Treating socio-technical systems as engineering systems: some conceptual problems

Peter Kroes, Maarten Franssen, Ibo van de Poel, Maarten Ottens

Dept. of Philosophy

Fac. Technology, Policy and Management

University of Technology Delft

Jaffalaan 5

2628 BX Delft, Nl

tel: (0)15 2785143

fax: (0)15 2786233

p.a.kroes@tbm.tudelft.nl

Treating socio-technical systems as engineering systems: some conceptual problems*

Peter Kroes, Maarten Franssen, Ibo van de Poel, Maarten Ottens

Dept. of Philosophy

Fac. Technology, Policy and Management

University of Technology Delft

Abstract

Systems engineering has been plagued by the problem of how to separate a system from its environment or context, in particular from its social context. We propose to include anything in the system that is necessary for performing its intended function and that may be the object of design. For certain engineering systems, such as civil aviation systems, this implies that human agents and social institutions have to be taken as integral parts of these systems. These 'socio-technical' systems are of a hybrid nature because they are constituted by different kinds of elements, intentional and non-intentional: social institutions, human agents and technical artefacts. This paper analyses two different roles that human agents, as elements of socio-technical systems, may play with regard to technical artefacts. Furthermore, it discusses some conceptual problems concerning the modelling of socio-technical systems that are due to the hybrid nature of these systems.

---

*Paper presented at the Engineering Systems Division Symposium, MIT, 31 Maart, 2004. Note that the notion of socio-technical system as it is used here is not to be confused with the notion of socio-technical systems theory which addresses issues about labour organization in relation to technology; cfr. Jackson (1991, p. 59 ff).

Key words: socio-technical systems, artefacts, human agents, social institutions, model.

1. Introduction: a conceptual problem

The field of systems engineering has inherited a conceptual problem from systems theory. Just as systems theory since its beginnings has been plagued by the question how to separate a system from its environment or context, the field of systems engineering has been confronted with a similar question about engineering systems. How are the boundaries of (engineering) systems to be drawn? What belongs to the (engineering) system under consideration and what to its environment? For engineering systems this problem manifests itself conspicuously with regard to the status of non-technical elements, such as social, political, economic and institutional ones. To what extent are these, or ought these elements to be considered to belong to engineering systems or to the environment or context?

The following remarks illustrate that this problem is still alive within systems engineering circles. One of the leading institutions in the field, the Engineering Systems Division (ESD) of the Massachusetts Institute of Technology defines an engineering system in the following way (ESD-WP-2003-01.20, p. 19):

> Engineering system – a system designed by humans having some purpose; large scale and complex engineering systems, which are of most interest to the Engineering Systems Division, will have a management or social dimension as well as a technical one.

This definition does not present a clear vision on the status of non-technical elements with regard to engineering systems. It leaves open what kind of elements are to be included in an engineering system, but strongly suggests that it may consist exclusively of technical elements and that only for a specific type of engineering system non-technical elements are a relevant dimension, viz. large scale and complex engineering systems. The definition has to be interpreted against the background of the Multidimensional View of Engineering Systems (ESD-WP-2003-01.20, p. 6 ff). A closer look at this view of engineering systems does not resolve our issue. On the contrary, it contains an ambiguity with respect to the status of non-technical elements. According to this view every engineering system is located in a three-dimensional space spanned by Technology, Management and Society. These three dimensions are the defining features of engineering systems, albeit that the technical dimension is the "central, defining feature" (p. 6). Nevertheless, the social dimension is characterized as the "context dimension" (p. 8). But how can this third feature of engineering systems be at the same time a defining feature, i.e. an intrinsic aspect of engineering systems, and its context?

The ambiguity of the Multidimensional View of Engineering Systems with regard to the status of non-technical elements is also illustrated by the following remark in the same document: "There are always key boundary issues associated with the contextual dimension. For some purposes, economic, legal and political factors are appropriately treated as exogenous factors impacting complex engineering systems. For other purposes, they are very much a part of the system under consideration." (p. 8). This immediately raises the question which kind of purposes allow these factors to be treated as exogenous and which as intrinsic to the system and on what grounds. And for multi-actor systems we have to add the question: whose purposes? Apparently, the status of social, institutional etc. elements depends on how the system boundaries are drawn, and there may be pragmatic reasons, depending on the specific problem and purpose at hand, to draw the boundaries such that these factors become either exogenous or intrinsic to the system under consideration. But pragmatic reasons have to be treated with care, since pragmatic reasons are determined to a high degree by professional practices and these in turn by prevailing institutionalised forms of the division of labour. And precisely existing forms of the division of labour may, from a systems perspective, be problematic.

