On direct and indirect probabilistic reasoning in legal proof[1]

Henry Prakken
Department of Information and Computing Sciences, Utrecht University &
Faculty of Law, The University of Groningen, The Netherlands

14 March 2014

Abstract

In the academic literature three approaches to rational legal proof are investigated, broadly speaking based, respectively on Bayesian statistics, on scenario construction and on argumentation. In this paper these approaches are discussed in light of a distinction between direct and indirect probabilistic reasoning. Direct probabilistic reasoning directly reasons from evidence to hypotheses, while indirect probabilistic reasoning reasons from hypotheses to evidence (and then back to the hypotheses). While statistical and story-based approaches usually model indirect probabilistic reasoning, argumentation-based approaches usually model direct probabilistic reasoning. It has been suggested that all legal probabilistic reasoning should be indirect, but in this paper it is argued that direct probabilistic reasoning has a rational basis and is, moreover, sometimes easier to perform for judges than indirect probabilistic reasoning. Moreover, direct probabilistic reasoning can be analysed in terms of standard probability theory, resulting in an alternative, non-Bayesian use of the terms "prior" and "posterior" probability and without the need to estimate unconditional probabilities of the hypotheses.

1 Introduction

Models of rational legal proof are usually of three kinds: statistical, story-based, and argument-based.[2] All three approaches acknowledge that evidence cannot provide watertight support for a factual claim but always leaves room for doubt and uncertainty. Statistical approaches (cf. e.g. Schum 1994; Fenton & Neil 2011; Fenton et al. 2013) account for this by applying the 'standard' theories for reasoning under uncertainty: statistics and probability theory. Often a Bayesian approach is taken, according to which domain experts provide conditional probabilities on hypotheses given the available evidence, while the fact-finder estimates the "prior" probabilities on the hypotheses. Bayesian probability theory then allows to compute the "posterior" probability of the hypotheses given the evidence.

Statistical approaches are by no means uncontroversial (a useful critical overview of the debate in the law is Lempert 1986; see also Thagard 2004 and more generally Pollock 2009). One objection is that in legal cases the required numbers are usually not available, either because there are no reliable statistics, or because experts or judges are unable or reluctant to provide reliable estimates of probabilities. Another objection is that statistics and probability theory impose a standard of rationality that cannot be attained in practice, so that their

---

[2] Parts of this introduction are adapted from the introduction of Kaptein et al. (2009).

application would lead to more instead of fewer errors. To overcome these and other claimed limitations of statistical models, other models have been proposed.

The story-based approach was initially proposed by the psychologists Bennett and Feldman (1981). Their main claim was not normative but empirical, being that the way lawyers make factual judgements is not by statistical reasoning but by constructing and comparing stories about what might have happened. Wagenaar et al. (1993) go a step further, arguing that this is in fact the only way for fact finders to reason about the facts of the case, given the cognitive limitations of humans. Their research then takes a normative twist, studying how story-based factual judgement may be more rational. The story that best explains the evidence must, if it does so to a sufficient degree, be adopted as true. Room for doubt remains since an as yet unknown story might be the true one or since new evidence might make another of the known stories the best one. An important point in this approach is that stories must be 'anchored' in factual generalisations, which may be less than certain, so that their applicability to a case must be critically examined. Recently, Van Koppen (2011) has further developed the 'anchored-narratives' approach and argued that it is a better model of legal proof than Bayesian approaches. The story-based approach has also been advocated by Allen Allen (2008) and has been embraced by some AI researchers, based on the idea that story-based reasoning is like abductive reasoning and inference to the best explanation, for which elaborate AI models exist (Josephson 2001; Poole 2001; Thagard 2005).

A third approach takes not stories but arguments as the main concept. This approach ultimately goes back to Wigmore's charting method of legal evidence, with which alternative arguments from evidence to hypotheses can be graphically displayed and sources of doubt in these arguments can be revealed (e.g. Wigmore 1931). Generalisations are important here as well, since they are the 'glue' that connects the evidence with the hypotheses, and since their general plausibility and possible exceptions are important sources of doubt. Wigmore's charting method has been extended and refined by the 'New Evidence Scholars' (e.g. Anderson et al. 2005). In AI and Law the argument-based approach has been developed in terms of logical frameworks for argumentation, which model reasoning as the construction and comparison of arguments and counterarguments (e.g. Verheij 2003; Bex et al. 2003; Prakken 2004), where the conclusions of the arguments that survive the competition with their counterarguments must be adopted as true. In this approach room for doubt remains since additional evidence might give rise to new arguments that defeat the currently undefeated arguments.

