

Boosting Legal Probabilism

Marcello Di Bello and Rafal Urbaniak

1 The Book

1.1 Brief Description

In one or two paragraphs, describe the work, including its rationale, approach, and pedagogy. (This book is... It does... Its distinguishing features are...)

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 and expected utility? 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 extensively debated in the literature in philosophy, law, forensic science and artificial intelligence. Legal probabilism is a research program whose supporters believe that the answer to these questions is, by and large, affirmative. Others, however, hold a less optimistic view. ‘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.

We first articulate the simple version of legal probabilism. This version has much to be recommended for, but also falls prey to several difficulties, including the problem of conjunction, puzzles of naked statistical evidence, 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 they cannot be addressed within the confines of simple legal probabilism. We then develop a more sophisticated version of the theory, what we call legal probabilism 1.02, which deploys Bayesian networks and takes advantage of 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 show that the more sophisticated version significantly improves on the simple version and 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 analytical argument is supplemented with an **R** code implementation.

‘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 draw attention outside philosophy from 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. So the book is aimed at scholars, advanced undergraduates and curious readers more generally. Some chapters present original research and require technical background in probability theory. Others are introductory, suitable for an advanced undergraduate course.

1.2 Outline

Part I

The first part of the book outlines simple legal probabilism (Chapter 1) and its foes (Chapter 2). The simple version 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, simple legal probabilism is liable to a host of conceptual difficulties: the conjunction problem, the problem of priors, paradoxes of naked statistical evidence. These difficulties are well-known. Others are less familiar: the problem of complexity, soft variables, the difficulty with corroboration.

Part I is meant to instill interest in the topic among unfamiliar readers and refresh seasoned readers about the main points of contention. Part I provides the essential background for a deeper examination of legal probabilism and the development of its more sophisticated version. The remaining two parts of the book covers two separate topics: evidence assessment (Part II) and decision-making (Part III). 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

The second part of the book discusses in great detail three formal tools used for the assessment of evidence from a probabilistic perspective: Bayes' theorem, likelihood ratios and Bayesian networks. This part discusses how these formal tools can help to assess, weigh and evaluate evidence at trial as well as what their limitations are. The objective of this part of the book is to articulate a more sophisticated version of legal probabilism for evidence assessment which overcomes the difficulties of the simpler version.

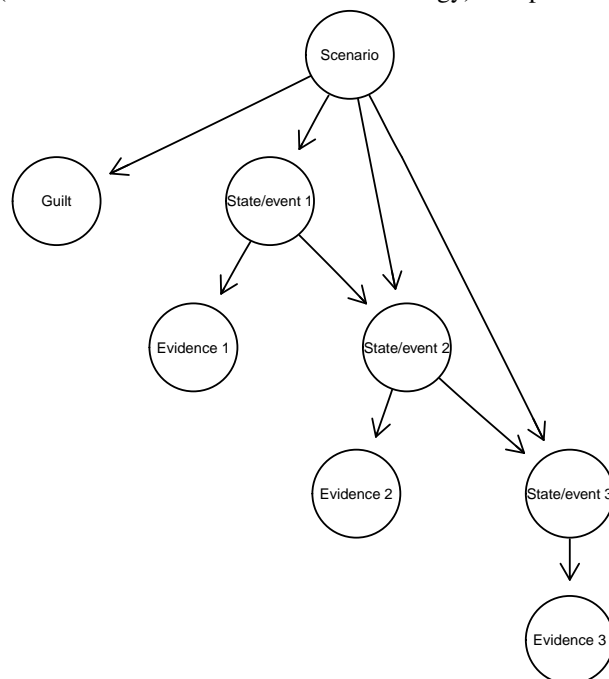
Chapter 3 begins with Bayes' theorem and illustrates how the theorem can be put to use in weighing pieces of evidence at trial. Bayes' theorem is useful in many ways, for example, to avoid reasoning fallacies such as the prosecutor's fallacy and the base rate fallacy. At the same time, its applications 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 a simple application of Bayes' theorem. A more sophisticated machinery for evidence assessment is needed.