One of the most obvious ways of approaching the problem of determining the boundaries of engineering systems is by including anything in the system that is necessary for performing its intended function and that can be the object of design. This is the approach taken in this paper. The intended function is taken to be the unifying principle of the system, that which makes the system into a 'unified whole'. On the basis of this starting point and an analysis of the notion of technical artefact, we will explore the status of agents and social factors with regard to engineering systems. We aim to show that at least three different situations have to be distinguished (see Table 1): (1) engineering systems which perform their function without agents and social institutions performing a sub-function within the system, (2) engineering systems in which some agent performs a sub function without social institutions playing a role, and (3) engineering systems which cannot perform their function without agents and some social/institutional infrastructure being in place. In the last case it seems appropriate to speak of socio-technical systems. The second aim of this paper is to show that from a systems-theoretic point of view engineering systems of the second and third type raise fundamentally different problems as compared to systems of the first type, since they are hybrid systems containing different kinds of elements and relations. In the final section we will indicate some of the problems engendered by this hybrid nature for the modelling of socio-technical systems. To be sure, the fact that these hybrid systems raise specific problems for engineering practice has already often been observed, e.g. in information systems design, and new approaches and tools for dealing with these problems pragmatically have been developed. In this paper, we are, however, primarily interested in spelling out some of the conceptual problems tied to the idea of hybrid, socio-technical systems, in particular conceptual problems with regard to modelling such systems.

	Without Agents	With Agents
Without Social Institutions	Tire of Landing gear	Airplane
With Social Institutions	?	Civil Aviation

Table 1 Type of Engineering Systems

We will start with some preliminary remarks on the relation between technical artefacts and engineering systems.

2. Technical artefacts versus engineering systems

The difference between technical artefacts and technical systems seems to be one mainly of scale and complexity. Civil aviation is, for instance, a complex public transportation system making use of all kinds of technological artefacts (airplanes, runways, luggage transport systems, communication equipment, fuel, etc.), all of which contribute to or are indispensable for the functioning of this transport system as a whole. This way of distinguishing between technical artefacts and systems is in line with the idea that the emergence of systems engineering as a separate engineering discipline is related to a specific phase in the development of technology, namely one in which large-scale, complex technological systems come to play an important role. Whereas the various branches of traditional 'artefact' engineering are focused on the design, development and production of isolated, stand-alone material objects with a specific technical function, systems engineering is focused on the design, development and implementation of complex combinations of technical artefacts (coming often from various engineering branches).[1]

A problem with this distinction between technical artefacts and systems is that the meaning of the notions of scale and complexity is highly context-dependent. What can be considered as a technical artefact from one perspective may appear as a complex technical system from another. Airplanes can be considered as technical artefacts

[1] See for instance Auyang (2004), p. 170 ff.

within a high-level analysis of civil aviation as a public transport system, but they become extremely complex technical systems by themselves when looked upon from the point of view of airplane design. And from the latter perspective, a tire from the landing gear may appear to be a 'simple' technical artefact, but it again becomes a complex system when looked upon from the point of its physical-chemical composition. Thus the technical system vs. technical artefact distinction is a relative one. When one tries to decompose, from a certain perspective, a technological system in terms of its constituent technical artefacts, a kind of Russian-doll effect occurs: from another perspective the constituent technical artefacts themselves may turn out to be complex technological systems. We end up with a picture in which technical systems are embedded, as technical artefacts, in higher level technical systems, which themselves are embedded, as technical artefacts, in again higher level systems, etc.

If indeed the system-artefact distinction is a relative one, it becomes difficult to understand what is so special about systems engineering *vis a vis* traditional forms of engineering. The latter are also dealing, from their own perspective, with complex technological systems. We think, however, that there are two reasons why the Russian-doll metaphor for the relation between technical systems and technical artefacts is misleading. Each Russian doll contains another doll of exactly the same shape, only smaller. This appears not to be true for the system-artefact distinction: going down the hierarchy of systems composed of artefacts which themselves may be considered to be systems, the nature of the systems involved appears to change in at least two ways.

