Article

Taking the dialectical stance in reasoning with evidence and proof

Floris J Bex
Department of Information & Computing Sciences, Utrecht University, Utrecht, The Netherlands

Douglas N Walton
CRRAR, University of Windsor, Canada

Abstract

We present a computational argumentation approach that models legal reasoning with evidence and proof as dialectical rather than probabilistic. This hybrid approach of stories and arguments models the process of proof in a way that is compatible with Allen and Pardo's theory of relative plausibility by adding arguments that can be used to show how evidence can support or attack explanations. Using some legal cases as examples, we show how criteria for assessing explanations connect arguments and evidence to story schemes. We show how this hybrid dialectical approach avoids the main problem of the probabilistic approaches, namely that they require precise numbers to be applied in order to decide legal cases. We provide an alternative method that allows fact-finders to reason with evidence holistically and not in the item-by-item fashion proposed by the probabilistic account.

For quite some time now, Professors Michael Pardo and Ron Allen have argued for the theory that reasoning in trials is typically based on the comparative plausibility of competing evidence-based explanations.[1] In their 2019 article, which this text comments on, Allen and Pardo reply to some

---

1. Between them they have many publications but the two central ones are Pardo and Allen (2008) and the article this comment is replying to, Allen and Pardo (2019).

Corresponding author:
Floris J Bex, Utrecht University, 3584, Netherlands.
E-mail: f.j.bex@uu.nl

powerful criticisms that have been directed against their theory, mainly by those of a traditional probabilistic inclination, and in so doing have investigated some new cases in which burdens and standards of proof are employed. This commentary, while in some parts criticising Allen and Pardo's latest article, in general supports and, we argue, deepens their relative plausibility account.

Allen and Pardo's four main criticisms of the probabilistic account are as follows.

1. Numerical probabilities (possibly subjective) are needed, but in practice these are hard to obtain or estimate.
2. Fact-finders process and reason with evidence holistically and not in the item-by-item fashion proposed by the probabilistic account.
3. Probabilistic thresholds apply to the case as a whole, whereas legal doctrine says that the standard of proof applies to the individual elements of a case (the conjunction paradox).
4. The probabilistic approach does not take into account the alternative explanations advanced by the other side in the case.

Criticisms 2–4 are only applicable to the so-called 'naive Bayesian' probabilistic approach, in which it is assumed that the probabilities of the individual pieces of evidence are independent of one another. This naive approach is the approach that is often advocated in a legal setting, either by scholars or by the ways in which, for example, forensic experts report likelihood ratios or random match probabilities. Scientific consensus, at least in the more mathematically inclined sciences, is that the naive Bayesian approach is not suitable for complex legal cases. However, there exist principled and broadly accepted ways of considering the probabilities of individual elements (events, evidence) in a broader context, taking the dependencies between the probabilities of different elements into account. One of these is through the use of so-called Bayesian Networks (BNs), in which the dependencies between different probabilistic variables are expressed as a graph.[2] These BNs allow for holistic reasoning with probabilistic explanations in that it provides principled ways to combine the probabilities of individual elements of a case into a single probability that can be compared to the (single) probability of another probabilistic explanation.

This leaves criticism 1: where do the numbers come from? In any probabilistic account, naive or not, a number of probabilities or likelihood ratios have to be estimated, which is enormously complex, not to say impossible. One of the core problems here is that the fact-finders or decision-makers estimating probabilities should be both experts in probabilistic models as well as experts on the particular evidence and reasoning used in a case. As an example, consider the situation where the probability that a certain motive of a suspect led to them committing the crime has to be estimated. Such an estimate requires knowledge of which kinds of motives would typically lead to certain behaviour by persons such as the suspect, which lies within the purview of experts in criminology or behavioural psychology. But we cannot expect the same experts to construct a complex probabilistic model, taking into account the probabilistic influences of completely different types of probabilities, such as those about DNA evidence and witness testimony credibility. And even if with the help of numerous different experts we succeed in building a complex probabilistic model, we cannot expect decision-makers, such as judges and juries, to make a decision based on a model which they themselves do not understand.