Before discussing this more sophisticated machinery, the book describes a formal tool that is, in some important way, an alternative to Bayes' theorem and that 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 offer a way to evaluate the evidence presented at trial without the need of assessing prior probabilities. Chapter 5 illustrates the many applications of likelihood ratios focusing on the debate about cold-hit DNA matches as an illustration. The chapter also examines the weaknesses of the likelihood ratio approach, in particular, the chapter details how the choice of the competing hypotheses to be compared in the likelihood ratio is often a source of confusion, manipulation and subjective judgment. A further problem is that, like Bayes' theorem, likelihood ratios are still unable to model complex bodies of evidence.

Chapters 3, 4 and 5 – combined – show that we need to move past simple legal probabilism. The journey toward legal probabilism 1.02 is accomplished in Chapters 6 through 10. 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, such as undercutting, rebutting, converging, corroborating evidence. We show that Bayesian networks constitute the formal machinery for developing a more sophisticated legal probabilism that is

able to formally capture these phenomena.

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. Simple graphical patterns called ‘idioms’ often appear while modeling the relationships between evidence and hypotheses. 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. Chapter 6 surveys this literature focusing on the work of Charlotte Vlek and Norman Fenton. 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. These correspond 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 dependency between the different parts, the more coherent the story. There may also be competing stories about what happened, say, one story was put forward by the prosecution and another by the defense. A Bayesian network can be built that comprises two competing stories and specifies that they are incompatible and cannot be true concurrently. Another approach to model 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 relative strength of the competing stories.

Chapter 7 and 8 articulate different lines of criticism of Vlek’s and Fenton’s formalizations. The criticisms are followed by our positive proposal that still relies, like Vlek’s and Fenton’s, on Bayesian networks. Chapter 7 focuses on weaknesses of the story node approach as an account of coherence. 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. Since these dependencies may or may not exist, they should be assessed on a case-by-case basis and modeled accordingly. In addition, the story node approach is overly simplistic as an account of coherence and fails to engage with the rich philosophical literature on the topic. After the critical part, Chapter 7 articulates a more adequate probabilistic account of the coherence of a story.

Instead of adding a story node, we show that it is more appropriate to assess the dependency between the different parts of a story on a case-by-case basis and build the Bayesian network accordingly. [THIS FEELS A BIT THIN. ANYTHING ELSE HERE TO DESCRIBE THE POSITIVE ACCOUNT OF COHERENCE? ANYTHING TO ADD?]

Chapter 8 and 9 focus on complex relationships between pieces of evidence, such as conflicting evidence and corroborating evidence. A criticism we level against Vlek's story node approach and Fenton's model comparison approach is that they both fail to capture crucial aspects of how pieces of evidence and competing stories conflict with one another at trial. It is too simplistic to posit that the complex adversarial dialectic that takes place in a trial could be modeled by 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 rather the reinterpretation of an existing story by supplying additional information. As chapter 8 illustrates, this process of reinterpretation can be represented formally as the refinement of an existing Bayesian network. We show that, by drawing additional arrows between evidence nodes and hypothesis nodes, conflicts between pieces of evidence such as undercutting and rebutting can be modeled adequately.

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 outside the literature in legal probabilism. Chapter 9 surveys the probabilistic literature on corroboration and the main difficulties that have been levelled against the proposed accounts. The chapter then formalizes a notion of corroboration based on Bayesian networks that overcomes most of the difficulties of the existing accounts. [ANYTHING ELSE HERE TO DESCRIBE THE POSITIVE ACCOUNT OF CORROBORATION?]

