17th BILETA Annual Conference

April 5th - 6th, 2002.
Free University, Amsterdam.

Use of Bayesian Belief Networks in legal reasoning

P.E.M. Huygen
Computer/Law Institute
Vrije Universiteit Amsterdam

Abstract

The principles of Bayesian probability and Bayesian belief networks, and their applicability for judicial reasoning are briefly explained. As an example, a Bayesian belief network model has been constructed that compares the probabilities of two hypothesised causes of a single-car accident. It is concluded that Bayesian belief networks can be a suitable tool for lawyers to analyse the evidence in judicial cases, as well as a suitable tool to train students to perform those analyses and to avoid statistical pitfalls.

1. Introduction

A murder has been committed. The police found a button in the fist of the victim. Later on, a suspect was identified, who was probably near the scene of the crime around te time that the crime took place. It turns out, that the suspect owns a coat with buttons of the same type as that in the fist of the victim. The prosecutor charges the suspect with murder, reasoning: ``The perpetrator of the crime owns a coat with the type of button that we found on the victim. The suspect owns such a coat, hence the suspect has committed the crime''. This reasoning scheme is called abduction. (Cause C causes effect E; E has been observed, hence C has occurred). Unfortunately for the prosecutor it is not logically valid. However, it is often the only way to reason about evidence. The problem is that, although an abductive argument does not prove that the proposition is true, it increases the probability of the proposition. Therefore, abductive reasoning should be performed with probabilistic techniques. A very suitable technique to do this is Bayesian probability. Bayesian probability calculates how the probability that a certain cause actually occurred increases, given the prior probability that this cause occurred and abductive evidence. Although non-statisticians may find Bayesian probability theory hard to understand, they can use the techniques, because nowadays there exist computer programs that perform the calculations and visualize the results without requiring detailed knowledge about how the calculations have to be performed. The use of these programs may also help to avoid statistical pitfalls that sometimes happen in courts. E.g. Wagenaar [9] presents a striking example of a known pitfall (the *prosecutors fallacy* [1], explained in section 1.1.1) that occurred in a Dutch court during a procedure where he performed an expert witness testimony.

1.1 Bayesian probability, the frequentist view and the Bayesian view

The most common view on probability is the statistical view, in which the probability of an event is derived from repeated experiments in which the frequency that interesting phenomena occur is analysed. This is called the *frequentist* view. An alternative view is the Bayesian view, that allows to reason with subjective belief estimations and provides a rule to revise ones belief when new evidence is presented. Although reverent Thomas Bayes formulated his rule about 250 years ago [2], the interest of scientists in the Bayesian view has only been widely adopted during the last decade, because scientists had trouble with the required subjective prior estimations of probability, and

because the computer power required to calculate probabilities in complex models was not available. Bayesian reasoning is very suitable to model judicial problems. Next sections will elaborate on both views of probability.

1.1.1 Frequentist view

Frequentists make predictions based on the observation of the frequency that the to be predicted event occurred in the past. For example, suppose it has been observed that 2% of a population has hair of a certain type and colour. We denote the observation ``*person x has matching hair''* with character E. Then the probability that a randomly chosen person has matching hair is P(E) = 0.02. Suppose such a hair has been found at the scene of a crime, and there is a suspect, whose hair type matches the hair found. If the suspect is guilty, the probability of a match is equal to unity, and otherwise it is equal to 0.02. These probabilities are called ``*conditional probabilities''.* Symbolising hypothesis ``*subject is guilty''* with G, we denote the conditional probability of a match, given that the person is guilty or given that he is innocent as P(E|G) = 1 resp. P(E|¬G) = 0.02. A judge or jury may be interested in the inverse probability, i.e. given a match, what is the probability that that the suspect is guilty (i.e. P(G|E). Calculation of the inverse probability is usually simple, but it is not trivial. A known fallacy in this respect is the *prosecutors fallacy* [1]. The prosecutor may reason: ``If the suspect did not commit the crime, the chance of a match is only 2%. Therefore the match is 98% proof that the suspect was the perpetrator''. That this reasoning is wrong, is best illustrated by another fallacy, the *defenders fallacy*, where the defender of the suspect reasons: ``The town where the crime took place has 40,000 inhabitants, so the evidence shows that 800 persons are as likely as my client to have committed the crime, Thus, the match contributes hardly anything to the evidence''. In fact the conditional probability is dependent of the number of suspects. So, if all the inhabitants of the town are as likely to have committed the crime as the suspect, then P(G|E) = 1/800 as the defender stated. However, if there are only 10 suspects, who are equally likely to have committed the crime as the given suspect, then the conditional probability is much higher: P(G|E) = 0.85, and the match of the hairs contributes a lot to the evidence. Bayes' rule (appendix A) allows to do these calculations.