The question is then whether any numerical or mathematical account based on probabilities or logic is a good interpretation of the reasoning in legal cases. Take, for example, the standard of beyond a reasonable doubt (BARD). According to *McCormick on Evidence*, 'The term reasonable doubt is almost incapable of any definition which will add much to what the words themselves imply' (Strong, 1992: 447). Courts have held that the legal concept of reasonable doubt itself needs no definition, the reason

---

2. BNs have been used to represent complex Wigmorean arguments, see Kadane and Schum (1996), as well as explanatory stories, see Vlek (2016).

being that any definition might have to be so subtle and technical that there would be dangers of misunderstandings if judges were to instruct juries with it. Allen and Pardo seem to agree with this view on evidence and proof: apart from general criteria for explanations (coherence, simplicity) which are not further explored or defined, they do not explain exactly what makes an explanation plausible, or more plausible than another explanation. This, we feel, makes their criticism of the precisely formally defined probabilistic account slightly unfair: by not committing to more explicit definitions of (relative) plausibility, Allen and Pardo do not expose their own theory to the kind of precise scrutiny to which they subject the probabilistic approach.

In this comment, we aim to give a more explicit and formal account of relative plausibility. As Tillers and Gottfried (2006) argued, even though there is a well-settled maxim supported by judicial wisdom that the beyond reasonable standard is not quantifiable by assigning probability values to it, it does not follow that this standard of doubt is not open to precise analysis based on a more formal model. Using Bex's hybrid theory (2011) as our starting point, we briefly discuss how the quality of explanations can be gauged and the how proof standards can be captured using the hybrid theory, and provide examples of both a criminal and a civil case.[3] Here, we take the *dialectical stance*,[4] viewing the process of proof as a dialectical process of critical discussion, asking critical questions and arguing for and against positions (van Eemeren and Grootendorst, 2004). This dialectical process is central in the study of argumentation, both informal (Walton, 1998) and formal (Dung, 1995), and has been applied to legal settings in many instances (cf. Verheij, 2003 and also Bex, 2011). Note that Pardo and Allen's relative plausibility approach is also inherently dialectical: it is possible to 'argue' for some position in a case by providing a particular alternative explanation. The probabilistic approach is harder to characterise as dialectical, as this approach has as its central tenet the probability of the truth of statements or events, whereas in dialectical approaches arguments for and against the possible probabilities of some statement or event being true are given. In other words, where probabilistic approaches rely on a correct and consistent probability distribution to determine the probability of some conclusion, dialectical approaches rely on inconsistent arguments and counterarguments to determine the correct probability distribution in a case (Prakken, 2018).

A hybrid theory of explanatory stories and arguments

In addition to the work by Pardo and Allen, there have been other authors that propose a theory of explanatory reasoning in the law (see Bex, 2011; Josephson, 2002; Pennington and Hastie, 1993; Schum, 2001; Wagenaar et al., 1993). One shortcoming of these theories of abductive explanation is that they for the most part refrain from giving exact, formal definitions of their central concepts. What is an explanation? How exactly can criteria such as 'consistency, coherence, fit with background knowledge, simplicity, absence of gaps, and the number of unlikely assumptions that need to be made' (see Allen and Pardo, 2019) be operationalised? The informal definitions of these concepts in the literature called for a more formal treatment of argument and explanation and this is why Bex proposed his *hybrid theory* (Bex, 2011).

