An Analysis of Model of Rational Proof: The Simonshaven Case

Name Student: Tessa van Dijk
First Monitor: Prof. dr. mr. H. (Henry) Prakken
Second Monitor: Dr. A. (Annemarie) Borg

Date: 26-06-2020
ECTS: 7.5
Study: Bachelor Artificial Intelligence, Utrecht University

Index

1. Introduction

In most courtrooms across the world, criminal cases use legal evidence to draw conclusions about the events that have taken place. The goal of the prosecution is to convince a judge or jury that the defendant is guilty of a crime. However, judges and juries are known to have difficulties when reasoning about legal evidence (Vlek, Prakken, Renooij, Verheij, 2014), they are human after all.

A well-known case of this happening is the Lucia de Berk case, where a Dutch nurse was convicted based on quasi-statistical considerations and was later acquitted of the charges.

To be able to remedy these common fallacies, ways have been created to formally represent the evidence and reasoning of a legal case. Within the formal study of legal rational reasoning, three approaches can be distinguished, focusing respectively on arguments, scenarios, and probabilities (Kaptein et al. 2009). In recent years, these three approaches have been combined and integrated to create new ways of formalizing reasoning and legal evidence. In 2019 a special issue of *Topics in Cognitive Science* was released dedicated to model rational proof in law (Prakken, Mackor, Bex, 2019)

In this issue, a variety of approaches were applied to a case study, this paper will be considering four of them. The first author, Prakken, applied an argumentation-based analysis to the Simonshaven case, Van Koppen and Mackor applied a scenario-based approach, Dahlman proposing a more basic application of Bayesian thinking, meant to help fact-finders to avoid certain fallacies and biases, and Fenton et al. applied a Bayesian network-based approach.

The aim of this paper is to establish common ground between the four approaches, point out their differences, what their outcomes are and how these differ among their peers. This will be done by means of studying some of the answers given to the questions posed at the beginning of the special issue of *TopiCS*, for which each author has written their respective answers, namely:

- To what extent is the analysis objective and to what extent is it based on subjective beliefs, assumptions, and choices?
- How natural is the analysis from a cognitive and legal point of view?
- Did you analysis identify errors or biases in the reasoning of the judge, persecutor, or defense? (Prakken, Mackor, Bex, 2019)

We will also look at the differences in granularity and precision of the approaches, and at common criticisms made about the approaches.

1.1. Theoretical imbedding

The approaches discussed in this paper have a theoretical background in AI. The Bayesian networks discussed in section 3 under Fenton et al.'s approach have been a part of AI systems and their reasoning. Also, the argumentation-based approach discussed by Prakken has been developed in terms of logical frameworks for argumentation in legal applications of AI.

Modelling these complex reasoning and legal evidence with these approaches and techniques, is relevant to the field of artificial intelligence because it offers insight into the complex structure of human reasoning regarding the field of law. It is helpful to construct these models, not only for the legal cases themselves, to make rationally sound decisions, but also to the further development of intelligent systems. Rational reasoning can be difficult for AI systems, therefore insight to human reasoning can help research. Hopefully, this paper will be able to

further the discussion about rational reasoning, not only in court cases but also in the field of artificial intelligence.

This paper is organized as follows. In Section 2, the facts and evidence of the case is summarized and briefly discussed. In Section 3, each of the four main articles' method and results are summarized. In Section 4, the results are compared and discussed in accordance with the questions posed in the introduction, also common criticisms on the approaches will be discussed here.

2. The Case

This case concerns the death of Jenny Lourens, who was murdered in a recreational area near the town of Simonshaven, on the night of the 11th August 2009. The initial prime suspect of the case, Ed Lourens (EL) , was the victim's husband. According to Ed, he and Jenny were walking in the area when a madman attacked them and killed Jenny. He then called his daughter-in-law and went to the road to stop a car. According to the prosecution there was no evidence that there had been a madman present (several eyewitnesses claimed they had not seen another man in the vicinity of the crime scene at that time) and the suspect had the opportunity, the means (a supposed gun that was not found), and a motive (a rough marriage) to kill the victim. This prompted the court to convict EL of the crime. Later, evidence was found that there had been, in fact, a madman active in the area and time of the crime. The second suspect was called Perry S., in his house a map was found with markings indicating spots near the place where the murder of Jenny Lourens had taken place. This prompted the case to go into appeal.

The articles discussed in this paper, considered the evidence and argumentation presented in the appeal.