Finally, some have studied combinations of approaches. For example, Floris Bex (Bex 2011; Bex et al. 2010) presents a hybrid narrative-argumentative theory of legal evidential reasoning, in which stories about what may have happened in a case can be constructed by applying a logical model of abductive explanation, while the quality of these stories can be discussed with argumentation schemes. Recently, Vlek et al. (2013a,2013b) has proposed a method for constructing Bayesian networks for forensic applications in terms of stories or scenarios, while Timmer et al. (2013) have studied how arguments and counterarguments can be extracted from a forensic Bayesian network.

Although the three approaches are different in important respects, they also share an important feature, namely, the explicit comparison of alternatives. Bayesian approaches compare how well alternative hypotheses explain the evidence, story-based approaches do the same for stories or scenario's about what may have happened in a case, and argumentation-based approaches compare arguments and counterarguments concerning a

claim. In all these cases the deeper reason for the need for comparisons is that evidence almost always leaves room for doubt: sources of evidence (such as witnesses) are fallible, and general world knowledge is uncertain and leaves room for exceptions.

The approaches can be compared and contrasted in a number of ways. For example, while Bayesian approaches use numbers, the argumentative and story-based approaches are (usually) qualitative. However, in this paper I will contrast Bayesian and story-based approaches on the one hand with argumentation-based approaches on the other on the following issue. Proponents of both Bayesian and story-based approaches have suggested that evidential reasoning in legal contexts should be from the hypotheses to the evidence (e.g. Evett 2009; Van Koppen 2011; Fenton & Neil 2011). However, the argumentation-based approaches as sketched above usually apply reasoning from the evidence to the hypotheses.

The main aim of this paper is to argue that
(1) Reasoning from evidence to hypotheses can be done in agreement with the axioms of probability theory, so that adopting probability theory as a foundation for evidential reasoning does not necessitate the use of Bayesian methods;
(2) there are cases in which reasoning from evidence to hypotheses in legal proof is just as rational as reasoning from hypotheses to evidence, but practically more feasible for lawyers.

The paper will be concluded by a brief discussion of Thagard's (2005) so-called dual pathway model of reasoning with testimony.

2 Direct and indirect probabilistic reasoning

I will use the term 'direct probabilistic reasoning' for directly reasoning from evidence to hypotheses and 'indirect probabilistic reasoning' for reasoning from hypotheses to evidence (and then back to the hypotheses).[3] Both Bayesian and scenario-based approaches usually advocate indirect probabilistic reasoning while argumentative approaches usually model direct probabilistic reasoning (although this is not enforced by the formal techniques used in these approaches).

Let me illustrate the difference with a very simple example involving a witness who states that something happened, for example, that the suspect was at the crime scene at the time of the murder. Let E (for 'evidence') stand for the fact that the witness made this testimony and H (for 'hypothesis') for the testified proposition.

A direct argument from evidence to hypothesis is then

The witness said that the suspect was at the crime scene at the time of the murder
<u>Witnesses usually speak the truth</u>
Therefore, presumably, the suspect was at the crime scene at the time of the murder.

In more stylised notation this becomes

---

[3] These terms are adapted from Derksen (2010, 193), who writes (the Dutch counterpart of) "argumentation" instead of "reasoning". I have made the change to "reasoning" to avoid the mistaken impression that these forms of reasoning only occur in argumentation-based approaches.

E
If E then usually H
Therefore, presumably, H

In Artificial Intelligence research on so-called commonsense reasoning this kind of reasoning is called 'default reasoning' and has been modelled in logics for so-called defeasible reasoning, of which argumentation logics are a subspecies. The reasoning here is very similar to classical-logic's modus ponens, except that the conditional is not material but defeasible, that is, it can have exceptions. In argumentation the possibility of exceptions is accounted for in lists of so-called critical questions to argument schemes. Let us specialise the above scheme for reasoning with generalisations for witness testimonies as follows.[4]

W testifies that H
W is in the position to know about H
Therefore (presumably), H

This scheme has the following critical questions:

CQ1: Is W sincere? ($\approx$ Schum's *veracity*)
CQ2: Did W's memory function properly? ($\approx$ Schum's *objectivity*)
CQ3: Did W's senses function properly? ($\approx$ Schum's *observational sensitivity*)