The hybrid theory is a mixture of argument-based and story-based reasoning. In *argument-based reasoning*, arguments are constructed by performing consecutive reasoning steps from evidence to conclusion. Each of these reasoning steps has an underlying evidential generalisation of the form '*e* is evidence for *p*'. Reasoning with arguments is dialectical because not only arguments for a conclusion but also counterarguments are considered. *Story-based reasoning* involves constructing stories about what (might have) happened in a case that explain the evidence. Reasoning with stories can be

---

3. The discussion of proof standards and the examples have been adapted from Bex and Walton (2012).
4. A stance is the level of abstraction on which certain concepts, decisions or behaviour is judged. Cf Dennett's *intentional stance*, which views the behaviour of an entity is judged in terms of the mental properties (intentions) of the entity (Dennett, 1996).
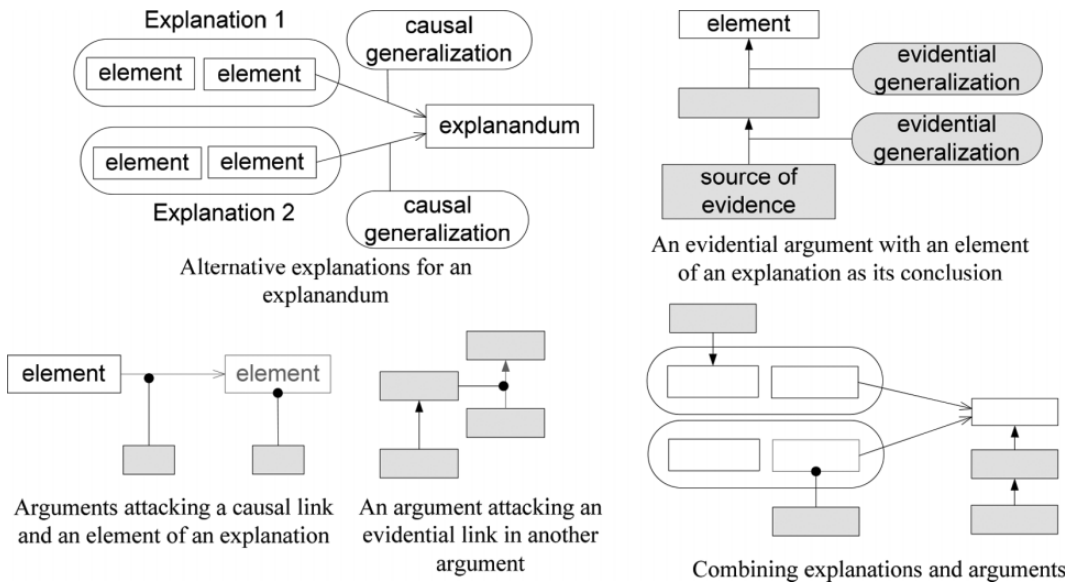
Figure 1. Visualising the hybrid theory.

characterised as causal reasoning: the relations between the events in a story and between the story and the evidence can be expressed as causal generalisations '*c* is a cause for *e*'. This approach is also dialectical: an explanation is defeasible, that is, it holds tentatively by placing a burden on an opponent to critically question it and to offer a possible alternative explanation.

The hybrid theory allows for both alternative explanations and arguments to reason about these explanations. For example, arguments can be used to support an explanation with evidence or to reason about the plausibility of the explanations. Figure 1 shows a visual overview of the hybrid theory. Here, ⟶ arrows denote causal links, ⟶ arrows denote evidential links and the ⟶● arrow denotes an attack of an argument (i.e. contradiction). Notice that some of the attacked elements or links are 'greyed out'. This represents the dialectical status of arguments: informally, elements or links can be either *justified*, which means that they are not defeated by other justified arguments, or *overruled* (grey in the picture), which means that they are defeated by other justified arguments.

Bex formulated several criteria that can be applied using is hybrid theory of stories and arguments (Bex, 2011: 94). Broadly speaking, these criteria fall into two categories:

1. whether the explanation conforms to the specific evidence in the case at hand, that is, how does the available evidence support and contradict the explanation, and which elements of the explanation are not supported by the evidence; and
2. whether the explanation conforms to general, common-sense knowledge of the world, that is, how plausible are the assumptions and generalisations the explanation makes and how complete, detailed and consistent is the explanation.