3. Approaches

3.1. Approach 1: Argumentation-based approach (Prakken)

3.1.1. Method description

This approach concerns arguments, in which arguments are defined as a series of inferences from evidence to conclusions (Prakken, Bex, Mackor). The argumentation model used in Prakken's paper, is the ASPIC+ framework (Modgil, 2014; Modgil & Prakken, 2013). In this framework, arguments are constructed as trees of inference steps. In these trees the leaves are the premises, the root is the conclusion and remaining nodes are the intermediate conclusions. Arguments can also have subtrees, containing the sub arguments.

There are three types of attacks on an argument:

- *Undermining attack*, where an argument is attacked on its premises.
- *Undercutting attack*, where an argument is attacked on its defeasible inference steps.
- *Rebutting attack*, where an argument is attacked on its derived conclusions.

Whether an attack survives or is defeated by these attacks, splits arguments into three categories:

- *Justified arguments*, which are arguments that survive all conflicts and thus can be accepted.

- *Overruled arguments*, which are arguments that are defeated by a justified argument and thus are rejected.
- *Defensible arguments*, which are arguments that are involved in irresolvable conflicts.

To determine the dialectical status of the arguments Prakken applied Dung's calculus (1995) in a labeling-based form. This meaning an assignment of either the label *in* or the label *out* to zero or more arguments of the framework, such that an argument is *in* iff all arguments defeating it are *out*, and an argument is *out* iff it is defeated by an argument that is *in*.

The labelling policy adapted by Prakken in his paper is the so-called *preferred labeling*, which maximizes the set of arguments labeled *in*.

3.1.2. Results

The analysis of the Simonshaven case done by Prakken resulted in the argumentation framework found in the Appendix A.1. Two labelling possibilities of the framework were given, one in which argument *A* (whether the suspect killed the victim) is labelled *in* (Appendix A.2) and one in which *A* is labelled *out* (Appendix A.3). Because *A* is not *in* in all labeling possibilities of the framework, *A* is a *defensible argument*. Therefore, the conclusion is, based on the present analysis, that a conviction was not justified. However, if *C*'s attacked sub arguments are preferred over their attackers *F* and *F'*, then argument *A* becomes justified.

Concluding, the main issues identified by Prakken are: (1) if the argument that Perry is a plausible alternative attacker of the victim is sufficiently strong to be able to defeat the argument for the opposite conclusion, (2) if a proper investigation was done into the traces of an alternative attacker, and (3) if the suspect's behavior after the crime can be explained if Perry is the attacker.

3.2. Approach 2: Scenario-based approach (van Koppen & Mackor)

3.2.1. Method description.

The scenario approach understands reasoning about evidence in terms of explanations. It states that fact finders should construct and compare alternative explanations of the evidence. The core question in the scenario approach is "Why should we believe it?" The answer to this question results in sub scenarios supporting the main scenario. Then the same question is applied to these sub scenarios, resulting in sub-sub scenarios. This process can go on indefinitely, but usually does not. When this process stops, it means that the analysis either arrived at a sub scenario that (1) is recognized as a fact by all parties, (2) that is not challenged by one or both parties, and/or (3) that cannot be further scrutinized because the necessary data is absent. Using scenarios, fact finders are able predict evidence. A distinction is made between two types of predictable facts: *new* facts (facts that will take place in the future) and *novel* facts (facts that are unknown but already exist when the prediction is made).

The scenario approach bases itself on six core ideas: scenarios, IBE, coherence, falsification, alternative scenarios and discriminating evidence, and the creation, accommodation, and prediction relations between scenarios (van Koppen & Mackor, 2019).

3.2.2. Results:

The main scenarios considered in van Koppen and Mackor's analysis are scenario S, the prosecutor's scenario and scenario AS, EL' scenario. Scenario S states that EL is the culprit whilst Ed's scenario states that a madman was the culprit.

In the defense original statement there was little to offer, there was little to no evidence of the AS scenario. In the appellate court both *new* and *novel* facts came into view that changed the initial ruling of the case. However, the AS scenario was not detailed enough to allow for precise predictions and therefore made, in the initial court ruling, scenario A the 'best' scenario.

3.3. Approach 3: Probabilistic-based approach (Dahlman)

3.3.1. Method description: Summary of the applied approach.

The probabilistic approach often entails using Bayes Rule to make estimations about the probability of certain hypotheses, given some evidence. More formally, in the odds version, Bayes Rule is stated as follows:

$$\frac{P(H)}{P(-H)} \times \frac{P(E|H)}{P(E|-H)} = \frac{P(H|E)}{P(-H|E)}$$

$P(H)/P(-H)$ are referred to as the *prior probabilities*, meaning the probability that the hypothesis is true before considering the evidence (Dahlman, 2019). With $P(H)$ being the probability of the hypothesis is true and $P(-H)$ the probability of the hypothesis being false. These two values are additive and therefore should add up to one.

