Extracting Legal Arguments from Forensic Bayesian Networks

Sjoerd T. TIMMER [a], John-Jules Ch. MEYER [a], Henry PRAKKEN [a b], Silja RENOOIJ [a], Bart VERHEIJ [c]

[a] *Department of Information and Computing Sciences, Utrecht University*
[b] *Faculty of Law, University of Groningen*
[c] *Institute of Artificial Intelligence, University of Groningen*

Abstract Recent developments in the forensic sciences have confronted the field of legal reasoning with the new challenge of reasoning under uncertainty. Forensic results come with uncertainty and are described in terms of likelihood ratios and random match probabilities. The legal field is unfamiliar with numerical valuations of evidence, which has led to confusion and in some cases to serious miscarriages of justice. The cases of Lucia de B. in the Netherlands and Sally Clark in the UK are infamous examples where probabilistic reasoning has gone wrong with dramatic consequences. One way of structuring probabilistic information is in Bayesian networks(BNs). In this paper we explore a new method to identify legal arguments in forensic BNs. This establishes a formal connection between probabilistic and argumentative reasoning. Developing such a method is ultimately aimed at supporting legal experts in their decision making process.

Keywords. Legal reasoning, Argumentation, Probablistic reasoning, Bayesian networks, ASPIC+, Defeasible reasoning, Evidential reasoning

Introduction

In recent years the rise of forensic sciences has posed the legal domain with the challenge of dealing with numerical, probabilistic evidence. Bayesian Networks (BNs) have become popular as a tool for modelling forensic evidence, and even complete legal cases [7,14,16,15,23,24]. Although Bayesian reasoning may come naturally to forensic experts familiar with mathematics and statistics, non-mathematical experts are prone to reasoning fallacies with statistics [8]. Moreover, the graphical structure of a BN, though intuitive for those familiar with the formalism, is often interpreted incorrectly [4]. Legal experts are often more familiar with argumentative models of proof. The different backgrounds of legal and forensic experts create a communication gap between them, which may lead to miscarriages of justice when undetected. Earlier attempts to explain probabilistic inferences have focussed on visual or textual explanation [13,12,5].

Different kinds of approaches to modelling legal evidence have been identified [9]. The field of *argumentation* offers formalisms to describe the chains of inference in legal arguments and studies the ways in which arguments support and attack each other. Secondly, *narrative* models of legal proof have been developed that study how evidence can be organised in scenarios or stories. Thirdly, *probabilistic* models have been introduced to provide very precise descriptions of uncertainty, probability of findings and (in-)dependencies between variables.

The present research is part of a research project[1] aiming for the integration of the three kinds of approaches [22]. In this paper we explore a method to extract arguments from BNs. This method offers the possibility of integrating BNs and argumentation, in order to ultimately bridge the communication gap between legal and forensic experts. An initial attempt at such a system was introduced in [21] but this had a number of serious shortcomings. Rules were given a strength that was not determined in the context of the given evidence. It is a well known fact from probability theory that the influence between variables heavily depends upon surrounding observations. It also did not allow rules to have more than one premise. In reality premises often corroborate to achieve a certain conclusion. Thirdly, and most importantly, no rules were extracted to capture synergistic effects among variables. As a result we were not able to capture some important aspects of Bayesian reasoning such as *explaining away*. By neglecting this inference a serious aspect of Bayesian reasoning was left out. We strongly improve on this result with a method that finds inference rules in a BN and ranks those rules using a probabilistic measure of strength. Those rules are then combined into arguments and formalised in an argumentation framework.

The rest of this paper is organised as follows. After some preliminaries in Section 1 we present our main contribution in Section 2 which is a method to extract rules and undercutters to those rules from a BN. In Section 3 we illustrate this using an example that shows the connection between the BN phenomenon of explaining away and the argumentation phenomenon of argument attack. Section 4 contains discussions and suggestions for future work.

