

Rolf Haenni  
Jan-Willem Romeijn  
Gregory Wheeler  
Jon Williamson

# Probabilistic Logics and Probabilistic Networks

## Probabilistic Logics and Probabilistic Networks

# SYNTHESE LIBRARY

STUDIES IN EPISTEMOLOGY,  
LOGIC, METHODOLOGY, AND PHILOSOPHY OF SCIENCE

*Editors-in-Chief:*

VINCENT F. HENDRICKS, *University of Copenhagen, Denmark*  
JOHN SYMONS, *University of Texas at El Paso, U.S.A.*

*Honorary Editor:*

JAAKKO HINTIKKA, *Boston University, U.S.A.*

*Editors:*

DIRK VAN DALEN, *University of Utrecht, The Netherlands*  
THEO A.F. KUIPERS, *University of Groningen, The Netherlands*  
TEDDY SEIDENFELD, *Carnegie Mellon University, U.S.A.*  
PATRICK SUPPES, *Stanford University, California, U.S.A.*  
JAN WOLEŃSKI, *Jagiellonian University, Kraków, Poland*

VOLUME 350

For further volumes:  
<http://www.springer.com/series/6607>

# Probabilistic Logics and Probabilistic Networks

by

**Rolf Haenni**

*Bern University of Applied Sciences, Switzerland*

**Jan-Willem Romeijn**

*University of Groningen, The Netherlands*

**Gregory Wheeler**

*Universidade Nova de Lisboa, Portugal*

*and*

**Jon Williamson**

*University of Kent, UK*



**Springer**

Prof. Rolf Haenni  
Bern University of Applied Sciences  
Department of Engineering and  
Information Technology  
Quellgasse 21  
CH-2501 Biel  
Switzerland  
rolf.haenni@bfh.ch

Prof. Jan-Willem Romeijn  
University of Groningen  
Faculty of Philosophy  
Oude Boteringestraat 52  
9712 GL Groningen  
Netherlands  
J.W.Romeijn@rug.nl

Dr. Gregory Wheeler  
Universidade Nova de Lisboa  
New University of Lisbon  
Quinta da Torre  
2829-516 Caparica  
Portugal  
grw@fct.unl.pt

Prof. Jon Williamson  
University of Kent  
School of European Culture & Languages  
Sec. Philosophy  
Cornwallis North West  
CT2 7NF Canterbury, Kent  
United Kingdom  
j.williamson@kent.ac.uk

ISBN 978-94-007-0007-9 e-ISBN 978-94-007-0008-6  
DOI 10.1007/978-94-007-0008-6  
Springer Dordrecht Heidelberg London New York

Library of Congress Control Number: 2010938961

© Springer Science+Business Media B.V. 2011

No part of this work may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, microfilming, recording or otherwise, without written permission from the Publisher, with the exception of any material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work.

Printed on acid-free paper

Springer is part of Springer Science+Business Media ([www.springer.com](http://www.springer.com))

*Chaos umpire sits,  
And by decision more embroils the fray  
By which he reigns: next him high arbiter  
Chance governs all.*

(John Milton, *Paradise Lost*)

*Chance, too, which seems to rush along  
with slack reins, is bridled and governed by  
law.*

(Boethius, *The Consolation of Philosophy*)

*Er glaubte nämlich, die Erkenntnis jeder  
Kleinigkeit, also zum Beispiel auch eines sich  
drehenden Kreisels, genüge zur Erkenntnis  
des Allgemeinen. Darum beschäftigte er sich  
nicht mit den großen Problemen, das schien  
ihm unökonomisch. War die kleinste  
Kleinigkeit wirklich erkannt, dann war alles  
erkannt, deshalb beschäftigte er sich nur mit  
dem sich drehenden Kreisel.*

(Kafka, *Der Kreisel*)<sup>1</sup>

*Lest men suspect your tale untrue,  
Keep probability in view.*

(John Gray, *The Painter who pleased  
Nobody and Everybody*)

---

<sup>1</sup> For he believed that the understanding of any trifle, so for example of a spinning top, would suffice for the understanding of everything. This is why he did not concern himself with the big questions, which seemed uneconomical to him. If the smallest detail was truly understood, then so was everything, hence he only busied himself with the spinning top.

# Preface

While in principle probabilistic logics might be applied to solve a range of problems, in practice they are rarely applied at present. This is perhaps because they seem disparate, complicated, and computationally intractable. However, we shall argue in this programmatic book that several approaches to probabilistic logic fit into a simple unifying framework: logically complex evidence can be used to associate probability intervals or probabilities with sentences.