1.1.2 Bayesian view

Many events cannot be predicted with frequentist methods, because the frequency with which they occur is too low to allow for statistical inferences. This is usually the case in judicial procedures. However, people can still have subjective beliefs about uncertain events. Many factors can serve to support this belief. For instance, John has a certain belief that football club Ajenoord will win the football match against Feyax next week. His belief is supported by many football games that he has seen in the past. We can symbolise the knowledge of John about the next game with the character K. If we denote the statement that Ajenoord will win with character A, then we denote P(A|K) as the subjective belief of a certain person that Ajenoord wins, given knowledge K. Another football aficionado, Jill, knows that the libero of Ajenoord has just got a flue. Her knowledge K' about the football match is different from that of John, and so is probably her subjective belief P(A|$K^{'}$). If Jill shares her knowledge with John, then John has to revise his belief. But even if John and Jill share the same knowledge, their beliefs may differ. This is the Bayesian view on probability. Bayes' rule is a method to revise ones belief in a certain event, given new knowledge. Many people may find it hard to express their belief in numbers. They have a feeling, that since it is not possible to indicate precisely the chance that a given statement is true, they might better hypothesise that the statement is true or hypothesise that it is false and then reason with this fictitious assumption. However, a rough indication of possibilities (like ``I think the chance that this statement is true is somewhere between 85% and 95%'') is much more accurate than a statement of belief (like ``this statement is beyond reasonable doubt''). People indicate their degree of belief in statements when they gamble. If John accepts a bet of 10 to 1 that Ajenoord wins next football match, he indicates that his belief in this event is better that 90%. If he does not accept a bet of 20 to 1, he indicates that his degree of belief is less than 95%. Wagenaar [9] describes how courts could use Bayesian calculations to use expert witness testimony to revise their opinion e.g. in the suspect being guilty. He states, that it is the task

of the court, and not of the expert, to form an opinion about the case. It is the task of the expert to indicate the ratio of conditional probabilities that the fact to be proven is true resp. false, given his expert knowledge and observations. The judge incorporates these ratios in his belief in the fact being true, using a Bayesian calculation.

1.2 Bayesian Belief Networks

The previous section elaborated on the influence that a single piece of evidence has on the probability of a single other event. However, Bayes' rule enables to do probability calculations on chains of events that influence each others probability of occurring. This is exploited in the *Bayesian Belief Network* (BBN). A BBN is a directed graph of connected nodes. Every node symbolises an event and contains a table of probability variables. Each node represents a variable. This variable can be binary (e.g. the suspect is guilty or not), multiple-valued (e.g. the number of goals that Ajenoord scores in a match) or continuous (e.g. the amount of time that the sun shines on a given day). The nodes are connected to each others with arrows, that indicate that they influence each other. BBN's enable us to reason about uncertainty. A decade ago the calculations needed to propagate the probabilities through the nodes of complex BBN were prohibitively time-consuming, but the increased capacity of modern computers has attracted interest in the implementation of BBN's.

2. Example of the use of Bayesian belief network The following example has been inspired by Prakken [7], who presented a case in which two alternative scenario's had been presented to the court as causes of a nocturnal car crash (Dutch Supreme-Court Decision HR 23 October 1992, NJ 1992, 813). The facts were as follows: In a single-car accident at about 3:00 am, a car crashed against a tree. After the crash, police officers entered the scene of the accident. They noted the following facts:

- The female driver was sitting on the ground, next to the car, and stated three times that ``he'' had pulled the handbrake.
- A heavily wounded passenger was sitting on the front passenger seat.
- The handbrake was in pulled position.
- The car had been driven through a curve in the road right before it crashed.
- There were tyre marks from blocked wheels in the curve of the road.
- There were tyre marks from a skidding (yawing) car. The marks led to the place of the accident.

Afterwards, both the driver and the passenger could not remember anything about the accident itself. The accident took place while they were returning from a party and the passenger had drunk alcoholic liquor. Car experts witnessed, that in a car of the type that was involved in the accident, the rear wheels would probably block if the handbrake was forcefully pulled during driving. The (insurance organisation of) the passenger sued the driver as being responsible for the accident and liable for the damages, because it is the duty of the driver to keep control over the vehicle. Although generally, in a single-car accident the driver is held responsible, in this case the driver stated that the passenger was responsible for the accident because he pulled the brake, making the car uncontrollable. The passenger had the burden of proof. He argued, that, under the Dutch law, in a single-car accident the responsibility of the driver is assumed to be proven, so that in this case the burden of proof goes to the driver. The supreme court concluded that, in this case, because of the probability of the passenger having pulled the handbrake, the single fact of this being a single-car accident cannot be considered a proof for the claimant.

