JON WILLIAMSON AND DAVID CORFIELD

INTRODUCTION: BAYESIANISM INTO THE 21ST CENTURY

1 BAYESIAN BELIEFS

Bayesian theory now incorporates a vast body of mathematical, statistical and computational techniques that are widely applied in a panoply of disciplines, from artificial intelligence to zoology. Yet Bayesians rarely agree on the basics, even on the question of what Bayesianism actually is. This book is about the basics — about the opportunities, questions and problems that face Bayesianism today.

So what is Bayesianism, roughly? Most Bayesians maintain that an individual's degrees of belief ought to obey the axioms of the probability calculus. If, for example, you believe to degree 0.4 that you will be rained on tomorrow, then you should also believe that you will not be rained on tomorrow to degree 0.6. Most Bayesians also maintain that an individual's degrees of belief should take prior knowledge and beliefs into account. According to the *Bayesian conditionalisation* principle, if you come to learn that you will be in Manchester tomorrow (m) then your degree of belief in being rained on tomorrow (r) should be your previous conditional belief on r given m: $p^{t+1}(r) = p^t(r|m)$. By *Bayes' theorem* this can be rewritten $p^t(m|r)p^t(r)/p^t(m)$.[1]

Although Bayesianism was founded in the eighteenth century by Thomas Bayes[2] and developed in the nineteenth century by Laplace,[3] it was not until well into the twentieth century that Frank Ramsey[4] and Bruno de Finetti[5] provided credible justifications for the degree of belief interpretation of probability, in the shape of their *Dutch book* arguments. A Dutch book argument aims to show that if an agent bets according to her degrees of belief and these degrees are not probabilities, then the agent can be made to lose money whatever the outcome of the events on which she is betting. Already by this stage we see disagreement as to the nature of Bayesianism, centring on the issue of *objectivity*. De Finetti was a *strict subjectivist*: he believed that probabilities only represent degrees of rational belief, and that an agent's belief function is rational just when it is a probability function — no further constraints need to be satisfied.[6] Ramsey, on the other hand, was a *pluralist* in that he also accepted objective frequencies. Further, he advocated a kind of calibration between degrees of belief and frequencies:

[1] [Howson & Urbach, 1989; Earman, 1992] and [Gillies, 2000] are good introductions to Bayesian thought.

[2] [Bayes, 1764].

[3] [Laplace, 1814].

[4] [Ramsey, 1926].

[5] [de Finetti, 1937].

[6] See Galavotti's paper in this volume.

> Thus given a single opinion, we can only praise or blame it on the ground of truth or falsity: given a habit of a certain form, we can praise or blame it accordingly as the degree of belief it produces is near or far from the actual proportion in which the habit leads to truth. We can then praise or blame opinions derivatively from our praise or blame of the habits that produce them.[7]

Such a view may be called *empirical Bayesianism*: degrees of belief should be calibrated with objective frequencies, where they are known.[8] Ramsey was cautious of too close a connection because of the reference class problem: Bayesian probabilities are single-case, defined over sentences or events, whereas frequencies are general-case, defined over classes of outcomes, and there may be no way of ascertaining which frequency is to be calibrated with a given degree of belief. The Principal Principle of [Lewis, 1980] aims to circumvent this problem by offering an explicit connection between degrees of belief and objective *single-case* probabilities. De Finetti shows that in certain circumstances, if degrees of belief are *exchangeable* then they will automatically calibrate to frequencies as Bayesian conditionalisation takes place.[9]

John Maynard Keynes advocated *logical Bayesianism*: a probability $p(b|a)$ is the degree to which a partially entails b, and also the degree to which a rational agent should believe b, if she knows a.[10] Thus for Keynes probability is truly objective — there is no room for two agents with the same knowledge to hold different belief functions yet remain perfectly rational. Moreover probability is fixed not by empirical frequencies but by logical constraints like the *principle of indifference*, which says that if there is no known reason for asserting one out of a number of alternatives, then all the alternatives must be given equal probability. There are problems with the principle of indifference which crop up when there is more than one way of choosing a suitable set of alternatives, but the *maximum entropy principle*, ardently advocated by Edwin Jaynes,[11] has been proposed as a generalisation of the principle of indifference which is more coherently applicable.