Specifically, we show in Part I that there is a natural way to present a question posed in probabilistic logic, and that various inferential procedures provide semantics for that question: the standard probabilistic semantics (which takes probability functions as models), probabilistic argumentation (which considers the probability of a hypothesis being a logical consequence of the available evidence), evidential probability (which handles reference classes and frequency data), classical statistical inference (in particular the fiducial argument), Bayesian statistical inference (which ascribes probabilities to statistical hypotheses), and objective Bayesian epistemology (which determines appropriate degrees of belief on the basis of available evidence).

Further, we argue, there is the potential to develop computationally feasible methods to mesh with this framework. In particular, we show in Part II how credal and Bayesian networks can naturally be applied as a calculus for probabilistic logic. The probabilistic network itself depends upon the chosen semantics, but once the network is constructed, common machinery can be applied to generate answers to the fundamental question introduced in Part I.

Bern,  
Groningen,  
Lisbon,  
Canterbury,

*Rolf Haenni*  
*Jan-Willem Romeijn*  
*Gregory Wheeler*  
*Jon Williamson*  
April 2009

# Acknowledgements

This research was undertaken as a part of the *proginet* international network on Probabilistic Logic and Probabilistic Networks. We are very grateful to The Leverhulme Trust for financial support.



# Contents

## Part I Probabilistic Logics

<b>1</b>	<b>Introduction</b> .....	3
1.1	The Fundamental Question of Probabilistic Logic .....	3
1.2	The Potential of Probabilistic Logic .....	4
1.3	Overview of the Book .....	5
1.4	Philosophical and Historical Background .....	7
1.5	Notation and Formal Setting .....	9
<b>2</b>	<b>Standard Probabilistic Semantics</b> .....	11
2.1	Background .....	11
2.1.1	Kolmogorov Probabilities .....	12
2.1.2	Interval-Valued Probabilities .....	13
2.1.3	Imprecise Probabilities .....	15
2.1.4	Convexity .....	16
2.2	Representation .....	18
2.3	Interpretation .....	19
<b>3</b>	<b>Probabilistic Argumentation</b> .....	21
3.1	Background .....	22
3.2	Representation .....	25
3.3	Interpretation .....	26
3.3.1	Generalizing the Standard Semantics .....	26
3.3.2	Premises from Unreliable Sources .....	28
<b>4</b>	<b>Evidential Probability</b> .....	33
4.1	Background .....	33
4.1.1	Calculating Evidential Probability .....	37
4.1.2	Extended Example: <i>When Pigs Die</i> .....	40
4.2	Representation .....	44
4.3	Interpretation .....	44

- 4.3.1 First-order Evidential Probability ..... 45
- 4.3.2 Counterfactual Evidential Probability ..... 46
- 4.3.3 Second-Order Evidential Probability ..... 46
- 5 Statistical Inference** ..... 49
  - 5.1 Background ..... 49
    - 5.1.1 Classical Statistics as Inference? ..... 49
    - 5.1.2 Fiducial Probability ..... 52
    - 5.1.3 Evidential Probability and Direct Inference ..... 55
  - 5.2 Representation ..... 57
    - 5.2.1 Fiducial Probability ..... 57
    - 5.2.2 Evidential Probability and the Fiducial Argument ..... 58
  - 5.3 Interpretation ..... 59
    - 5.3.1 Fiducial Probability ..... 59
    - 5.3.2 Evidential Probability ..... 60
- 6 Bayesian Statistical Inference** ..... 63
  - 6.1 Background ..... 63
  - 6.2 Representation ..... 65
    - 6.2.1 Infinitely Many Hypotheses ..... 66
    - 6.2.2 Interval-Valued Priors and Posteriors ..... 68
  - 6.3 Interpretation ..... 69
    - 6.3.1 Interpretation of Probabilities ..... 69
    - 6.3.2 Bayesian Confidence Intervals ..... 70
- 7 Objective Bayesian Epistemology** ..... 73
  - 7.1 Background ..... 73
    - 7.1.1 Determining Objective Bayesian Degrees of Belief ..... 74
    - 7.1.2 Constraints on Degrees of Belief ..... 75
    - 7.1.3 Propositional Languages ..... 76
    - 7.1.4 Predicate Languages ..... 77
    - 7.1.5 Objective Bayesianism in Perspective ..... 79
  - 7.2 Representation ..... 80
  - 7.3 Interpretation ..... 80
- Part II Probabilistic Networks**
- 8 Credal and Bayesian Networks** ..... 85
  - 8.1 Kinds of Probabilistic Network ..... 86
    - 8.1.1 Extensions ..... 87
    - 8.1.2 Extensions and Coordinates ..... 88
    - 8.1.3 Parameterised Credal Networks ..... 90
  - 8.2 Algorithms for Probabilistic Networks ..... 91
    - 8.2.1 Requirements of the Probabilistic Logic Framework ..... 91
    - 8.2.2 Compiling Probabilistic Networks ..... 92
    - 8.2.3 The Hill-Climbing Algorithm for Credal Networks ..... 94