2.1 Abductive-logical reconstruction

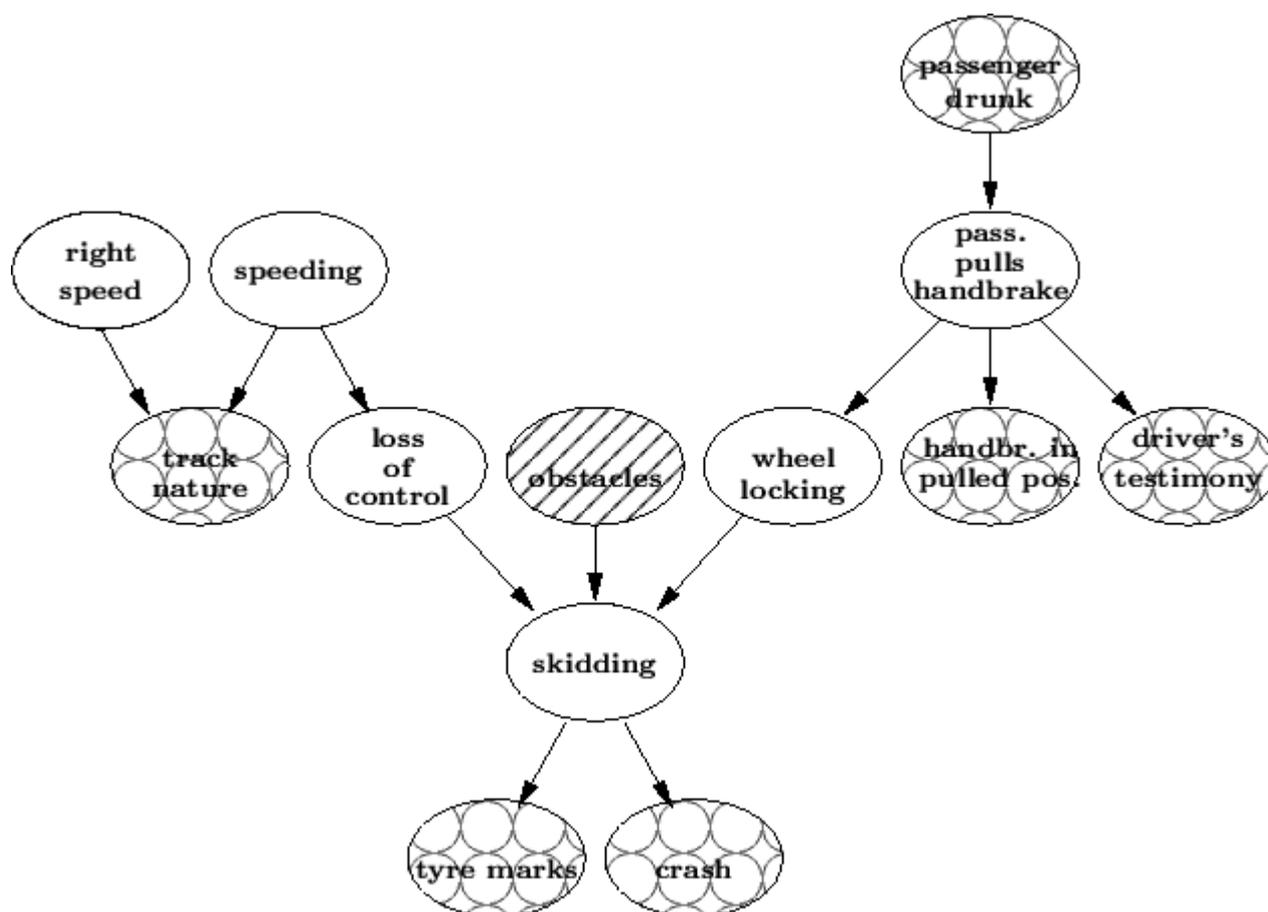Prakken built a causal structure model of this case (figure 1).

Figure 1: Causal structure for the example case, as depicted by [7].

Observations and hypotheses are depicted as ovals and arrows indicate whether facts are true or false match other observations/hypotheses or whether they contradict each other. True facts are indicated with a pattern of circles, and false facts with a pattern of lines. Prakken derived from this model that the plaintiffs claim (driver was speeding and lost control over the car) explained one observation (presence of tyre marks), but contradicts another observation (nature of the tyre marks) whereas the defenders claim (passenger pulled handbrake) explains three observations (tyre marks, driver's statement that "he" pulled the handbrake and the handbrake's pulled position after the accident)) and contradicts nothing. Prakken concludes that the defendant's solution is better than that of the plaintiff.

2.2 Bayesian Belief Model

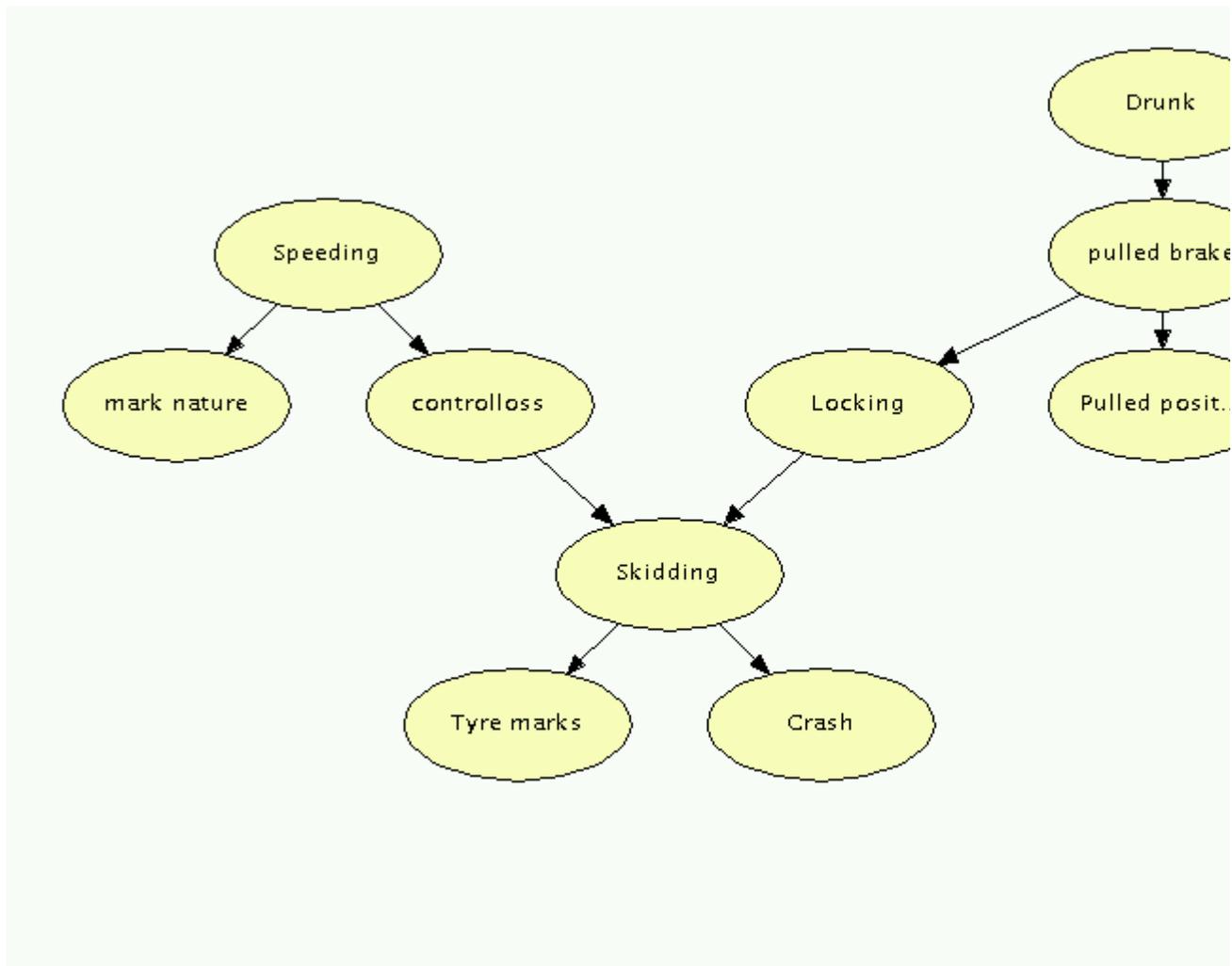Figure 2 shows a model that looks the same as that of figure 1, but in fact is a Bayesian belief network.

