

Boosting Legal Probabilism

bayesian network methods for narrations in legal fact-finding

Marcello Di Bello and Rafal Urbaniak

1 The Book

1.1 Brief Description

Can the evidence presented at trial be examined, weighed and assessed using probability theory? Can legal decision-making and standards of proof such as ‘preponderance of the evidence’ or ‘proof beyond a reasonable doubt’ be defined using the language of probability? Does the deployment of probability theory in assessing evidence and making decisions improve the accuracy and fairness of legal decision-making? Over the last fifty years, these questions have been debated in the literature in philosophy, law, forensic science and artificial intelligence. Legal probabilism is a research program that attempts to demonstrate that the answers to these questions should be, by and large, affirmative. Legal probabilists have made considerable progress, but also faced robust skepticism by legal theorists and philosophers. Our book ‘Boosting Legal Probabilism’ examines the most important objections to this research program and articulates a version of legal probabilism that is able to address many, if not all, of these objections. While the approach is normative, a strong emphasis is placed on being informed by the descriptive perspective—actual practice (1) is a source of problems and tasks that need to be addressed, (2) can inspire the formulation of various norms, (3) is useful in philosophical evaluation of various norms if their informal counterparts have already been discussed by the practitioners.

Added this
normative
sure if it s
here or la
look.

We begin with the simple version of the theory. It comprises two key elements: first, Bayes’ theorem for assessing the strength of the evidence, and second, probability thresholds as decision criteria. This simple version of legal probabilism is promising in many ways, but also falls prey to several difficulties, including the problem of conjunction, puzzles of naked statistical evidence, and the problem of priors. Some legal probabilists have attempted to dismiss these difficulties or downplay their significance. We confront them at face value and show that they cannot be addressed within the confines of simple legal probabilism. We then develop a more sophisticated theory, what we call legal probabilism 1.02, which takes advantage of Bayesian networks and other seminal ideas in the literature in forensic science and artificial intelligence. ‘Boosting Legal Probabilism’ articulates the first comprehensive philosophical analysis of whether—and if so, to what extent—legal probabilism 1.02 can overcome the limitations of simple legal probabilism. We also show that the more sophisticated version rivals in explanatory power two competing accounts of judicial fact-finding: argumentation theory and relative plausibility. To add precision to the claims made in the book, the analytic argument is supplemented by computer simulations and R code implementation.

The key positive contributions of the book are:

- A systematic introduction to the state of the art.
- A clear explanation of how the developments lead to the incorporation of insights from the narrative approach to legal fact-finding.
- The development of a rich formal framework based on Bayesian networks, meant to capture various features usually associated with narrations (such as coherence, completeness, or coverage).
- An extensive discussion of existing methods of handling multiple Bayesian methods and a proposal of improving on them for applications to legal fact-finding.
- An explication of various notions formulated within the argumentation theory in terms of the newly developed framework.
- An explanation of how the framework handles the philosophical objections to legal probabilism.
- A discussion connecting the framework to decision theory.
- An argument to the effect that when it comes to deployment of the framework, the tension accuracy and fairness is only apparent.
- A comparison of the approach to other philosophical accounts of legal fact-finding.

‘Boosting Legal Probabilism’ is aimed at philosophers with an interest in legal epistemology and epistemology more generally. Many of the difficulties of legal probabilism resemble difficulties faced by Bayesianism in epistemology. The book will also be of interest to legal scholars who have championed applications of probability theory to evidence law as well as scholars who have resisted this trend. Another target audience includes computer scientists and psychologists interested in studying evidential reasoning and decision-making under uncertainty. Besides contributing to the literature about legal probabilism, the book aims to introduce unfamiliar readers to the rich interdisciplinary debate on the topic, often scattered throughout journals and books in philosophy, law, computer science, forensic science and psychology. ‘Boosting Legal Probabilism’ can also be used by legal practitioners and reformers who aim to strengthen the protection for defendants against wrongful convictions, improve the accuracy of trial decisions and promote a fairer criminal justice system. So the book is aimed at scholars, advanced undergraduates, practitioners and reformers. Some chapters present original research and require technical background in probability theory. Others are introductory, suitable for an advanced undergraduate course.

1.2 Outline

The book is divided into three chapter clusters. Part I covers the current state of legal probabilism and the challenges it has faced. Part II introduces and motivates the philosophical moves and formal choices that we will make later on, by a focused discussion of the key formal tools used by legal probabilism so far and the issues that come up with their deployment in legal fact-finding. In Part III we develop our own approach and motivate it from the epistemological perspective. Part IV is devoted to a decision-theoretic perspective on our approach. Part V focuses on accuracy and fairness considerations, as pertinent to our approach.