In a formal argumentation-based account of legal evidential reasoning (Bex et al. 2003; Verheij 2003; Prakken 2004) negative answers to each of these three questions generate undercutting defeaters to arguments that use the scheme. If these undercutting defeaters are not themselves defeated, they prevent the drawing of inferences from the testimony. This process can be formalised in argumentation logics or in any other suitable nonmonotonic logic. The precise way in which this can be done is for present purposes irrelevant, except for one important feature: if no information is available at all on whether there is an exception, then the argument based on the scheme has no defeaters and its conclusion is warranted. Thus argumentation logics (and nonmonotonic logics in general) make it possible to draw inferences from incomplete (in contrast to uncertain) information.

A typical Bayesian analysis of the example is as follows[5]. Given the testimony E, a Bayesian reasoner is interested in P(H|E), that is, the conditional probability of H given E. According to Bayes' rule, this probability can be calculated as follows:

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

where P(E) can be further specified as

$$P(H) * P(E \mid H) + P(\neg H) * P(E \mid \neg H)^{6}$$

---

[4] This scheme plus its list of critical questions combine ideas from Walton (1996, 2008) with elements of Schum's work on witness credibility (e.g. Schum 1994, Kadane & Schum 1996).

[5] The state of the art in Bayesian thinking is so-called Bayesian networks (BN), which combine probability assignments with graphical representations of probabilistic dependencies. However, for present purposes it suffices to represent the example with more traditional Bayesian methods.

[6] The $\neg$ symbol means 'it is not the case that'.

Here P(H) is the unconditional probability of H. Note that P(H|E) positively depends on P(H), so if the unconditional probability of H is low, the conditional probability that the testified fact is true can also be low even if P(E|H) is high, that is, even if the witness is highly reliable (see also Schum 1994, pp. 203-4).[7]

By itself Bayes rule is just a mathematical equation, stating a constraint on joint probability distributions. However, it is often given a dynamic reading, in which P(H) is called the *prior probability* of hypothesis H and P(H|E) its *posterior probability* given the evidence E. The idea is that when evidence E becomes available, a Bayesian reasoner *updates* its prior belief P(H) in the hypothesis to its posterior belief P(H|E). More precisely, if evidence E becomes available, then first P(H|E) is computed and then, if new evidence E' arrives, P(H) is updated by setting it equal to P(H|E) and Bayes' rule is applied again to calculate P(H|E'). In this method, often called Bayesian updating, P(H) is called the "prior" and P(H|E) the "posterior" probability of H, where these temporal indicators refer to, respectively, the time before and after E becomes available. Hacking (2001, Chapter 21) makes this dynamic use of Bayes' rule explicit by giving the probability operator P explicit time stamps:

$$P_{t2}(H) = P_{t1}(H|E1) \qquad = \quad \frac{P_{t1}(E1|H) * P_{t1}(H)}{P_{t1}(E1)}$$

$$P_{t3}(H) = P_{t2}(H|E2) \qquad = \quad \frac{P_{t2}(E2|H) * P_{t2}(H)}{P_{t2}(E2)}$$

And so on …

Applying Bayes' rule to compute P(H|E) once E is known would at first sight seem to be a form of direct evidential reasoning, since we conclude something about the probability of H given E. However, upon closer inspection it turns out that to compute P(H|E), a Bayesian reasoner must provide the conditional probabilities of the evidence E given the alternative hypotheses H and ¬H (the reasoner must also estimate the unconditional probability of H, but does not make the reasoning direct).

Some claim that rational thinking about legal proof requires that judges reason in a Bayesian way (e.g. Sjerps & Berger 2011; Fenton & Neil 2011). With the above definition of indirect probabilistic reasoning this implies that they advocate that judges should always decide about matters of proof by way of indirect probabilistic reasoning (cf. Evett 2009). Van Koppen (2011), in his version of the scenario-based approach, makes the same recommendation, even though at other points his approach radically differs from the Bayesian ones: he does not use numbers since he claims that in legal cases it is impossible to come up with reliable numerical probabilities, and he does want to not take the prior probability of hypotheses into account since he thinks that they cannot be reliably estimated at all. However, for present purposes the similarities between his and a Bayesian approach are more relevant.

