Formalizations in dog cognition

Master thesis

J.N. de Vries
Formalizations in dog cognition

Master thesis

by

J.N. de Vries

to obtain the degree of Master of Science
at the Delft University of Technology
to be defended publicly on Friday July 12, 2019 at 10:00 AM.

Student number: 4357779
Project duration: November 12, 2018 – July 12, 2019
Thesis committee: Prof. dr. C.M. Jonker, TU Delft, supervisor
Prof. dr. T. Bosse, Radboud University Nijmegen
Dr. N. Yorke-Smith, TU Delft
Extra help from: Arjen van Alpen, de Roedel
Erica de Roode, Cattle Dog Coach
This thesis report shows the results of my graduation project on dog cognition. I choose this subject, since it combines my interest in animals with my computer science background. During this research, I learned a lot about dogs and how they think. Together with a logic-based checker I was able to show that there exist a correlation between paw departure and performance, as well as a correlation between performance and circumstances of a task. I created a new formalization for checking the route the dog takes. It showed that there exist a correlation between task circumstances and the route of the dog.

I want to thank my supervisor Catholijn Jonker for her guidance, advice and support during my study. Her feedback and support helped me through this with confidence and she taught me a lot about this subject. I would also like to thanks Arjen van Alpen and Erica de Roode for letting me attend the workshop on social learning and de roedelmethode. I learned a lot and got new insights in observing dog behaviour. I also want to thank them for their feedback during my project. Finally I want to thank Tibor Bosse and Neil Yorke-Smith for being in my committee and give feedback on my project.

I hope you enjoy reading this report.

J.N. de Vries
Delft, July 2019
# Contents

1 Introduction .................................. 1  
1.1 Motivation .................................. 1  
1.2 Structure of the thesis ....................... 1  

I Interobserver reliability ....................... 3  
2 Introduction to interobserver reliability ...... 5  
2.1 Motivation .................................. 5  
2.2 Research question and hypothesis .......... 5  
2.3 Method .................................... 6  

3 Methods for cognitive and social science ..... 7  
3.1 Observing behaviour .......................... 7  
3.1.1 Four questions for behaviour ............. 7  
3.1.2 Three levels of observation ................ 7  
3.1.3 Risk of interpretation ..................... 8  
3.1.4 Natural and experimental behaviour ....... 8  
3.2 Improving interobserver reliability .......... 8  
3.2.1 Observer ratings .......................... 8  
3.2.2 Citizen science ............................ 9  
3.2.3 Formal techniques ........................ 9  

4 Transcripts .................................. 11  
4.1 Formalization ................................ 11  
4.1.1 Ethogram ................................ 11  
4.1.2 XML format ............................... 12  
4.2 Creating transcripts .......................... 13  

5 Formal traces ................................ 15  
5.1 From transcript to trace ....................... 15  
5.1.1 Time .................................... 15  
5.1.2 Persistency rules .......................... 15  
5.2 Leads to trace ................................ 16  
5.3 TTL checker software ......................... 16  

6 Results comparison ............................. 19  
6.1 Hand made vs automatic performance ...... 19  
6.2 Differences ................................. 19  

II Dog cognition ................................ 21  
7 Introduction to dog cognition ................. 23  
7.1 Motivation .................................. 23  
7.2 Research question and hypothesis .......... 23  
7.3 Method .................................... 23  
7.3.1 Experiment C: Find the goody ............ 24  

8 Dog cognition ................................ 25  
8.1 Lateralization in humans ..................... 25  
8.2 Paw preference ............................... 25  
8.2.1 Methods ................................ 25  
8.2.2 Conflicting results ....................... 26  
8.2.3 Paw preference and behaviour .......... 27
# Experiment

## 9.1 Design of the experiment

- **9.1.1** Experiment A: Walking the dog.
- **9.1.2** Experiment C: Find the goody.
- **9.1.3** Experiment D: Alley of cones.
- **9.1.4** Experiment E: The final exercise

## 9.2 Aspects of performance

# Results paw departure and performance

## 10.1 Subjects

## 10.2 Analysis

# Dog Formalizations

## 11.1 Performance formalization

- **11.1.1** Definition
- **11.1.2** Semi formal notation
- **11.1.3** Formal notation

## 11.2 Route formalization

- **11.2.1** Definition
- **11.2.2** Semi formal notation
- **11.2.3** Formal notation

# Results

## 12.1 Subjects

## 12.2 Analysis

# Conclusion and future work

## 13.1 Conclusion

## 13.2 Future work

# Subject data

# Items needed for Experiments

# Parts of exercises

## C.1 Exercise A

## C.2 Exercise C

## C.3 Exercise D

## C.4 Exercise E

# Example route formalization

## D.1 Transcript

## D.2 Trace

## D.3 Checking

# Bibliography
1

Introduction

1.1. Motivation
While training people and their dogs, Arjen van Alpen observed that there is a meaning behind the paw use of dogs [13]. This and more observations led to the start of the Dog Cognition research led by Van Alpen and Prof. dr. Catholijn M. Jonker. The goal of this research is to get more insight in the cognitive abilities of dogs and get a better understanding of dog behaviour. The results can be used to improve the relation and communication between humans and dogs and improve the training of dogs.

Behavioural studies are vulnerable to observer biases. One set of observers can interpret the behaviours differently than another set of observers. The degree of agreement between different observers is called inter-observer reliability. This introduces differences in the observed behaviour and influence the reliability of the experiment. In his study for the relation between paw usage and performance, Arjen van Alpen also tried to tackle this problem. He did this by defining basic behaviours to be observed by humans and let the complex behaviours be observed by a computer [39]. Using this technique, he was able to improve observer reliability. In the study of the correlation between paw usage and performance, he found that in a search task for a goody (i.e., treat) hidden under a cone, dogs that depart with their left paw were finding the goody significantly more often than dogs that depart with their right paw [38]. These studies are the basis for this thesis.

1.2. Structure of the thesis
This thesis is split into 3 parts. The first part focuses on interobserver reliability I. It studies what is needed to set up a behavioural study as objectively as possible and also uses a formal checking technique to study more complex behaviours. In this part the research question: Is using an automated checking software improving the observer reliability compared to humans annotating complex behaviour? will be answered. The second part studies dog cognition. It dives into the research of [38] and expands upon this research. This part is used to answer the research question: Is the correlation between departing paw and performance influenced by circumstances? After this both parts will be combined to study dog behaviour using the formal checking tool described in part III. The research question answered in this part is: Is there a correlation between the circumstances of the task and the route the dog takes to the goody?
Interobserver reliability
Introduction to interobserver reliability

2.1. Motivation

In research within social and behavioural sciences, low interobserver reliability and reproducibility are well known problems. Several studies have tried tackling these problem. Interobserver reliability measures the agreement between subjects observing the same phenomenon or object [12]. Having a low interobserver reliability means that the research is hard to reproduce and makes it hard to judge the validity of the research. There are measure to indicate how good the inter observer reliability is. However a high inter observer reliability is not the same as a high accuracy. Accuracy means that the observed behaviour is the same as defined beforehand, while interobserver reliability only indicates the degree of agreement [17]. For example an observer can be accurate by observing the behaviour as defined by the pre-established standard, while they have a low interobserver reliability with another observer whose observations are not accurate. Or an observer could observe inaccurately and have a high interobserver reliability with other inaccurate observers. In research one wants to have both, high accuracy and high interobserver reliability. [17] says that the complexity of the observed behaviour influence the accuracy of the observations. The more complex the behaviour, the more it tends to be misinterpreted. So making the standard behaviours less complex will improve the accuracy. If behaviours are less prone to misinterpretation, the interobserver reliability will improve too. [39] proposed to use these ideas in his research. By using an Artificial intelligence technique. They defined basic behaviours for humans to observe, while letting the complex behaviours be decided by a computer. First they identify which behaviours are basic and which are complex. Next they split the complex behaviours into basic behaviours. These basic behaviours are formalized into basic scoring units. Then the experiment can be conducted and the results can be gathered. Using a computer tool, the complex behaviours can be checked. This thesis will use this technique in combination with the paw-preference research described above.

2.2. Research question and hypothesis

This part of the thesis tries to answer the following research question: Is using an automated checking software improving the observer reliability compared to humans annotating complex behaviour? ’

It is expected that the complex behaviours defined using the checking tool from the Leadsto software called the TTL checker would overall be the same as the hand written complex behaviours. The TTL checker is explained in Chapter 5. The transcripts are written by the same observers that observed which paw was used and how the performance was. So the annotations should describe the same behaviour. However a difference might occur when the observer might make a mistake in noting down the behaviour. For example writing the opposite of what the observer noted down in the complex behaviour. Furthermore it is expected that the TTL checker is more more accurate in checking the behaviour than human observers, since it is created to increase the interobserver reliability. The software checks the behaviour following simple rules, while observers might interpret the rule differently. The software will always follow the same rule, while one observer might interpret it a bit different than another observer checking the dog.
2.3. Method
The formal technique described in [39] will be used to get results based on the search task described above. The software used for this step is the Leadsto software [40]. From the experiments transcripts are made in XML format. Using a parser, we can enter this data in the Leadsto software. It is possible to used the traces created by Leadsto to check the more complicated behaviours. Using this technique we can answer the research question. Does the observed complex behaviour as annotated by humans match with what they noted down in the transcripts containing only the basic behaviour? There might be differences in the results of the automatically checked transcripts and what was done by hand. The differences in results might be due to biases. In this thesis we look into the differences and why these differences might occur. It will check this with the discussions between observers creating the transcripts.

First, Chapter 3 gives background information on interobserver reliability. Next Chapter 4 explains what is needed to set up a research correctly, it shows how transcripts are made based on an ethogram. Chapter 5 shows what input is used in the TTL checker that automatically checks complex behaviour. It also shows how the checker works. Finally, Chapter 6 compares the handwritten complex behaviours with the automatically checked behaviours and explains what causes the differences.
Methods for cognitive and social science

Cognitive and social science is an interdisciplinary study of mind and intelligence [34]. However it is a challenge to do research in this field, since it depends on observable behaviour. Interobserver reliability is important in this kind of research. Interobserver reliability measures the agreement between subjects observing the same phenomenon or object [12]. Having a low interobserver reliability means that the research is hard to reproduce and makes it hard to judge the validity of the research. This chapter discusses the difficulties of observing behaviour and some research that tried to improve interobserver reliability.

3.1. Observing behaviour
Observing behaviour is difficult for people, since they always consider things from a personal and human point of view. When observing other people, it is already hard to avoid this bias. Humans have their own experiences and emotions and this influences the way they look at other people. However when observing animals, this is even more of a problem. Humans tend to project their emotions, concerns, their human way of observing the world and their human way of reasoning about situations onto the animal [22]. The subjective experiences of an animal may be totally different from humans. This means that human interpretation of what is observed should not be based on extrapolations from humans but also on a good knowledge on the animals natural history and behaviour [22]. So in order to analyze, understand and explain non-human behaviour, one must remove the human need to interpret as much as possible. For example, when a person needs to write down characteristics of a situation, it is better to do this following a given set of measurements and items instead of letting the person simply write everything down. Instead of letting the person describe the situation, let the person write down the amount of wind, the terrain etc. this makes it easier to reproduce the experiment, but also it gives the observer a better understanding of what should be noted down. Furthermore, finding ways to measure instead of guess, excludes some biases.

3.1.1. Four questions for behaviour
In ethological studies Tinbergen’s four questions are used [1, 3, 35]. Instead of just asking why an animal behaves a certain way, one can split this question in four different questions. The first is, “what made it do it now?”, so the immediate causal control of the behaviour, the second is about the development and learning of the animal that led the animal behave in this way. The third question is “What use is this to the animal?” So what is the survival value for the animal. The final question is about the evolulional origins of the animal [15]. Using these question during observation can give insight in the origin and the reason behind the behaviour. Taking into account the evolulional line and learning process of the species will give more explanation than only focusing on the current behaviour. These questions can be used to set up an experiment to observe certain behaviour. For example, what is going to make a dog act at the moment needed during the experiment? In order to make a dog act when you want it to, a piece of food can be used to motivate the dog to start moving.

3.1.2. Three levels of observation
[18] claims it is possible to stay objective when taking into account the three levels of observation: The individual level, the educational level and the environmental level [4, 9]. The individual level indicated the
3. Methods for cognitive and social science

genetic makeup and development of the observed individual. Each individual has traits that belong to the species. For example German shepherds have a different appearance than a poodle. While other traits are from the individual dog and not specific to the species. One German shepherd might have a different height, weight or voice than another German shepherd.

On the educational level one looks at what the gained knowledge and skills are that individuals have learned during their life. These skills can be learned during training situations, for example training a dog to sit on command, through personal adaption needed for the dogs surroundings, for example learning how to open a cage when being locked up in it. Or skills that were learned by interacting with others. It is not possible to monitor everything that happens to the individual, however the more one knows about its background, the more one can understand its behaviour.

The final level is the environment level. This is about what the individual learns while adapting to the environment. Different social environments can create different behaviours. For example a dog will behave differently in a small closed area than when it is outside with no limitations. Taking these levels into account while observing behaviour, would not solve the bias of the human point of view. However it can help with explaining differences in behaviour.

3.1.3. Risk of interpretation
When people observe behaviour they already start interpreting this behaviour instead of just noting it down [22]. Instead of having annotations of interpreted behaviour, one wants to have the actual behaviour. In order to help with this, the observable behaviour should be defined in terms of observable facts, not interpretations. Therefore it is important to define a way to obtain and describe the observable behaviour beforehand. So define behaviour using measures. For example instead of writing down if the dogs is walking in the correct direction, it is useful to just write down the direction using the hours of a clock with respect to the experimental setup and define correct as walking in direction 12. Since what is the correct direction? Is it any direction that gets the dog to the final point or is it the fastest way it gets there? So defining walking without labels as good direction and bad direction and instead using the direction of a clock will reduce this bias. Using these direction of the clock we can then define which directions are allowed in the correct direction and which are not.

