Accepted Manuscript

Resolving the so-called "probabilistic paradoxes in legal reasoning" with Bayesian networks

Jacob de Zoete, Norman Fenton, Takao Noguchi, David Lagnado

PII: S1355-0306(18)30292-2
DOI: https://doi.org/10.1016/j.scijus.2019.03.003
Reference: SCIJUS 804

To appear in: *Science & Justice*

Received date: 7 October 2018
Revised date: 25 February 2019
Accepted date: 3 March 2019

Please cite this article as: J. de Zoete, N. Fenton, T. Noguchi, et al., Resolving the so-called "probabilistic paradoxes in legal reasoning" with Bayesian networks, Science & Justice, https://doi.org/10.1016/j.scijus.2019.03.003

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

Resolving the so-called “probabilistic paradoxes in legal reasoning” with Bayesian Networks

Jacob de Zoete[1,*] jcdezoete@gmail.com, Norman Fenton[1], Takao Noguchi[1], David Lagnado[2]

[1]School of Electronic Engineering and Computer Science, Queen Mary University of London

[2]Department of Experimental Psychology, University College London

[*]Corresponding author.

25 February 2019

Abstract

Examples of reasoning problems such as the *twins problem* and *poison paradox* have been proposed by legal scholars to demonstrate the limitations of probability theory in legal reasoning. Specifically, such problems are intended to show that use of probability theory results in legal paradoxes. As such, these problems have been a powerful detriment to the use of probability theory – and particularly Bayes theorem – in the law. However, the examples only lead to ‘paradoxes’ under an artificially constrained view of probability theory and the use of the so-called likelihood ratio, in which multiple related hypotheses and pieces of evidence are squeezed into a single hypothesis variable and a single evidence variable. When the distinct relevant hypotheses and evidence are described properly in a causal model (a Bayesian network), the paradoxes vanish. In addition to the twins problem and poison paradox, we demonstrate this for the food tray example, the abuse paradox and the small town murder problem. Moreover, the resulting Bayesian networks provide a powerful framework for legal reasoning.

ACCEPTED MANUSCRIPT

1 Introduction

The idea that there are fundamental limitations to the use of probability theory within the law was formalised in the work of Cohen (Cohen, 1977). Further concerns, with a special focus on the use of Bayesian probability and the likelihood ratio in the law, have been described in work such as (Park et al., 2010), (Engel, 2012), (Pardo, 2013) and (Sullivan, 2016). This body of work includes numerous examples of puzzles intended to demonstrate that probabilistic reasoning leads to errors or 'paradoxes' in the legal context. While work such as (Allen, 1993), (Allen & Carriquiry, 1997), (Dawid, 1987), (Fenton, Berger, Lagnado, Neil, & Hsu, 2013), (Lempert, 1977) (Picinali, 2012), (Redmayne, 2009), (Schweizer, 2013) and (Schwartz & Sober, 2017) have addressed and contested some of these so-called legal paradoxes, they continue to play a role in the strong resistance to the idea of using Bayesian probability in the law (Hastie, 2019). While it is primarily legal scholars involved in such discussions, there is no doubt that the concerns raised have influenced judges and practicing lawyers; for example, the paradoxes are discussed in standard textbooks on criminal evidence such as (Roberts & Zuckerman, 2010) and underlie judgements against the use of Bayes in the law such as in cases discussed in (Fenton, Neil, & Berger, 2016).

Our objective is to show that, not only is it incorrect to conclude that the puzzles and 'paradoxes' demonstrate probability theory is incompatible with legal reasoning, but also that a causal Bayesian modelling approach ***is*** naturally compatible.

We will show that what is common in all of the example problems – and this is what creates an apparent paradox – is a failure to disentangle distinct hypotheses and pieces of evidence. The urge to couch a problem in terms of a single Boolean hypothesis H (guilty/not guilty) and a single (but consolidated) set of evidence E is a natural response to the widespread use of the likelihood ratio as a measure of probative value of evidence, but it is this artificial simplification of the underlying problem that creates the so-called paradoxes. In Section 2 we summarise this likelihood ratio approach and explain why, when there are more than two hypotheses or conditionally dependent pieces of evidence, a simplistic application of the

ACCEPTED MANUSCRIPT

likelihood ratio approach causes problems. We explain how a causal model - a Bayesian network (BN) - linking the hypotheses and evidence can help resolve these issues. In Section 3 we review the discussion in (Park et al., 2010) in order to highlight the range of concerns and misunderstandings surrounding the use of Bayes and the law. In the subsequent sections we consider the main paradoxes and show that, in each case, by disentangling relevant hypotheses and evidence in a causal BN model, it is possible to ‘'resolve’ the paradoxes and avoid the underlying misunderstandings. Indeed, we demonstrate that the BN approach actually strengthens the argument for using Bayesian probability to evaluate evidence in a legal context in the law. Further examples are provided in the Supplementary material.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

2 The likelihood ratio, its limitations and the need for Bayesian networks

We start by briefly introducing some terminology and assumptions that we will use throughout (for more detailed discussion, see (Fenton et al., 2016)). A ***hypothesis*** is a statement which we seek to evaluate. In crime cases, typically two hypotheses are considered: one related to the standpoint of the defendant, and the other related to the standpoint of the prosecutor. For example, suppose that a DNA trace was found at the crime scene, and that a defendant has been arrested. For this situation, these standpoints can be summarized with “the defendant is the source of DNA found at the crime scene” and “the defendant is not the source of DNA found at the crime scene”. The Bayesian network representation for this hypothesis pair and evidence is shown in Figure 1.

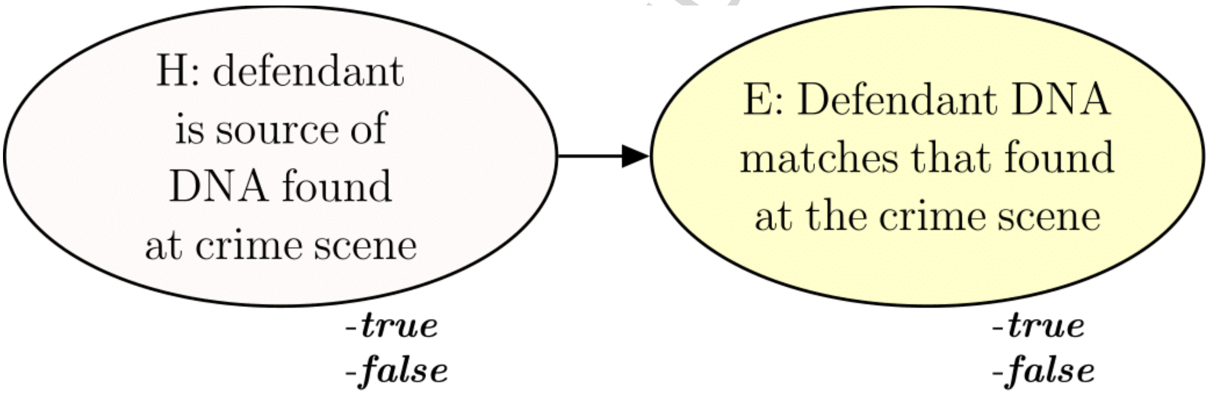

Figure 1 Causal view of evidence. This is a very simple example of a Bayesian Network (BN)

In the graphical representation in Figure 1, an arrow is drawn from the hypothesis node to the evidence node. The direction of this arrow indicates the dependency relation, for example due to causality: *H* being true (resp. false) can ***cause*** the evidence *E* to be true (resp. false).

Within this framework, the evidential value of an observation can be summarized as a likelihood ratio. The probability of observing the evidence given that a particular hypothesis is true is referred to as the *likelihood* of that observation given the hypothesis, i.e.

$$\Pr(\textit{evidence} \mid \text{prosecution hypothesis})$$

ACCEPTED MANUSCRIPT

The ratio of the two likelihoods is called the likelihood ratio (LR; (Aitken, Roberts, & Jackson, 2010)).

$$LR = \frac{\Pr(\text{evidence} \mid \text{prosecution hypothesis})}{\Pr(\text{evidence} \mid \text{defence hypothesis})}$$

A LR equal to 1 corresponds to evidence that is equally likely under both hypotheses, i.e. in isolation, it is "irrelevant" for distinguishing between these two hypotheses. A LR greater than 1 corresponds with evidence that it is more likely when the prosecution hypothesis is true than when the defence hypothesis is true. Similarly, a LR smaller than 1 corresponds to evidence that is more likely when the defence hypothesis is true than when the prosecution hypothesis is true.

In order to determine the value of the LR for the example from Figure 1, two questions need to be answered. (1) How likely is it to observe that the DNA profile of the defendant matches the DNA profile obtained from the crime stain *given that the defendant is the source of DNA found at the crime scene,* and, (2) How likely is it to observe that the DNA profile of the defendant matches the DNA profile obtained from the crime stain *given that the defendant is not the source of DNA found at the crime scene.* For illustrative purposes, assume the LR is equal to 1000.

$$LR = \frac{\Pr(\text{evidence} \mid \text{prosecution hypothesis})}{\Pr(\text{evidence} \mid \text{defence hypothesis})} = \frac{1}{1/_{1000}} = 1000$$

While the likelihood provides a measure of the probative value of the evidence in discriminating the defence hypothesis against the prosecution hypothesis, central to legal reasoning is the probability of a hypothesis: once we observe the evidence, we need to evaluate whether the defence or prosecution hypothesis is more likely. This probability of a hypothesis being true given the evidence is called the **posterior probability**. Bayes Theorem can be used to update prior beliefs regarding the prosecution and defence hypotheses into the posterior probability using the likelihood ratio. The odds form of Bayes theorem is,

ACCEPTED MANUSCRIPT

$$\text{posterior odds} = \text{prior odds} \times \text{likelihood ratio}$$

The prior odds, in terms of probabilities are equal to,

$$\frac{\Pr(\text{prosecution hypothesis})}{\Pr(\text{defence hypothesis})}$$

Assigning these prior probabilities is considered to be within the realm of the trier of fact, and correspond to answering *how likely these hypotheses are prior to considering any evidence.* These can, for the example from Figure 1, be based on an estimate regarding the number of people that could conceivably be the donor of the DNA found at the crime scene. If it is assumed that 100 people, including the defendant, could conceivably be the donor of the DNA found at the crime scene, and all of them are equally likely to be the donor, the prior probabilities are:

$$\Pr(\text{prosecution hypothesis}) = 0.01$$

$$\Pr(\text{defence hypothesis}) = 0.99$$

Hence, the prior odds are equal to,

$$\frac{\Pr(\text{prosecution hypothesis})}{\Pr(\text{defence hypothesis})} = \frac{0.01}{0.99} = \frac{1}{99}$$

And the odds form of Bayes Theorem tells us,

$$\text{posterior odds} = \frac{1}{99} \times 1000 = \frac{1000}{99}$$

In this case, where the hypotheses are exhaustive and mutually exclusive, the posterior probabilities can be retrieved from the posterior odds.

$$\Pr(\text{prosecution hypothesis}|\text{evidence}) = \frac{1000/99}{1 + 1000/99} = 0.91$$

And, similarly,

$$\Pr(\text{defence hypothesis}|\text{evidence}) = 0.09$$

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

It is important to note that, where the hypotheses are exhaustive and mutually exclusive it also follows from Bayes theorem (Fenton et al., 2016) that:

- The posterior probabilities of the hypotheses are unchanged from the priors if the $LR = 1$. In other words

$\Pr(\text{prosecution hypothesis}|\text{evidence}) = \Pr(\text{prosecution hypothesis})$ when $LR = 1$.

- The posterior probability of the prosecution hypothesis is greater than its prior if $LR > 1$.
- The posterior probability of the defence hypothesis is greater than its prior if $LR < 1$.

Hence, for exhaustive and mutually exclusive hypotheses, the LR is a genuine measure of ***probative value*** of the evidence in the sense that it really does tell us whether the evidence leads to a change in the posterior probabilities of the hypotheses. The fact that this is NOT true if the hypotheses are not exhaustive and mutually exclusive is important in the subsequent discussion.

