Evidence & decision making in the law: theoretical, computational and empirical approaches

Marcello Di Bello[1] · Bart Verheij[2]

Decisions of judges, court experts and lay jurors play an important role in the fabric of society. Since the price of decisional errors in civil or criminal cases can be significant, it is paramount to employ good methods for assessing the evidence on which decisions are based. This special issue addresses questions at the intersection of evidence assessment in court and legal decision-making. The special issue contains the postproceedings of the workshop 'Evidence & Decision Making in the Law: Theoretical, Computational and Empirical Approaches' that was held in conjunction with the 16th International Conference on AI and Law and took place on June 16th, 2017 at King's College London (https://icail2017evidencedecision.wordpress.com).[1] The workshop aimed to foster an interdisciplinary debate among researchers in AI & Law working on legal reasoning and argumentation theory, legal scholars, philosophers and empirically minded researchers. For some general references on the themes discussed during the workshop, we refer the reader to the list at the end of this introduction.

Below we provide a summary of the contributions in the special issue.

The paper 'Normative decision analysis in forensic science' by Alex Biedermann, Silvia Bozza and Franco Taroni shows how statistical decision theory can be fruitfully applied—as an analytical and a normative tool—to the decision problem that forensic experts routinely face. For example, when a forensic expert has compared two fingerprints—one print associated with the crime sample and the other associated with the suspect—the expert should decide whether the prints come from the same source or from different sources. The expert may decide to say that the prints

---

[1] The workshop was a successor of two earlier events, the symposium 'Trial With and Without Mathematics', which took place at Stanford Law School on May 30th, 2014 (https://law.stanford.edu/event/trial-with-and-without-mathematics-legal-philosophical-and-computational-perspectives), and the ICAIL 2015 workshop 'Studying Evidence in the Law', which took place at University of San Diego on June 12th, 2015 (https://icail2015evidence.wordpress.com).

✉ Marcello Di Bello
marcello.dibello@lehman.cuny.edu

[1] Department of Philosophy, Lehman College, City University of New York, New York, USA

[2] Bernoulli Institute for Mathematics, Computer Science and Artificial Intelligence, University of Groningen, Groningen, The Netherlands

🖄 Springer

come from the same sources, while in fact they do not (a false positive identification), or she may decide to say that they do not come from the same source, while in fact they do (a false negative identification). How should the expert decide in light of this uncertainty? The authors argue that this decision problem can be analyzed in terms of two ingredients: (1) a probabilistic assessment of the strength of the evidence for and against an identification; (2) an estimate of the relative losses that would result from a false negative and a false positive identification. These two ingredients can be combined in the formula for expected utility maximization (or expected loss minimization). Although statistical decision theory has been around for a while—since at least the work of de Finetti and Savage—and has been very influential in many disciplines, it is not the mainstream theoretical framework among forensic scientists. Biederman, Bozza and Taroni argue that forensic science can benefit, both conceptually and in practice, from taking more seriously the analytical insights of statistical decision theory.

The paper 'A new use case for argumentation support tools: supporting discussions of Bayesian analyses of complex criminal cases' by Henry Prakken also examines the role of experts in court. The focus in this paper is with how experts should assess the strength of the evidence presented in court. Some scholars believe that Bayesian, probabilistic methods are well suited for this task; others disagree, offering alternative, non-probabilistic methods. Although Prakken recognizes that probabilistic methods—especially because of the success of Bayesian networks—are gaining momentum in the courtroom, he remains neutral on their relative merits. Instead, he defends the claim that if one were to assess the strength of the evidence in a case by means of Bayesian, probabilistic methods, this approach would still need supplementation. Experts who use Bayesian methods will make modeling choices of various kinds, and these choices may be contested by other experts. This means that argumentation theory—a theory concerned with modelling the structure of reasons pros and cons—has a crucial role to play. Argumentation theory can be used as an 'add-on' to Bayesian, probabilistic analyses of the evidence. Prakken illustrates and defends this claim by examining two recent Dutch criminal cases for which he was appointed as an expert. While most of the paper is analytical and theoretical, Prakken also formulates requirements for an argumentation-based support system that could be added on top of a Bayesian network system for the assessment of the evidence in court cases.

The paper 'Group-to-individual (G2i) inferences: challenges in modeling how the U.S. court system uses brain data' by Valerie Gray Hardcastle is more empirically oriented. It examines how neuroscientific brain data influence decisions by judges in criminal cases, specifically how group-level brain data are used to make inferences about individuals (so-called Group-to-Individual, G2i, inferences). Hardcastle and her collaborators analyzed many recent appellate criminal cases that referenced brain data and G2i inferences. They concluded that judges assign culpability in ways that often depart from what our best science about human decision-making would recommend. They also found that brain data have an ambiguous and context-dependent effect on legal decision-making. These findings pose a challenge—Hardcastle argues—for formal and computational models of legal decision-making.

These models, on the one hand, should accommodate the nuances of how brain data influence decisions by judges, but they should also not merely reflect and repeat the biases that often inform these decisions. The challenge here is to strike the right balance between descriptive and normative adequacy.

Another challenge for formal and computational models of legal reasoning comes from questions of causation and legal responsibility. In the paper 'Arguing about causes in law: a semi-formal framework for causal arguments', Rūta Liepiņa, Giovanni Sartor and Adam Wyner show how 'causation talk' can be incorporated in an argumentation-based framework. Their concern is not so much with designing the right theory of legal causation and responsibility, but with how disputes about causation can be adequately modeled in a formal, argumentation-based framework. As the authors note, the process of modeling highlights a tension about the role of causation in law. Sometimes causation is understood as an evidentiary question (i.e., did X cause Y?) and sometimes it is understood as a policy question (should a certain relationship between X and Y be considered as a causal relationship that implies an attribution of responsibility?). Building on theoretical perspectives on causation—but-for, necessary element of a sufficient set, actual causation—the authors present a semi-formal perspective on causal argumentation and illustrate it in a vaccine injury case.