3.1.4. Natural and experimental behaviour
Doing experiments brings another issue with observing behaviour. Natural behaviour is different from experimental behaviour. During a controlled experiment the dog will adjust its behaviour and not show its natural behaviour like when the dog would be in its own known surroundings. This should be taken into account during the observation. Likewise observing natural behaviour in the wild might be difficult, since the observer can have an influence on the behaviour. The benefits of a controlled experiment however, is that the situation is controlled and thus can be replicated. So designing the experiment should be done with utmost care.

3.2. Improving interobserver reliability
As shown before, within the research of social and behavioural sciences, interobserver reliability and reproducibility are well known problems. The same holds for the dog cognition sciences used in this research. Several studies have tried tackling this problem.

3.2.1. Observer ratings
A review from Meagher [25] showed that observer ratings can be both be reliable and valid. With reliability defined as interobserver and intra-observer reliability. The agreement between multiple people independently rating the same individual and the agreement between ratings by the same individual on multiple occasions. Valid is defined as he accuracy in predicting scores on a criterion measure. In her paper she presents evidence from zoo, laboratory and farm animal studies that demonstrate that observer ratings can be both reliable and valid. She admits that biases are still a risk. Especially when these ratings can reflect the observer's own care of the animals. However this risk can be reduced by carefully phrasing the questions to be answered. Furthermore the experiment should be set up with care. For example when asking about punishment, a statement such as "it is not uncommon for people to correct their dog in some way" would be put before the question [44]. Regardless, when conducting such an experiment it is important to keep the risks in mind and be careful when setting up the experiment.
3.2. Improving interobserver reliability

3.2.2. Citizen science
Another approach to improve observer reliability is suggested by [32]. They propose citizen science as a new tool in Dog cognition research. They studied the biases introduced by letting dog owners do the experiments themselves at home instead of performing them with an experimenter in some study facility. First they did a pilot run, where the citizens did the experiments at home, while an experimenter was present. They showed the participants a video with the steps to take for each experiment and after each experiment a questionnaire was given to the participants. [32] found that the citizen scientist closely gave the same results as obtained using conventional laboratory approaches. They however found that citizens with more knowledge about the field already were biased during the experiment. This approach can be useful to gather more participants and do experiments in a shorter time, but it is important to keep in mind the biases this will give.

3.2.3. Formal techniques
The following approach tries to increase the reliability of observers by defining the behaviour that needs to be observed as simply as possible. [39] provided a way to reduce the bias by using Artificial intelligence techniques. By defining more complicated behaviour in basic score units, the chance of misinterpreting the behaviours is reduced. These basic behaviours are used as input in a formal checking tool that is able to check for the more complicated behaviours. First they identify which behaviours are basic and which are complex. Next they split the complex behaviours into basic behaviours. These basic behaviours will be formalized into basic scoring units. Now the experiment can be conducted and the results can be gathered. Using a computer tool, the complex behaviours can be checked. This method can help reduce the bias, but a good setup is needed. This is done by starting with defining the behaviours in a clear and objective way. So the annotators know how to judge behaviour. For example walking in a straight line could be defined as walking towards 12 when taking a clock as a way of measuring the direction the dog is taking. Besides a clear definition of the behaviour, the experiment itself should also be clear. It should be explained what each person needs to do, what is expected of the subject and what to do when a subject does not behave as expected. This research will use this approach to analyze the behaviour and answer the research questions.
4

Transcripts

Having the video of the experiments, is essential in transcribing the data. Having this visual data, together with the transcript makes it possible to analyze the experiment and behaviour afterwards. This chapter will describe the method used to describe the behaviour performed on the video.

4.1. Formalization

Transcripts are created by going over the video frame by frame and annotating all the behaviour that is visible on the video and defined in an ethogram. Annotating all behaviour compared to annotating only the behaviour needed for one hypothesis can help in the future to test other hypotheses without having to redo the whole experiment and to train computer vision programs to recognize behaviour. When there is no computer vision program to automatically annotate the behaviours in the ethogram, this needs to be done by hand. Before starting to annotate everything happening during the experiments, some vocabulary needs to be made and it should be decided in what format the transcript should be in. This format depends on what needs to be done with the annotations. Some tools for analyzing the data might require a certain format to be readable by the computer.

4.1.1. Ethogram

Before starting to annotate behaviour, the behaviours need to be defined. Ethograms are a way to do this. Ethograms are common in behavioural science [43]. An ethogram describes what kind of behaviours should be annotated and in what way. Although multiple ethograms for dog behaviour exist [24], not every ethogram can be used for any research. So creating a specific ethogram for ones research is needed. Once an ethogram has been created, it needs to be checked in order to make sure the behaviours are written down correctly. To do so, a pilot study needs to be done. During this pilot, one checks if behaviours are clear and unambiguous. The ethogram used in this thesis is a dog ethogram. It describes short unitary movements, like ‘dog licks nose’ or ‘dog moves right paw’. It also covers behaviours that take longer, for example ‘dog sits’ or ‘dog walks towards’. Besides the behaviours of the dog, the characteristics of the dog is annotated. For example ‘training level’. Next to the behaviour of the subject itself, behaviours from the people involved and other influences of the environment are also defined. Examples for this ethogram are ‘handler commands’ or ‘disturbing smell’. An ethogram describes first the title of the behaviour, then gives the format on how to write it down in the transcript and finally gives some extra information about the arguments. This extra information is also used to define the behaviour. An example from the ethogram used in this thesis is ear orientation. The definition is shown in figure 4.1. It shows first the name, then in blue two examples on how to use it with different arguments. After this the definition of the behaviour is given. Next the arguments are given with the possibilities and an explanation. As shown next to the title of the behaviour, there is a not. This note is added to show that this behaviour is annotated since 6/12/2018, because this behaviour is added later during the research. Having a clear and extended ethogram helps annotators be as precise as possible during annotation and improves observer reliability since definitions will be clear. This ethogram was already created before this thesis starts, but is always evolving. During annotation some behaviours might be unclear to a new annotator or missing. The ethogram will then be adjusted with more information or a new behaviour will be added. If a new behaviour is added, it is shown in the ethogram when this was first used, since some transcript might be
missing this information. So other researchers then know which transcripts to use or that they need to extend the transcript. Changing a behaviour is not done as that would require changing all transcripts already done. By only allowing additions to the ethogram (and improved explanations of the ethogram), the validity of old results is not affected. New behaviours are only added if approved by the responsible research staff.

**Figure 4.1:** Example from the ethogram.

```xml
<atomic arg1="left" arg2="towards_handler">dog_ear_orientation</atomic>
<atomic arg1="right" arg2="12">dog_ear_orientation</atomic>

{definition:}
ed_orientation: The dog might have its ears orientated in a certain direction. For example towards the handler.
argument1: left, right
argument2: place
(optional) arguments: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12

{definitions:}
place:
individual: handler, dog, experimenter, cameraman1, cameraman2, towards_experimenter,
towards_handler, away_from_handler, up_to_handler, at_handlers_hand,
at_experimenter_hand, at_handlers_leg, at_experimenter_leg, handlers_face,
handlers_hand, experimenters_hand.
object: at_surroundings, at_ground, down, alley, ground, mat, puddle, goody, barrel, thing,
handler_goody, wooden_board, flag, pylon
direction: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12
pylon: middle_pylon, left_pylon, right_pylon, goody_pylon, flag, start_pylon, starting_point
handler_goody: This expression indicates goodies offered by the handler to the dog. Handlers usually use
pylon to get the dog to sit.
flag: A flag which indicates the start position. This flag is sometimes exchanged for a pylon. The words
flag, start_pylon and starting_point are all used to indicate the start position. Starting_point is preferred.
start_pylon: A pylon which indicates the start position. This pylon is sometimes exchanged for a flag. The
words flag, start_pylon and starting_point are all used to indicate the start position. Starting_point is
preferred.
thing: a not specified object. The specification must be added in a comment: <comment>...</comment>.
arguments: The numbering can be used in combination with ‘pylon’, to indicate a pylon in the alley. See
appendix III.)
```

**4.1.2. XML format**

It is important to choose a format for annotating behaviours. The format needs to fit the needs of analyzing the data. If the data needs to be analyzed by a computer, the format needs to be readable by a computer. As can be seen in figure 4.1, the transcripts used in this research are written in XML format. This format was chosen, since it allows creating tags at will. Besides this, it is also a machine accessible format, which can later be converted into other formats, such as the Leadsto software [40] (see chapter 5). Next the structure of the XML file is given: a transcript consist of three elements, data, general circumstances and exercises. Data is the data about the dog, it contains information like name, age, breed and training level. An example is given in figure 4.2.
4.2. Creating transcripts

General circumstances describe the circumstances of the experiments like date, weather, area and surface. Figure 4.3 shows an example.

```
<general_circumstances>
  <date year="2004" month="1" day="25"/>
  <area>Open Field</area>
  <surface>sand</surface>
  <weather>
    <temperature>between 0 and 10</temperature>
    <wind_force>none</wind_force>
    <wind_direction>120</wind_direction>
    <sun>Shining</sun>
    <precipitation>none</precipitation>
  </weather>
</general_circumstances>
```

Figure 4.3: Example general circumstances part transcript

Finally the exercises describe the experiments conducted during this research. It consist of multiple exercise elements and these consist of time, comment and then a sequential element. The sequential exists of multiple parallel and or atomic behaviours. Atomic behaviours are the behaviours described in the ethogram. The exercises described are all the experiments related to the dog cognition research. Experiment A to E are described in Chapter 9. An example is shown in figure 4.4. As shown in the figure, it has an index VI. This indicates the part of the exercise, for example sitting, taking off or searching. A complete overview of each section per exercise is given in appendix C.

4.2. Creating transcripts

New annotators need to first study the ethogram and the XML format, before they can start transcribing the data. It is difficult to transcribe video data in plain text while keeping the time information. Stepping through the video frame by frame is helpful for determining the timing of the behaviours. Whilst transcribing, a person needs to be precise on what behaviour happens in parallel and what in sequence. Annotators can nest sequential and parallel behaviours. To help the annotators check their work, for each part of the experiment a timing is given for the point in the video where the data was annotated. See figure 4.4 where a time block containing the start and end time of this exercise for each camera is shown. These times are the exact times in the video.
Figure 4.4: Example exercise part transcript.
5

Formal traces

The transcripts are written in a machine readable way, so a computer can decide on more complex behaviours. This research uses a TTL checker. This is a formal technique to automatically check complex behaviours. In order to get the transcripts through the TTL checker, the transcripts are converted into Leadsto traces [40]. Leadsto is developed by the AI department of the VU Amsterdam. It consists of Leadsto and a TTL checker. The Leadsto traces can be used as input for the TTL checker, to check for formalized properties. Besides this, Leadsto offers a way of visualizing the traces. This helps understand the timeline of the exercise.

5.1. From transcript to trace
A parser is used to translate the transcripts into traces. The parser is a Prolog program that takes the XML file as input, parses it and gives a Leadsto trace as output. The parser works with steps.

1. The first step is to parse all the exercises. The information of the exercises is noted down. So the starting time, the end time and the comments for each exercise are translated.

2. Then it starts parsing the exercise content. The content exists of parallel, sequential and atomic items. For parallel items, it makes sure the next action has the same starting time as the current action. If it was a sequential item, then the parser makes sure the next action starts after the current action is finished. The atomic actions are parsed together with the related arguments.

3. After this the start and end points are fixed and the actions are checked to make sure the persistency rules hold. These persistency rules are explained in section 5.1.2.

4. When all actions are checked, the actions are written on separate lines. This is needed for the Leadsto trace file. Also the trace end time is noted down. The trace end time is based on the end time of the last exercise.

5. Finally the dog data and general circumstances information is parsed and added to the file.

5.1.1. Time
In the traces, no real time is used. The parser only shows which behaviour happens before or at the same time as another behaviour. It does not show how long it takes. However, due to step 1, there is a reference to the start and end time of an exercise. Figure 5.1 shows a section of an exercise, has the timing information for each camera. It shows times_on_camera('start min', 'start sec', 'end min', 'end sec', 'camera'). These times refer to the start and end time when this exercise is visible on the video belonging to that camera. For figure 5.1, this is that exercise c trial 2 part vi is visible on camera I from 1:54 to 1:65 and on camera II from 3:30 to 4:18. Preserving this information helps with matching the trace with the transcript and the video.

5.1.2. Persistency rules
Some behaviours cannot happen at the same time and some behaviour can mean the end of another behaviour. For example, ‘dog moves to 12’ and ‘dog moves to 6’ cannot happen at the same time. If ‘dog moves to 12’ holds from 10 till 20 and ‘dog moves to 6’ holds from 15 to 30, the behaviours overlap. Persistency rules
are created to keep such overlapping behaviours out of the trace. The rules define actions that terminate other actions. If the rules do not hold, the parser will fail and thus, one needs to solve this issue.

One has to think carefully when creating these rules. One has to consider all possible states and actions that can cause the termination of each other. Careful discussions have led to the rules used in this research. During the discussions it was made sure everyone gave their opinion and they all agreed on the rule. It is useful to have the persistency rules in mind while creating the transcripts, so no important actions are missed during annotating. However annotating is done by humans and humans make mistakes. So it is important that the rules are still checked in the parser.

A rule is formalized in the parser and looks as follows: exclusive(action 1, X). X can be a list of actions that terminates action 1 as soon as it occurs. X can also be 'all' which means that any other action following action 1 terminates action 1. Finally X can be a number. The number indicated how many time units action 1 holds. If no rule is given, then the action persist as long as the given exercise persist. table 5.1 shows a few examples of these persistency rules.

<table>
<thead>
<tr>
<th>Persistency rule</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>exclusive(action(dog_ear_stance, any), [action(dog_ear_stance, any),</td>
<td>the action ‘dog_moves_ear’ with any parameter is terminated by the</td>
</tr>
<tr>
<td>action(dog_moves_ear, any), action(dog_general_posture, any))]</td>
<td>occurrence of the action ‘dog_ear_stance’, ‘dog_moves_ear’ or</td>
</tr>
<tr>
<td></td>
<td>‘dog_general_posture’ with any parameter</td>
</tr>
<tr>
<td>exclusive(action(dog_moves_paw, any), all).</td>
<td>the action dog_moves_paw with any parameter is terminated by the</td>
</tr>
<tr>
<td></td>
<td>occurrence of any other action</td>
</tr>
<tr>
<td>exclusive(action(handler_hand_food_pose, any), 2).</td>
<td>the action ‘handler_hand_food_pose’ last for 2 time units</td>
</tr>
</tbody>
</table>

Table 5.1: Some examples of persistency rules used in the parser.