Now suppose there are more than two alternative hypotheses. For example, suppose, it is assumed that the brother of the defendant is among the 100 possible donors of the DNA trace. Then the hypothesis H "Source of DNA found at crime scene" should have three states: (1) *defendant*, (2) *brother of defendant* and (3) *unrelated other*. Since close relatives are more likely to share a particular DNA profile than unrelated people, these relatives should be considered separately when evaluating the evidence in situations where there is reason to believe that they are among the possible donors. The following probabilities are assigned, again based on the assumption that there are 100 possible donors where the defendant and his brother are part of this group,

$$\Pr(\text{defendant}) = 0.01$$

$$\Pr(\text{brother of defendant}) = 0.01$$

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

$$\text{Pr(unrelated other)} = 0.98$$

Subsequently, one needs the probability of observing the particular DNA profile given that the brother of the defendant was the donor. Here, it is assumed that it is 100 times more likely to observe the particular DNA profile when the donor was a sibling of the defendant than when the donor was an unrelated other, i.e. the likelihoods are,

$$\text{Pr(evidence | defendant)} = 1$$

$$\text{Pr(evidence | brother of defendant)} = 0.1$$

$$\text{Pr(evidence | unrelated other)} = 0.001$$

Now, because the defence hypothesis can be regarded as a combination of two sub-hypotheses, e.g. the brother of the defendant *or* an unrelated other is the source of the DNA found at the crime scene, the corresponding prior probabilities become part of the likelihood ratio. This is already something that can easily be overlooked, for examples see (de Zoete & Sjerps, 2018).

$$\text{Pr(evidence|prosecution hypothesis)} = 1$$

$$\begin{aligned}
&\text{Pr(evidence|defence hypothesis)} \\
&\quad = (\text{Pr(evidence|brother of defendant)} \times \text{Pr(brother of defendant)} \\
&\qquad + \text{Pr(evidence|unrelated other)} \times \text{Pr(unrelated other)}) \\
&\qquad \times \frac{1}{\text{Pr(defence hypothesis)}} \\
&\quad = \frac{0.1 \times 0.01 + 0.001 \times 0.98}{0.01 + 0.98} \\
&\quad = 0.002
\end{aligned}$$

$$LR = \frac{\text{Pr(evidence | prosecution hypothesis)}}{\text{Pr(evidence |defence hypothesis)}} = \frac{1}{0.002} = 500$$

And the posterior odds become,

ACCEPTED MANUSCRIPT

$$\text{posterior odds} = \frac{1}{99} \times 500 = \frac{500}{99}$$

Again, because the hypotheses are exhaustive and mutually exclusive, the posterior probability for the prosecution hypothesis can be retrieved from the posterior odds[1].

$$\Pr(\text{prosecution hypothesis}|\text{evidence}) = \frac{{}^{500}\!/_{99}}{1 + {}^{500}\!/_{99}} = 0.83$$

Similarly,

$$\Pr(\text{brother of defendant}|\text{evidence}) = \frac{{}^{50}\!/_{549}}{1 + {}^{50}\!/_{549}} = 0.08$$

and,

$$\Pr(\text{unrelated other }|\text{evidence}) = \frac{{}^{49}\!/_{550}}{1 + {}^{49}\!/_{550}} = 0.08$$

Although it is still possible to perform these calculations manually, it is substantially more challenging now that the prior probabilities for the sub-hypotheses of the defence hypothesis are explicitly present in the likelihood ratio. When additional pieces of evidence are evaluated in conjunction to the DNA evidence manually calculating these probabilities becomes practically infeasible. As an example, consider the situation presented in the BN in Figure 2 where, in addition to the DNA evidence, there is an eyewitness that claims that the brother was out of town on the day of the crime. Several dedicated software solutions (Agena Ltd, 2019; Hojsgaard, 2012; Hugin A/S, 2018; University of Pittsburg, 2018) have been developed that can help with constructing Bayesian networks and, subsequently, performing calculations with them. Using such a software solution, the posterior probability

[1] For this particular purpose, a generic formula can be used to retrieve the posterior probability, $\Pr(\text{prosecution hypothesis}|\text{evidence}) = \frac{1}{1+\frac{1}{10}+\frac{98}{1000}} = 0.83$. See (Balding & Steele, 2015).

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

that the defendant is the source of the DNA found at the crime scene is determined to be 0.90.

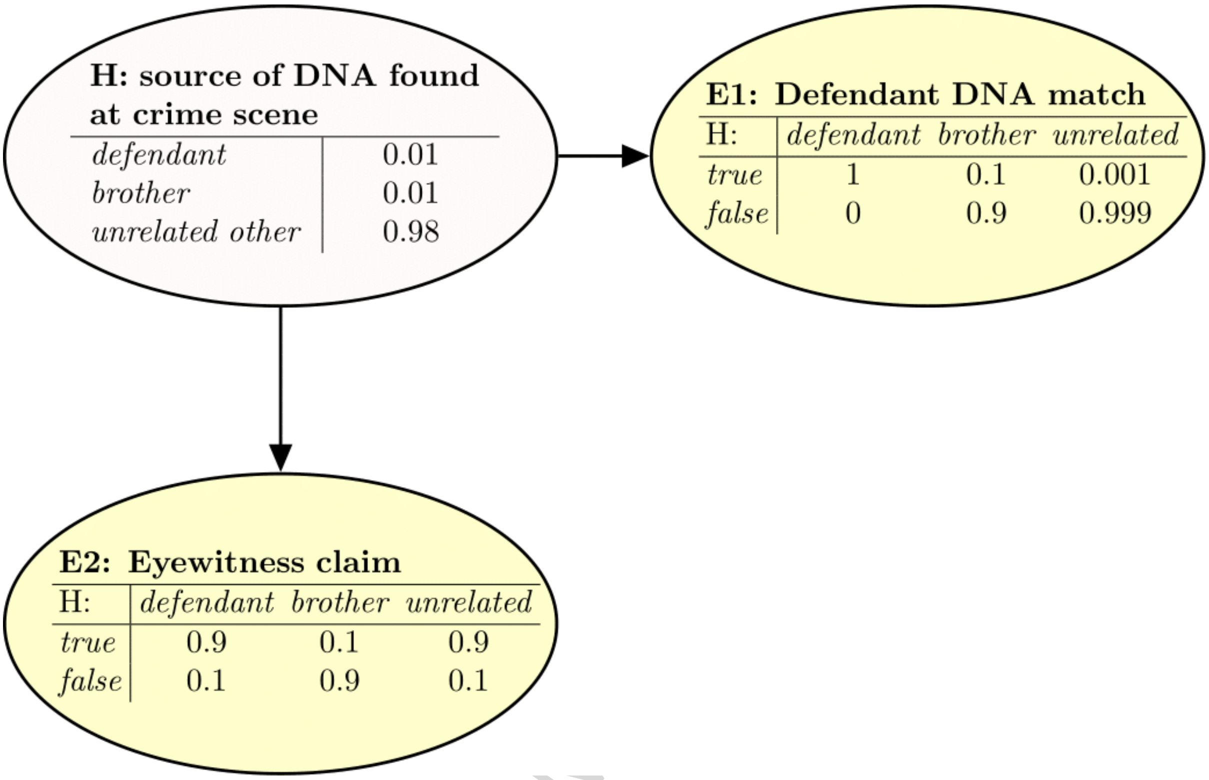

Figure 2 Bayesian network for two pieces of evidence with conditional probability tables

Furthermore, the likelihood ratio of the combined evidence can be retrieved by dividing the posterior odds by the prior odds (which are also computed automatically in the BN tool). For the example from Figure 2, this corresponds with,

$$LR = \frac{\text{posterior odds}}{\text{prior odds}} = \frac{0.9016/(1-0.9016)}{0.01/0.99} = 907$$

For illustrative purposes, the same results are manually derived in the Supplementary material, Section 1.1. In all of the BN examples that follow the probability calculations are performed using (Agena Ltd, 2019).

We believe that much of resistance to the use of Bayes is due to confusion, over-simplification and over-emphasis of the role of the LR. Namely, as can be seen from the examples presented in this paper, sceptics often present the LR in a simplistic form, e.g.

ACCEPTED MANUSCRIPT

What does this piece of evidence (in isolation) say about two (non-exhaustive) hypotheses? However, the true "power" of this probabilistic framework lies in the ability to take a more holistic view of the case, namely the hypotheses, the evidence and how they are interconnected. The issues are dealt with in depth in (Fenton et al., 2013, 2016; Fenton, Neil, & Hsu, 2014). While Bayes' Theorem and the LR provides a simple and natural match to intuitive legal reasoning in the case of a single Boolean hypothesis node *H* and a single piece of evidence *E*, practical legal arguments normally involve multiple hypotheses and pieces of evidence with complex causal dependencies. In such cases the simplistic LR approach does not provide the necessary overview, and this is the reason for the apparent 'paradoxes' described below. However, by using Bayesian networks to model the relevant hypotheses, evidence and causal dependencies it is possible to resolve the paradoxes and provide coherent and consistent conclusions about the probative value of evidence.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

3 The key issues arising from the 'Small town murder' problem

In the discussion paper *Bayes Wars Redivius– An exchange* (Park et al., 2010), Allen presents the following example (which we will refer to as the 'small town murder' problem) to claim that the LR approach does not accurately capture the concept of relevance in legal trials.

> *A person accused of murder in a small town was seen driving to the small town at a time prior to the murder. The prosecution's theory is that he was driving there to commit the murder. The defense theory is an alibi: he was driving to the town because his mother lives there to visit her. The probability of this evidence if he is guilty equals that if he is innocent, and thus the likelihood ratio is 1, and under what is suggested as the "Bayesian" analysis, it is therefore irrelevant. Yet, every judge in every trial courtroom of the country would admit it (…). And so we have a puzzle.*

Hence, specifically, the puzzle considers the problem that evidence with a likelihood ratio of 1, which occurs when it does not favour one hypotheses (prosecution) over the other (defense), is labelled irrelevant. However, as Kaye pointed out in the exchange, the problem with this conclusion is that it makes the mistake of evaluating the evidence in isolation and fails to take account of the impact of the evidence on other relevant hypotheses in the case. In other words (as is pointed out in (Fenton et al., 2013)), for such a piece of evidence it is meaningless to speak of "the likelihood ratio". The value, and therefore the degree of support, is dependent on one's assumptions with regards to the considered hypotheses and background information.

Much of the exchange focuses around disagreements about the notion of when evidence is "relevant". From the legal perspective, evidence is relevant if it has any tendency to make a fact more or less probable than it would be without the evidence.

The relevance of a piece of evidence based on the LR value only refers to the relevance in distinguishing between the considered hypotheses, i.e. the evidence is not unequivocally

ACCEPTED MANUSCRIPT

relevant (or irrelevant), it is relevant specifically with these hypotheses in mind. Hence, whether a piece of evidence is “relevant” (according to the LR approach), depends on the standpoints of the prosecution and the defence. So, as long as there is uncertainty with regards to the contents of these standpoints, all evidence can be treated as potentially relevant and can therefore be admitted. Only in situations where one cannot recognize it as having any influence on the case whatsoever (e.g. there were seven trees in the street of the crime scene) or when the “evidence” is considered to be common knowledge that does not alter the narrative of the case (e.g. the defendant has brown hair) one could deem it “irrelevant” without knowledge of the (to be) presented standpoints. Furthermore, the notion that “*if the evidence is a critical part of both parties’ case, it’s not relevant at all”* is a simplification of the issue. Even though evidence could fit within both parties’ narrative, that does not mean that it is equally likely under both hypotheses.

Gross presents such an example in (Park et al., 2010).