$P(E|H)$ is the probability of the evidence, given the hypothesis is true, this value is also referred to as the *"hit rate".* Analogously, $P(E|-H)$ is the probability of the evidence, given that the hypothesis is false, this value is also referred to as the "*miss rate"*. These values divided are known are the *likelihood ratio*.

One of the main characteristics of evidence this approach takes is mind is dependence. If two pieces of evidence support a hypothesis independently, meaning that their probabilities have no influence on each other, the likelihood ratio of one piece of evidence is multiplied by the second, formally:

$$\frac{P(E_1|H)}{P(E_1|-H)} \times \frac{P(E_2|H)}{P(E_2|-H)}$$

If the two pieces of evidence are dependent, meaning that they influence each other's probabilities, the value is calculated as follows:

$$\frac{P(E_1|H)}{P(E_1|-H)} \times \frac{P(E_2|H,E_1)}{P(E_2|-H,E_1)}$$

Dahlman also discussed in his article three common fallacies/biases in probabilistic thinking. Firstly, there are *false dichotomies*. This meaning that when deciding two events are treated as mutually exclusive when they are not (Cederblom & Paulsen, 2011). In legal evidence this occurs when a fact-finder incorrectly treats the prosecutor's hypothesis (H) and one of the variety of alternative hypotheses (e.g. h_1), as mutually exclusive, so $P(H) + P(h_1) = 1$ and

therefore disregarding other (non-zero) alternative hypotheses. Secondly, there is the *dependence neglect*. This happens in a faulty use of *Bayes theorem*, where it is assumed that pieces of evidence are independent when they are not. Lastly, there is the *miss rate neglect*. This happens when there is an attentional bias towards the *hit rate*, and therefore the *miss rate* is neglected.

3.3.2. Results

In his analysis, Dahlman attempted to approximate the likelihood ratios with regards to the prosecutor's hypothesis (H), thus the assumption that EL is guilty.

1. *Motive.* According to Dahlman, the evidence is more likely if EL is guilty, than if he is innocent. Likelihood ratio of 2
2. *Opportunity.* To calculate the miss rate here it was assumed that in 20% of the cases the husband is present when the wife is murdered, and the husband is innocent. Likelihood ratio of 5.
3. *Character.* Because character in most courts is not accepted as evidence as guilt a conservative estimation was made. Likelihood ratio of 2.
4. *Evidence contradicting EL's story*. The miss rate was calculated based on the probability (with and without considering the evidence) of each alternative attacker, in this case, a robber, a serial killer, a crime of passion or a gang related member. Likelihood ratio of 50.
5. *Strange behavior*. Dahlman concluded that the bundle of evidence was to be more likely if EL were the culprit, rather than if EL were innocent. Likelihood ratio of 5.
6. *Blood splatter*. The probability according to Dahlman of EL having blood on his shoes was equal in both the case that EL were innocent, then if he were guilty. Likelihood ratio of 1.
7. *Gunshot residue*. The gunshot residue found on Jenny's head did not belong to Ed's gun, however the gunshot residue found on EL did come from Ed's gun. This evidence is more likely if EL were innocent. Likelihood ratio of 0.25.
8. *Gun cartridges*. DNA on the gun cartridges found at the crime scene did not belong to Ed. Likelihood ratio of 0.20.
9. *No evidence against other suspects*. Equally likely. Likelihood ratio of 1.

The combining force of this evidence is therefore,

$$2 \times 5 \times 2 \times 50 \times 1 \times 1 \times 0.25 \times 0.2 \times 1 = 250$$

This means that it would be 250 times more likely that all the evidence was present in the case that EL is guilty, rather than innocent. However, the likelihood of the prosecutor's hypothesis overall depends on the prior likelihoods, the bigger they are, the more likely it is that EL is guilty. If this prior likelihood were to be calculated in the was Dawid (1993) calculates it, meaning, dividing 1 by the amount of possible perpetrators, it would mean that there would not be enough evidence to convict Ed, according to Dahlman.

Dahlman also discusses three main fallacies/biases with regards to the case. In his opinion the judge had applied all of them. Firstly, because there was no one else near the crime scene at the time of the crime, the court concluded that EL must have been guilty. This is a case of *miss rate neglect.* Secondly, the court did not consider EL's hypothesis plausible, and thus assumed the hypothesis of the prosecutor to be true. This is an example of a *false dichotomy*.