---

[7] Well-known examples of this phenomenon involve the so-called base rate fallacy (cf. Tversky & Kahneman 1974). Imagine a town with 100 taxis of which 85 blue and 15 green, and suppose that an 80% reliable witness (i.e. P(E|H) = 0.8) identifies a given taxi as green. Let E = "the witness identifies the taxi as green" and H = "the taxi is green". Since P(H) = 0.15, we have that P(H|E) = 0.41 even though the witness is 80% reliable.

3 Direct probabilistic reasoning versus probability theory

Is direct probabilistic reasoning at odds with probability theory? This question becomes relevant for those who claim that legal proof should be Bayesian and that Bayesian probabilistic reasoning is always indirect. As we described direct probabilistic reasoning above, it is qualitative, that is, it does not use numbers. This is in line with the general motivation behind nonmonotonic logics as being formalisms for "quick and dirty reasoning" in circumstances where reliable statistics are not available and reliable subjective estimates of probabilities cannot be given. Thus qualitative direct probabilistic reasoning can at best be an approximation of standard probabilistic reasoning. Now is this good or bad? It would be bad if it could be shown that this form of reasoning is prone to making probabilistic reasoning fallacies. This is a complex issue (see e.g. Pearl 1988) but I want to argue that at least arguments as above can very well be given a standard probabilistic interpretation.

Let us probabilistically interpret the above argument as follows:

$P(E) = 1$
$P(H|E)$ is x (where x is high)
Therefore (deductively!) $P(H)$ is x (where x is high)

This argument, which in fact is a qualitative version of the so-called statistical syllogism (cf. Pollock 1995) is deductively valid given the axioms of probability theory. It is not Bayesian but it is just as consistent with the laws of probability theory as the method of Bayesian updating. Interestingly, in this reconstruction it is $P(H|E)$ that can be called the prior probability of H and $P(H)$ that can be called the posterior probability of H (relative to E): as long as we do not know anything about $P(E)$, all we know about H is $P(H|E)$, so $P(H)$ could be anything. But once we know that $P(E) = 1$, we can deduce that $P(H) = P(H|E)$, so $P(H)$ is the probability of H after receiving the evidence E. This also shows that the terms "prior" and "posterior probability" are not inherent to probability theory but belong to specific *uses* of probability theory.

Now very interestingly, in this reconstruction the problem of the priors disappears, since $P(H|E)$ is often much easier to estimate than $P(H)$. Below I will say more about how this is possible, but first I want to delve deeper into the probabilistic analysis of default reasoning.

A common situation in default reasoning is that of conflicting defaults. Consider the following well-known example from AI: Quakers are usually pacifist while Republicans are usually not pacifist. Now president Nixon was both a quaker and a republican. What can we conclude about whether Nixon was a pacifist or not without knowing anything else about him? Two conflicting arguments are possible:

Argument A:
Quakers are usually pacifist
Nixon was a quaker
Therefore, Nixon was a pacifist.

Argument B:
Republicans are usually not pacifist
Nixon was a republican
Therefore, Nixon was not a pacifist.

In Modgil & Prakken (2013, pp. 381-2) such cases are analysed as follows. It is consistently possible to say at the same time that Quakers are usually pacifists while Republicans are usually not pacifists and that president Nixon was both a quaker and a republican. From this it follows that the arguments cannot be deductively valid, since if they were, the fact that their conclusions contradict each other would imply that not all premises can be true. So on this account the arguments are defeasible, applying a defeasible form of modus ponens.

However, in the probabilistic version of this example according to our above method we obtain

Argument A (statistical version)
$P(Q) = 1$
$P(P|Q)$ is x (where x is high)
Therefore (deductively!) $P(P)$ is x (where x is high)

Argument B (statistical version)
$P(R) = 1$
$P(\neg P|R)$ is y (where y is high)
Therefore (deductively!) $P(\neg P)$ is y (where y is high)

Now assuming that a high probability is always higher than 0.5, then given the axioms of probability theory the two conclusions contradict each other so, since the arguments are deductively valid, not all their premises can be true. Does this mean that the analysis of Modgil & Prakken (2013) is wrong? The answer is no, since in translating the natural-language example into its probabilistic version we in fact left one step implicit. We translated "Quakers are usually pacifists" as "$P(P|Q)$ is high" but the conditional probability is much more specific than the natural-language generalization, since $P(P|Q)$ is the probability that *Nixon* was a pacifist given that *Nixon* was a quaker, while the generalization speaks of quakers *in general*. Inferring "the probability that *Nixon* was a pacifist given that *Nixon* was a quaker is high" from "Quakers are usually pacifists" is a defeasible inference. Qualitative modellings in nonmonotonic logics compress these two steps into one, which therefore has to be defeasible. By contrast, the above probabilistic modelling left out the defeasible inference and instead used it as a heuristic for finding the conditional probabilities in the probabilistic modelling of the problem.[8] It thus made the argument deductive but at the cost of misrepresenting the second premise: this should instead have been that $P(P|Q)$ is *usually* high, but this cannot be expressed in standard probability theory. The refined version of argument A is as follows:

