

The use of Bayesian networks for combining forensic evidence in a  
Dutch criminal case

Sonja Scheer

January 30, 2006

# Contents

<b>Introduction</b>	<b>4</b>
<b>1 Bayesian Networks</b>	<b>5</b>
1.1 Definitions and properties	5
1.2 Hierarchy of propositions	8
1.3 How to construct a Bayesian Network	8
<b>2 Bayesian network of the Robbery</b>	<b>11</b>
2.1 Case description	11
2.2 Source level	11
2.2.1 Hairs	11
2.2.2 Scent	16
2.3 Activity level, stage 1	19
2.3.1 Hairs/Hat	20
2.3.2 Scent	21
2.3.3 Network stage 1	25
2.4 Activity level, stage 2	25
2.4.1 Hairs/Hat	26
2.4.2 Network stage 2	27
2.5 Activity level, stage 3	28
2.5.1 Hairs/Hat	28
2.5.2 Network stage 3	31
<b>3 Discussion</b>	<b>34</b>
<b>Bibliography</b>	<b>35</b>

This report was written during an internship at the Netherlands Forensic Institute, The Hague in the Netherlands in the period from October 2005 to January 2006. It was supervised by:

- dr. M.J. Sjerps, Netherlands Forensic Institute, The Hague, The Netherlands.
- Prof. dr. R.D. Gill, Mathematical Institute, University Utrecht, Utrecht, The Netherlands.



# Introduction

Investigating a crime and as a consequence convicting a suspect is a long process, which requires solid evidence against the suspect. Forensic scientists help the police and the judicial system with the investigation by searching the crime scene, the suspect and the victim, for the presence of evidence. This evidence can be fingerprints, hairs, fibres and a lot more. When they find something, it will be investigated, and determined from what source it is. This can be a person, animal or an object. One of the most important kinds of evidence is DNA, which can be determined from a bloodstain, hairs, saliva or semen. When we have the evidence and it points to a specific person, we want to know what the probability is of the suspect being guilty, given this evidence. But we usually have different kinds of evidence, one maybe weighing more than one piece and then we have to combine these. The aim of this report is to see, if in criminal cases different pieces of evidence can be combined into one value of evidence. We will investigate, if the use of Bayesian Networks is in any way helpful to do this. Bayesian Networks are used in forensic science, because they provide a clear way of representing uncertainty, unpredictability and imprecision between variables of interest to the scientist, see [17], [6] and [2] for the theory on this. Bayesian Networks combine probability theory and graph theory, in order to represent the situation in a logical and intuitive way. In this report, I will use the software program HUGIN [8], to build and analyse the Bayesian networks. We are interested in the likelihood ratio (LR), which is the ratio of the probability of the evidence given one hypothesis, usually of the prosecution, and the probability of the evidence given the other hypothesis, of the defence. We can do this by reading off the posterior probabilities in the hypotheses node in Hugin, after setting the priors and instantiating the evidence. Then we can use the odds form of Bayes' Theorem to determine the LR.

In Chapter 1, I will give the definitions and ideas of Bayesian Networks that we need, to understand the application, which will be handled in Chapter 2. The background on mathematics and genetics falls beyond the scope of this report, so for further reading we will refer to some books and articles. The application in Chapter 2 is an old Dutch robbery case, which had three types of evidence, namely a mtDNA profile, a partial autosomal DNA profile and scent identification line-ups by dogs. We will try to combine these pieces of evidence, with the help of Bayesian Networks, and investigate if we can give a reasonable value to its worth. Chapter 3 will be a discussion, containing ideas for further research and observations made during the process of investigating this topic.

# Chapter 1

## Bayesian Networks

In this chapter, first we will look at definitions and properties from the theory of probability and graphs, which we need to understand Bayesian Networks and their construction. On the basis of an example, we will explain these properties. For further references on graph theory, see [11], and probability theory, see [14]. In the second paragraph, we will explain the underlying ideas of the evaluation of scientific evidence and the last paragraph will contain an example of how to construct a Bayesian network.

### 1.1 Definitions and properties

First, we will give some properties from probability theory. In forensic science the way to weigh the evidence is done by means of something called *degree of belief*. This is an expert's belief if an event is more probable to occur under one of the two (or more) different hypotheses. The problem now is to give a measure of strength to that belief. In Bayesian inference this is done by assigning numbers and a set of rules, which have to hold for them. At the basis lies probability theory, given by a number of laws. I shall give a summary of this, in the context of degrees of belief, following [1]:

Take  $A$  and  $B$  to be two propositions, then it holds:

- $0 \leq \Pr(A) \leq 1$ .
- If  $A$  and  $B$  are mutually exclusive propositions (ie. only one of them can be true at the same time), then  $\Pr(A \text{ or } B) = \Pr(A) + \Pr(B)$ .
- $\Pr(\neg A) = \Pr(\text{not } A) = 1 - \Pr(A)$ .
- $B$  is relevant to  $A \Leftrightarrow \Pr(A | B, I) \neq \Pr(A | I)$ , where  $\Pr(A | B, I)$  denotes the degree of belief that proposition  $A$  is true, given that proposition  $B$  is assumed to be true, and given the background information  $I$ .
- $\Pr(A, B | I) = \Pr(A | I) \Pr(B | A, I) = \Pr(B | I) \Pr(A | B, I)$ .
- $\Pr(A | I) = \Pr(A, B | I) + \Pr(A, \neg B | I)$ , this rule is called *the extension of the conversation rule*.
- $B$  is relevant to  $A \Leftrightarrow \Pr(A | B, I) \neq \Pr(A | \neg B, I)$

**Odds form of Bayes' Theorem:**  $E$  is the evidence,  $H_i$  are the hypotheses.

$$\frac{\Pr(H_1 | E, I)}{\Pr(H_2 | E, I)} = \frac{\Pr(E | H_1, I)}{\Pr(E | H_2, I)} \times \frac{\Pr(H_1 | I)}{\Pr(H_2 | I)} \quad (1.1.1)$$

The left hand side is the odds in favour of  $H_1$ , conditional on the evidence  $E$ , this is what we call the *posterior odds* in favour of  $H_1$ . The first term on the right hand side is what we call the *likelihood ratio* (LR), the second term is the *prior odds* in favour of  $H_1$ . The likelihood ratio, also called  $V$ , is used to measure the value of the evidence with respect to two alternative hypotheses. This is used to update the prior odds,

when there is evidence  $E$  available. If  $V = 1$ , the evidence  $E$  has no effect and thus is not relevant.  $V > 1$  implies that the evidence  $E$  is more likely if  $H_1$  has occurred, than if  $H_2$  has occurred.  $V < 1$  implies that the evidence  $E$  is more likely if  $H_2$  has occurred, than if  $H_1$  has occurred.

A property we will use throughout this report is, that if we set the a priori probabilities of both the hypotheses equal to each other, we will get prior odds equal to one. This then implies that the posterior odds are equal to the likelihood ratio.

In the rest of this report, many propositions will concern values taken by random variables. If  $X$  is a random variable, then  $X = x$  is the proposition that  $X$  takes the value  $x$ . Furthermore, we write  $\mathbf{Y}$  for a set of random variables  $Y_1, \dots, Y_n$ , where  $\mathbf{Y} = \mathbf{y}$  stands for the proposition  $Y_1 = y_1, \dots, Y_n = y_n$ .

Next, we give some definitions and properties from graph theory. We will use the notation following [1] and [2].

**Definition 1.1.1** A *Directed acyclic graph or DAG* is a graph with all edges directed and no directed cycles, where a cycle is a directed path beginning and ending with the same node.

Here, a **path** is a sequence of edges not violating the directions, and a **directed path** is a path, with all edges directed, which we have in this context. So, this means that in a DAG, nodes are connected with arrows and you can't go from any point to itself, when following the direction of the arrows.

Now we are ready to combine the graph theory and the probability theory described here, and go to our main subject, Bayesian networks. In a Bayesian network the nodes in a corresponding DAG are random variables and an edge between two nodes determines a dependency of one of the variables on the other, this leads to the following definition:

**Definition 1.1.2** A DAG with a joint probability distribution  $\Pr(\cdot)$  over its variables, is a **Bayesian Network (BN)**  $\Leftrightarrow$  for all variables  $X$  in the DAG, for all sets  $Y$  s.t.  $\mathbf{DE}(X) \cap \mathbf{Y} = \emptyset$  and for all  $x, \mathbf{y}$  and  $z$ :

$$\Pr(X = x \mid \mathbf{PA}(X) = z, \mathbf{Y} = \mathbf{y}) = \Pr(X = x \mid \mathbf{PA}(X) = z),$$

where  $\mathbf{DE}(X)$  is the set of descendants of  $X$ , and  $\mathbf{PA}(X)$  is the set of graphparents of  $X$ . By a joint probability distribution over the variables, say  $A, B, C, D$  and  $E$ , we mean  $\Pr(A = a, B = b, C = c, D = d, E = e)$  such that  $\sum_{\text{all}(a,b,c,d,e)} \Pr(A = a, B = b, C = c, D = d, E = e) = 1$ .

Graphparents of node  $X$  are defined as all nodes with direct edges going into  $X$ . By the descendants of  $X$ , we mean the nodes which have a path from  $X$  to that node. The relationships between the nodes are specified by node probability tables for every variable. When a node has no parents, this will be an unconditional table, which we also will call the prior for this variable. When a node does have parents, the probability table will be conditional on the state the parents are in.

### Example of a Bayesian network

In Figure 1.1.1, we see an example of a Bayesian network. Here  $A, B, C, D$  and  $E$  are random variables. We see that  $A$  and  $B$  are graphparents of  $D$ . And conversily,  $D$  is a graphdescendent of  $A$  and  $B$ , but also  $E$  is a graphdescendent of  $A$  and  $B$ , because there is a path from  $A$  and  $B$  to  $E$ .

So for a Bayesian Network we have to determine the variables of interest (the nodes) and the relationships between them, using links between the nodes. Also the absence of a link has to be justified from Definition 1.1.2. A consequence of the definition of Bayesian networks is that calculation of the joint probability distribution becomes much easier. For a BN with a set of variables  $A_1, \dots, A_n$ , then it is a theorem (see [11]) that the joint probability distribution can be written as:

$$\Pr(A_1 = a_1, \dots, A_n = a_n) = \prod_i \Pr(A_i = a_i \mid \mathbf{PA}(A_i) = p_i),$$

for all  $a_1, \dots, a_n, p_1, \dots, p_n$  Using this in our example will give:

$$\begin{aligned} \Pr(A = a, B = b, C = c, D = d, E = e) &= \Pr(A = a) \Pr(B = b) \Pr(C = c \mid A = a) \Pr(D = d \mid A = a, B = b) \\ &\times \Pr(E = e \mid D = d), \end{aligned}$$

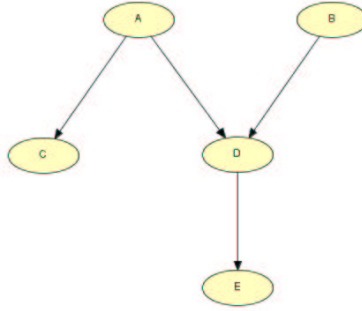


Figure 1.1.1: Example of a Bayesian Network.

for all  $a, b, c, d$  and  $e$ , where the first two terms on the right hand side are priors of  $A$  and  $B$ , respectively. The other ones are conditional on the parents. When  $A, B, C, D$  and  $E$ , are binary random variables, so they take values zero or one, we can also look at them as propositions. So ‘ $A = 1$ ’ corresponds to an underlying proposition, say  $P$ , being true, and ‘ $A = 0$ ’ corresponds to the same proposition being false.