1. Preliminaries

1.1. Bayesian networks

Bayesian networks (BNs) are a well-known tool in the analysis of complex probabilistic systems [19]. A BN models a probability distribution $\Pr(\boldsymbol{V})$ over a number of discrete random variables $\boldsymbol{V}$. For our purposes we restrict our attention to Boolean variables. BNs exploit conditional independences between variables to efficiently represent the full joint probability distribution. More specifically, a BN captures the (in)dependences between variables in a directed acyclic graph in which every random variable is associated with a node. Every node specifies the conditional probabilities of the outcome of that variable given any configuration of its parents in the graph. Observations ($\boldsymbol{e}$ from here onwards) of values for variables

[1]This work is part of the project "Designing and Understanding Forensic Bayesian Networks with Arguments and Scenarios", supported by the Forensic Science research program financed by the Netherlands Organisation for Scientific Research (NWO).

can be entered into the network and the posterior probability distribution $\Pr(\boldsymbol{V}|\boldsymbol{e})$ of all the other nodes conditioned on this evidence can be computed.

We want to highlight one important aspect of BN reasoning. Through a head-to-head connection (two incoming arcs on an undirected path) an interaction between the parents can occur. Such a situation often models two causes of a shared effect. Observing the effect raises the belief in either of the causes. If then one of the parents is observed this *explains away* the other cause, again decreasing its probability. The example that we will show further on is a minimal case in which this kind of reasoning is present.

1.2. Argumentation

In this paper we choose to use an instantiation of the ASPIC+ framework [17] for structured argumentation, which allows us to deal with defeasible rules that can be argued against with undercutters. It also enables us to reason with preferences. ASPIC+ specifies how arguments attack and defeat each other. It is a special case of a Dung argumentation framework [6]. The following definitions summarize the parts of an argumentation system as described in [17].

Definition 1 (Argumentation System)**.** *An argumentation system is a tuple $AS = (\mathcal{L}, \bar{\ }, \mathcal{R}, n)$ where:*

- *$\mathcal{L}$ is a logical language*
- *$\bar{\ }$ is a contrariness function from $\mathcal{L}$ to $2^{\mathcal{L}}$*
- *$\mathcal{R} = \mathcal{R}_s \cup \mathcal{R}_d$ is a set of strict ($\mathcal{R}_s$) and defeasible ($\mathcal{R}_d$) inference rules of the form $\phi_1, \ldots, \phi_n \rightarrow \phi$ and $\phi_1, \ldots, \phi_n \Rightarrow \phi$ respectively (where ϕ, ϕ_i are meta-variables ranging over wff in $\mathcal{L}$), and $\mathcal{R}_s \cap \mathcal{R}_d = \emptyset$.*
- *n is a naming convention ($n : \mathcal{R}_d \rightarrow \mathcal{L}$)*

An argumentation system in ASPIC+ consists of a logical language ($\mathcal{L}$) that describes the basic elements that can be argued about. A contrariness function maps elements to a set of incompatible elements. As opposed to strict rules $\mathcal{R}_s$, defeasible rules $\mathcal{R}_d$ can have exceptions [20]. A naming function $n(r)$ maps defeasible rules to elements of the language. This means that defeasible rules themselves can be argued about to determine the exceptions to those rules. The contrariness function captures a generalization of negation that does not need to be symmetric.

To reason with an argumentation system, a knowledge base $\mathcal{K} = \mathcal{K}_n \cup \mathcal{K}_p$ is required for which $\mathcal{K}_n \cap \mathcal{K}_p = \emptyset$. The knowledge base differentiates between certain knowledge $\mathcal{K}_n$ and presumed knowledge $\mathcal{K}_p$.