Empirical and logical Bayesianism may be grouped together under the banner of *objective Bayesianism*. Objective Bayesians may adopt a mixed approach: for example Rudolf Carnap had a position which incorporated both empirical and logical constraints on rational belief.[12] Objective Bayesians disagree with a strict subjectivist like de Finetti, since they claim that it is not sufficient that a belief function satisfies the axioms of probability — it must satisfy further constraints before it can be called rational. But objective Bayesianism harbours many views and proponents often disagree as to which extra constraints must be applied. Also, unlike Keynes many objective Bayesians accept

[7] [Ramsey, 1926] 51.
[8] See [Dawid, 1982].
[9] [de Finetti, 1937]. See also [Gaifman & Snir, 1982].
[10] [Keynes, 1921].
[11] [Jaynes, 1998].
[12] [Carnap, 1950], [Carnap & Jeffrey, 1971].

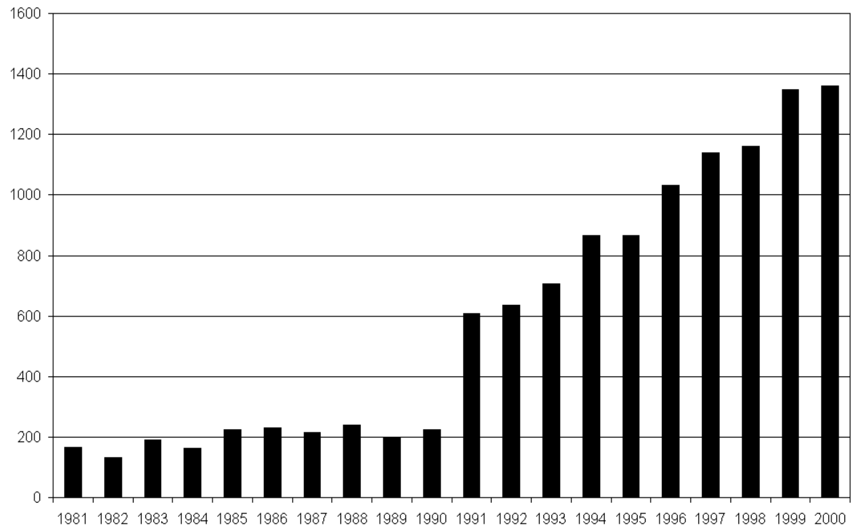

Figure 1. Number of Bayesian articles by year.

that in some situations there may be more than one rational probability function — two rational agents may have the same background knowledge but different belief functions.[13]

The question of objectivity remains an important issue for Bayesians today, and one that will crop up in several papers in this book.

2 BAYESIANISM TODAY

The last decade of the twentieth century has witnessed a dramatic shift in the profile of Bayesianism. Bayesianism has emerged from being thought of as a somewhat radical methodology — for enthusiasts rather than research scientists — into a widely applied, practical discipline well-integrated into many of the sciences. A search of the Web of Science database for articles whose subject contains the word or prefix 'Bayes' shows a dramatic upturn in the number of Bayesian papers in the 1990s — see Figure 1. A search for Bayesian books on the British library catalogue tells a similar story, as do other searches,[14] and the rise in the number of Bayesian meetings and the success of new organisations like the International Society for Bayesian Analysis[15] provide further evidence.

[13] See [Williamson, 1999] and the paper of Paris and Vencovská in this volume.
[14] [Berger, 2000] §2.1.
[15] ISBA was established in 1992. See www.bayesian.org.

This renaissance has occurred largely thanks to computational and sociological considerations. The calculation of the posterior probability of a hypothesis given data can require, via Bayes theorem, determining the values of integrals. These integrals may often have to be solved using numerical approximation techniques, and it is only recently that computers have become powerful enough, and the algorithms efficient enough, to perform the integrations. The sociological changes have been on two main fronts. First, scientific researchers, who are usually taught to present their work as objectively as possible, were often discouraged from applying Bayesian statistics because of the perceived irreducible subjectivity of Bayesianism. This has changed as objective Bayesian techniques have become more popular. Second, Bayesian statistics has to a certain extent unified and absorbed classical techniques. Any religion worth its salt absorbs the gods of its competitors, and 'Bayesianity' is no different:[16] the diverse and seemingly unrelated techniques of classical statistics have been viewed as special-case approximations to Bayesian techniques, and Bayesianism has been invoked to shed light on the successes as well as the failures of classical statistics.[17] Present-day statistics is often a half-way house between the classical and Bayesian churches: increasingly one finds that Bayesian techniques are used to select an appropriate statistical model, while the probabilities within the model are tacitly treated as being objective.