These criteria are can be defined in more detail using the machinery in the hybrid theory. The *evidential support* and *evidential contradiction* of an explanation are, respectively, the number of pieces of evidential data that support or contradict an explanation through arguments. Note that only arguments which are justified support or contradict an explanation: if an argument is overruled, the evidence does not support the explanation. Arguments can also be used to reason about the plausibility of an explanation, as the plausibility of elements of an explanation or the explanation as a whole can become the

subject of an argumentation process. Arguments about the plausibility of explanations are based on assumptions, as reasoning about plausibility is done using common-sense knowledge about how the world generally works. For plausibility, we do not just consider the individual elements of an explanation (events, causal relations), but also the story as a whole by seeing if the story fits a plausible *story scheme* or script (Schank and Abelson, 1977), a general sequence of events as we expect them to normally happen.

The hybrid theory uses *critical questions* that point to typical sources of doubt in a case (Bex and Verheij, 2012). Some of these questions concern the arguments in the case—these are the questions associated with argumentation schemes (Walton et al., 2008), stereotypical ways of reasoning in a case. For example, given an Argument from Witness Testimony, we can ask 'Does the witness accurately remember what happened?' and 'Does she have a reason to lie about what happened?' Other critical questions in the hybrid theory concern the causal coherence of the explanations. For example, we can ask what kind of general scheme the explanation adheres to—for example, if an explanation speaks of a restaurant visit which lasted only two minutes, this is not clear. Furthermore, we can ask whether the explanation is complete—for example, why does a restaurant story not mention any of the actors eating food? Finally, there are general questions that concern the case and comparison of explanations on a higher level. For example, we can ask whether the search for alternative explanations was thorough enough, how decisively some explanation surpasses the alternatives, or how good an explanation is independently of the alternatives. The latter two concern the quality of explanations, and we can use the criteria for explanations given above, taking into account the various standards of proof given by the law.

Standards of proof in the hybrid theory

In explanation-based accounts of judicial proof, such as the hybrid theory, different explanations have to be compared. The law further gives standards of proof, which can be used to indicate not only when one explanation is stronger or better than another but also by what margin they are better and how good they are in themselves. These standards of proof can be captured using the hybrid theory as follows.[5] An explanation meets the preponderance of evidence (PE) standard if it is better than each alternative explanation (i.e. the explanation is supported by more evidence, contradicted by less evidence and/or more plausible). An explanation meets the beyond a reasonable doubt standard if it is strong and much stronger than its competing explanations.[6] For the BARD standard, we can say that each competing explanation has to be very weak, so weak as to be highly implausible. A plausible explanation consistent with innocence creates a reasonable doubt, so for each competing explanation, either the evidential contradiction should be high and the plausibility should be low if the 'guilt' explanation is to meet the standard of proof. Note that the evidential support of the competing explanations is not tied to extra requirements, as these explanations only have to be *consistent* with the evidence.
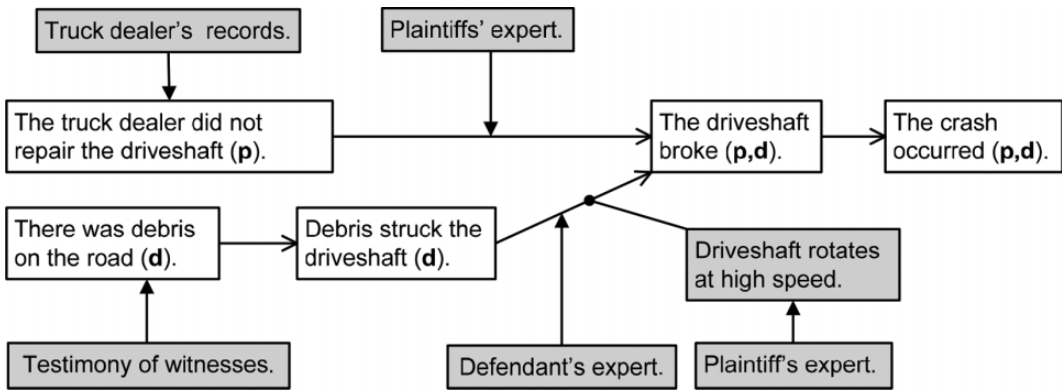
The case of *Anderson* v *Griffin*