Chapter 10 summarizes our proposal for legal probabilism 1.02, limited to the realm of evidence assessment at trial, draws some general morals and compares our proposal to competing accounts. In preceding chapters, we show that legal probabilism 1.02 – roughly, simple legal probabilism supplemented by Bayesian networks – can represent complex relationships between pieces of evidence such as undercutting and corroboration (Chapters 8 and 9), as well as holistic notions such as coherence (Chapter 7). The key idea is that the coherence of a story as well as conflicts between pieces of evidence can be modeled formally in terms of corresponding properties of, operations on, and relations between Bayesian networks. Our formal analysis was guided by the following working hypothesis: An accusation of liability must be substantiated by providing a well-specified account – story, 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. Judges and jurors are tasked with assessing how well the claim that the defendant is civilly or criminally liable is supported by the evidence and how well it stands up against criticism. It is this complex dynamics that we attempted to formalize in the preceding chapters.

Chapter 10 compares our proposed version of legal probabilism 1.02 with two competing accounts of judicial fact-finding: argumentation theory and relative plausibility. We suggest that our theory outperforms these competing accounts in some respect, although we do not claim that it outperforms them on all counts. First, argumentation theory does not offer a clear account of degrees of evidential strength. Evidence may support a hypothesis more or less strongly, or two pieces of evidence may conflict with one another more or less strongly. Not only does legal probabilism 1.02 offer an account of evidential support, conflict and convergence, but the account is nuanced enough to capture degrees of support, conflict and convergence. This is marked improvement. Second, relative plausibility is often criticized because the defense need not always present a full-fledged alternative story or alternative explanation of the evidence. The criticism is that it should be enough for the defense to challenge the story or explanation proposed by the party, say the prosecution in a criminal case. We do not take a stance on this controversy, but we note that legal probabilism 1.02 is flexible enough to model competing stories or explanations (in agreement with relative plausibility) and alternatively model conflicts between

pieces of evidence without the need to construct a full-fledged alternative story or explanation (as critics of relative plausibility prefer). There remains, however, a fundamental difficulty with legal probabilism 1.02, its cognitive plausibility. It is not clear that the process of evidence assessment naturally conforms to the probabilistic machinery. In fact, it most likely does not. This difficulty does not undermine legal probabilism 1.02 as a theory of evidence assessment, but it certainly questions its empirical adequacy as a theory of evidence assessment at trial. In this respect, relative plausibility – and perhaps argumentation theory – fares much better.

MAYBE ADD OUR ACCOUNT CAN CAPTURE THINGS LIKE COMPLETENESS OF EVIDENCE, SPECIFICITY OF A NARRATIVE, WHETHER A PIECE OF EVIDENCE IS INDEPENDENTLY SECURED, CROSSWORD PUZZLE ANALOGY AND SUSAN HAACK.

Part III

The final 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.

Table of contents

- I Legal probabilism 1.01 and its foes
 - 1 The emergence of legal probabilism
 - 1.1 Famous cases
 - 1.2 Probabilistic evidence
 - 1.3 Trial by mathematics
 - 1.4 Some history
 - 2 A skeptical perspective
 - 2.1 The difficulty about conjunction
 - 2.2 The problem of priors
 - 2.3 Naked statistical evidence
 - 2.4 The complexity problem
 - 2.5 Soft variables
 - 2.6 Corroboration
 - 2.7 The reference class problem
 - 2.8 Non-probabilistic theories
- II Evidence assessment
 - 3 Bayes’ Theorem and the usual fallacies
 - 3.1 Assuming independence
 - 3.2 The prosecutor’s fallacy
 - 3.3 Base rate fallacy
 - 3.4 Defense attorney’s fallacy
 - 3.5 Uniqueness fallacy
 - 3.6 Case studies
 - 4 Complications and caveats
 - 4.1 Complex hypotheses and complex bodies of evidence
 - 4.2 Source, activity and offense level hypotheses
 - 4.3 Where do the numbers come from?
 - 4.4 Modeling corroboration
 - 4.5 Stories, explanations and coherence
 - 5 Likelihood Ratios and Relevance
 - 5.1 Likelihood ratio as a measure of evidence strength
 - 5.2 The risk of false positive and its impact
 - 5.3 Hypothesis choice
 - 5.4 Levels of hypotheses and the two-stain problem
 - 5.5 Relevance and the small-town murder scenario