- 8.2.4 Complex Queries and Parameterised Credal Networks . . . . . 96
- 9 Networks for the Standard Semantics . . . . . 99**
  - 9.1 The Poverty of Standard Semantics . . . . . 99
  - 9.2 Constructing a Credal Net . . . . . 100
  - 9.3 Dilation and Independence . . . . . 104
- 10 Networks for Probabilistic Argumentation . . . . . 107**
  - 10.1 Probabilistic Argumentation with Credal Sets . . . . . 107
  - 10.2 Constructing and Applying the Credal Network . . . . . 108
- 11 Networks for Evidential Probability . . . . . 111**
  - 11.1 First-Order Evidential Probability . . . . . 111
  - 11.2 Second-Order Evidential Probability . . . . . 113
  - 11.3 Chaining Inferences . . . . . 116
- 12 Networks for Statistical Inference . . . . . 119**
  - 12.1 Functional Models and Networks . . . . . 119
    - 12.1.1 Capturing the Fiducial Argument in a Network . . . . . 119
    - 12.1.2 Aiding Fiducial Inference with Networks . . . . . 120
    - 12.1.3 Trouble with Step-by-Step Fiducial Probability . . . . . 122
  - 12.2 Evidential Probability and the Fiducial Argument . . . . . 123
    - 12.2.1 First-Order EP and the Fiducial Argument . . . . . 123
    - 12.2.2 Second-Order EP and the Fiducial Argument . . . . . 124
- 13 Networks for Bayesian Statistical Inference . . . . . 125**
  - 13.1 Credal Networks as Statistical Hypotheses . . . . . 125
    - 13.1.1 Construction of the Credal Network . . . . . 126
    - 13.1.2 Computational Advantages of Using the Credal Network . . . 127
  - 13.2 Extending Statistical Inference with Credal Networks . . . . . 128
    - 13.2.1 Interval-Valued Likelihoods . . . . . 129
    - 13.2.2 Logically Complex Statements with Statistical Hypotheses . 131
- 14 Networks for Objective Bayesianism . . . . . 133**
  - 14.1 Propositional Languages . . . . . 133
  - 14.2 Predicate Languages . . . . . 135
- 15 Conclusion . . . . . 139**
- References . . . . . 141**
- Index . . . . . 153**

**Part I**  
**Probabilistic Logics**

# Chapter 1

## Introduction

### 1.1 The Fundamental Question of Probabilistic Logic

In a non-probabilistic logic, the fundamental question of interest is whether a proposition  $\psi$  is entailed by premise propositions  $\varphi_1, \dots, \varphi_n$ :

$$\varphi_1, \dots, \varphi_n \models \psi?$$

A *probabilistic logic* (or *progic* for short) differs in two respects. First, the propositions have probabilities attached to them. Thus the premises have the form  $\varphi^X$ , where  $\varphi$  is a classical proposition and  $X \subseteq [0, 1]$  is a set of probabilities, and each premise is interpreted as ‘the probability of  $\varphi$  lies in  $X$ ’.<sup>1</sup> Second, the analogue of the classical question,

$$\varphi_1^{X_1}, \dots, \varphi_n^{X_n} \models \psi^Y?$$

is of little interest, because while there is often a natural conclusion  $\psi$  under consideration, there is rarely a natural probability set  $Y$  presented by the problem at hand since there are so many possible candidates for  $Y$  to choose from. Rather, the question of interest is the determination of  $Y$  itself:

$$\varphi_1^{X_1}, \dots, \varphi_n^{X_n} \models \psi? \tag{1.1}$$

---

<sup>1</sup> This characterisation of probabilistic logic clearly covers what are called *external* progics in (Williamson, 2009b, §21)—the probabilities are metalinguistic, external to the propositions themselves. But it also covers *internal* progics, where the propositions involve probabilities (discussed in (Halpern, 2003), for example), and *mixed* progics, where there are probabilities both internal and external to the propositions.

That is, what set  $Y$  of probabilities should attach to the conclusion sentence  $\psi$ , given the premises  $\varphi_1^{X_1}, \dots, \varphi_n^{X_n}$ ? This is a very general question, which will be referred to as the *Fundamental Question of Probabilistic Logic*, or simply as Schema (1.1).<sup>2,3</sup>

Part I of this book is devoted to showing that the fundamental question outlined above is indeed very general, providing a framework into which several common inferential procedures fit. Since the fundamental question of probabilistic logic differs from that of non-probabilistic logic, different techniques may be required to answer the two kinds of question. While proof techniques are often invoked to answer the questions posed in non-probabilistic logics, in Part II we show that probabilistic networks can help answer the fundamental question of probabilistic logic.

