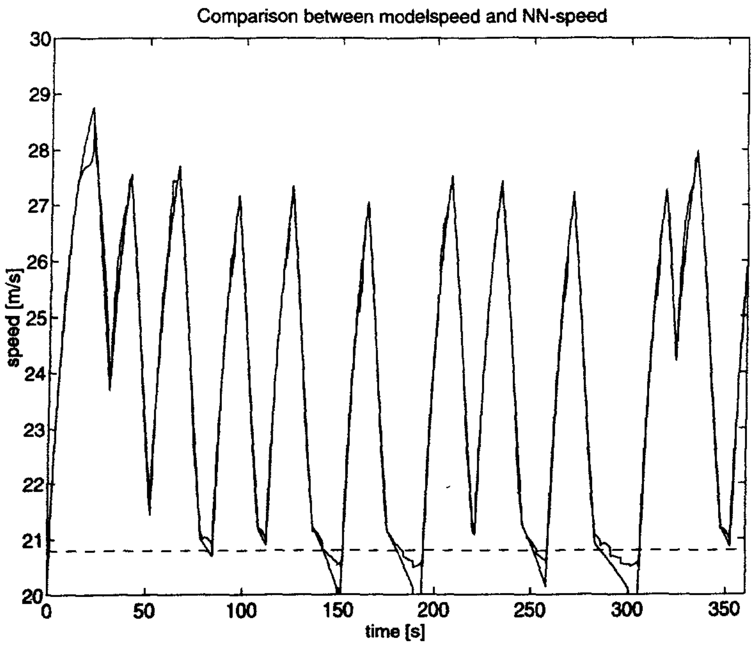


Fig. 9. Overtaking task results in terms of the driving speed

Safety margins in car-following

Monique van der Hulst
University of Groningen
Traffic Research Centre
P.O. Box 69
9750 AB HAREN
The Netherlands

Introduction

In order to assist drivers in coping with the high demands of driving on today's busy roads, intelligent electronic co-pilots are being developed. Intelligent co-pilots monitor vehicle control performance continuously, and can advise and warn the driver or correct his or her actions autonomously. Little is known about the actual behavioural effects of these systems. Therefore, methods and models have to be developed, which allow for the study and the description of these effects. At the Traffic Research Centre of the University of Groningen, a research project has started on the behavioural effects of intelligent co-pilots that support lateral and longitudinal control.

Present vehicle control models usually describe steering rather than speed control. Lateral vehicle control can be described as a serial alternation between passive (open loop) and active (closed loop) strategies (Godthelp et al., 1984). During open-loop control, steering errors are neglected, and attention can be directed at other aspects of driving. Uncertainty about the position of the vehicle on the road determines when the driver has to direct attention to the steering task again, and decide whether it is necessary to switch to a closed-loop steering strategy. Thus, lateral vehicle control may be described as a tracking task in which uncertainty thresholds determine switches of attention. In this study, it is assumed that uncertainty has an important role to play in longitudinal control as well.

Uncertainty thresholds in steering have been studied in occlusion experiments. In an experiment described by Godthelp et al. (1984), subjects had to steer a car on a straight road while visual feedback was occluded. Subjects could control occlusion duration and were instructed to look whenever they felt it was necessary. It was expected that uncertainty about the lateral position on the road would determine the accepted duration of occlusion. Results indicated that subjects' self-chosen occlusion times were related to the time-to-line-crossing (TLC). This concept represents the time necessary for the vehicle to reach either edge of the lane. In another experiment (Godthelp, 1988) the limits of path error-neglecting were studied. Subjects had to initiate a steering correction at the last moment they thought it was still possible to correct the heading error comfortably. The time-to-line-crossing at the moment the driver decided to switch to closed-loop control (minimal TLC) was found to be constant. Thus, there are two thresholds involved in steering control. The uncertainty threshold plays a role in the absence of visual feedback, and determines switches of attention to the steering task. The correction threshold plays a role in the presence of visual feedback and determines the onset of compensatory steering actions.

There is some evidence that analogous criteria occur in longitudinal tracking. Van der Horst (1991) has studied the correction thresholds in distance-keeping, analogous to the steering studies. He measured time-to-collision (TTC), a measure that incorporates speed differences and indicates urgency of conflict situations. Subjects were approaching a stationary object and were instructed to start braking at the last moment they thought they could come to a complete stop without hitting the object. TTC at the onset of braking increased with speed, but not enough to keep a constant deceleration at all speeds: at higher speeds deceleration was higher. The minimum TTC reached was independent of speed. Stroboscopic occlusion was found to deteriorate braking performance. It was concluded that the direct perception of TTC from the optic flow field plays an important role in the decision to start braking and in the control of braking. Thus, the correction threshold in longitudinal control seems to be related to TTC. The present paper describes the preliminary results of an experiment that was carried out to analyse the effect of uncertainty on safety margins in longitudinal vehicle control. The time available for error-correcting and error-neglecting was manipulated by the application of an occlusion technique.

Method

Subjects

Twelve subjects were recruited from the TRC subject pool. The age of the subjects was between 25 and 45 years. All participants had their driving licence for more than 3 years and drove at least 5,000 kilometres annually.

Apparatus

The experiment was carried out in the TRC driving simulator. This is a fixed base driving simulator consisting of an adapted car with original controls. The car is dynamically interfaced to a virtual environment, which is projected on a large panoramic screen. All driving variables can be sampled with a desired frequency, usually 10 Hz.

Procedure

After arrival at the institute, subjects filled out a six-item questionnaire about age, driving experience, etc. Subjects were told that the goal of the experiment was to investigate car-following behaviour. The experimental sessions started with a training session. The goal of the first training session was to familiarise subjects with operating the simulator car. After the training session, normal following performance was measured at three different velocities: 50, 80, and 110 km/h. Subjects were instructed to follow a car as closely as they would do in heavy traffic, but to avoid unsafe distances. After training the occlusion session followed. The experimental session consisted of three training trials and nine experimental trials. The training trials were used to familiarise subjects with occlusion. In the nine experimental trials, velocity (50, 80, and 110 km/h) and occlusion time (1, 2, and 3 sec) were varied. The order of the three velocities was balanced across subjects, and the order of the occlusion conditions was balanced within velocity conditions. The experimental session was followed by another session of car-following without occlusion at all three velocities.

Experimental task

Each trial took about 3 minutes and consisted of an initial part, which was necessary to reach the trial velocity, and a 2.5-minute measurement part, in which occlusion was applied. At the start of the trial, the subject's car was on a straight road behind a simulated car which accelerated from 0 km/h to the trial speed, being 50, 80, or 110 km/h. The subject was instructed to follow this car and to reproduce its speed variations. When the trial velocity was reached, the occlusion started. Occlusion meant hat the lead car was invisible, but road markings could still be seen, in order to provide subjects with sufficient information to steer normally. The time during which the car was visible was always 1 second, the time during which the car was invisible was 1, 2, or 3 seconds, depending on the trial. The speed of the lead car varied slowly within margins of 10% around the trial speed. Sessions without occlusion had the same structure.

Hypotheses

It was assumed that occlusion would increase uncertainty about the position and acceleration of the lead car. The driver was expected to anticipate on this higher level of uncertainty by choosing a larger safety margin. Speed changes of the lead car would on average be perceived later, which would lead to a delay of the subject's response. When occlusion duration is too long compared to view-time duration, the driver's strategy might even change from anticipation to simple stimulus-driven reaction. Therefore, car-following performance was expected to deteriorate with increasing occlusion duration. Minimal time-to-collision might also become shorter as a consequence of delayed response.

Results

Subjects

Twelve subjects, six male and six female, participated in the study. One female subject suffered from simulator sickness, and could not finish the experimental sessions. Therefore, analyses were performed the data set of the remaining eleven subjects. Mean age of the participant was 30 years. Subjects had their driving licence for ten years on average, and drove about 16,000 kilometres annually. Total driving experience was about 215,000 kilometres driven.

Anticipation

Figure 1 presents time-headway (THW) for different speeds and occlusion durations, figure 2 presents minimal time-headway (THW min). THW and THW min were used as measures of anticipation. Multivariate analysis revealed no significant effect of trial velocity (Wilks' lambda = .38, NS). Multivariate tests of the expected influence of occlusion on THW and THW min revealed a significant effect (Wilks' lambda = .07, p = .009). Univariate tests show that this effect can be explained by a linear trend in the effect of occlusion on THW ($F(1,10) = 8.73$, $p = .014$). Thus, subjects chose a longer time-headway when occlusion duration increased. Univariate F-tests showed an interaction between the effects of trial velocity and occlusion on THW min ($F(1,10) = 9.65$, $p = .011$). When trial velocity is 110 km/h, THW min decreases when occlusion duration increases, when trial velocity is 50 km/h, THW min increases.

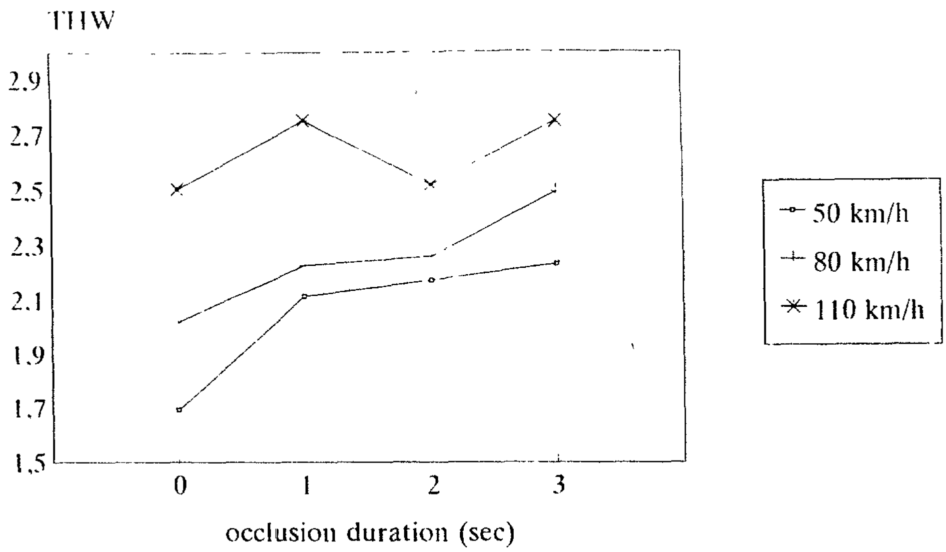


Figure 1. Average time-headway

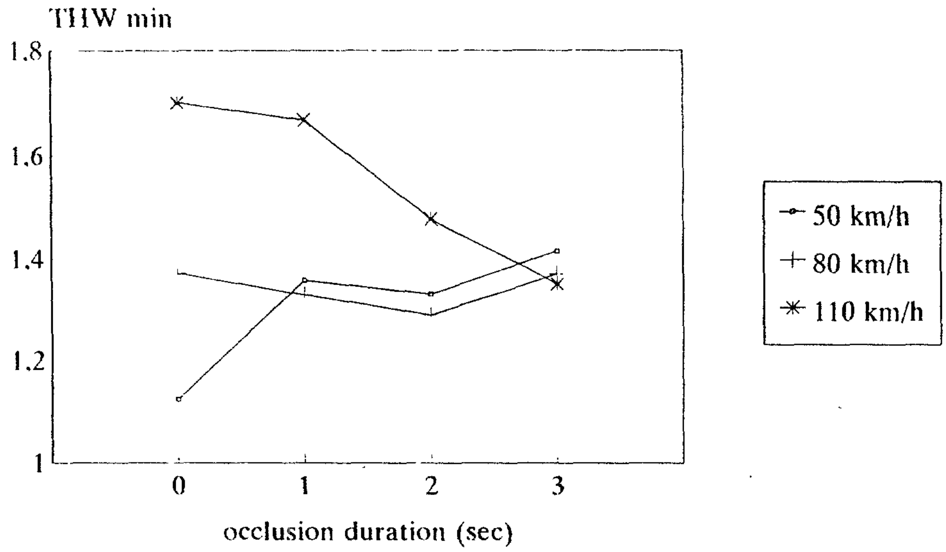


Figure 2. Minimal time-headway

Car-following performance

Figure 3 shows variations of time-headway (SD THW). Trial velocity affected SD THW (Wilks' lambda = .29, p = .004). Univariate F-tests revealed that variation of THW increased linearly when trial velocity increased (F(1,10) = 20.32, p = .001). There was a significant effect of occlusion (Wilks' lambda = .32, p = .022), which could also be described by a linear trend (F(1,10) = 20.92, p = .001). Thus, occlusion resulted in deteriorated car-following performance. There was no interaction between speed and occlusion (Wilks' lambda = .31, NS).