5.2. Leadsto trace

After parsing the transcript, a Leadsto trace is produced. A trace is a sequence of states and actions using time points. The Leadsto software can visualize these traces in a sort of time line. On the horizontal axes there is the time line with time points. As said before, these are no real time points. On the vertical axes, the actions are shown. When an action hold, the time points are marked blue. An example is given in figure 5.2. In this example, the action ‘dog_moves_to(11,trot)’ holds from time point 0 to 20. Then the dog sniff at the middle cone from 20 to 25. This is shown using the blue rectangles on the time line.

5.3. TTL checker software

By translating the transcripts into Leadsto trace format one can get a better understanding of the results. Besides this, the Leadsto traces can be used as input for the TTL checker software [40]. This software is used...
to automatically check for more complex behaviours in these traces. In order to check for more complex behaviours, formalizations need to be made. The TTL software offers a graphical interface to enter the formalizations. A property name needs to be defined. Then one uses logic to make the formalization. For example in figure 5.3 shows the definition for a straight route. Based on a given exercise, it determines if the dog starts straight between certain timepoints. More information about this formalization is given in section 11.2. Once a formalization is created, one can use it to check. In order to do so, one has to load the traces into the TTL software and perform the check. Checking is done by choosing the `check property verbose` or `check property quiet` commands. The checker will compile the formula. Then the checker loads the traces and checks interval by interval whether the given property holds. If this is the case, the formula is satisfied. If no interval can satisfy the property, the result will be not satisfied. The result is shown as an output in the interface.

Figure 5.2: Example trace for exercise c trial 3
Figure 5.3: example visualization straight route formalization
6 Results comparison

This chapter shows the results of the comparison between humanly-annotated complex behaviour and automatic checking. The comparison is based on the dog cognition data. More information about the dog cognition research is given in part II.

6.1. Hand made vs automatic performance
For this comparison a subset of 32 dogs were used. This subset has an even distribution according to age, gender and training level. But more important, all these dog have a transcript that can be transformed in a trace needed for the TTL checker. Appendix A table A.1 summarizes the data of the selected dogs. It shows their age in months during the experiment.

The formalization used for this comparison is given in chapter 11.1, it was determined if the dog started with left or right and if the performance was good or bad. The exact definition of these behaviours is not needed for this comparison, but the definition can be found in chapter 11.1. The results of the automatically checked behaviours are compared with the humanly-annotated complex behaviours.

6.2. Differences
For each dog a total of 7 results were gathered. For experiment C1, C2, C3, D1, D2, D3 and E1. C1 stands for experiment C trial 1, C2 stands for experiment C trial 2 and so on. More details about the experiment are given in Chapter 9. In total there are 224 results for the TTL checker and 224 results are humanly-annotated. From these 224 results, there were 29 incorrect matches. This is 13%. Table 6.1 shows the kind of differences that occurred between handwritten and automated checks.

<table>
<thead>
<tr>
<th>Difference</th>
<th>Number Mismatch</th>
</tr>
</thead>
<tbody>
<tr>
<td>Left good Left bad</td>
<td>9</td>
</tr>
<tr>
<td>Right good Right bad</td>
<td>3</td>
</tr>
<tr>
<td>Good left Good right</td>
<td>5</td>
</tr>
<tr>
<td>Bad left Bad right</td>
<td>9</td>
</tr>
<tr>
<td>Left good Right bad</td>
<td>3</td>
</tr>
<tr>
<td>Left bad Right good</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 6.1: Number of mismatch per difference

These differences can occur for several reasons. Using the discussion for each dog and the belonging transcripts, it was possible to determine why there was a difference. 6 of the differences had a mismatch, because in the transcript the first paw used was not the first paw used according to the handwritten check. 4 of these mismatches showed misbehaving in the transcript, while the annotator did not think it was misbehaving. 2 of the mismatches showed both problems. This shows that annotators can make mistakes in checking. But also are not consistent while determining misbehaving. The other 17 difference where cause by discussions on which paw was used and if the dog misbehaved or not. 8 discussions where about which paw was used. This discussion was caused by which speed the dog was going. While in the end of the discussion, the annotators
agreed, the transcript was not adjusted to the right gate and thus a different paw was used. 8 of the discussions were about misbehaving or not. These discussed if going to the experimenter was correct or not. In the definition it is defined that going straight to the experimenter is considered correct. So solving this with the TTL checker would have solved this discussion. The last 2 differences had discussions about both the performance and the paw used. These discussions show that determining the speed of the dog is important for knowing which paw is used. But if annotators disagree, the transcript should be adjusted once there is an agreement. The TTL checker shows its profit when one looks at the performance. As shown above, annotators have problems with being consistent even when the definition is clear. The TTL checker will always use the same definition no matter which dog is being checked.
Dog cognition
Introduction to dog cognition

7.1. Motivation
Since the discovery of lateralized behaviour in non-human species [29], multiple studies have investigated lateralisation in dogs and other animals. Studies show conflicting results with the relation between paw preference and sex. So instead, it is more interesting to study the meaning behind paw uses. Besides checking the distribution of dogs with a paw preference and dogs without a paw preference, research also focuses on the behavioural differences between these dogs. In humans it is shown that the left hemisphere is used when we are in a new situation and we have to solve an unfamiliar problem, while the right hemisphere is active when we are familiar and confident [8]. So studying the behavioural differences compared to the paw uses, can give more insight into the meaning of the paw use. It also gives more insight in the lateralization of the dog. It might work the same as with humans. A study done by van Alphen et al. [38] studied the correlation between paw usage and task performance. In a search task for a goody hidden under a cone, they found that dogs that depart with their left paw were finding the goody significantly more often than dogs that depart with their right paw. They also found that there was no gender differentiation in their data. Previous work [2] and this research both suggest that paw-preference might be task specific and correlated with performance [38]. From this point forward, the term paw usage or paw departure is used instead of paw preference, since it is not really a preference if it is task specific. This thesis will investigate if the same correlation between paw usage and performance consists in tasks with different circumstances.

7.2. Research question and hypothesis
In the second part of this thesis, we want to answer the question: Is the correlation between departing paw and performance influenced by circumstances? It will compare the results with the results of [38]. A different result could be due to the cones either guiding or distracting the dog.

It is expected to have a similar results as shown in [38]. We expect to see a correlation between paw usage and task performance. So that dogs that depart with the left paw will significantly perform better. Performing better means, more dogs perform well compared to the number of dogs that perform poorly. Besides this, the overall performance will provide more insight into the way dogs react to their environment. If they see the alley as a guide, then the performance should be better than in the previous task. The experiments are explained in Chapter 9, but for the setup of the experiments, see figures 7.1, 7.2 and 7.3. It is expected that the performance in experiment E is better than in experiment C and D, due to the lack of pressure. The performance in experiment D will be worse, due to more pressure on the dog by adding more cones. However dogs could perform better in D compared to C if the alley is guiding the dog to the goody.

7.3. Method
To answer the research question, we will apply the same approach from [38]. Instead of only focusing on experiment C, we also look into experiment D and E. To see if the correlation persists and if the circumstances causes any differences. We will look into the search task with an alley of cones. See figure 7.2 for the layout of the experiment. The task used in the paper from van Alpen et al. is also a search task. It has three cones next to each other and one of them contains a goody. The dog need to find this goody. For the layout see figure 7.1.
The difference between the exercise used in this thesis and the one used in [38] is that for this thesis there is an alley of cones (12 pieces of orange cones 45 x 20cm) evenly distributed over 5 meters from the starting point of the handler and the dog to the position of the middle cup in the front. The width of the alley is the size of the dog + 40 cm. The experiment goes as follows: The dog is sitting and it is blindfolded. The experimenter lets the dog sniff the goody and puts the goody under cup1. The blindfold is removed and the dog is allowed to go find the goody. The owner is instructed to stand still and not walk with the dog. After the dog has found the goody, the experiment is finished. This is repeated 3 times.

7.3.1. Experiment C: Find the goody
The goal of this experiment is to test the hypotheses that paw preference correlates to performance in a simple search task.

To help answer this question, Chapter 8 gives more information on the current research in dog cognition. Then the experiments used for this research are described in more detail in Chapter 9. Finally, Chapter 10 shows the analysis of paw departure and performance. It shows if the correlation between paw departure and performance persists and what the influence of adding an alley of cones is.
This chapter will provide more information about the current research in paw preference and lateralisation in dogs.

8.1. Lateralization in humans
It has been shown that the left-anterior region of the brain is associated with anger and fear, which leads to higher levels of activity that are clearly focused and directed [11]. It is also shown that the left hemisphere is more involved in the expression of facial emotions. In general, the right-hemisphere of the brain is focused on experience, while the left hemisphere is used when there is a lack of experience [8]. Motivation plays an important role when performing a task. It has been shown that higher motivation leads to a better performance [28]. Not only is brain lateralization connected to motivation, it is also connected with emotion and behaviour. It is observed that brain-damaged patients with left-hemisphere lesions where more likely to show catastrophic behaviour such as crying, guilt, complaints, pessimistic statements and worries about the future, while patients with right-hemisphere lesions showed more indifferent reactions [8]. Learning a new skill, being in a novel situation will activate the left hemisphere of the brain. Once the skill is learned or the novel situation becomes more familiar, the activity in the left brain will decrease, while the right hemisphere will increase in action [31]. So during a novel situation, the left hemisphere is more active and when we know what to do, the right hemisphere is more active. Research suggest that to a large degree, one side of the body is linked to the opposite side of the hemisphere of the brain. So the left-hemisphere controls the right side of the body and the left hemisphere controls the left side of the body [20]. Combining this information suggests that when there is a positive motivation and the subject is confident on what to do, it is likely that a movement is done with the left part of its body. While when there is a negative motivation and the situation is new and uncertain, the subject is likely to move with the right part of its body.

8.2. Paw preference
Since the discovery of lateralized behaviour in non-human species (see for example [29, 37]), multiple researches have investigated lateralisation in dogs and other animals.

8.2.1. Methods
In order to test the paw preference, there are multiple tests. One well, known and often used method is the Kong™ test. Another way of testing lateralization in dogs is called the first-stepping test.

Kong™ test
The Kong™ test is done using a Kong™. This is a hollow cylindrical rubber dog toy, See figure 8.1 for example. It was first developed by Branson et al. [5] to determine paw preference in dogs. During the test, the toy is filled with food. Sometimes they put some moist food and than froze that so that it is difficult for the dog to get the food out [41]. First the dog is allowed to sniff the toy. Then the toy is placed in front of the dog. Now the paw uses are counted for left and right. A paw use is when a dog has one or both paw on the toy, regardless of the duration. When the paw is removed and again put on the toy, it is a new paw use. When both paws
are used to stabilize the Kong™, it was recorded, but did not count as a paw use. Examples of paw uses are shown in figures 8.1, 8.2 and 8.3. This test is ended when a total of 100 paw uses were recorded [41]. If the toy was empty before 100 paw uses, it was refilled. The experiment took place on a grass area close to the home environment of the dog during the experiment of Branson et al. [5], while it took place in the home environment of the dog during the experiment of Wells et al. [41]. Two experimenters were in the same area in order to record the paw uses of the dog. Since the dog moved around in the area, the experimenters did not have fixed positions, but moved in order to see the paw movements. However they remained between one and two meters in front of the dog. The dogs were not familiarized with the experimenter and during the test, no interaction between dog and human took place. The owners stayed outside of the testing area. Most dogs had not eaten for 12-18 hours before the test [5].

**First-stepping test**
The stepping test is explained in [36]. The idea behind this test is that in this case one does not need the dog to be hungry. Since the Kong™test need the dog to be interested in food, this might influence the result. Some dogs are not that interested in the food and thus getting 100 paw uses might take a long time. This is not the case during the stepping test, since no food is used.

During the first-stepping test, the first paw used to step-off from a stand position is recorded. The handler is holding the dog and they are standing on top of a staircase. The handler will either stand on the left or right side of the dog. Both the handler and experimenter would stay stationary during the experiment. The experimenter would call the dog and the paw used to step-off was recorded. The dog is rewarded with a small food treat. The handler will then call the dog to return to the stairs. This is repeated 10 times while the handler is standing on the left side of the dog and 10 times while the handler is standing on the right side of the dog. The next 15 times, the handler would be on the left and the final 15 times the handler would be on the right. This order was kept for each dog. Between each set, the dog was walked and allowed to drink before the next set. The relationship between experimenter and dog is not known, except that the dogs first performed The Kong™test with the experimenter. Before the test, all the dogs were checked by a veterinarian to ensure that they were in good health and no underlying condition that might influence their results [36]. This test tried to remove the reliability on food, however each time the dog finished the step, it was rewarded with food. By repeating this 50 times, a dog could learn and know which paw to use. So it is not paw preference, but learned behaviour. If the dog learns to step with left and gets rewarded, it might always choose left. Instead less steps should have been annotated, to reduce the change of learning. Besides this, the same order in which the handler stands is used for each dog, so an order bias can be introduced. Instead it might be better to alternate the starting position of the handler per dog. As paw preference, the first paw lifted was used, but they do not consider it paw preference when the dog jumps down with both paws. Is this considered no paw preference? Maybe it is better to not focus on paw preference, but paw usage.

**8.2.2. Conflicting results**
Multiple studies have recorded the paw preference of dogs and the relation between paw preference and sex. However these results are conflicting. [33] reported 75.1% right preference, 17.9% left preference and only 25.0% no preference in dogs. This was tested by letting dogs remove an adhesive plaster from the eyes. Research done by Branson and Rogers [6] found 44% of the dogs being left preferent, 33% being right-preferent and 23% having no preference. His test was done using The Kong™test. Other studies show no difference in the distribution of lateralized and non-lateralized dogs. [21] found a more even distribution of 46.5% of lateralized and 53.5% of ambilateral dogs using The Kong™test. [23] found that 37% of the dogs had a prefer-
ence, while 63% were ambilateral. They also found that there was a significant bias for left-paw use by male dogs and a right-paw use by female dogs. MacGrevey et al. used The Kong™ test as well for determining the paw preference. So even when the same test is used, the results are not the same. The lack of motivation for eating the food might cause the differences. [36] tried to solve the dependence on food by introducing the first stepping method.