> *Defendant is stopped in his car three minutes after an aborted bank robbery, 1/2 a mile and speeding away from the site. Prosecution says it's relevant to guilt: it shows he was escaping. Defendant says it is relevant to innocence: no escaping bank robber would speed and attract attention. I used to be a criminal defense lawyer, so I think the defendant's argument is quite a bit more specious than the prosecutor's.*

In other words, even though the evidence is a critical part of both parties’ case, Gross believes that this piece of evidence better fits with the prosecutor’s argument than the argument of the defense, which translates to a likelihood ratio greater than 1. However, once again, it is important to stress that one cannot speak of ‘the’ LR. Especially with this example, the evidential value of the speeding evidence is dependent on the answers to sub-questions like “how likely is it that a bank robber would be speeding away from a crime scene” or “was there a police chase going on”. Given that there most likely will be a disagreement over the “answers” to such questions, it is fair to state that there cannot be a conclusive LR that defines the relevance of the evidence.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Both Gross and Allen suggest that evidence, although "irrelevant" with respect to a LR of 1 can still be relevant for the case as a whole. This is correct, mostly because pieces of evidence will usually have a (conditional) dependency relation with other pieces of evidence. Since the presented hypotheses (standpoints) will disagree on at least one aspect, it is likely that the relevance of a piece of evidence is not necessarily based on their evidential value with regards to the hypotheses "directly" but rather for establishing the evidential value of another piece. In other words, it is often insufficient to evaluate pieces of evidence in isolation since the interdependency between them says so much more. Hence, it is possible that a piece of evidence that, on its own, would be labelled irrelevant, i.e. a LR of 1, is relevant when evaluated together with another piece of evidence. We will show this in the Abuse example in Section 4.3 Similarly, it is possible that a piece of evidence with a very discriminating LR becomes "irrelevant" when evaluated together with other pieces of evidence. Consider the following example

> *At a crime scene where a fight took place, a wall is covered with blood spatters. DNA profiles are obtained from multiple blood spatters, all of them match with the DNA profile of the defendant. Furthermore, a blood spatter analyst reports that the pattern was most likely caused due to an assault with a blunt object.*

For such a situation, if the prosecution's hypothesis states that the defendant was one of the people present at the crime scene during the fight and the defence disputes this by stating that the defendant was not present at the crime scene during the fight, the DNA profiles evidence obtained from the blood spatters is very discriminating for establishing that the defendant was recently at the crime scene, and, therefore, relevant. However, for this set of hypotheses, the report of the blood spatter analyst, when evaluated in isolation of the other evidence, is irrelevant; the presence of the defendant does not change our belief in what type of pattern we expect to observe. Nonetheless, when evaluated together, the DNA profiles become relevant specifically with regards to being present during the fight due to the blood pattern report. Furthermore, the evidential value of additional reports on individual

ACCEPTED MANUSCRIPT

blood spatters diminishes for every added spatter. After "observing" that the first 10 matched the profile of the defendant, we already suspect that the 11th will do so as well. Hence, at some point, yet another report on the DNA profile of a blood spatter will become practically irrelevant, given all the other evidence, even though the piece of evidence in isolation suggests it is highly relevant.

In (Park et al., 2010) Kaye suggested that BNs could help evaluate evidence to address the issues above. Most importantly, such a presentation forces one to evaluate the evidence on the basis of multiple hypotheses and the (assumed) interdependency between pieces of evidence and hypotheses becomes explicit. There has been much concern and debate about the practicalities of constructing BNs and assigning the necessary probabilities in order to perform calculations. This is certainly a limiting factor of bringing BNs into the courtroom. Furthermore, due to the fact that, potentially, there could be countless possible scenarios that describe what caused the declared evidence it is unlikely that all of them can be satisfyingly accounted for in a single model. Nonetheless, the notion that BNs will not overcome all of the potential hurdles of a full criminal trial is no reason for them to be disregarded as helpful tools in analysing situations and evidence in general. As we show later, BNs can be helpful when determining the relevancy of particular pieces of evidence, or highlighting what is at the core of an apparent paradox and, subsequently, resolving this. Also, even without specifying definite probabilities, a BN can help in evaluating evidence.

As a very basic example, consider the BN in Figure 3 for the small town murder problem. Even without the necessary probabilities to perform calculations, the relation between hypotheses and evidence is apparent and, due to the very straightforward structure, it is even possible to formalize the relation between prior beliefs, the likelihood ratio of the evidence and the posterior probabilities. For more complex situations, this can be very difficult, but theoretically it is possible.

ACCEPTED MANUSCRIPT

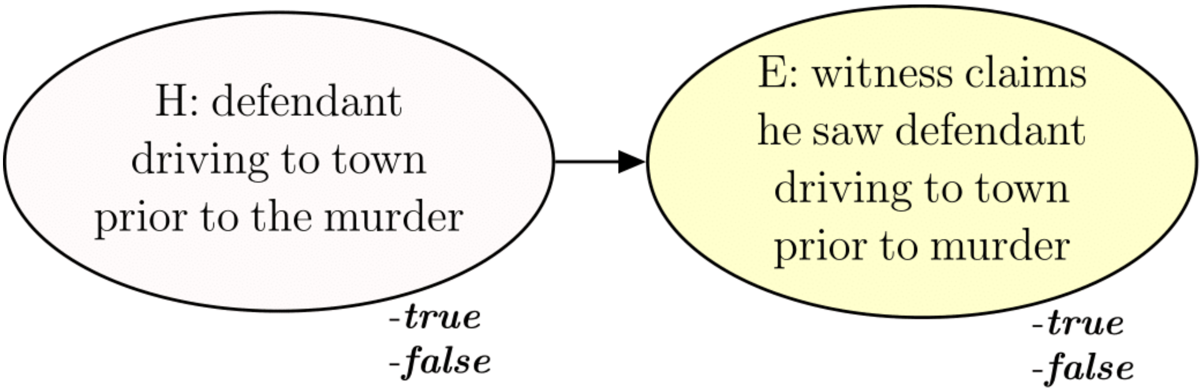

Figure 3 Very basic example of a Bayesian network for the "small town murder" problem

Nonetheless, the key point following from the small town murder example was not satisfyingly resolved with the responses of Gross and Kaye (Park et al., 2010). Allen states:

> *[Kaye] doesn't address the second point (…) that the same piece of evidence can support both guilt and innocence, making the pertinent likelihood ratio 1.0. In fact, many if not most trials have massively overlapping evidence. The actual differences between the evidentiary proffers of the opposing sides often come to only a few points, yet judges consistently let all this overlapping evidence in for just the reason Sam identifies. Thus, if the likelihood ratio approach to relevance were true in some sense, that means the trial judges throughout the country have been admitting massive amounts of irrelevant evidence.*

The notion that overlapping evidence is necessarily similar to evidence with a LR of 1 is incorrect. This links to the previous discussion that pieces of evidence, when evaluated in isolation of the other evidence could suggest that they are irrelevant when distinguishing between the competing hypotheses but could be highly relevant in the bigger picture.

As the LR is determined by two hypotheses, a different hypothesis can result in a drastic change in the likelihood ratio. To illustrate, consider the *suspect driving to town prior to the murder* example. For the hypotheses pair Hp: *Defendant (D) was in town and had the*

ACCEPTED MANUSCRIPT

opportunity to kill the deceased and Hd: *D was in town to visit his mother and was with her at the time of the murder* and the evidence E: *witness claims he saw defendant driving to town prior to murder* the LR is equal to 1, since the hypotheses both state that the defendant was in town. However, these two hypotheses present a very restricted view of the case. Essentially, one is explicitly assuming that the defendant was in town when evaluating the evidence that a witness saw him driving to town prior to the murder. If one considers this a valid assumption, i.e. one firmly believes that the defendant was in town at the time of the murder, the evidence provides no reason for accepting either hypothesis.

Alternatively, if the defence disputes that the defendant was in town, i.e. they present Hd: *D was out of town*, the likelihood ratio will be discriminative towards the prosecution hypothesis.

The hypotheses presented in those hypothesis pairs are not necessarily exhaustive, i.e. it is possible that neither of the presented hypotheses is true. During a trial, it is not only important to evaluate which presented narrative is the more likely one, given the evidence. One should also evaluate whether the more likely narrative is probable at all. Hence, a more inclusive approach would consider all three hypotheses, as in Figure 4.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

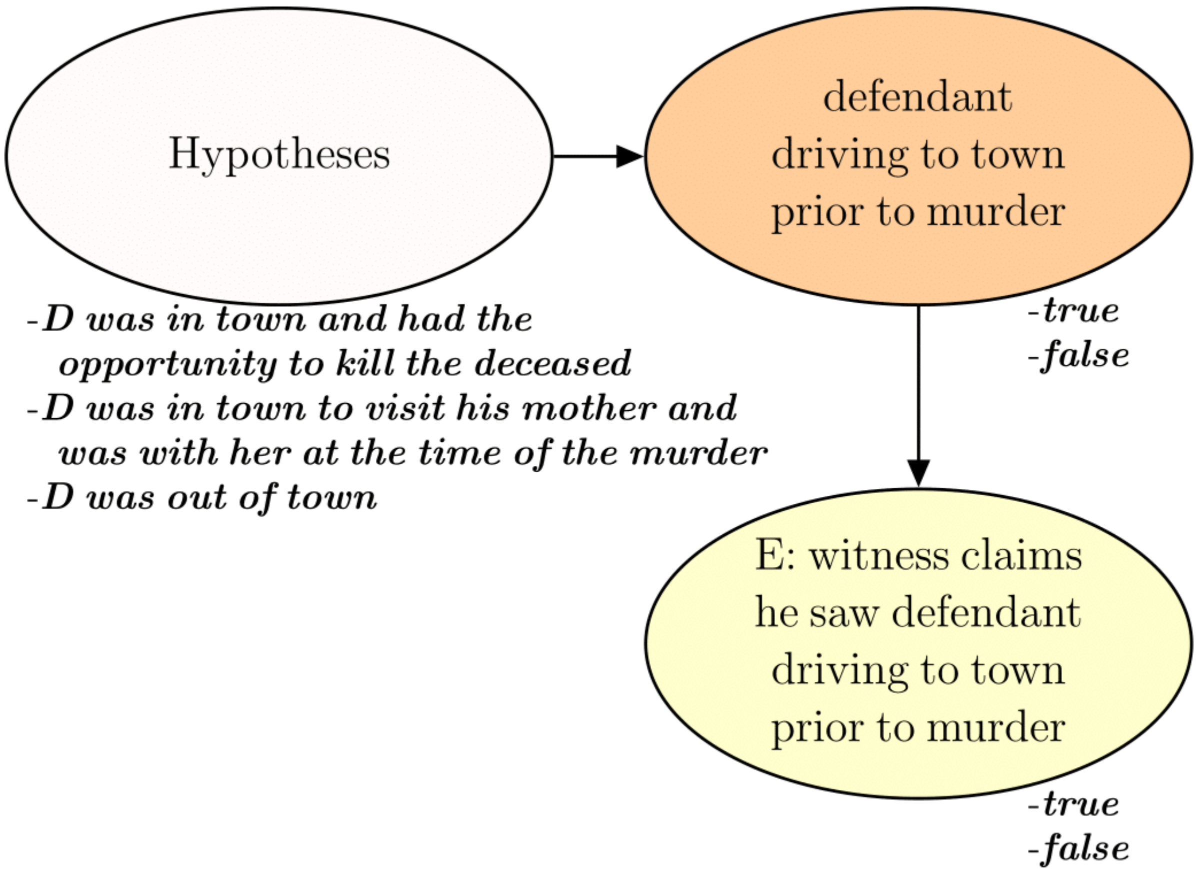

Figure 4 Defendant driving to town - exhaustive set of hypotheses

For this BN, the node "defendant driving to town prior to murder" only serves to disentangle the hypotheses node into relevant sub-hypotheses and could potentially be left out. Nonetheless, this BN may result in yet another LR. Perhaps more importantly, because the evidence is evaluated based on more than two hypotheses, the prior probabilities assigned to these hypotheses become part of the LR, see (Aitken & Taroni, 2004; de Zoete & Sjerps, 2018). Hence, in this instance, it is impossible to determine the value of the LR without specifying the prior probabilities of the hypotheses. Since it is highly uncommon that these prior probabilities are specified within a trial, it would usually be impossible to determine the value of the LR. Still, this does not imply that the LR is unfit to evaluate the "relevance" of evidence in legal trials. Consider, for example, the model in Figure 4, with unspecified probability tables as in Figure 5. As long as the prior probability for *D was out of town* is nonzero (i.e. this scenario is not impossible prior to observing any evidence), and the LR of the witness statement with regards to whether the defendant was driving to town prior to the

ACCEPTED MANUSCRIPT

murder supports that he was (i.e. a LR >1), it follows that the evidence supports the prosecution hypothesis. Furthermore, the witness statement evidence *also* supports the statement that D was in town to visit his mother, while it decreases the probability that D was out of town. Hence, the evidence is relevant.