The programme of this book—namely that of showing how the fundamental question can (i) subsume a variety of inferential procedures and (ii) be answered using probabilistic networks—we call the *progicnet programme*. We view this as a logical programme. While it is possible to hold a narrow view of logic as being concerned primarily with proof—i.e., with the task of developing sound and complete axiomatisations—the progicnet programme is intended to fit with a broad view of logic as being concerned on the one hand with semantics—i.e., with specifying which inferences are in principle condoned by the logic—and on the other hand with calculi—i.e., with specifying one or more practical procedures for answering a given inferential question (e.g., truth-tables, semantic trees, proof systems, or, in our case, probabilistic networks).

## 1.2 The Potential of Probabilistic Logic

Due to the generality of Schema (1.1), many problem domains would benefit from an efficient means to answer its question—any problem domain whose structure has a natural logical representation and whose observations are uncertain in some respect. Here are some examples. In the philosophy of science we are concerned with the extent that a (logically complex) conclusion hypothesis is confirmed by a range of premise hypotheses and evidential statements which are themselves uncertain. In bioinformatics we are often interested in the probability that a complex molecule  $\psi$  is present, given the uncertain presence of molecules  $\varphi_1, \dots, \varphi_n$ . In natural language processing we are interested in the probability that an utterance has semantic structure  $\psi$  given uncertain semantic structures of previous utterances and uncertain

<sup>2</sup> In asking what set of probabilities *should* attach to the conclusion, we are restricting our attention to logic rather than psychology. While the question of how humans go about ascribing probabilities to conclusions in practice is a very interesting question, it is not one that we broach in this book.

<sup>3</sup> In our notation, the probabilities are attached to the propositions  $\varphi_1, \dots, \varphi_n, \psi$ , *not* to the entailment relation. However, in the literature one sometimes sees expressions of the form  $\varphi_1^{X_1}, \dots, \varphi_n^{X_n} \models^Y \psi$  (Williamson, 2002, §2.2–2.3). Our choice of notation is largely a question of convenience: in our notation the premises and conclusion turn out to be the same sort of thing, namely propositions with attached probabilities, and there is a single entailment relation rather than an uncountable infinity of entailment relations; but of course from a formal point of view the two kinds of expression can be used interchangeably.

contextual factors. In robotics we are interested in finding the sequence of actions of a robot that is most likely to achieve a goal given the uncertain structure of the robot's surroundings. In expert systems we are interested in the probability to attach to some prediction or diagnosis given statistical knowledge about past cases. The list goes on.

Unfortunately, this potential of probabilistic logics has not yet been exploited. There are a number of reasons for this. First, current probabilistic logics are a disparate bunch—it is hard to glean commonalities to see how they fit into a general framework, and hard to see how a solution to the general problem of probabilistic logic would specialise to each individual logic (Williamson, 2002, 2009b, §21). Second, probabilistic logics are often hard to understand: while probabilistic reasoning is well understood and so is logical reasoning, when these two components interact in formalisms that combine them, their complexities compound and a great deal of theoretical work is required to determine their properties. Third, probabilistic logics are often thwarted by their computational complexity. While they may integrate probability and logic successfully, it may be very difficult to determine an answer to a question such as that of Schema (1.1). Sometimes this is because a probabilistic logic seeks more generality than is required for applications; but often it is no fault of the logic—probabilistic and logical reasoning are both computationally infeasible in the worst case, and their combination is no more tractable.

### 1.3 Overview of the Book

In this book we hope to address some of these difficulties. In Part I we show how a range of alternative inferential procedures fit into a general framework for probabilistic logic. We will cover the standard probabilistic semantics for probabilistic logic in §2, in §3 the support-possibility approach of the probabilistic argumentation framework, evidential probability in §4, inference involving statistical hypotheses in §5 and §6, and objective Bayesian epistemology in §7. The background to each procedure will be discussed in §X.1, where  $X$  ranges from 2 to 7; note that §2.1 contains prerequisites for the other sections and should not be skipped on a first reading. In §X.2 we show how a key question of each inferential procedure can be viewed as a question of the form of Schema (1.1). In §X.3 we will show the converse, namely that each inferential procedure can be viewed as providing semantics for the entailment relation  $\approx$  found in this schema.

We permit a generic notion of entailment  $\approx$  which is weaker than that of classical logic. Generally, the entailment  $\varphi_1^{X_1}, \dots, \varphi_n^{X_n} \approx \psi^Y$  holds iff all models of the left-hand side satisfy the right-hand side, where suitable notions of *model* and *satisfy* are filled in by the semantics in question. We say that a semantics for the entailment relation yields a *probabilistic logic* if (i) models are probability functions (satisfying certain conditions that are specified by the semantics) and (ii) probability function  $P$  satisfies  $\psi^Y$  iff  $P(\psi) \in Y$ . In this book we distinguish between *non-monotonic* and *monotonic* entailment relations. Monotonicity holds where  $\varphi_1^{X_1}, \dots, \varphi_n^{X_n} \approx \psi^Y$