Argument A (refined statistical version)
$P(Q) = 1$
$P(P|Q)$ is usually high
Therefore (presumably) $P(P|Q)$ is high
Therefore (deductively) $P(P)$ is high

Likewise for argument B.

---

[8] Hacking (2001, p. 138) calls this heuristic the *frequency principle*.

The same analysis applies to the witness testimony example. Let E again stand for "The witness testified that the suspect was at the crime scene at the time of the murder" and H for "The suspect was at the crime scene at the time of the murder", and let E' stand for "The witness is not sincere". Then we have two probabilistic arguments:

Argument C
P(E) = 1
P(H|E) is high
Therefore (deductively) P(H) is high

Argument D
P(E) = 1
P(E') = 1
P(H|E & E') is not high
Therefore (deductively) P(H) is not high

Again the joint premises of these two arguments are inconsistent, since their conclusions contradict each other and the arguments are deductively valid. But again we can turn this into two defeasible arguments by replacing "P(H|E) is high" with "P(H|E) is usually high" and "P(H|E & E') is not high" with "P(H|E & E') is usually not high":

Argument C refined
P(E) = 1
P(H|E) is usually high
Therefore (presumably) P(H|E) is high
Therefore (deductively) P(H) is high

Argument D refined
P(E) = 1
P(E') = 1
P(H|E & E') is usually not high
Therefore (presumably) P(H|E & E') is not high
Therefore (deductively) P(H) is not high

On this analysis we can still consistently say at the same time that the witness makes a certain statement, that what witnesses say is usually true, and that what witnesses who are not sincere say is not usually true. We can then build two conflicting defeasible arguments from these jointly consistent premises. In the usual argumentation approaches (e.g. with the above argumentation scheme for witness testimony) the first of these arguments will again compress the defeasible plus subsequent deductive inference into a single defeasible inference:

Argument C 'quick-and-dirty':
Witness E says that H
Witnesses usually speak the truth
Therefore (presumably) H

The second argument can then be modelled as an undercutter in the sense of Pollock (1995):

Argument D 'quick-and-dirty':
Witness E is not sincere
<u>Witnesses who are not sincere do not usually speak the truth</u>
Therefore the premises of argument A do not support its conclusion

We can conclude that AI models of default reasoning (that is, of direct probabilistic reasoning) are on the one hand consistent with probability theory (as least in the above argument forms) but on the other hand account for the defeasible nature of inferences where in probabilistic approaches this defeasible nature is hidden in the conditional probability estimates that are used as input to the theory.

But when is direct probabilistic reasoning appropriate? Let me discuss this by looking at why this form of reasoning does not involve explicit estimates of unconditional (in Bayesian terms 'prior') probabilities of hypotheses.

With Bayes' rule we know that "P(H|E) is usually high" is equivalent to "usually the value of the following expression is high:"

$$\frac{P(E|H) * P(H)}{P(H) * P(E \mid H) + P(\neg H) * P(E \mid \neg H)}$$

And this implies that P(H) cannot be too low (unless the denominator of this fraction is very low). So we could say that the phrase "P(H|E) is usually high" makes some implicit assumptions about the various probabilities in the above fraction, such as about P(H). In a 'deep' Bayesian modelling these assumptions are made explicit by decomposing P(H|E) into the above fraction and then computing it from the probabilities occurring in the fraction. Providing the other probabilities is often hard and sometimes impossible. Only if the statement "P(H|E) is usually high" is controversial, can it be worthwhile delving deeper into the details and trying to make the implicit probability assumptions explicit. But as long as there is no reason to do so, we simply accept such generalisations as true and reason defeasibly with them. Situations where in an explicit Bayesian modelling this would lead to base rate fallacies are then captured as exceptions to the generalisation. The important point is that as long as there is no evidence for such exceptions, we simply assume that they are not present, without having to make all kinds of explicit probability estimates. Such 'quick-and-dirty' probabilistic reasoning is, as explained above, still rational, and is preferable if the time or information to do a detailed probabilistic analysis is lacking.