Part I - Legal Probabilism and Its Foes

The first part of the book will instill interest in legal probabilism among unfamiliar readers and refresh seasoned readers about the main points of contention. **Chapter 1** outlines simple legal probabilism which comprises a familiar repertoire: Bayes' theorem, likelihood ratios, probability thresholds, expected utility maximization. This repertoire has proven useful in several ways, especially in the assessment of explicitly quantitative evidence such as DNA matches and other expert evidence. At the same time, as **Chapter 2** shows, simple legal probabilism is liable to a host of conceptual difficulties: the conjunction problem, the problem of priors, and the paradoxes of naked statistical evidence. These difficulties are well-known. Others are less familiar: the problem of complexity, soft variables, the problem of unanticipated possibilities, and the difficulty with corroboration.

The first two chapters provide the essential background for a deeper examination of legal probabilism and the development of its more sophisticated version. The remaining parts of the book cover two distinct topics: evidence assessment (Part II and Part III, Chapters 3 through 10) and decision-making (Part IV and Part V, Chapters 11 through 17). This distinction reflects the fact that legal probabilism is both a theory of evidence assessment (or evidence evaluation, evidence weighing) as well as a theory of decision-making at trial. These two topics are obviously intertwined in important ways, but are best kept separate for analytical clarity.

Part II - Evidence Assessment First Pass

The second part of the book discusses in great detail two formal tools that are essential for the legal probabilist: Bayes' theorem and likelihood ratios. We focus in particular on how these tools can help to assess, weigh and evaluate evidence at trial as well as what their limitations are.

Chapter 3 begins with Bayes' theorem and surveys many of its applications, for example, as a tool to avoid reasoning fallacies such as the prosecutor's fallacy and the base rate fallacy. At the same time, its applications are also limited. As discussed in **Chapter 4**, court cases often require fact-finders to weigh several pieces of evidence, sometimes conflicting and susceptible to different interpretations. The hypotheses that the fact-finders are asked to evaluate in light of the evidence are structured stories or explanations constituted by several sub-propositions. This level of complexity can hardly be modeled by successive discrete applications of Bayes' theorem. A more sophisticated machinery for evidence assessment is needed.

Chapter 5 describes a formal tool distinct from Bayes' theorem which many legal probabilists have found useful: likelihood ratios. Bayes' theorem requires one to assess the prior probabilities of the hypothesis of interest. The problem is that assessing prior probabilities is notoriously difficult. Likelihood ratios, instead, offer a way to evaluate the evidence presented at trial without the need of assessing prior probabilities. We illustrate the applications of likelihood ratios focusing on the debate about cold-hit DNA matches and the impact of false positive probability. The chapter also examines the weaknesses of this approach. The choice of the competing hypotheses compared in the likelihood ratio is often a source of confusion, manipulation and subjective judgment. Another problem is that, like simple applications of Bayes' theorem, likelihood ratios are still unable to model complex bodies of evidence. A further problem, as critics have alleged, is that likelihood ratios face difficulties in modeling the notion of evidential relevance in certain cases.

Chapters 3 and 4 show that we need to move past simple legal probabilism. Chapter 5 shows that likelihood ratios, while useful in many ways, are still an unsatisfactory approach overall. The journey toward legal probabilism 1.02 is accomplished in Part III, Chapters 6 through 10.

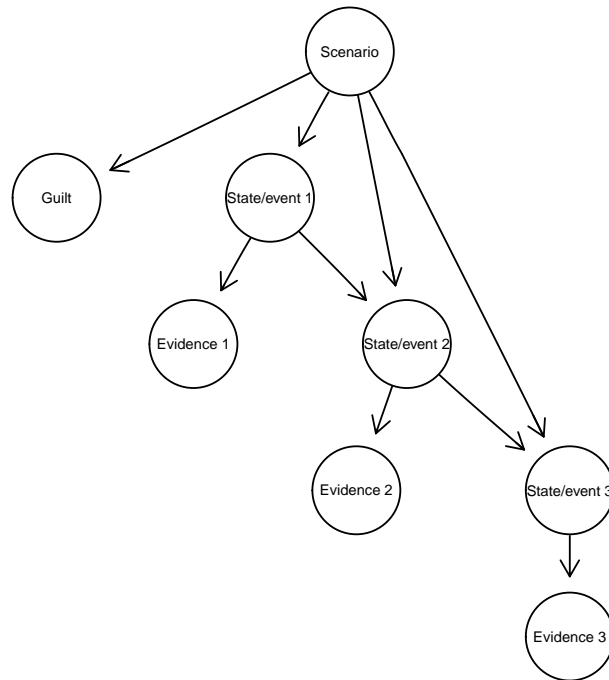
Part III - Evidence Assessment More and Better

The analysis in Part III is informed by the following working hypothesis. An accusation of liability must be substantiated by providing a well-specified account—a story or narrative—of the alleged illegal act committed by the defendant. This account consists of several moving parts, each supported by different items of evidence. The defense may respond by attacking the supporting evidence, the internal consistency of the account, or by offering an alternative account. It is this complex dynamics that we aim to formalize in the following chapters.