In the field of artificial intelligence (AI) Bayesianism has been hugely influential in the last decade. Expert systems have moved from a logical rule-based methodology to probabilistic techniques, largely involving the use of Bayesian networks.[18] Statistical learning theory has helped integrate machine learning techniques into a probabilistic framework,[19] and Bayesian methods are often now used to ascertain the parameters of machine learning models, and to determine the error between model and data.[20] Applications in industry have followed quickly: Bayesian networks are behind several recent expert systems including the print trouble-shooter of Microsoft's Windows '95 (and, alas, the paperclip of Office '97);[21] Bayesian reasoning is widely implemented using neural networks, forming the core of Autonomy's software for dealing with unstructured information (which made Autonomy's director, Mike Lynch, Britain's first dollar-billionaire);[22] other graphical models also form the basis of applications of Bayesian statistics to med-

[16]The almost religious fervour with which Bayesians pursue the cause of Reverend Bayes, and with which non-Bayesians undergo the conversion to Bayesianism, has occasionally been noted. Jaynes appears to have coined the term 'Bayesianity'.

[17][Jaynes, 1998].

[18]See [Pearl, 1988] and the website of the Association for Uncertainty in AI at www.auai.org.

[19][Vapnik, 1995].

[20]See for example [Bishop, 1995], [Jordan, 1998] and Williams' paper in this volume.

[21]See research.microsoft.com/dtas/ and [Horvitz *et al.*, 1998].

[22]See the technology white paper at www.autonomy.com. Peter Williams reported at the conference Bayesianism 2000 that neural network based Bayesian reasoning also proved successful (and lucrative!) when applied to gold prospecting.

ical expert systems[23] and health technology assessment.[24]

These developments in AI and other sciences have stimulated work on more traditional philosophical issues. Bayesian networks integrate causality and probability in a particular way, and the question naturally arises as to how exactly Bayesian probability is related to causality, and whether techniques for learning Bayesian networks from data can be applied to the problem of discovering causal structure.[25] Probability logics and their AI implementations have prompted renewed investigations into the relationship between Bayesian probability and logic.[26] Objective Bayesian methods, often involving the use of the maximum entropy principle, have been successfully applied in physics,[27] and this has led to debate about the validity of objective Bayesianism[28] and further applications of maximum entropy.[29] Probabilistic decision-theoretic techniques have now been widely adopted in economics, and this has stimulated research in the foundations of Bayesian decision theory.[30] On the other hand, the application of Bayesianism to scientific methodology may lead to a corresponding application to mathematical methodology.[31]

In the context of this recent Bayesian upswell, it is all the more important to avoid complacency: criticisms of Bayesianism must be given due attention,[32] and the key messages of the early proponents of Bayesianism must be better understood.[33]

3 PROSPECTS FOR BAYESIANISM

Judging by the papers in this book, the future of Bayesianism will depend on progress on the following foundational questions.

- Is Bayesianism to be preferred over classical statistics?
- If so, what type of Bayesianism should one adopt — strict subjectivism, empirical objectivism or logical objectivism?
- How does Bayesian reasoning cohere with causal, logical, scientific, mathematical and decision-theoretic reasoning?

---

[23] [Spiegelhalter *et al.*, 1993].

[24] [Spiegelhalter *et al.*, 2000].

[25] See [Spirtes *et al.*, 1993], [McKim & Turner, 1997], [Hausman & Woodward, 1999], [Hausman, 1999], [Glymour & Cooper, 1999], [Pearl, 2000] and Pearl's, Dawid's and Williamson's papers in this volume.

[26] See [Williamson, 2000] and the papers of Cussens, Gabbay, Howson, and Paris and Vencovská in this volume.

[27] [Jaynes, 1998].

[28] See Howson's and Paris and Vencovská's papers.

[29] See Williamson's paper.

[30] See the papers of Mongin, McClennen, Bradley and Albert in this volume.

[31] See Corfield's paper in this volume.

[32] See the papers of Mayo and Kruse, Albert and Gillies.

[33] See Galavotti's paper.

These questions are, of course, intricately linked. The first two are well-worn but extremely important: much progress has been made, but it would be foolhardy to expect any conclusive answers in the near future. The last question is particularly pressing, given the recent applications of Bayesian methods to AI. AI is now faced with a confusing plethora of formalisms for automated reasoning, and unification is high on the agenda. If Bayesianism can provide a framework into which AI techniques slot then its future is guaranteed.