Definition 2 (Arguments)**.** *ASPIC+ defines arguments as one of the following constructs:*

1. *ϕ if $\phi \in \mathcal{K}_n \cup \mathcal{K}_p$.*
2. *$A_1, \ldots, A_n \rightarrow \psi$ if $A_1, \ldots, A_n$ is an argument such that $\mathrm{Conc}(A_1), \ldots, \mathrm{Conc}(A_n) \rightarrow \psi$ is a strict rule in $\mathcal{R}_s$.*
3. *$A_1, \ldots, A_n \Rightarrow \psi$ if $A_1, \ldots, A_n$ is an argument such that $\mathrm{Conc}(A_1), \ldots, \mathrm{Conc}(A_n) \Rightarrow \psi$ is a defeasible rule in $\mathcal{R}_d$.*

Here Conc(A) is the conclusion of an argument A. For the first type of argument its conclusion is simply the element itself. For the latter two types of argument, the conclusion is the consequent of the rule. The intuitive understanding of arguments is that they form a graph of subarguments. Elementary arguments are elements from the knowledge base. Composite arguments are constructed by the application of a rule. The immediate subarguments of a composite argument must together satisfy all antecedents of the rule. In ASPIC+, three modes of attack are possible. Arguments can attack each other by negating a premise or a conclusion from another argument which is called *undermining* and *rebutting* respectively. The third way of attack is to negate the applicability of a rule thereby *undercutting* the conclusions that were derived using that rule. Attack can be resolved into defeat using the weakest argument ordering strategy that relies on an ordering of rules.

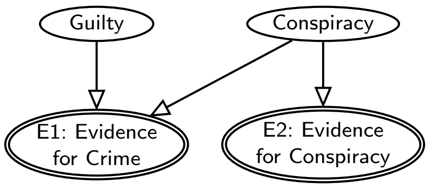
The ASPIC+ framework is abstract in the sense that it specifies (or requires) very little of the actual instantiation of the argumentation system. The instantiator is free to choose a language to represent domain knowledge and the actual implementation of contrariness and rules has no further restrictions than the ones mentioned above. Different instantiations of the framework can exist for specific argumentation tasks. In the next section we will introduce an instantiation of ASPIC+ tailored to reasoning about evidence in BNs.

1.3. Example: explaining away and argument attack

In order to illustrate our approach we use the example BN displayed in Figure 1. The observed evidence nodes have a double outline. The main evidence (on the lower left in Figure 1) is some trace of a crime. Besides the crime there is another explanation for this evidence, namely that there is a conspiracy against the suspect and that he is being framed. Either explanation could bring about the evidence. If one of the two causes is observed the other is explained away by this fact. If we enter evidence for a conspiracy in the BN this explains the evidence and the belief in the guilt of the suspect drops. This is exactly the phenomenon of *explaining away* that we mentioned earlier.

Probabilistic inference and explaining away can be related to an argumentation setting. We can show this in the example. There are two pieces of evidence that can be represented by information in the knowledge base of the argumentative approach. Both evidence variables are probabilistically correlated to the variable for which they are evidence so we expect to find some rules that can be used to derive from that knowledge base an argument for the guilt of the suspect and an argument for a conspiracy. In the BN there is the typical explaining away effect from the conspiracy to the guilt of the suspect. In an argumentation system this can be modelled by undercutting the argument for guilt. All of this is shown in the form of a small argument graph in Figure 2. It shows two straightforward inferences (from evidence to a corresponding conclusion) where the second conclusion undercuts the first by giving an alternative explanation for the evidence.

Guilty	
true	0.1
false	0.9



Conspiracy	
true	0.01
false	0.99

Evidence for Crime

Guilty	true		false	
Conspiracy	true	false	true	false
true	0.9	0.9	0.8	0.001
false	0.1	0.1	0.2	0.999

Evidence for Conspiracy

Conspiracy	true	false
true	0.9	0.1
false	0.1	0.9

Figure 1. An example BN.

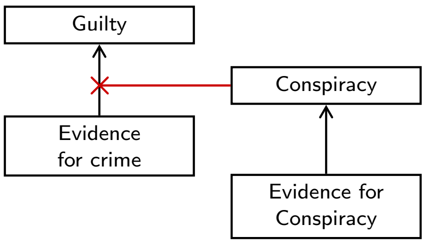


Figure 2. Inferences in argument graphs are displayed as normal arrows and the undercutting attack is visualised by a cross-headed arrow.