Lastly, because the court failed to consider the fact that EL being present at the crime scene and EL being the victim's husband are dependent of each other and is therefore a case of *dependence neglect.*

3.4. Approach 4: Probabilistic-based approach: Bayesian Networks (Fenton et al.)

3.4.1. Method description

The following approach is a more advanced way of applying Bayes rule, by means of a *Bayesian network*. A Bayesian network (BN) is a compact representation of a joint probability distribution, where probabilities are calculated based of their dependencies. They primarily consist of nodes, representing variables (which can be true or false) and arcs representing the dependency relationships between these variables (Fenton et al.,2019). The arcs are directional, and the graph does not contain any cycles. The strength of the relationships is represented by joint probability distributions. These JPDs are calculated by allotting the value true or false to each variable and calculating its outcome. An example a small BN with JPDs is given in Fig 1.

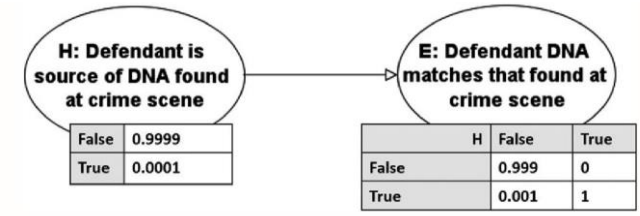


Fig. 1: Causal view of evidence, with prior probabilities shown in tables. This is a very simple example of a Bayesian network. (Fenton et al., 2019)

Bayes rule is used to calculate the probability of the hypothesis given the evidence. In this case the classical definition of the rule is used, defined as follows:

$$P(H|E) = \frac{P(E|H)P(H)}{P(E)}$$

To construct the BN for the Simonshaven case idioms were used. These were used to institute a systematic approach to easily develop a generally approved model for complicated legal cases. Three idioms were applied, the first one being the *evidence accuracy idiom*, which considers the accuracy of legal evidence. Secondly, the *cause-consequence idiom*, where different instances of idioms are joined. Lastly to consider are the *opportunity* and *motive* idioms. It bases itself around three main hypotheses, the main hypothesis of the case and the two hypotheses that are assumed to be its necessary cause, so "opportunity" and "motive". In the Simonshaven case also "capability" was added as a necessary cause.

3.4.2. Results

The resulting model can be found in appendix B. The model was built in AgenaRisk and can calculate the JPDs effectively. Before entering any evidence into the model, the probability that EL is the culprit is 1.08%, after all evidence is entered, it becomes about 74%.

A sensitivity analysis was done on the model and it identified the defendant's credibility as the source of evidence the case relies most heavily on. Followed by, forensic credibility and passing witness credibility for the pumping station.

According to Fenton et al. the model is robust, since out of the 99 independent parameters in the model, only seven can, on their own, push the posterior probabilities of guilt above 95%.

4. Results

4.1. Extent of objectivity of each approach.

Prakken rephrased the question by asking how likely it is that others applying the same method will produce the same or a similar outcome. The analysis done by Prakken in his paper is not very reliable in this sense, different analysts have different opinions on which pieces of evidence have the upper hand in the model how these pieces of evidence relate to each other. Another interesting insight mentioned in his paper is the observation that, rather than asserting that, in the case a suspect had the opportunity, means and motive to commit a crime and there is no other possible offender, the suspect must be guilty; it asserts that a more accurate summary of the argument scheme also includes the observation that if several observations can be explained if the suspect is the offender, then the suspect is guilty. Prakken considered Perry S to be the only plausible attacker. Interestingly, in Dahlman's analysis, other culprits were also considered plausible when calculating the likelihood ratio for point 4, the evidence contradicting EL's story, namely the serial killer scenario, the robber scenario, the other mobster scenario and the crime of passion scenario.

Dahlman's approach is not meant to model an entire case but rather it is designed as a tool to help factfinders avoid common fallacies and biases by establishing dependency relations and calculating exact values for likelihood ratios. These techniques might help factfinders, but they are by no means objective, the odds allotted to each piece of evidence are based on the creator's opinion, even if they resemble Bayes rule. That is why this approach cannot be called objective in any way.