4 THIS VOLUME

The fifteen chapters of this book have been arranged in four parts. The first of these parts is entitled 'Bayesianism, Causality and Networks' and consists of four chapters. What unites the authors of the first three contributions is an eagerness to clarify the relationship between causal and probabilistic reasoning, two of them by way of the use of directed acyclic graphs. The author of the fourth chapter, on the other hand, reports on research on a different category of network — neural networks. In the opening chapter, Pearl proceeds from the fundamental idea of Bayesianism that we should integrate our background knowledge with observational data when we reason. He then argues that our everyday and scientific knowledge is largely couched in causal, rather than statistical, terms, and that as such it is not readily expressible in probabilistic terms. Now, clearly it would preferable to be able to feed background knowledge directly into our reasoning calculus, and so, if possible, we should devise a new mathematical language in which we can represent causal information and reason about it. The article advertises Pearl's exciting new research programme, detailed in his book 'Causality', whose central aim is the mathematisation of causality via directed graphs.[34] The key questions to be addressed then concern the benefits of adopting such a radically new language and the safety of the reasoning it warrants. Pearl himself says that is possible to cast his causal models in terms of probabilities using hypothetical variables, but then argues that the only purpose in doing so is to avoid confrontation with the consensus position in the statistics community, which sees no limitations to the expressiveness of probability theory. Indeed, for Pearl, there is a definite disadvantage in a choice of language which gives counterfactual propositions precedence over more readily comprehensible causal ones.

So Pearl's idea is that the previous failure to construct a mathematical system capable of integrating background causal knowledge has led to much of this most important way of encoding our beliefs about the world being overlooked. As such he has located a novel way in which we may take the Bayesian to be failing to act in as rational as possible a manner. On the other hand, a long-standing complaint of irrationality made against Bayesianism, one which will recur through the chapters of this volume, alleges that a Bayesian's tenets do not force her to test whether her degrees of belief are, in some sense or other, optimal. These two themes intertwine

[34] [Pearl, 2000].

in Dawid's article. Dawid is well known for his 'Popperian' Bayesianism which aims to assess an agent's degrees of belief by a process of calibration, where, for example, weather reporters are to be congratulated if it rains on roughly 30 percent of the occasions they give 0.3 as the probability that it will rain. This concern with testability recurs in Dawid's contribution to the volume. While he agrees with Pearl that statisticians have largely ignored causality and have been wrong to do so, still he finds some elements of Pearl's new thinking problematic. What is at stake here is the Popperian belief that anything worthy of scientific consideration is directly testable. For Dawid some of the counterfactual reasoning warranted by Pearl's calculus (and by statisticians adopting other schemes, such as Rubin's potential-outcome approach) just is untestable. An example illustrating this key difference between Dawid and Pearl is their respective treatments of counterfactual questions such as whether my last headache would have gone had I not taken an aspirin, given that I did take one and it did go. How should knowledge of the effects of aspirin on other headache incidents of mine bear on this question? Pearl says that, without evidence to the contrary, we should presume that such knowledge does have a bearing on the counterfactual statement. By contrast, Dawid claims that singular counterfactual statements are untestable and therefore should not be accepted by the scientifically minded.[35] For Dawid what may be justifiably said about counterfactuals does not involve their essential use.

As Dawid is a self-professed Popperian, a comparison that comes to mind is to think of Pearl as a Lakatosian. While Popper's philosophy allowed that metaphysical principles might guide the generation of novel scientific theories, thereby restoring some worth to them after the Logical Positivists had dismissed them as 'meaningless', still they accrued no further value even when those theories passed severe tests. Where Lakatos went further than Popper was to allow metaphysics to be an integral part of a research programme, which was to be assessed by its theoretical and empirical success as a whole. Similarly, we could say that Pearl has devised a research programme with a powerful heuristic and a new mathematical language. There is a metaphysical belief on Pearl's part in the regularity of a world governed by causal mechanisms which is integrated into this programme, hence his turn to structural equation models. Dawid, meanwhile, views the presuppositions behind the use of these models as unwarranted — the world for him is not so easily tamed.

In the third chapter Williamson questions the validity of the causal Markov condition, an assumption which links probability to causality and on which the theory of Bayesian networks and Pearl's recent account of causality depends. He argues that the causal Markov condition does not hold for an empirical account of probability, or for a strict subjectivist Bayesian interpretation, but does hold for an objective Bayesian interpretation, i.e., one using maximum entropy methods. If it can be established that the causal Markov condition does not hold with respect

[35]See the comments and rejoinder to Dawid's paper in the Journal of American Statistical Association 95 (June 2000), pages 424-448.

to a notion of empirical probability, this means that causal networks must be restructured if they are to be calibrated with frequency data. This leads Williamson to propose a two-stage methodology for using Bayesian networks: first build the causal network out of expert knowledge and then restructure it to fit observational data more closely. The validity of stage 1 of this methodology depends on the validity of a maximum-entropy based objective Bayesian interpretation of probability and so would not appeal to subjectivists like de Finetti or Howson (see below), while the validity of stage 2 depends on acceptance of the idea that one ought to calibrate Bayesian beliefs with empirical data.