This analysis is similar to Thagard's (2005) "dual pathway" model of evaluating testimony. In the next section I briefly comment on this similarity.

5 Thagard on testimony evaluation

Thagard (2005) presents a "dual pathway" model for assessing testimonies (not just witness testimonies but any human source of any kind of information). He distinguishes a "default pathway" in which people almost automatically accept a testimony and a "reflective pathway" in which people build a causal model of the relevant knowledge and decide whether to believe the testimony by inference to the best explanation (IBE). People shift from the default to the reflective pathway when the content of the testimony is inconsistent with their current beliefs or when there is reason to doubt the credibility of the source.

Thagard does not say how the default pathway looks like but it seems plausible to model it as AI-style default reasoning as we described it above. Thagard shows in detail how the reflective pathway can be modelled in his theory of explanatory coherence and how it can be implemented in his ECHO system. Essentially, a testimony is positively linked with the event or state of affairs to which it testifies, and negatively linked to other reasons for the testimony, such as a motive to lie, or circumstances that can cause flawed perception or memory.

Thagard does not model how the change from the default pathway to the reflective pathway is triggered; he confines himself to some informal remarks. In my opinion, it would be very valuable to develop a formal account of when and how a shift from the default pathway to the reflective pathway occurs. In fact, this holds irrespective of whether one wants to model the reflective pathway as IBE or in some other way, such as in a deep Bayesian analysis.

6 Conclusion

In this paper I have contrasted two forms of legal-evidential probabilistic reasoning: direct probabilistic reasoning, which directly reasons from evidence to hypotheses, and indirect probabilistic reasoning, which reasons from hypotheses to evidence (and then back to the hypotheses). While some suggest that all legal probabilistic reasoning should be indirect, I have argued that direct probabilistic reasoning has a rational basis and is, moreover, sometimes easier to perform for judges than indirect probabilistic reasoning. I have also argued that direct probabilistic reasoning can be defined in terms of standard probability theory, resulting in an alternative, non-Bayesian use of the terms "prior" and "posterior" probability and without the need to estimate unconditional probabilities of the hypotheses. I then concluded that argumentation-based approaches are suitable for modelling direct probabilistic reasoning.

A question that I have not discussed is under which conditions direct or indirect probabilistic reasoning is more suitable. All I have argued is that *sometimes* the direct form is more suitable, namely, when is safe to make general default assumptions about the conditional probability of a hypothesis given the evidence. In cases where such assumptions are not safe, some 'deeper' form of probabilistic reasoning must be applied. Then two important questions are how the need to shift from a quick-and dirty to a deeper form of reasoning can be recognised and whether this deeper form of reasoning should (always) be Bayesian. But these questions have to be left for future research.

References

Allen, L.J. (2008), Explanationism all the way down. *Episteme* 5, 320-328.

Anderson, T.J., Schum, D. A., and Twining, W. L. (2005) *Analysis of Evidence*, 2nd edition, Cambridge University Press, Cambridge.

Bennett, W. L. and Feldman, M. S. (1981), *Reconstructing Reality in the Courtroom: Justice and Judgment in American Culture*. London: Methuen-Tavistock.

Bex, F.J. (2011) *Arguments, Stories and Criminal Evidence: A Formal Hybrid Theory*. Dordrecht: Springer.

Bex, F.J., Koppen, P.J. van, Prakken, H. and Verheij B. (2010) A hybrid formal theory of arguments, stories and criminal evidence. *Artificial Intelligence and Law* 18, 123-152.

Bex, F.J., Prakken, H., Reed, C. and Walton, D.N. (2003), Towards a formal account of reasoning about evidence: argumentation schemes and generalisations. *Artificial Intelligence and Law* 11, 125 – 165.

Derksen, T. (2010), *De Ware Toedracht. Praktische Wetenschapsfilosofie voor Waarheidszoekers*. Diemen: Uitgeverij Veen Magazines B.V. (In Dutch)

Evett, I. (2009), Evaluation and professionalism, *Science and Justice* 49, pp. 159-160.

Fenton, N. E., Lagnado, D. and Neil, M. (2013). A general structure for legal arguments using Bayesian networks." *Cognitive Science* 37, 61-102.