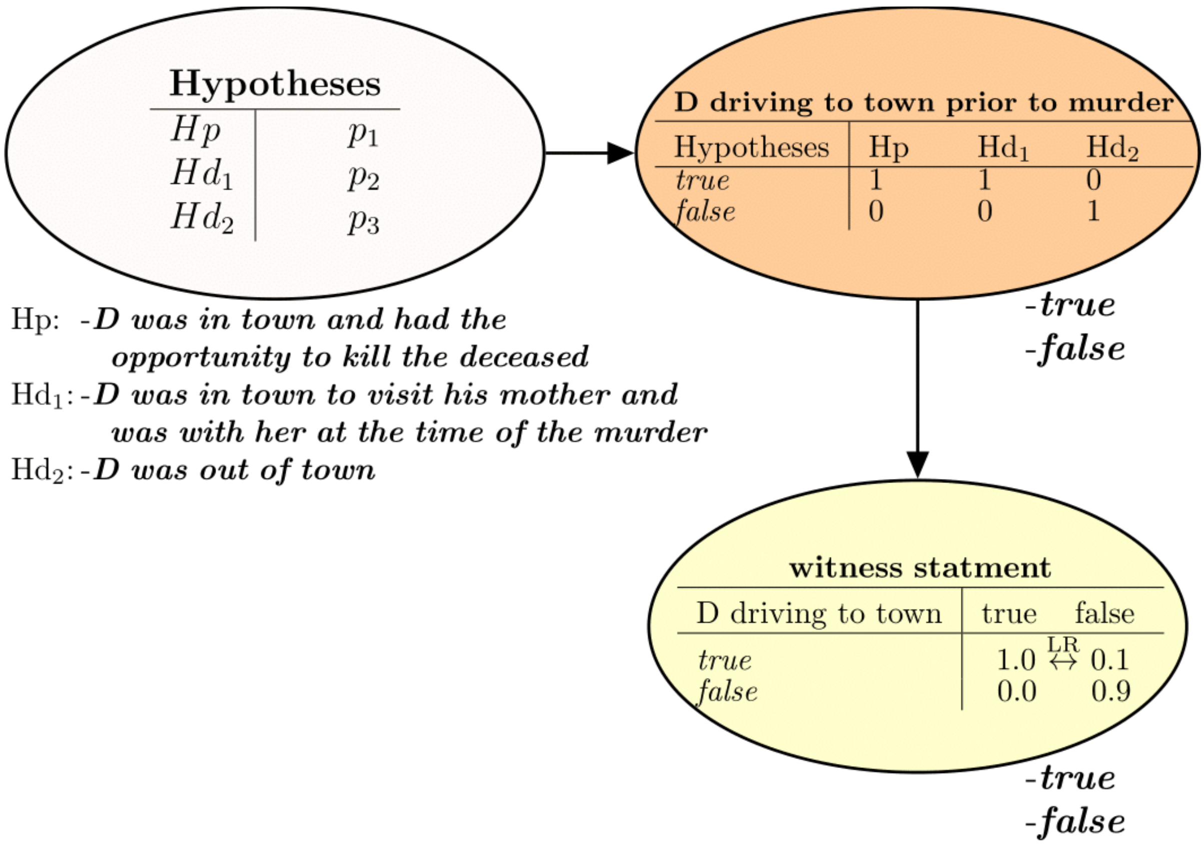

Figure 5 Defendant driving to town - probability tables

The three presented models might all result in different and possibly practically indeterminable LR values, but they also present three different scenarios under which the evidence is evaluated. The relevance of a piece of evidence according to a LR approach should only be regarded within the narrative of the evaluated hypotheses and possibly accompanying evidence. Hence, for the first model in Figure 3 a LR equal to 1 should only make the witness statement "irrelevant" when evaluating it in isolation of other evidence with regards to differentiating between the "*Defendant (D) was in town and had the opportunity to kill the deceased"* and "D *was in town to visit his mother and was with her at the time of the murder".* In several of the discussed legal "paradoxes" the observation that the LR is 1 for

ACCEPTED MANUSCRIPT

one set of hypotheses is used to label the piece of evidence as being irrelevant according to the LR approach. Subsequently, this conclusion is labelled paradoxical since the evidence is intuitively relevant for the case as a whole. For example, because the evidence is a key element of both the prosecution and the defense standpoints, i.e. because it either strengthens the belief that either of these represent what actually happened over alternative, non-mentioned, scenarios (see for example the Twins problem in Section 4.1) or because it should be regarded relevant in combination with other pieces of evidence (see the Abuse paradox in Section 4.3).

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

4 Bayesian networks for probabilistic paradoxes in legal reasoning

As a follow up to the discussions in (Park et al., 2010), (Pardo, 2013) argued that probabilistic conception of evidence produces many theoretical and practical problems and should not be used in the court. To illustrate, Pardo discussed a number of example problems. We review these problems and, in each case, identify the misunderstandings that result in the apparent paradox in legal reasoning. We then show that a correct representation with a Bayesian network avoids the paradox. Four of the problems ("Twins", "Food tray", "Poison" and "Abuse") are reviewed here while three more ("Lottery", "Liberal candidates" and "Typewriter" are worked out in a similar way and are available in the Supplementary material, Section 2.

4.1 Twins problem

The so-called Twins problem is stated in (Pardo, 2013) as:

> *A witness testifies that someone matching the defendant's appearance was seen fleeing a crime scene. The defendant claims that it was his identical twin and introduces evidence establishing the twin's existence. Suppose there is no reason to believe the testimony distinguishes the defendant from his twin.*

Pardo notes that

> *If we are comparing the likelihood of the defendant's guilt versus his twin, then (…) there does not appear to be any reason to think the likelihood ratio is different from 1. Nevertheless, the evidence is relevant.*

And on a probabilistic interpretation of this evidence,

> *Of course, the probabilist has a rejoinder as to why the evidence is also relevant under a probabilistic interpretation: namely, it eliminates everybody except the defendant and his twin, and by eliminating everyone else it thereby increases the probability the defendant is guilty. The rejoinder is correct—but notice the tension*

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

> *between this conclusion and the implications of the likelihood-ratio view. Although the evidence is relevant because it eliminates all other suspects, it technically fails to fit the likelihood-ratio conception as soon as evidence about the twin is introduced. As soon as the twin evidence is introduced, the probability of the evidence, given the defendant's guilt, is exactly the same as the probability of the evidence, given the defendant's nonguilt (assuming this is equivalent to the probability of the twin's guilt). If that is so, then under this interpretation the likelihood ratio implies that the witness's testimony should be excluded as irrelevant.*

The analysis by Pardo presents a misunderstanding with how one should incorporate a `likelihood ratio approach' when dealing with such evidence. This probabilistic approach requires clear definitions on what hypotheses are evaluated. In Pardo's analysis, the exact hypotheses that are compared change multiple times. Namely, Pardo notes that the evidence eliminates everybody except the defendant and the twin. Hence, here three possibilities are considered with regards to the person fleeing the crime scene:

1. The defendant
2. The twin of the defendant
3. Someone else

However, when evaluating the eyewitness evidence, in terms of relevance, using a likelihood ratio approach, only two are considered.

1. The defendant is guilty
2. The defendant is not guilty (assuming this is equivalent to the probability of the twin's guilt).

It is important to highlight that, as with the "small town murder problem, the notion of 'the likelihood ratio' as described by Pardo is at the core of the misunderstanding. Indeed, for distinguishing between the twin and the defendant, the LR is 1 and the evidence is irrelevant. However, it is incorrect to therefore conclude that the evidence is irrelevant for the

ACCEPTED MANUSCRIPT

case as a whole. The LR only allows one to distinguish between the associated hypotheses. In the twin example, the evidence is relevant because it distinguishes between people that look like the defendant and people who do not. Hence, by explicitly incorporating other people in the analysis, the LR will differ from 1 and hence, be relevant when distinguishing between these hypotheses.

If we ignore details such as whether the witness was accurate, whether people other than the twins would match the same description, and whether fleeing the scene is the same as guilty (our later numeric example *does* consider a Bayesian network where these are taken into account), then a Bayesian Network representation of the problem is the two node network shown in Figure 6.

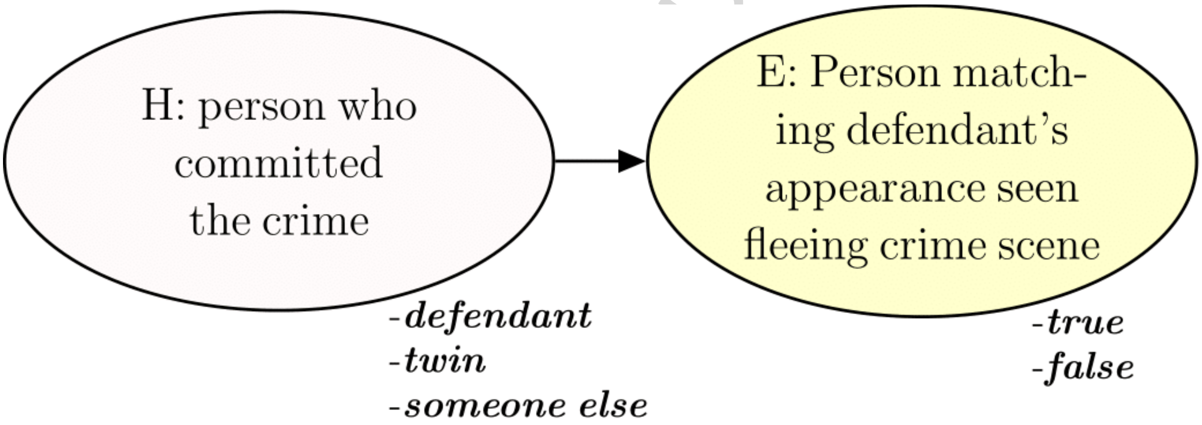

Figure 6 Simple formulation of twins problem

In this analysis, the possibility that 'someone else' committed the crime is not ruled out. For illustration purposes equal probabilities are assigned to each of the states of H, i.e. 1/3 each. The conditional probability table of the evidence node is defined as shown in Table 1.

ACCEPTED MANUSCRIPT

Table 1 CPT for evidence node E: Person matching defendant's appearance seen fleeing crime scene given H

H: person who committed the crime	defendant	twin	someone else
true	1	1	0
false	0	0	1

("person who committed the crime")

For this model, the prior probabilities for the different hypotheses are updated as in Table 2,

H: person who committed the crime	Prior probability / Pr(H)	Posterior probability / Pr(H|E)
defendant	1/3	½
twin	1/3	½

Table 2 Prior and posterior probabilities for simplified twin example, N=3

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

someone else	1/3	0

Crucially, this model presents the following (non-paradoxical and consistent) facts:

1. The evidence does not help to distinguish the guilt of the defendant and the twin since the likelihood ratio (see Table 1):

$$\frac{\Pr(E|\text{Defendant committed the crime})}{\Pr(E|\text{twin committed the crime})} = 1$$

   Hence, the posterior odds between of defendant guilty and twin guilty are equal to the prior odds.

2. However, the likelihood ratio for the exhaustive pair of hypotheses "defendant guilty" and "defendant not guilty" is easily determined by dividing the posterior odds by the prior odds (see Table 2, the same result is derived in Supplementary material, Section 1.2).

$$LR = \frac{\frac{1}{2} \Big/ \frac{1}{2}}{\frac{1}{3} \Big/ \frac{2}{3}} = 2,$$

   which confirms that the evidence is relevant (since the LR is not 1) for this set of hypotheses. More specifically, the evidence ***does*** support the hypothesis that the defendant is guilty. This is also confirmed by the fact that the posterior probability for "defendant committed the crime" increases compared to the prior probability from 0.33 to 0.50.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

So, while the evidence is not 'probative' in distinguishing between whether the defendant or their twin committed the crime it certainly ***is*** probative in distinguishing between the defendant committing the crime or the defending being innocent. And the model shows both of these assertions. Note that, especially for more complicated situations, expressing the likelihood ratio as a formula (see Supplementary material, Section 1.2) of all the relevant probabilities will become practically infeasible. Instead we use the Bayesian network tool and simply divide posterior and prior odds.

A Bayesian network representation can be used to include other uncertainties associated with such a case. For example, the Bayesian network in Figure 7, incorporates the accuracy of the witness, the size of the offender population as a parameter to determine the prior probabilities for the different hypotheses and the reliability of the evidence that establishes that the defendant has a twin.