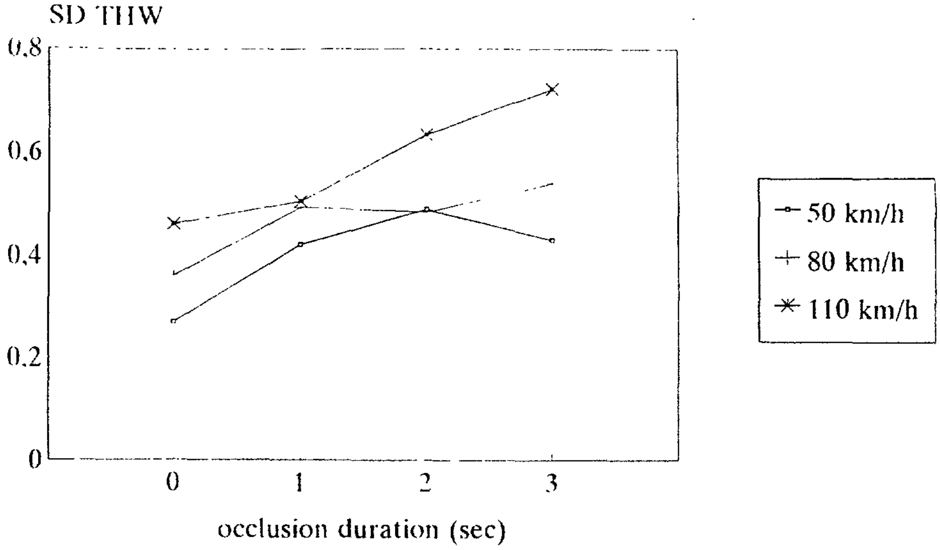


Figure 3. Standard deviation time-headway

Delayed response to speed changes of the lead car may be reflected in minimal time-to-collision (TTC min). Time-to-collisions shorter than four seconds are usually regarded to be critical. Figure 4 presents TTC min during all trials. A relation was found between speed and TTC min (Wilks' lambda = .43, p = .022). Univariate tests showed that this was due to a linear relation; higher trial velocity was associated with a longer TTC min (F(1,10) = 13.23, p= = .005). Occlusion had no clear association with TTC min (Wilks' lambda = .75, NS). There was a significant interaction between speed and occlusion effects (Wilks' lambda = .089, p = .016). This multivariate interaction could be explained as a difference between 50 km/h-trials and 110 km/h trials. In 50 km/h-trials, TTC min decreases when occlusion duration increases, whereas in 110 km/h-trials, TTC min increases when occlusion duration increases from 1 to 3 seconds.

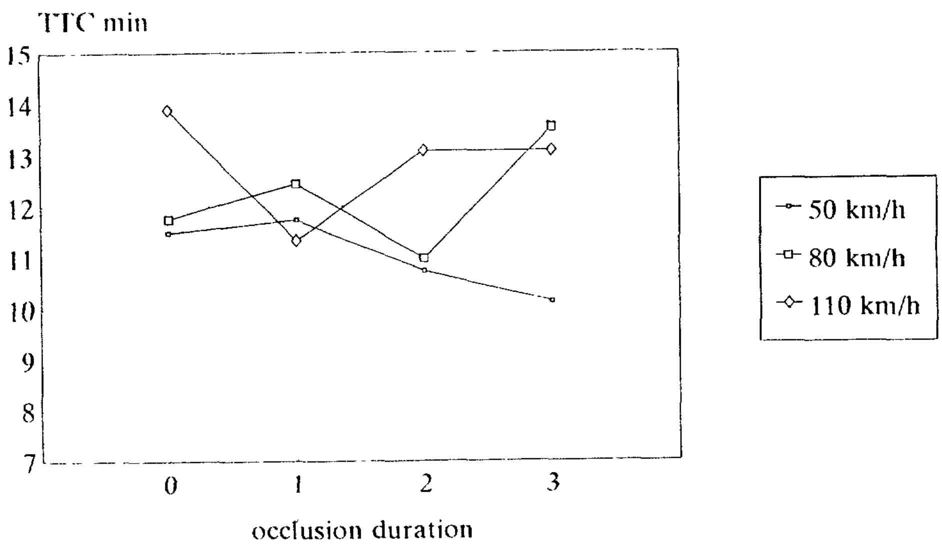


Figure 4. Minimal time-to-collision

Discussion

The main research question in this study was how safety margins in car-following are influenced by uncertainty. From the relation between occlusion and time-headway it can be concluded that subjects anticipate on the uncertainty that is caused by occlusion by adjusting the safety-margin. However, variations in time-headway increased when occlusion duration increased. Thus, subjects were not able to maintain a constant safety margin. This indicates that car-following performance deteriorated. Coherence between the speed patterns of the lead car and the simulator car should be calculated in a further analysis, in order to confirm this degraded performance. The adjustment of safety-margins prevented minimal time-to-collision from getting critical. Although car-following performance was worst in the 110 km/h-conditions, time-to-collision was highest in these conditions. An explanation might be that safety-margins were very large in these conditions, which prevented speed differences from getting conflicts. In this study, speed variations of the lead car were proportional to the trial velocity. Acceleration and deceleration of the lead car were equally fast in all conditions. The dynamics of the simulator car may have complicated accurate car-following at high velocities. This may have influenced the results, especially concerning interactions between the effects of trial speed and occlusion. A suggestion for further studies would be to apply fixed speed variations of five or ten km/h. Furthermore, duration of the vision period should be varied, in order to determine how much time subjects need to perceive velocity changes accurately.

References

Godthelp, H., Milgram, P. & Blaauw, G.J. (1984). The development of a time-related measure to describe driving strategy. Human Factors, 26, 257-268.

Godthelp, H. (1988). The limits of path error-neglecting in straight lane driving. Ergonomics, 31, 609-619.

Van der Horst. R. (1991). Time-to-collision as a cue for decision-making in braking. In: A.G. Gale et al. (Eds.). Vision in Vehicles III. Amsterdam: Elsevier Science Publishers B.V.

SESSION 9

Human-Machine Interfaces

Chairman: P.A. Wieringa, The Netherlands

Software Technologies for Designers of Human-Machine Interfaces

Elena A. Averbukh

Laboratory for Man-Machine Systems, IMAT-MMS, GhK, University of Kassel
Mönchebergstr.7, D-34109 Kassel, Germany
Tel.: +49 561 804 2754 Fax.: +49 561 804 3542
e-mail: elena@imat.maschinenbau.uni-Kassel.de

Abstract. The increasing complexity of automation along with high requirements to its safety and overall quality, as well as competition and frequent deversification of the market demand software technologies which support effective design and redesign of these systems and, particularly, of their human-machine interfaces. This paper specifies methodological guidelines, as well, as practical design solutions and software implementation for adaptive intefaces in control applications. Practical examples are discussed.

Keywords: software technologies, human-machine interfaces, knowledge-based systems, design of interfaces, user model

1. INTRODUCTION

Along with increasing demands for safety, quality, ecological friendliness and effectiveness of human-machine systems (HMS), the problem of their usability in terms of users acceptance and reusability attracts more and more attention of research and industrial communities, especially in the field of industrial control (Johannsen, 1995).

Efficient design of adaptive human-machine interfaces (HMI) which can be easily tailoured to the needs of different user categories and/or individual users in specific task situations becomes nowadays a fashionable field of research and development activities (Rasmussen, 1986, Kawai, 1992). Such interfaces can be regarded as distributed knowledge-based software systems which comprise knowledge about the end-users of the system, i.e., user models, knowledge about technical system, i.e., application or technical system (TES) model and also different kind of interaction models, i.e., dialogue-, task-, presentation-models etc. High dimensionality and also uncertainty of these „heterogeneous" knowledge-structures, as well as required cross-disciplinary design solutions make HMI design process even more complicated than for traditional knowledge-based systems. Systematic user-centred approaches to knowledge-based interface design and software toolkits which effectively support both usability and reusability of HMI, as well as creativity and overall job satisfaction of the designers as end-users of these toolkits become of high importance.

The paper is focused on one approach to the design and practical implementation of such software technology, so-called AMICA-technology. The development as well as trial applications of this technology are carried out within the on-going BRITE/EURAM Project „AMICA-Advanced Man-Machine Interfaces for Industrial Control Applications" by the team of University of Kassel together with four industrial partners[1].

In the next section the conceptual and logical design of adaptive interfaces are discussed. Task-orientation of related knowledge is specified in the third section. Practical examples are given in the fourth section.

2. LOGICAL MODEL OF ADAPTIVE INTERFACES

Modern concurrent engineering paradigms of HMI life-cycle differ significantly from traditional so-called „waterfall" and also „spiral" models (Boehm, 1988) towards „star" models (Hix and Harston, 1993) with multiple local loops of re-design and re-engineering with the help of rapid prototyping and usability testing techniques(Schneiderman, 1992).

The whole life-cycle of knowledge-based adaptive interfaces is accompanied by the task- and knowledge-acquision and elicitation procedures (Averbukh, 1994a). These procedures are extremely time- and labour consuming and highly demand specific support technologies for their possible intensification.

Conceptual and logical design of HMI are one of the first design phases requested by ESA Standards (1991) and are rather explicitely reviewed in the literature (see, e.g., Johannsen, 1995). Usually logical models provide the decomposition of main functionalities without specifying how to realise them, as conceptual design focuses on the implementation strategies.

Conceptual design of adaptive HMI presented in this paper is one possible extension of the so-called UIMS (User Interface Management System) design where dialogue and presentation are clearly separated from each other concept (see, e.g., Johannsen,1994). It refers to user centered approaches and realised through embedding user- and application models into HMI system as one of the most significant functionalities for managing the whole interaction process.

The explicite logical model of adaptive HMI, as well as its main managing functionalities and knowledge sources are presented (see Fabiano et al, 1994).

By analogy with theory of adaptive and supervisory control (see, e.g., Sheridan, 1992), the main activities in adaptive HMI organisation and operation are qualified in relation to monitoring, supervising, quality assuarance (control) and learning, e.g., via explanation etc. Thus Dialogue is considered as a main managing media which

- provides monitoring and „supervisory control" of the whole interaction process between human and machine. Monitoring is executed via tracking interaction context and -content. The controlling of interaction process is realised, e.g., by devising appropriate schemes and dynamical plans of the dialogue which fit to the current interaction context and to the users goal and task, and provides „in-process" verification and validation of embedded models and overall usability of the interface system.

Presentation media deals with handling and generating inputs and outputs to/from the user and technical system and, hense demands comparatively less reasoning. Within such concept it is managed by the dialogue media, particularly by matching actual user-, TES- and interaction models with expected ones and selecting appropriate dialogue plans and related presentation modes (see, e.g., Averbukh et al, 1994). Alternative paradigms which are based, e.g., on flow modelling (Lind, 1981), means-ends hierarchies and ecological interface design concepts (Kim, 1993) accumulate more reasoning facilities and are concurrently designed within the AMICA-technology (see, R.van Paassen, 1995).

Specific communication activities, such as explanation, justification etc. comprise both dialogue and presentation media in a very specific, application oriented forms and thus can be designed in both generic and domain-dependent ways. Dynamical updating of User-Model and subordinate Interaction- and TES-Models which provides a basis for effective managing and adaptation of interaction process to the concrete user needs is discussed.

3. TASK ORIENTATION AND DISTRIBUTED ARCHITECTURAL DESIGN

3.1. Task- and Situation- analysis

Task orientation is widely recognised as an efficient way to provide HMI systems' flexibility and easy adaptation to specific users' needs in typical task situations. This approach is also quite effective in increasing systems' software extendability and reusability (Lind, 1981, Mitchel and Miller, 1987, Sundstrom, 1993). It becomes possible particularly due to the generalisation of typical task situations and near optimal strategies for selection of the appropriate interaction „services", as well as by the self-consistent task-oriented HMI design (Averbukh, 1994b, Kenji Kaijiri, 1994).

Efficient execution of such complicated multistage cognitive process, as task-orientation is only possible within systematic approach. The latter should consider in a unified manner the whole domain of tasks and subtasks allocated between different systems' agents, i.e., man, machine or technical system (TES) and man-machine interface and also between TES and HMI modules.

Intensive task-analysis and -evaluation techniques adapted to the purposes of HMI design are discussed. The main sources for intensification are considered as following

- unified task-oriented knowledge representations for the basic models embedded into interface, i.e., user-, TES- and interaction-models, e.g., in the form of so-called „bottle diagrams" i.e., symmetric knowledge hierarchies *„goals-functions-tasks-means-resources-criteria"*, (see, e.g., Averbukh and Johannsen, 1995),
- analysis and „dynamic" consideration of cognitive models of the experts involved in the task analysis,
- situation-analysis and -design, i.e., analysis of application-dependent task-situations or

scenarios and design of „simulated" situations for a spectrum of design and evaluation goals, e.g., for preliminary testing the end-users and their behavioural characteristics, for usability testing of HMI design solutions etc., and

- identification of so-called „critical task situations" which comprise significant attributes of any task situation that affects controlling activities within the HMI, such as dynamical adaptation of user model, updating current interaction knowledge structures, e.g., interaction context, dialogue goals-plans and related „interface services" presented to the user (Averbukh et al, 1994).

3.2. Task-oriented Architectural Design

Concrete architectural design of HMI Toolkit normally compromises between the requirements of the quite ambitious conceptual and logical designs, the established list of priorities and also by actual resources, including available tools for software development. One of possible approaches which reflect the discussed above concepts logical decomposition is schematically shown in Fig.1 (extended after Averbukh, 1994b). It shows necessity and possibility effectively to realise both goal- and data-driven reasoning mechanisms within the proposed symmetrical task-oriented human-computer interaction modelling. Here interaction process comprises dialogue, presentation, explanation, justification and other communication forms. The goals of end-users are normally determined by the goals of Technical System, i.e., safety, quality, ecological goals etc., and also by their internal or so-called egoistic or usability goals, e.g., understandability of the HMI, acceptance etc. The goals of designers of human-machine interfaces are determined by the goals of end-users, as well as by their personal egoistic goals and also by the goals of the design technological processes, e.g., quality, effectiveness of design etc. The functions for achieving these partially intersected goal-domains and appropriate tasks, needs, resources and criteria are specified during task- and knowledge-analyses.

4. SOFTWARE TECHNOLOGIES FOR HMI DESIGN IN CONTROL APPLICATIONS

After explicit market analysis two software development tools, GMS and COGSYS have been selected as a basis for HMI Toolkit architectural design and implementation on the UNIX platform for real-time control applications (Kwaan and Averbukh, 1993). The programming languages C(C++) are also widely used in addition to these development tools. ANSI C compiler allows running of two selected software tools, i.e., GMS and COGSYS on the same UNIX platform. GMS, the Graphical Modeling System, is a tool for graphical design screen of layouts by simple connecting graphical objects. It can be responsible for the realization of every presentation function, particularly by mediating the communication between the user and the COGSYS side of the HMI and for managing all the graphical aspects related to the presentation, like the managing of the GMS graphic models as well.