8.2.3. Paw preference and behaviour

So instead of focusing on research of the distribution of dogs with a paw preference and dogs without a paw preference, research has shifted towards investigating the behavioral differences between these dogs. A study done by Brandon and Rogers [6] studied the relation between paw preference and noise phobia. Noise phobia is the expression of excessive fear in response to a sound stimulus [6]. He tested paw preference using The Kong™ test and he showed that the strength of paw preference regardless of the side are associated with reduced arousal and calmer responses in dogs. So if a dog had a strong preference for left, the dog would be less likely to express fear as response to a sound stimulus. The same hold for dog with an right paw preference. If the dog has no paw preference, it is more likely to have a noise phobia [6].

Marshall-Pescini et al. tested the relationship between paw preference and proactivity in novel situations [21]. They tested paw preference using The Kong™ test and let the dog do a problem solving task using a puzzle box with food. They found a more proactive inclination of non-lateralized dogs compared to lateralized dogs. So dogs with no paw preference are faster at settling down in a resting position when in a novel situation and are faster at catching a novel object [21].

Using both The Kong™ test and the adhesive tape test, Batt et al. researched the relationship between paw preference, salivary cortisol concentrations and behaviour [2]. During the adhesive tape test, a piece of tape is put on top of the dog’s nose. It then needs to take it off using its paw. The paw used is the preferred paw. Batt et al. showed that stronger preference is associated with more confident and relaxed behaviour in a novel situation. Dogs with a strong paw preference, are likely to be more confident and relaxed in a novel environment [2].

A study done by Wells et al. [42] showed no association between paw preference and behaviour problems. Paw preference was tested using The Kong™ test, while the behaviour profile was determined using the C-BARQ, a psychometric tool designed to provide standardized evaluations of canine behaviour [14]. They did find that dogs with a left paw preference have lower score than dog with right paw preference or dogs with no preference on the score of stranger-directed aggression [42]. The higher the score, the more aggression the dog showed.

A study done by van Alphen et al. [38] studied the correlation between paw departure and task performance. Using the first paw to take of during a search test as the departing paw. They found that dogs departing with left were performing significantly better than dogs that depart with right. Meaning that dogs departing with left usually had a good performance, while dog departing with right had a bad performance. They also found that there was no gender differentiation in their data. This suggest that paw use might be task specific and correlated with performance. In this thesis we investigate if this is indeed the case and replicate this research using this and other experiments conducted for the Dog Cognition research program.
To test the hypotheses, several sub-experiments are created. Below each experiment is explained. In total there are 205 dogs of various ages, training levels and breeds who participated in these experiments. The experiments were carried out at different locations and areas like sand fields or grass fields. Before the experiments started, an intake form was filled in by the owner of the dog.

9.1. Design of the experiment

The experiment consists of 5 sub-experiments. Figure 9.1 shows the general overview of the experimental area. Experiment A is outside of the experimental area. The field is built with respect to the wind direction, in order to have the scent of the goody in the right way. The location of the experiment and the terrain is different. These factors are noted down, however the influence of these factors are not studied in this thesis. In this thesis only experiment C, D and E are used. Experiment A is used to let the dog get familiar with the area. For example the smell, the noises and other stimulants of the area. More information about experiment A and B can be found online [16].

For the experiments, several items are needed as shown in the general overview of the experiment. The most important items are listed below. For a complete overview, see appendix B.

- 2 cameras in order to record the experiment.
- flags to indicate the borders of the area and the positions needed.
- cups, which are cones to hide the goody.
- cones, needed for the alley in experiment D.
- Goodies, sausages for the dog to find.
- leash, long leash of 5 meters with a collar.

The cones are numbered. The goody cones are separated from the other cones, to limit the scent of the sausage to the goody cones only. A plate is used to lay the sausage on, to avoid the scent getting on the ground.

9.1.1. Experiment A: Walking the dog

The goal of this experiment is to let the dog get familiar with the area. For example the smell, the noises and other stimulants of the area.

Figure 9.2 shows the set up of experiment A. There are two trials. One in which the handler starts walking with their left foot and one where the handler starts walking with their right foot. At the start the dog is sitting to the left of the handler. To the left or to the right of the dog the experimenter calls the dog. This causes the dog to look left or right accordingly. Now the handler is starting to walk. The order of the foot the handler used to start walking and the order when the experimenter calls the dog on the left or right side is changed for each dog. Differing the order of using the left or right foot of the handler and sitting left or right from the dog rules out an order affect. Since the order is different every dog.
9.1.2. Experiment C: Find the goody
The goal of this experiment is to test the hypotheses that paw preference correlates to performance in a simple search task.

Figure 9.3 shows the set up of experiment C. This experiment has 3 trials. In each trial the goody is hidden under one of the three cups. The handler blindfolds the dog. If the dog resists the blindfold, a wooden board is used to obstruct the dog's view on the cups. While the dog is blindfolded, the experimenter holds the goody under the dogs nose. When the dog has the smell of the goody, the goody is hidden under one of the three cones. Then the blindfold is removed and the dog is asked to search for the goody. The experiment is done when the dog finds the goody or goes straight towards the experimenter. Going to the experimenter is also considered as finding the goody, since the experimenter also has a smell. When the experimenter hold the goody under the dog's nose, the dog can both smell the goody and the experimenter. So if the dog is following its nose, it can also go to the experimenter.

9.1.3. Experiment D: Alley of cones
The goal of this experiment is the same as experiment C, to check if there exist a correlation between paw departure and performance. But it is also to check if changing the environment with more cones affect the behaviour of the dog. Can the additional cones help guide the dog to the goody? Or is it giving the dog more pressure and limiting the environment for the dog.

Figure 9.4 shows the set up of experiment D. There are two differences with experiment C. First of all an
9.1. Design of the experiment

Figure 9.2: Set up of experiment A

Figure 9.3: Set up of experiment C

Figure 9.4: Set up of experiment D
alley of cones is added between the dog and the goody cups. This alley consist of 12 cones evenly distributed. The approximate width of the alley is the width of the dog plus 40 cm. The 40 cm is chosen to give the dog enough space to walk in between the cones. The second difference is that the goody is hidden under the middle cup in all three trials.

9.1.4. Experiment E: The final exercise
The goal of this exercise is to test the influence of pressure on the performance of the dog, by comparing its behaviour with that in experiments C and D. The final exercise is experiment E. Figure 9.5 shows the set up of experiment E. In this exercise, all pressure is removed. The dog is not blindfolded anymore and the goody is not hidden under a cone anymore, but placed directly visibly for the dog on the ground.

9.2. Aspects of performance
In order to test the validity of the hypothesis, the notion performance needs to be defined in objective and observable terms. A perfect performance is that a dog goes in a near straight line from the starting point to the correct cone. Bad performance is defined as misbehaving or taking too long to find the goody. These definitions were obtained in a number of brainstorm sessions with domain experts in dog training. The definitions were formally defined before the start of the experiments. During the experiments, not only the correct cone had the smell of the goody, also the experimenter had his own smell. When the experimenter hold the goody under the dog's nose, the dog can both smell the goody and the experimenter. If the dog is following its nose, it can also go to the experimenter. So it is also considered to be good performance when a dog goes straight towards the experimenter. Moving away from the goody cone or the experimenter, stopping, biting, peeing, sniffing at wrong cones, moving backward, passing the goody cone, going to the cameraman and going to the handler during the trial is considered bad behaviour, while moving towards the experimenter or the goody is considered good behaviour.

Paw departure is defined as the first paw used to depart. In order to have a clear paw departure, the dog has to stand still for at least two consecutive images on the video at the start of each trial. Meaning that its paws do not move in those two images. Furthermore, the paw noted down as the departing paw is defined in relation to the gait in which the dog departs. If the dog departs in a walk or a trot, then the first paw to move forward is taken as the paw used. If the dog departs in a gallop, in which both paws move forward, the paw that reaches most forward is taken as the paw used.
Results paw departure and performance

For the second part of this thesis an analysis was made to check if the correlation between paw departure and performance persisted over multiple tasks. It uses the annotated data to determine if the dogs used their left or right paw and if they performed well. The exercises tested in this hypothesis are experiment C, D and E (see Chapter 9).

10.1. Subjects
205 dogs participated in the experiments. From these dogs 105 could not be used for analysis due to mistakes with equipment or mistakes during experiments, for example dog did not stand still before departing or the handler pulled the leash at moments that influenced the results. Some of these dogs could also not be used if they did not complete all exercises. So from the remaining 100 dogs, a subset of 48 dogs was selected with the aim of having a fair distribution with respect to training level, age and gender. A low level of training means, none at all, puppy-or basic levels and a high level means GG1 or higher in Dutch training terms. Age was taken as adult or non-adult. For males adult means 2 years or older, for females this is 1.5 years. This is conform to [7, 30]. So there were 24 males and 24 females, each group consisting of 12 dogs with low and 12 dogs with high training level. Non-adult dogs with high training level are hard to come by. Therefore 4 out of 24 non-adult dogs had a high training level and 20 out of 24 adult dogs had a high training level. Appendix A table A.2 summarizes the data of the selected dogs. It shows their age in months during the experiment. Of the 48 dogs used, there were 6 dogs, who used a wooden board to hide the goody instead of the blindfold. This was because there was a strong resistance towards the blindfold. Using the board instead of the blindfold might influence the results. However this is not investigated in this research.

10.2. Analysis
For each dog, both videos were used to annotate the behaviour in a formal language. The annotation was created with at least 2 people viewing the video images. If the 2 people disagreed, 2 more people would view the video. After this, all people involved discussed the images. These disagreements mostly concerned the paw departure. These were solved by agreeing on the gait the dog used to depart, since trot and gallop can be misleading. The results were statistically analyzed using the $\chi^2$ test. This test is chosen because we have categorical data and we want to do a hypothesis test. The null hypothesis is that paw departure is independent of performance. The alternative hypothesis is that paw departure and performance are dependent and there exist some kind of correlation. For a chi square test, the expected frequency count for at least 80% of the cells is 5 or more [27].

These results are shown in the table 10.1. It contains the $\chi^2$ analysis of the paw-performance variables. During this study a p value smaller than 0.05 is considered significant. The second column indicates the experiment and trial number. Each Experiment had 3 trials. The next four columns give the number of observations of the different combinations. The last two columns show the $\chi^2$ value and the significance of the $\chi^2$ tests. For example, the first row show the results of experiment C trial 1, where 13 dogs departed with their left paw and performed good, 2 departed with left and performed bad. 7 departed with their right paw and performed good and 26 dogs departed with right and performed bad. The last two columns show that there is significant
correlation between paw performance and paw departure ($\chi^2 = 18.18, p<0.001$)

<table>
<thead>
<tr>
<th>First paw</th>
<th>Trial</th>
<th>Left/Good</th>
<th>Left/Bad</th>
<th>Right/Good</th>
<th>Right/Bad</th>
<th>$\chi^2$</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vs</td>
<td>C1</td>
<td>13</td>
<td>2</td>
<td>7</td>
<td>26</td>
<td>18.18</td>
<td>2.01E-05 *</td>
</tr>
<tr>
<td>Performance</td>
<td>C2</td>
<td>11</td>
<td>12</td>
<td>3</td>
<td>22</td>
<td>7.44</td>
<td>0.006371 *</td>
</tr>
<tr>
<td></td>
<td>C3</td>
<td>7</td>
<td>14</td>
<td>5</td>
<td>22</td>
<td>1.38</td>
<td>0.239639</td>
</tr>
<tr>
<td></td>
<td>C all</td>
<td>31</td>
<td>28</td>
<td>15</td>
<td>70</td>
<td>19.51</td>
<td>1.0E-05 *</td>
</tr>
<tr>
<td></td>
<td>D1</td>
<td>8</td>
<td>11</td>
<td>0</td>
<td>29</td>
<td>14.64</td>
<td>0.000129 *</td>
</tr>
<tr>
<td></td>
<td>D2</td>
<td>13</td>
<td>11</td>
<td>5</td>
<td>19</td>
<td>5.69</td>
<td>0.017073 *</td>
</tr>
<tr>
<td></td>
<td>D3</td>
<td>15</td>
<td>7</td>
<td>6</td>
<td>20</td>
<td>9.85</td>
<td>0.001697 *</td>
</tr>
<tr>
<td></td>
<td>D all</td>
<td>36</td>
<td>29</td>
<td>11</td>
<td>68</td>
<td>27.88</td>
<td>1.29E-07 *</td>
</tr>
<tr>
<td></td>
<td>E</td>
<td>29</td>
<td>5</td>
<td>8</td>
<td>6</td>
<td>4.45</td>
<td>0.034924 *</td>
</tr>
<tr>
<td></td>
<td>CDE all</td>
<td>75</td>
<td>40</td>
<td>29</td>
<td>96</td>
<td>43.06</td>
<td>5.3E-11 *</td>
</tr>
</tbody>
</table>

Table 10.1: Results test first paw and performance, (* significant p value)

As shown in the table 10.1, over all C trials there is a significant correlation between paw departure and performance ($\chi^2 = 19.51, p<0.001$). The same hold for all D trials ($\chi^2 = 27.88, p<0.001$) and for all trials including E ($\chi^2 = 43.06, p<0.001$). All p values are below 0.05 and thus we can reject the null hypothesis. So one can conclude that there is indeed a significant correlation between paw departure and performance. As already shown in the study from van Alphen et al. [38] and thus confirming the hypothesis that there is a correlation between paw departure and performance.