The main advantage of the use of Bayesian Networks is that we can, in a coherent way, combine simpler parts, also Bayesian Networks, into a large Bayesian Network with the help of probability theory. In this way, we can look at a certain piece of evidence and its influence on the hypothesis of interest, and then putting it together with other pieces of evidence. For constructing these parts, we have three main connections in the BN’s, which can be combined to get bigger BN’s. These three connections are: Let  $X, Y$  and  $Z$  be three nodes in a DAG. Then  $X, Y$  and  $Z$  are in

- **serial connection** if  $X \rightarrow Y \rightarrow Z$
- **diverging connection** if  $X \leftarrow Y \rightarrow Z$
- **converging connection** if  $X \rightarrow Y \leftarrow Z$

Applying this to example 1.1.1, gives that:

- $A, D$  and  $E$  are in serial connection, i.e.  $A \rightarrow D \rightarrow E$
- $C, A$  and  $D$  are in diverging connection, i.e.  $C \leftarrow A \rightarrow D$
- $A, D$  and  $B$  are in converging connection, i.e.  $A \rightarrow D \leftarrow B$

**Definition 1.1.3** Let  $\mathbf{X}, \mathbf{Y}$  and  $\mathbf{Z}$  be disjoint subsets of variables in a DAG. Then  $\mathbf{X}$  and  $\mathbf{Y}$  are **d-separated** given  $\mathbf{Z} \Leftrightarrow \forall$  path between a variable  $X$  in  $\mathbf{X}$  and a variable  $Y$  in  $\mathbf{Y}$ , the path contains:

- either a serial connection  $\rightarrow Z \rightarrow$ , or a diverging connection  $\leftarrow Z \rightarrow$ , such that the middle node  $Z$  belongs to  $\mathbf{Z}$ , i.e.  $Z$  is instantiated, which means it changes from unknown to known.
- or a converging connection  $\rightarrow W \leftarrow$ , such that the middle node  $W$  does not belong to  $\mathbf{Z}$  and no descendant of  $W$  belongs to  $\mathbf{Z}$ , i.e.  $W$  and its descendants have not received evidence.

We have followed here the notations of [18]. In example 1.1.1, we have the following d-separation properties:

- $B$  and  $C$  are d-separated
- $B$  and  $C$ , **given**  $D$ , are **not** d-separated
- $B$  and  $C$ , given  $D$  **and**  $A$ , are d-separated

**Theorem 1.1.1** Given a DAG and  $A_1, \dots, A_n$  random variables, it holds: Sets of nodes  $\mathbf{X}$  and  $\mathbf{Y}$  are d-separated given  $\mathbf{Z} \Leftrightarrow \forall$  choices of  $\Pr(A_i \mid \mathbf{PA}(A_i))$ :  $\mathbf{X}$  is independent of  $\mathbf{Y}$  given  $\mathbf{Z}$ .

For the proof see [11].

We will see an example of the use of d-separation in section 1.3, where we use it to construct a Bayesian network.

## 1.2 Hierarchy of propositions

Now we look at the forensic point of view of the use of Bayesian networks in the matter of evaluating the evidence and the question whether or not the suspect is guilty. According to [1], there are three principles which we have to take into account when we are evaluating scientific evidence, namely that:

- we always have to have at least one alternative hypotheses when we are evaluating evidence.
- this evaluation considers the probability of the evidence given these hypotheses, so we are interested in the following:  $\Pr(E | H_p)$  and  $\Pr(E | H_d)$ , with  $E$  the evidence and  $H_p$  and  $H_d$  the hypotheses, where in general  $H_p$  stands for the hypothesis of the prosecution, and  $H_d$ , for the defence hypothesis.
- this evaluation always works within a framework of circumstances  $I$ , which contains the case circumstances and the available background information. In the notation of the conditional probabilities it is usually omitted.

When we have a case with 2 propositions, it has to hold that they are mutually exclusive, this means that there is no overlap in the definition of the two. But they don't have to be mutually exhaustive, which means that their union doesn't have to contain all the possibilities. We can expand this to more than two propositions in a logical way.

In [1], they use an approach to construct Bayesian networks, which we will use here also. This is based on the idea that we are interested in hypotheses, which can go from 'the suspect left the stain', or not, to 'the suspect is the offender', or not. So we have different kinds of hypotheses, each of which describes a different type of situation. We call this the **hierarchy of propositions**. There are three levels in this hierarchy:

**Source level:** At this level, the only thing you have to investigate are the samples from the scene of the crime and the control samples of the suspect(s). The hypotheses on this level only take into account the evidence you get from investigating this. For example, 'the stain came from the suspect', or not.

**Activity level:** At this level, the proposition has to contain an action, so on the activity level you get extra conditions to consider, such as contact and transfer of material to and from the suspect and/or scene. For example, 'the suspect left the hat at the scene of the crime', or not.

**Crime level:** This is the level of interest to the judge, but for a scientist to conclude on this level, he has to have a solid body of circumstantial information, which often is not the case, and he has to make assessments of a juridical nature, which is not allowed. So scientists are mostly restricted to the source or activity level. For example, 'the suspect committed the robbery', or not.

This will give us a tool to start at the bottom, with only the evidence that has to be linked to the source, and work our way up to the question the judge has to answer, namely whether or not the suspect committed the crime.

## 1.3 How to construct a Bayesian Network

Before we can construct a network, we have to note that a Bayesian Network does not represent the flow of information, but serves as a direct representation of a part of the real world. For the construction of Bayesian Networks it is very important how the hypotheses are formulated, because a little change in the definition can change the whole network. Furthermore it is important in what context the network is placed. Another feature is, that the construction depends on the person who constructed it. It is influenced by the expert's personal perception and his experience. This will be implemented in the network and in the (conditional) probability tables.

We will now look at an example of how to construct a network. From the previous section, we have seen that the hypothesis of interest to the judge is at the crime level, for example, whether or not the suspect is the offender. This is the Hypothesis node, denoted  $\mathbf{H}$ , with states  $H_p$  and  $H_d$ , respectively. Now we want to make a link to the stain found at a crime scene, for example. How can we do this? First, we have to know if the stain came (or not) from the *offender*, this is what we call the *relevance* of the stain. Denote this node

by  $\mathbf{R}$ , with states  $R$  if it is true and  $\neg R$  if it's not true. Secondly, we have to look at the probability that the stain came from the *suspect*, to determine this we have to know whether or not the suspect is the offender **and** whether or not the stain was left by the offender. So if we call the event, that the suspect left the stain or not,  $\mathbf{A}$ , with states  $A$  if it is true and  $\neg A$  if it's not true, then there has to be a link from  $\mathbf{H}$  and  $\mathbf{R}$  to  $\mathbf{A}$ . There will be no link between  $\mathbf{H}$  and  $\mathbf{R}$ , because knowing whether or not the stain came from the offender will not give any information about the event of the suspect being the offender, or vice versa. Finally, we want to put the evidence in the network. Suppose we took a DNA sample of the crime stain and from the suspect, the evidence node  $\mathbf{E}$  could have states: 'the DNA-profile of the suspect matches the DNA-profile of the crime stain', called  $E$  or not, called  $\neg E$ . This node depends on the node  $\mathbf{A}$ , and not directly on  $\mathbf{H}$  and  $\mathbf{R}$ . This means that if we know the state  $\mathbf{A}$  is in, knowing the states of  $\mathbf{H}$  and  $\mathbf{R}$  will not add any information on whether or not  $\mathbf{E}$  holds, this is called the d-separation property, explained in section 1.1. So the network will be as shown in Figure 1.3.1.

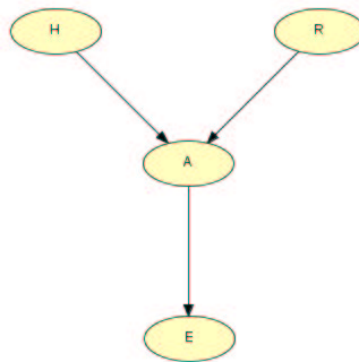


Figure 1.3.1: Another example of a Bayesian Network.

The probability tables will now follow from the network. Because we are interested in the likelihood ratio of the evidence given the hypotheses, we set the prior probabilities of  $H_p$  and  $H_d$  both to 0.5. Now the likelihood ratio will be equal to the posterior odds of  $H_p$  versus  $H_d$ . So when the evidence is instantiated, which means that the evidence node  $\mathbf{E}$  is set to what we know, we can read off the posterior odds, and thus the likelihood ratio, in the node  $\mathbf{H}$ .

The table of the hypothesis node is given in Table 1.3.1.

$H_p$	0.5
$H_d$	0.5

Table 1.3.1: Prior probabilities of the hypotheses node H, 'the suspect is the offender'.

In the node of the relevance we have the parameter  $\Pr(R) = r$ , which we have to determine. So the priors of this node  $\mathbf{R}$  are given in Table 1.3.2

$R$	$r$
$\neg R$	$1 - r$

Table 1.3.2: Prior probabilities of the relevance of the stain.

For the node  $\mathbf{A}$  we have to determine four conditional probabilities:

- probability that the stain came from the suspect, given that the stain came from the offender **and** the suspect is the offender, this is 1.

- probability that the stain came from the suspect, given that the stain **didn't** come from the offender **and** the suspect is the offender, this is 0.
- probability that the stain came from the suspect, given that the stain came from the offender **and** the suspect is not the offender, this is also 0.
- probability that the stain came from the suspect, given that the stain **didn't** come from the offender **and** the suspect is not the offender, this is the probability that the stain is from the suspect by accident, call this  $p$ .

This will give the following table with  $p$  unknown:

<b>H</b>	$H_p$		$H_d$	
<b>R</b>	$R$	$\neg R$	$R$	$\neg R$
$A$	1	0	0	$p$
$\neg A$	0	1	1	$1 - p$

Table 1.3.3: Conditional probability table of **A** given **H** and **R**.

The conditional probability table of the node **E** given **A**, will contain the following:

- probability that the DNA-profiles of the stain and the suspect match, given that the suspect left the stain, this will be 1 (assuming there occur no typing errors etc.).
- probability that the DNA-profiles of the stain and the suspect match, given that the suspect didn't leave the stain, this means the probability that someone with the same profile left the stain, this is the profile probability  $\gamma$ .

<b>A</b>	$A$	$\neg A$
$E$	1	$\gamma$
$\neg E$	0	$1 - \gamma$

Table 1.3.4: Conditional probability table of the evidence given **A**.

In a specific case, we can estimate  $\gamma$  from population genetic databases, so the only two unknown parameters are  $r$  and  $p$ . Now, playing with the values of these two, we can investigate with Hugin the influence of them on the likelihood ratio, which we won't do in this example. Now that we have seen an (easy) example of constructing a Bayesian Network, the next chapter will continue with a real criminal case, where we want to find out if Bayesian networks can be used to combine different kind of pieces of evidence to get just one likelihood ratio. We will do this in steps, following the hierarchy of propositions.

## Chapter 2

# Bayesian network of the Robbery

In this chapter we will give an insight in the use of Bayesian networks in Hugin on the basis of an example. This example describes a real criminal case from the Netherlands, in which there are a number of different kinds of evidence, which have to be combined. Then the likelihood ratio has to be determined.

### 2.1 Case description

A robbery took place at a company in the Netherlands. The robbers, there were two seen on the monitor at the company, took a bulldozer, which was stolen a weekend before the crime, and drove it through the wall of the building. Next, they drove a Mercedes truck, also stolen, through the hole, and within minutes they packed it with sacks containing money, of a total of 13 million Dutch guilders. The robbers drove off in the Mercedes, and the police pursued them. This resulted in a dangerous chase, with the robbers shooting at the police cars, injuring one police woman in the shoulder. The Mercedes was found with the engine still running. The robbers got into an other car, with this they crashed a roadblock, but got away. The police later found the car, with most of the money still there, but the robbers nowhere near. The police found several items of evidence: a hat in the bulldozer, with hairs in it and a glove. There were six suspects apprehended, but there were three main suspects, the others were released. The investigation concentrated on these three suspects. They knew each other and bloodsamples were taken from them, to compare with the DNA of the hairs in the hat. Also the help of search dogs was used on the seats of the bulldozer and Mercedes.