In the first place, lower in the hierarchy the functioning of engineering systems becomes less dependent on non-technical aspects. Consider again civil aviation. The functioning of this transport system is not only highly dependent on the functioning of technical artefacts, but also on the functioning of social (legal, institutional, economic) elements and on the behaviour of various actors. When we put our focus on airplanes, the technological nature of the system comes much more to the fore; nevertheless non-technical elements still play an important role with regard to the functioning of this system (e.g. the behaviour of pilots). Going still further down the line, we reach technical systems whose functioning at first sight does not involve non-technical elements at all, for instance landing systems or turbine engines. Thus, the nature of the systems involved changes and related to that the meaning of the notion of the functioning of the system. At the highest level technological systems are embedded in social systems and the functioning of the system is not purely a technical matter. At the lower levels technical systems are as much as possible isolated from their social context and their functioning seems at first sight to be a matter of technology alone.

A second significant change concerns the variety of engineering disciplines involved in designing and developing systems at the various levels. The design, manufacture and operation of the technological infrastructure of a civil aviation system involve almost all traditional engineering disciplines. This is not to say that each technological element requires the cooperation of all these disciplines, but in many cases close cooperation between engineers from different disciplines is necessary. Far down in the hierarchy (especially at the level of technical components) systems tend to become much more mono-disciplinary in nature (a mechanical, electrical, hydraulic, etc. system).[2]

So if we enlarge the boundaries of engineering systems, the heterogeneity of the system increases in a double sense: the system tends to contain technical artefacts from different engineering disciplines and additionally non-technical elements beside technical artefacts. In a way, this will increase the complexity of engineering systems, which may make systems engineering higher up in the hierarchy different from traditional forms of engineering. The latter appear to be dealing with technological objects proper; that is, with objects whose function is completely detached from any social/institutional context. At first sight it seems that the functions of these purely technological artefacts can be described and analysed exclusively on the basis of their physical make-up. However, the idea that the function of a technical artefact may be understood in terms of its physical, more generally, material make-up, is problematic.

3. The dual nature of technical artefacts

We claim that technical artefacts have a dual nature: the function of a technical artefact is grounded on the one hand in its physical structure, on the other hand in a context of

[2] The mono-disciplines involved are the result of historically grown patterns of division of labour in engineering practices; the viability of these patterns of division of labour may, of course, be put into question.

intentional human action.[3] So the idea that the function of technical artefacts can be understood by looking only at their physical make-up has to be rejected.

Before presenting this dual nature view of technical artefacts, we will briefly discuss Herbert Simon's theory on artificial things as exposed in his classic *The sciences of the artificial* (in the following text, page numbers refer to (Simon, 1996)). This theory proves to be a nice steppingstone to the dual-nature view.

For Simon the science of the artificial will closely resemble the science of engineering because engineering deals with the synthesis of things. In contrast to the scientist, the engineer and more in particular the designer is "concerned with how things *ought* to be – how they ought to be in order to *attain goals*, and to *function*" (p. 4-5). One of the striking features of (technical) artefacts is precisely that they can be characterized in terms of functions and goals. Functions and goals are analysed by Simon in the following way (p. 5):

> Let us look a little more closely at the functional or purposeful aspect of artificial things. Fulfillment of purpose or adaptation to a goal involves a relation among three terms: the purpose or goal, the character of the artifact, and the environment in which the artifact performs.

For instance, the purpose of a clock is to tell time and the character of the clock refers to its physical makeup (gears, springs, etc. for a mechanical clock). Finally, the environment is important because not every kind of clock is useful in every

---

[3] See also Kroes (2002).

environment; sundials can only perform their function in sunny climates. Simon's analysis of artefacts is represented in a schematic way in figure 1.[4]

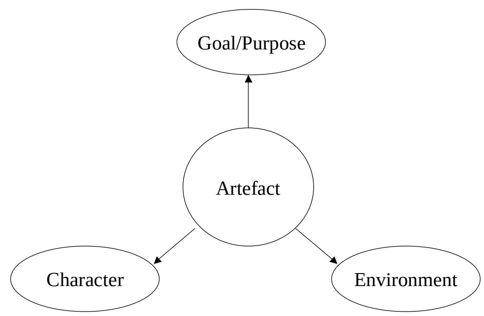

Figure 1 Simon's analysis of artefacts

According to Simon the environment of an artefact is very important because it moulds the artefact. He considers the artefact to be a kind of 'interface' between "an 'inner' environment, the substance and organization of the artifact itself, and an 'outer' environment, the surroundings in which it operates" (p. 6). The inner environment of the artefact, its character, is shaped in such a way that it realizes the goals set in the outer environment (p. 10). Therefore, the science of the artificial has to focus on this interface, since the "artificial world is centered precisely on this interface between the inner and outer environments; it is concerned with attaining goals by adapting the former to the latter" (p. 113).