This is something Fenton et al. improves upon by means of idioms. These idioms are set rules for the creation of BNs. An indication of its effectiveness is the fact that it was mentioned in their article that multiple people had also constructed a BN for the same case to compare. Fenton at al. claimed that besides differences in granularity, the models produced were very similar. Indicating a somewhat higher level of objectivity for this approach. However, they have not established the reproductivity of their model formally. A simple claim of the model being somewhat similarly reproduced by a handful of analysts does not back up the argument that a BN model is reproducible, and therefore subjective. What might only seem a small difference by them, might be a big difference to somebody else. These differences between models might even get bigger if they are used improperly by legal professionals with no background in mathematical calculations. Otherwise, Fenton et al. mentions, that what has most influence on these models are the prior probabilities, which initially would be subjective even though the model is robust with only seven of the 99 variables being able to lift the probability of guilt over 95%.

Van Koppen & Mackor mention in their article that, their scenario-based analysis of the Simonshaven case is neither purely objective nor purely subjective. It could be argued that their analysis approximates subjectivity more than objectivity. Even though their analysis is not based on purely arbitrary or random choices, it is based on the authors' mental representation of the Simonshaven case in a scenario-based approach, therefore it suffers from similar problems as the argumentation-based approach applied by Prakken. Similarly, to the argumentation-based approach, different analysts might have different reasoning on which scenarios to elaborate and in which way, and where to stop the analysis altogether. What might help these approaches to gain more reproducibility, is to create amore standardized way of creating a model based off their approach. This could be done by creating rules similar to the idioms used by Fenton et al.in their approach, standardized rules by which to create their models, that are less left to the interpretation of the reader.

4.2. Naturalness from a cognitive and legal point of view.

Prakken compared his analysis to that of the court. Naturally, the court explains how it reached it decision, which evidence it used and which it dismissed. His analysis based itself more on how the court could justify it decisions rather than on how it did so. Prakken argues in his paper that his analysis of the case does not stray far from the one described by the court in the court documents. Van Koppen and Mackor differentiate their scenario-based analysis by commenting that the key difference between the scenario- and argumentation-based approach is that the scenario-based approach reasons from the hypothesis to the evidence, and the argumentation-based approach the other way around,. This is an interesting observation, and evokes the question, how do people reason about hypotheses and evidence in the first place? Two main approaches are discussed in van Koppen and Mackors paper, but they do not offer concurring explanations. It could be argued that both the argumentation- and scenario-based approaches each offer one dimension to the complicated process that is human reasoning.

Dahlman mentions in his article that critics argue that Bayesian thinking cannot mirror the complexity of our world. As a counter argument, Dahlman offers that models of Bayesian thinking are not models of the world, they are models of our reasoning about the world. Coincidentally, Fenton et al. make a similar claim in their paper, stating that a BN is a natural representation of a legal case and its evidence. This might be case; however, this does not negate the fact that this way of Bayesian thinking does not come natural to most legal professionals. Take for example the answer to the Mony Hall problem, where three doors are offered, one which has a prize behind it. The contestant can choose one door, followed by the host who opens a door which has no prize. The question then is, is it better to switch or to stay with the door initially chosen? Instinct drives most people to argue that it does not matter, that the chances are equal. However, using Bayesian thinking, we can get to the conclusion that switching is the smarter move. This example implying that Bayesian thinking and BNs are not immediately natural, at least not cognitively. It therefore could be argued that without background knowledge on probabilities and the functionality of Bayes rule, approaches like Dahlman's and Fenton et al.'s might prove quite difficult for legal professionals to learn and adapt, certainly if they have no prior knowledge on mathematics.

4.3. Errors and/or biases in reasoning

Interestingly, each of the four approaches identified slightly different fallacies or biases in the case.

Prakken identified two main fallacies with the court's reasoning. The first one being the fact that the suspect's statement was disregarded, because it was contradicted by the statements of four witnesses. This way of reasoning disregards the fact that, just because no other possible culprit was seen at the crime scene, does not mean there was not one. The second one being the fact that the court fails to consider that the existence of another possible culprit (Perry S.) diminishes the probability of EL lying.

Interestingly, another bias/fallacy can be found looking at the case in the way Prakken does. The court regards the absence of evidence in the argument of EL's testimony as conclusively meaning that there exists none. However, the court does not regards the absence of a gun on the crime since the same way, in this case it concludes the opposite, that, just because no gun was found, doesn't mean none did exist.

Van Koppen & Mackor, identify two main errors within the case, the first one concerning ELs statement (resembling Prakken's concern with the case), but also the courts acceptance of the blood found on ELs shoes as evidence of his guilt; which could be explained by EL's statement of events.

Dahlman identified three completely different fallacies and biases in the case, such as the use of false dichotomies, dependence neglect and miss rate neglect within the case. Which are more fallacies in overall reasoning, than fallacies uniquely attributed to this case. Whilst Fenton et al. did not identify concrete fallacies or gaps in reasoning more than a stronger insight into the probabilities of reasoning with the evidence in this case.