In this report I will treat a simpler form of this crime. I will focus on one suspect, called A, and if he committed the robbery, sat in the bulldozer and left the hairs and scent in the bulldozer. So I will leave the Mercedes, the glove and the other suspects out of consideration.

### 2.2 Source level

I will begin at the source level. In this case the traces are hairs found in a hat, which was left on the seat of the bulldozer, and a scent which was left in the bulldozer.

#### 2.2.1 Hairs

Here we will concentrate on the hairs found in the hat. We are at the source level, so the hypotheses have to describe an event only taking in account the origin of the hairs found. So take as the hypotheses:

$H_1$ : the hairs in the hat are from A

$H_2$ : the hairs in the hat are from a person who is related through the maternal line to A

$H_3$ : the hairs in the hat are from a person who is not related through the maternal line to A

So this will give the hypothesis node, **Hairs from A** in the Bayesian network. The evidence on this level consists of a partial DNA profile and a mitochondrial DNA, also called mtDNA, profile of the hairs.

So first I will give a little bit of genetics to understand what is happening. Each human being has 23 pairs of *chromosomes* in a cell, 22 of which are called autosomes, and one pair are the sex chromosomes. For a given individual, one chromosome of the pair comes from the DNA of the mother and the other one comes from the DNA of the father. A *locus* is a specific place at a chromosome, the DNA at this locus we will call his *gene*. DNA can take different forms at a specific locus, and these options we will call *alleles*. So a locus corresponds to a (random) variable and the allele to its realised state. A The *genotype* will then be the combination of the 2 alleles at a specific locus, with the *phenotype* what you physically see. In a DNA laboratory they use DNA markers, this is a known locus where the allele can be identified. Short Tandem Repeats (STR) are markers with alleles given by integers. So when we have 11/13 at a specific locus, this means that a specific sequence of base pairs is 11 times repeated at one chromosome and 13 times at the other chromosome. For further reference on DNA, we refer to [4].

The reason we have to include the second hypothesis, is because mtDNA inherits from mother to all her children, without changes (assuming there are no mutations), so everyone in the maternal line will have (almost) the same mtDNA profile. The mtDNA of 3 hairs found are the same, and we will call this string  $\Gamma$ . The DNA profile was found at 2 markers, namely HumFES 11/13, and HumTHO1 8/9.3. The comparison of the hairs found and the bloodsample that A gave showed that:

1. the mtDNA of the suspect A has no differences with the mtDNA of string  $\Gamma$
2. A fits into the partial DNA profile found in the hairs

The police didn't compare specific hair characteristics, probably because they already had DNA evidence, but we don't know for sure why they didn't do this. Now we can define our evidence nodes, with their states:

**DNA trace:**

$x_d$ : partial autosomal DNA profile of the hairs found is HumFES 11/13, and HumTHO1 8/9.3

$\neg x_d$ : partial autosomal DNA profile of the hairs found has any other profile

**mtDNA trace:**

$x_m$ : mtDNA profile of the string  $\Gamma$

$\neg x_m$ : mtDNA profile of any other string

**DNA from A:**

$x_d$ : partial autosomal DNA profile of the suspect A is HumFES 11/13, and HumTHO1 8/9.3

$\neg x_d$ : partial autosomal DNA profile of the suspect A has any other profile

**mtDNA from A:**

$x_m$ : A has mtDNA profile of the string  $\Gamma$

$\neg x_m$ : A has mtDNA profile of any other string

The hypothesis  $H_2$  will complicate the determination of the partial DNA profile given this hypothesis. So first I will assume that all the persons related through the maternal line to A are excluded. Now the hypotheses will become:

$H_1^*$ : the hairs in the hat are from A

$H_2^*$ : the hairs in the hat are not from A

Having defined all the ingredients, we can construct the Bayesian network, so first we have to determine the links between the three nodes. We assume that the DNA and the mtDNA evidence are independent. This may not strictly be true, but from a biological theory we expect it to be true to a close approximation, see [4], page 317. (Investigating this dependence may be of interest for further research). As a result, there will be no link between the two nodes, **DNA** and **mtDNA**. To determine the probabilities of the **DNA** and **mtDNA** nodes, we have to know, whether or not the hairs are from A. So there is a link from **Hairs from A** to **DNA**, and from **Hairs from A** to **mtDNA**. The network can be seen in Figure 2.2.1.

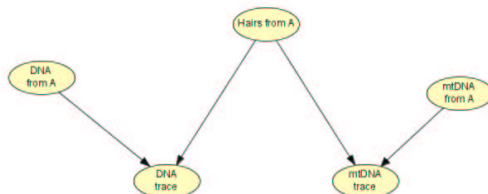


Figure 2.2.1: Bayesian Network for source level hypothesis of the hairs.

We now have to fill in the (conditional) probability tables, starting with the priors for the event that the hairs found in the hat are from A or not. These we will set on 0.5 for both events, so we have no background of whether or not the hairs are from A, an advantage of this is, that it holds that the posterior probabilities are equal to the likelihood ratio, which we are interested in. When we take the priors unequal to 0.5, we can use the odds form of Bayes' theorem (section 1.1), to determine the likelihood ratio.

Given that the hairs are not from A, we have to determine the probability that another person has the same profile as the hairs found. This is called the *profile probability* which means the probability with which an unknown person would have a particular profile in a relevant population, where the relevant population depends on the background of the offender. Here we have to determine the profile probability  $\rho_1$  of the partial profile HumFES 11/13, and HumTHO1 8/9.3, and the profile probability  $\rho_2$  of the mtDNA of string  $\Gamma$ . From research results, we know that the frequency of HumFES 11/13 is about 4%, and of HumTHO1 8/9.3 is about 7%. The frequency of the mtDNA of the string  $\Gamma$  has been checked with 2 groups, one being a group of 163 British Caucasians and 132 North-American Caucasians. I will take the group of British Caucasians as a reference, of this group 10 people had the same mtDNA as the string  $\Gamma$ . (These numbers are taken from the report of this case, and could have been changed). Assuming linkage equilibrium, we get  $\rho_1 = 0.07 \times 0.04 = 0.0028$  and  $\rho_2 = 10/163 = 0.06135$ . Now we can set the tables:

$H_1^*$	0.5
$H_2^*$	0.5

Table 2.2.1: Prior probabilities of **Hairs from A**.

$x_d$	1
$\neg x_d$	0

Table 2.2.2: Prior probabilities of **DNA from A**.

$x_m$	1
$\neg x_m$	0

Table 2.2.3: Prior probabilities of **mtDNA from A**.

<b>Hairs from A</b>	$H_1^*$		$H_2^*$	
<b>DNA from A</b>	$x_d$	$\neg x_d$	$x_d$	$\neg x_d$
$x_d$	1	0	$\rho_1 = 0.0028$	$\rho_1 = 0.0028$
$\neg x_d$	0	1	$1 - \rho_1 = 0.9972$	$1 - \rho_1 = 0.9972$

Table 2.2.4: Conditional probability table of **DNA trace** given **Hairs from A** and **DNA from A**.

<b>Hairs from A</b>	$H_1^*$		$H_2^*$	
<b>mtDNA from A</b>	$x_m$	$\neg x_m$	$x_m$	$\neg x_m$
$x_m$	1	0	$\rho_2 = 0.0028$	$\rho_2 = 0.0028$
$\neg x_m$	0	1	$1 - \rho_2 = 0.9972$	$1 - \rho_2 = 0.9972$

Table 2.2.5: Conditional probability table of **mtDNA trace** given **Hairs from A** and **mtDNA from A**.

We set the evidence, so we know that  $x_d$  and  $x_m$  hold, both for the trace as for the suspect A, and the values of  $\rho_1$  and  $\rho_2$  are fixed. We are interested in the likelihood ratio, that is  $\frac{\Pr(E|H_i)}{\Pr(E|H_j)}$ , and we can determine this with the help of the odds form of Bayes' Theorem (section 1.1). Rewriting this, gives us the following:

$$\text{Likelihood ratio}(LR) = \frac{\text{Posterior odds}}{\text{Prior odds}},$$

with the **prior odds** =  $\frac{\Pr(H_i)}{\Pr(H_j)}$ , and the **posterior odds** =  $\frac{\Pr(H_i|E)}{\Pr(H_j|E)}$ . If there are only two hypotheses, which must be exhaustive, we get:

$$LR = \frac{\Pr(H_1 | E)}{1 - \Pr(H_1 | E)} / \frac{\Pr(H_1)}{1 - \Pr(H_1)}.$$

The posterior probability in favour of **Hairs from A**, we read off in the network, using the program Hugin ([8]). For these fixed values of  $x_d$ ,  $x_m$ ,  $\rho_1$  and  $\rho_2$ , the LR is fixed. Calculating the LR will give:

$$LR = \frac{\Pr(x_m, x_d | \text{Hairs are from A})}{\Pr(x_m, x_d | \text{Hairs are not from A})} \approx \frac{1}{0.0028 \times 0.06135} \approx 5821$$

When we read off the posterior probability in Hugin, (when we take prior probabilities of 0.5) we see that it is 0.99982825 in favour of the hairs being from A, and 0.00017175 in favour of the hairs being not from A. This will give a LR of  $\frac{0.99982825}{0.00017175} \approx 5821$ . But if we don't take as many decimals, but instead, say 0.9998, then we get  $LR \approx 4999$ . This is a common error, but what happens here? For the determination of the LR, we have to have not only the posterior probability, but also its complement. So we have to be careful that we have enough significant numbers after the nines, and as a consequence for the complement, when we want to compute the LR to a certain number of digits. In the remainder of this report we acknowledged this fact and used accurate enough posterior probabilities, read off in Hugin, to determine the LR.

Now we shall look at what happens if we drop the assumption that all the persons related through the maternal line to A are excluded, here we assume Hardy-Weinberg and linkage equilibrium. We get the following three hypotheses:

$H_1$ : the hairs in the hat are from A,

$H_2$ : the hairs in the hat are from a person who is related through the maternal line to A,

$H_3$ : the hairs in the hat are from a person who is not related through the maternal line to A.

$H_1$	$1/3$
$H_2$	$1/3$
$H_3$	$1/3$

Table 2.2.6: Prior probabilities of **Hairs from A**.

<b>Hairs from A</b>	$H_1$		$H_2$		$H_3$	
<b>DNA from A</b>	$x_d$	$\neg x_d$	$x_d$	$\neg x_d$	$x_d$	$\neg x_d$
$x_d$	1	0	$p$	$p$	$\rho_1 = 0.0028$	$\rho_1 = 0.0028$
$\neg x_d$	0	1	$1-p$	$1-p$	$1-\rho_1 = 0.9972$	$1-\rho_1 = 0.9972$

Table 2.2.7: Conditional probability table of **DNA trace** given **Hairs from A** and **DNA from A**.

We take all the priors equal, this has as a consequence that we can read off the posterior odds of a combination of two out of three, and this then will be equal to the LR. Now we have to redetermine the conditional probability tables of **DNA** and **mtDNA**, given **Hairs from A**.