Simon's distinction between inner and outer environment points to two different ways of looking at technical artefacts. Looked upon from the outer environment, the

[4] The arrows stand for conceptual implication: the notion of an artefact conceptually implies the notion of a character, a goal or purpose and an environment.

technical artefact presents itself primarily as something, whatever its inner environment, that fulfils a certain goal, purpose or function. From this perspective the artefact is characterized primarily in a functional way; the inner environment remains a black box. Looked upon from the inner environment, the artefact is described as some kind of physical system; from this perspective, the goal that it fulfils in the environment remains a black box. As Simon remarks (p. 7) "*Given* an airplane, or *given* a bird, we can analyze them by the methods of natural science without any particular attention to purpose or adaptation, without reference to the interface between what I have called the inner and outer environments." These two different ways of characterizing artefacts, in terms of its inner and outer environment, correspond closely to what we call the dual nature of technical artefacts.

Our starting point is the following characterization of technical artefacts: *technical artefacts are objects with a technical function and with a physical structure consciously designed, produced and used by humans to realize this function*. In short, a technical artefact is a physical object with a technical function. In so far as technical artefacts are physical objects, they can be described in terms of physical properties and the way they work can be explained in terms of causal processes. But as mere physical objects, they are not technical artefacts.[5] Their function turns them into technical artefacts and it is their function that ties technical artefacts to human action, because it makes no sense to speak about technical functions without reference to a context of human action. Functional discourse is part of the intentional conceptualisation of the world in which

[5] Note that in the following we will make a distinction between an artefact and a technical artefact: an artefact is an object that has come into being as the result of intentional human action; a technical artefact is an artefact that performs or is ascribed a practical function.

reference to goals and ends makes sense and in which means-end relations play an important role. It is meaningless to speak about technical functions without a context of intentional (human) action. This characterization of a technical artefact makes it an object with a dual nature: it is an object with a function, which on the one hand is closely related to the physical structure of the object, and on the other hand is closely related to intentional human action.

According to the above line of thought, the notion of technical artefact is related to three key notions, namely the notion of a technical function, which itself is related to the notion of a physical structure and of a context of intentional human action (see figure 2).

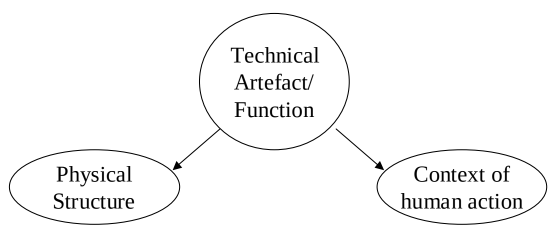

Figure 2 The dual nature of technical artefacts

There are some notable differences between our analysis of technical artefacts and Simon's. Simon's notion of goal or purpose has been replaced by the notion of function. This may seem an insignificant move, but it is an important one, because we may attribute functions to technical artefacts, but not goals (in the sense of an aim or an end (*telos*)). The notion of goal or end refers to a context of intentional human action;

within such a context a means used to achieve a goal (end, aim) is attributed a function. Thus, Simon's analysis implicitly refers to a context of human action by referring to goals and purposes. Furthermore, the notion of environment has been replaced by the notion of context of human action. It could be argued that this is also a minor change, because one form of environment is a context of human action. Simon's claim that the artefact has to be adapted to its environment then boils down to the, rather obvious, claim that the design of the artefact has to take into account the context of human action in which it is used. Nevertheless, this is a noticeable change, because it brings to the fore that not any kind of environment is relevant for the analysis of technical artefacts; only references to environments comprising a context of human action are appropriate. In his example of the sundial, for instance, Simon interprets the environment in a physical way (sunny climates are the required environment for sun dials). But this is problematic. It is not this physical environment that turns the object involved, a stick that casts a shadow on a surface, into an artefact of the type sundial. Only within a context of human action (e.g. of ordering events or comparing time intervals) this physical object acquires a function and becomes a technical artefact (a time-keeping device or clock).

Note that the above characterization of a technical artefact involves intentional processes in an essential way: without some context of human action (activity, processes) the notion of function loses its meaning, and what is left of a technical artefact without its function is merely some physical object. This interpretation of technical artefacts is confirmed by the observation that with most, if not all, technical artefacts is associated a (implicit or explicit) user manual. A user manual has at least two functions: it is a means to communicate the intended function to the user as well as a way to make this function accessible to the user by prescribing which actions have to be performed in order to realize the intended function. In other words, the user manual contains a use plan. The assumption that user manual and technical artefact are inseparable is not as innocent as it may seem. It leads to an important shift in our conception of a technical artefact, which is in line with the dual-nature interpretation. Instead of looking upon a technical artefact as a material object with a practical function, this assumption favours a view of a technical artefact as an object with a specific material structure embedded in a use plan and it is by virtue of its material structure *and* its use plan that it has a technical function.[6] In the same vein, engineering design may be considered not to be primarily about man-made material objects as such, but about devising use plans (activities) together with the material objects referred to in those use plans.