Figure 2: Causal structure for the example case as a Bayesian Belief Mo

The model has been constructed with Hugin Light (Hugin Expert, Denmark), a small version of a commercial software system that can be freely downloaded from `www.hugin.com`. Every top node contains the prior beliefs that the fact or hypothesis in the node is true or false. Daughter nodes contain the prior conditional beliefs that they are true for every combination of truth/falsehood of the parent nodes. Two nodes that are present in Prakken's model have been left out of the BBN model. According to the text of the verdict, the node ``Slowing down in S-curve'' seems to be no more than the negation of the node ``Speeding in S-curve''. ``Speeding'' means in this example ``driving with a speed that causes the driver to loose control over her vehicle'' and ``Slowing down ...'' means ``Adjusting the speed so that the driver keeps in control over her vehicle''. The node ``Obstacles'' has been left out because obstacles are not mentioned in the verdict. Two sets of prior belief variables have been used: one set with values that seem the most likely by the writer, and another set, in which the values have been set in the advantage of the plaintiff (table 1).

Table 1: Prior probabilities of the nodes.

Node	Y/N	O	Plaint
Drunk	Yes	0.05	0.02
	No	0.95	0.98
Pulled │ Drunk	Yes	0.03	0.00001
	No	0.97	0.99999

Pulled | ¬Drunk	Yes	0.0	0.00001
	No	1	0.99999
Speeding	Yes	0.00001	0.0001
	No	0.99999	0.9999
Tyre Mark nature | Speeding	Yes	0.2	0.4
	No	0.8	0.6
Tyre Mark nature | ¬Speeding	Yes	0.7	0.6
	No	0.3	0.4
Control-loss | Speeding	Yes	1	1
	No	0	0
Control-loss | ¬Speeding	Yes	0.99999	0.9999
	No	0.00001	0.9999
Wheels-lock | pulled	Yes	0.8	0.5
	No	0.2	0.5
Pullpos | pulled	Yes	0.99	0.90
	No	0.01	0.1
Pullpos | ¬pulled	Yes	0.001	0.05
	No	0.999	0.95
Testimony | pulled	Yes	0.9	0.8
	No	0.1	0.2
Testimony | ¬pulled	Yes	0.03	0.1
	No	0.97	0.9
Skidding | (Control loss ∧ lock)	Yes	1	1
	No	0	0
Skidding | (Control loss ∧ ¬lock)	Yes	1	1
	No	0	0
Skidding | (¬Control loss ∧ lock)	Yes	1	1
	No	0	0

Skidding $\mid$ ($\neg$Control loss $\wedge$ $\neg$lock)	Yes	0	0
	No	1	1
Tyre marks $\mid$ skidding	Yes	1	1
	No	0	0
	No	1	1
Tyre marks $\mid$ $\neg$ skidding	Yes	0	0
	No	1	1
Crash $\mid$ Skidding	Yes	0.2	0.2
	No	0.8	0.8
Crash $\mid$ $\neg$ Skidding	Yes	0.0001	0.0001
	No	0.9999	0.9999

A set in the advantage of the defendant has not been implemented, because both sets resulted in outcomes that were highly in favour of the defendant.

2.3 Comments on the prior probabilities

The prior probabilities indicate the amount of belief of the truth of the hypothesis expressed by the node, when the evidence presented in the case is not considered. For instance, in the node ``Crash'' the estimated possibility must be given, that a car crashes, given that it started to skid, and the estimated possibility must be given that a car crashes, given that it did not skid. The prior likelihood that a car crashes in a single-car accident in a given section of a road is extremely unlikely, and more so, if the car didn't skid. Actually, the odds of this happening are not important in our problem, because both the alternative explanations of the accident assume the car to skid. The top node *drunk* contains the belief that this passenger was drunk, given that he sat in the passenger's seat in the middle of the night, returning from a birthday party. Considering the time of the night this is not extremely improbable. *Speeding*, as we defined it in our problem, is a highly improbable event. Possibly one out of 100.000 or one out of a million cars skid in a given stretch of the road. The belief in a skidding event can be based on statistics on single-car accidents, but a judge could possibly modify the statistic probability, e.g. when he thinks that the driver in this case is more (or less) likely to become involved in accidents than the mean driver, e.g. given the age and gender, the appearance and the time of the day.