We focus on two aspects of the evaluation of trial evidence which simple legal probabilism is unable to model. First, judges and jurors often think holistically about the evidence, say in terms of coherent stories or explanations, without assessing the evidence by discrete applications of Bayes' theorem or likelihood ratios. Second, different pieces of evidence interact in complex relationships. They undercut, rebut or corroborate one another. We show that Bayesian networks constitute the

formal machinery necessary for developing a more sophisticated legal probabilism that is able to capture these phenomena. The coherence of a story as well as conflicts between pieces of evidence can be modeled formally by corresponding properties of, operations on, and relations between Bayesian networks.

Chapter 6 offers a crash course on Bayesian networks with a focus on the assessment of legal evidence. A Bayesian network comprises a directed acyclic graph (called a DAG) that represents relations of dependence between variables, along with conditional probability tables corresponding to these relations. In the last decade, the literature in artificial intelligence and forensic science has made significant progress in modeling holistic notions such as the coherence of a story and argument-based notions such as conflicts between pieces of evidence. We survey this literature. Charlotte Vlek, together with Bart Verheij and Henry Prakken, proposed to model the coherence of a story by adding a node in the Bayesian network, call it a ‘story node.’ The story node has other nodes as its children corresponding to the events that make up the story. In turn, these events are linked to their supporting evidence. An example of a Bayesian network with a story node (or scenario node to use Vlek’s terminology) is depicted below:



Since the story node unifies the different parts of a story, changes in the probabilities of these parts can be used to model the notion of coherence. The stronger the (positive) probabilistic dependence between the different parts, the more coherent the story. To model conflicts of evidence, a Bayesian network can be built that comprises two competing stories, say, one story put forward by the prosecution and another by the defense, each supported by their own evidence. The network would specify that these stories are incompatible and cannot be true concurrently. Another approach to model conflicting evidence and competing stories was developed by Norman Fenton and his research group. Separate stories are represented by separate Bayesian networks, and Bayesian model comparison is then used for assessing the comparative evidential support of the competing stories.

Chapter 7 assesses Vlek’s story node approach as an account of coherence. The critical argument is followed by a positive proposal. We show that adding a story node by fiat—without any good reason for supposing that the different parts of the story are connected other than being part of one story—introduces unnecessary probabilistic dependencies between the elements of a story. In addition, the story node approach is simplistic as an account of coherence and fails to engage with the rich philosophical literature on the topic. After the critical argument, the chapter articulates a more adequate probabilistic account of coherence. We define a formal notion of ‘story coherence’ that reflects properties of the Bayesian network used to model the evidence. We show that our formal notion of coherence that incorporates background causal knowledge addresses the objections against probabilistic accounts of coherence in the philosophical literature.

Chapter 8 focuses on conflicts between pieces of evidence. Neither Vlek’s story node approach nor Fenton’s model comparison approach adequately capture how pieces of evidence and competing stories may conflict with one another. The complex adversarial dialectic that takes place in a trial cannot be modeled by simply averaging different Bayesian networks (Fenton) or postulating relationships of incompatibility between different story nodes (Vlek). We need an account of more fine-grained notions, such as undercutting and rebutting evidence, and more generally we need an account of how cross-examination operates at trial. What cross-examination often accomplishes is not so much the creation of an alternative story, but the reinterpretation of an existing story by supplying additional information. We show that this process of re-interpretation can be represented formally as the refinement of an existing Bayesian network. Conflicts between pieces of evidence such as

added bit
ground ca
knowledg
a look

undercutting and rebutting can be modeled by drawing additional arrows between evidence nodes and hypothesis nodes.

The reverse of the phenomenon of conflicting evidence is that of converging evidence, in particular, the fact that one piece of evidence corroborates another. Corroboration has been the focus of extensive scholarly debate often independently of the debates within legal probabilism. **Chapter 9** surveys the literature on corroboration and the main difficulties that have been leveled against proposed probabilistic accounts. The chapter then formalizes a notion of corroboration using Bayesian networks that overcomes most of the difficulties of existing accounts.

Chapter 10 draws some general morals after comparing legal probabilism 1.02—as formulated in the previous two chapters—to other accounts of judicial fact-finding, in particular, argumentation theory and relative plausibility. Argumentation theory is well suited to model conflicts between evidence, but cannot easily model the fact that evidence may conflict more or less strongly with other evidence. Unlike argumentation theory, legal probabilism 1.02 offers an account of evidential support, conflict and convergence that captures how these relations come in degrees of strength. The other competing theory we consider, relative plausibility, is often criticized because the defense lawyer need not present a full-fledged alternative story. Without settling this controversy, we note that legal probabilism 1.02 is flexible enough to model competing stories (in agreement with relative plausibility) or model conflicts without the need to construct a full-fledged alternative story (as critics of relative plausibility prefer).