Against the background of the dual nature of technical artefacts we will now reconsider the issue about the boundaries of engineering systems.

4. Engineering systems as systems of technical artefacts

At first sight our analysis appears to have far reaching consequences for how the boundaries around (systems of) technical artefacts may be drawn. The idea that technical artefacts with their functions can be considered on their own, irrespective of a

[6] See Houkes et. al. (2002).

context in which agents play a role, has to be rejected. Functional properties are simply not intrinsic properties of technical artefacts. They are relational properties, which are on the one hand connected to properties of material structures, on the other to properties of agents. Without material structures there are no technical artefacts, but without agents there are also no technical functions. Man-made material objects, i.e. artefacts, become technical artefacts, that is, acquire a technical function, through their embedding in means-end relations by agents, that is, through their embedding in a context of human action.[7] This context of intentional action is constitutive for the object being a technical artefact.

The implications of this analysis for the status of social factors with regard to engineering systems have to be drawn with care. Intentional human action is often associated with social aspects, but intentional action may not necessarily presuppose a social context. In order to avoid misunderstanding let us begin with distinguishing the following two roles agents may play with regard to technical artefacts.

Certain technical artefacts, such as a car or an airplane, explicitly require for their functioning the presence of agents. These agents perform a function with regard to the artefact: anyone who wants to make use of a car or an airplane as a means of transportation needs a driver or a pilot.[8] So the car itself or the plane itself is an incomplete means of transportation. When designing a car or an airplane, engineers

[7] We leave open the issue whether natural, i.e. non-man-made objects are turned into technical artefacts by embedding them in means end relations.

[8] Note that when somebody uses a car to go from A to B and drives the car herself, then that person plays a double role with regard to the car: she is the user of the car, and at the same time the driver. As the driver, she is part of the transportation system and fulfils a (sub)function within that system, as the user she is not.

have to include agents in the system to be designed, since these agents are supposed to perform a function that is crucial for the functioning of the technical artefact. They also have to take into account the properties of these agents; it would be no use designing a car (airplane) nobody would be able to drive (fly). For this type of technical artefact our claim that the boundaries around the artefacts have to be drawn such that the technical artefacts include agents is not very remarkable: the inclusion of agents in the system to be designed is rather obvious. The agents involved may be considered to be subsystems, with specific sub-functions, of the whole system to be designed (although these subsystems are not designed but at most 'redesigned', so to speak, by being trained to operate the technical artefact).

Other technical artefacts, such as a television set or a lamp, do not need agents to perform their core function. Of course, these devices have to be switched on by humans, but once this is done, they continue to function without human assistance. When someone who is watching television leaves the room, the television set and the lamp lightening the room will continue to operate as if nothing had happened.[9] Clearly, agents perform an important role as users of these technical artefacts: someone has to switch on the television set or the light. But once this has been done, the technical artefact operates without the help of agents, i.e., without agents performing some sub-function such that the thing keeps on going. Our claim is that even this kind of technical artefact has a dual nature and that this dual nature implies that as a technical artefact it cannot be isolated conceptually from agents as users.

[9] It can be questioned whether a television set is still functioning when nobody is watching it. The television set is no longer embedded within an actual use plan (means end relations) and therefore strictly cannot be said to be still functioning, since nobody attributes any function to the artefact. So, some technical artefacts may be operating without functioning.

In order to clarify further the different roles of agents with regard to technical artefacts, it may be remarked that in the first type of technical artefacts, agents, in so far as they are part of the technical system and perform a specific function, may be eliminated, e.g. by the development of auto-drivers or auto-pilots. Given that we are dealing here with transport systems, it would be quite absurd to propose to replace the passengers with auto-passengers. It simply does not make sense to eliminate the role of agents as users with regard to either kind of technical artefacts.