Williams rounds out Part 1 of the book by offering us an overview of research carried out by the neural network community to provide a principled way of using data to fashion an accurate network. All forms of machine learning must find a way to reconcile the demands of accuracy and the risks of overfitting data. This relates to a long-standing debate in the philosophy of science about the desirability of choosing as simple as possible a model to represent empirical data. Now, some Bayesians, including those working in the tradition of Harold Jeffreys, claim to have found a principled way to effect this reconciliation by according a higher prior probability to a model with fewer free parameters. The potential for increased accuracy provided by an extra parameter will then be balanced by a lower prior probability for the more complicated model. Neural network researchers are now invoking these Bayesian notions to arrive at optimal network configurations and settings of connection strengths.

A frequently encountered point of disagreement between the different approaches to artificial intelligence concerns the need to represent data and inference in propositionally encoded form. Neural networks come in for criticism for acting like black boxes. They may work well in many situations, the thought is, but we do not really understand why. Thus, unlike in the case of Bayesian networks, they offer no insight to the expert hoping to use them to support decision-making. Of course, one might respond to this criticism by making the point that accuracy, not transparency, is the most important quality of a decision-making process, especially in critical situations such as medical diagnosis. In the context of Williams' chapter the lack of transparency relates to the fact that the space of weight configurations of a network bears no straightforward relation to an expert's qualitative understanding of a domain. Thus background knowledge cannot be encoded directly into a prior distribution over possible networks, but only through the mediation of real or simulated data. Perhaps this difficulty is the reason that we find such a great range of techniques employed by the neural network community, even though, in the case of the ones described by Williams at least, Bayesian principles are guiding them.

We turn next to the second part — Logic, Mathematics and Bayesianism. Here the five authors wish to investigate the relationship between Bayesian probabilistic reasoning and deductive logic. In two chapters (Howson, Paris & Vencovská) we find probability theory presented as an extension of deductive logic, while in two others (Galavotti, Corfield) it appears in the guise of realistic personalist degrees of belief. Finally, Cussens discusses his use of stochastic logic programming, an artificial intelligence technique, to encode probabilistic reasoning.

Howson views Bayesianism as an extension of deductive logic in the sense that, just as the use of deductive logic provides rules to ensure a consistent set of truth values for the statements of language, so the probability theory axioms ensure consistent degrees of belief. In doing so he rules out three widely held, yet disputed, aspects of Bayesian reasoning: its inextricable link to utility theory; the principle of indifference, along with any other notion of objective priors; and, conditionalisation. Justification of this logical core of Bayesianism is provided by the idea of probability as expected truth value, using the device of the indicator function of a proposition, where a proposition is taken *à la* Carnap as the set of structures in which a sentence is true. Howson's belief that probability theory is a form of logic sets him against decision theorists, such as Herman Rubin, who believe that 'you cannot separate probability from utility'.[36] Thus he aims to provide a justification for the probability axioms foregoing the use of Dutch Book arguments, thereby avoiding reliance on the notion of the desirability of acquiring money.

Howson also rejects the strain of Bayesianism which hopes to arrive at some values to enter into the probability calculus through the use of the principle of indifference or of maximum entropy. More radically still, he continues by arguing that conditionalisation has no place in a Bayesian logic, since it is a rule relating

[36] [Rubin, 1987].

truth values held at different times. He illustrates this thesis in parallel deductive terms: if you held 'A implies B' to be true yesterday, then find out today that A is true, you are not now forced to accept B, since you may no longer believe that A implies B. Similarly, if yesterday you have $p(A|B) = x$, and today $p'(B) = 1$, this does not mean you need have $p'(A) = x$. Here the reader might wonder about the status of the commonly held notion that, unless you have good reason for this change of heart, you should stick to your original beliefs. Is it just an extra-logical rule of thumb that $p'(A|B) = p(A|B)$ unless there is good cause to change one's mind?