In the case of *Anderson* v *Griffin*[7] (a civil case), the driveshaft broke on a tractor-trailer proceeding down an interstate highway disconnecting the brakes. At the same time, debris kicked up from the surface of the highway and struck a pickup truck behind the tractor-trailer. The pickup truck crashed into the tractor-trailer and the collision injured two people in a car behind the pickup truck. These two

5. Here we only discuss the PE and BARD standards. See Bex and Walton (2012) for the other standards. Our account is based on earlier work by Gordon and Walton, which defines standards of proof for purely argument-based approaches: Gordon and Walton (2009).

6. We do not say exactly what a 'strong' explanation is, as idea is not to precisely calculate the strength of explanations, but rather to give the decision-maker dialectical tools they can use to determine for themselves whether an explanation is strong enough.

7. 397 F. 3d 515, Court of Appeals, 7th Circuit, 2005.

**Figure 2.** Two explanatory stories in the *Anderson* v *Griffin* case.

people, who were injured, sued the truck dealer who was supposed to be responsible for the mainte-
nance of the trailer.

The plaintiffs' explanation for the crash was one of negligence (Figure 2): defendant (the truck
dealer) had failed to tighten the middle joint on the driveshaft. This caused the driveshaft to break,
setting in motion the accident. The explanation was supported by evidence from the truck dealer's record
stating that the repairman did not repair that joint, and an expert witness who states that the crash was
caused by the fact that defendant did not repair the driveshaft. Defendant gave an alternative explana-
tion, claiming that the cable was broken by debris that struck the driveshaft. He supported his claim that
there was debris on the road by testimony of witnesses. According to the PE standard, plaintiff still wins
at this point: plaintiffs supported their explanation with both the defendant's own records and an expert
witness, while defendant only supported his explanation with a witness testimony. Defendant therefore
added a new testimony by an expert supporting his own explanation, namely a claim that the cable was
broken by debris that struck the driveshaft. The plaintiffs' expert countered that a piece of road junk
would be highly unlikely to strike the driveshaft with enough force to break it, because of the speed at
which the driveshaft rotates (27 times a second). Depending on how we would resolve the competing
arguments about whether the debris could have caused the driveshaft to break, either the plaintiffs'
explanation is slightly better (if we decide to believe the plaintiffs' expert) or the explanations are
equally good (if we believe the defendant's expert). In the case the jury ruled for the defendant. For
some reason, they must have found that the attacking argument based on plaintiffs' expert was not
convincing enough and because the two explanations were then equally good, decided that the PE
standard had not been met by the plaintiffs.

The case of *Jackson* v *Virginia*

Our second example concerns a criminal case, *Jackson* v *Virginia*.[8] The case concerns the death of Mary
Cole, who had been a member of staff at the county jail where she had befriended James Jackson, an
inmate. After his release, Cole and Jackson stayed in contact. A few days before Mary Cole was found
dead, Cole and Jackson were seen drinking by two sheriffs at a diner. As the two were preparing to leave
the diner in Cole's car the sheriff testified that he had offered to keep Jackson's revolver until he sobered
up, but that Jackson had said this would be unnecessary since he and Cole were about to engage in sexual
activity. The same evening, Jackson drove from Virginia to North Carolina. A day and a half later,