- 5.6 The cold-hit confusion
- 5.7 Likelihood ratio and cold-hit DNA matches
- 6 Bayesian Networks
 - 6.1 Bayesian networks to the rescue
 - 6.2 Legal evidence idioms
 - 6.3 Scenario idioms
 - 6.4 Modeling relevance
 - 6.5 Case study: Sally Clark
 - 6.6 DNA evidence
- 7 Conflicts
 - 7.1 Argumenation theory
 - 7.2 Undercutting and rebutting evidence
 - 7.3 Cross-examination
 - 7.4 Conflicting evidence in Bayesian networks
- 8 Corroboration
 - 8.1 Boole's formula and Cohen's challenge
 - 8.2 Modeling substantial rise in case of agreement
 - 8.3 Ekelöf's corroboration measure and evidentiary mechanisms
 - 8.4 General approach with multiple false stories and multiple witnesses
- 9 Coherence
 - 9.1 Existing probabilistic coherence measures
 - 9.2 An array of counterexamples
 - 9.3 Coherence of structured narrations with Bayesian networks
 - 9.4 Application to legal cases
- 10 New legal probabilism
 - 10.1 Desiderata
 - 10.2 A probabilistic framework for narrations
 - 10.3 Probabilistic explications of the desiderata
 - 10.4 Bayesian network implementation
- III Trial Decisions
 - 11 The functions of the proof standards
 - 11.1 Conceptual desiderata
 - 11.2 Protecting defendants
 - 11.3 Error reduction and error distribution/allocation
 - 11.4 Dispute resolution and public deference
 - 11.5 Justification and answerability
 - 12 Standards of proof
 - 12.1 Legal background
 - 12.2 Probabilistic thresholds
 - 12.3 Theoretical challenges
 - 12.4 Specific narratives
 - 12.5 The comparative strategy
 - 12.6 The likelihood strategy
 - 12.7 Challenges (again)
 - 12.8 Probabilistic thresholds revised
 - 12.9 Bayesian networks and probabilistic standard of proof
 - 13 Accuracy and the risk of error
 - 13.1 Minimizing expected costs
 - 13.2 Minimizing expected errors
 - 13.3 Expected v. actual errors
 - 13.4 Competing accounts of the risk of error

- 13.5 Bayesian networks and the risk of error
- 14 Fairness in trial decisions
 - 14.1 Procedural v. substantive fairness
 - 14.2 Competing measures of substantive fairness
 - 14.3 Bayesian networks and fairness
- 15 Alternative accounts and legal probabilism
 - 15.1 Baconian probability
 - 15.2 Relative Plausibility
 - 15.3 Arguments
 - 15.4 Sensitivity
 - 15.5 Normic Support
 - 15.6 Justification/foundherentism
 - 15.7 Completeness
 - 15.8 Relevant alternatives
 - 15.9 Knowledge
- 16 Conclusions

1.3 Outstanding Features of the Book

- (First) comprehensive sustained philosophical discussion of legal probabilism.
- Multi-faceted in its incorporation of insights from various discussions present in legal, philosophical, and forensic research.
- With a practical accent, due to the implementation of the conceptual points by means of bayesian networks and **R** programming language.

what else?

1.4 Apparatus

a. Will the book include photographs, line drawings, cases, questions, problems, glossaries, bibliography, references, appendices, etc.?

Yes, the book will contain various plots, either of Bayesian networks, or some other data visualisations generated by `ggplot2`. The book also will contain bibliography.

b. If the book is a text, do you plan to provide supplementary material to accompany it? (Teacher's manual, study guide, solutions, answers, workbook, anthology, or other material.)

The book will be accompanied by an online-only appendix detailing the use of the R code in the book and the source code we used.

1.5 Competition

a. Consider the existing books in this field and discuss specifically their strengths and weaknesses. Spell out how your book will be similar to, as well as different from, competing works.