So technical artefacts involve, by definition, agents, either as users or as designers or makers or operators or what have you. In practice, however, engineers proceed as if agents are not there. They treat functional properties as if they are properties of the objects themselves (when listing their products or components with their functions they do not bother to describe the agents who attribute these functions to the objects involved). This does not lead to any problems as long as users make ‘proper’ use of technical artefacts, that is, use technical artefacts according to the use plan intended by their designers (Houkes, 2005). In those cases, the functions attributed to technical artefacts by their designers and users are the same, and since the physical structure of the technical artefact was specifically designed to perform this function it is a small step to conclude that this function is a property of this physical structure itself. In daily engineering practice the assumption of proper use is taken for granted, which makes it easy to slip into an objectivistic language with regard to functions.

Problems arise, however, in cases of ‘improper’ use of technical artefacts. Then the intimate relationship between functions and agents comes to the fore. Suppose that somebody uses a screwdriver as a chisel to remove some wood, and when hitting the screwdriver with a hammer the shaft of the screwdriver breaks and the user gets injured. Suppose, furthermore, that the user files a complaint against the seller of the screwdriver, who keeps the maker/designer responsible. The obvious response of the designer is that the use of the screwdriver as a chisel is improper; the user has embedded the object in a wrong user plan, so he is himself to blame. The proper function of the screwdriver, he claims, is to drive screws and not to chisel wood; that is what the object has been designed and made for.

But how are we to decide what constitutes proper or improper use of a technical artefact or what the *proper* function of an object is? Can this be decided solely in terms of the physical structure of the object involved, or are the intentions of the designer or the user also relevant? The idea that the proper function of an object can be determined solely on the basis of its physical constitution, its physical capacities, appears highly problematic. Different capacities may form the basis for performing different functions. Take again our example of the screwdriver; suppose that it turns out that the screwdriver broke because of a fault in the shaft and that, if it would have been used as a screwdriver, the shaft would have broken all the same. Suppose, furthermore, that without this fault, the screwdriver could have been used without any trouble as a chisel for removing the wood. It seems that without a recourse to agents, in particular the intentions of users or designers, it will not be possible to determine which function (screwdriver or chisel) is the proper function of the artefact. It may well be that what is a proper function is partly a social (institutional) matter (in which conventions play an important role).

We may conclude that when engineering systems are considered to be systems of technical artefacts, agents may play two different roles with regard to engineering systems. Agents may be involved functionally in the sense that agents perform a subfunction, which is necessary for the functioning of the engineering system as a whole, given its current design. In this role, agents are not *conceptually* necessary, since in principle they can be eliminated by automation, involving a redesign of the artefact. The other role that agents play is that they are *conceptually* necessary for attributing functions to technical artefacts. In this role they do not perform a subfunction within the system. The consequences of this conceptual role for the status of contextual (social, legal, economic) factors with regard to engineering systems are limited. It does not imply that contextual factors are also constitutive of technical artefacts and that technical artefacts always have to be conceived and analysed in combination with social, legal, economic, etc. aspects. The agent does not appear in the form of a social, legal or economic agent, but only in the form of an intentional agent (who intentionally attributes a function to an object).

5. Engineering systems as socio-technical systems

Down at the lowest levels, engineering systems may be conceived as consisting of technical artefacts, which, although agents play a constitutive role for these technical artefacts, perform their function without the explicit intervention of agents. But higher up in the hierarchy agents play a much more prominent role, in the sense that they perform all kinds of (sub-)functions which are crucial for the functioning of the whole complex network of technical artefacts. In the case of civil aviation one can think of pilots, check-in personnel, air-traffic controllers, maintenance staff, etc. But more is involved than only agents who are using or operating technical artefacts. Agents are also involved in the sense that they produce and keep functioning social, economic, legal, etc. institutions that are necessary for the functioning of the whole system. Without education programs for pilots and institutions for licensing pilots, without banks, insurance companies, laws, regulations, etc. modern civil aviation systems could not function (or at least would not function as they currently do).

Thus, at the higher levels, engineering systems become much more heterogeneous, in the sense that, apart from technical infrastructure (hardware), agents and social infrastructure (software, institutions) become an integral part of the overall system too. Those elements become essential for the functioning of the system as a whole. This is why it seems more appropriate to call such systems socio-technical systems. One way to keep out all non-technical elements (agents, social, legal, institutional elements) from the domain of study of engineering systems would be to take into consideration only the hardware of such systems (that is, the technological artefacts proper) and to put all other elements into their context. This, however, does not seem to be a fruitful approach. Take again the civil aviation transport system and reduce this system to the system consisting of only the technological artefacts involved. It may be questioned whether the 'technological' civil aviation transport system defined in this way is a system at all. How are the relations to be conceived between the various technical artefacts that are supposed to glue together these artefacts into a system, a unified whole? What kinds of technical relations exist between such diverse technical artefacts as a luggage handling system, a runway, an air-traffic control system, an airplane, a passenger booking system, etc? It is not clear what binds these elements together.