Legal probabilism 1.02 can still be challenged because of its questionable empirical adequacy since judges and jurors hardly follow probability theory. Nevertheless, we emphasize how legal probabilism 1.02 offers a richer account of evidential support beyond what critics have recognized. Susan Haack, for example, criticized legal probabilism for its monodimensional account of evidential support. This is true of simple legal probabilism in which evidential support is modeled by the posterior probability of a hypothesis given the evidence or the likelihood ratio. Legal probabilism 1.02, however, offers a richer account. In it, evidential support also depends on the degree of specificity, completeness, and coherence of the story put forward, and the extent to which the supporting evidence withstands objections. These features can serve to formulate decision criteria that are not unidimensional thresholds. Criteria for trial decisions are discussed more extensively in the next part of the book.

Part IV - Decision-making

The third part of the book examines trial decision-making, specifically, to what extent standards of proof such as ‘preponderance of the evidence’ and ‘proof beyond a reasonable doubt’ can be understood through the lenses of probability theory. There has been a spur of research arguing that legal probabilism is unfit to model standards of proof. But this research often holds a narrow view of legal probabilism. We show that the version formulated in the previous chapters of the book provides a more promising framework for theorizing about standards of proof.

Chapter 11 examines different strategies for theorizing about standards of proof using probabilistic language. We begin with the most natural decision criterion, a probability threshold whose stringency is determined by expected utility maximization. This account falls prey to well-known objections, most notably, the puzzles of naked statistical evidence and the difficulty with conjunction (discussed in greater detail in later chapters). We then turn to alternative accounts, in particular, the comparative strategy (Cheng) and the likelihood ratio strategy (Dawid, Kaplow, Sullivan). Finally, we present our own proposal. That is, the standard of proof is a function of several criteria: the probability of liability, the specificity and coherence of the accusatory story, the comprehensiveness of the supporting evidence and its ability to withstand objections. We emphasize that these criteria—probability, specificity, etc.—can be modeled using Bayesian networks. So our proposal lies within the confines of legal probabilism, though not the narrow version its critics have in mind. In the following chapters, we illustrate the theoretical payoffs that come from endorsing our proposal.

Chapter 12 tackles the puzzles of naked statistical evidence. This is a topic of enormous scholarly attention in the recent philosophical literature. Our solution rests on two premises. First, an accusation of liability should be substantiated by a well-specified account—or story, narrative—whose moving parts are each supported by adequate evidence. In cases of naked statistical evidence, the probability of liability is high, but the specificity of the accompanying narrative is suspiciously low. The second premise is that the supporting evidence should typically be ‘causally grounded.’ This grounding contributes to a well-specified story and is achieved during cross-examination by eliciting additional information about the relation between the evidence and the alleged facts—for example, information about the visibility conditions; the academic credential of an expert witness; the chain of custody of a document. Interestingly, naked statistical evidence blocks cross-examination because no undercutting evidence can in principle be brought against it.

Chapter 13 tackles another central problem for legal probabilism, the difficulty with conjunction. We first provide a detailed argument for why previous attempts in the literature on legal probabilism have failed, focusing on the likelihood ratio approach by Dawid and the comparative strategy by Cheng. Our proposal follows the holistic approach by Allen and Pardo. We show how legal probabilism can address the difficulty with conjunction so long as it can account for holistic notions such as coherence and specificity. As noted already, our working hypothesis is that the prosecution (or the plaintiff in a civil trial) should aim to establish a well-specified accusatory narrative whose moving parts are supported by adequate evidence. Once the prosecution has accomplished that—and its case withstands criticism—each element of the accusation is established to the required standard if and only if the conjunction of these elements (that is, the story as a whole) is established to the required standard.

Chapter 14 compares our probabilistic account of decision-making and standards of proof to other accounts in the literature. We advance two main points of criticism. First, other accounts are not necessarily incompatible with legal probabilism 1.02 which may in fact provide a more rigorous way to express their insights. This point applies to foundeherentism (Hack), normic support (Smith), argumentation theory (Sartor and Prakken) and to some extent relative plausibility (Allen and Pardo). The second criticism we make is that other accounts are engaged with what we might call ‘epistemology fetishism’—that is, they borrow ideas in contemporary analytic epistemology and force them onto legal-decision making. This criticism applies in particular to knowledge accounts of legal proof. To be sure, we might ourselves be accused of ‘probability fetishism’—that is, unquestionably taking probability theory as a paradigm of rationality and force it onto legal-decision making. Admittedly, there is no empirical evidence that a probabilistic turn in legal decision-making would improve trial decisions, if only because what it would mean to ‘improve’ trial decisions is unclear. A set of criteria are needed for assessing the desired improvements. We tackle this normative question in the final part of the book and show how probability theory can be of service here.