Here  $p$  is the probability that someone related to A through the maternal line, has the same partial DNA profile. We will only allow a brother of A in this case. The probability that a brother has genotype AB, when A has genotype AB at a specific locus is:  $\frac{2p_A p_B + p_A + p_B + 1}{4}$ , where  $p_A$  and  $p_B$  are the allelfrequencies of A and B, respectively [4]. In our case, we have two loci, so we will multiply the results of the two, assuming that they are independent. We use the allelfrequencies used in the report. These are for the marker FES,  $p_{11} = 0.492$  and  $p_{13} = 0.033$ , and for the marker THO1,  $p_8 = 0.111$  and  $p_{9.3} = 0.310$ . Then we get:

$$\begin{aligned}
p &= \text{Pr}(\text{brother also has partial profile FES 11/13, and THO1 8/9.3}) \\
&= \frac{2p_{11}p_{13} + p_{11} + p_{13} + 1}{4} \times \frac{2p_8p_{9.3} + p_8 + p_{9.3} + 1}{4} = 0.389368 \times 0.372455 \\
&= 0.145022 \approx 14.5\%
\end{aligned}$$

We assume that there are no mutations in the mtDNA, so the brother of A, will have the same mtDNA profile as A, with probability one.

<b>Hairs from A</b>	$H_1$		$H_2$		$H_3$	
<b>mtDNA from A</b>	$x_m$	$\neg x_m$	$x_m$	$\neg x_m$	$x_m$	$\neg x_m$
$x_m$	1	0	1	0	$\rho_2 = 0.0028$	$\rho_2 = 0.0028$
$\neg x_m$	0	1	0	1	$1-\rho_2 = 0.9972$	$1-\rho_2 = 0.9972$

Table 2.2.8: Conditional probability table of **mtDNA trace** given **Hairs from A** and **mtDNA from A**.

With these changes, we have the following LR's:

- LR  $H_1/H_2 \approx 7$
- LR  $H_1/H_3 \approx 5821$
- LR  $H_2/H_3 \approx 844$

We see here that the LR of the hypothesis **Hairs are from A** versus **Hairs are from someone not related to A through the maternal line**, is the same with or without taking into account that the hairs could have come from a brother. But here we can see, that the brother also has a large LR in favour of the hairs being from him versus from someone who is not related through the maternal line. In [4], we find all kinds of formulas for the dependence of the genotypefrequencies on the relatedness of persons. We won't treat any other situations in this report. In the rest of the report, we will exclude that a person related to

A through the maternal line, has anything to do with the crime, but changing the states and probability of mtDNA in this matter as described above, will give a solution for this.

### 2.2.2 Scent

Before we can construct the Bayesian network of the scent on the source level, we have to know something about the way a scent identification test is performed, or more importantly, the assumptions which are made. We can find this in [15] and [16].

In this case we want to compare the scent on the seat of the bulldozer with the scent of the suspect A. This was done with two dogs. They used a scentcloth and rubbed it over the seat of the bulldozer, then the dog got to smell at the cloth and he has to pick out the origin of the scent by walking along a line of tubes with the cloths of the scents of six different people, among which the suspect's. Both dogs picked out the suspect as the origin of the scent, so this is our evidence. We now have to be careful that we don't overestimate the result that both dogs picked him out, because the 2 dogs are not independent. But two dogs picking him out, should make the result more reliable. They make a number of assumptions in [15] and [16], which we will now state:

- every person has a unique scent, in any case that the probability of 2 people in an arbitrary group of 6, with the same scent, is sufficiently low.
- the dog is well trained in smelling differences.
- the scents are stable enough to link a scentcloth to a suspect at some time after the crime.
- different parts of a human body smell the same.

We will define the hypotheses we need to construct the network: we are at the source level, so the hypotheses have to contain a source of the scent. We take following source node: **Scent from A**, with states:

**Yes:** the scent is from A

**No:** the scent is not from A

The dogs are trained to compare the scents on the scentcloth and the suspect, so the evidence nodes will be **Dog i**,  $i=1,2$ , with states:

**Yes:** Dog i points out the suspect as the source of the scent in the bulldozer

**No:** Dog i doesn't point out the suspect as the source of the scent in the bulldozer

#### Dogs independent

This means that there will be no link between the two dogs. There will be links from the node **Scent from A** to the two dogs, because we have to know if the scent is from A or not, to use the results described in Table 2.2.9. The Bayesian network will then look like this:

For the prior probabilities of the **Scent from A**, we again will take 0.5. For the determination of the conditional probabilities of the dogs given the scent is from A or not, there has been done some research. The conditions were the same as the way they use it now with an investigation. They found the following result showed in Table 2.2.9

	suspect is the offender	suspect is not the offender
Dog points out the suspect	31% correct positive	7.3% false positive
Dog points out an extra	60.6% error	80.5% error
Dog points out nobody	8.4% miss	12.2% correct negative

Table 2.2.9: Results of the research of the reactions of the dogs.

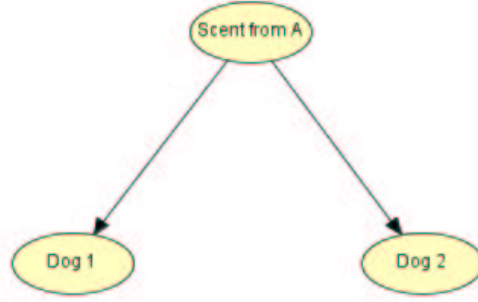


Figure 2.2.2: Bayesian network when the dogs are independent.

With these results we can fill in the conditional probability tables of Dog 1 and Dog 2, and the unconditional probability table of **Scent from A**, as can be seen in Tables 2.2.10 and 2.2.11. Where Dog 2 has the same conditional probabilities as Dog 1, in the case they are independent.

Yes	0.5
No	0.5

Table 2.2.10: Prior probabilities of **Scent from A**.

<b>Scent from A</b>	Yes	No
Yes	0.31	0.073
No	0.69	0.927

Table 2.2.11: Conditional probability table of **Dog 1** given **Scent from A**.

**Results:**

Letting the dogs compare the scent of the scentcloth contained from the bulldozer to the scent of A, gave for both dogs a recognition. So the evidence consists of two dogs matching the scent of the cloth to the scent from A. Knowing this we can instantiate the two evidence nodes to ‘Yes’.

For the determination of the LR, we do the following:

$$LR = \frac{\Pr(\mathbf{Dog\ 1} = \text{Yes}, \mathbf{Dog\ 2} = \text{Yes} \mid \text{Scent is from A})}{\Pr(\mathbf{Dog\ 1} = \text{Yes}, \mathbf{Dog\ 2} = \text{Yes} \mid \text{Scent is not from A})} = \frac{0.31 \times 0.31}{0.073 \times 0.073} \approx 18$$

$$LR = \frac{\Pr(\mathbf{Dog\ 1} = \text{Yes} \mid \text{Scent is from A})}{\Pr(\mathbf{Dog\ 1} = \text{Yes} \mid \text{Scent is not from A})} = \frac{0.31}{0.073} \approx 4.25$$

In the second equation we see what happens if we only instantiate the evidence that Dog 1 matches the scent of the scentcloth with the scent of the suspect. We can conclude that **Dog 2** really increases the LR in favour of the scent is from A.

**Dogs (strongly) dependent**

When Dog 2 is dependent on what Dog 1 does, conditional on the scent from A, for example they respond on the same scent, which is not the typical scent of A. Then Dog 2, given that Dog 1 has pointed out the suspect will have a larger probability than Dog 1 of pointing out the suspect. Analogously, given that Dog 1 did not point out the suspect, Dog 2 will have a larger probability of not pointing out the suspect than Dog

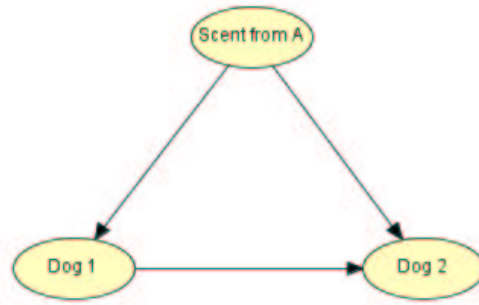


Figure 2.2.3: Bayesian network when the dogs dependent on each other.

Scent from A	Yes		No	
	Yes	No	Yes	No
Dog 1	$p$	$q$	$r$	$s$
No	$1 - p$	$1 - q$	$1 - r$	$1 - s$

Table 2.2.12: Conditional probability table of **Dog 2** given **Scent from A** and **Dog 1**, with  $p > 0.31$ ,  $1 - q > 0.69$ ,  $r > 0.073$ ,  $1 - s > 0.927$

1. This gives the following Bayesian network, displayed in Figure 2.2.3, with the table for Dog 2 adjusted, as can be seen in Table 2.2.12.

**Results:**

In Table 2.2.13 we see the results of the influence of variation of the values of the parameters on the LR. The LR when we only instantiate the evidence that dog 1 matches the scent of the scentcloth with the scent of the suspect, is  $\approx 4.25$ .

$p$	$q$	$r$	$s$	LR of 2 dogs
0.45	0.2	0.1	0.05	19.108
0.6	0.2	0.1	0.05	25.4760
0.8	0.2	0.1	0.05	33.9773
0.45	0.1	0.1	0.05	19.1086
0.45	0.05	0.1	0.05	19.1086
0.45	0.3	0.1	0.05	19.1086
0.6	0.3	0.1	0.05	25.4760
0.45	0.2	0.2	0.05	9.5552
0.45	0.2	0.5	0.05	3.8218
0.45	0.2	0.45	0.05	4.2466
0.45	0.2	0.1	0.07	19.1086
0.45	0.2	0.1	0.01	19.1086

Table 2.2.13: Effect of  $p$ ,  $q$ ,  $r$  and  $s$  on the likelihood ratio, when the 2 dogs are dependent.

We see here that the  $q$  and  $s$  don't have an influence on the LR, and for  $p$  it holds that the bigger  $p$ , the bigger the LR. For  $r$  it holds that the bigger  $r$ , the smaller the LR becomes. Furthermore, we see that  $p$  and  $r$  are dependent on each other. If  $p = r$ , the LR determined above, call this LR(2), is equal to the LR of only instantiating Dog 1 as evidence, call this LR(1). And if  $p < r$ ,  $LR(2) < LR(1)$ , also if  $p > r$ , then  $LR(2) > LR(1)$ . Does this make any sense? Let us determine the LR of this situation, we see:

$$\begin{aligned}
LR(2) &= \frac{\Pr(\mathbf{Dog\ 1} = \text{Yes}, \mathbf{Dog\ 2} = \text{Yes} \mid \text{Scent is from } A)}{\Pr(\mathbf{Dog\ 1} = \text{Yes}, \mathbf{Dog\ 2} = \text{Yes} \mid \text{Scent is not from } A)} \\
&= \frac{\Pr(\mathbf{Dog\ 1} = \text{Yes} \mid \text{Scent is from } A, \mathbf{Dog\ 1} = \text{Yes}) \Pr(\mathbf{Dog\ 2} = \text{Yes} \mid \text{Scent is from } A, \mathbf{Dog\ 1} = \text{Yes})}{\Pr(\mathbf{Dog\ 2} = \text{Yes} \mid \text{Scent is not from } A, \mathbf{Dog\ 1} = \text{Yes}) \Pr(\mathbf{Dog\ 1} = \text{Yes} \mid \text{Scent is not from } A, \mathbf{Dog\ 1} = \text{Yes})} \\
&= \frac{p \times 0.31}{r \times 0.073} \\
&= \frac{p}{r} LR(1).
\end{aligned}$$

So, here we see that the LR doesn't depend on  $q$  and  $s$ , and also that  $LR(2)$  depends on the fraction  $\frac{p}{r}$  and  $LR(1)$ . In the remainder of this report we will not go into the situation in which the dogs are dependent, but only look at the situation assuming the dogs act independently of each other. This situation can easily be modified to the situation with dependent dogs by adding an arrow from Dog 1 to Dog 2, and adjusting the probability table of Dog 2.

### 2.3 Activity level, stage 1

Now moving on to the activity level, we have to take into account that our hypotheses have to contain something an offender did. We will concentrate on the node **Bulldozer**, with states:

$H_1$ : A sat in the bulldozer during the crime,

$H_2$ : A never sat in the bulldozer.