Galavotti has provided a largely historical piece on the Bayesianism of Bruno de Finetti. De Finetti is famous for his assertion that 'probability does not exist', preferring to see probabilities as subjective degrees of belief, rather than something inherent in the universe. But while he was keen to stress his disapproval of an objectivism which sees probabilities as simply out there in the world, this did not entail a disregard for objectivity. Empirical frequency data might be integrated into one's degrees of belief by the subjective judgement of the exchangeability of the data sequence. Moreover, and this may be a surprise for readers who share the commonly held impression that de Finetti was the arch-subjectivist Bayesian, he had a considerable interest in scoring rules used to judge the success of one's personal probability assignments. Comparisons of the accuracy of one's own previous probability judgements with those of others were to be integrated into one's current personal degrees of belief.

Corfield bases his paper on the ideas of the Hungarian mathematician George Pólya, who in his description of plausible mathematical reasoning, which he interpreted by means of probabilistic degrees of belief, discerned what he took to be the common patterns of everyday reasoning. Corfield argues that no attempt to construe mathematical reasoning in Bayesian terms can assume logical omniscience — the requirement that rational agents accord the same degree of belief to any two logically equivalent statements. In the absence of this principle, logical and mathematical learning become thinkable in Bayesian terms. The idea that Bayesians should put logical and empirical learning on an equal footing goes back at least as far as de Finetti, and would seem to set Corfield against Howson who frames his Bayesian logic in such a way that logical omniscience comes already built in.

One could argue that a Bayesian reconstrual of mathematical reasoning as it occurs in practice is likely to be a largely empty exercise. Certainly, Bayesian reconstructions of scientific reasoning have come in for this kind of criticism. One may be able to explain why observing a white tennis shoe provides no support for the law 'all ravens are black', despite being an instance of the logically equivalent 'all non-black things are not ravens', these critics say, but it offers very little by way of insight into the rationality of decision making in science. However, one might reply that it has led Corfield to consider the rationality of certain overlooked styles of mathematical reasoning: use of analogy, choice of proof strategy, large scale induction. Regarding the latter, for instance, to date very little attention has been paid by philosophers of mathematics to the rationality of mathematicians raising their degrees of belief in conjectures due to confirmations. For example, should the computer calculation which shows that the first 1.5 billion nontrivial zeros of the Riemann zeta function have real part equal to $\frac{1}{2}$ be thought to lend support to the Riemann hypothesis, which claims that all of the infinitely many zeros lie on this line in the complex plane?

Paris and Vencovská share Howson's vision of probability theory as a logic, but unlike him they seek to isolate and justify principles which will allow the agent to select her priors rationally. In an earlier paper[37] they showed that the probability function which maximises entropy is the only choice if certain intuitively plausible constraints on objective Bayesian reasoning are to be respected. This was a significant result, but with one drawback: in their framework background knowledge is assumed to be encapsulated in a set of linear constraints. This rules out knowledge of, say, independencies amongst variables. In this chapter Paris and Vencovská extend their result to deal with non-linear constraints in the agent's background knowledge. There is now some room for subjectivity since there may be more than one most rational (i.e., maximum entropy) probability function. A point to note is that the framework adopted here is in the propositional calculus. This may be adequate for many AI applications, but it is not clear how it could be extended to the predicate calculus. If different reasoning principles are required for predicate reasoning, how does the resulting formalisation cohere with the propositional approach given here?

As uncertainty is now treated probabilistically by the majority of AI practitioners, those adopting a logic based approach who wish to discuss uncertain reasoning are faced with the thorny problem of integrating logic and probability. Philosophers have worked hard on this problem for many years with no consensus emerging. The line of thought that takes Bayesianism to be an extension of deductive logic would suggest that this should not be very problematic for a degree of belief interpretation of probability. However, this has not turned out to be the case — a very large number of disparate techniques have proposed by the AI community. Cussens bases his attempt to integrate probability theory and logic on what are called 'stochastic logic programs'. Stochastic logic programs (SLPs) originated in the inductive logic programming (ILP) paradigm of machine learning.[38] When presented with data, an ILP program will attempt to generate a logic program (essentially, a set of Horn clauses) which includes as successful goals as many positive examples as possible, while excluding as many negative examples. In cases where only positive examples are available, a common situation in science, to prevent overfitting, it was found necessary to generate a distribution over all possible ground instances. Muggleton did this by labelling the clauses of a proposed logic program with probabilities generated from the data. Elsewhere, Cussens has extended this idea to apply it to natural language processing, where a successful parsing of a sentence will be accorded a probability depending on the

[37] [Paris & Vencovská, 1990].
[38] [Muggleton & de Raedt, 1994].

ways it may be generated by the grammar encoded by the logic program. In the present article, he takes SLPs to be capable of representing a very wide range of AI techniques, in particular showing how Bayesian networks may be encoded in its terms. He then compares his SLP approach to other techniques.