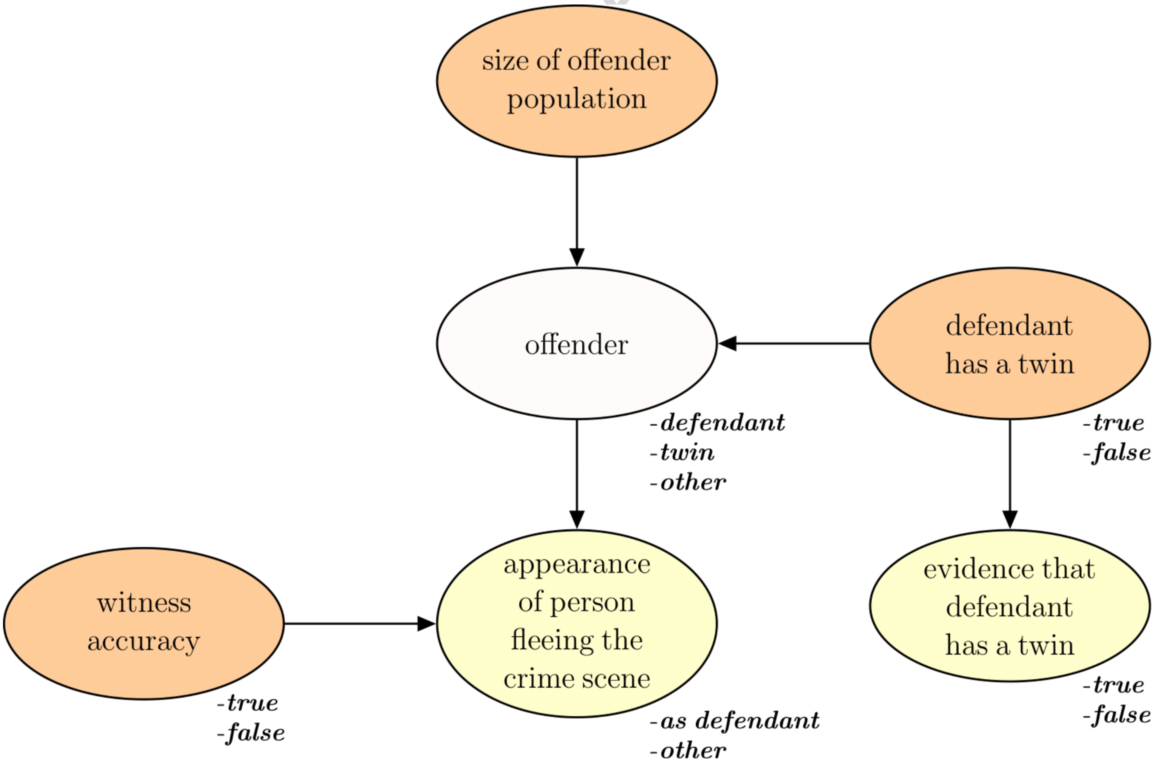

Figure 7 Bayesian network for twin example

ACCEPTED MANUSCRIPT

When using the probability assignments from Table 3 for the conditional probability tables, the prior probability and posterior probabilities are as in Table 4. By setting both the “appearance of person fleeing the crime scene” to “as defendant” and “evidence that defendant has a twin” to “true” the posterior probabilities are obtained using a Bayesian network tool. By dividing the posterior odds and the prior odds, the likelihood ratio of the combined evidence can be retrieved. In this case, the LR is approximately 42.

Table 3 Probability assignments for Bayesian network from Figure 7

Parameter	Assignment
size of offender population	1000
Pr(defendant has a twin)	0.01
Pr(similar appearance as defendant | someone else fled the crime scene)	0.02
Pr(twin evidence | twin exists)	1.00
Pr(twin evidence | twin does not exist)	0.05
Pr(witness accurate)	0.85

Table 4 Probability assignments for Bayesian network from Figure 12

H: person who committed the crime	Prior probability (Pr(H))	Posterior probability (Pr(H|E))

ACCEPTED MANUSCRIPT

Defendant	0.1%	4.07%
Twin	0.001%	0.68%
Someone else	99.899%	95.25%

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

4.2 Food tray example

The following example based on *People v. Johnson* presented in (Allen, Kunhs, Swift, Schawartz, & Pardo, 2011) and discussed in (Pardo, 2013) further extends the need to evaluate the evidence with regards to a bigger set of uncertain events.

> *The defendant, an inmate at a maximum-security prison, was charged with two counts of battery on prison guards. The charges arose from an altercation between the defendant and guards after the defendant refused to return a food tray in his cell. The prosecution's theory was that the defendant battered the officers when they opened the cell door to retrieve the tray. The defendant testified that one of the guards rushed in and began hitting him first, and his attorney argued that, even if the defendant made contact first with the officer, the defendant was acting in self-defense.*
>
> *(…) The attorneys discussed (…) that the defendant had not received a package sent to him by his family, and that after several weeks and several attempts to speak with a sergeant about it, the defendant refused to return his food tray. (…) Each side used this evidence to support its competing theory: (1) the defendant was frustrated and angry about not receiving the package, withheld his tray, and charged the guard, and (2) the defendant was frustrated about not receiving the package, withheld the tray to get a sergeant's attention about the matter, and in response the guards attacked him (to retaliate or punish him for this behaviour).*

Pardo notes (Pardo, 2013),

> *The evidence does not appear to distinguish between the two theories; […] there is no reason to believe that this evidence supports one theory over the other. In other words, the likelihood ratio is 1:1. Under the likelihood-ratio theory, the evidence is irrelevant (and a fortiori has no probative value), and, thus, should have been excluded.*

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

This example highlights a limitation of the simple likelihood ratio approach, which considers only one piece of evidence at one time based on one uncertain event, like in the Bayesian network of Figure 8. In particular, the evidence, that the defendant did not receive a package, does not reject either of the theories on its own. When this evidence is considered in conjuncture with other possible pieces of evidence, however, the evidence can provide stronger support to one of the theories. To illustrate, see the Bayesian network proposed in Figure 9.

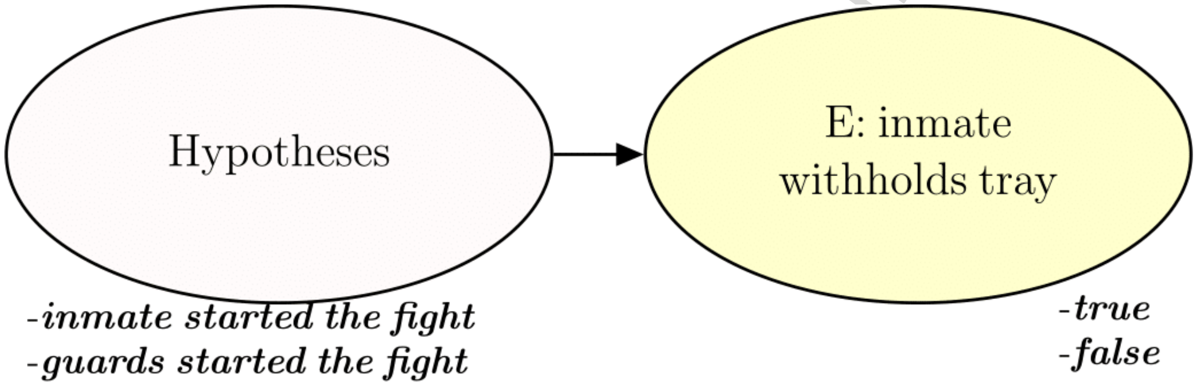

Figure 8 Simple Bayesian network for foodtray example

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

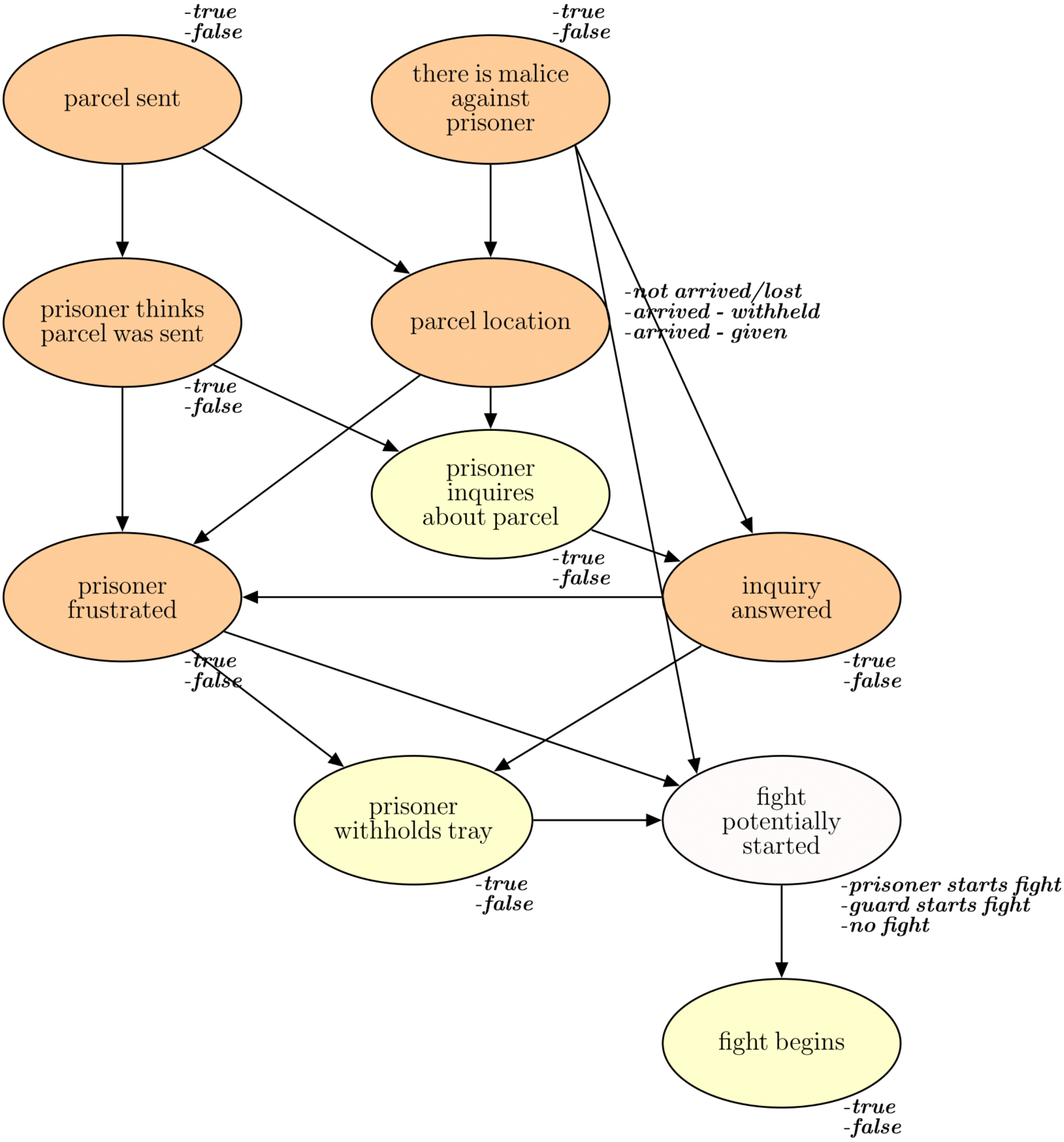

Figure 9 Bayesian network for the foodtray example

This network is a representation of the defendant's and guards' theories. This network has 10 nodes, indicating that 10 pieces of facts should be examined to validate the theory: for example, the location of the parcel, and whether there is malice among the guards against the prisoner. All of these, together with the evidence that the prisoner did withhold his tray and a fight started, influence the belief with regards to who started the fight. For example, if one is assigning a very high probability to the guard having malice against the prisoner, this

ACCEPTED MANUSCRIPT

will increase the belief that they withheld the parcel, that the prisoner is frustrated because of that and therefore withholds the tray. Through all of this, it will increase the probability that the guard started the fight. Again, establishing a concrete value of the likelihood ratio is practically infeasible. First of all, it requires one to assign probabilities to all of the nodes, and furthermore, one should unanimously agree that the model from Figure 9 exactly captures the situation. Nonetheless, the model shows that the evidential value of "prisoner withholds tray" with regards to who started the fight depends on a whole range of uncertain events and that it is practically impossible that one's combined beliefs in these will result in a likelihood ratio of 1. Furthermore, because answers to the questions represented by nodes will presumably be discussed in a trial, i.e. "was a parcel sent?", it is impossible to assign the evidential value of "prisoner withholds tray" before the actual trial.