2.4 Results

Figure <u>3</u> is a figure of the BBN model, after the ``objective'' values have been entered.
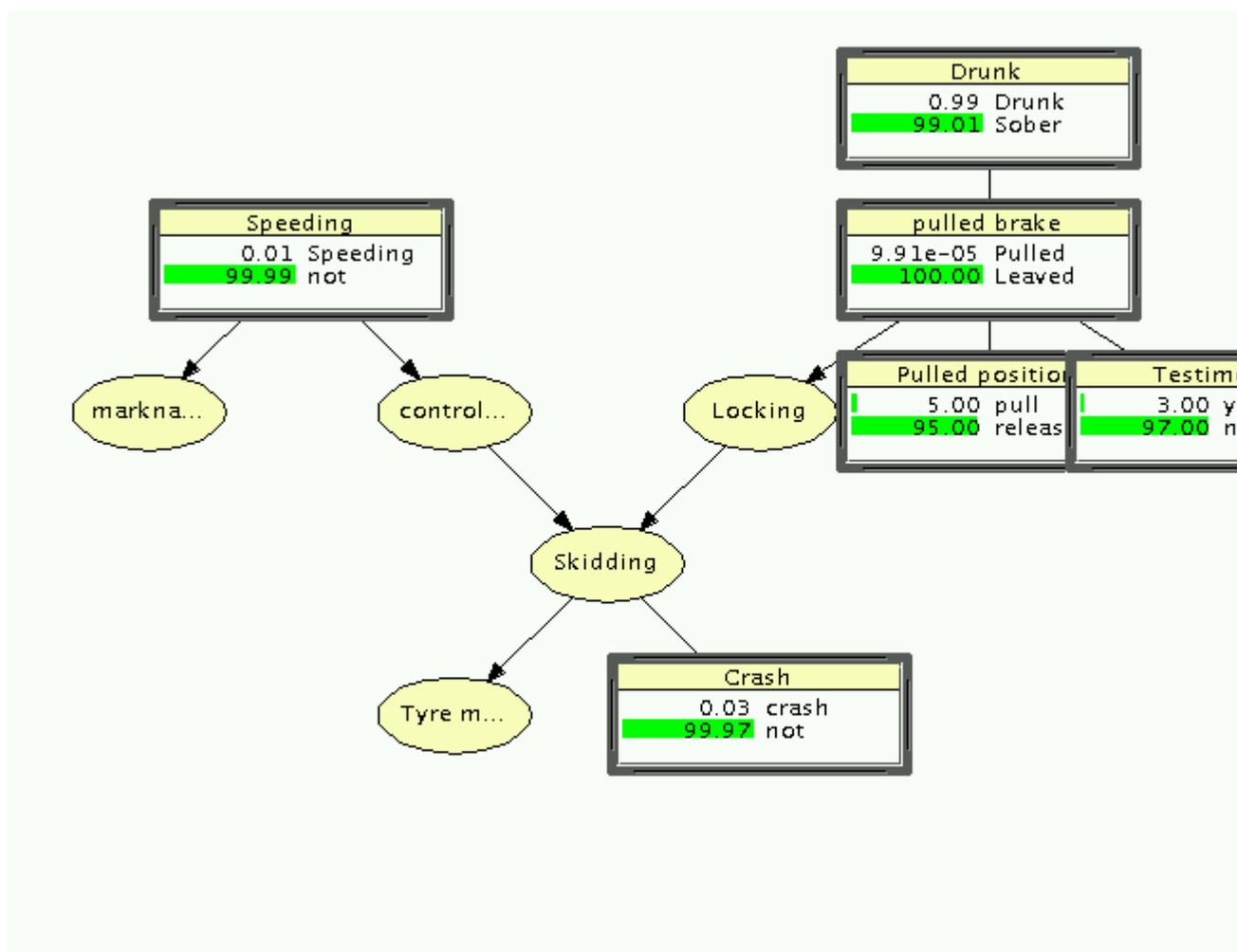
Figure 3: Initial probabilities after entering the "objective" values.