COGSYS is a blackboard based tool for the development of real-time knowledge-based systems. The main reasoning functions take place around the blackboard as a shared data structure and their effects are communicated to the user through the functionalities implemented with GMS.

The Toolkit architecture is designed as an open system architecture. The basic data driven control approach to the architectural design realised on the COGSYS basis is extended with concurrent architectural design concepts (Averbukh et al, 1994, Hsu et al, 1994). Independent software C and C++ modules which do not use COGSYS blackboard but communicate with other Toolkit modules through the appropriate inputs/outputs are normally developed after rapid prototyping of some concepts with the help of COGSYS, e.g., in order to improve real-time characteristics etc.

The prototype of the discussed Toolkit is developed within the AMICA project[1] and consists of two basic types of software modules, i.e., generic or application domain-independent modules and domain-dependent ones. Two target application domains are used as a „poligon" for rapid prototyping, particularly of domain-dependent software modules with further analysis of the possibility for there generalisation, i.e., power plants- and cement plants-domains. Industrial experts from both application domains, particularly from ENEL and Santa Gilla Power Plant (Italy) and from FLS/A and Aalborg Portland Cement Plant (Denmark) are involved in task analysis and requirements engineering, as well as in rapid prototyping and evaluating of the results during the development of both Toolkit and specific for each domain prototypes. The concrete examples for software implementation are further specified.

5. DISCUSSION

Different design paradigms of adaptive human-computer interfaces are analysed as a trade-off between their complexity und multidimensionality, on the one hand, and limited resources within systems life-cycle, on the other hand. Unified task-oriented approaches to the design of distributed knowledge-based interface systems are discussed. Their practical implementation is examplified for two industrial control applications in cement industry and for power plants during different design stages.

6. REFERENCES

Averbukh, E.A. (1994a). Multidisciplinary Task Knowledge Engineering, *Proceedings IFAC Conference on Integrated Systems Engineering*, Baden-Baden, Germany, pp. 47-53, September 27-29.

Averbukh, E.A. (1994b). Task-Orientation of Human-Machine Interaction in Distributed Systems. *Proceedings of 3rd IEEE International Workshop on Robot and Human Communication,* Nagoya, Japan, July 18-20.

Averbukh , E.A. and G. Johannsen. (1995). Human Performance Models in Control. *IEEE Transactions on Systems, Man and Cybernetics (extended paper from the Intern. Conference IEEE/SMC'93*, Le Touquet, France, October 17-20 , P. 397 - 402), to appear.

Averbukh, E.A., G. Johannsen, F.Mletzko, R. van Paassen, J. Rudewig and V. Stefanuk (1994).Toolkit Architectural Design, Issue 2. *Internal Report IR-03-BRITE/EURAM,* Project 6126, December.

Boehm, B.W (1988). A Spiral Model of Software Development and Enhancement. *IEEE Computer,* Vol.21, N.5, pp. 61-72.

ESA Standards (1991). Guide to the software requirements definition phase. *ESA PSS-05-03 Issue 1*, October.

Fabiano, A.S., C. Lanza and J.Kwaan (1994). Toolkit Architectural Design. *Internal Topical Report D4 -2-BRITE/ EURAM,* Project 6126, April.

Johannsen, G. (1994). Design of Intelligent Human-Machine Interface, *In Proc. of 3rd IEEE International Workshop on Robot and Human Communication (RO/Man'94),* Nagoya, July.

Johannsen, G. (1995). Knowledge-based Design of Human-Machine Interfaces, *Control Eng. Practice,* Vol.3, No.2, pp. 267-273.

Johannsen, G. and E.A. Averbukh (1994).Human Interaction for Process Supervision Based on End-User Knowledge and Participation, *Proceedings IFAC/IFIP /IMACS Symposium on Artificial Intelligence in Real Time Control,* Valencia, Spain, October 3-5.

Hix , D. and H.R.Hartson (1993). *Developing User Interfaces. Ensuring Usability Through Product and Process,* John Wiley and Sons, Canada. 381p.

Hsu, C., L.Gerhardt, D.Spooner and A.Rubenstein (1994). Advanced Integrated Manufacturing Enterprises: Information Technology for the Next Decade, *IEEE Transactions on Systems, Man and Cybernetics*, Vol.24, N.5, pp. 828-837.

Kaijiri, K. (1994). Support of Plan Library Construction, *In Proc. Japan-CIS-Symposium on Knowledge Based Software Engineering,* JCKBSE' 94 (H.Ueno, V.L.Stefanuk Eds.), Pereslavl-Zalesski, p.203-208.

Kawai, K. (1992). Human Interface Technology in power Plant Monitoring and Control System, In K.Hiroi (Ed.), *Theory and Application of Control System Technology*, Denkishoin.

Kim, W. (1993). Knowledge-based framework for an automated user interface presentation design tool, Ph.D. Thesis, Dept. of Electrical Eng. and Comput. Science, George Washington University, Washington, D.C.

Kwaan, J. and E.A. Averbukh (1993). Market Analysis of Tools, *Internal Report IR-03-BRITE/EURAM,* Project 6126, June.

Lind M. (1981). The Use of Flow Models for Automated Plant Diagnosis, in *Human Detection and Diagnosis of System Failures*, J. Rasmussen and W.B.Rouse, eds. London, Plenum Press, p.411-432.

Mitchell, C.M. and R.A. Miller (1987). A Descrete Control Model of Operator Function. A Methodology for Information Display Design, *In IEEE Transactions on Systems, Man and Cybernatics*, 17, p.573-593.

Nielsen, J. (1990). Designing User Interfaces for International Use, *Elsevier Science Publishers,* Amsterdam, The Netherlands.

Rasmussen, J. (1986). *Information Pro-cessing and Human- Machine Interaction.* New York: North-Holland.

R. van Paassen. (1995). New Visualisation Techniques for Industrial Process Control, In *Proc. 6th IFAC/IFIP/IFORS/IEA Symposium on Analysis, Design and Evaluation of Man-Machine Systems,* Cambridge, June 27-29.

Sheridan, T. (1992). *Telerobotics, Automation, and Human Supervisory Control.* Cambridge, MA: MIT Press.

Shneiderman, B. (1992). *Designing the User Interface. Strategies for Effective Human-Computer Interaction.* Addison-Wesley, 573.

Sundstrom, G.A.(1993). Model-Based User Support: Design Principles and an Example, In *Proc. IEEE/SMC'93*, Le Touquet, p.355-360.

[1] The presented work is partly supported by the Commission of the European Union under the BRITE/EURAM II Programme in Project 6126 (AMICA: Advanced Man-Machine Interfaces for Process Control Applications). The consortium partners are: CISE (Italy), ENEL (Italy), FLS Automation A/S (Denmark), University of Kassel (Germany), Marconi Simulation and Training (Scotland).

TRACKING HUMAN-COMPUTER DIALOGUES IN PROCESS CONTROL APPLICATIONS

René van Paassen

Universität Gesamthochschule Kassel, Laboratory for Human-Machine Systems
Mönchebergstraße 7, 34109 Kassel, e-mail repa@imat.maschinenbau.uni-kassel.de

Abstract. Communication acts exchanged between a human operator and a machine may be seen as forming part of dialogues, coherent sets of information exchanged to reach one or more goals in the Man-Machine System. In process control applications, where many click-and-point type interfaces are used, dialogues may easily be interrupted temporarily. This paper describes a dialogue tracking system, that recognises human-computer dialogues and keeps track of their progress. Results from the tracking system may be used for supporting functions in the Man-Machine System.

Keywords. Process control, man/machine interface, computer interface, human/computer dialogue

1 INTRODUCTION

In order to be able to cope with the complexity of many modern process control applications, the requirements posed on the economical operation, and the safety requirements, an increased attention must be devoted to the design of the Man-Machine Interface (MMI).

In an MMI, two main components may be distinguished (Johannsen, 1995). One is the presentation system, which is concerned with the format of the information that is presented to the user. The second is the dialogue system, which controls the flow of information between operator and MMI.

A dialogue system cannot always be seen as one integral component. Often software modules from different generations and different manufacturers have to cooperate. Via parallel channels, these components are then competing for the operator's attention and for room on the terminal screen. When this is the case, a dialogue system cannot provide support for the user in keeping track of the dialogues she or he is maintaining with the various components in the MMI.

These dialogues between the user and the controlled process or the support functions consist of logically coherent sequences of actions and reactions, called interaction acts. Examples of such interaction acts are the reporting of an alarm, control actions by the HO, requests for information, etc.

This paper describes a *dialogue tracking system*, which is able to keep track of the dialogues between various components in the MMI and the human operator. For this function, the tracking system interprets the log of the interaction between the Human Operator (HO) and the MMI. The tracking system is in first instance developed for integration in the software toolkit of the AMICA[1] project (Fabiano *et al.*, 1994; Averbukh *et al.*, 1995). This toolkit contains a system for logging the interaction acts that are exchanged between the user and the MMI, and these logged data are used as the input data for the tracking system.

The construction of this dialogue tracking system is based on a pragmatic approach. The possible dialogues between users and the MMI are described in so-called dialogue schemes. These are prototype dialogues, that describe of which interaction acts a typical dialogue consists. Additional information stored in a dialogue scheme describes in how far actual dialogues may deviate from the scheme. Such deviations may take the form of a allowable differences in order of the interaction acts, whether acts may be skipped, or repeated and at what points it is possible to interrupt the dialogue and continue with another dialogue.

[1]Brite/EuRam project 6126 AMICA, Advanced Man-machine Interfaces In process Control Applications, supported by the European Commission

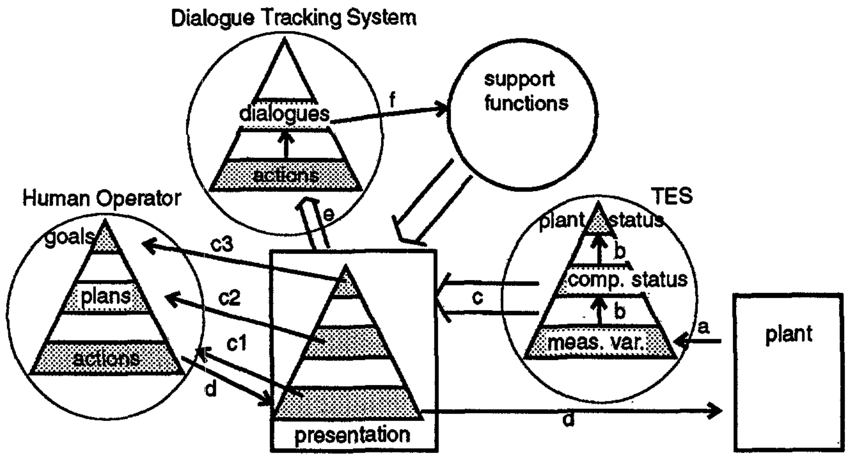


Fig. 1. An information flow view of the dialogue tracking system in relation with other components in the interface. The information flow from the HO to the MMI takes place at the level of actions. At the hand of information flows in the presentation system, the tracking system infers which dialogues the HO is working on. Information at the base of the pyramids is raw, unprocessed data, in the middle and at the top information is processed and more abstract.

The output of the dialogue tracking system is a list of current dialogues that can be presented directly to the HO. At the top of this list the *active* dialogue is presented, this is the dialogue that the HO is currently working on. The other dialogues in the list are dialogues that have been interrupted and are not yet completed. The output of the tracking system can also be used to check the consistency and completeness of the human's actions. When then one or more essential acts in a dialogue are forgotten, or the HO starts a dialogue that is probably not the answer to urgent problems diagnosed by the MMI, a warning may be issued. This would be similar to the concept of pilot assistants (Wittig and Onken, 1992; Le Fort *et al.*, 1993) that are being developed for aeronautical applications.

2 INFORMATION FLOW VIEW OF THE DIALOGUE TRACKING SYSTEM

It is also possible to view the dialogue tracking system at a more abstract level. A schematic representation of this view is given in Figure 1.

Information from the plant is transmitted as raw data, i.e. measured values, to the man-machine interface (path a). By the Technical System (TES), this raw data is processed into integrated data that provides, at a more abstract level, information about the functioning of the plant (path b). The TES comprises the plant with traditional automation, and a supervisory control system, and may provide analysed data and generate alarms. An alternative for the conventional alarm analysis may be a system that performs a means-ends based analysis of the data, which results in indications of how various sub-goals and goals in the plant are reached. The HO may select at which level, or combination of levels, the information is presented on the screen (paths c1, c2, c3).

Information coming from the HO, i.e. his control actions, information requests, operation and usage of support system, is usually passed directly to the appropriate component, either the plant, or a supporting function (path d).

There is thus a disparity between the information coming from the plant, that is available as both low-level (raw) data and as higher-level (processed) data, and the information about the human's intentions and actions, that is only available as low-level data.

The dialogue tracking system attempts to detect a coherence in the user's actions, and to detect which dialogues are maintained by the HO and the MMI. It provides therefore information at a more abstract level about the tasks and goals that the user is working on (path e). This information, together with the information from the TES can be used by supporting functions in the MMI, and fed back to the user.

3 APPLICATIONS OF A DIALOGUE TRACKING SYSTEM

Interpreting the user's actions and determining what dialogue (s)he is working on, does in itself not constitute a contribution to the safety, economy or usability of a Man-Machine System. One of the first applications that comes to mind is the presentation of the current dialogues in a list. This allows the user to keep track of all dialogues that have been started. This is especially useful in situations where the human operator has to manage a large number of possibly parallel tasks. A task analysis for one of the example applications in the AMICA project, a cement mill, indicated that operators would like to have a reminder of pending tasks that have been interrupted when an alarm or other urgent situation had to be handled (Heuer *et al.*, 1993). An example of how such a support would function is given in Table 1.