2. Argument Extraction

2.1. Rules and rule strengths

The first step in constructing arguments is to define which inference rules apply and how they are ordered. We extract inference rules from a BN by looking at a measure of strength that is based on conditional probabilities. Evidence entered in the BN is maintained and used during the extraction of rules, which guarantees that the strengths of the rules are calculated in the context of the available proof. Due to the dynamic nature of BNs, dependencies and independences (and therefore also the measure of inferential strength that we propose) may change with every new observation. Therefore it is important to assess the strength of rules in the context of the available evidence.

As elements (antecedents and consequent) of our rules we take propositions that are simply assignments $A = a_1$ of values to variables from the BN; Candidate rules for every consequent $A = a$ are identified by enumerating value assignments to non-empty subsets of parents, children and parents of children of A in the BN graph. Parents and children can have a direct effect on the consequent node because there is a statistical correlation between those nodes. Parents of children can explain away the consequent. This means that they often, but not necessarily, have a negative effect on that node. These nodes together (parents, children and parents of children) form the minimal set that suffices to explain statements about the node. For each candidate rule we compute its strength using the so-called normalised likelihood [3]. The strength of a candidate rule $p_1, \ldots, p_n \Rightarrow c$ is evaluated in a particular context $\boldsymbol{\epsilon}$, which is a set of assignments.

Definition 3 (strength)**.** *A rule is defined to have strength:*

$$\text{strength}\,(p_1, \ldots, p_n \Rightarrow c \,|\, \boldsymbol{\epsilon}) = \frac{\Pr\,(\ c \mid p_1, \ldots, p_n \wedge \boldsymbol{\epsilon}\ )}{\Pr\,(\ c \mid \qquad\qquad\qquad \boldsymbol{\epsilon}\ )}$$

in the evidential context $\boldsymbol{\epsilon} = \boldsymbol{e} \setminus \{p_1, \ldots, p_n, c\}$*.*

When the strength of a rule is tested it may be that an antecedent was already set as an observed variable. As a result, this premise is present in the conditions for both the numerator and the denominator. This means that we can no longer measure the actual influence of this premise on c. The same applies when c is among the observed values. To resolve this issue, we temporarily remove any evidence in the BN for antecedent variables or the consequent variable of the rule under consideration. We do this by removing those assignments from the context. All actual evidence in the BN is thus considered during the calculation of the strengths, excluding assignments to the variables that occur in the rule itself. Various measures of probabilistic strength exist, but typically result in the same ordering of rules, despite numerical differences [3].

2.2. Exceptions to rules

Since we extract defeasible rules, we wish to know under which circumstances there are exceptions to those rules. In a BN setting, inferences can often be undone or weakened by observing further evidence, particularly in the case of explaining away. Therefore, we identify undercutting variable assignments by checking if the measure of strength drops below one given a potentially undercutting variable assignment.

Definition 4 (undercutter)**.** *An undercutter to a rule* $p_1, \ldots, p_n \Rightarrow c$ *in the context* $\boldsymbol{\epsilon}$ *is a variable assignment* u *such that* $\text{strength}(p_1, \ldots, p_n \Rightarrow c|\boldsymbol{\epsilon}, u) \leq 1$

Note that we abuse terminology when we use the term *undercutter* for a variable assignment that invalidates a rule rather than an argument that uses such an assignment to attack another argument, which is also called undercutting.

2.3. ASPIC+ instantiation

We now construct an instantiation of ASPIC+ that uses knowledge represented as a BN.

Definition 5 (Rule extraction)**.** *Given a BN with variables* $\boldsymbol{V}$ *and evidence* $\boldsymbol{e}$*, let*

- c *be an assignment to some variable* $C \in \boldsymbol{V}$ *in the BN*
- $\boldsymbol{p}$ *be a set of assignments to parents, children and parents of children of* C*.*

$\mathcal{R}_d$ *consists of all rules of the following form:*

$$r = (\boldsymbol{p} \Rightarrow c) \quad \textit{iff} \;\; strength(r|\boldsymbol{\epsilon}) > 1$$