Some of the nodes have been replaced by their probability tables. Now we enter the fact that a crash actually happened by fixing its probability to unity. (figure 4),
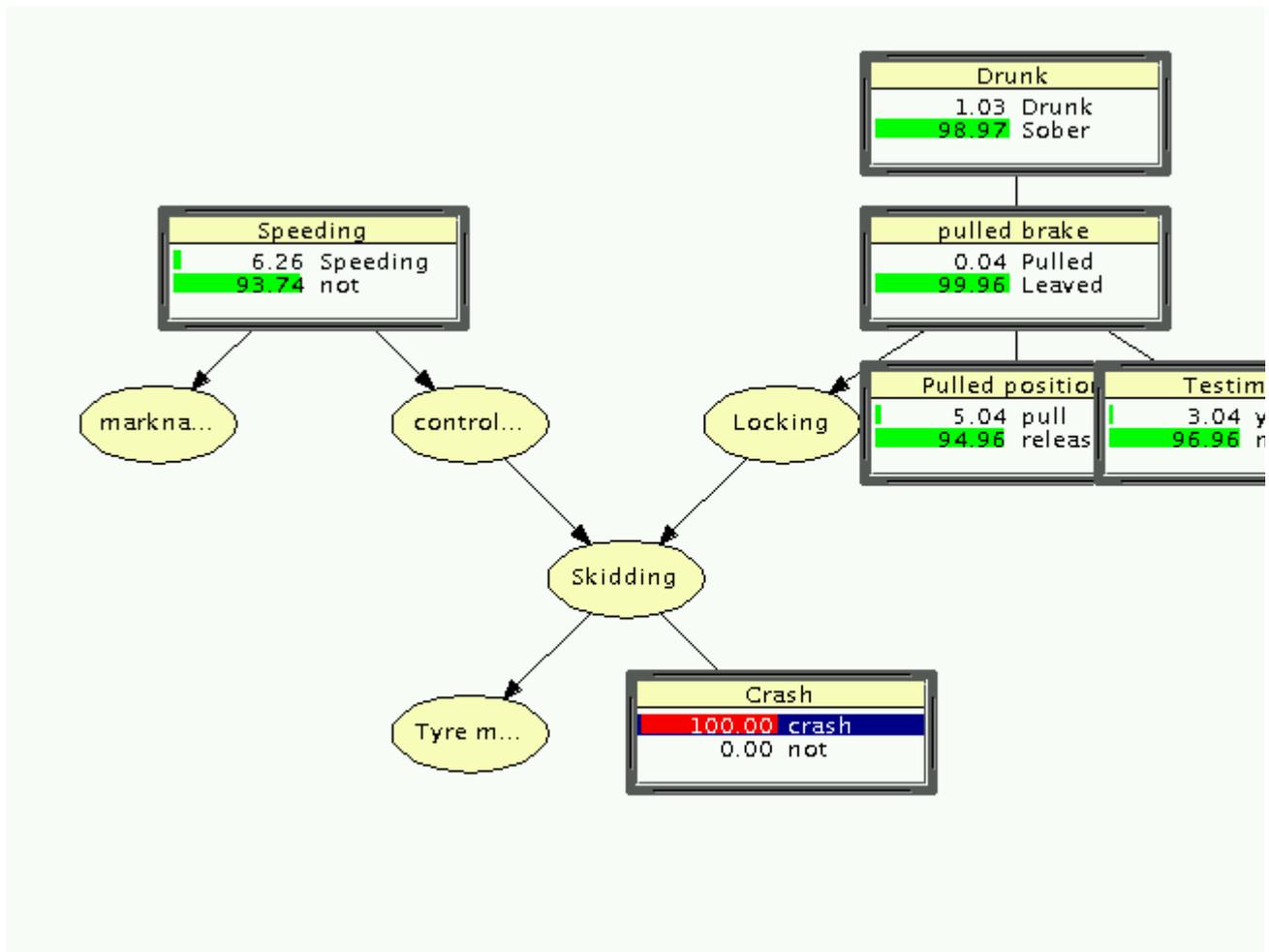
Figure 4: Initial probabilities after entering the fact ``crash''.

According to the model, it is more probable that the crash has been caused by speeding than by pulling the handbrake. If, however, we include other pieces of evidence (i.e. the driver's testimony that the passenger pulled the handbrake), the observed pulled position of the handbrake and the fact that the passenger was drunk, pulling the handbrake turns out to be a much more likely cause of the accident than speeding is (figure 5).
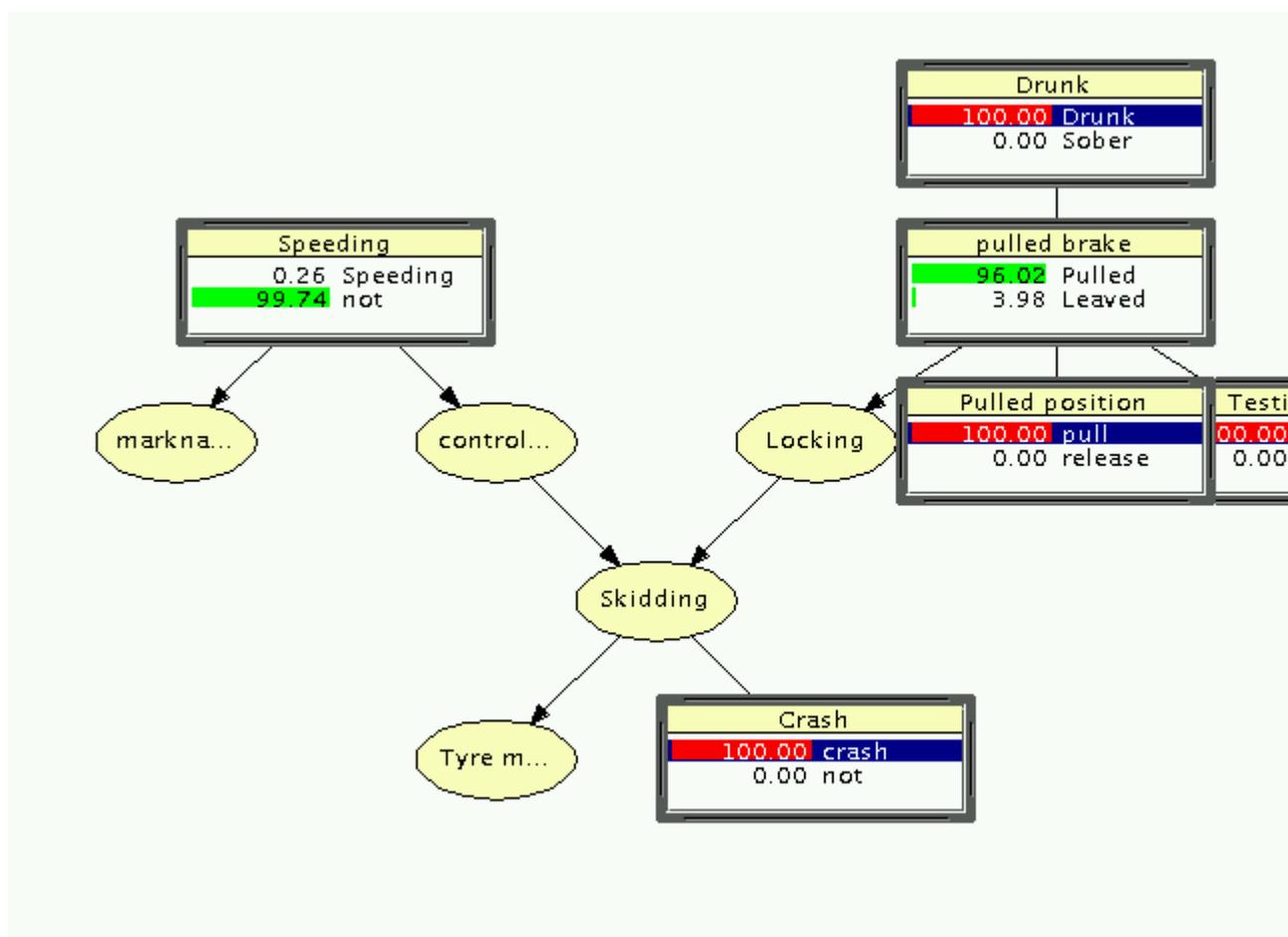
Figure 5: Probabilities after entering the facts of the case.