Table 10.2 shows the correlation between paw use and gender ($\chi^2=0.43, p=0.51$), gender and performance ($\chi^2=0.80, p=0.37$) and adult and performance ($\chi^2=0.20, p=0.65$). As shown in these tables, the p values are above 0.05 and thus we cannot reject the null hypothesis that there is no correlation between paw use and gender, gender and performance and age and performance. This shows that non-adult dogs do not perform significantly different from adult dogs. Also female dogs do not perform significantly different than male dogs and finally it seems that there is no paw preference depending on gender. This is in line with what van Alphen et al. showed in their research [38]. These results show that the tasks are indeed simple and no training is needed. Otherwise one would expect to have a correlation between training level and performance. So the tasks are indeed simple for each dog. Each dog is able to follow the scent of the goody no matter its age or training level.

Table 10.2: Correlations between paw use and gender, gender and performance and age and performance

<table>
<thead>
<tr>
<th>Variables</th>
<th>Trial</th>
<th>Left/Female</th>
<th>Left/Male</th>
<th>Right/Female</th>
<th>Right/Male</th>
<th>$\chi^2$</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>First paw and gender</td>
<td>All</td>
<td>82</td>
<td>76</td>
<td>86</td>
<td>92</td>
<td>0.43</td>
<td>0.51</td>
</tr>
<tr>
<td>Gender and Performance</td>
<td>All</td>
<td>69</td>
<td>61</td>
<td>99</td>
<td>107</td>
<td>0.80</td>
<td>0.37</td>
</tr>
<tr>
<td>Adult and Performance</td>
<td>All</td>
<td>67</td>
<td>101</td>
<td>63</td>
<td>105</td>
<td>0.20</td>
<td>0.65</td>
</tr>
</tbody>
</table>

Table 10.2: Correlations between paw use and gender, gender and performance and age and performance

<table>
<thead>
<tr>
<th>Variables</th>
<th>Trial</th>
<th>Left/Female</th>
<th>Left/Male</th>
<th>Right/Female</th>
<th>Right/Male</th>
<th>$\chi^2$</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Good</td>
<td>C</td>
<td>46</td>
<td>98</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>47</td>
<td>97</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>E</td>
<td>37</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 10.3: Correlation between trial conditions and performance

Table 10.3 show the correlation between trial conditions and performance. The null hypothesis is that the tasks are independent of each other, while the alternative hypothesis is that there is a correlation between
the tasks. It shows that the correlation between the trials is significant ($\chi^2=34.81$, $p=<0.001$). However when looking at the correlations separately, it is shown that between trial C and D there is no significant correlation ($\chi^2=0.02$, $p=0.90$). We cannot reject the null hypothesis, since $p > 0.05$. While between C and E and D and E there is a significant correlation ($\chi^2=29.89$, $p=<0.001$) and ($\chi^2=28.90$, $p=<0.001$). So in both cases we can reject the null hypothesis and accept that there is significant correlation between the circumstances in exercise C and E and here is significant correlation between the circumstances in exercise D and E. This suggest that adding the alley does not significantly influence the performance of the dog. It does not increase or decrease the performance significantly. However removing all pressure by removing the pylons, not blindfolding the dog and place the goody visible on the ground indeed significantly increases the performance of the dog.
III

Dog Formalizations
Formalizing and Checking Dynamic properties

After gathering data, transcribing it and translating it into traces, it is time to use the checker software. The software checks the validity of formalized properties. In this research, two formalizations are created. The first checks the performance of the dog together with the first paw used. This formalization is used in the comparison between handwritten and automated defined complex behaviour in chapter 6. The second formalization is used to answer the research question: Is there a correlation between the circumstances of the task and the route the dog takes to the goody? It is expected that adding the alley in experiment D changed the route of the dog. It is expected that the number of zigzag routes will increase in D compared to C, if the alley is distracting the dog. However if the alley is guiding the dog, than we expect to see more straight routes. In experiment E, we expect to have more straight routes, since all pressure is gone and the dog can see where the goody is placed. The number of backward routes will be reduced too. Below the formalizations are explained.

11.1. Performance formalization

In order to know if there is a correlation between paw departure and performance, we need to check which paw the dog uses and how the dog performs. In order to do so, several formalization’s were created. One checks if the dog departs with left and performs good, one checks if the dog departs with left and performs bad, one checks if the dog departs with right and performs good and finally one checks if the dog departs with right and performs bad. Departing with left or departing with right is clearly defined. But it is sometime difficult to see. For example when the camera did not film the paws correctly. A dog can move in different speeds, for example walk, trot or gallop. This can also make it difficult to determine the paw used to depart, since one needs to determine the speed used. This is the first movement of the dog. However good and bad performance needs a clear definition.

11.1.1. Definition

Bad performance is defined as misbehaving or taking too long to find the goody. Good performance is defined as not bad performance and eating or taking the goody or going to the experimenter. These definitions are the guideline to the more precise definitions used in the formalization.

Bad performance

A bad performance is when the dog misbehaves during searching or takes to long to find the goody.

- The dog moves backwards. Moving backwards is defined as moving into directions 4 till 8.
- the dog sniffs at cones that do not have the goody. Sniffing means sniffing the object while staying at the object.
- The dog passes the goody or the cone that hides the goody.
- The dog is standing still
- The dog marks an object. Marking an object can be done by urinating or defecate on the object or by touching it with its paw, body or tongue.
11. Formalizing and Checking Dynamic properties

- The dog goes to the cameraman
- The dog goes to the handler
- The dog looks up to the handler, as if waiting for commands, confirmation or instructions. Except when this happens right after taking off to go search at the goody. Or when the dog is at the goody cone, before getting the goody.
- The dog avoids the alley in experiment D
- The dog leaves the alley before getting to the end in experiment D.
- The experimenter ends the trial before the dog finds the goody.

These definitions are used to define misbehaving in the formal notation.

Taking too long to find the goody is when the trial takes much longer than it would take in an perfect scenario. Taking longer than expected can happen because the dog is new in this situation and just does not know how to act. For example it could happen that the dog keeps searching while it already knows for certain where the goody is hidden. The experimenter can decide to help the dog by leading the dog to a place where it can catch the smell of the goody. But if the experimenter thinks that the dog already knows where the goody is, but keeps searching, the experimenter can terminate the trial. Determining how long a perfect scenario takes is not possible. So instead experiment E is used to determine this time. In experiment E, there are no cones and the goody is not hidden anymore. In this exercise the dogs usually go straight to the goody. To determine if a trial takes too long, the duration of experiment E is used and multiplied by 4. If the trial takes longer, then it takes too long.

**Good performance**

For good performance, the dog need to behave in the following way:

- The dog goes to the experimenter
- The dog is not showing bad behaviour
- The dog eats or takes the goody.
- The dog goes through the alley in experiment D.

During the exercises, the experimenter is also a large body of smell and presented to the nose of the dog while it is blindfolded. So if the dog follows the scent, going to the experimenter is also correct.

11.1.2. Semi formal notation

Each experiment is split into sections. Each sections shows a certain part of the exercise, for example sitting, taking off or searching. A complete overview is shown in appendix C. In experiment C, D and E, the dog takes off in section v. The searching takes place in section vi. As described before, performance is split into 4 separate definitions: (left, good), (left, bad), (right, good), (right, bad). So the property performance(left,good) declares that the dog takes off with its left paw in section v and then performs good in section vi of the particular experiment and trial. Below an example definition for exercise C, trial 1.

**Performance(left, good)**

For all time points

| if | the examined section is exercise(c, 1, v) |
| and | the dog steps with its left front paw forward during this section |
| then | during the vi section of exercise c trial 1, the dog shows good performance. |

Similar definitions hold for Performance(left, bad), Performance(right, good) and Performance(right, bad) and for the different experiments. Figure 11.1 shows the time line of the definition above. First the start and end points of each section is specified. At time point t, the dog steps forward with its left paw. During the next section, the dog has a good performance.

11.1.3. Formal notation

The TTL checker uses logic to check the formalizations. The formal notations needed for the performance property are given below.

**Performance(left, good, exercise c, trial 1)**

\[
\forall m \exists t_0, t_e : t_0 \leq t \leq t_e, \forall t_0, t_1 : t_e \leq t_0, t_1 \leq t_e.
\]
11.1. Performance formalization

Figure 11.1: example time line for performance (left, good)

\[
\begin{align*}
\text{trace_part_of_interest}(m, \text{exercise}(c, 1, v), t, t_b, t_e) \land \\
\text{first_step_of_dog_left_forward}(m, t, t_b, t_e) \land \\
\text{trace_part_of_interest}(m, \text{exercise}(c, 1, v), t, t_b, 1, t_e) \land \\
\text{good_performance}(m, t_b, 1, t_e)
\end{align*}
\]

\[
\begin{align*}
\forall t_1 & \quad t_b \leq t_1 \leq t_e \text{ state}(m, t_1) \models E \land \\
& \quad t_1 < t_b \text{ state}(m, t_1) \models \neg E \land \\
& \quad t_e < t_1 \text{ state}(m, t_1) \models \neg E
\end{align*}
\]

\[
\begin{align*}
\text{first_step_of_dog_left_forward}(m, t, t_b, t_e) \\
& \forall t_1, \forall V_1: \text{LeftRight}, t_b \leq t_1 < t : \text{state}(m, t_1) \models \neg \text{dog_moves_paw}(V_1, \text{forward})
\end{align*}
\]

\[
\begin{align*}
\text{good_performance}(m, t_b, t_e) \\
& \neg \text{misbehaving_during_exercise}(m, t_b, t_e) \land \\
& \neg \text{trial_takes_too_long}(m) \land \\
& \exists t : t_b \leq t \leq t_e \land \\
& \quad \text{state}(m, t) \models \text{dog_eats}(\text{goody}) \lor \\
& \quad \text{state}(m, t) \models \text{dog_takes}(\text{goody})) \lor \\
& \exists t_2 : t_b \leq t_2 \leq t_e, ds : \text{dogspeed} \\
& \quad \neg \text{misbehaving_during_exercise}(m, t_2, t_e) \land \\
& \quad (\text{state}(m, t_2) \models \text{goes_to}(\text{dog}, \text{experimenter}) \lor \\
& \quad \text{state}(m, t_2) \models \text{goes_towards}(\text{dog}, \text{experimenter}) \lor \\
& \quad \text{state}(m, t_2) \models \text{dog_sniffs}(\text{at_experimenter_s_hand}) \lor \\
& \quad \text{state}(m, t_2) \models \text{dog_moves_to}(\text{experimenter}, ds))
\end{align*}
\]

\[
\begin{align*}
\text{bad_performance}(m, t_b, t_e) \\
& \neg \text{good_performance}(m, t_b, t_e)
\end{align*}
\]

\[
\begin{align*}
\text{misbehave_during_exercise}(m, t_b, t_e) \\
& \exists t : t_b \leq t \leq t_e, ds : \text{dogspeed} \\
& \quad \text{dog_moves_backwards}(m, t_b, t_e) \lor \\
& \quad \text{dog_sniffs_at_wrong_pylon}(m, t_b, t_e) \lor \\
& \quad \text{state}(m, t) \models \text{passes}(\text{dog, goody}) \lor \\
& \quad \text{state}(m, t) \models \text{passes_pylon}(\text{dog, goody_pylon}) \lor \\
& \quad \text{state}(m, t) \models \text{dog_stands} \lor \\
& \quad \exists py : \text{pylon state}(m, t) \models \text{dog_bites_pylon}(py) \lor \\
& \quad \exists o : \text{object, s : substance state}(m, t) \models \text{dog_marks}(o, s) \lor 
\end{align*}
\]
11. Formalizing and Checking Dynamic properties

\[
state(m, t) \models \text{goes_to(dog, handler)} \vee
\text{looking_at_handler_at_wrong_time}(m, t_p, t_e)
\]

\[
\text{dog_moves_backward}(m, t_p, t_e)
\exists t, d, ds : \text{dogs_speed} : t_p \leq t \leq t_e, 4 \leq d \leq 8
\quad state(m, t) \models \text{dog_moves_to}(d) \vee
\quad state(m, t) \models \text{dog_moves_to}(d, ds)
\]

\[
\text{dog_sniffs_at_wrong_pylon}(m, t_p, t_e)
\exists t, p : \text{pylon} : t_p \leq t \leq t_e
\quad state(m, t) \models \text{dog_sniffs_at_pylon}(p) \land
\quad p \neq \text{good_pylon}
\]

\[
\text{looking_at_handler_at_wrong_time}(m, t_p, t_e)
\exists t, t_1 : t_p < t < t_1 \leq t_e
\quad trace\_part\_of\_interest(m, \text{dog_looks(up_to_handler)}, t, t_1) \land
\quad \neg \text{at_good_pylon}(m, t)
\]

\[
\text{at_good_pylon}(m, t)
\exists t_1 : t < t_1 \leq t_1 \leq t_e
\quad state(m, t_2) \models \text{goes_to_pylon(dog, good_pylon)} \lor
\quad state(m, t_2) \models \text{dog_sniffs_at_pylon(good_pylon)}
\]

\[
\text{trial_takes_too_long}(m)
\exists \text{smi, ssec, emin, esec, smin}_1, \text{ssec}_1, \text{emin}_1, \text{esec}_1
\quad \text{time_points}(m, \text{exercise}(c, l, vi), \text{smi}, \text{ssec}, \text{emin}, \text{esec}) \land
\quad \text{time_points}(m, \text{exercise}(e, l, vi), \text{smi}_1, \text{ssec}_1, \text{emin}_1, \text{esec}_1) \land
\quad \text{unacceptable_duration}(\text{smi}, \text{ssec}, \text{emin}, \text{esec}, \text{smi}_1, \text{ssec}_1, \text{emin}_1, \text{esec}_1)
\]

\[
\text{time_points}(m, E, \text{smi}, \text{ssec}, \text{emin}, \text{esec})
\exists t
\quad state(m, t) \models E \land
\quad state(m, t) \models \text{times_on_camera}(\text{smi}, \text{ssec}, \text{emin}, \text{esec}, \text{cam}_{-II})
\]

\[
\text{unacceptable_duration}(\text{smi}, \text{ssec}, \text{emin}, \text{esec}, \text{smi}_1, \text{ssec}_1, \text{emin}_1, \text{esec}_1)
\exists \text{diff_m}, \text{diff_s}, \text{diff_m} \text{diff_s}_1
\quad \text{duration}(\text{smi}, \text{ssec}, \text{emin}, \text{esec}, \text{diff_m}, \text{diff_s}) \land
\quad \text{duration}(\text{smi}_1, \text{ssec}_1, \text{emin}_1, \text{esec}_1, \text{diff_m} \text{diff_s}_1) \land
\quad \text{more_than_four_times}(\text{diff_m}, \text{diff_s}, \text{diff_m}_1, \text{diff_s}_1)
\]

\[
\text{duration}(\text{smi}, \text{ssec}, \text{emin}, \text{esec}, \text{diff_m}, \text{diff_s})
\quad (\text{emin} - \text{smi}) = \text{diff_m} \land
\quad (\text{esec} - \text{ssec}) \geq 0 \rightarrow (\text{esec} - \text{ssec} = \text{diff_s}) \land
\quad (\text{esec} - \text{ssec}) < 0 \rightarrow (\text{esec} - \text{ssec} + 60 = \text{diff_s})
\]

\[
\text{more_than_four_times}(\text{diff_m}, \text{diff_s}, \text{diff_m}_1, \text{diff_s}_1)
\exists k, \text{dmin}, \text{dsec}
\quad k = (4 \ast \text{diff_s}_1) \land
\quad (k \geq 60 \rightarrow \text{dmin} = ((4 \ast \text{diff_m}_1) + 1) \land \text{dsec} = (k - 60)) \land
\quad (k < 60 \rightarrow \text{dmin} = (4 \ast \text{diff_m}_1) \land \text{dsec} = k) \land
\quad (\text{diff_m} > \text{dmin} \lor (\text{diff_m} = \text{dmin} \land \text{diff_s} > \text{dsec})
\]

11.2. Route formalization

Besides checking if the dog performs good or bad, it is also interesting to check which route the dog took to get to the goody. A route formalization was created for experiment C,D and E to see how the dog got to the goody.
The route can give insight in the differences between experiment C, D and E and the influence of pressure in the task. For example the alley added in experiment D makes dogs go straight, while in experiment C they might have taken a left or right turn.