The paper 'Assessment criteria or standards of proof? An effort in clarification' by Giovanni Tuzet examines the distinction between criteria and standards of decision. The former are guidelines for how to assess the evidence; how much weight a piece of evidence should be given compared to others; how to respond to conflicting pieces of evidence; etc. Standards, on the other hand, are rules of decision. They determine what one should decide in light of the evidence available. As Tuzet shows, some legal systems, predominantly those in continental Europe, are explicit about evidence assessment criteria but are silent about standards of proof. Other legal systems, predominantly those in the common law tradition, are explicit about standards of proof but silent about evidence assessment criteria. One might think that assessment criteria and standards of decision are two faces of the same coin, and thus one of the two can be dispensed with. Tuzet, however, argues that criteria and standards are not functionally equivalent. They are both necessary ingredients for the correct functioning of a legal system. So, whenever a legal system is not explicit about one or the other, Tuzet argues that legal practice will implicitly shape the system's conception of the missing standard or criterion.

The paper 'Proof beyond a context-relevant doubt. A structural analysis of the standard of proof in criminal adjudication' by Kyriakos N. Kotsoglou offers a novel answer to the question of how we should understand the standard 'proof beyond a reasonable doubt'. The paper begins with a discussion of the skeptical challenge in epistemology. No matter one's evidence, the skeptic can always raise doubts and questions one's claims to knowledge. Following Wittgenstein's insights in 'On Certainty', the author argues that the skeptic can raise doubts only if they are justified. But when is a doubt justified? Kotsoglou's answer is complex, but ultimately hinges on the claim that the relevance of a doubt depends on contextual factors. It is a matter of practice, not of mere theoretical analysis. This framework—Kotsoglou

argues—can be used to understand what it means to establish guilt beyond a reasonable doubt. It does not mean to answer all doubts, nor does it mean to establish guilt to a certain degree of probability, nor does it mean to answer some doubts and regard others as unworthy of consideration. What Kotsoglou aims to offer is a structural analysis of the proof standard, one rooted in our epistemic practices, not a conceptual clarification of what the proof standard requires in the abstract.

The paper 'A system of communication rules for justifying and explaining beliefs about facts in civil trials' by João Marques Martins defends a dialectical and communicative theory of how beliefs about facts should be justified in civil trials. The paper presents formal rules that judges can deploy as they justify their reasoning. These rules are similar, to some extent, to the rules of probability theory, but they have the advantage—according to the author—of being simpler and easier to apply. Some of these rules require the judge to assign a degree of support to hypotheses put forward by the parties and to make updates as new evidence is presented, in a manner similar to Bayesian conditionalization. For the purpose of illustration, these formal rules are applied to a recent tort case discussed by a Portuguese appellate court.

The last paper in the special issue—'Interactive Virtue and Vice in Systems of Arguments: A Logocratic Analysis' by Scott Brewer—tackles fundamental questions about the nature of argument and evidence. The paper defends a so-called logocratic theory of evidence and argument. This theory conceives of arguments as playing different functions, in terms of their internal logical cogency (or epistemic strength), dialectical strength and rhetorical persuasion. The conception of evidence in the paper is closely intertwined with the notion of argument. What arguments do is to provide evidence and, conversely, what evidence does is to make arguments possible. Thus, different types of arguments give rise to different types of evidence—deductive, inductive, analogical and abductive evidence. The paper also offers a novel analysis of abduction in general and legal abduction in particular, and it articulates the idea of dynamic interactive virtue—that is, the strength of a system of arguments considered as a whole is a function of the strength of the component arguments that make up the system of arguments. The paper then illustrates these theoretical concepts—especially the notion of a system of arguments—by analyzing a court case in US contract law.

We are grateful to our submitters, speakers, participants, reviewers and the conference organisers, without whose support the workshop would not have been possible.

References

Anderson, T., D. Schum, and W. Twining. 2005. *Analysis of Evidence*, 2nd ed. Cambridge: Cambridge University Press.

Dawid, A.P., W. Twining, and M. Vasiliki (eds.). 2011. *Evidence, inference and enquiry*. Oxford: Oxford University Press.

Di Bello, M., and Verheij, B. 2018. Evidential Reasoning. In *Handbook of Legal Reasoning and Argumentation* (eds. Bongiovanni, G., Postema, G., Rotolo, A., Sartor, G., Valentini, C., & Walton, D.), 447–493. Dordrecht: Springer.

Fenton, N.E., and M.D. Neil. 2013. *Risk assessment and decision analysis with Bayesian networks*. Boca Raton, FL: CRC Press.

Hacking, I. 2001. *An introduction to probability and inductive logic*. Cambridge: Cambridge University Press.

Jensen, F.V., and T.D. Nielsen. 2007. *Bayesian networks and decision graphs*. Berlin: Springer.

Kaptein, H., H. Prakken, and B. Verheij (eds.). 2009. *Legal evidence and proof: statistics, stories, logic* (Applied legal philosophy series). Farnham: Ashgate.

Schum, D.A., and S. Starace. 2001. *The evidential foundations of probabilistic reasoning*. Evanston, Il.: Northwestern University Press.

Taroni, F., A. Biedermann, S. Bozza, P. Garbolino, and C. Aitken. 2014. Statistics in practice. In *Bayesian networks for probabilistic inference and decision analysis in forensic science*, 2nd ed. Chichester: Wiley