Part V - Accuracy and Fairness

What values and objectives should trial decisions seek to realize? What criteria and principles should guide trial decisions so that they can further these values and objectives? The fourth part of the book addresses these questions by focusing on improving the accuracy and fairness of trial decisions.

As a preliminary step, **Chapter 15** surveys different values and objectives that may inform the design of the trial system. In response to criminal and civil wrongs, we assume that the state has, in principle, the authority to impose punishments and decide about monetary compensations. Given this assumption, we survey the most common values and objectives that trial decisions should conform with: trial decisions should be accurate; they should be fair; they should be accompanied by a justification that is public and subject to scrutiny; they should be reversible under appeal; they should contribute to further social cohesion, compliance and deterrence; they should be humane and respectful of people’s dignity. We also explore to what extent these values and objectives may be in tension with one another and to what extent they can be subject to restrictions. The subsequent chapters focus on two values and objectives: accuracy and fairness. We focus on them, not because the others are not important, but because we believe they are foundational for the others.

Chapter 16 examines what it means for trial decisions to be accurate and how accuracy can be promoted. There are different ways to understand accuracy in trial decisions. We first distinguish accuracy in the single instance and accuracy in the long run. Accuracy in the single instance consists in the conviction of a defendant that is factually guilty and the acquittal of a defendant that is factually innocent. A similar definition applies to civil liability. Some have criticized the notion of ‘factual guilt’ on the ground that guilt is a judgment based on evidence, not a state of affairs that exists objectively. We argue that this view deprives trial decisions of objectivity and legitimacy. We then turn to accuracy in the long run. In this sense, accuracy can be understood as predictive or diagnostic. The former is the probability that, if the defendant is found (not) liable at trial, the defendant is actually (not) liable; the latter is the probability that if the defendant is (not) liable, the defendant would (not) be found liable at trial. The two notions are related but are not equivalent. Distinguishing them has implications for how we should understand standards of proof. To this end, we ask whether the standard of proof as we defined in earlier chapters—consisting of multiple criteria such as high probability of liability, narrative specificity and resistance to objections—does in fact promote accuracy in the long run. We devise a computer simulation that compares two models of the standard of proof. On the first model, the standard of proof simply requires that the defendant’s liability be established with a high probability. On the second model, the standard of proof requires, in addition to high probability, that the narrative presented by the prosecution or the plaintiff be reasonably specific. We compare the performance of the two models. The first model prevails in terms of predictive accuracy, while the second prevails in terms of diagnostic accuracy. We argue that diagnostic accuracy, rather than predictive accuracy, is the most adequate notion of accuracy for assessing trial decisions. We conclude that understanding the standard of proof as the combination of a ‘high probability’ and ‘narrative specificity’ (as we have argued in earlier chapters) is a more rational model of the standard of proof.

Chapter 17 turns to the fairness of trial decisions. What does it mean for trial decisions to be fair? In a formalistic sense, fairness merely requires that every rule be applied to all defendants in the same way. In a substantive sense, fairness requires, first, that trial defendants not be subject to excessive burdens, and second, that burdens (and benefits) be roughly equally distributed across different defendants. Call the latter substantive comparative fairness. It can be understood as the requirement that all defendants be exposed to roughly the same risks of mistaken conviction. Equipped with this notion of fairness, we examine whether admitting certain types of evidence makes trial decisions unfair, and to what extent trial decisions are inevitably unfair because of structural inequalities in society. We argue that admitting certain forms of evidence, specifically profile evidence and group-based generalizations, tends to burden certain groups of defendants more heavily than others, thus deteriorating the fairness of trial decisions. We also compare, via a computer simulation, different models of the standard of proof and examine which model is conducive to fairer trial decisions. We show that a standard of proof that incorporates multiple criteria—such as high probability of liability, specificity, resistance to objections and comprehensiveness of the evidence—allocates burdens and benefits more fairly across defendants than a standard of proof that simply requires that the probability of liability be above a fixed threshold. Thus, the multidimensional account of the standard of proof we have proposed fares relatively well in light of normative criteria such as accuracy (previous chapter) and fairness (this chapter).

Chapter 18 draws some conclusions and points to open problems for legal probabilism. We emphasize that our aim in the book is to show that legal probabilism is a richer and more flexible theory than it might seem at first. It is a theory that is able to incorporate insights from other, possibly rival theories. Legal probabilism 1.02 can satisfactorily address long-standing conceptual difficulties: the problem of naked statistical evidence, the difficulty with conjunction, corroboration and cross-examination. We also made some progress on other conceptual difficulties, such as the problem of priors and the problem of unanticipated possibilities. We underscore how, besides furthering the academic debate, ‘Boosting Legal Probabilism’ can be used by legal practitioners and reformers who aim to strengthen the protection for defendants against wrongful convictions, improve the accuracy of trial decisions and promote a fairer criminal justice system. We sketch how our theoretical findings can be implemented in the trial setting, but leave a more precise articulation of the challenges of implementation for future work.