The reader might be interested to know of two other approaches to the integration of logic and Bayesian networks. Williamson develops ‘logical Bayesian networks’ (as opposed to causal Bayesian networks) whose nodes are sentences (rather than causes and effects) and whose arrows correspond to the logical implication relation (rather than the causal relation).[39] Meanwhile, Dov Gabbay is working on a way of representing Bayesian networks in the framework of his labelled deductive systems. His results were not ready in time for this volume, but they will appear in the near future.

Turning now to the third part we find the contributions of three Bayesian decision theorists. Probabilistic decision theory has a long heritage, stretching back to Pascal’s Wager, but there still rage many disputes over its fundamental principles. Here, two of the contributors, Mongin and McClennen, scrutinise the acceptability of particular axioms, while in the first chapter of this part Bradley discusses the problem of the measurement of belief.

Bradley’s claim is that the resources for resolving the issue of how to assess the strengths of beliefs and desires of an agent are to found in the writings of Ramsey from the 1920s. While decision theorists have followed the lead of Savage, Ramsey has largely been overlooked. However, as Bradley points out, Savage relied on the assumption of state-independent utility, where the desirability of an outcome is independent of the state of the world in which it occurs. This assumption has come in for a great deal of criticism, which has given rise to highly complex theories of state-dependent utility. Bradley argues that if we revive Ramsey’s notion of ‘ethically neutral events’, ones to whose outcome the agent is indifferent, we gain a means to access the strength of an agent’s beliefs and desires without the need to invoke these complex theories.

The Independence Principle, and the closely related Sure-Thing Principle, are central to Bayesian decision theory. The independence principle states that if an agent shows no preference between two gambles, P and P', then, for any $0 < \alpha \leq 1$ and any gamble Q, she will also show no preference between the composite gambles $R = \alpha P + (1-\alpha)Q$ and $R' = \alpha P' + (1-\alpha)Q$. While it appears to be a highly plausible principle, McClennen investigates various arguments put forward to support it, both directly and via the Sure-Thing Principle, and finds them all wanting. One might have supposed that the independence principle holds, since these composite gambles are disjunctive in the sense that in the case of R the final outcome will either be the outcome of P or the outcome of Q but not both. Still, McClennen argues, there may be an interactive effect making the agent prefer R to R'. This may occur, for instance, if Q more closely resembles P than P' and

[39] [Williamson, 2001]. See also [Williamson, 2000] where these logical networks form the basis of a proof theory of a probabilistic logic.

the agent has a preference for a less varied composite gamble.

In the final chapter of Part 3, Mongin takes on the task of examining the difficulties created by the simultaneous assumption of Bayesian and Paretian principles. The latter refer to assumptions about preferences of outcomes in the light of group consensus about preferences. For example, for a group of experts working with different utilities or different probabilities in the framework of state-independent utility theory, there will not in general be a way to select a utility function and probabilities such that one outcome is preferred over another whenever all the experts agree to this ordering. This result lends support to the move towards state-dependent utility theory mentioned above. However, a pure form of this theory entails the undesirable consequence that subjective probabilities are not in general uniquely determined. Mongin then proceeds to scrutinise a form of state-dependence which entails unique probabilities. He shows, however, that the assumptions of this theory still conflict with Paretian principles.

The fourth and final part consists of three contributors' criticisms of Bayesianism. As the chapters up to this point amply demonstrate, Bayesians disagree amongst themselves about all manner of issues: the extent of rationality constraints, the link to utility theory, the role of conditionalisation, etc. This being so, the critic's task is made harder. Whatever principle she attacks, some Bayesian may claim not to hold to it.

Albert's criticism is aimed at the use of Bayesian principles by decision theorists. Adopting a line reminiscent of Popper's critical attitude towards psychoanalysis, he claims that for the Bayesian 'there is no such thing as irrational behavior' — any set of actions can be construed as satisfying the constraints of Bayesianism. Albert argues for this conclusion by discussing a situation involving a chaotic clock which outputs a sequence of 0s and 1s. The agent must judge the likelihood of these digits occurring, based on the sequence to date, to help him win as much money as possible. What Albert shows is that, even with the agent's utility function given, whatever he does one can reconstruct it as rational according to some choice of prior distribution over the hypothesis space. Now, in response one may argue that the chaotic clock situation does not resemble the everyday conditions met with in economic life, but Albert argues that his example is sufficiently generic in this sense. Objective Bayesians may also claim that knowledge of the chaotic clock set up provides the rational agent with an obvious unique choice of prior. On the other hand, subjectivists might question whether the ways of recording the agent's behaviour are sufficient to pick up fully his belief structure. For example, the chaotic clock situation would not allow you to discover the incoherence of an agent who is certain that 0 will appear next, but also certain that 1 will appear next.