On the basis of the results from the dialogue tracking system it is also possible to check whether dialogues have been completely handled, and that the user did not forget any essential acts. Such a slip is an error of omission from the user. When the dialogue describes a control operation on the plant, providing an early warning may also prevent the plant from getting into an undesirable state.

Analysis of the plant's state with techniques such as Multilevel Flow Modelling (Lind, 1990) may indicate which goals of the plant are not met, and thus need the user's attention. These goals may be matched against the activities of the operator, to determine whether he or she is devoting attention to the important goals and that the proper strategies are chosen.

In a multi-user system, an overview of the dialogues that the different users are working on can in addition provide a tool to optimise the coordination of the crew.

The following applications were considered of most use and will be the first to be developed:

- Checking that the user does not omit essential acts from dialogues.
- Ensuring that time-critical dialogues are not left unattended for too long.

4 INTERACTION LOGGING IN THE AMICA SYSTEM

To explain the functioning of the dialogue tracking system, it is necessary to give a small description of the way interaction acts are logged in the AMICA toolkit. Interaction acts are for example notifications of alarms, inputs by the user to be sent to the plant, or to supporting functions in the MMS, changes in the display layout etc. These acts, whether they come from the user or from the various components in the MMS, are recorded and logged by the interaction logging facility in the AMICA toolkit (Fabiano *et al.*, 1994). All logged acts have a name, and if applicable information about the topic is also logged. The topic is, in analogy to normal communication, the subject of the conversation. In this case the topic always refers to a part of the controlled plant.

Not all the information in the interaction log is necessary to *recognise* an interaction act. In many cases knowing the name of the act suffices to uniquely identify the act, sometimes the topic is also required.

5 DESCRIPTION OF DIALOGUES; DIALOGUE SCHEMES

At the start of the AMICA toolkit, the dialogue tracking system reads a file with descriptions of typical dialogues. These descriptions, the dialogue schemes, contain general information about the dialogue, such as its name, an associated goal, and the description of the interaction acts in the dialogue, in the form of *act_atoms*, descriptions of an interaction act that contain the necessary information to uniquely distinguish this act from the other possible acts. A section of the file with dialogue schemes is given in Figure 2.

Additionally, the dialogue scheme contains information that describes how the actual dialogues may differ from the dialogue described in the dialogue scheme. The following specifiers may be given to the act_atoms:

INTERRUPTIBLE It is possible that the dialogue is interrupted after this act.

SKIPPABLE This act may be omitted.

REPEATABLE To indicate that an act may be repeated.

Table 1 A hypothetical scenario for a cement mill, and the reactions of the dialogue tracking system. The user has a synoptic presentation of the plant in the background, and pop-up windows for the electronic diary, for taking a sample of the milled product and for inspecting the dialogue stack.

Action of user or plant	Dialogue Tracking System
The user opens the sample window	This action may belong to the dialogues "take sample" and "inspect samples". Both dialogues are started on a tentative basis, but not yet shown on the stack.
The user issues the command to take a sample.	"take sample" is made to be the active dialogue.
The user closes the sample window	The stack remains unchanged.
The user opens the diary window.	"diary entry" is made to be the active dialogue, "take sample" moves to place two on the stack.
The user enters some text.	The stack remains unchanged.
The TES reports an alert for the roller press (a press that crushes the cement clinker before it is fed to the mill proper).	The dialogue tracking system starts up tracking of the dialogue "roller press adjustment", but since the user has not yet worked on it, it is not placed on the stack.
The user closes the diary window.	The stack remains unchanged.
The user adjusts the roller press speed.	"roller press adjustment" becomes the active dialogue.
The TES reports that the roller press values are OK again.	"roller press adjustment" is considered complete and marked for ejection.
The user opens the diary window again.	"diary entry" becomes the active dialogue again, "roller press adjustment" is removed from the stack.
The sample results come in.	The tracking system takes notice of this act, but the stack remains unchanged, because no participation of the user is involved.
The user submits the diary entry and closes the diary window.	"diary entry" is marked for ejection.
The user checks the dialogue stack and sees that "take sample" is still there.	There is no dialogue scheme defined for "inspection of the dialogue stack", so the dialogue tracking system does not react.
The user opens the sample window.	"take sample" becomes the active dialogue, "diary entry" is removed.

Furthermore it is possible to indicate how the ordering of the interaction acts in a real dialogue may differ from the order in which they appear in the dialogue scheme. For this the acts in the scheme are grouped into blocks, see Figure 3. Block 1 contains the acts which normally indicate the start of the dialogue, the last block usually contains the acts that indicate the end of the dialogue. The strictness with which the user will follow the block order can be specified (`STRICT_ORDER`). A new block starts with an interaction act that has a `NEXT_ACTS_BLOCK` specifier, the number following this specifier indicates the order strictness of the acts within the block.

Block 0 is a special case, it is reserved for "wild-card" acts, acts that may occur anywhere in the dialogue. The example dialogue in Figure 2 uses two such wild-card acts.

The division of a dialogue scheme into blocks gives sufficient freedom to specify the dialogue that results when the user has to fill out the questions in a number of consecutive windows, thus with `STRICT_ORDER 1.0`, which specifies that the order in which the windows are presented is fixed, and `NEXT_ACTS_BLOCK 0.0`, which specifies that the order in which the questions in a window are answered is determined by the user.

Using the order information, the dialogues tracking system can resolve conflicts that arise when the same interaction act occurs at more than one place in a dialogue, or when it occurs in more than one dialogue.

The example scheme given in figure 2 describes the making of an entry in an electronic diary (Fabiano *et al.*, 1993). Such a dialogue starts with opening of a diary window on the computer's screen, then the user fills out a number of items, of which some are optional (`SKIPPABLE 1.0`). The order in which

```
DIALOGUE "diary entry"
{
  GOAL "administrative work"
  GOAL_IMPORTANCE 0.3
  NORMAL_COMPLETION_TIME  60.5
  STRICT_ORDER 0.1
  MUST_BE_COMPLETED 1.0
  EJECTION complete_last_block
  # wild-card acts, may occur anywhere in the diary dialogue
  # these acts occur when a user closes the diary window
  # while the process of making an entry is not complete
  ACT_ATOM WaOpen
  {
    WA_ID diary
    REPEATABLE 1.0
  }
  ACT_ATOM WaClose
  {
    WA_ID diary
  }
  # first block, open
  ACT_ATOM WaOpen
  {
    WA_ID diary
    NEXT_ACTS_BLOCK 1.0
    REPEATABLE 0.0
  }
  # second block, all possible acts, unordered,
  # may be repeated, may be skipped
  ACT_ATOM DiaryEnterSystem
  {
    NEXT_ACTS_BLOCK 0.0
    SKIPPABLE 1.0
    REPEATABLE 0.2
  }
  ACT_ATOM DiaryEnterComponent
  ACT_ATOM DiaryEnterVariable
  ACT_ATOM DiaryEnterAction
  ACT_ATOM DiaryEnterVarious
  ACT_ATOM DiaryEnterTime
  # sentinel block. submitting an entry and closing the
  # diary window completes the dialogue
  ACT_ATOM DiarySubmitEntry
  {
    NEXT_ACTS_BLOCK 1.0
    SKIPPABLE 0.0
    REPEATABLE 1.0
  }
  ACT_ATOM WaClose {WA_ID diary}
}
```

Fig. 2. A section of the datafile with the dialogue schemes, description of a diary entry

the acts are performed is only determined by the user, so the strictness of ordering in the second block is set to 0. When an entry has been prepared, the user can press a button to submit this entry, and then he can close the diary.

The tracking system considers the diary work finished when at least one entry has been submitted and the window is closed. The way to determine the end of the dialogue is specified with `EJECTION_METHOD complete_last_block`. In this case the last block contains the act SubmitDiaryEntry, which is specified as not skippable, and WaClose, which signifies the closing of the working area (=window) of the diary. The ordering in this block is strict. When the user temporarily interrupts the work on the diary, the tracking system will not assume that the WaClose act in the last block has occurred, because the SubmitDiaryEntry act is "blocking the way". Instead it assumes that the user has temporarily closed the window, and that the WaClose act in the block with wild-cards is the appropriate act. When SubmitDiaryEntry has occurred, and then the working area is closed, the tracking system assumes that the WaClose act in the last block has occurred, because that fits better than a mere wild-card act.

WaOpen	Block 0, wild-card acts
WaClose	
WaOpen	Block 1, start
DiaryEnterSystem	Block 2, loose order, acts are repeatable and skippable.
DiaryEnterComponent	
DiaryEnterVariable	
DiaryEnterAction	
DiaryEnterVarious	
DiaryEnterTime	
DiarySubmitEntry	Block 3, last block, when both acts occurred, dialogue is finished.
WaClose	

Fig. 3. Division of the acts into blocks. Block 0 is the block with wild-card acts, that may occur anywhere within the dialogue. The dialogue starts with acts from block 1. Block 2 contains the "body" of the dialogue, the last block contains the actions that close off the dialogue. Strictness of order within the blocks, as well as the strictness with which the order of the blocks is followed, can be specified.

6 STAGES IN TRACKING DIALOGUES

6.1 *The dialogue stack*

The tracking dialogue system maintains a stack of current dialogues. The first place in this stack is occupied by the dialogue that is believed to be the active dialogue, i.e. the dialogue that the user is working on. Other places on the stack contain the dialogues that have been interrupted.

The tracking system often will have to deal with more or less ambiguous situations, where it is not clear to which dialogue an incoming interaction act belongs. In order to express this ambiguity in the information about the dialogue stack, a truth value `active` is maintained. When a new interaction act comes in, the `active` property is updated, in accordance with the goodness with which the act fits in the history of acts in the dialogue, and in accordance with the fact whether that act is specific for this one dialogue, or whether it may fit in more dialogues.

6.2 *Look-up of an interaction act*

At the initialisation of the tracking system, when the dialogue schemes are read in, the tracking system indexes the acts in the dialogues according to their name. Each entry in this index array contains possible extra specifications (topic, wa_id, etc.), and references to the places in the dialogue scheme or schemes where the act was specified.

The tracking system interprets an incoming interaction act in three stages. In the first stage all the dialogues and all the places where this act occurs are looked up in the index array, by comparing the act_name of the interaction act, with the act_name specified in the dialogue scheme and, if that information is available, by comparing `topic`, `actor`, `recipient` and `wa_id`. The result of this stage is a list of dialogue schemes were this interaction act may fit, and for each scheme a list of one or more places where the act may occur.

6.3 *Fitting the act in a dialogue scheme*

A "goodness of fit" is calculated for all possible places where the act may fit into a dialogue. When this dialogue has already been started, and thus the previous acts are known, the relation between the new act and the previous acts is considered, and one or more of the following cases may apply:

1. The present act follows the previous the previous act, not only in the real-life dialogue, but also in the description in the dialogue scheme, therefore the fit is optimal.
2. Both acts do not follow each other in the dialogue scheme but are in the same block of acts. The goodness of fit is 'punished' according to the order strictness in the block. If a low value for order strictness was given, goodness stays high, otherwise goodness will be lowered.
3. When an act was repeated, the goodness will be corrected by the value found after the `REPEATABLE` specifier.
4. When the previous and present act are not from the same block, the goodness is corrected with the order strictness for the blocks. It is possible that the correction from Item 2 and this correction are accumulated.
5. Lastly, the evaluation of the goodness of fit considers the possibility that one or more acts were skipped. The goodness is each time corrected for the `SKIPPABLE` property of the possibly skipped acts.

When the dialogue is new, it is checked whether it is feasible that the received interaction act signals the start of the dialogue. In principle only acts in the first block start the dialogue, but it is possible to tune the tracking behaviour so that is also accepts wild-card acts or acts that are specified elsewhere in the dialogue scheme as starting acts.

6.4 *Ascribing acts to a dialogue*

After looking up the acts and checking the fit of these acts at each of the places in the dialogue schemes where they are specified, a decision is made how to update the information on the running dialogues. Three cases are possible:

1. The lookup of the act did not provide any results, i.e. there were no dialogue schemes specified with this act, or the act was evaluated to have a very low goodness of fit. In this case no change is made to the dialogue stack.
2. The search for dialogues resulted in only one dialogue, or the goodness of fit is only good for one dialogue. In that case the act is ascribed to that dialogue.
3. Multiple dialogues proved to have a claim on the act, and for more than one of them a reasonable goodness of fit was calculated. In this case the act is tentatively described to *all* these dialogues. The runtime information for all these dialogues is updated as if they received the act.

A tentative ascription is made undone at the next cycle of the tracking program. When a new act comes in, and it this act is ascribed to one dialogue, then it is checked whether this dialogue has a tentative ascription. This tentative description is then made definitive for this dialogue, and taken away from the others. If the next incoming act is not a definite ascription, but a tentative ascription, the older tentative description will be made definitive for the dialogue that is the most likely candidate to be active.

7 DISCUSSION

The dialogue tracking system as described here is a first, practice oriented solution for interpreting the user's actions. Solutions based on other principles are also imaginable, such as a knowledge-based system with rules describing effect of actions in view of system goals, or a self-learning system that generates its knowledge from actual dialogues and questions an expert when it does not properly recognise the current dialogue. Compared to such systems the proposed solution is relatively simple and hopefully easily maintainable and adaptable.

It is possible to describe schemes in which the acts have a strict or a loose order, to specify acts that may be omitted or repeated, and acts may also occur in more than one dialogue. The possibilities offered for specifying the behaviour of the tracking system should be flexible enough to allow descriptions of commonly occurring dialogue types, such as question and answer and point and click dialogues. The dialogue schemes are entered in a data file with a lexically free format, to allow users of the MMI to make adjustments to the schemes if performance of the tracking system is not satisfactory or when operational procedures change.

The inventory of dialogues produced by the tracking system provides support for the operator when he has to switch between multiple concurrent tasks. Further operator support can be developed with the output of the tracking system as a basis.