If the probabilities that are most favourable for the plaintiff are filled in, it turns out that the likelihood of the crash being caused by pulling of the handbrake (30.60%) is still comparable to that of speeding (27.82%). The node ``mark nature'' has not been used. In the sentence it was not explicitly stated that the nature of the marks was proof for not speeding, but that ``[the tyre marks] give insufficient support for the suggestion of [plaintiff] that [defendant] had speeded'' and that ``[the tyre marks] suggest that [defendant] had adjusted her speed to the situation and had kept her vehicle under control''. To use this kind of information the causality would have to be changed in the model.

3. Discussion

In this paper a Bayesian belief network has been presented as a possible tool to support legal reasoning, as an alternative or as a supplement to logics. The advantage is, that the plausibility of the argumentation can be quantified. The major difference between the Bayesian model of the car accident and the argumentation model is, that the argumentation model provides for both hypotheses the facts that they explain and the facts that they don't explain or that they even contradict, whereas the Bayesian model provides for both hypotheses the likelihood that they have occurred, given the facts (hence, probides more abductive reasoning). The argumentation model does not give insight in how much the explanation of a fact contributes to the evidence of the case. For instance, if it is considered highly unlikely that after an accident the handbrake is in pulled position, the fact that the handbrake *was* in pulled position contributes a lot to the credibility of a hypothesis that explains this fact. If, on the other hand, a judge considers it very likely that the driver has a very strong habit to pull the handbrake when she leaves the car, so that she even does it right after an accident as the present one, the pulled handbrake contributes less to the credibility of a hypothesis that explains the pulled handbrake. The BBN model reflects this in a natural way. Surprising facts contribute more evidence than unsurprising facts. With a Bayesian model it is easy to explore the case, e.g. to analyse

what the outcome would be when all the probabilities are adjusted in favour of one or the parties, or when they are adjusted in favour of the other party. This results in an upper and lower limit of the probability that the event to be proven had actually happened, and this may help the judge to form his opinion on the case. The model could be extended with a ``decision node'' that calculates the ``cost'' of a decision. In this case, the cost means the harm that a wrong decision would inflict. In many cases the harm done by an unjustified positive decision is different from the harm done by an unjustified negative decision [5]. For instance, when a large and wealthy company claims a large sum of money from a private individual, the damage done to the individual by an unjust allotment is much larger than the damage done to the company by an unjust dismissal of the claim. The decision node could calculate for both the decision for allotment and for denial the product of the probability that the decision is wrong and the cost of the wrong decision. In this way the computer calculates the optimal decision in which the expectation value of the outcome has its maximum. To be really useful for lawyers, a specialised Bayesian tool would have to be made that facilitates easy construction of models of judicial cases, provides support for the users and probably contains statistical material that can be used as basis for the prior probability estimations. Such a tool would also be very useful to train law students in exploring and evaluating judicial cases, to get a feeling for probability and statistics, and to avoid common pitfalls in reasoning with probability.

3.1 Objections against use of probability in judicial fact finding

The use of mathematical/statistical tools in sentencing has for a long time been a topic of discussion. A full discussion about this issue is not within the scope of this paper. However, some points can be raised here. One of the issues is, that use of statistical methods by non-statisticians can lead to errors and hence, false decisions. Kerkmeester [6] uses this as an argument to avoid use of statistical methods in court. However, in the first place, the use of argumentation of a statistical character is sometimes unavoidable, and does already sometimes lead to errors [9,3]. A well-supported BBN could be of help to recognize and avoid the statistical pitfalls. By explicitly modelling the problem, the computer can clearly indicate where possible pitfalls are, and invite the user to motivate her assumptions as clearly as possible. Another objection against use of statistics in judicial decisions is that it is hard to express probabilities in numbers, and that, as a result, emphasis will be put on those factors that are easy to quantify and that other factors, that may be really important, are neglected. However, a Bayesian model as presented in this paper enforces the user to enter probabilities for all the factors that play a role in the causality. Therefore, it is not expected that this would cause problems.