Three types: BNs in the law, Philosophy & law, Statistics in law and forensics

- "Bayesian Networks and Probabilistic Inference in Forensic Science" by Taroni, Aitken, Garbolino and Biedermann.
- "Risk Assessment and Decision Analysis with Bayesian Networks" by Fenton and Neil.
- "Bayesian Networks With Examples in R" by Marco Scutari and Jean-Baptiste Denis.
- Alex Stein, foundations of evidence law
- Nance, Burdens of proof
- Schauer, Profiles, ...
- Ho, Philosophy of evidence law
- Robertson, Vignaux
- Lucy Dawid,
- Statistics for Lawyers etc.

For now, let's list competition, and discuss key differences

- b. Consider what aspects of topical coverage are similar to or different from the competition. What topics have been left out of competing books and what topics have been left out of yours?
- c. Please discuss each competing book in a separate paragraph. (If possible, please provide us with the publisher and date of publication as well.) This information will provide the reviewers and the publisher a frame of reference for evaluating your material. Remember, you are writing for reviewers and not for publication, so be as frank as possible regarding your competition. Give credit where credit is due, and show how you can do it better.

2 Market Considerations

2.1 The Primary Market

1. What is the major market for the book? (Scholarly/professional, text, reference, trade?)
2. If this is a text, for what course is the book intended? Is the book a core text or a supplement? What type of student takes this course? What is the level? (Major or non-major; freshman, senior, graduate?) Do you offer this course yourself? If so, how many times have you given it? Is your text class-tested?
3. If the market is scholarly/professional, reference, or trade, how may it best be reached? (Direct mail, relevant journals, professional associations, libraries, book or music stores?) For what type of reader is your book intended?

3 Status of the Work

1. Do you have a timetable for completing the book?
 - a. What portion or percentage of the material is now complete?
 - b. When do you expect to have a complete manuscript?
2. What do you estimate to be the size of the completed book?
 - a. Double spaced typewritten pages normally reduce about one-third when set in type; e.g., 300 typewritten pages make about 200 printed pages. There are about 450 words on a printed page.
 - b. Approximately how many photographs do you plan to include?
 - c. Approximately how many line drawings (charts, graphs, diagrams, etc.) will you need?
 - d. Do you plan to include material requiring permission (text, music, lyrics, illustrations)? To what extent? Have you started the permissions request process?
3. Do you plan to class-test the material in your own or other sections of the course? (Any material distributed to students should be protected by copyright notice on the material.)

4 Sample Chapters

Select one or two chapters of the manuscript that are an integral part of the book. They should be those you consider the best-written ones, and do not have to be in sequence. For example, you might submit chapters 3, 7, and 14 of a 20-chapter book, so long as these chapters represent the content and reflect your writing style and pedagogy in the best possible light. It is also advisable to submit any chapter that is particularly innovative or unique. Sample chapters should contain rough sketches, charts, hand-written musical examples or xerox reproductions, and description of photographs to be included. The material need not be in final form, although it should be carefully prepared and represent your best work. In your preparation, emphasis should be on readability. Please do not bind your manuscript, as we will have to unbind it in order to make photocopies for reviewers. Also be sure all pages are numbered either consecutively or double-numbered by chapter.

5 Reviews

If we are interested in your project, we will commission outside reviewers to read and evaluate your proposal. We will, of course, obtain the best available reviewers to consider your work. If you wish to suggest the names of experts in your field whom you believe to be ideally suited to evaluate your proposal, you may provide their names, titles, and email addresses. While we are unlikely to approach these scholars to act as reviewers themselves, we may ask them for their suggestions for peer readers. Naturally, we do not reveal the names of reviewers without their permission.

6 Author Background

Please include a current CV or brief biography of your writing, teaching, and/or educational background and experience. Be sure to list any books that you have previously published, and any other information about yourself on why you are qualified to write this book.

7 Response Time

Please allow at least 6-10 weeks for the manuscript proposal evaluation and review process. We will contact you as soon as we have had a chance to thoroughly examine your manuscript proposal. Thank you for your interest in Oxford University Press. We look forward to reading your materials.