Table of contents

I Legal probabilism 1.01 and its foes	10 Towards legal probabilism 1.02
1 The emergence of legal probabilism	10.1 Outperforming competing accounts
1.1 Famous cases	10.2 Empirical adequacy
1.2 Probabilistic evidence	10.3 Specificity and coherence
1.3 Trial by mathematics	10.4 Resistance against objections
1.4 Some history	10.5 Comprehensive evidence
2 A skeptical perspective	10.6 Bayesian network implementation
2.1 The difficulty about conjunction	IV Standards of proof
2.2 The problem of priors	11 Are standards of proof thresholds?
2.3 Naked statistical evidence	11.1 Legal background
2.4 The complexity problem	11.2 Probabilistic thresholds
2.5 Soft variables	11.3 Theoretical challenges
2.6 Corroboration	11.4 The comparative strategy
2.7 The reference class problem	11.5 The likelihood strategy
2.8 Unanticipated possibilities	11.6 Probabilistic thresholds revised
2.9 Non-probabilistic theories	11.7 Bayesian networks and probabilistic standard of proof
II Evidence assessment (first pass)	12 Naked statistical evidence
3 Bayes' theorem and the usual fallacies	12.1 Forty years of hypothetical
3.1 Assuming independence	12.2 Specific narratives
3.2 The prosecutor's fallacy	12.3 Cross-examination and causal grounding
3.3 Base rate fallacy	12.4 Specificity, causality and Bayesian networks
3.4 Defense attorney's fallacy	12.5 Are cold-hit DNA matches naked statistics?
3.5 Uniqueness fallacy	13 The Difficulty with Conjunction
3.6 Case studies	13.1 The problem
4 Complications and caveats	13.2 The likelihood strategy
4.1 Complex hypotheses and complex bodies of evidence	13.3 The comparative strategy
4.2 Source, activity and offense level hypotheses	13.4 The holistic strategy
4.3 Where do the numbers come from?	13.5 Complex bodies of evidence and structured narratives
4.4 Modeling corroboration	14 Other accounts
4.5 Stories, explanations and coherence	14.1 Baconian probability
5 Likelihood ratios and relevance	14.2 Sensitivity
5.1 Likelihood ratio as a measure of evidence strength	14.3 Normic Support
5.2 The risk of false positive and its impact	14.4 Foundherentism
5.3 Hypothesis choice	14.5 Relevant alternatives
5.4 Levels of hypotheses and the two-stain problem	14.6 Knowledge
5.5 Relevance and the small-town murder scenario	14.7 Relative Plausibility
5.6 The cold-hit confusion	14.8 Arguments
5.7 Likelihood ratio and cold-hit DNA matches	V Accuracy and Fairness
III Evidence assessment (more and better)	15 The objectives of trial decisions
6 Bayesian Networks	15.1 Conceptual desiderata
6.1 Bayesian networks to the rescue	15.2 Accurate and fair decisions
6.2 Legal evidence idioms	15.3 Protecting defendants
6.3 Scenario idioms	15.4 Public justification
6.4 Modeling relevance	15.5 Revision and appeal
6.5 Case study: Sally Clark	15.6 Dispute resolution and public deference
6.6 DNA evidence	16 Accuracy and the risk of error
7 Coherence	16.1 Single case accuracy
7.1 Existing probabilistic coherence measures	16.2 Minimizing expected errors
7.2 An array of counterexamples	16.3 Expected v. actual errors
7.3 Coherence of structured narrations with Bayesian networks	16.4 Predictive and diagnostic accuracy
7.4 Application to legal cases	16.5 Error reduction and error distribution
8 Conflicts	16.6 Accuracy, high probability and specificity
8.1 Argumentation theory	17 Fairness in trial decisions
8.2 Undercutting and rebutting evidence	17.1 Procedural v. substantive fairness
8.3 Cross-examination	17.2 Absolute v. comparative fairness
8.4 Conflicting evidence in Bayesian networks	17.3 Measures of fairness and measures of accuracy
9 Corroboration	17.4 Profile evidence and group-generalizations
9.1 Boole's formula and Cohen's challenge	17.5 Structural inequalities
9.2 Modeling substantial rise in case of agreement	17.6 A fairer standard of proof
9.3 Ekelöf's corroboration measure and evidentiary mechanisms	18 Conclusions
9.4 General approach with multiple false stories and multiple witnesses	18.1 Assessing the progress made
	18.2 Open questions problems
	18.3 To reformers and practitioners