11.2.1. Definition
During the search of the goody, the dog needs to walk towards the goody. This formalization decides which route the dog takes to the goody. To be able to compare, we only considers routes that end by finding the goody at the cone. Since walking to the experimenter will give a different route than walking to the cones. A dog can move in different speeds: walk, gallop, trot and an irregular speed. The direction in which the dog walks is determined using a clock. Figure 11.2 shows how this is used. A left movement is a movement in direction 9 or 10. A straight movement is in direction 11, 12 or 1. right is in direction 2 or 3. All movements in direction 4, 5, 6, 7 or 8 are considered backward movements. Using these definitions, we can determine how the dog started and how the dog ended its route. If the dog starts left, the first movement of the dog is left or the first movement is straight followed by a left movement. If the dog starts right, the first movement of the dog is right or the first movement is straight followed by a right movement. Starting straight is by starting with a straight movement.

![Figure 11.2: clock used to determine direction](image)

With these starting definitions, routes can be defined. For a left route, the starting movement should be left and the direction can only be changed once. A change in direction means going from a left movement to a right movement or going from right to left. If the direction is changed another time, the route is transformed to a zig zag route, since the dog goes from left to right to left. Going backwards is not allowed, since this changes the route to a backward route. For a right route, the starting movement should be right and the direction can only be changed once. For a straight route, the starting movement should be straight, and all other movements should be straight too. For a zig zag route, the dog can start left or right, but there need to be at least 2 direction changes. A change in direction means going from a left movement to a right movement or from right to left. A backward route is any route that contain a backward movement. Example routes are given in figure 11.3.

11.2.2. Semi formal notation
Below an example definition of a left route in exercise c trial 1 is given.

**Left route**

For all time points

- if the examined section is exercise(c, 1, vi)
- and the dog starts with a left movement
- and there exists a time point t, such that there is no right movement before t
  - and no left movement after t.
- and the dog finds the goody under the cone
- then during the vi section of exercise c trial 1, the dog takes a left route to the goody

Figure 11.4 shows the line for the left route definition. At time point T the dog does a first left movement. This movement can be preceded by multiple straight movements. At time point T3, the dog has its final movement, this can be straight, right or left. The blue area indicated all time points in between these movements. At some point in this blue area there exist a time point T1. In the blue area before T1, there is
Formalizing and Checking Dynamic properties

(a) examples of left routes

(b) examples of right routes

(c) examples of straight routes

(d) examples of zig zag routes

Figure 11.3: Example routes

no right movement. In the blue area after T1 there is no left movement. At time point Te3 the dog finds the
goody. Similar definitions hold for the right, straight, zigzag and backward routes.

Figure 11.4: example time line for Left route.

To make this definition clear, a state machine was created. This is shown in figure 11.5. It shows how
different movements lead to the different routes. S is the straight route, B is the backward route, Z is the
zigzag route. sL is the start of the left route and eL is the end of the left route. sR is the start of the right route
and eR is the end of the right route. s is a straight movement between 11 and 1. l is a left movement, between
9 and 10. r is a right movement between 2 and 3. b is a backward movement between 4 and 8. From this state
machine regular expressions are created as follows:

- Left route \( L_r \) = \( \{ S^* L; S, L \}^* \{ S, R \}^* \).
- Right route \( R_r \) = \( \{ S^* R; S, R \}^* \{ S, L \}^* \).
- Straight route = \( \{ S^* \} \).
- ZigZag route = \( \{ [L_L^*; R_L^* L_L^*], (R_L^* L_L^* R_L^*)]^* \}\).
- All movements \( M \) = \{ L, R, S, B \}.
- Backward route = \( \{ M^* BM^* \} \).

Where S stand for a straight movement, L stands for a left movement, R stands for a right movement, \( L_r \) is a
left route, \( R_r \) is a right route, M are all movements and B is a backward movement.
11.2.3. Formal notation

The TTL checker uses logic checks. The formal notations needed for the route properties are given below. To create a better understanding on the use of this formalization, a step by step example is given in Appendix D.

Left_route(exercise c, trial 1)

\[ \forall m \exists t_0, t, t_3, t_1, t_2 : t_0 \leq t, t \leq t_3, t_0 \leq t_1, t_1 \leq t_3, t_0 \leq t_4, t_4 \leq t_3, t_0 \leq t_3 \land \]

\[ \text{trace_part_of_interest}(m, exercise(c, 1, vi), t_0, t_3) \land \]

\[ \text{dog_starts_left}(m, t, t_0, t_3) \land \]

\[ \forall t_2, d : \text{rightmove}, d_2 : \text{leftmove}, ds, ds_2 : \]

\[ (\text{begin}(t_0) \leq \text{begin}(t_2) \leq \text{end}(t_2) \leq \text{begin}(t_4), t_1 \neq t_2 \rightarrow \]

\[ \neg (\text{state}(m, t_2) \models \text{dog_moves_to}(d)) \lor \neg (\text{state}(m, t_2) \models \text{dog_moves_to}(d, ds)) \land \]

\[ (\text{end}(t_1) \leq \text{begin}(t_2) \leq \text{end}(t_2) \leq \text{end}(t_3) \rightarrow \]

\[ \neg (\text{state}(m, t_2) \models \text{dog_moves_to}(d_2)) \lor \neg (\text{state}(m, t_2) \models \text{dog_moves_to}(d_2, ds_2)) \land \]

\[ \neg \text{dog_moves_backward}(m, t_0, t_3) \land \]

\[ (\text{state}(m, t_4) \models \text{dog_eats}(	ext{goody})) \lor \]

\[ (\text{state}(m, t_4) \models \text{dog_takes}(	ext{goody})) \]

Right_route(exercise c, trial 1)

\[ \forall m \exists t_0, t, t_3, t_1, t_2 : t_0 \leq t, t \leq t_3, t_0 \leq t_1, t_1 \leq t_3, t_0 \leq t_4, t_4 \leq t_3, t_0 \leq t_3 \land \]

\[ \text{trace_part_of_interest}(m, exercise(c, 1, vi), t_0, t_3) \land \]

\[ \text{dog_starts_right}(m, t, t_0, t_3) \land \]

\[ \exists t_2, d : \text{leftmove}, d_2 : \text{rightmove}, ds, ds_2 : \]

\[ (\text{begin}(t_0) \leq \text{begin}(t_2) \leq \text{end}(t_2) \leq \text{begin}(t_4), t_1 \neq t_2 \rightarrow \]

\[ \neg (\text{state}(m, t_2) \models \text{dog_moves_to}(d)) \lor \neg (\text{state}(m, t_2) \models \text{dog_moves_to}(d, ds)) \land \]

\[ (\text{end}(t_1) \leq \text{begin}(t_2) \leq \text{end}(t_2) \leq \text{end}(t_3), t_1 \neq t_2 \rightarrow \]

\[ \neg (\text{state}(m, t_2) \models \text{dog_moves_to}(d_2)) \lor \neg (\text{state}(m, t_2) \models \text{dog_moves_to}(d_2, ds_2)) \land \]

\[ \neg \text{dog_moves_backward}(m, t_0, t_3) \land \]

\[ (\text{state}(m, t_4) \models \text{dog_eats}(	ext{goody})) \lor \]

\[ (\text{state}(m, t_4) \models \text{dog_takes}(	ext{goody})) \]
Straight_route\( (\text{exercise c}, \text{trial 1}) \)
\[
\forall m \exists t_0, t, t_0, t_2, t_3 : t_0 \leq t, t \leq t_0, t_0 \leq t_2, t_0 \leq t_3, t_3 \leq t_2 \wedge \\
\text{trace_part_of_interest}(m, \text{exercise}(c, 1, vi), t_0, t_2) \wedge \\
\text{dog_starts_straight}(m, t, t_0, t_2) \wedge \\
\forall t_1, d : \text{notstraight}, ds : \\
(\text{end}(t) \leq \begin{state}(t_1), \begin{state}(t_1) \leq \begin{state}(t_2), t_1 \neq t_2 \rightarrow \\
\neg (\begin{state}(m, t_1) \models \text{dog_moves_to}(d)) \lor \neg (\begin{state}(m, t_1) \models \text{dog_moves_to}(d, ds)) \wedge \\
\neg \text{dog_moves_backward}(m, t_0, t_2) \wedge \\
(\begin{state}(m, t_3) \models \text{dog_eats(goody)}) \lor \\
(\begin{state}(m, t_3) \models \text{dog_takes(goody)})
\]

ZigZag_route\( (\text{exercise c}, \text{trial 1}) \)
\[
\forall m \exists t_0, t, t_2 : t_0 \leq t, t_2 \leq t_0 \\
\neg \text{straight_route}(\text{exercise}(c, 1, vi)) \wedge \\
\neg \text{right_route}(\text{exercise}(c, 1, vi)) \wedge \\
\neg \text{left_route}(\text{exercise}(c, 1, vi)) \wedge \\
\neg \text{dog_moves_backward}(m, t_0, t_0) \wedge \\
(\begin{state}(m, t_2) \models \text{dog_eats(goody)}) \lor \\
(\begin{state}(m, t_2) \models \text{dog_takes(goody)})
\]

trace_part_of_interests\( (m, \text{exercise}(c, 1, vi), t_0, t_e) \)
\[
\forall t_1 \\
\begin{state}(t_0) \leq t_1 \leq t_e \rightarrow \begin{state}(m, t_1) \models \text{exercise}(c, 1, vi) \wedge \\
\begin{state}(t_1) < \begin{state}(t_0) \rightarrow \neg (\begin{state}(m, t_1) \models \text{exercise}(c, 1, vi)) \wedge \\
\begin{state}(t_0) < \begin{state}(t_1) \rightarrow \neg (\begin{state}(m, t_1) \models \text{exercise}(c, 1, vi)) 
\]

dog_starts_left\( (m, t, t_0, t_0) \)
\[
\exists d : \text{leftmove}, ds : \begin{beginning}(t_0) \leq \begin{beginning}(t), \begin{end}(t) \leq \begin{end}(t_0) \\
\neg \text{notstraight}(t_0, t) \wedge \\
(\begin{state}(m, t) \models \text{dog_moves_to}(d) \lor \begin{state}(m, t) \models \text{dog_moves_to}(d, ds))
\]

dog_starts_right\( (m, t, t_0, t_0) \)
\[
\exists d : \text{rightmove}, ds : \begin{beginning}(t_0) \leq \begin{beginning}(t), \begin{end}(t) \leq \begin{end}(t_0) \\
\neg \text{notstraight}(t_0, t) \wedge \\
(\begin{state}(m, t) \models \text{dog_moves_to}(d) \lor \begin{state}(m, t) \models \text{dog_moves_to}(d, ds))
\]

dog_starts_straight\( (m, t, t_0, t_0) \)
\[
\exists d : \text{straight}, ds : \begin{beginning}(t_0) \leq \begin{beginning}(t), \begin{end}(t) \leq \begin{end}(t_0) \\
\neg \text{no_movement}(t_0, t) \wedge \\
(\begin{state}(m, t) \models \text{dog_moves_to}(d) \lor \begin{state}(m, t) \models \text{dog_moves_to}(d, ds))
\]

not_straight\( (t_0, t) \)
\[
\exists t_1, d : \text{notstraight}, ds \\
\begin{beginning}(t_0) \leq \begin{beginning}(t_1), \begin{end}(t_1) \leq \begin{end}(t) \wedge \\
(\begin{state}(m, t) \models \text{dog_moves_to}(d) \lor \begin{state}(m, t) \models \text{dog_moves_to}(d, ds))
\]

no_movement\( (t_0, t) \)
\[
\exists t_2, d : \text{nomove}, ds \\
\begin{beginning}(t_0) \leq \begin{beginning}(t_2), \begin{end}(t_2) \leq \begin{end}(t) \wedge \\
(\begin{state}(m, t) \models \text{dog_moves_to}(d) \lor \begin{state}(m, t) \models \text{dog_moves_to}(d, ds))
\]

dog_moves_backward\( (\text{exercise}(c, 1, vi)) \)
\[
\forall m \exists t_0, t_0 : t_0 \leq t_0 \wedge \\
\text{trace_part_of_interest}(m, \text{exercise}(c, 1, vi), t_0, t_0) \wedge \\
\text{dog_moves_backward}(m, t_0, t_0) \wedge \\
\exists t, d : \text{backwardmove}, ds : t_0 \leq t \leq t_e,
\begin{align*}
\text{state}(m, t) &= \text{dog\_moves\_to}(d) \\
\text{state}(m, t) &= \text{dog\_moves\_to}(d, ds)
\end{align*}

\text{nottstraight} is a movement from the set: [2,3,4,5,6,7,8,9,10].
\text{nomove} is a movement from the set: [1,2,3,4,5,6,7,8,9,10,11,12].
\text{leftmove} is a movement from the set: [9,10].
\text{rightmove} is a movement from the set: [2,3].
\text{backwardmove} is a movement from the set: [4,5,6,7,8].
\text{straight} is a movement from the set: [11,12,1].
\text{ds} is a dog speed from the set: [walk, trot, gallop, irregular].
Results

The route formalization is used to answer the research question: Is there a correlation between the circumstances of the task and the route the dog takes to the goody? It is expected that adding the alley in experiment D changed the route of the dog. If the alley is guiding, dogs will choose a straight route. If the alley adds more pressure, the route will avoid the alley and thus less straight routes are expected.