Here, we make a lot of assumptions, which we will now state:

- 1 There is only one offender.
- 2 The offender sat in the bulldozer.
- 3 If A is innocent, he never sat in the bulldozer.
- 4 The hat is from the offender.
- 5 The hat has been worn by one person.
- 6 The offender left the hat in the bulldozer.
- 7 Every person loses the same amount of hair in the same period of time.

Then in the next section, we will look at the consequence on the LR when we remove one assumption at a time. We will see three stages of this and ideas for further research will be stated in the discussion.

Now in stage 1, when we have all these assumptions, we want a relation between the bulldozer and the hairs, and between the bulldozer and the scent. This we will work out in the following two paragraphs. Because of assumptions 2 and 3, we can change the node **Bulldozer** to **Robbery+Bulldozer**, with states:

$H_p$ : A committed the robbery and sat in the bulldozer during the robbery,

$H_d$ : A didn't commit the robbery and he never sat in the bulldozer.

Doing so, we already have hypotheses on the crime level, which is of interest to the judge.

### 2.3.1 Hairs/Hat

The evidence we have are the partial DNA profile and the mtDNA of the found hairs, and the bloodsample of A. To go from the evidence to the hypotheses node, we have to look at what we know and have to know to get there. With the assumptions given above, the only thing open is, whether or not the hairs found in the hat are from A. We again will assume that all the persons related through the maternal line to A are excluded, so we don't have a complication with the mtDNA.

Putting all the characteristics of the suspect A in one node will give the node **X**, with two states:

- x**: Characteristics of the known suspect A such as the DNA-profile and the mtDNA profile, call this M
- **x**: All other characteristics

Then we have to have a node containing all the characteristics of the hairs found in the hat at the scene of the crime. This is the node **Y Hairs**, consisting of four states:

- x**: Observed characteristics of the trace hairs and the partial DNA-profile; HumFES 11/13, HumTHO1 8/9.3, and the mtDNA profile of string  $\Gamma$
- **x**: All other characteristics

Now we have to determine the links between the three variables. There has to be a link from **Robbery+Bulldozer** and **X**, to **Y Hairs**, because for knowing which kind of hairs are found, we have to know if A is the offender or not and the characteristics of A. We know with the assumptions, that the hairs found have to be from the offender. There isn't a link between **Robbery+Bulldozer** and **X**, the characteristics of A are fixed and don't depend on whether or not A committed the crime. This gives the network shown in Figure 2.3.1.

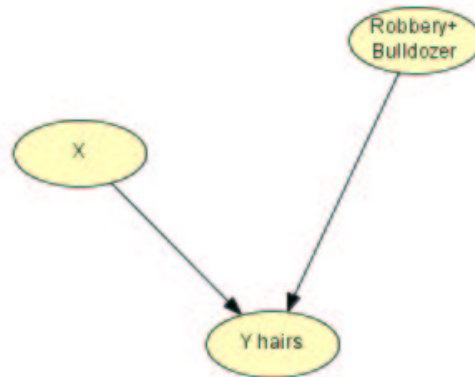


Figure 2.3.1: Bayesian network of the hairs at stage 1.

Next, we have to fill in the tables. Because we are again interested in the LR of  $H_p$  versus  $H_d$ , we set the priors equal to one half. Further we know that the characteristics of A are  $x$ , so we set the prior of  $x$  on 1. The tables will be as shown in Tables 2.3.1 and 2.3.2.

$H_p$	1/2
$H_d$	1/2

Table 2.3.1: Unconditional probability table of **Robbery+Bulldozer**

Now, we only have to determine the conditional probabilities of **Y hairs** given **Robbery+Bulldozer** and **X**.

x	1
¬ x	0

Table 2.3.2: Unconditional probability table of **X**

- When we know that A is the offender and the characteristics of A are x, then the hairs found have characteristics x, with probability 1.
- When we know that A is the offender and the characteristics of A are ¬x, then the hairs found have characteristics ¬x, with probability 1.
- When we know that A is not the offender and the characteristics of A are x, then the hairs found have characteristics x and ¬x with the profile probabilities.
- When we know that A is not the offender and the characteristics of A are ¬x, then the hairs found have characteristics x and ¬x with the profile probabilities.

<b>Robbery+Bulldozer</b>	$H_p$		$H_d$	
	x	¬ x	x	¬ x
x	1	0	0.000179	0.000179
¬ x	0	1	0.999821	0.999821

Table 2.3.3: Conditional probability table of **Y hairs** given **Robbery+Bulldozer** and **X**.

All the (conditional) probabilities are fixed, so we can calculate the LR of  $H_p$  versus  $H_d$ .

$$\begin{aligned}
\text{LR hairs} &= \frac{\Pr(\mathbf{Y \text{ hairs} = x} \mid \mathbf{Robbery+Bulldozer = Yes})}{\Pr(\mathbf{Y \text{ hairs} = x} \mid \mathbf{Robbery+Bulldozer = No})} \\
&= \frac{\Pr(\mathbf{Y \text{ hairs} = x} \mid \mathbf{X = x, Robbery+Bulldozer = Yes})}{\Pr(\mathbf{Y \text{ hairs} = x} \mid \mathbf{X = x, Robbery+Bulldozer = No})} \\
&= \frac{1}{0.000179} \approx 5586.6
\end{aligned}$$

### 2.3.2 Scent

In order to combine the hypothesis node **Robbery+Bulldozer**, to the network defined in section 2.2.2, we have to had transfer of scent from A to the seat of the bulldozer. We assume that there are no matching background scents, which is a huge assumption, but there are at this moment little records of research on the relevance and influence this has on the dogs.

#### Dogs independent

For the dogs to smell a scent there must have been a transfer of that scent from A, or not, to the seat of the bulldozer. Furthermore, we have to define the evidence nodes. This gives the following nodes:

- **Transfer scent** with the states:  
There has been a transfer, persistence and recovery of an amount 's' of scent from A, where s can be in the states high, low and nothing.
- **Dog i (i=1,2)** with the states:  
**Yes:** Dog i recognizes the scent of A on the seat of the bulldozer.  
**No:** No scent recognition.

We now want to get the links between the different nodes. Because we assumed independence between the dogs, there will be no link between them. Whether or not the dogs recognize the scent of A, depends on how much sent has been transferred to the seat of the bulldozer. So there has to be links from the transfer node to the two dogs. The **Transfer scent** node itself depends on whether or not A sat on the seat of the bulldozer, so a link is needed from the **Robbery+Bulldozer** node to the **Transfer scent** node.

This gives the Bayesian network, as displayed in Figure 2.3.2.

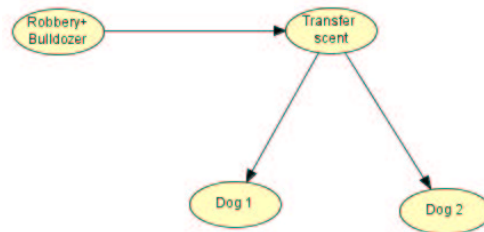


Figure 2.3.2: Bayesian network of the scent at stage 1, when the dogs are independent.

Our next task is to fill in the (conditional) probability tables. Following our goal, we want to determine the likelihood ratio of A sitting in the bulldozer, after the evidence is instantiated. So we set the priors of **Robbery+Bulldozer** both to  $1/2$ . This will give prior odds equal to one, and after instantiating the evidence we can read off the posteriori odds, and thus the likelihood ratio. For the **Transfer scent** node we get the following table, displayed in Table 2.3.4.

<b>Robbery+Bulldozer</b>	$H_1$	$H_2$
high	$t$	0
low	$1 - t$	0
nothing	0	1

Table 2.3.4: Conditional probability table of **Transfer scent**, given **Robbery+Bulldozer**.

Looking at the dogs, they have to have the same probability table because they are independent and we have assumed that they had the same training and as a consequence, they can be seen as the same. In Table 2.2.9, we saw the results of a test to determine the probabilities of a dog recognizing a scent. This we will use here again. Now we have a distinction between high, low and nothing, as an amount of scent. Assuming the test with the dogs, described in table 2.2.9, was done with a high amount of scent, we get the following table for the dogs, displayed in Table 2.3.5.

<b>Transfer scent</b>	high	low	nothing
Yes	0.31	$s$	0.073
No	0.69	$1 - s$	0.927

Table 2.3.5: Conditional probability table of **Dog i**, with  $i=1,2$ , given **Transfer scent**.

here we have  $0.31 > s > 0.073$ . We have 2 parameters,  $t$  and  $s$ , which we can vary.

**Results:**

We see that the higher the transfer of scent, the higher the LR becomes, meaning that the hypotheses of A being in the bulldozer and committed the crime becomes more likely. When we vary  $s$ , we see that when  $s$  increases and comes near 0.31, the LR increases also.

Writing out the LR, using the probability laws described in section 1.1, gives:

$t$	$s$	LR scent of <b>Robbery+Bulldozer</b>
0.1	0.2	8.558829
0.3	0.2	10.66429
0.5	0.2	12.76975
0.7	0.2	14.87521
0.9	0.2	16.98067
0.3	0.1	6.723588
0.3	0.15	8.365547
0.3	0.25	13.61982
0.3	0.3	17.23213

Table 2.3.6: Results of the influence of  $s$  and  $t$  on the LR, when the dogs are independent and only the scent is taking into account.

$$\begin{aligned}
LR &= \frac{\Pr(\mathbf{Dog\ 1} = \text{Yes}, \mathbf{Dog\ 2} = \text{Yes}, \mathbf{TransferScent} \mid \mathbf{Robbery+Bulldozer} = \text{Yes})}{\Pr(\mathbf{Dog\ 1} = \text{Yes}, \mathbf{Dog\ 2} = \text{Yes}, \mathbf{TransferScent} \mid \mathbf{Robbery+Bulldozer} = \text{No})} \\
&= \frac{(0.31 \times 0.31 \times t) + (s \times s \times (1 - t)) + (0.073 \times 0.073 \times 0)}{(0.31 \times 0.31 \times 0) + (s \times s \times 0) + (0.073 \times 0.073 \times 1)} \\
&= \frac{(0.31)^2 t + s^2 (1 - t)}{0.073^2}
\end{aligned}$$

Let us verify one case and see how accurate Hugin really is, take  $t = 0.3$  and  $s = 0.2$ . Then calculating the LR gives:  $LR = \frac{(0.31)^2 \times 0.3 + (0.2)^2 \times (1 - 0.3)}{0.073^2} = 10.66429$ . As one can see, Hugin is in this case very accurate, the other cases will be left to the reader.

### Dogs (strongly) dependent

Take for example that Dog 2 depends on what Dog 1 does. In the same way as in section 2.1.2, we have to have a link from **Dog 1** to **Dog 2**. This gives the following network with an adjusted table for Dog 2:

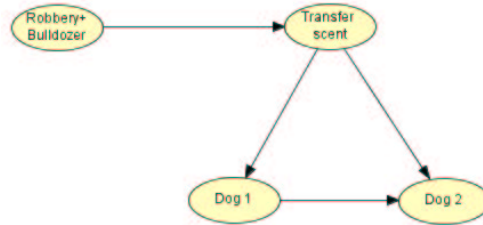


Figure 2.3.3: Bayesian network of the scent at stage 1, when the dogs are dependent.

<b>Transfer scent</b>	high		low		nothing	
	Yes	No	Yes	No	Yes	No
<b>Dog 1</b>						
Yes	0.9	0.1	0.9	0.1	0.9	0.01
No	0.1	0.9	0.1	0.9	0.1	0.99

Table 2.3.7: Conditional probability table of **Dog 2** given **Transfer scent** and **Dog 1**.