Importantly, an interaction of these facts can help us to distinguish the defendant's and guards' theories. If it is established, for example, that the guards generally hold malice against prisoners, then the evidence that the defendant did not receive the package implies malice against the defendant among the guards. Hence, this evidence provides a stronger support for the defendant's theory than the guards'. Therefore, by highlighting a limitation of a simple, straightforward, likelihood ratio approach (as in Figure 8), this example suggests that a more elaborate probabilistic approach is necessary. The limitation can be overcome with Bayesian networks.

The Bayesian network representation allows for a more careful examination of the influence of some probability assignments on the question of interest, i.e. who started the fight. For example, how does uncertainty about whether the parcel was sent in the first place affect the probability that the defendant was the one starting the fight? In Table 5 two different probability assignments are given representing two different "stories". In the first probability assignment, it is assumed that it is very likely that the parcel was sent and, similarly, that there is a malice against the prisoner. In the second, an opposite scenario is assumed. Ideally these (prior) probability assignments are based on further evidence, e.g. statements

ACCEPTED MANUSCRIPT

from other inmates or a paper trail for the parcel. The posterior probability given the Bayesian network representation from Figure 9 that the defendant started the fight for the first set of probability assignments is 19%. In the second scenario this posterior probability is 74%. Assigning fixed, final, probabilities to these events can be practically impossible, and any assignment can be contested on the value, the underlying evidence and reasoning or even on whether the underlying uncertainty can be captured as a single probability. Hence, one should not focus solely on the resulting posterior probabilities but concentrate on the model structure and the fact that the "relevance" of a piece of evidence is based on a much larger set of (unknown) events. Even though one can criticize the structure, the probability assignments and the considered set of evidence, the fact that one cannot simply regard the "prisoner withholds tray" evidence as irrelevant evidence with a LR of 1 because it fits both stories is clear from the network structure. Furthermore, a "sensitivity analysis" can be run on a Bayesian network structure like the one in Figure 9. Such an analysis provides insight with regards to the more influential probability assignments or evidence nodes.

Table 5 Probability assignments for Foodtray example

Parameter	**Probability assignment - 1**	**Probability assignment - 2**
$\boldsymbol{Pr}(\textbf{parcel sent})$	0.9	0.1
$\boldsymbol{Pr}(\textbf{malice against prisoner})$	0.9	0.1
$\boldsymbol{Pr}(\textbf{prisoner thinks parcel was sent} \mid \textbf{parcel sent})$	1.0	0.9
$\boldsymbol{Pr}(\textbf{prisoner thinks parcel was sent} \mid \textbf{parcel not sent})$	0.0	0.1
$\boldsymbol{Pr}(\textbf{parcel lost} \mid \textbf{parcel sent})$	0.1	0.1

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

$Pr($parcel withheld | parcel sent, malice$)$	0.8	0.8
$Pr($parcel withheld | parcel sent, no malice$)$	0.1	0.1
$Pr($inquires about parcel | parcel not with prisoner$)$	0.9	0.9
$Pr($inquiry answered | malice$)$	0.1	0.1
$Pr($frustrated | inquiry not answered, parcel not with prisoner$)$	0.9	0.9
$Pr($frustrated | inquiry answered, parcel not with prisoner$)$	0.5	0.5
$Pr($withholds tray | inquiry answered, not frustrated$)$	0.0	0.0
$Pr($withholds tray | inquiry not answered, not frustrated$)$	0.1	0.1
$Pr($withholds tray | inquiry answered, frustrated$)$	0.5	0.5
$Pr($withholds tray | inquiry not answered, frustrated$)$	0.9	0.9
$Pr($prisoner starts fight | withholds tray, malice against prisoner$)$	0.2	0.2
$Pr($prisoner starts fight | withholds tray, no malice against prisoner$)$	1.0	1.0
Posterior probability – prisoner starts fight	19%	74%

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

4.3 Abuse example

This example was originally presented in (John William Strong, Kenneth S. Broun, George E. Dix, Edward J. Imwinkelried, & D. H. Kaye, 1999) and concerns ``a behavioural pattern said to be characteristic of abused children'' (also of relevance to this example is (Lyon & Koehler, 1996)). Once again, a likelihood ratio of 1 is at the root of the paradox. However, for this example, similar to the Poison example presented in Section 4.4 the apparent paradox is due to evaluating the evidence in isolation contrary to a combined evaluation.

> *If research established that the behaviour is equally common among abused and non-abused children, then its likelihood ratio would be 1, and evidence of that pattern would not be probative of abuse'' (…) And if it were a thousand times more common among abused children, its probative value would be far greater.*

Pardo notes (Pardo, 2013)

> *(…) Even if the behaviour is equally common among both groups of children, it might nevertheless be highly probative in a given case if, for example, abused children exhibiting this behaviour also possess, and non-abused children lack, an additional characteristic and the particular child at issue possesses (or lacks) this characteristic*

The probabilistic fallacy here is that one should not evaluate the evidence sequentially but simultaneously. This fallacy can be exposed by structuring the problem and evaluating the evidence using a Bayesian network. Furthermore, Pardo recognizes a reference class problem:

> *(…) the probative value may nevertheless be minimal if the child possesses (or lacks) an additional characteristic that places the child in the group of non-abused children who exhibit the behaviour.*

Hence, three groups of children are recognized:

1. Abused children
2. Non-abused children

ACCEPTED MANUSCRIPT

a. Non-abused children - exhibiting abuse-related behaviour

b. Non-abused children - not exhibiting abuse-related behaviour

By distinguishing between these groups in the analysis or Bayesian network, one can observe that two pieces of evidence that are individually uninformative with regards to the question of whether a child was abused can be very discriminative when evaluated together. A Bayesian network structure for this example is given in Figure 10. The (conditional) probability tables should account for the assumption that the behavioural pattern said to be characteristic of abused children is equally common among abused and non-abused children. In other words, observing this behaviour should not alter one's belief in whether the child was abused. Only when evaluated in concurrence with an additional characteristic, the behaviour becomes highly probative. This can be mimicked in the probabilistic model from Figure 10 by setting the (conditional) probabilities to the values from Table 6 (for the equations that should be satisfied see Supplementary material, Section 1.3). Both pieces of evidence are, individually, uninformative with regards to whether a child was abused. They do, however, alter the posterior distribution among non-abused children exhibiting the abuse related behaviour. If one wouldn't distinguish between non abused children that do exhibit this behaviour and only focus on the "ultimate" hypothesis, *was this child abused*, the Bayesian network representation is as in Figure 11.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

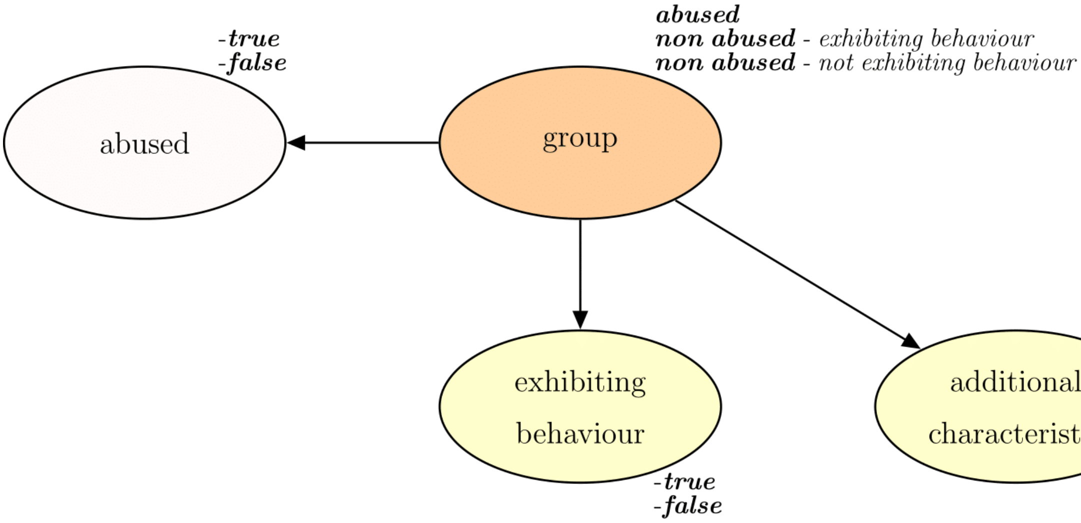

Figure 10 Bayesian network for abuse example

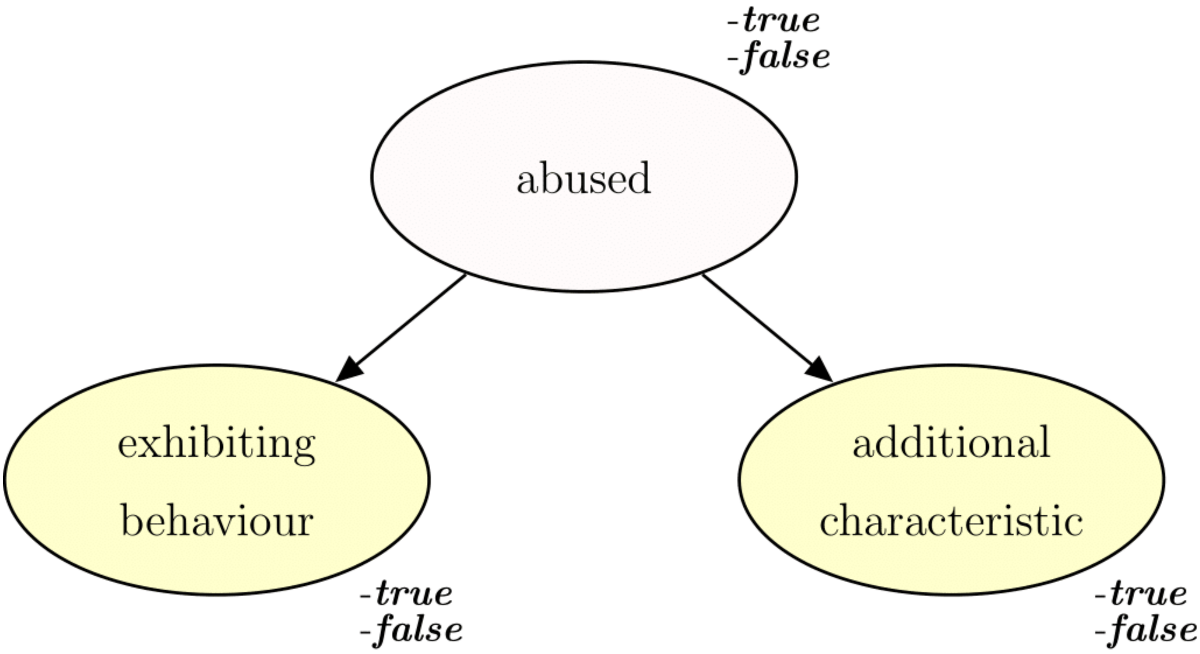

Figure 11 Bayesian network for abuse example, restricted view

The Bayesian network in Figure 11 presents a restricted view of the abuse example and is at the core of the apparent paradox. Indeed, when evaluating the evidence based on the same (conditional) probabilities, the evidence, individually but also combined, suggests a LR of 1, while the "complete" overview shows the correct evaluation.

ACCEPTED MANUSCRIPT

The (conditional) probabilities from Table 6 capture the essence of this example. For the Bayesian network representing the restricted view from Figure 11, inserting the evidence will not alter the prior belief that a child was abused. For the "complete" representation in Figure 10 does show the influence of evaluating the joint evidence with respect to the known sub-categories of the non-abused group. The results are summarized in Table 7.