12.1. Subjects

205 dogs participated in the experiments. From these dogs, we had 95 dogs which had transcripts for the experiments C, D and E. From these 95 dogs a subset of 32 dogs was selected with the aim of having a fair distribution with respect to training level, age and gender. A low level of training means, none at all, puppy-or basic levels and a high level means GG1 or higher in Dutch training terms. Age was taken as adult or non-adult. For males adult means 2 years or older, for females this is 1.5 years. This is conform [7, 30]. So there were 16 males and 16 females, each group consisting of 8 dogs with low and 8 dogs with high training level. Non-adult dogs with high training level are hard to come by. Therefore 3 out of 16 non-adult dogs had a high training level and 13 out of 16 adult dogs had a high training level. Appendix A table A.1 summarizes the data of the selected dogs. It shows their age in months during the experiment. All dogs of this subset have used the same blindfold. There was no use of a wooden board for any of these dogs. So this analysis is free of the influence of using different blindfolds.

12.2. Analysis

For each dog, it was determined which route the dog took according to the formalization created in section 11.2. If the dog never got to the goody, it has no route. The results are shown in table 12.1. For each route is given the number of dogs that took this route during the trial. From these results are $\chi^2$ analysis was done. The results were statistically analyzed using the $\chi^2$ test. This test is chosen because we have categorical data and we want to do a hypothesis test. The null hypothesis is that paw departure is independent of performance. The alternative hypothesis is that paw departure and performance are dependent and there exist some kind of correlation. During this study a p value smaller than 0.05 is considered significant.

<table>
<thead>
<tr>
<th></th>
<th>left</th>
<th>right</th>
<th>straight</th>
<th>zig zag</th>
<th>backwards</th>
<th>none</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>4</td>
<td>2</td>
<td>12</td>
<td>26</td>
<td>26</td>
<td>9</td>
</tr>
<tr>
<td>D</td>
<td>5</td>
<td>6</td>
<td>12</td>
<td>40</td>
<td>12</td>
<td>8</td>
</tr>
<tr>
<td>E</td>
<td>0</td>
<td>2</td>
<td>18</td>
<td>8</td>
<td>3</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 12.1: Number of dogs taking a route per exercise

The null hypothesis used during the analysis is that the routes are independent of the task. The alternative hypothesis is that there is a dependence on the task. It shows that there is indeed a correlation between route and task circumstances. Between all task the significant correlation is ($\chi^2$=49.23, $p$<0.001). Between C and E there is a significant correlation ($\chi^2$=42.27, $p$<0.001). Between D and E there is a significant correlation ($\chi^2$=39.50, $p$<0.001). In table 12.1 it shows that the number of straight routes is much higher in experiment
E compared to C and D. So one can conclude that the reduced pressure in experiment E causes the dog to go straight towards its goal. There is also a significant correlation between exercise C and D ($\chi^2=16.69$, $p<0.05$). As shown in table 12.1, more zigzag routes are taken in experiment D. However, the number of dogs taking a left or straight route are roughly the same. The number of right routes are reduced. So the alley is reducing the number of backward routes and right routes, but introduces more zigzag routes. This indicates that the alley is indeed distracting and not guiding the dog straight towards the goody. This can be explained by the increase in pressure by adding the alley.
Conclusion and future work

This concluding chapter will go over the results and their meaning. It will also show future work that can be done to continue and improve this research and get more insight into dog cognition.

13.1. Conclusion

During the first part of this thesis, we tried to answer the question: Is using an automated checking software improving the observer reliability compared to humans annotating complex behaviour? The result show that there is indeed an improvement in interobserver reliability by using a formalized tool to check for complex behaviours. Important is the use of a tested ethogram and defining the basic behaviours clear and unambiguously. It shows that letting humans interpret complex behaviour introduces biases caused by misinterpreting the behaviours. Discussions on the interpretation can happen, but the result should be put into the transcript or this will introduce more differences. The TTL checker solves this problem by consistently checking the complex behaviour the same way based on basic behaviours. It shows that letting humans interpret complex behaviour introduces biases caused by misinterpreting the behaviours. Discussions on the interpretation can happen, but the result should be put into the transcript or this will introduce more differences. The TTL checker solves this problem by consistently checking the complex behaviour the same way based on basic behaviours. It should be noted that these basic behaviours are still not free of biases. They are still annotated by humans and thus susceptible to biases.

During the second part of this thesis, we tried to answer the question: Is the correlation between departing paw and performance influenced by circumstances? The result support the hypothesis that paw departure correlates to task performance as shown previously in [38]. Dogs departing with left perform significantly better than dogs departing with right ($\chi^2 = 38.83$, $p<0.001$). Meaning dogs that depart with their left paw, tend to have a good performance, while dogs departing with right have a bad performance. This holds for all tasks separately as well as the separate trials. Task dependency did not show in the task C and D ($\chi^2=0.09$, $p=76$). This shows that adding an alley to the search task does not significantly increase or decrease the performance. However not hiding the goody and showing the dog where the goody is, indeed correlates significantly with a better performance as expected. The results are ($\chi^2=20.36$, $p=<<0.001$) and ($\chi^2=22.44$, $p=<<0.001$) for a comparison between C and E and a comparison between D and E respectively. So to answer the question: the correlation itself is not influenced by the circumstances of the task. For each task, the left paw usually indicates a good performance and a right paw departure leads to bad performance. However the number of good and bad performances are influence by the circumstances of the task. It shows that removing all pressure of the blindfold and not hiding the goody under a cone will significantly improve the performance. However the results also showed that adding an alley did not significantly influence the performance compared to experiment C.

The final part of this thesis combined the previous parts into a new research in dog behaviour. It tried to answer the question: Is there a correlation between the circumstances of the task and the route the dog takes to the goody? The results show that the circumstances of each task indeed influenced the route the dog took. It showed similar results as the correlation between task and performance. Removing the pressure in experiment E causes dog to take more straight routes to the goody ($\chi^2=42.27$, $p=<<0.001$, $\chi^2=39.50$, $p=<<0.001$), which is expected, since they are not distracted. The influence of the alley in experiment D showed to be on the route dog takes and not on the performance. In experiment D, more dog took a zig zag route to the goody ($\chi^2=16.69$, $p=<<0.05$). This can be explained by the fact that the alley might be distracting the dog and not guiding the dog in a straight line to the goody. More research is needed to explain the influence of the alley.
on the dog behaviour.

13.2. Future work
Using the TTL checker software showed to be a fast way of checking behaviours, once one has the traces. It also showed to be more consistent in checking the same behaviour over and over again. This reduces biases introduced by humans. However, annotations of basic behaviours are still needed. Annotating every single behaviour takes a lot of time and is prone to biases by humans. For a future project, a automatic video translator can be made, who analysis video data frame by frame and annotates all behaviours visible. This will make annotating faster and more objective.

During the analysis it showed the correlations between paw departure and performance in exercise C, D and E. It also shows a correlation between circumstances and performance. Adding the alley in experiment D did not show a significant correlation with experiment C, however the circumstances in experiment E showed a strong correlation. Researching the route to dog takes, gave more insight in the influence of the alley on the dog. It shows the dog took more zigzag route, indicating that the alley might be viewed as distracting. The exact nature of the correlations is not explained by the data. Many other intermediate concepts may play a role, which can be considered in future research. These concepts can be analyzed using the TTL checker, since one can reuse the existing data on new research questions.
Subject data

This appendix will show the data of the selected dogs for each hypothesis.

<table>
<thead>
<tr>
<th>Dog number</th>
<th>Dog name</th>
<th>Breed</th>
<th>Gender</th>
<th>Age in months</th>
<th>Training level</th>
</tr>
</thead>
<tbody>
<tr>
<td>33</td>
<td>Max</td>
<td>Wirehaired Dachshund</td>
<td>male</td>
<td>38</td>
<td>high</td>
</tr>
<tr>
<td>36</td>
<td>Liza</td>
<td>Labrador Retriever</td>
<td>female</td>
<td>128</td>
<td>low</td>
</tr>
<tr>
<td>39</td>
<td>Tommie</td>
<td>Maltese</td>
<td>male</td>
<td>46</td>
<td>low</td>
</tr>
<tr>
<td>41</td>
<td>La Chouffe</td>
<td>Belgian Shepherd Malinois</td>
<td>female</td>
<td>48</td>
<td>high</td>
</tr>
<tr>
<td>74</td>
<td>Jan van het Heksewiel</td>
<td>Bullmastiff</td>
<td>male</td>
<td>60</td>
<td>high</td>
</tr>
<tr>
<td>81</td>
<td>Shadow</td>
<td>Australian Shepherd</td>
<td>male</td>
<td>37</td>
<td>high</td>
</tr>
<tr>
<td>83</td>
<td>Nando</td>
<td>German Dog</td>
<td>male</td>
<td>79</td>
<td>high</td>
</tr>
<tr>
<td>85</td>
<td>Amber</td>
<td>German Shepherd Dog</td>
<td>female</td>
<td>63</td>
<td>high</td>
</tr>
<tr>
<td>86</td>
<td>Ginger vom Finkenschlog</td>
<td>German Shepherd</td>
<td>female</td>
<td>36</td>
<td>high</td>
</tr>
<tr>
<td>165</td>
<td>Jenny</td>
<td>Galgo - Labrador</td>
<td>female</td>
<td>32</td>
<td>low</td>
</tr>
<tr>
<td>168</td>
<td>Erik</td>
<td>Boxer</td>
<td>male</td>
<td>13</td>
<td>low</td>
</tr>
<tr>
<td>169</td>
<td>Jeanny</td>
<td>Golden Retriever</td>
<td>female</td>
<td>16</td>
<td>low</td>
</tr>
<tr>
<td>171</td>
<td>Dana</td>
<td>Deutscher Schäferhund</td>
<td>female</td>
<td>15</td>
<td>low</td>
</tr>
<tr>
<td>173</td>
<td>Fatu</td>
<td>Rodesion Ridgeback</td>
<td>male</td>
<td>6</td>
<td>low</td>
</tr>
<tr>
<td>176</td>
<td>Donald</td>
<td>Boxer</td>
<td>male</td>
<td>73</td>
<td>high</td>
</tr>
<tr>
<td>177</td>
<td>Nick</td>
<td>Labrador-Silver-Boy</td>
<td>male</td>
<td>17</td>
<td>low</td>
</tr>
<tr>
<td>179</td>
<td>Jimmy</td>
<td>weisser schweizer Schäferhund</td>
<td>male</td>
<td>11</td>
<td>low</td>
</tr>
<tr>
<td>182</td>
<td>Othello</td>
<td>alt deutscher Schäferhund</td>
<td>male</td>
<td>8</td>
<td>low</td>
</tr>
<tr>
<td>184</td>
<td>Bernabas</td>
<td>Mali-chow-chow Mix</td>
<td>male</td>
<td>39</td>
<td>high</td>
</tr>
<tr>
<td>186</td>
<td>Ijosha</td>
<td>Enthubcher Sennenhond</td>
<td>male</td>
<td>25</td>
<td>low</td>
</tr>
<tr>
<td>189</td>
<td>Kira</td>
<td>Beagle</td>
<td>female</td>
<td>133</td>
<td>high</td>
</tr>
<tr>
<td>190</td>
<td>Roxy</td>
<td>Podenco Mix</td>
<td>female</td>
<td>53</td>
<td>high</td>
</tr>
<tr>
<td>192</td>
<td>Emma</td>
<td>Bernasennenhond</td>
<td>female</td>
<td>11</td>
<td>low</td>
</tr>
<tr>
<td>197</td>
<td>Bella</td>
<td>Westhighland white terrier</td>
<td>female</td>
<td>129</td>
<td>high</td>
</tr>
<tr>
<td>198</td>
<td>Mocca</td>
<td>Labrador retriever</td>
<td>female</td>
<td>136</td>
<td>high</td>
</tr>
<tr>
<td>199</td>
<td>Rowdy</td>
<td>Jack Russel/ Bordercollie/Australien Shepard</td>
<td>male</td>
<td>70</td>
<td>high</td>
</tr>
<tr>
<td>201</td>
<td>Cora</td>
<td>Labrador/Rodgeback</td>
<td>female</td>
<td>13</td>
<td>low</td>
</tr>
<tr>
<td>202</td>
<td>Louis</td>
<td>Bearded Collie</td>
<td>male</td>
<td>23</td>
<td>high</td>
</tr>
<tr>
<td>203</td>
<td>Ben</td>
<td>hybrid</td>
<td>male</td>
<td>13</td>
<td>low</td>
</tr>
<tr>
<td>204</td>
<td>Jule</td>
<td>hybrid</td>
<td>female</td>
<td>13</td>
<td>low</td>
</tr>
<tr>
<td>210</td>
<td>Layla</td>
<td>Farm Collie</td>
<td>female</td>
<td>15</td>
<td>low</td>
</tr>
<tr>
<td>213</td>
<td>Sam</td>
<td>Puggle</td>
<td>male</td>
<td>13</td>
<td>low</td>
</tr>
</tbody>
</table>