Just as the technical infrastructure, the social infrastructure of socio-technical systems is man-made. The social institutions may be the result of intentional action or emerge unintentionally. In so far as they are the result of intentional action, the design and implementation practices of these institutions appear to be very different from engineering practice (lawyers and policy makers instead of engineers play a crucial role in these practices). The reason for this is rather obvious: the functioning of these social institutions is not intimately related to material structures, as it is the case for technical artefacts. Instead of knowledge of the material world, knowledge of the social world is primarily relevant for institutional design.

The idea of engineering systems as socio-technical systems is by no means new. It plays an important role in the work of Thomas Hughes on large technical systems (LTS) (Hughes, 1987). Hughes conceives of large technical systems, such as power networks, as seamless webs of social and technical elements. He distinguishes the following elements of such systems: physical artefacts, organisations, scientific components, legislative artefacts and natural resources (Hughes, 1987). Such systems are characterized by a common system goal and all elements of the system in one way or the other functionally contribute to that goal. While social and organisational elements are part of the system, contextual or environmental elements are not. In Hughes' conceptualisation, the environment of the system consists of those elements with which the system has no interaction, but with which nevertheless unilateral relations may exist. The system thus will depend on certain elements in the environment and will influence others.

Human actors have a special place in the system: "Inventors, industrial scientists, engineers, managers, financiers, and workers are components of but not artefacts in the system" (Hughes, 1987, p.54). They are – unlike the other elements of the system - free to act. Human actors have a number of roles with respect to, and within, technological systems. One role is that they often complete the feedback loop between system performance and system goal; they cybernetically correct system failures. Another role is that they are the inventors, designers and developers of the system. In this respect, one category of actors is especially important according to Hughes: the system builders, i.e. the people that consciously build and further develop large technical systems.

It is one thing to observe that engineering systems and large technical systems are socio-technical systems. It is another thing to spell out the concept of a socio-technical system in more detail in system-theoretic terms. What kinds of elements constitute socio-technical systems and what kinds of relations between those elements exist? In technology studies there is a tendency to treat social and technical elements of socio-technical systems as similar; they are seen as similar nodes in a seamless web or an actor-network (cfr. Hughes, 1987, Bijker & Law, 1992; Latour, 1993). From a philosophical point of view, such conceptualisations are unsatisfactory and hinder rather than further a useful conceptualisation and modelling of socio-technical systems, as we will explain in the final section.

6. Discussion: modelling socio-technical systems

Our starting point was that anything has to be included in the engineering system that is necessary for performing its function and that is open to design. The previous analysis has shown that this implies that agents and social, economic, legal, etc. institutions have to be included within the engineering system.[10] As far as agents are concerned, we have distinguished between two different roles. In the first place, they may have a particular sub-function to perform with regard to technical artefacts within the system. Concerning this sub-function they are in principle replaceable by other, presumably engineered hardware systems. In this role, agents should certainly be considered as parts of the overall engineering system. Second, agents are constitutive for any system as a technical artefact, in ascribing a function to it and in using it, or part of it, to achieve some goal. In this role, it makes little sense to consider them as elements of the system; they are not replaceable by engineered hardware. Here agents play a role that is conceptually necessary, but they do not perform certain specific sub-functions within the system necessary for the functioning of the system as a whole.

[10] It may be questioned whether agents as such are open to design; their behaviour, however, may be said to be open to redesign through, for example, training.

Apart from these two roles, agents may be involved with technical artefacts and engineering systems in still two other ways. Agents draw up and work out the specifications of a technical artefact or engineering system, either as designers or as those that specify norms for designers to work with. Again, in this role, it makes just as little sense to think of replacing agents by engineered hardware as in the previous case, and it seems therefore that they are not to be treated as part of the system. This is completely in line with traditional engineering practice in which agents in this role are, more or less by definition, considered not to be part of the technical artefacts under consideration. But a fourth role of agents with regard to engineering systems should be distinguished, and it is this role that marks a decisive break with traditional engineering practice. Agents within the system, who perform a sub-function, may change or redesign the system 'from within'. In other words, the (re)design of the system no longer takes place from a central point outside the system, as it is the case in traditional engineering, but becomes decentralized. This is an important aspect of socio-technical systems and should be taken into account. Thus in modelling socio-technical systems, certain agents have to be included within the system with a double role, as performers of sub-functions and as (re)designers of the system. Of course in practice, things are not so neat as this fourfold role model suggests. Agents will often be involved in several roles at once.