8. 443 U.S. 307 (1979).
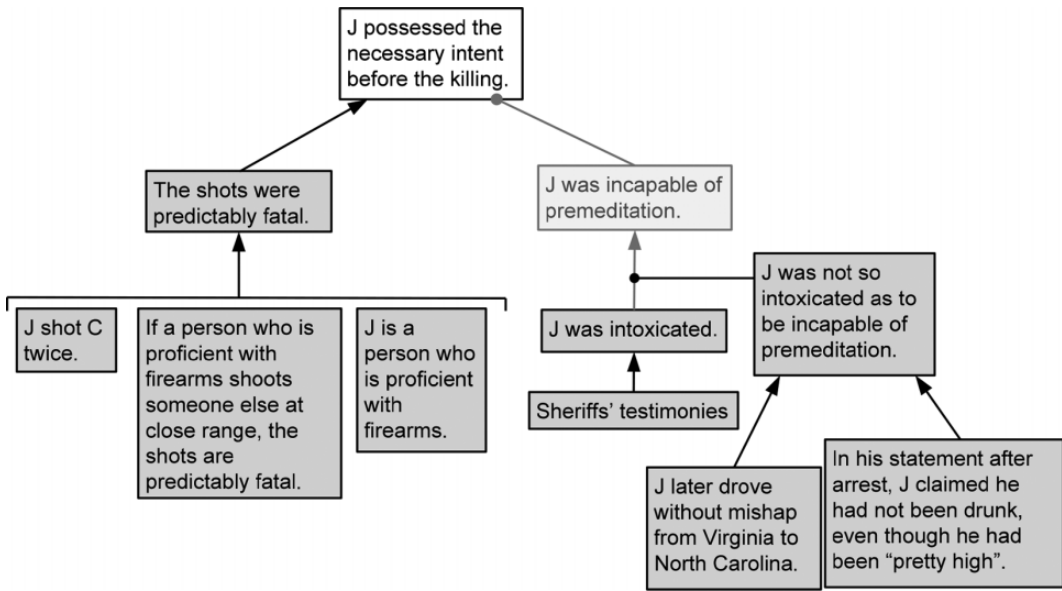
Figure 3. Prosecution arguments attacking Jackson's Story about No Premeditation.

Cole's body was found in a secluded parking lot, naked from the waist down, her slacks beneath her body; Jackson was arrested a few days later.

Jackson was convicted of first-degree murder, and appealed to the Supreme Court arguing that the beyond reasonable doubt standard had not been met, that is, it was questionable whether there was sufficient evidence to justify a rational trier of fact to find guilt beyond a reasonable doubt. One of Jackson's arguments was that there was insufficient evidence to support the finding that he had intended to kill Mary Cole.

The prosecution's explanation first recounted the events at the diner and then, without mentioning a specific motive, argued that Jackson intended to kill Cole, that he shot her with his revolver, killing her, and that he then drove to North Carolina. The story was supported by the sheriffs' statements of Jackson's behaviour at the diner, as well as expert medical evidence that Cole had been shot twice at close range with Jackson's revolver. Because Jackson further admitted he had shot Cole, the main factual dispute at the trial was whether there was sufficient evidence to prove Jackson's *intention* to kill Cole. Evidence that he had so intended was that Jackson had fired two shots at Cole at close range, shots that were predictably fatal given that he was a person experienced in the use of firearms. The left side of Figure 3 shows the evidential argument for the claim that Jackson intended to kill Cole.

The prosecution had thus met the defence's objections to the quality of the prosecution's explanation, and there was no evidential contradiction and no competing explanation. Jackson then attacked the prosecution's explanation using the sheriffs' testimonies to argue that he had been too intoxicated to form the specific intent necessary to make him guilty of the crime of first-degree murder. The prosecution countered this by arguing that Jackson was not so intoxicated as to be incapable of intent, based on the fact that Jackson drove without mishap from Virginia to North Carolina and on a statement Jackson made after his arrest, where he claimed he had not been drunk. This new argument then overrules Jackson's argument against the prosecution's explanation (see right side of Figure 3).