Definition 6 (BN Argumentation System)**.** *Suppose a BN is given with nodes* $\boldsymbol{V}$*. Let* $vals(V_i)$ *denote the values that variable* $V_i \in \boldsymbol{V}$ *can take on and let* $\boldsymbol{e}$ *denote the evidential context defined by the available evidence. We define the following instantiation of ASPIC+:*

$$\mathcal{L} \;=\; \left(\bigcup_{V_i \in \boldsymbol{V}} \; \bigcup_{v_{ij} \in vals(V_i)} \{V_i = v_{ij}\} \right) \cup \{n(r) | r \in \mathcal{R}_d\}$$

$\overline{x}$ *for all* $x \in \mathcal{L}$ *such that:*

$$\overline{V_i = v_{ij}} \;= \{V_i = v_{ik} \,|\, j \neq k\}$$
$$\overline{n(r)} \;= \{u | u \in \mathcal{L} \setminus \mathcal{R}_d, u \textit{ undercuts } r \textit{ as in Definition 4}\}$$

$\mathcal{R}_s \;= \emptyset$

$\mathcal{R}_d$ *as in Definition 5*

$\mathcal{K}_p \;= \emptyset$

$\mathcal{K}_n \;= \{V_i = v_{ij} | (V = v_{ij}) \in \boldsymbol{e}\}$

$\leq$ *s.t.* $r_1 \leq r_2$ *iff* $\text{strength}(r_1) \leq \text{strength}(r_2)$

Informally, the language consists of variable assignments together with rule names. The contraries of a variable assignment are other assignments to the same variable and the contraries to rules are the undercutters as described above. Furthermore we have modelled observed variable instantiations as necessary knowledge. The preference ordering on arguments can very conveniently be defined in terms of preordering $\leq$ on rules $\mathcal{R}_d$ using the weakest-link[17] principle. This ordering prefers argument A over B iff at least one rule in B is weaker than all rules in A and at least one premise in B is weaker than all premises in A.

There is one caveat with the method as described so far. In a BN two parents (say V_j and V_k) of a node V_i, are independent of each other unless a descendant of V_i is observed. If the latter is the case the composite effect of the inference from V_j to V_i and V_i to V_k does not follow logically from the two separate inferences. We have already discussed the possibility of explaining away in such circumstances. To prevent reasoning errors in such cases we have to exclude any chain of inference that passes through a head-to-head node. Even if a descendant is observed, we will find the correct influence because there is a rule possible directly from V_j to V_k. Excluding these arguments that do use a head-to-head connection can be done in a post-processing step or on the fly by labelling all intermediate conclusions that have already made one inference along the direction of an edge in a way similar to Pearl's CE-Logic [18]. In the latter case we know that if we never apply a rule against the direction of an edge to a premise that was derived with a rule along the direction of an edge we exclude exactly these fallacious reasoning steps.

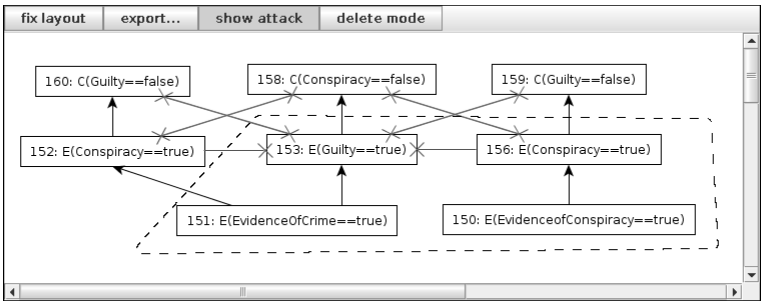


Figure 3. Part of the argument graph resulting from the network in Figure 1. Sub arguments are linked with regular arrows whereas defeat is displayed by cross-headed arrows. The $C()$ and $E()$ notations can be ignored, these are a detail of the CE-logic that we applied to filter rules that pass through head-to-head connections.