With these numbers written in the table, we investigate the effect on **Robbery+Bulldozer** when the two dogs are **strongly** dependent. We then can compare this with the case when the dogs are independent

to see if it has a huge effect on the hypothesis node **Robbery+Bulldozer**.

**Results:**

$t$	$s$	LR scent of <b>Robbery+Bulldozer</b>
0.1	0.2	2.890411
0.3	0.2	3.191781
0.5	0.2	3.493151
0.7	0.2	3.794521
0.9	0.2	4.09589
0.3	0.1	2.232877
0.3	0.15	2.712329
0.3	0.25	3.671233
0.3	0.3	4.150685

Table 2.3.8: Results of the influence of  $s$  and  $t$  on the LR, when the dogs are strongly dependent

In Table 2.3.8 we see the results of the influence of  $s$  and  $t$  on the LR. The LR, given in column 3, is the same as the LR when we only instantiate dog 1, this means that dog 2 doesn't add any value to the LR. So, we can only take dog 1 as evidence. We could have argued this in advance, because we instantiate **Dog 1** on 'Yes', the probabilities of **Dog 2** are for all values of **Transfer scent** the same, so the arrow from **Transfer scent** to **Dog 2** disappears, and **Dog 2** only depends on **Dog 1**. Let's now see what happens if we change this, take the following table for **Dog 2**, as shown in Table 2.3.9.

<b>Transfer scent</b>	high		low		nothing	
	Yes	No	Yes	No	Yes	No
Yes	0.9	0.1	0.5	0.1	0.2	0.01
No	0.1	0.9	0.5	0.9	0.8	0.99

Table 2.3.9: Conditional probability table of **Dog 2** given **Transfer scent** and **Dog 1**.

We get the following results, displayed in Table 2.3.10. As we can see, these results lie somewhere between

$t$	$s$	LR scent of <b>Robbery+Bulldozer</b>
0.1	0.2	8.075342
0.3	0.2	10.5274
0.5	0.2	12.979945
0.7	0.2	15.43151
0.9	0.2	17.88356
0.3	0.1	8.130137
0.3	0.15	9.328767
0.3	0.25	11.72603
0.3	0.3	12.92466

Table 2.3.10: Results of the influence of  $s$  and  $t$  on the LR, when the dogs are strongly dependent

the case where the dogs are independent and the case that the dogs are strongly dependent. This is what we would expect.

### 2.3.3 Network stage 1

Putting the hairs and dogs (when they are independent, we won't treat the case when the dogs are strongly dependent) together, will give the network displayed in Figure 2.3.4.

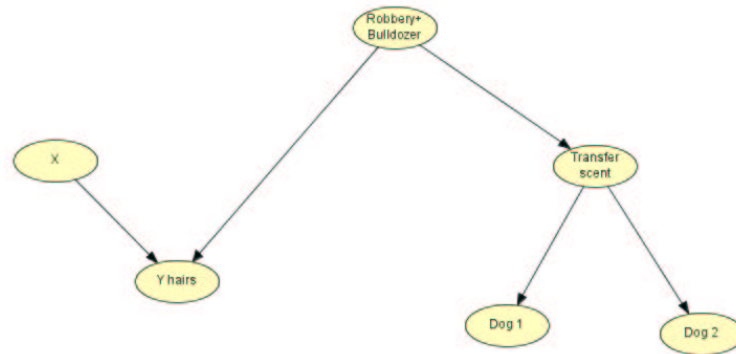


Figure 2.3.4: Bayesian network of the situation at stage 1, when the dogs are independent.

Looking at the influence of the parameters on the LR, we see that only the scent part has parameters. Varying these, give the following results, displayed in Table 2.3.11 With LR hairs as calculated in section

$t$	$s$	LR of <b>Robbery+Bulldozer</b>	LR hairs×LR scent
0.1	0.2	47823	47815
0.3	0.2	59594	59577
0.5	0.2	71326	71339
0.7	0.2	83125	83102
0.9	0.2	94876	94864
0.3	0.1	37565	37561
0.3	0.15	46728	46735
0.3	0.25	76103	76088
0.3	0.3	96245	96269

Table 2.3.11: Results of the influence of  $s$  and  $t$  on the LR in the whole network, when the dogs are independent

2.3.1, and LR scent as the results from Table 2.3.6. We see that there are little differences in the values and increasing the accuracy will give values even closer to the values in column 4. The accuracy of the posterior probabilities is crucial if we want to get the right LR. Furthermore, we see that when increasing  $t$ , the LR also increases, and the same holds for  $s$ .

## 2.4 Activity level, stage 2

In the previous section we assumed the hairs were from the offender, now we omit this assumption. So, this gives us the following remaining assumptions:

- 1 There is only one offender.
- 2 The offender sat in the bulldozer.
- 3 If A is innocent, he never sat in the bulldozer.
- 4 The hat has been worn by one person.

5 The offender left the hat in the bulldozer.

6 Every person loses the same amount of hair in the same period of time.

The network of the scent will remain the same as in section 3.3.2 and we won't repeat it here.

### 2.4.1 Hairs/Hat

If we don't know whether or not the hairs are from the offender, we have to introduce a new node, which will show if the hairs are relevant as evidence or not. Because if the hairs are not from the offender, it is not connected with the crime, and then we aren't interested in the source of the hairs. So we define the new node **Relevance hairs** with the states:

**Yes:** The offender left the hairs in the hat at the scene of the crime.

**No:** The offender did not leave the hairs in the hat at the scene of the crime.

For this to work, we have to introduce an extra node, called **Hairs from A**, with states **Yes** and **No**. We still have the nodes **Robbery+Bulldozer**, **Y hairs** and **X**. Determining the links, we see that if we want to know **Y hairs**, we have to know whether or not the hairs are from A and the profile of A, so links from **Hairs from A** and **X** to **Y hairs**. For the determination of **Hairs from A**, we have to know if the offender left the hairs or not, and whether or not A is the offender, so links from **Robbery+Bulldozer** and **Relevance hairs** to **Hairs from A**. We get the following:

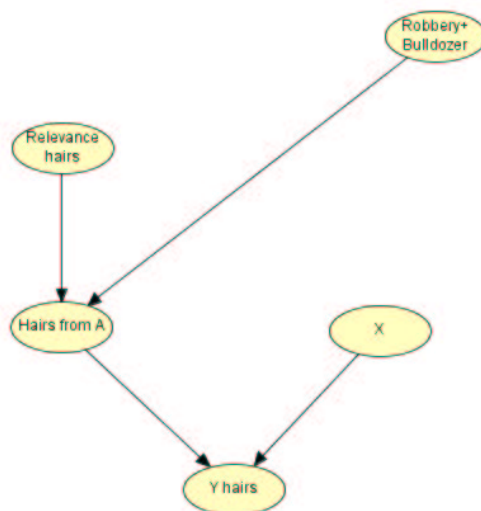


Figure 2.4.1: Bayesian network of the hairs at stage 2.

The tables of **Robbery+Bulldozer** and **X** will remain the same. **Relevance hairs** is an unconditional table, which we will vary, so let's set this at  $r$ .

Yes	$r$
No	$1 - r$

Table 2.4.1: Unconditional probability table of **Relevance hairs**

Here,  $q$  in Table 2.4.3 is the probability that there are hairs found of A, but A didn't commit the crime. This value can be set at one's own insight. Here we set it to 0.1. So, the only parameter we have is  $r$ . Varying  $r$ , will give the results stated in Table 2.4.4.

Hairs from A	Yes		No	
	x	$\neg x$	x	$\neg x$
x	1	0	0.000179	0.000179
$\neg x$	0	1	0.999821	0.999821

Table 2.4.2: Conditional probability table of **Y hairs** given **Hairs from A** and **X**.

Robbery+Bulldozer	$H_p$		$H_d$	
	Yes	No	Yes	No
Yes	1	0	0	$q$
No	0	1	1	$1 - q$

Table 2.4.3: Conditional probability table of **Hairs from A** given **Relevance hairs** and **Robbery+Bulldozer**.

$r$	LR
0.99	839.8237
0.9	88.43464
0.7	23.20085
0.5	9.967889
0.3	4.277332
0.1	1.110891

Table 2.4.4: Results of the influence of **Relevance hairs** on the LR

Here we see that the relevance has a great impact on the LR of **Robbery+Bulldozer**. This is a logical result, because when the hairs are not that relevant as evidence, they can't have much value. Furthermore, we see that even if  $r$  is close to one, namely 0.99, we still have a big difference with the LR of 5586.6 we found in section 2.3.1.

## 2.4.2 Network stage 2

Putting the 2 networks, the independent dogs of section 2.3.2 and the hairs of the previous section, together, will give the network displayed in Figure 2.4.2.

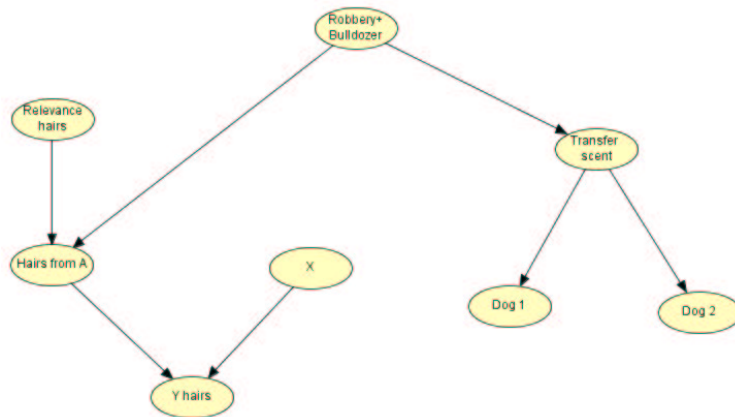


Figure 2.4.2: Bayesian network of the situation at stage 2, when the dogs are independent.

### Results:

In this network, we have three parameters we can vary. Again, we will set two out of three fixed and look at the influence on the LR. The results are shown in Table 2.4.5.

$t$	$s$	$r$	LR	LR hairs $\times$ LR scent
0.1	0.2	0.9	756.8973	756.8970
0.3	0.2	0.9	943.0909	943.0926
0.5	0.2	0.9	1129.288	1129.288
0.7	0.2	0.9	1315.482	1315.484
0.9	0.2	0.9	1501.675	1501.679
0.3	0.1	0.9	594.5997	594.5981
0.3	0.15	0.9	739.8066	739.804
0.3	0.25	0.9	1204.458	1204.464
0.3	0.3	0.9	1523.925	1523.917
0.3	0.2	0.99	8956.363	8956.123
0.3	0.2	0.7	247.4207	247.4206
0.3	0.2	0.5	106.3005	106.3005
0.3	0.2	0.3	45.6147	45.6147
0.3	0.2	0.1	11.8469	11.8469

Table 2.4.5: Results of the influence of  $t$ ,  $s$  and  $r$  on the LR

In the fifth column, we took the product of the two separate LR's of the hairs and the scent, which we determined in section 2.4.1 and 2.3.2, respectively. We see that as we expect that the two values in the fourth and fifth column are very close, and the differences we explain by the level of accuracy. Furthermore, the LR increases if  $t$ ,  $r$  or  $s$  increases. This can be explained, while a bigger transfer probability or higher relevance the hairs, will give a higher probability that the suspect is guilty.

## 2.5 Activity level, stage 3

In this section, we want to leave the opportunity open for hairs being in the hat before the crime was committed. But we assume that the hairs can be from not more than two persons. So we have two stages in time, namely before and during the crime. We still hold on to the assumption that the offender left the hat, but we now have to specify which person(s) wore the hat. The network of the scent will still remain the same as in the previous section.