Jackson next move was to try to cast doubt on the prosecution's explanation by proposing a sufficiently plausible alternative explanation consistent with the evidence. He claimed that after they left the
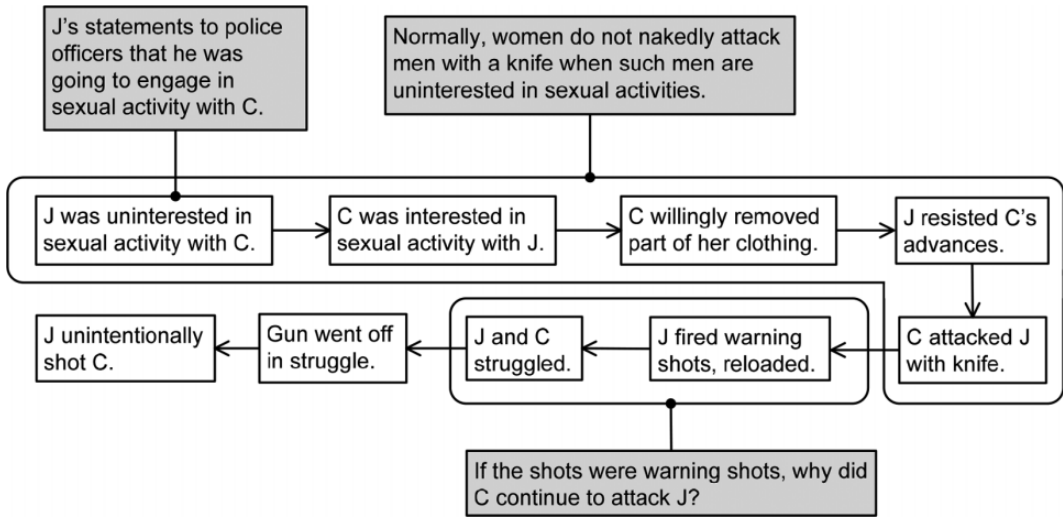
Figure 4. Jackson's story attacked by arguments.

diner, Cole had made sexual advances towards Jackson. When he had resisted her sexual advances, she had attacked him with a knife. He defended himself by firing warning shots into the ground, and then reloaded the weapon. Cole attempted to take the gun away from him and it went off during the struggle. He claimed that he had fled without seeking help for Cole because he was afraid. Later, during the trial he claimed that he had acted in self-defence. The prosecution countered Jackson's story in various ways. First, Jackson said to the sheriffs he was going engage in sexual activity with Cole but later supposedly resisted Cole's sexual advances, which is implausible. Furthermore, it was found implausible that Cole first willingly removed part of her clothing and then attacked him with a knife when he resisted her advances, and continued to attack even though he was armed with a loaded revolver that he had just demonstrated he knew how to use. Figure 4 shows these attacks on the plausibility of Jackson's story. The court found that the BARD standard had been met, because the prosecution's explanation was supported by evidence and not contradicted by justified arguments, and Jackson's story was deemed insufficiently strong to cast doubt on the conviction.

Concluding remarks

In this comment, we have presented a dialectical approach to evidence and proof compatible with Allen and Pardo's theory of relative plausibility. Our approach supports and further specifies Allen and Pardo's work. First, we further explore the dialectical principles underlying reasoning with and about explanations by adding arguments that can be used to reason about explanations, and critical questions that point to typical sources of doubt in a case. Second, we make explicit the criteria for assessing explanations by defining the *evidential support*, *evidential contradiction* and *plausibility* of an explanation given a number of arguments and story schemes in a case. Third, using these criteria we specify two oft-used standards of proof in the law.

Our approach more clearly defines the reasoning structures underlying the process of proof. At the same time, the approach does not suffer from what we feel is the main problem of the probabilistic approaches, namely that precise numbers are needed in order to reason and decide in legal cases. Put even more strongly, our approach shows that reasoning with evidence and proof is in fact not inherently probabilistic, but rather dialectical: while the probability or likelihood of the conclusion in a case plays an important role, an important aspect of the process of proof is that it can be characterised as a rational

procedure (Rescher, 1977). A decision is rational if it is in agreement with the knowledge that has been considered (or should have been considered) in a dialectical procedure. Because humans are limited by cognitive and practical constraints in their consideration of knowledge, especially in complex legal cases, determining a consistent probability distribution is next to impossible, but arguing to determine (or at least approach) the correct probability distribution in a case is not.