4.4. Granularity/precision

The granularity in the approaches differs from each other. Dahlman and Fenton et al.'s approach seems more precise; however, this seemingly precise answer is based on subjective estimates of probability. These estimates are often just that, estimates and are, in some cases, even though based on rational though, created by the author. Because of this lack of objectivity these approaches are prone to seem precise, but they could be very far from the truth.

The interesting fact that each approach identifies errors in different aspects of the case indicates a lack of precision in all of them. They all focus on a different way of reasoning regarding the case, which is why they this is the case.

5. Conclusion

In the introduction four main areas of difference between the approaches were identified, (1) objectivity, (2) naturalness, (3) capability to identify errors and/or biases in reasoning, and (4) precision and granularity.

With regards to objectivity each approach differs, but they all suffer from subjectivity in one way or another. The approaches based on Bayesian thinking have to base their prior probabilities and likelihood ratios on personal estimates, making chances small that, were another analyst to apply the same method, that it would result in the same or similar outcome. Making

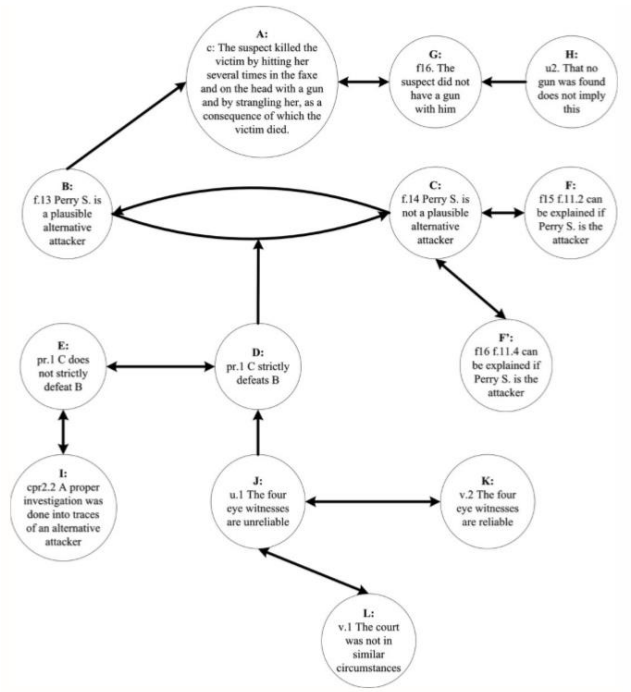
these approaches sensitive to subjectivity. Fenton et al. remedies these concerns with the help of idioms and proves in his paper that the BN build by them is robust to changes, attempting to somewhat remedy lack of objectivity in their approach. The other two approaches, the scenario- and argumentation-based approach suffer from a lack objectivity in similar ways, where reproduced of their models might be similar but rarely the same. The naturality of the scenario- and argumentation-based approaches is significantly bigger than that of the approaches using Bayesian thinking. Bayesian thinking just is not something that comes natural to people with no background in mathematics, such as legal professionals. Interestingly, all approaches identified different types of biases and fallacies in the case, although most of them concerning EL's testimony or the interpretation of evidence on probabilities. Precision is implied to be low since each approach identified different fallacies.

The logical frameworks offered in this paper might help understand human reasoning about evidence and hypotheses. However, is could also, most importantly, aid the creation of new higher functioning reasoning frameworks. Understanding human cognition and its shortcoming help the creation of models which lack these disadvantages, certainly when these do not come natural to people.

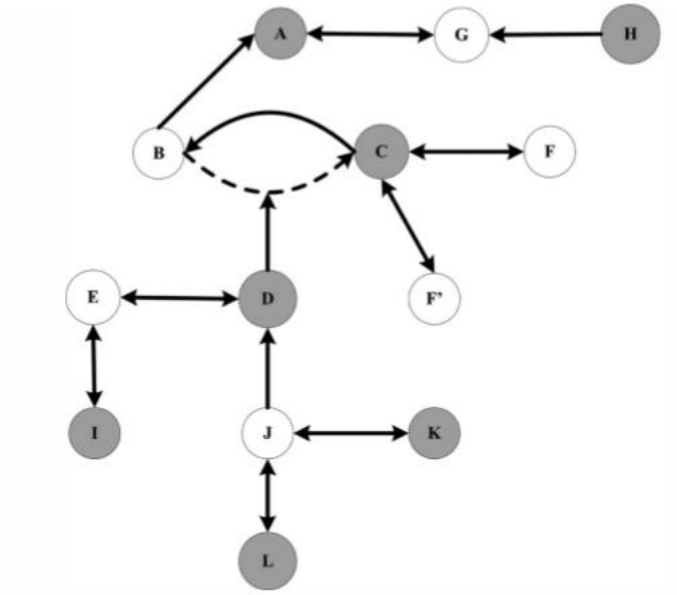
As established, each approach has its shortcomings and benefits, a way to remedy these might to combine some of these approaches. An example of this is the article by Vlek et al. (2014) where the argumentation method and a BN were combined to create a new type of model, aided with the addition of some new idioms. Increased specificity in the creation of the models, aids the reproducibility of them and therefore their objectivity. It could also aid precision, because probabilistic reasoning and the argumentation-based approach look to different methods of reasoning in the case.