1.3 Outstanding Features of the Book

- ‘Boosting Legal Probabilism’ is the first comprehensive sustained philosophical examination of legal probabilism and how it fares against well-known objections. It is also conciliatory in admitting that many insights delivered by the critics of legal probabilism are valid and should be incorporated within the probabilistic approach to legal fact-finding.
- The book is interdisciplinary. It closely engages with the literature in philosophy (see, for example, the discussion about coherence and corroboration) as well as literature outside philosophy, in artificial intelligence and in forensic science (see, for example, the discussion of Vlek’s story node approach and Fenton’s averaging).
- The philosophical argument in defense of legal probabilism is accompanied by mathematical proofs as well as an **R** code implementation. This underscores the theoretical, practical and computational aspiration of ‘Boosting Legal Probabilism.’
- The book is suitable for different audiences with different interests. It is partly introductory and partly describing original research by the authors. Instead of reading the entire book, one could follow different tracks. One could read the book to learn about the proof paradoxes (Chapter 2, 11, 12 and 13), Bayesian networks for evidence assessment and decision-making (Chapters 6 through 11, legal probabilism and its difficulties (Chapter 1, 2, 3, 4 and 11), the accuracy and fairness of trial decisions (Chapters 14, 16 and 17), etc. We will make sure to describe several tracks that readers could follow depending on their interests.
- ‘Boosting Legal Probabilism’ is divided into different parts, each tackling a distinct domain and level of analysis. The first half (Part II and Part III) examines how the evidence presented at trial should be assessed, weighed and evaluated using the tools of probability theory. The second half examines decisions at trial. The book offers a probability-based explication of legal standards of proof that is theoretically plausible (Part IV). In addition, the book also provides a normative justification of the proposed theory based on how standards of proof would perform in terms of accuracy and fairness under competing definitions (Part V).

1.4 Apparatus

Besides the customary list of references at the end, the book will contain multiple graphs and plots, Bayesian networks, and other data visualizations generated via the **R** package `ggplot2` as is standard in publications in statistics and machine learning.

Since computer simulations play an integral part in the argument, the book will be accompanied by supplementary materials available on-line. These materials will contain the **R** source code for the simulations along with tutorials for readers interested in learning the technical details. Some of the derivations and results will be also supported by associated **Mathematica** notebooks, which will also be made freely available.

1.5 Competition

Several books in print cover topics at the intersection of evidence, law and probability. They, however, do not have the same aims as our book, and since they can be sensibly grouped with respect to how our book will differ from them, we’ll discuss them in batches accordingly.

- Consider first the key books written by legal theories and philosophers, such as Stein (2005), *Foundations of Evidence Law*, Oxford University Press; Ho (2008), *A Philosophy of Evidence Law: Justice in the Search for Truth*, Oxford University Press; Haack (2014), *Evidence Matters: Science, Proof, and Truth in the Law*, Cambridge University Press; Nance (2016), *The Burdens of Proof: Discriminatory Power, Weight of Evidence, and Tenacity of Belief*, Cambridge University Press; Dahlman, Stein, and Tuzet (eds.) (2021), *Philosophical Foundations of Evidence Law*, Oxford University Press. These books offer original theories of evidence law, the standards of proof and legal decision-making under uncertainty. They address, from a philosophical perspective, many of the conceptual difficulties that we examine in our book. But they do not develop a probabilistic theory of evidence and decision-making that crucially relies on Bayesian networks. Nor do they combine analytic arguments, mathematical proofs and computer simulations. They also focus on rather abstract examples and their use of real-life cases and discussions between the practitioners is rather limited.
- Many books by forensic scientists and computer scientists have championed applications of Bayesian networks to legal evidence evaluation, for example, Taroni, Aitken, Garbolino and Biedermann (2006) *Bayesian Networks and Probabilistic Inference in Forensic Science*, Wiley; Taroni, Bozza, Biedermann, Garbolino and Aitken (2010), *Data Analysis in Forensic Science: A Bayesian Decision Perspective*, Wiley; Fenton and Neil (2012/2018) *Risk Assessment and Decision Analysis with Bayesian Networks*, Chapman and Hall/CRC Press. These books are groundbreaking in that they offer a detailed examination of how expert evidence can be assessed using Bayesian networks. They, however, do not address philosophical problems such as the problem of naked statistical evidence or the difficulty of conjunction, nor do they aim to offer a novel theory of the standard of proof, and they make no connection to the narrative approach or evidence evaluation criteria brought up by the critics of legal probabilism. Unlike our book, they do not address

double ch
all referen
publishers

normative questions such as the accuracy or fairness of trial decisions.