The behaviour of these agents is partly governed by rules and regulations, i.e. social institutions, and in so far these as rules and regulations are implemented specifically to guarantee the functioning of the system as a whole they may be taken to be part of the system. But these rules and regulations come in various sorts: rules that are either intended to govern the behaviour of agents directly or indirectly by being translated into hardware by designers, rules that direct the way users make use of the system, and rules that govern the functional behaviour of human subsystems. Thus, various kinds of social institutions may be part of the overall system.

On the basis of the foregoing, it is possible to distinguish categorically different elements – non-intentional and intentional elements, object-like and rule-like elements – within socio-technical systems, testifying to their truly hybrid character. This poses numerous challenges to the modelling of such systems. The modelling of the hardware components of socio-technical systems does not differ fundamentally from the way they are modelled in stand-alone artefacts, i.e. through law-like and functional relations governing their material components. For human actors, on the other hand, no unified approach exists. Roughly two major approaches can be distinguished. One describes agents as governed by behavioural dispositions, triggered by circumstances, contexts, roles, and so forth. The other sees agents as acting according to a single fundamental disposition; to act so as to choose at each moment from all available options the one that seems the best in terms of expected outcome. The former is a dominant approach in sociology and social psychology; the latter has a virtual monopoly in economics but has been gaining ground in other social sciences. Both for the purpose of explaining and understanding the functioning of socio-technical systems as for the purpose of designing such systems, ways to model the relations between the different elements of the system, notably the hardware components and the human agents in their various roles, have to be found.

Let us close with a critical note on our ‘functional approach’ to the demarcation problem of engineering systems. We have taken as our starting point that anything has to be included within the boundaries of an engineering system that is both open to design and necessary for performing its intended function. But this approach has its limitations. As long as the function (or functions) of the system under consideration is (are) well defined, this approach may work. Down at the bottom of the hierarchy of engineering systems (at the level of components or simple technical artefacts) this is usually the case, but higher up in the hierarchy this becomes questionable. What is the function of an airport or a civil aviation system? Does the idea of proper use of an airport or civil aviation system make sense at all? At the level of socio-technical systems many stakeholders may be involved, each of which has its own vision on the function of the system. Generally speaking, none of these stakeholders is in a position to impose its definition of the function of the system on all the others. So the idea that this kind of system can be designed, made and controlled from some central view of *the* function of the system has to be given up. This reinforces the point made above about the system being changed from within. Many actors *within* the socio-technical system are continuously changing (re-designing) the system. This, probably, is one of the most significant differences between traditional engineering design practices and the ‘design’ and development of socio-technical systems: socio-technical systems, in so far as the notion of (re)design is applicable to these systems at all, are not only (re)designed from a vantage point outside these systems, where their functions are defined once and for all, but also from the inside, which makes the idea of ‘total design control’ problematic.

References

Auyang, Sunny Y. (2004) *Engineering – An endless Frontier*. Cambridge (MA), Harvard University Press.

Bijker, W. and J. Law, Eds. (1992) *Shaping Technology/Building Society; Studies in Sociotechnical Change*. Cambridge (MA), MIT Press.

ESD-WP-2003-01.20, 'ESD Symposium Committee Overview: Engineering systems research and practice', MIT/ESD Working Paper Series ESD-WP-2003-01.20 *ESD Internal symposium*, 2002, p. 19 (*ESD Terms and Definitions (Version 13*) (http://esd.mit.edu/WPS/wps.html; accessed February 2005)

Houkes, W., P.E. Vermaas, K. Dorst and M.J. de Vries (2002) 'Design and use as plans: an action-theoretical account' *Design Studies*, 23, pp. 303-320.

Houkes, W. (2005) 'Knowledge of artefact functions' *Studies in History and Philosophy of Science Part A*, to appear in 2005.

Hughes, T. (1987) 'The Evolution of Large Technological Systems'. In: *The Social Construction of Technological Systems; New Directions in the Sociology and History of Technology*. W. Bijker, Th. Hughes and T. Pinch (eds.) Cambridge, MIT Press: 51-82.

Jackson, M.C. (1991) *Systems Methodology for the Management Sciences*, New York: Plenum Press.

Kroes, P.A. (2002) 'Design methodology and the nature of technical artifacts', *Design Studies*, 23, pp. 287-302.

Latour, B. (1993) *We Have Never Been Modern.* New York, Harvester Wheatsheaf.

Simon, H.A. (1996) *The sciences of the artificial*, Cambridge (MA), MIT Press, third edition.