The application of a dialogue tracking system will keep the operator "in the loop". The system will not automise tasks or dialogues, and the operator will remain aware of all his actions.

Currently, the software of the dialogue tracking system is being tested with a simulated interaction log. Integration of the software into a demonstrative example is planned, this can then be used to evaluate the tracking system in a more realistic environment, and interaction with users.

8 ACKNOWLEDGEMENTS

This work has been carried out as part of the European Community research training project "Visualisation of Process Means-Ends Hierarchies for Advanced Man-Machine Interfaces" (ERB-BRE2CT943085). The work has been done in co-operation with the partners in the Brite/EuRam II Programme in Project 6126 (AMICA: Advanced Man-Machine Interfaces for Process Control Applications). The consortium partners are: CISE (Italy), ENEL (Italy), FLS Automation A/S (Denmark), University of Kassel (Germany), Marconi Simulation and Training (Scotland).

9 REFERENCES

Averbukh, E.A., G. Johannsen, F. Mletzko, R. van Paassen, J. Rudewig, T. Gavrilova and A. Voinov (1995). Toolkit detailed architectural design issue 2 (software modules group 1). AMICA Brite/EuRam Project 6126 D5-1. University of Kassel, KONSYTEC GmbH.

Fabiano, A.S., C. Lanza and J. Kwaan (1994). Toolkit architectural design. AMICA Brite/EuRam Project 6126 D4-2. CISE and MARCONI Simulation.

Fabiano, A.S., C. Lanza, B. Mulvihill, G. Zardetto, M. Abram and F. Calabretta (1993). Architectural design for application a prototype. AMICA Brite/EuRam Project 6126 D8-A. CISE, ENEL.

Heuer, J., S. Borndorff-Eccarius and Averbukh E.A. (1993). Task analysis for application b. AMICA Brite/EuRam Project 6126 IR1-04. University of Kassel, Laboratory for Human-Machine Systems.

Johannsen, G. (1995). Knowledge based design of human-machine interfaces. *Control Eng. Practice* **3**(2), 267–273.

Le Fort, N., O. Aboukhaled and D. Ramamonjisoa (1993). A co-pilot architecture based on a multi-expert system and a real-time environment. In: *IEEE International Conference on Systems, Man and Cybernetics*. Le Touquet, France.

Lind, M. (1990). Representing goals and functions of complex systems - an introduction to Multilevel Flow Modelling. Technical Report 90-D-381. Institute of Automatic Control Systems, Technical University of Denmark.

Wittig, T. and R. Onken (1992). Knowledge based cockpit assistant for controlled airspace operation. In: *5th IFAC Symposium on Analysis, Design and Evaluation of Man-Machine Systems*. The Hague.

A MAN-MACHINE INTERFACE FOR THE REAL-TIME CONTROL OF ROBOTS AND THE SIMULATION OF SATELLITE TELEOPERATIONS

D. Garton[1], R. Rehm[2], B. Schäfer[2]

[1]Computer Anwendung für Management (CAM) GmbH
Rudolf-Diesel-Str. 7a, 82205 Gilching, Germany
[2]German Aerospace Research Establishment (DLR), German Space Operations Center,
82234 Oberpfaffenhofen, Germany

Abstract

The Servicing Test Facility Pilot (STFP) is a test and simulation facility being developed by the DLR in Oberpfaffenhofen. The facility will be used in preparing for the remote tending or teleoperation of space segment subsystems and, in particular, to investigate the support needs of the teleoperator in order for him to perform his task. This paper will concentrate on one aspect of this work: the man-machine interface provided to the operator for control of the robot manipulator and known as the Teleoperator Interface (TI). The development of the TI is one of the major aims of the STFP project. The TI provides the operator with a comfortable and simple set of tools to control the system.

Key words: man-machine interface, real-time control, robotics, simulation, satellite, teleoperations.

Introduction

The use of automation and robotics (A&R) technology in space systems has gained increased importance during the past decade. The driving forces behind this are clearly not very much different from those of terrestrial applications: cost decrease, routine operations with both high precision and repeatability, human safety and operations in hostile and inaccessible environments. In space applications the expected reduction in the immense operational costs and the less stringent safety requirements compared to manned systems are the main drivers. Moreover, looking at planetary rover systems equipped with one or more small manipulators, continuous monitoring and possibly telemanipulation from Earth would reduce tremendously the operational costs compared to manned systems and could prepare and initialize manned missions if these were still required.

In view of the above, A&R systems will play a key role in future space missions. This applies not only to unmanned missions but also to manned missions in order to relieve the astronaut from routine tasks. Up to now the only flight-proven space robotics system developed in Europe was the Robotics Technology Experiment, ROTEX. which flew on the German Spacelab Mission D2 in April 1993[1]. ROTEX was an important step towards the development and successful demonstration of an operational robotic system with high on-board intelligence given by a multi-sensorial end-effector. The operational modes were: an automatic mode, a telesensor programming mode and a telemanipulation mode via both an astronaut on-board and, much more spectacular and challenging, a human teleoperator from the ground. One of the main objectives was to demonstrate novel telecontrol components such as the 6 degrees-of-freedom (6D) control ball and the integrated co-operation of the complex systems consisting of advanced Man-Machine Interfaces (MMI), stereo 3D computer graphics animation, stereo image generation and the 6D control ball.

The ground-based telerobotic operational mode is certainly the most challenging for future operational systems, not only for platforms in low earth orbit but even more so for those in geostationary orbit. Projects currently under study include the minisatellite programme, MINISAT/MINIMAN, in Spain, the Experimental Servicing Satellite (ESS) in Germany and the Geostationary Servicing Vehicle (GSV) at ESA. These missions will rely on robotics for performing servicing tasks such as maintenance, repair and inspection as well as payload handling, be it either in an automated or in a teleoperational mode.

One of the most significant factors guaranteeing mission success is the development of an effective and powerful Man-Machine Interface which can enhance the human virtual presence in space. This article therefore focuses on the development of an effective human teleoperator environment. The MMI telerobotic developments are part of a simulator, the Servicing Test Facility (STF), at the DLR in Oberpfaffenhofen.

The Simulation Facility

The STF is to be used in preparing for the remote manipulation or teleoperation of space segment subsystems. The reasons for developing such a simulation facility mainly derive from the novel technological challenges inherent in the remote operation of space robots:

- there is a considerably long time delay for uplink/downlink data transmission that makes continuous telemanipulation nearly impossible;
- the space robot and its environment are not directly visible for the human teleoperator on the ground;
- the teleoperator's perception of the scene is restricted by the limited number of camera positions and viewing angles available on-board;
- further deterioration of the view may arise from the necessity of compressing the video data;
- the possibility of transmission errors in the uplink of commands must be considered.

The STF will provide a realistic simulation of satellite-based manipulators that are controlled from the ground. The STF must thus consist of a computer system capable of full simulation of the manipulator, its space environment, ground link communications and a teleoperator workstation. The simulation must cover the dynamics, kinematics and sensorics of the operational system.

The STF was originally intended to prepare for remote tending of Columbus space segment systems. The specification of the requirements and implementation of the STF therefore depended heavily on the requirements of this mission which, as is well known, are far from being finally decided. The solution to this problem was to introduce an initial, "pilot" phase. In this phase a reduced function facility can be developed quickly and cost effectively to investigate technical problems in and gain experience of the full system. The Servicing Test Facility **Pilot** (STFP) will provide a subset of the functions planned for the final STF system and will be used to investigate:

- the support needs of the teleoperator in order for him to perform his task;
- the problems associated with real-time control of a manipulator;
- the use of real-time animated graphics and the necessary support hardware;
- the man-machine interface requirements in all the systems;
- which elements of the final, operational system are important and thus must be simulated;
- the computer power necessary to perform the various system functions.

The heavy investment in software and hardware in the pilot will not be wasted as they can be reused in the final implementation.

The basic approach adopted for the STFP is to guide and control the space manipulator using computer graphics animation. Within this telerobotic philosophy, the space manipulator behaves like a slave, with the computer simulated manipulator as the master. Since the space robot will react some seconds later, depending on the time delay, the simulated robot can used as a predictive model. The quality of the model strongly depends on the system knowledge that has been gained so far, on the availability of efficient modelling tools and of course on the performance and power of the computer systems. These aspects in conjunction with the requirement for simulating in real-time have driven the design of the facility[2]. In addition to using real manipulators in the simulation (hardware-in-the-loop or HIL), software dynamic modelling tools will also be used (software-in-the-loop or SIL). Dynamic modelling is especially useful for large manipulators and Rendez-vous and Docking installations.

The STFP therefore provides a realistic simulation environment (see figure 1) for satellite-based manipulators that are controlled from the ground. The facility consists of

- an industrial robot and a satellite mock-up both capable of serving as the relevant flight segment;
- kinematic and dynamic modelling software for the flight components;
- Teleoperator Interface (TI) workstation including the entire process control environment and the MMI components.

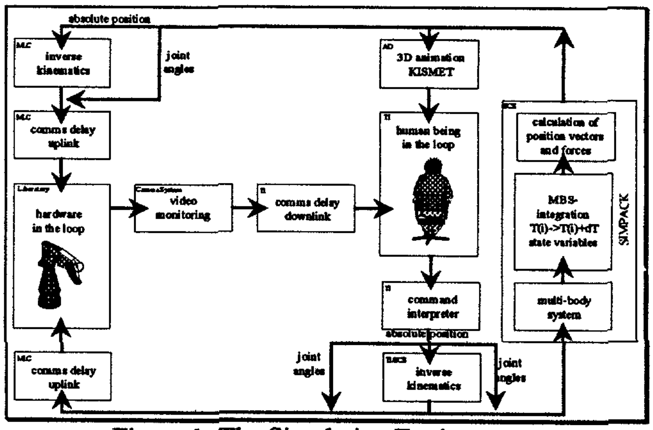


Figure 1. The Simulation Environment

The Teleoperator Interface

The development of user interfaces for teleoperations has the ultimate aim of using human skills and intelligence to support a remote manipulator system having only limited intelligence in the best possible manner.

The basic TI can be considered as consisting of the set of hardware input devices which control the robot or simulation system, a user interface, a 3D kinematic animation display and one or more video monitors.

Requirements on the TI

The TI must provide the means by which:

- the teleoperator can control, continuously if necessary:

- ♦ the position of the end-effector in all 6 degrees of freedom;
  - ♦ the manipulator's joints;
  - ♦ the end-effector functions such as grasping;
  - ♦ each camera's pan, tilt, focus and zoom.
- position, velocity and sensor data are displayed in real-time, preferably in graphical form;
- the whole system is started, configured or reconfigured and closed down;
- the machines in the network can communication with one another;
- experiments can be archived and later retrieved for post-run evaluation or fault finding.

Choice of Control Devices

The requirements section above listed the types of control that are needed for the TI. The most basic decision made at the outset was which, if any, of these would be fulfilled by attaching additional peripherals over and above the now standard devices of keyboard, monitor and mouse. The main advantage of using a screen-based approach is its flexibility. It is far simpler to change the number of buttons provided for a given task than provide new hardware. External devices can, however, provide better control than solutions using the standard devices alone, especially when more than 2 degrees of freedom in the control are necessary. As screens are limited in size, using external devices also avoids screen clutter and the risk of confusing the operator.

Controlling 6 degrees of freedom simultaneously is basically possible using the standard user interface items. It is also possible using a pair of joy-sticks or similar devices. A much better solution, however, is to use a device such as the DLR 6D sensor ball. This responds to 3 translational and 3 rotational impulses and has 8 programmable buttons. The operator can thus easily control the absolute position and orientation of the robot or the position of each joint separately.

The main concern with providing the end-effector control is flexibility. The number of functions required varies greatly with the particular end-effector in use. This situation may change even during operations if, for example, an end-effector tool is exchanged. This implies that the number and interpretation of the control keys must be software controllable. Two possible solutions are

- an external digitizer pad with an overlay sheet to show the key functions;
- the control pad can be displayed on the user interface as a menu and the mouse or keyboard used to select options.

This question has not been completely resolved yet in the STFP. The original designs called for a programmable control pad. Investigations must be made to see if better solutions are currently available. It should be noted here that, whatever solution is adopted, the teleoperator will need careful training in the use of the control devices.

The camera control could also be implemented using an input device such as a joy-stick or tracker ball (since not so many degrees-of-freedom are needed) or via the user interface. Similar arguments to those for the end-effector control apply here. As this kind of control does not need to be as interactive as that for an end-effector it is probably best implemented via a menu on the user interface.

The problem as to whether control options should be offered on the user interface or via programmable keys on input devices such as the 6D ball is a difficult one and may only eventually be solved after operational experience is gained. It is important to note that a flexible development environment is imperative to allow such changes to be made easily. In view of these discussions, all the necessary control features are offered as menus on the user interface screen. It should also be noted here that there is always the third option available of controlling the devices via typed-in commands. This method was used effectively for controlling the robot in the INTA experiment (see below).

Choice of Hardware and Software

In view of the use of animation displays and simulators as well as the user interface itself, several computers will be needed for the system and these must be networked together to provide the necessary close co-operation. The choice of hardware and operating system was dictated by the need for a multi-tasking system with networking capabilities and a sensible user interface. There was originally no realistic alternative to the use of Unix or VAX workstations and the X Window system.

Silicon Graphics (SGI) workstation were selected for the animation display due to their excellent graphics performance and the existence of a 3D animation system specially suited for robotics, KISMET, which runs on SGI machines. For the TI machine itself, the choice of workstation is not so critical and the original specification called for a 486 based PC running SCO Unix. At the moment a second SGI machine is in use. For the simulation machine a more powerful SGI was selected for running the commercially available simulation system, SIMPACK.