### 2.5.1 Hairs/Hat

At this point, we have to define a number of extra nodes, for the derivation of this network. The nodes **Robbery+Bulldozer** and **X** will be the same. We begin with the time periods at hand. We define **Hairs from offender** for during the crime and **Background hairs** for before the crime, both with states:

**O1 resp. B1:** Hairs from A.

**O2 resp. B2:** Hairs from someone else.

**O3 resp. B3:** No hairs.

We have to know whether or not the hat was worn, so we define the nodes **R wearing before**, with states:

**Yes:** Someone wore the hat before the crime.

**No:** No one wore the hat before the crime.

and **R wearing during**, with states

**Yes:** The offender wore the hat.

**No:** The offender did not wear the hat.

From wearing the hat to the hairs found, there had to be transfer of the hairs to the hat, so define for both time periods, **T hairs before**, with states:

**Yes:** There was transfer, persistence and recovery of hairs to the hat **before** the crime,

**No:** There was **no** transfer, persistence or recovery of hairs to the hat **before** the crime,  
and **T hairs during**, with states:

**Yes:** There was transfer, persistence and recovery of hairs to the hat **during** the crime,

**No:** There was **no** transfer, persistence or recovery of hairs to the hat **during** the crime.

Because we now can have two groups of hairs found in the hat, we have to adjust the definition of **Y hairs**, we now get the following states:

**x:** one or two groups of hairs with characteristics x are found,

$\neg$  **x:** one or two groups of hairs with characteristics  $\neg$  x are found,

**two groups:** two groups of hairs are found with different characteristics,

**no groups:** no hairs are found from before and during the crime.

Now we have all the ingredients to form the Bayesian network of this particular situation. As we have seen above there has to be links from **R wearing before** to **T hairs before**, and from **R wearing during** to **T hairs during**. To determine the probabilities of **Hairs from offender** and **Background hairs**, we have to know if A was the offender or not and whether or not there was transfer of hairs to the hat. So also links from **Robbery+Bulldozer** and **T hairs during** to **Hairs from offender**, and from **Robbery+Bulldozer** and **T hairs before** to **Background hairs**. The final links will be from **Hairs from offender**, **Background hairs** and **X** to **Y hairs**. The network will be given in Figure 2.5.1.

Filling in the tables will be the same for **Robbery+Bulldozer** and **X** as in the previous stage. The nodes **R wearing before** and **R wearing during** are priors, call these  $r_b$  and  $r_d$ , respectively. We get:

	<b>R wearing before</b>	<b>R wearing during</b>
Yes	$r_b$	$r_d$
No	$1-r_b$	$1-r_d$

Table 2.5.1: Unconditional probability table of **R wearing before** and **R wearing during**

Determining the conditional probabilities of the transfer nodes given the hat was worn or not, we assume that there is no secondary transfer, which means if the hat was not worn, there is no transfer of hairs. When the hat was worn, there was transfer with a large probability. Further we assume that this probability will be larger in the case the hat was worn before the crime. We take the following values, as can be seen in Tables 2.5.2 until 2.5.5, with the assumption that  $t_b > t_d$ .

<b>R wearing before</b>	Yes	No
Yes	$t_b$	0
No	$1-t_b$	1

Table 2.5.2: Conditional probability table of **T hairs before** given **R wearing before**

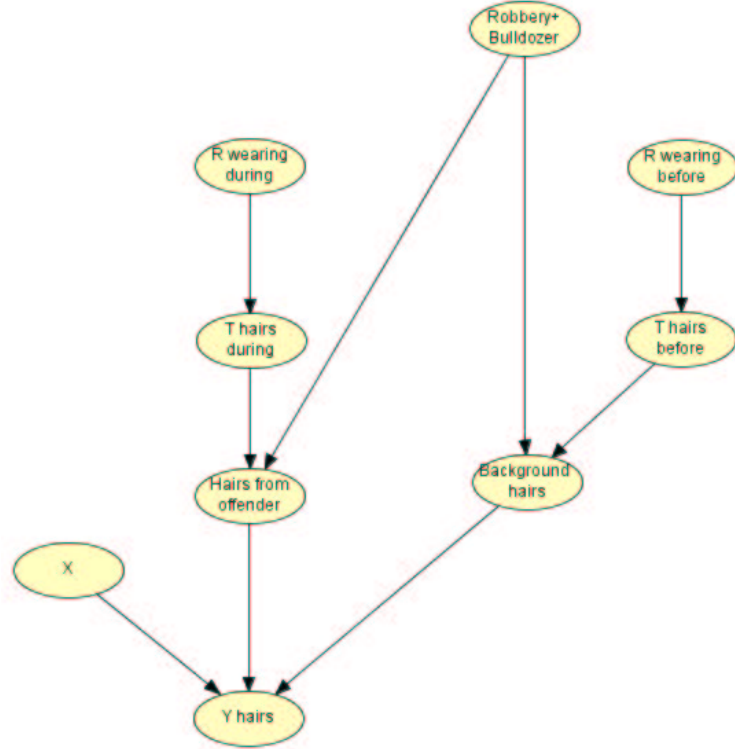


Figure 2.5.1: Bayesian network of the hairs at stage 3.

<b>R wearing during</b>	Yes	No
Yes	$t_d$	0
No	$1-t_d$	1

Table 2.5.3: Conditional probability table of **T hairs during** given **R wearing during**

<b>Robbery+Bulldozer</b>	Yes		No	
<b>T hairs during</b>	Yes	No	Yes	No
O1	1	0	0	0
O2	0	0	1	0
O3	0	1	0	1

Table 2.5.4: Conditional probability table of **Hairs from offender** given **T hairs during** and **Robbery+Bulldozer**

Furthermore, in Table 2.5.5,  $1-q$  is the probability that someone else wore the hat, and  $p$  is the probability that the hairs from A are in the hat for ‘innocent’ reasons. In this probability we find the possibility that the hat was planted by someone who dislikes A, and the possibility that the hat was from someone A knew and that A wore it, but didn’t commit the crime. The probability  $p$  will be small, and  $q$  we will vary and look at the influence of it. What is left is the probability table of **Y hairs** given **Hairs from offender**, **Background hairs** and **X**, this is given in the following Tables 2.5.6 and 2.5.7. We will explain a number of probabilities, first of all  $\Pr(\mathbf{Y hairs} = x \mid \mathbf{Background hairs} = \mathbf{B2}, \mathbf{Hairs from offender} = \mathbf{O2}, \mathbf{X} = x) = p^2$ , this holds because we can have two groups of hairs with characteristics  $x$ . For the probability

<b>Robbery+Bulldozer</b>	Yes		No	
	Yes	No	Yes	No
<b>T hairs before</b>				
B1	$q$	0	$p$	0
B2	$1 - q$	1	$1 - p$	0
B3	0	0	0	1

Table 2.5.5: Conditional probability table of **Background hairs** given **T hairs before** and **Robbery+Bulldozer**

$\Pr(\mathbf{Y hairs} = \neg x \mid \mathbf{Background hairs} = \text{B2}, \mathbf{Hairs from offender} = \text{O2}, \mathbf{X} = x)$ , we take a small number but bigger than  $p$ , say 0.0002.

<b>Background hairs</b>	B1						B2			
	O1		O2		O3		O1		O2	
<b>Hairs from offender</b>	x	$\neg x$	x	$\neg x$	x	$\neg x$	x	$\neg x$	x	$\neg x$
<b>X</b>	x	$\neg x$	x	$\neg x$	x	$\neg x$	x	$\neg x$	x	$\neg x$
x	1	0	0.000179	0	1	0	0.000179	0	$0.000179^2$	0
$\neg x$	0	1	0	0	0	1	0	0	$\approx 0.0002$	0
two groups	0	0	0.999821	1	0	0	0.999821	1	$\approx 0.999799968$	1
no groups	0	0	0	0	0	0	0	0	0	0

Table 2.5.6: Conditional probability table of **Y hairs** given **Hairs from offender**, **Background hairs** and **X**.

<b>Background hairs</b>	B2				B3				
	O3		O1		O2		O3		
<b>Hairs from offender</b>	x	$\neg x$	x	$\neg x$	x	$\neg x$	x	$\neg x$	
<b>X</b>	x	$\neg x$	x	$\neg x$	x	$\neg x$	x	$\neg x$	
x	0.000179	0.000179	1	0	0.000179	0.000179	0	0	
$\neg x$	0.999821	0.999821	0	1	0.999821	0.999821	0	0	
two groups	0	0	0	0	0	0	0	0	
no groups	0	0	0	0	0	0	1	1	

Table 2.5.7: Conditional probability table of **Y hairs** given **Hairs from offender**, **Background hairs** and **X**.

In this situation, we have six parameters,  $p$ ,  $q$ ,  $r_b$ ,  $r_d$ ,  $t_b$  and  $t_d$ . Varying these give the results displayed in Table 2.5.8.

In Table 2.5.8 we see that increasing  $q$ ,  $r_d$  or  $t_d$ , will increase the LR, and increasing  $p$ ,  $r_b$  or  $t_b$  will decrease the LR. This is what we would expect, looking at the definitions of the parameters. Furthermore, when varying  $r_b$  or  $t_b$ , we get the same LR, for the same values, when we hold the rest fixed. Also we see that the highest LR is reached when either  $r_b$  or  $t_b$ , is 0.1. But what happens now if we take both  $r_b$  and  $t_b$  to be 0.1, then we get a LR of 1766.

## 2.5.2 Network stage 3

Putting the network of the previous section together with the network of the scent in section 2.3.2, we get the network at stage 3, displayed in Figure 2.5.2.

This will give us a total of eight parameters to vary, the results of this are given in Table 2.5.9.

Again we see that increasing  $q$ ,  $r_d$  or  $t_d$ , will increase the LR, and increasing  $p$ ,  $r_b$  or  $t_b$  will decrease the LR. And increasing  $s$  and  $t$  will also increase the LR. Furthermore, when varying  $r_b$  or  $t_b$ , we get the same LR, for the same values, when we hold the rest fixed. Also we see that the highest LR is obtained when either  $r_b$  or  $t_b$ , is 0.1. Taking both of them equal to 0.1, will only increase the LR more.

$p$	$q$	$r_d$	$r_b$	$t_d$	$t_b$	LR
0.1	0.8	0.9	0.9	0.8	0.9	34.5
0.4	0.8	0.9	0.9	0.8	0.9	8.64
0.6	0.8	0.9	0.9	0.8	0.9	5.76
0.9	0.8	0.9	0.9	0.8	0.9	3.84
0.1	0.1	0.9	0.9	0.8	0.9	9.58
0.1	0.4	0.9	0.9	0.8	0.9	20.3
0.1	0.6	0.9	0.9	0.8	0.9	27.4
0.1	0.9	0.9	0.9	0.8	0.9	38.1
0.1	0.8	0.1	0.9	0.8	0.9	8.89
0.1	0.8	0.3	0.9	0.8	0.9	11.2
0.1	0.8	0.5	0.9	0.8	0.9	14.9
0.1	0.8	0.7	0.9	0.8	0.9	21.1
0.1	0.8	0.9	0.1	0.8	0.9	275
0.1	0.8	0.9	0.3	0.8	0.9	96.7
0.1	0.8	0.9	0.5	0.8	0.9	59.5
0.1	0.8	0.9	0.7	0.8	0.9	43.5
0.1	0.8	0.9	0.9	0.1	0.9	9.01
0.1	0.8	0.9	0.9	0.3	0.9	11.8
0.1	0.8	0.9	0.9	0.6	0.9	20.1
0.1	0.8	0.9	0.9	0.9	0.9	51.9
0.1	0.8	0.9	0.9	0.8	0.1	275
0.1	0.8	0.9	0.9	0.8	0.3	96.7
0.1	0.8	0.9	0.9	0.8	0.5	59.5
0.1	0.8	0.9	0.9	0.8	0.7	43.5

Table 2.5.8: Results of the network in stage 3, only the hairs/hat.