- Several books exist about Bayesian networks that are more general in content. Some of them focus on the technical dimension, such as Scutari and Denis (2014), *Bayesian Networks With Examples in R*, (Neapolitan 2004), *Learning Bayesian Networks* or Fenton, Neli and Martin (2013), *Risk Assessment and Decision Analysis with Bayesian Networks*. Some are more psychological and use Bayesian networks to examine how our minds rely on evidence to understand the world and solve problems, such as Spirtes, Glymour and Scheines (2000), *Causation, Prediction and Search*, Glymour (2001), *The mind's arrows. Bayes Nets and Graphical Causal Models in Psychology*, or Lagnado (2021), *Explaining the Evidence: How the Mind Investigates the World*, Cambridge University Press. These general books lack focus on legal applications (some legal applications are discussed in Fenton, but the coverage is fairly narrow as compared to our intent—the authors focus on covering on presenting a few technical representation methods of various types of evidence).
- There are books that focus exclusively on the evaluation of statistical and probabilistic evidence at trial. They include Robertson and Vignaux (1995), *Interpreting Evidence: Evaluating Forensic Science in the Courtroom*, Wiley [second edition: Robertson, Vignaux and Berger, published in 2016]; Lucy (2005), *Introduction to Statistics for Forensic Scientists*, Wiley; Finkelstein (2010), *Statistics for Lawyers*, Springer; Balding and Steele (2015), *Weight-of-Evidence for Forensic DNA Profiles*, Wiley; Buckleton, Bright and Taylor (eds.) (2016) *Forensic DNA Evidence Interpretation*, CRC Press. These books differ from our own because they are more specific, often focusing on just DNA evidence or select forms of expert evidence. Besides DNA and expert evidence, our book also covers common forms of evidence, such as eyewitness testimony. In addition, these books are written from the perspective of forensic science and do not examine the philosophical issues that are in the focus in our book, and they do not develop a general approach to the probabilistic modelling of narrations.
- Finally, a few books focus on the legal and ethical problems that arise from relying on statistical and actuarial evidence at trial and in the criminal justice system more generally. These books include Schauer (2006), *Profiles, Probabilities and Stereotypes*, Belknap Press; Harcourt (2008), *Against Prediction: Profiling, Policing, and Punishing in an Actuarial Age*, University of Chicago Press. Some chapters of our book discuss actuarial and statistical evidence, but this is not our sole focus. In addition, the authors of these books do not rely on legal probabilism as their guiding theoretical framework.

More needed
said about
as he does
legal evid

2 Market Considerations

2.1 The Primary Market

The target audience comprises scholars in various disciplines, primarily philosophers and legal theorists with an interest in epistemology, evidence and decision-making under uncertainty. Computer scientists and psychologists who work on similar topics will also be interested in reading the book.

Selection of chapters from the book are suitable to be used as primary readings in advanced undergraduate courses or graduate seminars with titles such as Legal Probabilism, Legal Epistemology, Probability and the Law, Bayesian Epistemology, Bayesian Networks in Philosophy, Statistics in the Law, Math on Trial. We have ourselves taught similar courses in the past and felt the need of a book such as ours. We believe other instructors felt the same.

3 Status of the Work

We plan 16 chapters, two of which already exist, so 14 remain. We predict each chapter will take around 2 months to complete, which results in 28 months of work. We add eight months for proofreading, delays, revisions and running various versions of the chapters by our colleagues, so overall the fairly lenient schedule suggests completion after three years.

With appendices the existing chapters have around 32k words each. We expect the early chapters (including appendices) to be shorter, and the later chapters to be of approximately that length, the whole book being around 350k words.

It is not clear how many graphs will be needed. Early chapters won't need many, the existing chapters have around 10-15 each (including appendices). Supposing this is the number for eight chapters, we seem to be looking at 80-100 graphs and visualizations.

We have already used some of the materials on which this book is based in graduate seminars, for example, in the seminar Bayesian Networks in the Philosophy that one of us taught at Arizona State in Spring 2021. We plan to use select chapters to teach graduate seminars and advanced undergraduate courses. By the time the book is published, we hope to have student-tested a significant portion of the book (in compliance with copyright restrictions as needed).

To better reach our intended audience, we plan to present portions of the book at international conferences in the coming years. We have been invited to an international conference in Girona, Spain on Evidence Law in Spring 2022. This conference

gathers some of the leading scholars in law, philosophy, forensic science and psychology interested in topics at the intersection of law, evidence, probability and decision-making.

We do not plan to include any material requiring additional copyright permissions.

4 Sample Chapters

Two sample chapters with their appendices are attached to this submission.

5 Reviews

- Floris Bex
- Alex Biedermann
- Edward Cheng
- Christian Dahlman
- Norman Fenton
- David Lagnado
- Anne Hsu
- Brian Hedden
- Silja Renooij
- Frederick Schauer
- Silvia Bozza
- Bart Verheij

6 Author Background

Two CVs are attached to this submission.