In order to provide a user interface easily and be able to modify it readily according to changing needs or new ideas the use of a user interface building tool is essential. These tools allow complex user interfaces to be put together interactively so that the developer can see the results of his work immediately. In addition, many systems allow application code to be generated from which a program can be produced which will display and control the user interface. Following a comparison of several systems, the decision was made to choose the

Transportable Application Environment, TAE+, originally developed by NASA but now available commercially. TAE+ has many advantages over other systems including ease of building and modifying a user interface, implementations for many systems such as VAX, Silicon Graphics, SCO Unix and Solaris and many varieties of display type including dynamic data objects (DDO) which are very useful for the graphical display of information.

Description of the Teleoperator Interface

Figure 2 shows the current system at DLR. The computers in use at the moment are as follows:

- Teleoperator Interface computer (TI). This is a Silicon Graphics Iris 4D/420 VGX with two 40Mhz CPUs, 64Mb memory and running Irix 4.0.5c.
- Animation Display computer (AD). This is a Silicon Graphics Iris 4D/310 VGX with one 33MHz CPU, 16 Mb memory and running Irix 4.0.1.
- Manipulator Laboratory Control computer (MLC). This is a MVE165 based VMEbus computer running VxWorks 5.1. It uses as a host processor a Sun Spark Classic running Solaris 2.3.
- Simulation Computer System computer (SCS). This is a Silicon Graphics "Jurassic Classic" with one 150MHz CPU, 64Mb memory and running Irix 5.2.

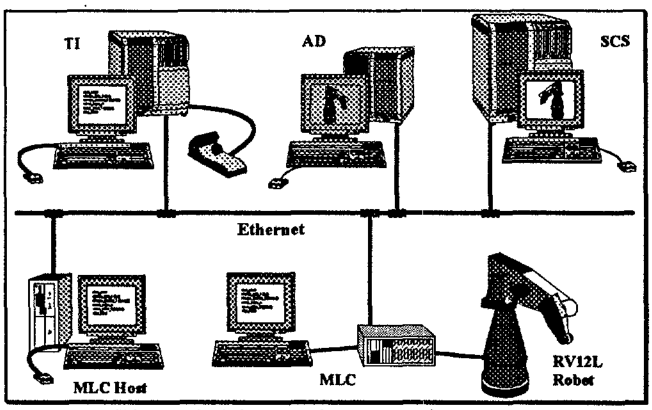


Figure 2. The Hardware Environment.

The manipulator is controlled in real-time by the control software running on the MLC. This receives commands from the operator at the TI who uses the control devices or the menus to control the robot. The commands are also sent to the animation display so that the operator can see the results of his actions in real time. Command entry is via the 6D ball or menu. This can be done either in local mode by controlling each joint individually or in global mode by controlling the position and orientation of the end-effector directly. Apart from being able to view the effects of commands via the AD, the operator also has menus which show the command issued and the actual measured values of the joint angles or positions. These menus use the TAE+ DDO features to great effect to display the data graphically as sliders or rotators. The measured values are continually update from 'telemetry' received from the MLC. Menus are also available which show the velocities of each joint or the end-effector. Note that the 6D ball is physically connected to the TI machine. The reason for this is to avoid overloading the relatively slow AD machine.

It has been found necessary in practice to use an X terminal as an extra control device for starting up and monitoring the MLC and manipulator. This is useful for avoiding screen clutter both on the TI and AD machines. The terminal also serves to monitor the progress of other, subsidiary tasks and the networking system in use.

The manipulator control software for the STFP has limited intelligence. In particular, no path planning is implemented. The manipulator is thus not able to plan tasks from start to finish or find routes through a workspace containing obstacles. The aim is to provide this intelligence via the teleoperator who drives the manipulator to perform its task. There are, of course, built in safety features such as collision avoidance and velocity limits to protect the hardware in case of such eventualities as misuse, poor planning or command errors due to transmission failures.

Experience with the Teleoperator Interface

The performance of the MMI layout was successfully demonstrated in November 1994 by *the Teleoperations Demonstration Experiment* (TDE) between DLR and the Spanish Aerospace Institute, INTA, in Madrid. For this experiment, a ground version of the MINISAT/MINIMAN flight segment located in Madrid was remotely controlled (teleoperated) from a ground station located in Oberpfaffenhofen via the HISPASAT telecommunications satellite[3]. Communication was established by making use of the SLIP protocol over the satellite link.

The function of the set-up at DLR was that of a Teleoperator Station (TOS) controlling a PUMA robot through the set-up at INTA which represented the flight segment. The TOS controlled the robot using typed-in commands and the 6D ball, an animation display running KISMET and video pictures of the scene at INTA. Voice contact with INTA staff was also available. The aims of the experiment were to demonstrate the feasibility of teleoperations with this configuration, improve the computer systems on both sides, verify the command interface description and determine the extent and effect of transmission errors.

The experiment was successful in that the robot at INTA could be controlled from DLR to approach a desired target, the communication links functioned satisfactorily and the command interface could be verified. The communications delay was around 800ms with approximately 2% packet loss. These losses did not seem to degrade the system appreciably. The situation can be improved by using a better protocol, PPP, for example. The operator preferred the 'dominant' mode of the 6D ball whereby only the strongest signal is used implying that one joint or axis is controlled at once. The tests also showed up several ways in which the systems could be improved:

- Because of the lack of collision avoidance features, the PUMA could only be moved at a low speed. This meant that the volume of commands had to be restricted in order not to overload the flight segment software.
- Problems in the ball control software and the inverse kinematics were corrected.
- Problems with the animation display were corrected.
- Problems with the way in which movement commands were processed in the flight segment software pointed the way to changes there.
- The need for a graphical representation of the commands issued and the current position was highlighted (this has since been implemented).

The TI set-up shown in Figure 3 is currently used for testing at DLR. This represents a bare system without end-effector, camera system, collision avoidance, SIL simulation or time delay. The experiments have centred around operating the robot in real-time in both local and global modes in order to improve the software and hardware configurations and the user interface. Several results are worthy of mention:

- The use of the 6D ball to control more than one degree-of-freedom is difficult. Use of other devices must be investigated.
- Menus must be carefully designed to avoid screen clutter and present only the most essential information to the operator.
- Division of tasks between the various machines and terminals must be carefully determined to avoid scheduling problems and screen clutter.
- The operator's work area must be carefully designed.
- Software process interaction must be carefully tuned to avoid scheduling and resource clash problems.

Concluding Remarks and Further Activities

The most important conclusion from the work is that the devices and methods offered to the teleoperator must be improved. This involves, in particular, further investigations into the effectiveness of and alternatives to the 6D control ball.

The experiments have also shown up deficiencies in the user interface builder in use. These relate particularly to resizing of menus, setting up a hierarchy of menus and about how to avoid screen clutter. A review of current user interface builders is thus necessary with a view to the possible replacement of TAE+.

The work on developing a teleoperator interface reported here can be considered to be a prototyping exercise to define the **requirements** on such an interface. Once this exercise is complete, the requirements can be used as input to the design process of developing an integrated system. This design process can be based on more modern methods of task description and task level planning[4] or the establishment of environment models for known work areas and the involvement of graphical object-oriented operator methods.

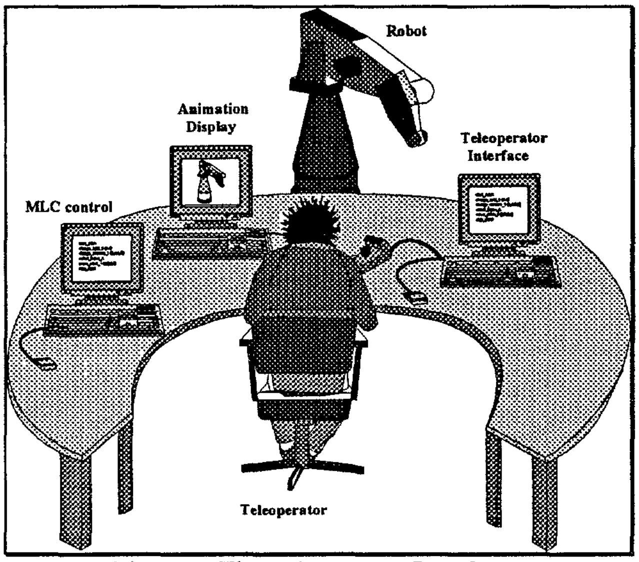


Figure 3. The Teleoperator Interface.

References

1 G. Hirzinger, B. Brunner, J. Dietrich, J. Heindl. *Sensor-Based Space Robotics - ROTEX and its Telerobotic Features.* IEEE Transactions on Robotics and Automation, Vol. 9, No. 5, 649-663, Oct 1993.

2 B. Schäfer, G. Burghaus, M. Schlenger, A. Seth. *Telerobotic Test and Training Facility for Ground-Based In-Orbit Servicing.* 2nd Workshop on Simulators for European Space Programmes, 25-27 Nov 1992, ESTEC, Noordwijk, NL.

3 J. Gomez-Elvira, J. Santo-Tomas, T. Munoz, E. de la Fuente, B. Schäfer. *Platform for Robotic Simulation for LEO Satellites.* 3rd Workshop on Simulators for European Space Programmes, 15-17 Nov 1994, ESTEC, Noordwijk, NL.

4 A. Elfving, U. Kirchhoff. *Design Methodology for Space Automation and Robotic Systems.* ESA Journal, Vol. 15, 149-164, 1991.

INTEGRATED DESIGN OF A PILOT VEHICLE INTERFACE FOR 4-D GUIDANCE AND NAVIGATION: FROM CONCEPT TO IN-FLIGHT DEMONSTRATION

Erik Theunissen
Delft University of Technology, Faculty of Electrical Engineering
P.O. Box 5031, 2600 GA Delft, The Netherlands, (e.theunissen@et.tudelft.nl)

Abstract. To efficiently utilize the flexibility in data presentation offered by today's programmable display systems, an integrated approach to the design of the Man-Machine Interface is required, necessitating a seamless fusion of knowledge from the different disciplines involved in the design process. This paper describes the development of the DELPHINS Tunnel-in-the-Sky display, and shows how an integrated approach has been applied to answer many of the design questions.

Keywords. Integrated design, perspective displays, guidance, navigation, man-machine interface

1. INTRODUCTION

The introduction of digital datalinks between aircraft and Air Traffic Control (ATC) and the advent of highly accurate positioning systems, offers the possibility to increase airspace capacity by decreasing separations between aircraft. By using flexible curved approach procedures, ATC has more freedom in managing the traffic flow, resulting in a better utilization of airway and runway capacity. The resulting increase in requirements on position and velocity control of the aircraft and the fact that approach paths may contain curved segments, will certainly increase the pilot's workload and reduce his ability to maintain an adequate level of spatial and navigational awareness. This can be compensated for by providing the required data in such a way that the effort for interpretation, integration, and evaluation is reduced.

Conventional guidance displays employ a very simple presentation, e.g. a moving bar indicating a deviation to be zeroed. The design of algorithms driving the guidance display is a typical control engineering problem. The introduction of programmable display systems on the flightdeck offers almost unlimited flexibility in the presentation of guidance and navigation data, and as a result the possibility to improve the information transfer is available.

The development of advanced display formats requires consideration of perceptual and cognitive aspects. Due to the interdependency of requirements and constraints from the different disciplines involved, and the fact that margins exist, trade-offs are possible. The efficiency of the design process is largely determined by the ability to mediate requirements and constraints between the different disciplines, while the quality of the final product is significantly influenced by the trade-offs which have been made to satisfy the requirements within the constraints. As a result, it is very important that the consequences of trade-offs are clear for all disciplines involved in the design process. An approach is needed which allows potential concepts to be qualitatively evaluated against certain predefined criteria with respect to possibilities for interpretation, integration, and evaluation of the presented data.

In 1990 the Delft Program for Hybridized Instrumentation and Navigation Systems (DELPHINS) was initiated at the department of Telecommunication and Traffic Control Systems of the Faculty of Electrical Engineering. In the context of DELPHINS, research is performed into presentation methods for guidance and navigation data to improve the information transfer from machine to man. An example of a potential display concept for four-dimensional (4-D) navigation and guidance is the DELPHINS Tunnel-in-the-Sky display, which is characterized by a perspective presentation of the future flightpath.

This paper describes the design of perspective flightpath displays for aircraft guidance and navigation in a control-theoretical, cognitive, and perceptual context, while taking into account current and expected future technical possibilities and limitations.

2. GUIDANCE AND NAVIGATION

Navigation can be defined as "to direct and control the course of an aircraft". To fulfil the navigation task, guidance is required. This comprises control of elevator, aileron, rudder, and thrust. It can be performed manually, or automatically. In the latter case, since humans possess invaluable qualities in coping with unpredictable situations, the pilot functions as a supervisor. His role is to compensate for the limited flexibility and adaptability of automated systems in the event of an unforeseen circumstance for which the system was not designed. To exploit the flexibility and adaptability of the human operator, the system must be designed so that the pilot is able to quickly detect anomalies and to safely and rapidly take over full control of the aircraft. For the safe execution of the guidance and navigation task, it is important that the pilot is able to determine the relation between his Ego-centered Reference Frame (ERF) and the World Reference Frame (WRF), thus establishing an adequate level of spatial awareness. Furthermore, in order to be able to anticipate changes, it is important that the pilot is able to predict the future required ERF-WRF relation, which is determined by his navigational awareness.

The Navigation Error (NE) of an aircraft consists of a Positioning Error (PE) and a Flight Technical Error (FTE). The PE is the difference between the true position of the aircraft and the position reported by the positioning system. The FTE represents the difference between the desired position of the aircraft and the position reported by the positioning system. The pilot is only aware of the FTE, and a change in PE will be perceived as a change in FTE.

Today's aircraft displays mostly employ singular and sometimes dual dimensional data presentation methods for guidance and navigation data. The integration of the data which is required to obtain spatial and navigational awareness has to be performed by the pilot. This process involves mental rotation and scaling operations, which costs time and may introduce errors. With one-dimensional (1-D) and two-dimensional (2-D) guidance and navigation displays, position and orientation data is either presented separately, or combined into one parameter. The Navigation Display (ND) presents a plan view of the flightpath relative to the aircraft position (Figure 1).