3. Application of the method: explaining away and argument attack

Applying the method as described above to our example, we find among others the arguments as presented in Figure 3. Recall that we had two pieces of evidence for a crime and a conspiracy. From these observations, two basic arguments from knowledge can be constructed (150 and 151 in Figure 3). Our rule extraction method yields a number of rules which include the rule "Evidence for Crime ⇒ Guilty" with a strength of 5.79 and "Evidence for Crime ⇒ Conspiracy" with a slightly lower strength of 5.37. The lower strength is explained by the slightly lower probability of finding the evidence given that there was a conspiracy (0.8) than that the suspect is actually guilty (0.9). These rules can be applied to the premises that we just established which we can see as arguments 153 and 156 in the figure. This means that we do indeed find the arguments that we anticipated in Figure 2. To be precise, the four arguments from that figure correspond directly to the four arguments in the dashed trapezoid in Figure 3.

We can also see that argument 156 attacks argument 153. This is an undercutting attack. In ASPIC+ an undercutting attack is targeted at the applying argument rather than the inference rule, so the attack from Conspiracy to Guilty corresponds to the undercutter of the rule that we anticipated. This means that not just the arguments but also the attack relation is exactly as we would expect. The conspiracy explains away the crime and in the argumentation setting this is captured by the undercutting attack between the two corresponding arguments.

In the figure we can observe that these are not the only arguments that can be constructed. We find, for instance, that evidence for a crime also supports a conspiracy (argument 152). This is logical if we consider that the evidence of a crime can indeed be a sign of a conspiracy. Furthermore, we see that arguments for a conspiracy can support arguments against guilt (160 and 159) and arguments for guilt support an argument against a conspiracy (158). This can be explained by the fact that these conclusions are probabilistically strongly correlated. If there was a crime, then there probably was not a conspiracy and vice versa. The arguments that have opposite conclusions also show rebutting attack in both directions. This corresponds exactly to the intuition of rebutting arguments.

4. Conclusions

4.1. Related research

The current research is part of a project in which three approaches to legal reasoning are combined [22]. In previous work Bex has investigated how argumentation and narrative models can be integrated [2,1]. Currently research is also targeted at integrating narrative and Bayesian models [24,23] and in this paper we introduce a method that combines BNs with argumentation.

Earlier attempts to explain probabilistic inferences have focussed on visual or textual explanation. See for instance the work of Lacave and Díez [13,12], Koiter [11] and Druzdzel [5]. All of these methods try to make Bayesian networks easier to interpret by visualising correlations or by verbally presenting the relations between variables. However, none of these methods provide (structured) arguments. In our method we obtain formally well-defined arguments from the Bayesian network. Vreeswijk [25] has proposed a similar but much simpler method to construct rules from BNs to form arguments. His approach, however, only respects the independence properties of a BN under a number of limiting constraints on the design of the network that make explaining away impossible.

Another approach to formalise BN reasoning was taken by Keppens [10] who extracts Argumentation Diagrams from BNs. An Argumentation Diagram is a graphical structure that informally represents support between statements, but does not allow one to identify possible counter-arguments. In Argumentation Diagrams, therefore, only one side of the story is highlighted. In a persuasive setting such as legal reasoning, the argumentation approach can be more expressive because it allows for reasoning about the other parties' arguments as well. Attacking the opponent's counterargument may reinstate your own argument which is a well-known concept in argumentation.

4.2. Discussion

A problem that arises with any automated approach is the computational complexity of combining inference rules. There are numerous ways in which rules can be combined. Doing probabilistic inference in BNs is already computationally hard. However, we find that combining rules in all possible ways is in practice far more time consuming than calculating the probabilities required to identify the rules. To make the process more feasible on large scale networks is subject of future research.

We have formally defined how to extract ASPIC+ arguments from BNs. By doing so we establish a formal connection between concepts from Bayesian reasoning and argumentative reasoning. In particular we have shown the connection between explaining away and argument attack. By extracting rules and undercutters from BNs we were able to construct arguments that support and attack each other. With an example we showed that the characteristic features of probabilistic inference and explaining away are captured by argumentative inference and undercutters. This methodology enables further investigation of how argumentation can deal with probabilities which also opens possibilities for future investigation.

References

[1] F. J. Bex. Towards an integrated theory of causal scenarios and evidential arguments. *Proceedings of COMMA 2014*, 2014. to appear.