Fenton, N.E. and Neil, M. (2011), Avoiding legal fallacies in practice using Bayesian networks', *Australian Journal of Legal Philosophy* 36, 114-151.

Hacking, I. (2001), *An Introduction to Probability and Inductive Logic*. Cambridge: Cambridge University Press.

Kadane, J.B. and Schum, D.A. (1996), *A Probabilistic Analysis of the Sacco and Vanzetti Evidence*. New York: John Wiley & Sons.

Kaptein, H., Prakken, H. and Verheij, B., eds. (2009), *Legal Evidence and Proof: Statistics, Stories, Logic*. Farnham: Ashgate Publishing.

Josephson, J.R. (2001), On the proof dynamics of inference to the best explanation, *Cardozo Law Review* 22, 1621–1643.

Lempert, R. (1986), The new evidence scholarship: analyzing the process of proof, *Boston University Law Review* 66, 439-477.

S.J. Modgil & H. Prakken, A general account of argumentation with preferences. *Artificial Intelligence* 195 (2013): 361-397.

Pearl, J. (1988), Embracing causality in default reasoning. *Artificial Intelligence* 35, 259 – 271.

Pollock, J.L. (1995), *Cognitive Carpentry. A Blueprint for How to Build a Person*. Cambridge, MA: MIT Press.

Pollock, J.L. (2009), Problems for Bayesian Epistemology. Department of Philosophy, University of Arizona.

Poole, D.L. (2001), Logical argumentation, abduction and Bayesian decision theory: a Bayesian approach to logical arguments and its application to legal evidential reasoning, *Cardozo Law Review* 22, 1733-1745.

Prakken, H. (2004), Analysing reasoning about evidence with formal models of argumentation. *Law, Probability & Risk* 3:1, 33-50.

Schum, D.A. (1994), The *Evidential Foundations of Probabilistic Reasoning*. New York: John Wiley & Sons.

Sjerps, M. & Berger, C.E.H. (2011), Het Bayesiaanse model biedt een helder zicht op een complexe werkelijkheid. Den Haag: Nederlands Forensisch Instituut. (In Dutch)

Thagard, P. (2004) Causal inference in legal decision making: explanatory coherence vs. Bayesian networks. *Applied Artificial Intelligence* 18: 231-249.

Thagard, P. (2005) Testimony, credibility, and explanatory coherence. *Erkenntnis* 63: 295-316.

Timmer, S., Meyer, J.-J.Ch., Prakken, H., Renooij, S. & Verheij, B. (2013), Inference and Attack in Bayesian networks. In *Proceedings of the 25th Benelux Conference on Artificial Intelligence (BNAIC 2013)*, 199-206.

Tversky, A. and Kahneman, D. (1974), Judgment under uncertainty: heuristics and biases. *Science*, vol. 185, no. 4157, 1124-1131.

Van Koppen, P.J. (2011), *Overtuigend Bewijs. Indammen van Rechterlijke Dwalingen*. Amsterdam: Nieuw Amsterdam. (In Dutch)

Verheij, B. (2003) Dialectical argumentation with argumentation schemes: An approach to legal logic. *Artificial Intelligence and Law* 11, 167 – 195.

Ch. Vlek, H. Prakken, S. Renooij and B. Verheij (2013a), Modeling crime scenarios in a Bayesian network. *Proceedings of the 14th International Conference on Artificial Intelligence and Law*. New York: ACM Press, 150-159.

Ch. Vlek, H. Prakken, S. Renooij and B. Verheij (2013b), Unfolding crime scenarios with variations: a method for building a Bayesian network for legal narratives. In K.D. Ashley (ed.), *Legal Knowledge and Information Systems. JURIX 2013: The Twenty-sixth Annual Conference*. Amsterdam etc, IOS Press, 145-154.

Wagenaar, W.A., Van Koppen, P.J. and Crombag, H.F.M. (1993) *Anchored Narratives: The Psychology of Criminal Evidence*, St. Martin's Press, New York (New York).

Walton, D.N. (1996), *Argumentation Schemes for Presumptive Reasoning*. Mahwah, N. J.: Erlbaum.

Walton, D.N. (2008), *Witness Testimony Evidence. Argumentation, Artificial Intelligence and Law*. Cambridge University Press.

Wigmore, J.H. (1931), *The Principles of Judicial Proof*, 2nd ed., Boston: Little, Brown and Company.