As a result, it contains 2-D (lateral) position information, and 1-D orientation information (heading). Depending on the mode, a WRF (North Up) or an ERF is used (Track or Heading up). The Attitude Indicator (AI) presents the pitch and bank

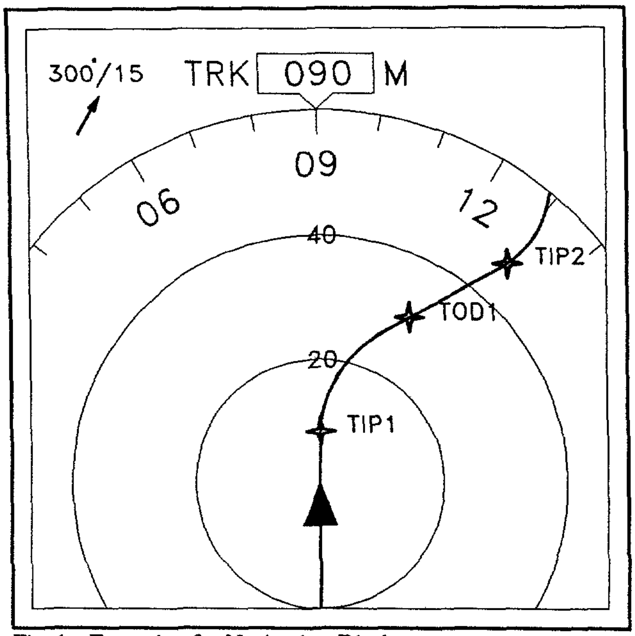


Fig. 1. Example of a Navigation Display

of the aircraft relative to a depiction of the horizon. In general a so-called inside-out frame-of-reference is used (fixed airplane symbol against a moving horizon), although Russian aircraft employ a hybrid solution, in which the aircraft symbol rolls but is fixed in the vertical direction, and the artificial horizon translates in the vertical direction to convey pitch information. By allowing the aircraft symbol to roll against a fixed background, the principle of control display motion compatibility (Johnson and Roscoe, 1972) is satisfied. The altimeter presents 1-D position information, and can also be used to indicate the desired altitude. The glideslope and localizer indicators present 1-D position error information, while a flight director presents guidance commands. Figure 2 presents an example of a conventional guidance display.

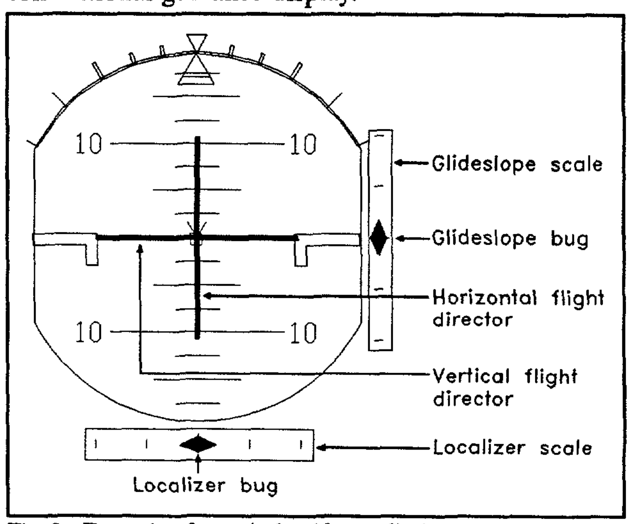


Fig. 2. Example of a typical guidance display

By integrating the information conveyed by the ND, the AI, and the altimeter the pilot is able to obtain a certain level of spatial and navigational awareness. Conventional flight directors are based on a weighted combination of position- and angular errors. In the horizontal dimension, Cross-Track Error (XTE) and Track-Angle Error (TAE) are used to calculate the deflection of the vertical flight director bar. In the vertical dimension, Flight path Angle Error (FPAE) and Vertical Error (VE) are used to calculate the deflection of the horizontal flight director bar. As a result of the integration of multiple parameters into a single dimension, the pilot is unable to extract information about the specific errors from the flight director display. Furthermore, since the error gains of the display are determined by the flight director algorithms, the possible bandwidth the pilot can apply for scanning and executing the flight director commands is rather limited. In situations where the required performance is less than the performance for which the gains have been determined, the pilot is forced to maintain the higher gain, and the possibility to neglect errors for a certain time is very limited. Finally, the flight director does not present the pilot with preview on the future desired trajectory which is required for anticipatory control. The ND presents the pilot with trajectory preview in the horizontal dimension, required for lateral navigational awareness. However, the resolution of this data is too low to be useful for anticipatory control. As a result, the pilot is forced to apply a continuous compensatory control strategy.

3. DESIGN QUESTIONS

The goal of the design process is to optimize the information transfer from machine to man. One of the most effective mechanisms for the simplification of complex visual scenes is the human perceptual system (Garner, 1970). This simplification mechanism is developed in humans through years of repeated confrontation with the rules of perspective scenes. With this system, the human is capable of rapid interpretation of otherwise complex visual scenes. To capture this simplification capability in man-machine systems requires the use of pictorially realistic information presentation (Jensen, 1978)

The advancements in the area of computer graphics make it technically and economically feasible to present an abstract, dimensionally and dynamically compatible analogy of the spatial environment in real-time. Such Computer Generated Imagery (CGI) can be used to emphasize important features in the outside world scene, de-emphasize or eliminate unimportant features, and introduce artificial cues.

To reduce the required effort for interpretation and evaluation, emergent features can be used to exploit certain cognitive abilities which are involved in the early stages of perceptual processing. The Proximity Compatibility Principle (PCP) predicts that tasks requiring the integration of information across sources benefit from more integrated displays (Wickens and Andre, 1990). By presenting the data so that the presentation is compatible with the user's expectation, semantic distance can be minimized (Norman, 1989). The spatial presentation of the imaginary flightpath in the 3-D environment can be used to alleviate the pilot from performing the mental integrations of the separately displayed position and orientation data into a spatially coherent picture.

For the design of a 3-D guidance and navigation displays, questions regarding the contents and representation of the real-world analog must be answered. The following first three questions address the contents, while the latter six address the representation.

- How to determine which objects in the visual environment contribute, and should be emphasized, and which objects mainly cause clutter?
- When to employ representations of imaginary elements?
- How to determine whether and when additional data presentation is necessary?
- How can the objects be represented and to what abstraction level can the representation be reduced?
- How to emphasize important objects?
- How to employ representations of imaginary elements?
- How to integrate additional data into the presentation?
- How to select the perspective design parameters and the frame of reference?
- How to select the presentation medium?

For the implementation and the integration in a target environment the following additional questions must be addressed:

- What are the system performance requirements in terms of memory, speed, and display resolution?
- What data is required?
- What are the requirements with respect to data latency, update-rate, accuracy, noise?

Addressing these questions requires a more detailed analysis of the specific properties of spatial data presentation in relation to the anticipated tasks to be performed. Such an analysis also allows the comparison with findings from other studies related to a specific aspect. Important questions which must be addressed are:

- What are the specific properties of spatially integrated environment and trajectory presentation, and what are the similarities and fundamental differences with 1-D and 2-D datapresentation?
- What are the consequences/possibilities of spatially integrated data presentation with respect to perception, interpretation, evaluation, and action?
- What are the consequences of a mismatch between the presented and perceived virtual space?
- What is the influence of data latency, limited update-rate, limited accuracy, noise?
- What is the influence of non-ideal operating conditions like turbulence, crosswind?
- What are the specific advantages and disadvantages of spatially integrated datapresentation?
- What are possibilities to compensate for deficiencies, limitations and disadvantages?

4. DESIGN

A perspective flightpath displays presents a spatially integrated view of the future 3-D trajectory on a 2-D display (Figure 3).

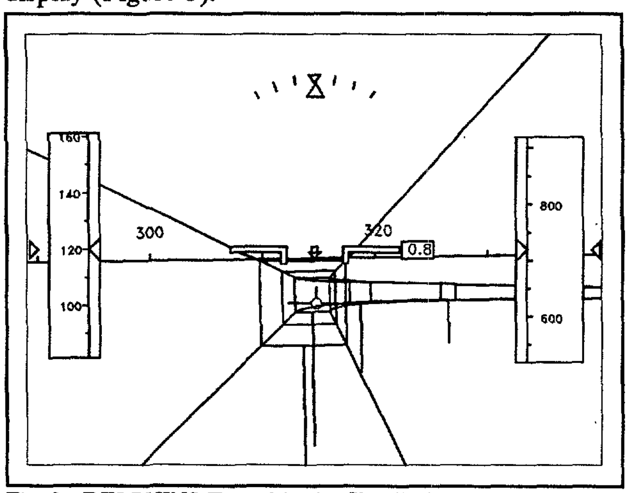


Fig. 3. DELPHINS Tunnel-in-the-Sky display

4.1. Frame of reference

Based on the frame of reference used for the projection, these displays can be divided into egocentric and exocentric ones. In an egocentric perspective flightpath display, the 3-D world is depicted as seen from the aircraft. In an exocentric display, the situation is viewed from another position. With an egocentric perspective projection, information about position and orientation errors is conveyed through a distortion of the natural symmetry of the presented trajectory. Since the detection of symmetry takes place in the early processing cycles of visual information, this feature can be exploited to reduce the required effort for interpretation and evaluation. Any other frame-of-reference than an ego-centered one cannot exploit this advantage, and will require additional mental processing. Therefore, an egocentric projection was selected.

4.2. Design parameters

Position and orientation errors distort the symmetry of the representation of the tunnel. It is the distortion of the symmetry that is perceived, and not separate position or orientation errors. Theunissen (1994b) describes the relation between the distortion of the symmetry and the position and orientation errors as a function of the design parameters of the perspective display.
The motion of the aircraft relative to the virtual tunnel allows the extraction of error rates and produces additional cues which are conveyed through the presentation of successive snapshot images of the situation. In Theunissen and Mulder (1995a) it is discussed how data about position errors, and rotation rates are present in the visual flow field. These dynamic cues give the pilot a sense of egospeed. Besides a cue for egospeed, pilots can extract temporal range information from the display. Temporal range judgements are based on global optical flow rate, which must exceed a certain threshold to allow accurate estimates to be made. Temporal range information is often used to determine the moment to initiate certain anticipatory control actions. Theunissen and Mulder (1994, 1995a) studied the relation between the moment an error-correcting control action is initiated and temporal range information in a perspective flightpath display. In Theunissen and Mulder (1995b), some requirements on the design parameters to generate adequate temporal range cues are discussed.

4.3. Representation of the flightpath

Just as with real-world objects, the meaning of an imaginary element should be intuitively apparent from the representation. Since the real-world 2-D counterpart of a 3-D trajectory is a road, the desired flightpath is often visualized as a 3-D road. Various representations have been tried in the past, resulting in designations such as Flightpath Channel (Wilckens and Schattenmann, 1968), Pathway-in-the-Sky (Hoover *et al.*, 1983) and Tunnel-in-the-Sky (Grunwald, 1984). In Theunissen (1994b), the representation of a flightpath is divided into a flightpath element, cross sections, and altitude poles based on the following three different functions:

- provide position and orientation information
- resolve ambiguities in the trajectory
- resolve ambiguities towards other objects

Position and orientation errors are provided by all elements of the flightpath. The ambiguity within the representation is resolved through the presentation of cross-section frames, which in combination with the observers expectation about the shape of the object provide a cue for resolving ambiguities. As a result of the apparent motion of the cross-section frames towards the observer, and the resulting optic flow field, the feeling of three-dimensionality increases, and ambiguities are further reduced. The ambiguity towards other objects, notably the ground, is resolved through the presentation of altitude poles. The altitude poles also provide a possibility to temporarily use a very high lateral error gain, which will be discussed later.

Various representations of the flightpath have been tried in the past. It must be realized that especially in the early period of research into perspective flightpath displays, the representation was dictated by the limitations of the available means to generate perspective images in real-time. Wilckens and Schattenmann (1968) used dots to indicate the corners of cross-section frames in his 'channel display'. Hoover *et al.* (1983) represented their 'pathway-in-the-sky' by means of tiles. Jensen (1978) used 'telephone poles' to visualize the desired trajectory. None of these formats did employ a continuous presentation of the flightpath, i.e. no interconnections existed between the references. In the absence of such interconnections, the error gains in the display are determined by the positions of these trajectory frames. Grunwald (1984) and Wickens *et al.* (1989a) both used interconnections, yielding a continuous presentation of the desired trajectory, and as a result of the error gains. As discussed previously, the height and width of the tunnel determine the position-error gain. Sometimes, it is desirable to also have a source of a very high position error gain which can be used for temporal fine-tuning. Reducing the tunnel size to obtain this high gain would force the pilot to continuously apply a high control gain, which reduces the flexibility. This problem can be solved by presenting references indicating the center of the tunnel sections. In this way, horizontal and vertical error gain can be used separately. In fact, the altitude poles already provide such information for lateral control. During experiments performed in the flight simulator of the Delft University of Technology, pilots mentioned that in the final approach they used the alignment of these poles for accurately positioning the aircraft on the centerline. An alternative might be to present a diamond shaped cross-section. This, however, introduces a number of drawbacks of which the discussion goes beyond the scope of this paper.

4.4. *Identification of objects to be presented*

The identification of objects which are to be displayed requires a method to identify which objects in the visual environment contribute to the tasks to be performed, and which objects mainly cause clutter. With respect to the guidance and navigation task, objects which function as an important reference for spatial orientation and/or navigation in the 3-D world are considered relevant. Examples are objects with a known geographical location, and objects with a familiar shape and/or size, allowing the observer to estimate his relative position. With respect to collision avoidance, the presentation of objects which might constitute a potential hazard is desired. The two most important objects of the latter category are terrain and other aircraft. An imaginary element is the position predictor, which depicts the future estimated position of the aircraft.