Table A.1: Demographics of the 32 dogs used for checks using the TTL software.
<table>
<thead>
<tr>
<th>Dog number</th>
<th>Dog name</th>
<th>Breed</th>
<th>Gender</th>
<th>Age in months</th>
<th>Training level</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>Tobias</td>
<td>Basset Artesien Normand</td>
<td>male</td>
<td>5</td>
<td>low</td>
</tr>
<tr>
<td>23</td>
<td>Tess</td>
<td>Border Collie</td>
<td>female</td>
<td>4</td>
<td>low</td>
</tr>
<tr>
<td>25</td>
<td>Cato</td>
<td>Labrador Retriever</td>
<td>female</td>
<td>23</td>
<td>low</td>
</tr>
<tr>
<td>29</td>
<td>Tjaarda</td>
<td>Standard Schnauzer</td>
<td>female</td>
<td>47</td>
<td>high</td>
</tr>
<tr>
<td>33</td>
<td>Max</td>
<td>Wirehaired Dachshund</td>
<td>male</td>
<td>38</td>
<td>high</td>
</tr>
<tr>
<td>36</td>
<td>Liza</td>
<td>Labrador Retriever</td>
<td>female</td>
<td>128</td>
<td>low</td>
</tr>
<tr>
<td>39</td>
<td>Tommie</td>
<td>Maltese</td>
<td>male</td>
<td>46</td>
<td>low</td>
</tr>
<tr>
<td>41</td>
<td>La Chouffe</td>
<td>Belgian Shepherd Malinois</td>
<td>female</td>
<td>48</td>
<td>high</td>
</tr>
<tr>
<td>52</td>
<td>Tasja</td>
<td>?</td>
<td>female</td>
<td>153</td>
<td>low</td>
</tr>
<tr>
<td>55</td>
<td>Anko</td>
<td>Jack Russell Terrier</td>
<td>male</td>
<td>41</td>
<td>low</td>
</tr>
<tr>
<td>57</td>
<td>Maia</td>
<td>Setter/Silken Windhound</td>
<td>female</td>
<td>47</td>
<td>high</td>
</tr>
<tr>
<td>60</td>
<td>Foemel</td>
<td>Wirehaired Dachshund</td>
<td>male</td>
<td>123</td>
<td>low</td>
</tr>
<tr>
<td>63</td>
<td>Fax</td>
<td>hybrid</td>
<td>female</td>
<td>36</td>
<td>low</td>
</tr>
<tr>
<td>66</td>
<td>Camille</td>
<td>Standard Poodle</td>
<td>male</td>
<td>38</td>
<td>low</td>
</tr>
<tr>
<td>68</td>
<td>Nina</td>
<td>Labrador Retriever</td>
<td>female</td>
<td>50</td>
<td>high</td>
</tr>
<tr>
<td>69</td>
<td>Gunnar</td>
<td>Broholmer</td>
<td>male</td>
<td>44</td>
<td>high</td>
</tr>
<tr>
<td>71</td>
<td>Muis</td>
<td>Standard Poodle</td>
<td>female</td>
<td>26</td>
<td>high</td>
</tr>
<tr>
<td>74</td>
<td>Jan van het Heksewiel</td>
<td>Bullmastiff</td>
<td>male</td>
<td>60</td>
<td>high</td>
</tr>
<tr>
<td>76</td>
<td>Astra</td>
<td>Dutch Shepherd</td>
<td>female</td>
<td>69</td>
<td>high</td>
</tr>
<tr>
<td>81</td>
<td>Shadow</td>
<td>Australian Shepherd</td>
<td>male</td>
<td>37</td>
<td>high</td>
</tr>
<tr>
<td>83</td>
<td>Nando</td>
<td>German Dog</td>
<td>male</td>
<td>79</td>
<td>high</td>
</tr>
<tr>
<td>86</td>
<td>Ginger vom Finkenschlog</td>
<td>German Shepherd</td>
<td>female</td>
<td>36</td>
<td>high</td>
</tr>
<tr>
<td>88</td>
<td>Ynka vom Finkenschlag</td>
<td>German Shepherd</td>
<td>female</td>
<td>64</td>
<td>high</td>
</tr>
<tr>
<td>94</td>
<td>Veitha</td>
<td>nova scotia dutch tolling retriever</td>
<td>female</td>
<td>25</td>
<td>high</td>
</tr>
<tr>
<td>96</td>
<td>Mitral</td>
<td>Border Collie</td>
<td>male</td>
<td>28</td>
<td>high</td>
</tr>
<tr>
<td>97</td>
<td>Beau</td>
<td>Podenco hybrid</td>
<td>male</td>
<td>55</td>
<td>high</td>
</tr>
<tr>
<td>98</td>
<td>Twister</td>
<td>Border Collie</td>
<td>male</td>
<td>55</td>
<td>high</td>
</tr>
<tr>
<td>100</td>
<td>Misty</td>
<td>Border Collie</td>
<td>female</td>
<td>43</td>
<td>high</td>
</tr>
<tr>
<td>165</td>
<td>Jenny</td>
<td>Galgo - Labrador</td>
<td>female</td>
<td>32</td>
<td>low</td>
</tr>
<tr>
<td>168</td>
<td>Erik</td>
<td>Boxer</td>
<td>male</td>
<td>13</td>
<td>low</td>
</tr>
<tr>
<td>169</td>
<td>Jeanny</td>
<td>Golden Retriever</td>
<td>female</td>
<td>16</td>
<td>low</td>
</tr>
<tr>
<td>171</td>
<td>Dana</td>
<td>Deutscher Schäferhund</td>
<td>female</td>
<td>15</td>
<td>low</td>
</tr>
<tr>
<td>173</td>
<td>Fuat</td>
<td>Rodesien Ridgeback</td>
<td>male</td>
<td>6</td>
<td>low</td>
</tr>
<tr>
<td>176</td>
<td>Donald</td>
<td>Boxer</td>
<td>male</td>
<td>73</td>
<td>high</td>
</tr>
<tr>
<td>177</td>
<td>Nick</td>
<td>Labrador-Silver-Boy</td>
<td>male</td>
<td>17</td>
<td>low</td>
</tr>
<tr>
<td>179</td>
<td>Jimmy</td>
<td>weisser schweizer Schäferhund</td>
<td>male</td>
<td>11</td>
<td>low</td>
</tr>
<tr>
<td>182</td>
<td>Othello</td>
<td>alt deutscher Schäferhund</td>
<td>male</td>
<td>8</td>
<td>high</td>
</tr>
<tr>
<td>184</td>
<td>Bernabas</td>
<td>Mali-chow-chow Mix</td>
<td>male</td>
<td>39</td>
<td>high</td>
</tr>
<tr>
<td>186</td>
<td>Iljoshia</td>
<td>Entlebucher Sennenhond</td>
<td>male</td>
<td>25</td>
<td>low</td>
</tr>
<tr>
<td>192</td>
<td>Emma</td>
<td>Bernasennenhond</td>
<td>female</td>
<td>11</td>
<td>low</td>
</tr>
<tr>
<td>197</td>
<td>Bella</td>
<td>Westhighland white terrier</td>
<td>female</td>
<td>129</td>
<td>high</td>
</tr>
<tr>
<td>199</td>
<td>Rowdy</td>
<td>Jack Russel/ Bordercollie/Australien Shepard</td>
<td>male</td>
<td>70</td>
<td>high</td>
</tr>
<tr>
<td>201</td>
<td>Cora</td>
<td>Labrador/Rodgeback</td>
<td>female</td>
<td>13</td>
<td>low</td>
</tr>
<tr>
<td>202</td>
<td>Louis</td>
<td>Bearded Collie</td>
<td>male</td>
<td>23</td>
<td>high</td>
</tr>
<tr>
<td>203</td>
<td>Ben</td>
<td>hybrid</td>
<td>male</td>
<td>13</td>
<td>low</td>
</tr>
<tr>
<td>204</td>
<td>Jule</td>
<td>hybrid</td>
<td>female</td>
<td>13</td>
<td>low</td>
</tr>
<tr>
<td>210</td>
<td>Layla</td>
<td>Farm Collie</td>
<td>female</td>
<td>15</td>
<td>low</td>
</tr>
<tr>
<td>213</td>
<td>Sam</td>
<td>Puggle</td>
<td>male</td>
<td>13</td>
<td>low</td>
</tr>
</tbody>
</table>

Table A.2: Demographics of the dogs used for testing the correlation between paw departure and performance.
Items needed for Experiments

This appendix contains a list of all items needed for the experiments. This appendix can be used as a checklist.

Items needed for the experiments

- 2 Cameras with extra batteries.
- Electricity, extension cord.
- 2 Tripods for the cameras.
- Measuring tape for measuring the experimental field.
- 3 cones used as cups to hide the goody.
- 12 cones numbered for the alley in experiment D.
- Cleaning supplies to clean the cones
- 7 flags, in order to restrict the experimental area and give the starting and end points of the experiment.
- sausages, enough for all dogs.
- sausage plate.
- Barbecue tongs to place the sausage.
- towel, to be used as a blindfold. Need to be big enough for every dog.
- Intake forms printed.
- pens.
- pencils.
- Long dog leash.
- Hands Free Belt (HFB), so people do not touch the leash.
- clipboard, to fill in the intake form.
- Score board for the trial indication.
- all trial and exercise numbers.
- daily schedule, when are people expected.
- 2 umbrella’s when it is raining to cover the camera.
- 2 chairs for the cameraman.
- small present for each dog and owner.
- poop bags

Besides items, it is also important to think about the people you need during the experiment. Keep in mind the weather on the day of recording and adjust the clothing. Also make sure there is a lunch for the people involved.

Persons needed

- 1 experimenter, who is responsible for the whole experiment.
- 2 persons who can operate the cameras.
- 1 person who receive the participants and explains the experiment before it is their turn to do the experiment.
This appendix shows for each exercise a list of sections.

C.1. Exercise A
- I: before sitting
- II: sitting
- III: handler steps with leg and back
- IV: before sitting
- V: sitting
- VI: handler steps with leg, all three steps and back.

C.2. Exercise C
- I: before sitting
- II: sitting, eyes covered and then uncovered
- III: before sitting
- IV: sitting
- V: taking off, the first step of the dog only
- VI: searching

C.3. Exercise D
- I: before sitting
- II: sitting, eyes covered and then uncovered
- III: before sitting
- IV: sitting
- V: taking off, the first step of the dog only
- VI: searching

C.4. Exercise E
- I: before sitting
- II: sitting
- III: before sitting
- IV: sitting
- V: taking off, the first step of the dog only
- VI: searching
Example route formalization

This appendix gives a step by step example on how the route formalization works. First it shows an example exercise as a transcript, then as a trace. Finally this appendix explains each step of the checking.

D.1. Transcript
Figure D.1 shows the route a dog takes during exercise C. During part V of the exercise, the first step of the dog is annotated. In this case this is the right paw. In part VI of the exercise, the searching is described. It shows that after walking in different directions and sniffing on the ground, the dog walks to the goody pylon and eats the goody.

D.2. Trace
From the transcript a trace is created. Figure D.2 shows the visualization of the trace. It shows that during exercise C part V, the dog moves his right paw forward. It also shows the direction changes of the dog during exercise C part VI.

D.3. Checking
Using the trace as input, we can now check which route the dog took. In this appendix we will check if the dog takes a left route step by step.

From section 11.2, we can see that to check the left route, we first call trace part of interest for part VI, to find values for \( tb_0 \) and \( te_3 \), which are the begin and end point of this part of the exercise.

In trace part of interest we find a time point \( tb_0 \) for which it holds that all time points before \( tb_0 \) are not part of exercise C part VI, looking at the trace, this would be time point 10, since everything before belongs to exercise C part V. Then we find timepoint \( te_3 \), by finding a point for which it holds that all timepoints after \( te_3 \) are not part of exercise C part VI. In the trace it shows to be timepoint 80.

Now we have the begin and end of the exercise. Next we need to know if the dog starts left at a certain time point \( t \) between \( tb_0 \) and \( te_3 \). To find this \( t \), we go over all timepoints between \( tb_0 \) and \( te_3 \) to find a point \( t \) where the dog moves left. If we find a point \( t \), we check if all timepoints between \( tb_0 \) and \( t \) it holds that the dog does not move backward, left, or right. Since this would cause the dog to start in a different direction. In this example, \( t \) would be time point 20, since the dog moves towards 10, which is considered a left movement. The timepoints before this is timepoint 10, where the dog moves towards 12. This is a straight movement and allowed. So the dog indeed starts left.

We have the begin and end of the exercise, and timepoint \( t \) where the dog starts left. Subsequently we have to make sure the dog is not zigzagging in between. So we search for a time point \( t_1 \) for which holds that for all timepoints before this, the dog only moves left or straight and after \( t_1 \) the dog only moves right or straight. If such timepoint does not exist, then this is not a left route. In our example we can find a timepoint \( t_1 \), namely \( t_1 = 50 \). After this the dog moves towards 2, which is right, but no other movements. Before this the dog moves to 12, 10, 9 and 11, which are only straight and left movements.

So, the dog is not zigzagging in between. We have to check if the dog does not move backwards during the exercise, since this would change the left route to a backward route. To check, we go over all timepoints be-
between tb0 and te3 and check if the dog does not move in a backward direction. In our example this is the case, the dog never moves backward.

Finally, to make sure the dog finish the exercise, we need to know if at some timepoint t4 between tb0 and te3, the dog either takes or eats the goody. In our case t4 is at 70, where the dog eats the goody.

Now we have successfully go through all the steps of the left route and we can thus conclude that our example was a left route. The other checks go in a similar way. If we would have checked for the right route, we had to check if the dog starts right, which is not the case. If we checked for a straight route, we had to check if the dog starts straight, this is true, since the first movement is in the direction 12. But then we had to check all timepoints in between to see if the dog moves straight. This is not the case, the dog also moves towards 10, 9 and 2. Finally the dog does not move backwards and thus this is not a backward route.

Figure D.1: Example transcript exercise C V and C VI.
Figure D.2: Example trace for exercise CV and CVI.
Bibliography