4. Appendix: Bayesian probability

Bayes' formula can be explained as a formula to inverse conditional properties. Suppose A and B are two events that may occur. Define $P(A)$ as the probability that event A occurs and define $P(B)$ as the probability that event B occurs. Suppose further that the fact that one of the events did actually occur influences the possibility that the other event did also occur. Let $P(A,B)$ be the probability that both events occured, and let $P(A|B)$ be the conditional probability that event A occured, given that event B did actually occur. The probability that A and B both occur is equal to the product of the probabilities that B occurs and the conditional probability that A occurs:

(1)

$$P(A,B) = P(A|B)P(B)$$

We can interchange A and B in the previous equation:

$$P(A, B) = P(B|A)P(A). \tag{2}$$

Hence, combining equations 1 and 2:

$$P(A|B)P(B) = P(B|A)P(A) \tag{3}$$

and thus:

$$P(A|B) = P(B|A)\frac{P(A)}{P(B)} \tag{4}$$

This is Bayes" equation. The simplest explanation of this equation is, that it derives the inverse conditional probability. A more ``Bayesian'' interpretation is, that it derives how a piece of evidence should modify one's believe in the occurrence of event A. The following case example illustrates this: A crime has been committed, and a carpet fibre has been found at the crime scene. It is estimated that five percent of the households has a carpet with the same fibres. A suspect has been identified. The best estimation of the likelihood that this suspect committed the crime was 30%. It turns out, that the household of the suspect contains a carpet of which the fibres match the fibre found at the scene of the crime. Therefore, the belief in the likelihood that this subject committed the crime must be revised. Define A as: ``the suspect committed the crime'', and $B$ as ``the fibre matches those of a carpet in the suspects household''. Assuming that B is a certainty if A has occurred (i.e. $P(B|A) = 1$), the prior probabilities are:

$$P(A) = 0.3 \tag{5}$$

$$P(B) = P(B|A)P(A) + P(B|\neg A)P(\neg A) \tag{6}$$

$$= 1 \times 0.3 + 0.05 \times 0.7 \tag{7}$$

$$= 0.335 \tag{8}$$

The posterior probability that the suspect committed the crime is:

$$P(A|B) = P(B|A)\frac{P(A)}{P(B)} \tag{9}$$

$$= 1 \times \frac{0.3}{0.335} \tag{10}$$

$$= 0.8955 \tag{11}$$

This means, that the new evidence material adds a lot to the probability that the suspect actually committed the crime. However, when, later on, it turns out that the prior probability of guilt has no foundations (e.g. because it was based on what later turns out to be a false witness testimony), the calculation has to be revised. Suppose, there live 10000 persons in the neighbourhood who are equally likely to have committed the crime as our suspect, then:

$$P(A) = 0.0001 \tag{12}$$

$$P(B) = P(B|A)P(A) + P(B|\neg A)P(\neg A) \tag{13}$$

$$= 1 \times 0.0001 + 0.05 \times 0.9999 \tag{14}$$

$$= 0.0501 \tag{15}$$

and

$$P(A|B) = P(B|A)\frac{P(A)}{P(B)} \tag{16}$$

$$= 1 \times \frac{0.0001}{0.0501} \tag{17}$$

$$= 0.002 \tag{18}$$

The new evidence does hardly contribute anything to the probability of guilt. In the previous example the evidence has a discrete, binary character (e.g. evidence is present or it is not present). However, Bayes' equation can also be used for continuous variables that assume multiple values (e.g. the outcome of a thrown dice can be any integer value between 1 and 6) or continuous variables (e.g. the speed of a given car on a given road at a given moment).

Bibliography

1   C. Aitken.
Lies, damned lies and expert witnesses.
Mathematics today, 32(5/6):76-80, 1996.

2   Th. Bayes.
An essay toward solving a problem in the doctrine of chance.
Philosophical Transactions of the Royal Society, 1763.

3   N.E. Fenton and M. Neil.
The jury observation fallacy and the use of bayesian networks to present probabilistic legal arguments.
Mathematics Today, 36(6):180-187, 2000.

4   M.O. Finkelstein and W.B. Fairley.
A bayesian aproach to identification evidence.
Harvard Law Review, 83:489-517, 1970.

5   J. Kaplan.
Decision theory and the fact finding process.
Stanford Law Review, page 1065, 1968.

6   H.O. Kerkmeester and R.V. de Mulder.
Statistiek en bewijsrecht.
Nederlands Juristenblad, 62(3):73-81, 1987.

7   H. Prakken and S. Renooij.
Reconstructing causal evidence: a case study.
In B. Verheij, A.R. Lodder, R.P. Loui, and A.J. Muntjewerff, editors, *Jurix 2001: the*

fourteenth annual conference, pages 131-142. Jurix, IOS press, 2001.

8   L.H. Tribe.
        Trial by mathematics: Precision and ritual in the legal process.
        Harvard Law Review, 84:1329-1393, 1971.

9   W.A. Wagenaar.
        The proper seat: a bayesian discussion of the position of expertwitness.
        Law and Human Behavior, 12:499-510, 1988.

Paul Huygen
2002-04-28