4.5. *Presentation of objects*

For the presentation of objects, the question regarding the level of detail of the representation must be addressed. In this context, the highest level of detail is considered a representation which is visually indistinguishable from the real-world analogy. Besides the fact that this would be a computational extremely expensive operation, in most cases such a high level of detail is likely to result in clutter, and hence not desirable. Thus, the question is: 'to what abstraction level can the object representation be reduced?'. However, the question is not complete yet, since an important constraint regarding the required effort for interpretation must still be specified. This constraint is formulated as: 'the real-world objects must be intuitively recognized from the abstract representation'. With the current version of the display, terrain is depicted as a 3-D mesh, in which the height of each point is determined by the maximum altitude within a predefined range. Color coding is used as an additional means to convey terrain altitude. Other traffic is presented as aircraft symbols, similar to the symbology used by Ellis *et al.* (1987) in their perspective Cockpit Display of Traffic Information (CDTI) studies. In certain situations it might be necessary for the pilot to focus his attention on a specific object, for example in case the object poses a potential hazard. Attributes such as color, intensity, blinking, and magnification can be used to emphasize such an object. Since the attention of the pilot is influenced by his expectations and motivation, features must be used that are strong enough to attract his attention regardless of a certain bias. With the current display format, two types of objects, representing two different types of threats (terrain and other aircraft), can be emphasized by a change in color and by blinking. To exploit the common population stereotype of red for danger, terrain which is below the aircraft altitude and aircraft which constitute a potential collision hazard are colored red. When the time to collision reaches a certain minimum threshold, the representation of the corresponding object(s) starts to blink. To present the future predicted position of the aircraft to the pilot, an abstract presentation of an aircraft is used. Position ambiguity is resolved by presenting the imaginary cross-section of the tunnel at the future position of the aircraft. This cross-section is transparently highlighted, which in turn avoids occlusion of other objects.

4.6. Disadvantages and compensations

A spatially integrated presentation is only beneficent when integration of information from the three spatial dimensions is required. With 1-D and 2-D datapresentation methods it is possible to use a constant scaling for the depiction of the desired data. With 3-D displays the accuracy with which a singular parameter can be determined is often a function of position, orientation, and velocity of the viewpoint. 3-D displays suffer several other limitations which must be taken into account. As a result of the integration of the third dimension, the resolution of the information along the viewing axis decreases with increasing distance from the viewpoint. Furthermore, due to the integration of multiple parameters into a single object, it is often harder to estimate the value of a parameter in a single dimension (Wickens *et al.*, 1989b). Also, angular distortion occurs, which makes it very hard to estimate angles in planes which are not perpendicular to the viewing direction (McGreevy and Ellis, 1986), and finally objects which are close to the observer might mask objects which are further away.

From the previous discussion, two drawbacks of perspectively projected spatially integrated data can be identified which might need to be compensated for: the lack of an angular reference in curved sections, and the reduced accuracy with which single spatial parameters can be estimated. The former problem can be compensated for by presenting a position or track prediction relative to the desired track. The latter problem, resulting from the perspective projection, can partly be compensated for by integrating virtual metrical aids, or by separately presenting the required data. The warping of virtual metrical references is equal to the warping of the other data, which reduces the errors resulting from this distortion.

4.7. Integration of additional data

As indicated in Section 4.6, a disadvantage of perspective data presentation is that the integration makes it harder to estimate singular parameters, and the fact that the accuracy is determined by the position, orientation, and velocity of the viewpoint. By analysing the information which is required for the tasks to be performed with respect to accuracy, and comparing this with the way this information is conveyed through the perspective presentation, assumptions can be made about the necessity of additional information. Examples are the presentation of airspeed, roll angle, and altitude. To maximize spatial and representational consistency with current displays, the additional data about altitude, airspeed, and roll angle is integrated in a way which is equivalent both in location and representation with today's PFD.

4.8. Dealing with constraints

A major difference between command displays such as the flight director, and perspective flightpath displays such as the Tunnel-in-the-Sky, is that the former is based on the presentation of a weighted sum of position and angular errors and error rates, whereas the latter presents an abstraction of the real world, and thus is based on position and attitude. To avoid information conflicts, visual stimuli obtained through the perspective flightpath display must be compatible with visual stimuli from the outside world and the motion cues obtained through the vestibular system. In order for the pilot to believe the flight director, the commands must have a certain degree of consistency with the other information available. The fact that a flight director command is not required to have a one-to-one relation with any other perceivable cue, allows for certain differences in the update-rate of the required data. The data which is required for the closure of the inner control loop (attitude) has to satisfy more stringent requirements with respect to latency and update-rate as compared to the data required for the closure of the outer loop (position) (Hess, 1987). With a perspective flightpath display, the information is not combined into a single parameter. As a result, both position and attitude data must satisfy update-rate requirements which yield a smoothly animated display.

To achieve such a smoothly animated display, the data update-rate must exceed a certain threshold. Update-rates in the order of 20 to 30 Hz prove to be adequate. As a result of the limited bandwidth of the carrier tracking loop in GPS receivers (typically about 16 Hz), these receivers output position data at an update-rate below that required for smooth animation. In case it is impossible to oversample the position data, inter- or extrapolation techniques are needed to increase the position information update-rate. Interpolation introduces latency, which reduces the stability of the control loop due to a decrease in phase-margin. Thus, interpolation is only acceptable in case the position update-rate is sufficiently high. With extrapolation, the prediction, which is based on position data and models which use other elements of the state vector such as velocity, attitude, and heading is inevitably accompanied by a prediction error which is corrected at each new position update. These corrections, however, can be perceived as a sudden change in FTE, and introduce a noise component in the optical flow field with the same frequency as the position information update-rate, which can become very distracting. Therefore, the prediction algorithm must apply some form of error smoothing to avoid a distortion of the dynamic cues. An in-depth discussion of position prediction techniques is beyond the scope of the paper, however, more information about position prediction can be found in Mulder (1992). For the in-flight testing of the Tunnel-in-the-Sky display, a Kalman predictor with a circular-path message model was used.

5. RESULTS

In 1990, an initial concept for a perspective flightpath display was specified in the context of DELPHINS. In parallel, based on the anticipated system requirements, development of a display design system and target hardware for simulator and in-flight evaluation commenced (Theunissen 1994a). A first laboratory concept demonstration was given in the beginning of '91, and at the end of '91 the flight simulator at the Faculty of Aerospace Engineering was equipped with a programmable display system developed in the context of DELPHINS. Display format evaluations were performed in '91 and '92, and in '93 a study was performed into pilot closed-loop control behaviour (Theunissen, 1993). In 1994, open-loop control strategies were investigated (Theunissen and Mulder, 1994a). Furthermore, a concept for the integration of terrain and traffic information was developed and implemented. An in-flight concept demonstration with the laboratory aircraft of Delft University followed in december '94 (Theunissen, 1995). For this purpose, an airborne version of the display system and a simple experimental Flight Management System (XFMS) have been developed. The system is based on commercial of-the-shelf components. Position data is obtained from a GPS receiver, and through a datalink with a ground reference station, DGPS corrections are obtained, resulting in sub-meter accuracy. A simple XFMS and a database with the runway coordinates and the ILS approach path is used for the generation of the required trajectories. From the data of the XFMS and the actual position and attitude of the aircraft, the Display Electronics Unit (DEU) generates the perspective flightpath, which is presented on the Display Unit (DU). To execute a curved approach procedure, ATC vectors the aircraft towards an arbitrary point on the ILS path. The XFMS calculates a route from the current position of the aircraft to this point, and the DEU generates a perspective flightpath (Figure 4), allowing a smooth intercept of the final straight segment.

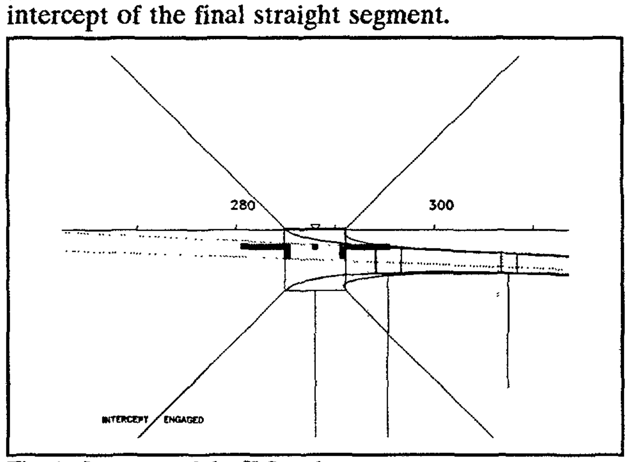


Fig. 4. Intercept of the ILS path

The radius of the curvature between the intercept segment and the ILS segment is determined by aircraft velocity and the desired bank angle.

6. CONCLUSION

As indicated in the introduction, the large degree of freedom resulting from the flexibility in data presentation with programmable displays poses the designer with new problems. An example is the design of a perspective flightpath display, which requires the specification of numerous parameters. An approach was needed which allows some kind of qualification of potential concepts with respect to the different domains involved in the design process. By means of a structured analysis of the specific properties of perspectively projected spatially integrated data, and by identifying the strengths and weaknesses, it is possible to:

- reduce the large number of degrees of freedom in the design,
- compare the possibilities with respect to interpretation, evaluation and action with conventional presentation methods,
- allow trade-offs to be made,
- compensate for deficiencies,
- define system requirements,
- justify design decisions.

The result of such an integrated approach is more than the sum of its parts.

8. REFERENCES

Ellis, S.R., McGreevy, M.W., Hitchcock, R.J. (1987) 'Perspective Traffic Display Format and Airline Pilot Traffic Avoidance' *Human Factors* Vol. 29 No. 4 (1987) pp. 371-382.

Garner, W.R. (1970) 'Good Figures have Few Alternatives', *American Scientist*, Vol. 58, No. 1, pp. 34-42.

Grunwald, A.J. (1984), *Tunnel Display for Four-Dimensional Fixed-Wing Aircraft Approaches*, Journal of Guidance, Vol. 7, No. 3, pp. 369-377.

Hess, R.A. (1987) 'Feedback Control Models', *In: Handbook of Human Factors*, Eds. Salvendy, G., Wiley, Chapter 9.5.

Hoover, G.W., Cronauer, V.T., Shelley, S., Dittenhauser, J.N., Eulrich, B.J., Reynolds, P.A. (1983), *The Command Flight Path Display, phase 1 and 2*, Final Technical Report, System Associates, Inc. Long Beach, CA.

Jensen, R.S. (1978), *The Effects of Prediction, Quickening, Frequency Separation, and Percent of Pursuit in Perspective Displays for Low-Visibility Landing*, Proceedings of the Human Factors Society - 22nd Annual Meeting, pp. 208-212.

Johnson, S.L. and Roscoe, S.N. 'What Moves; the Airplane or the World?' *Human Factors* Vol. 14 No. 2 (April 1972) pp. 107-130.

McGreevy, M.W. and Ellis, S.R. (1986) 'The Effect of Perspective Geometry on Judged Direction in Spatial Information Instruments' *Human Factors* Vol. 28 No. 4 pp. 439-456

Mulder, M. (1992) *Aviation Displays and Flightpath Predictors.* Master's Thesis, Delft University of Technology, Faculty of Aerospace Engineering.

Norman, D.A. (1989) 'Cognitive Engineering' In: *User Centered System Design - New Perspectives on Human-Computer Interaction* eds. Norman, D.A. and Draper, S.W., Lawrence Erlbaum Associates, Inc.

Theunissen, E. (1993). 'A Primary Flight Display for Four-Dimensional Guidance and Navigation: Influence of Tunnel Size and Level of Additional Information on Pilot Performance and Control Behaviour' *Proceedings of the AIAA Flight Simulation Technologies Conference*, August 9-11, Monterey, CA.

Theunissen, E. (1994a) 'The Development of the Delphins Display Design System' *Proceedings of the International Training and Equipment Conference*, pp. 583-588. The Hague, The Netherlands.

Theunissen, E. and Mulder, M. (1994) 'Open and Closed Loop Control With a Perspective Tunnel-in-the-Sky Display' *Proceedings of the AIAA Flight Simulation Technologies Conference*, August 1-3, Scottsdale, AZ.

Theunissen, E. (1994b) 'Factors influencing the design of perspective flight path displays for guidance and navigation' *Displays*, Vol. 15, No. 4, pp. 241-254.

Theunissen, E. (1995) 'In-flight application of 3-D guidance displays: problems and solutions' *Proceedings of the '95 IFAC MMS Conference*, Cambridge, MA.

Theunissen, E. and Mulder, M. (1995a) 'Error-Neglecting Control with Perspective Flightpath Displays' Proceedings of the Eight International Symposium on Aviation Psychology, Columbus, OH.

Theunissen, E. and Mulder, M. (1995b) 'Availability and Use of Information in Perspective Flightpath Displays' *Proceedings of the AIAA Flight Simulation Technologies Conference*, August 7-9, Baltimore, MD.

Wickens, C.D., Haskell, I., and Harte, K. (1989a) *Ergonomic Design for Perspective Flight Path Displays*, IEEE Control Systems Magazine, Vol. 9, No. 4, pp. 3-8.

Wickens, C.D., Todd, S., Seidler, K. (1989b) *Three-dimensional displays: Perception, implementation, applications* University of Illinois, Aviation Research Laboratory.

Wickens, C.D. and Andre, A.D. (1990) 'Proximity compatibility and information display: Effects of color, space, and objectness on information integration' Human Factors, 32, pp. 61-77.

Wilckens, V. and Schattenmann, W. (1968) 'Test Results with New Analog Displays for All-Weather Landings' *AGARD Conference Proceedings No. 55*, pp. 10.1-10.31

Index