[2] F. J. Bex, P. J. van Koppen, H. Prakken, and B. Verheij. A hybrid formal theory of arguments, stories and criminal evidence. *Artif. Intell. Law*, 18(2):123–152, 2010.

[3] V. Crupi, K. Tentori, and M. Gonzales. On Bayesian measures of evidential support: Theoretical and empirical issues. *Philosophy of Science*, 74(2):229–252, april 2007.

[4] A. P. Dawid. Beware of the DAG! In I. Guyon, D. Janzing, and B. Schlkopf, editors, *NIPS Causality: Objectives and Assessment*, volume 6 of *JMLR Proceedings*, pages 59–86, 2010.

[5] M. J. Druzdzel. Qualitative verbal explanations in Bayesian belief networks. *Artificial Intelligence and Simulation of Behaviour Quarterly*, 94:43–54, 1996.

[6] P. M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, 77:321–357, 2005.

[7] N. E. Fenton, M. Neil, and D. A. Lagnado. A general structure for legal arguments about evidence using Bayesian Networks. *Cognitive Science*, 37(1):61–102, 2013.

[8] D. Kahneman. *Thinking, fast and slow.* Farrar, Straus and Giroux, 2011.

[9] H. Kaptein, H. Prakken, and B. Verheij, editors. *Legal Evidence and Proof: Statistics, Stories, Logic.* Applied Legal Philosophy. Ashgate Publishing, 2009.

[10] J. Keppens. Argument diagram extraction from evidential Bayesian Networks. *Artificial Intelligence and Law*, 20(2):109–143, 2012.

[11] J. R. Koiter. Visualizing inference in Bayesian Networks. Master's thesis, Delft University of Technology, 2006.

[12] C. Lacave and F. J. Díez. A review of explanation methods for Bayesian Networks. *Knowledge Engineering Review*, 17(2):107–127, 2002.

[13] C. Lacave, M. Luque, and F. J. Díez. Explanation of Bayesian Networks and influence diagrams in elvira. *Systems, Man, and Cybernetics, Part B*, 37(4):952–965, 2007.

[14] D. A. Lagnado, N. E. Fenton, and M. Neil. Legal idioms: a framework for evidential reasoning. *Argument and Computation*, pages 1–18, 2012.

[15] K. B. Laskey and S. M. Mahoney. Network fragments: Representing knowledge for constructing probabilistic models. In *Uncertainty in Artificial Intelligence*, pages 334–341. Morgan Kaufmann, 1997.

[16] T. S. Levitt and K. B. Laskey. Computational inference for evidential reasoning in support of judicial proof. *Cardozo Law Review*, 22:1691–1731, 2000.

[17] S. Modgil and H. Prakken. A general account of argumentation with preferences. *Artificial Intelligence*, 195:361–397, 2013.

[18] J. Pearl. Embracing causality in default reasoning. *Artificial Intelligence*, 35:259–271, 1988.

[19] J. Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann, 1988.

[20] J. L. Pollock. Defeasible reasoning. *Cognitive Science*, 11:481–518, 1987.

[21] S. T. Timmer, J.-J. C. Meyer, H. Prakken, S. Renooij, and B. Verheij. Inference and attack in Bayesian networks. In *Benelux Conference on Artificial Intelligence 2013*, pages 199–206, 2013.

[22] B. Verheij. To catch a thief with and without numbers: arguments, scenarios and probabilities in evidential reasoning. *Law Probability & Risk*, 13(3-4):307–325, 2014.

[23] C. Vlek, H. Prakken, S. Renooij, and B. Verheij. Modeling crime scenarios in a Bayesian network. In *ICAIL*, 2013.

[24] C. S. Vlek, H. Prakken, S. Renooij, and B. Verheij. Building Bayesian networks for legal evidence with narratives: a case study evaluation. *Artificial Intelligence and Law*, 2014.

[25] G. A. Vreeswijk. Argumentation in Bayesian belief networks. In I. Rahwan, P. Moraïtis, and C. Reed, editors, *Argumentation in Multi-Agent Systems*, volume 3366 of *Lecture Notes in Computer Science*, pages 111–129. Springer Berlin / Heidelberg, 2005.