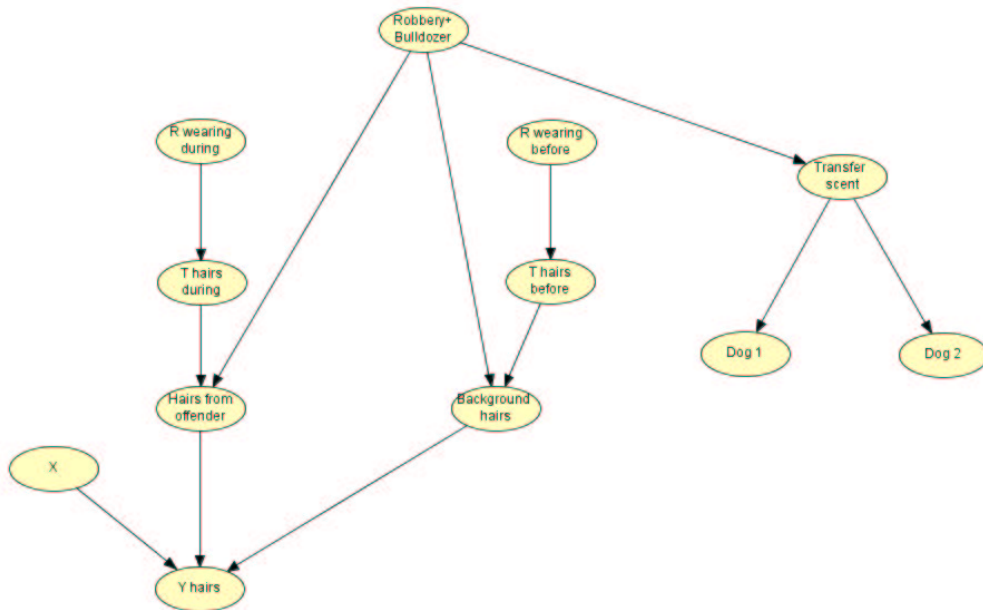


Figure 2.5.2: Bayesian network of the situation at stage 3, when the dogs are independent.

$p$	$q$	$r_d$	$r_b$	$t_d$	$t_b$	$s$	$t$	LR
0.1	0.8	0.9	0.9	0.8	0.9	0.2	0.3	368
0.4	0.8	0.9	0.9	0.8	0.9	0.2	0.3	92.2
0.6	0.8	0.9	0.9	0.8	0.9	0.2	0.3	61.5
0.9	0.8	0.9	0.9	0.8	0.9	0.2	0.3	41.0
0.1	0.1	0.9	0.9	0.8	0.9	0.2	0.3	102
0.1	0.4	0.9	0.9	0.8	0.9	0.2	0.3	216
0.1	0.6	0.9	0.9	0.8	0.9	0.2	0.3	292
0.1	0.9	0.9	0.9	0.8	0.9	0.2	0.3	408
0.1	0.8	0.1	0.9	0.8	0.9	0.2	0.3	94.8
0.1	0.8	0.3	0.9	0.8	0.9	0.2	0.3	120
0.1	0.8	0.5	0.9	0.8	0.9	0.2	0.3	159
0.1	0.8	0.7	0.9	0.8	0.9	0.2	0.3	225
0.1	0.8	0.9	0.1	0.8	0.9	0.2	0.3	2934
0.1	0.8	0.9	0.3	0.8	0.9	0.2	0.3	1031
0.1	0.8	0.9	0.5	0.8	0.9	0.2	0.3	635
0.1	0.8	0.9	0.7	0.8	0.9	0.2	0.3	463
0.1	0.8	0.9	0.9	0.1	0.9	0.2	0.3	96.1
0.1	0.8	0.9	0.9	0.3	0.9	0.2	0.3	126
0.1	0.8	0.9	0.9	0.6	0.9	0.2	0.3	214
0.1	0.8	0.9	0.9	0.9	0.9	0.2	0.3	553
0.1	0.8	0.9	0.9	0.8	0.1	0.2	0.3	2934
0.1	0.8	0.9	0.9	0.8	0.3	0.2	0.3	1031
0.1	0.8	0.9	0.9	0.8	0.5	0.2	0.3	635
0.1	0.8	0.9	0.9	0.8	0.7	0.2	0.3	463
0.1	0.8	0.9	0.9	0.8	0.9	0.1	0.3	232
0.1	0.8	0.9	0.9	0.8	0.9	0.15	0.3	289
0.1	0.8	0.9	0.9	0.8	0.9	0.25	0.3	470
0.1	0.8	0.9	0.9	0.8	0.9	0.3	0.3	594
0.1	0.8	0.9	0.9	0.8	0.9	0.2	0.1	295
0.1	0.8	0.9	0.9	0.8	0.9	0.2	0.5	440
0.1	0.8	0.9	0.9	0.8	0.9	0.2	0.7	513
0.1	0.8	0.9	0.9	0.8	0.9	0.2	0.9	586

Table 2.5.9: Results of the network in stage 3.

## Chapter 3

# Discussion

Looking at all the results we have till now, we see that it is very important what the values of the parameters are. This makes a huge difference to the value of the LR and thus on the question whether or not the suspect is guilty. So it is very important that research has to be done on a way to determine these, otherwise it is not possible to use the networks in real cases. This is only one of the problems we encountered during this report. Also the formulation of the hypotheses is very important, as also of the states of the nodes. When changing a little in the definitions, the whole network changes. We saw that it is easier to make a lot of assumptions first and then drop them one by one, than to try to put the whole situation directly in a Bayesian network. The first way allows to investigate the influences of the assumptions on the network, while the second way, gets complicated really fast. The risk of forgetting a particular situation is huge, and also taking a situation twice, for example the suspect being in the background is covered in two different nodes, isn't impossible. To put the whole situation in a network requires a lot of nodes but also with a lot of states, as a consequence you get huge conditional probability tables, which are prone to (human) error, because in this situation they have to be filled in by hand.

In this particular case we had three pieces of evidence, two of which are either very controversial or difficult to use. First, we got the situation with the scent identification line-ups by dogs. There are people who are very critical in the use of this kind of evidence as can be seen in [5] and [10]. Here they make it clear that it is imperative to do a second test with a different dog when the first one points out the suspect, this can also be seen in our networks, that having a second dog will increase the LR, if the second dog makes the same decision. The problem is that there are no results available on how independent two dogs are in this situation, when we know this, we can determine the probabilities with more accuracy. Also, what to do if the two dogs point out a different person, has to be investigated. Secondly, we have the mtDNA, which is practically the same for every member related through the maternal line to the suspect, so it can never be said with great probability, that if the mtDNA matches, that then the suspect is the offender. Here we only looked in the simplest case at the possibility of a brother, but this can be extended to all relatives at the maternal line by using formulas from [4], and worked through all the way to the big networks. An other way is putting the relatives you want to include in the network as nodes.

Also the use of Hugin gave some problems. First of all the accuracy of the posterior probabilities you read off in the hypothesis node, is of real importance, this has to be very high, to get the same LR as in an exact calculation. We saw an example of this in section 2.2.1, where there was a huge difference in the LR, when taking the posterior probabilities more or less accurate. Secondly, in this case we had to fill in all the probabilities by hand, which has as a consequence that it is very prone to error when you get to larger probability tables. This problem is made worse, because when you fill in a number wrongly, Hugin won't give you a warning that the probabilities don't add up to one, but just normalizes the numbers you have written in that particular column.

Returning to the problem of the determination of the parameters, some research has been done, especially in the field of the transfer probabilities for example in [7] and [13], so we could look at more specific values of the transfer probabilities and also take into account the possibility of secondary transfer, this means that there are hairs on the hat from A, but A didn't wear the hat. For example, A has been in contact with the offender, and there have been hairs transferred from A to the offender and the offender transferred the hairs

from A to the hat. Furthermore, we can in our case try to do sensitivity analysis on the parameters, which means that we have to vary all the parameters at once, instead of what we did here, varying one parameter and holding the other ones fixed. In [3], this has been done with two parameters, and here we sometimes have five or six parameters, so this will become much harder to do. Another way of doing this is putting the parameter as an extra node in the network, and defining a probability distribution on it. So then the LR will also incorporate the uncertainty about the parameter.

I think that this method with Bayesian networks is not applicable in the near future in real cases. In this report we tried to put a real case in a Bayesian network, but simplified it a considerable amount, real cases involve much more information and situations we have to take into account, such as multiple offenders and suspects, much more evidence, which may depend on each other, and these are just a number of complications. But I do think that doing this very often, will make you more skilled working with Hugin and it does have some advantages for a statistician to explain with the help of these networks the case at hand to a lawyer or judge. It can give a clear picture of the situation at hand, when the situation is not too complicated. In my opinion, Bayesian reasoning is a very helpful tool in forensic science, and in the future it can be good base for the evaluation of evidence when presented to the courtroom, but lawyers and judges then have to have some basic knowledge in statistics.

# Bibliography

- [1] Aitken, C.G.G., Taroni, F. (2004): *Statistics and the evaluation of evidence for forensic scientists*, 2nd ed., John Wiley and Sons Ltd.
- [2] Aitken, C.G.G. et al (2006, to appear): *Bayesian Networks and probabilistic inference in forensic science*, John Wiley and Sons Ltd.
- [3] Biedermann, A., Taroni, F. (2005): ‘Bayesian Networks and probabilistic reasoning about scientific evidence when there is a lack of data’, *Forensic Sci. Int.*
- [4] Buckleton, J., Triggs, C.M., Walsh, S.J. (editors)(2005): *Forensic DNA evidence interpretation*, CRC Press.
- [5] Frijters, J.E.R., Boksem, J. (2004): ‘Een positieve geuridentificatie proef dient vrijwel altijd in een tegenonderzoek worden herhaald!’, *Nederlands Juristen Blad*, **14**.
- [6] Garbolino, P., Taroni, F. (2002): ‘Evaluation of scientific evidence using Bayesian networks’, *Forensic Sci. Int.* **125** 149-155.
- [7] Gaudette, B.D., Tessarolo, A.A. (1987): ‘Secondary transfer of human scalp hair’, *Journal of Forensic Science*, **32**, no.5, 1241-1253.
- [8] HUGIN: free version available at <http://www.hugin.com>
- [9] Jensen, F.V. (2001): *Bayesian Networks and Decision Graphs*, Springer-Verlag, New York
- [10] Jong,de, S. (2003): ‘De natte neus als rechter’. *HP de Tijd*.
- [11] Lauritzen, S.L. (1996): *Graphical Models*, Clarendon Press, Oxford.
- [12] Lauritzen, S.L., Sheehan, N.A. (2003): ‘Graphical Models for Genetic Analysis’, Department of Mathematical Science, Aalborg University
- [13] Peabody, A. J. and Thomas, K. (1984): ‘On the nature of human head hairs shed onto various types of headgear’, *Journal of the Forensic Science Society* **24** 383-384
- [14] Rice, J.A. (1995): *Mathematical Statistics and Data Analysis*, 2nd ed., Duxbury Press.
- [15] Schoon, G.A.A. (1998): ‘A first assessment of the reliability of an improved scent identification lineup’. *Journal of Forensic Sciences* **43** (1): 70-75
- [16] Schoon, G.A.A. en Koppen, P.J. (2002): ‘Identificatie door honden’. In: *Het recht van binnen. Psychologie van het recht*, ed: van Koppen, P.J., Hessing, D.J., Merkelenbach, H.L.G.J. en Crombag, H.R.M.. Kluwer, Deventer, pp. 597-622.
- [17] Taroni, F. et al. (2004): ‘A general approach to Bayesian Networks for the interpretation of evidence’, *Forensic Sci. Int.* **139** 5-16.
- [18] <http://www.andrew.cmu.edu/user/scheines/tutor/d-sep.html>