Table 6 Probability assignments for Abuse example

Parameter	**Probability assignment**
$\Pr(\text{abused})$	0.4
$\Pr(\text{non abused, exhibiting behaviour})$	0.5
$\Pr(\text{non abused, not exhibiting behaviour})$	0.1
$\Pr(\text{exhibiting behaviour} \mid \text{abused})$	0.5
$\Pr(\text{exhibiting behaviour} \mid \text{non abused, behaviour})$	0.6
$\Pr(\text{exhibiting behaviour} \mid \text{non abused, not behaviour})$	0.0
$\Pr(\text{additional characteristic} \mid \text{abused})$	0.1
$\Pr(\text{additional characteristic} \mid \text{non abused, behaviour})$	0.0
$\Pr(\text{additional characteristic} \mid \text{non abused, not behaviour})$	0.6

The Bayesian network simultaneously visualizes how to evaluate such a problem with subcategories for certain hypotheses (two groups of non-abused children) and allows for the effortless evaluation of the combined evidential value. Although it might be challenging to

ACCEPTED MANUSCRIPT

assess the necessary probabilities, it does identify the equalities that must hold and focuses on the dependency structure of the problem.

Table 7 Posterior probabilities restricted and complete model

	Posterior probability abused	
Evidence	**Restricted model (Figure 11)**	**Complete model (Figure 10)**
none	40%	40%
Exhibiting behaviour	40%	40%
Additional characteristic	40%	40%
Exhibiting behaviour AND additional characteristic	40%	100%

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

4.4 Poison example

This example is based on a similar example in (Achinstein, 2001). Here, the wording from (Pardo, 2013) is used. Like the food tray example, a situation is described in which a straightforward, simple, analysis of the evidence is insufficient to evaluate the situation as a whole. Furthermore, contrary to the previous examples, the presented paradox does not rely on an apparent likelihood ratio value equal to 1.

> *The Prosecution alleges that Victim died of poisoning, and Defendant contends that Victim died from some other cause. There is evidence that at 12:00 p.m. on the day he collapsed and died, Victim's lunch contained a poison that is fatal for ninety percent of the people who ingest it. Suppose there is also evidence that at 12:30 p.m., Victim ingested a second poison concealed in a drink that completely counteracts the first poison; however, it is fatal for eighty percent of the people who ingest it.*

Pardo notes (Pardo, 2013),

> *Is evidence of the second poison relevant for proving that Victim died of poisoning? Yes, of course. Articulating exactly why, however, is critical for understanding the potential analytic gap between epistemic relevance and probability. (…)*
>
> *(…) First, because the evidence lowers the probability [that Victim died of poisoning], it is also relevant for disproving that Victim died of poisoning. (…) but this, by hypothesis, was not the Prosecutor's theory of relevance for seeking to admit the evidence.*

First of all, it is insufficiently clarified in the example that both the eighty and the ninety percent should be regarded as prior probabilities that someone would die after ingesting the poison. This prior should be updated after observing the "evidence" that the victim died. Furthermore, in order to determine the posterior probability that the victim died of poisoning

ACCEPTED MANUSCRIPT

a prior probability for dying due to some other cause is required. For example, if we assume that the probability of dying due to some other cause is 10% (which presumably is rather large), the posterior probability that the victim died of poisoning after "inserting" the evidence that the victim died is 99% (based on a probability of 90% of dying due to ingesting the poison. Similarly, when the probability that one dies due to ingesting the poison is 80%, this posterior probability drops to 98%.

This does not solve the issue with the fact that the posterior probability drops after "inserting" the evidence that the second meal also contained a poison. However, the notion that the second poison is "of course" relevant for proving that Victim died of poisoning and, furthermore, that is should be relevant in terms of supporting the Prosecution's theory requires a careful consideration of what should be treated as "uncertain". An example of such a situation is presented in Figure 12.

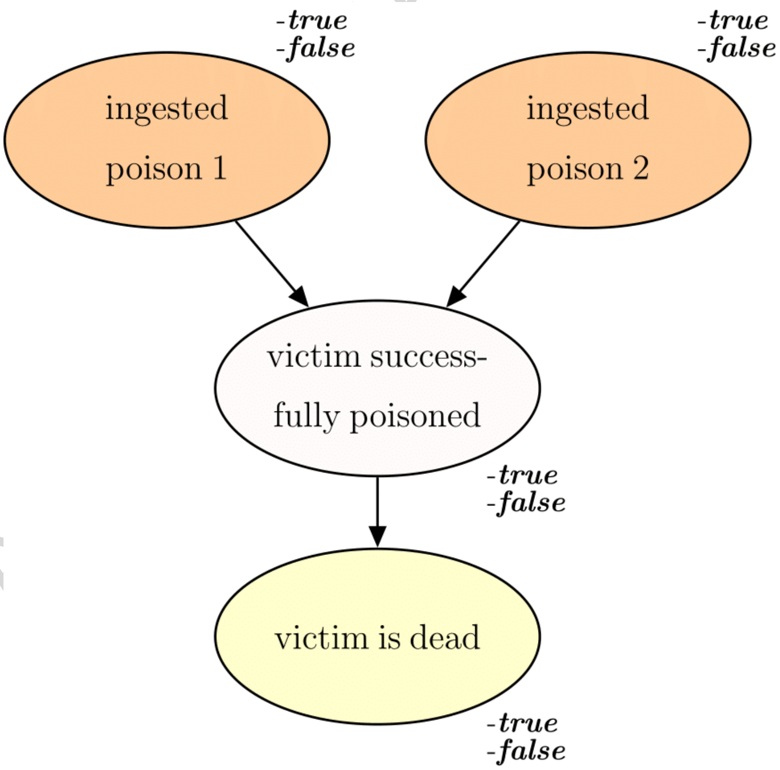

Figure 12 Victim dies of poisoning basic network

Pardo (Pardo, 2013) further states (page 584):

ACCEPTED MANUSCRIPT

> *Alternatively, the probabilist defender may also attempt to recharacterize the example so that it supports the Prosecution's theory while also resulting in an increase in probability. For example, we might separate the two effects of the second poison as two distinct pieces of evidence: counteracting the first poison and causing death. Under this reinterpretation, the first piece of evidence lowers the probability to zero percent and, then, the second piece of evidence raises the probability to 0.8, thus making the evidence relevant and raising the probability. This type of ad hoc recharacterization suggests that there may indeed be creative ways to make the probabilistic conception fit with epistemic relevance.*

Here, it is suggested that it should be possible to treat the different pieces of evidence sequentially, i.e. by first evaluating the change in posterior probability of the first piece of evidence, determining whether it is `relevant' based on the influence it has on the probability distribution and repeating this for the next piece of evidence. However, this often does not contribute to a clear understanding of the joint evidential value of the pieces of evidence. In this example, if it is absolutely certain that the victim ingested both poisons, then the probability that those combined poisons are lethal is 80%. If it is known that the first poison is completely counteracted, it is nonsensical to consider the probability of 90% for the first meal as a relevant probability, i.e. any other probability assignment would lead to the same result. Pardo's discussion seems to conflate and confuse two different hypotheses:

- determining whether poison was the cause of death (normally the domain of a coroner's court)
- determining whether the defendant intended to poison the victim (the domain of a criminal court)

These are, of course, different. If we were to focus on the second of these (which, for simplicity, we will not do in what follows) then having the two pieces of poison evidence is clearly relevant even though the first may be irrelevant in determining cause of death.

ACCEPTED MANUSCRIPT

If one is certain that the evidence should be probative for establishing the prosecution hypothesis, a very careful consideration is needed. In (Pardo, 2013) it is stated, in relation with the *explanatory conception* method that,

> *The second poisoning is part of the prosecution's explanation of what occurred. Even if the evidence lowers the probability of poisoning from the probability prior to its introduction, it nonetheless provides evidence that supports, or provides a reason to believe, the prosecution's explanation. It is relevant.*

This can definitely be the case and, furthermore, can be made visible using a probabilistic model. However, such a model requires careful consideration of the relevant uncertainties. A Bayesian network structure can help create awareness for the necessity of these parameters and it forces us to specify why and how certain pieces of evidence are relevant in establishing a certain hypothesis.

If it is certain that both meal 1 and meal 2 contained poison, but there exists uncertainty regarding whether the victim ate those meals, like in Table 8, the probability that the victim was poisoned increases once one introduces the second meal as evidence.

Table 8 Probability assignments for numeric example

Parameter	**Assignment**
Pr(**victim is dead | victim not poisoned**)	0.10
Pr(**victim ate meal 1**)	0.80
Pr(**victim ate meal 2**)	0.80
Pr(**meal 1 contained poison**)	1.00
Pr(**meal 2 contained poison**)	1.00
Posterior probability	97%

ACCEPTED MANUSCRIPT

Using the (conditional) probabilities from Table 8 for the model in Figure 13, the posterior probability that the victim was poisoned is 96% without the second meal as evidence. Once the second meal is introduced, the posterior probability increases to 97%. Hence, for this formalization of the example and the associated uncertainties, the second meal is, of course, relevant for proving that the victim died of poisoning. Here, key is that it is unequivocally clear why the second meal increases one's belief. The uncertainties and the relation between them are formalized.

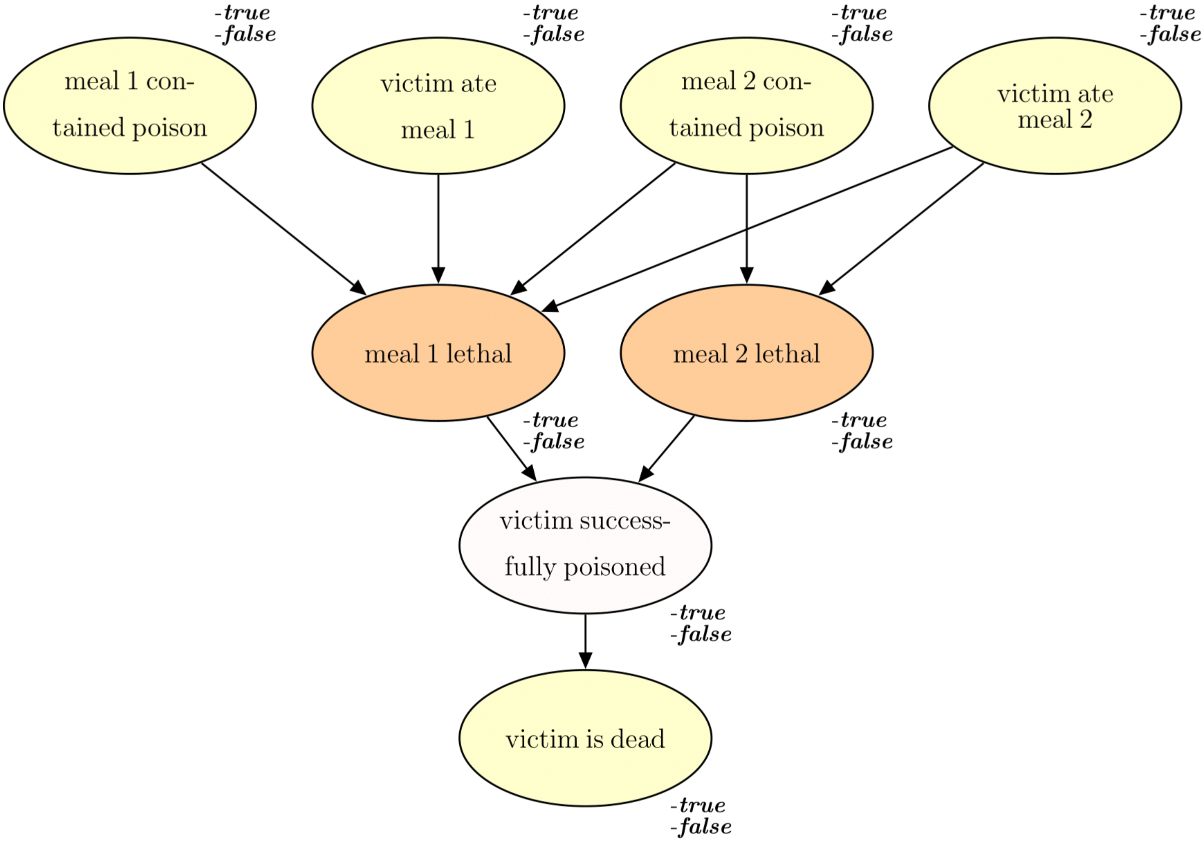

Figure 13 Poison network - additional uncertainties

Note that the Bayesian network structure in Figure 13 represents one of the many possible models for this problem. As previously discussed, if the focus was to determine whether the defendant intended to poison the victim (as opposed to simply determining the cause of death) one could include: the intent of the person that placed the poisons and whether the same person was responsible for the first and the second poison as nodes to the network.