Author's Note

Dr. Bex is also affiliated to Tilburg Institute for Law, Technology and Society Tilburg University, Tilburg, The Netherlands.

Declaration of Conflicting Interests

Funding

References

Allen RJ and Pardo MS (2019) Relative plausibility and its critics. *International Journal of Evidence and Proof* 23(1-2): 5–59.

Bex FJ (2011) *Arguments, Stories and Criminal Evidence—A Formal Hybrid Theory*. Law and Philosophy Library, Vol 92. Dordrecht: Springer.

Bex F and Verheij B (2012) Solving a murder case by asking critical questions: An approach to fact-finding in terms of argumentation and story schemes. *Argumentation* 26(3): 325–353.

Bex F and Walton D (2012) Burdens and standards of proof for inference to the best explanation: Three case studies. *Law, Probability and Risk* 11(2–3): 113–133.

Dennett DC (1996) *The Intentional Stance*. Cambridge, MA: MIT Press.

Dung PM (1995) On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n–person games. *Artificial Intelligence* 77(2): 321–357.

Gordon TF and Walton D (2009) Proof burdens and standards. In Rahwan I and Simari G (eds) *Argumentation and Artificial Intelligence*. Berlin: Springer, 239–260.

Josephson JR (2002) On the proof dynamics of inference to the best explanation. In: MacCrimmon M and Tillers P (eds) (2002) *The Dynamics of Judicial Proof—Computation, Logic and Common Sense*. Berlin: Physica Verlag, 287–306.

Kadane JB and Schum DA (1996) *A Probabilistic Analysis of the Sacco and Vanzetti Evidence*. New York: John Wiley & Sons.

Pardo MS and Allen RJ (2008) Juridical proof and the best explanation. *Law and Philosophy* 27(3): 223–268.

Pennington N and Hastie R (1993) Reasoning in explanation-based decision making. *Cognition* 49(1–2): 123–163.

Prakken H (2018) A new use case for argumentation support tools: Supporting discussions of Bayesian analyses of complex criminal cases. *Artificial Intelligence and Law*. Epub ahead of print 9 October. https://doi.org/10.1007/s10506-018-9235-z.

Rescher N (1977) *Dialectics: A Controversy-Oriented Approach to the Theory of Knowledge*. Albany, NY: State University of New York Press.

Schank RC and Abelson RP (1977) *Scripts, Plans, Goals and Understanding: An Inquiry into Human Knowledge Structures*. Hillsdale, NJ: Lawrence Erlbaum.

Schum DA (2001) Species of abductive reasoning in fact investigation in law. *Cardozo Law Review* 22(5–6): 1645–1681.

Strong JW (1992) *McCormick on Evidence*. 4th ed. St Paul, MN: West Publishing Co.

Tillers P and Gottfried J (2006) Case comment—United States v. Copeland, 369 F. Supp. 2d 275 (E.D.N.Y. 2005): A collateral attack on the legal maxim that proof beyond a reasonable doubt is unquantifiable? *Law, Probability and Risk* 5(2): 135−157.

Van Eemeren FH and Grootendorst R (2004) *A Systematic Theory of Argumentation. The Pragma–dialectical Approach*. Cambridge: Cambridge University Press.

Verheij B (2003) Dialectical argumentation with argumentation schemes: An approach to legal logic. *Artificial Intelligence and Law* 11(1–2): 167–195.

Vlek CS (2016) *When stories and numbers meet in court: Constructing and explaining Bayesian networks for criminal cases with scenarios*. PhD Thesis, University of Groningen, The Netherlands.

Wagenaar WA, van Koppen PJ and Crombag HFM *Anchored Narratives: The Psychology of Criminal Evidence*, St. Martin's Press, New York

Walton DN (1998) *The New Dialectic*. Toronto: University of Toronto Press.

Walton D, Reed C and Macagno F (2008) *Argumentation Schemes*. Cambridge: Cambridge University Press.