Appendix

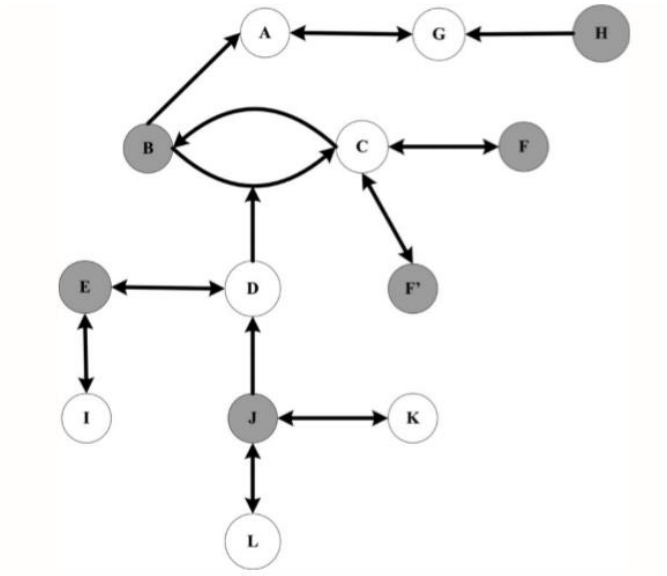
A.1: An extended argumentation framework (Prakken, 2019)



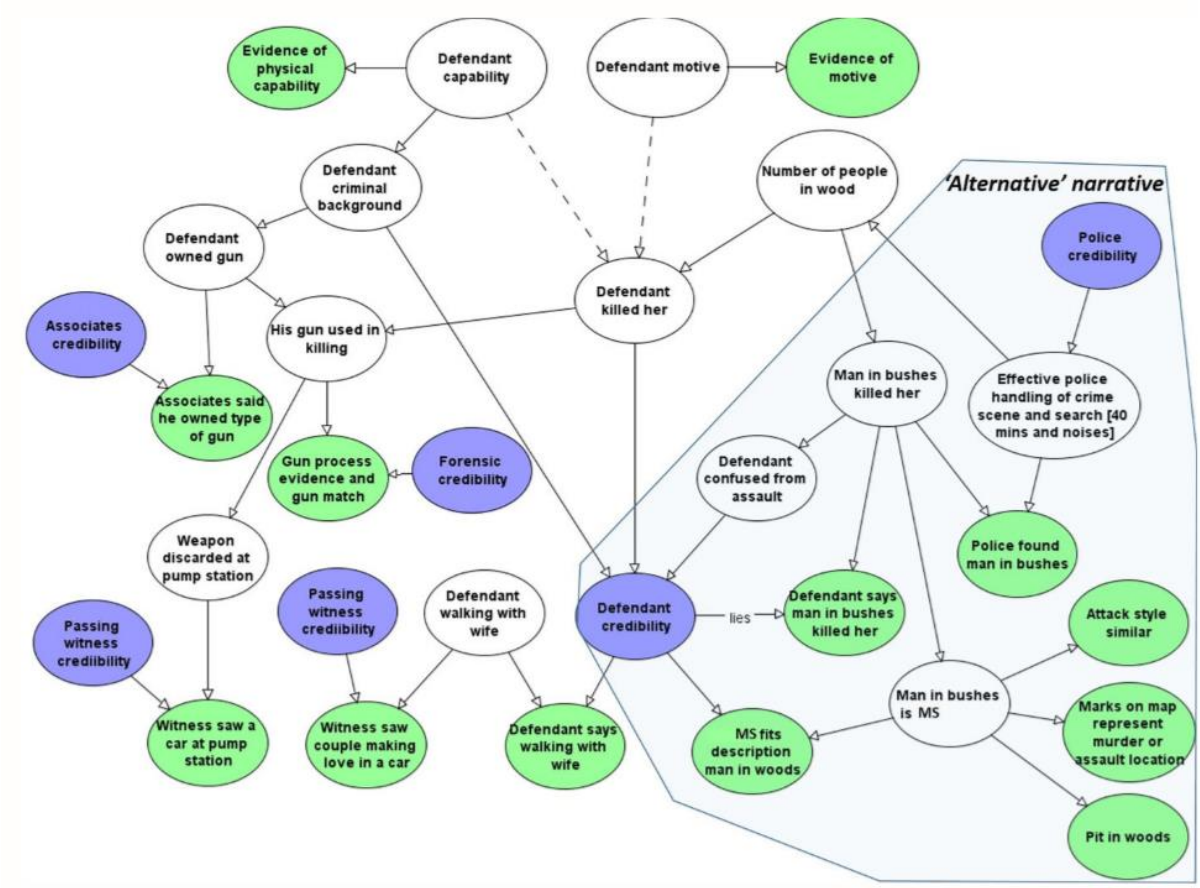
A.2: A labelling of the framework A.1, in which argument *A* is accepted (Prakken, 2019).