ACCEPTED MANUSCRIPT

Again, a Bayesian network could serve as the model that presents the assumed relation between relevant hypotheses and evidence.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

5 Conclusions and recommendations

The arguments that have been used to support the idea that the various puzzles produce supposed probability paradoxes are based on the following fundamental misunderstandings:

1. That it is possible to evaluate the evidence in isolation without taking into account the impact of the evidence on other relevant hypotheses in the case.
2. That evidence that is useful for each of two contradictory hypotheses is not relevant. (This is false because it could be more useful for one hypothesis than the other.)
3. Speaking of “the LR” as if there is only one LR for each item of evidence.
4. Equating LR = 1 for a certain set of hypotheses with a claim that the evidence is irrelevant for the case as a whole.

According to Pardo (Pardo, 2013), an ideal methodology for handling evidence appropriately must satisfy: the “micro level”- that of individual evidence; the “macro” level- that of narrative or story; and the “integration constraint”- individual evidence must be integrated into their wider story context. Bayesian networks satisfy these constraints, and so provide an appropriate formal framework for use in a legal setting. We have shown that, by modelling the puzzles as Bayesian networks, the claimed probabilistic ‘paradoxes’ in each case are easily discredited. Moreover, when these models are used properly they can help prevent logical blunders commonly made when reasoning with evidence.

It is also desirable that a method for handling evidence is flexible- allowing one to try out different stories, to change assumptions, and to refine and develop a model for a given set of evidences. Bayesian networks provide this flexibility. Moreover, Bayesian networks are being increasingly used in practice to help forensic scientists assess the impact of their evidence - see, for example (Kokshoorn, Blankers, de Zoete, & Berger, 2017; Taroni, Aitken, Garbolino, & Biedermann, 2014; Taylor, Biedermann, Hicks, & Champod, 2018) - and to help legal practitioners understand the overall impact of combined evidence – see for example (de Zoete, Sjerps, & Meester, 2017; Edwards, 1991; Lagnado, Fenton, & Neil, 2013; Taylor et al., 2018).

ACCEPTED MANUSCRIPT

As with any methodology, Bayesian networks have not been perfected to the point where they can adequately model all legal situations. However, this need not deter us from attempting such a formal framework for evidence. Without such a framework, it is easy to ignore implicit assumptions, and we would have little basis beyond untutored intuition for combining and weighing multiple items of evidence such as we see in many, if not most, cases. Furthermore, by explicitly framing what evidence and which hypotheses are considered one does not have to speak of "the" LR in broad terms because the underlying assumptions for "their" LR are explicit.

Some legal professionals may feel discouraged from using any kind of probability theory in legal cases because they do not wish to "put a number" on doubt or belief. It is worthwhile recalling that Bayesian networks are useful primarily as models of how events relate to one another, rather than as a guilt-calculator, throwing out an infallible number for judgement. Also, this method is tractable, and accommodates uncertainty; it is unnecessary to commit to a single "point value" for a probability when this is not appropriate. Furthermore, if a line of legal reasoning does not make use of such a formal framework, this does not prevent the necessity of assumptions or banish uncertainty.

A particular advantage of using Bayesian nets is that they are visual, making this methodology more intuitive to non-mathematicians. Importantly, for any user of Bayesian nets, the process of building invites interrogation at every stage of construction, and assumptions at each step are more easily identified than with non-visual methods. They are useful as maps of how events are related; as maps of belief and doubt; and as a tool for considering a case fully, integrating story, real-world context, and evidence.

We have shown that by using a Bayesian network to structure these legal paradoxes evaluating the combined evidential value can be done effortlessly. Furthermore, when evidence is only indirectly relevant for the hypothesis of interest, i.e. when it is relevant for another, related, pair of hypotheses, a Bayesian network can be used to make this

ACCEPTED MANUSCRIPT

connection visual. Even in situations where exact probability assignments are difficult or even impossible to assign due to the nature of the evidence or a disagreement amongst the involved parties, the structured probability model does allow users to establish whether a piece of evidence is relevant regardless of the exact values. Most importantly, by disentangling the dependency relations between distinct hypotheses and pieces of evidence, it can be shown that common examples of probabilistic paradoxes in legal reasoning only exist due to the restricted view with which they are approached and not because of the underlying probabilistic concept of “relevant evidence”.

Acknowledgments

This work was supported by the ERC project BAYES_KNOWLEDGE (ERC-2013-AdG339182) and the Leverhulme Trust Grant RPG-2016-118 CAUSAL-DYNAMICS.

Declarations of interest: none

6 References

Achinstein, P. (2001). *The Book of Evidence*. Oxford University Press.

Agena Ltd. (2019). AgenaRisk. http://www.agenarisk.com. Retrieved from http://www.agenarisk.com

Aitken, C. G. G., Roberts, P., & Jackson, G. (2010). *Fundamentals of Probability and Statistical Evidence in Criminal Proceedings, Practitioner Guide No 1*. Royal Statistical Society’s Working Group on Statistics and the Law. Retrieved from http://www.rss.org.uk/uploadedfiles/userfiles/files/Aitken-Roberts-Jackson-

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Practitioner-Guide-1-WEB.pdf

Aitken, C. G. G., & Taroni, F. (2004). *Statistics and the evaluation of evidence for forensic scientists (2nd Edition)*. Chichester: John Wiley & Sons, Ltd.

Allen, R. J. (1993). Factual Ambiguity and a Theory of Evidence. *Northwestern University Law Review, 88*. Retrieved from http://heinonline.org/HOL/Page?handle=hein.journals/illlr88&id=620&div=&collection=journals

Allen, R. J., & Carriquiry, A. (1997). Factual Ambiguity and a Theory of Evidence Reconsidered: A Dialogue between a Statistician and a Law Professor. *Israel Law Review, 31*. Retrieved from http://heinonline.org/HOL/Page?handle=hein.journals/israel31&id=462&div=&collection=journals

Allen, R. J., Kunhs, R., Swift, E., Schawartz, D., & Pardo, M. S. (2011). *Evidence : text, problems, and cases*. Wolters Kluwer Law & Business.

Balding, D. J., & Steele, C. D. (2015). *Weight-of-evidence for Forensic DNA Profiles*. Wiley-Blackwell.

Cohen, L. J. (1977). *The Probable and the Provable*. Oxford: Clarendon Press.

Dawid, A. P. (1987). The Difficulty About Conjunction. *Journal of the Royal Statistical Society. Series D (The Statistician)*, *36*(2/3), 91–97.

de Zoete, J., & Sjerps, M. (2018). Combining multiple pieces of evidence using a lower bound for the LR. *Law, Probability & Risk*, *17*(2), 163–178.

de Zoete, J., Sjerps, M., & Meester, R. (2017). Evaluating evidence in linked crimes with multiple offenders. *Science & Justice*, *57*(3), 228–238. https://doi.org/10.1016/J.SCIJUS.2017.01.003

Edwards, W. (1991). Influence diagrams, Bayesian imperialism, and the Collins

ACCEPTED MANUSCRIPT

case: an appeal to reason. *Cardozo Law Review*, *13*, 1025–1079.

Engel, C. (2012). Neglect the Base Rate: It's the Law! *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2192423

Fenton, N. E., Berger, D., Lagnado, D. A., Neil, M., & Hsu, A. (2013). When 'neutral' evidence still has probative value (with implications from the Barry George Case). *Science and Justice*. https://doi.org/http://dx.doi.org/10.1016/j.scijus.2013.07.002

Fenton, N. E., Neil, M., & Berger, D. (2016). Bayes and the law. *Annual Review of Statistics and Its Application*, *3*(1), 51–77. https://doi.org/10.1146/annurev-statistics-041715-033428

Fenton, N. E., Neil, M., & Hsu, A. (2014). Calculating and understanding the value of any type of match evidence when there are potential testing errors. *Artificial Intelligence and Law*, *22*, 1–28. https://doi.org/http://dx.doi.org/10.1007/s10506-013-9147-x

Hastie, R. (2019). The case for relative plausibility theory: Promising, but insufficient. *The International Journal of Evidence & Proof*, 136571271881674. https://doi.org/10.1177/1365712718816749

Hojsgaard, S. (2012). Graphical Independence Networks with the gRain Package. *Journal of Statistical Software*, *46*(10), 1–26.

Hugin A/S. (2018). Hugin Expert. Retrieved from http://www.hugin.com

John William Strong, Kenneth S. Broun, George E. Dix, Edward J. Imwinkelried, & D. H. Kaye. (1999). *McCormick on Evidence, Fifth Edition, Vol. 1 (Practitioner's Treatise Series)* (5th ed.). West Group. Retrieved from https://www.amazon.co.uk/McCormick-Evidence-Practitioner-Practitioners-1999-01-30/dp/B01JXT9FZY

ACCEPTED MANUSCRIPT

Kokshoorn, B., Blankers, B. J., de Zoete, J., & Berger, C. E. (2017). Activity level DNA evidence evaluation: on propositions addressing the actor or the activity. *Forensic Science International*, *278*, 115–124.

Lagnado, D. A., Fenton, N. E., & Neil, M. (2013). Legal idioms: a framework for evidential reasoning. *Argument and Computation*, *4*(1), 46–63. https://doi.org/dx.doi.org/10.1080/19462166.2012.682656

Lempert, R. O. (1977). Modeling Relevance. *Michigan Law Review*, *75*(5/6), 1021. https://doi.org/10.2307/1288024

Lyon, T. D., & Koehler, J. J. (1996). The relevance ratio: Evaluating the probative value of expert testimony in child sexual abuse cases. *Cornell Law Review*, *82*, 43–78.

Pardo, M. S. (2013). The Nature and Purpose of Evidence Theory. *Vanderbilt Law Review*, *66*.

Park, R. C., Tillers, P., Moss, F. C., Risinger, D. M., Kaye, D. H., Allen, R. J., … Kirgis, P. F. (2010). Bayes Wars Redivivus — An Exchange. *International Commentary on Evidence*, *8*(1), 1–38.

Picinali, F. (2012). Structuring inferential reasoning in criminal fact finding: an analogical theory. *Law, Probability and Risk*, *11*(2–3), 197–223. https://doi.org/10.1093/lpr/mgs006

Redmayne, M. (2009). Exploring the Proof Paradoxes. *Legal Theory*, *14*, 281–309. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1324102

Roberts, P., & Zuckerman, A. A. S. (2010). *Criminal evidence*. Oxford University Press. Retrieved from https://global.oup.com/academic/product/criminal-evidence-9780199231645?cc=gb&lang=en&

Schwartz, D. S., & Sober, E. R. (2017). The Conjunction Problem and the Logic of

ACCEPTED MANUSCRIPT

Jury Findings. *William & Mary Law Review*, *59*. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2927252

Schweizer, M. (2013). The Law Doesn't Say Much About Base Rates. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2329387

Sullivan, S. P. (2016). A Likelihood Story: The Theory of Legal Fact-Finding. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2837155

Taroni, F., Aitken, C. G. G., Garbolino, P., & Biedermann, A. (2014). *Bayesian Networks and Probabilistic Inference in Forensic Science* (2nd ed.). Chichester, UK: John Wiley & Sons.

Taylor, D., Biedermann, A., Hicks, T., & Champod, C. (2018). A template for constructing Bayesian networks in forensic biology cases when considering activity level propositions. *Forensic Science International: Genetics*, *33*, 136–146. https://doi.org/10.1016/J.FSIGEN.2017.12.006

University of Pittsburg, D. S. L. (2018). GeNIe: Graphical Network Interface. Retrieved from http://genie.sis.pitt.edu/

Highlights

- A comprehensive review of common probabilistic paradoxes in legal reasoning.
- Probabilistic paradoxes like the *twins problem* are resolved using Bayesian Networks.
- The resulting Bayesian networks provide a powerful framework for legal reasoning.
- Also considered are the poison, the lottery and the abuse paradox.
- We also consider the typewriter, the food tray and the liberal candidates example.