A.3: A labelling of the framework A.1, in which argument *A* is not accepted (Prakken, 2019).



B: Fenton et al. Bayesian network model of the Simonshaven case.



6. Literature

Bex, F. J. (2011). Arguments, Stories and Criminal Evidence (2011 ed.). New York, Verenigde Staten: Springer Publishing. https://doi.org/10.1007/978-94-007-0140-3

Cederblom, J., & Paulsen, D. (2011). Critical Reasoning (7th ed.). Boston, USA: Wadsworth.

Dahlman, C. (2019). De-Biasing Legal Fact-Finders With Bayesian Thinking. Topics in Cognitive Science, 1–17. https://doi.org/10.1111/tops.12419

Dawid, P. (1993). The island problem - coherent use of identification evidence (115). University College London. Retrieved from https://www.ucl.ac.uk/statistics/research/pdfs/rr115.pdf

Fenton, N, & Neil, M. (2000). The jury observation fallacy and the use of Bayesian networks to present probabilistic legal arguments. . Mathematics Today - Southend on Sea, 36(6), 180–187. Retrieved from https://www.eecs.qmul.ac.uk/~norman/papers/jury_fallacy.pdf

Fenton, Norman, Neil, M., & Lagnado, D. A. (2012). A General Structure for Legal Arguments About Evidence Using Bayesian Networks. Cognitive Science, 37(1), 61–102. https://doi.org/10.1111/cogs.12004

Fenton, Norman, Neil, M., Yet, B., & Lagnado, D. (2019). Analyzing the Simonshaven Case Using Bayesian Networks. Topics in Cognitive Science, 1–23. https://doi.org/10.1111/tops.12417

Gill, R. D., Groeneboom, P., & de Jong, P. (2018). Elementary Statistics on Trial—The Case of Lucia de Berk. CHANCE, 31(4), 9–15. https://doi.org/10.1080/09332480.2018.1549809

Modgil, S., & Prakken, H. (2013). A general account of argumentation with preferences. Artificial Intelligence, 195, 361–397. https://doi.org/10.1016/j.artint.2012.10.008

Modgil, S., & Prakken, H. (2014). TheASPIC+framework for structured argumentation: a tutorial. Argument & Computation, 5(1), 31–62. https://doi.org/10.1080/19462166.2013.869766

Prakken, H. (2009). Legal Evidence and Proof (New edition). Abingdon, Verenigd Koninkrijk: Routledge.

Prakken, H, Mackor, A. R., & Bex, F. (2019). Models of Rational Proof in Criminal Law. Topics in Cognitive Science, 1–11. Retrieved from http://www.staff.science.uu.nl/~prakk101/papers/Topics_Intro2.pdf

Prakken, Henry. (2019). An Argumentation-Based Analysis of the Simonshaven Case. Topics in Cognitive Science, 24. https://doi.org/10.1111/tops.12418

van Koppen, P. J., & Mackor, A. R. (2019). A Scenario Approach to the Simonshaven Case. Topics in Cognitive Science, 20. https://doi.org/10.1111/tops.12429

Vlek, C. S., Prakken, H., Renooij, S., & Verheij, B. (2014). Building Bayesian networks for legal evidence with narratives: a case study evaluation. Artificial Intelligence and Law, 22(4), 375–421. https://doi.org/10.1007/s10506-014-9161-7