In his chapter, Gillies wants to argue that Bayesianism is appropriate only in a restricted range of situations, indeed, 'only if we are in a situation in which there is a fixed and known theoretical framework which it is reasonable to suppose will not be altered in the course of the investigation'. As soon as the reasoner departs from the current theoretic framework, Bayesianism is of no assistance. In saying this, Gillies appears to be aligning himself with one side of an argument heard before in the philosophy of science about the room for Bayesian principles to operate when a new theory is proposed. Earman, for instance, who is very sympathetic to Bayesianism, argues that changes of conceptual framework require the resetting of priors, which occurs in a non-algorithmic fashion by plausibility arguments.[40] He also indicates that these exogenous redistributions of prior probabilities will occur frequently, perhaps within the course of an everyday conversation. Now, Howson's Bayesianism might have no problem with this — after all it is not a diachronic theory — but Bayesians of a different stripe, wishing to salvage a substantial role for conditionalisation, might prefer to concede that the advent of novel ideas will have their effects not through conditionalisation, but still look to 'normal' science to see Bayes' theorem at work. However, the two examples put forward by Gillies could hardly be considered revolutionary changes. Rather they appear to involve the kind of reasoning that any statistician will have to perform in the course of their work, i.e., the assessment of the validity of the current model. The indication is that where the error statistician is always eager to challenge the current model and put it to severe tests, the Bayesian has neither the means nor the incentive to look beyond the current framework. It is true that Dutch Book arguments by themselves do not require an agent to take the slightest trouble to make an observation or to challenge a modelling assumption that would be beneficial to a bet they are making. However, under reasonable assumptions, one can show that it is always worth seeking cost-free information before making a decision. It is therefore not surprising to find Bayesian statisticians engaging in what Box and Tiao call 'model criticism'.[41] Readers may care to see how a Bayesian statistician works on a very similar problem to Gillies' second example in [Gelman *et al.*, 1995], 170-171.

In their article, Mayo and Kruse take on Bayesian statistics on the issue of stopping rules. For the error statistician the conditions stipulated before the start of an experiment as to when it will be deemed to have ended will usually be relevant to the significance of the test. For instance, when testing for a proportion in some population, even if the data turns out the same, it makes a difference whether it has been generated by deciding to stop after a fixed number of positive cases have been observed or whether it has been generated by deciding to stop after a fixed number of trials. For most Bayesians (see [Box & Tiao, 1973], 44-46 for an exception), on the other hand, acceptance of the likelihood principle entails that such considerations should play no part in the calculation of their posterior distributions. This marks a very significant difference between the schools. *Pace* Gillies, as Mayo has said elsewhere, 'one cannot be just a little bit Bayesian'. Most Bayesians would agree.

The Bayesian position has a considerable plausibility to it. Should we condemn an experimenter who intends to test 100 cases, finds half way through that the proportion of positive cases is low, decides then to wait for 10 such cases to occur, which duly happens after precisely 100 trials, and then writes up the experiment as

[40] [Earman, 1992].

[41] [Box & Tiao, 1973].

originally planned? Are all experiments called off early because of funding problems worthless? There appears to be something magical occurring in that the 'real' intentions of the experimenter make a difference. On the other hand, as critics of the Bayesian indifference to stopping rules, Mayo and Kruse need only point out the problematic consequences of operating with a single rule of their choosing. What then if an experimenter in a binomial situation with, say, $p = 0.2$ is to continue testing until this value of p becomes unlikely to a specified degree, putting some upper limit on the number of trials to make it a proper rule? Here at least the Bayesian can provide bounds for the likelihood of a test achieving this end. But Mayo and Kruse go on to discuss an experimental situation where due to the stopping rule the Bayesian statistician will necessarily reason to a foregone conclusion on the basis of the likelihood principle. This type of situation arises when an improper prior, failing to satisfy countable additivity, is employed. One could, of course, maintain that countable additivity be enforced, but improper priors are commonplace in the Bayesian literature. Mayo and Kruse present several quotations revealing that some Bayesian statisticians are struggling to come to terms with this apparent paradox. Readers who wish to read more on this topic may well enjoy the discussion of this phenomenon by the Bayesians Kadane, Schervish and Seidenfeld.[42] Not wishing to forego the use of improper priors, these authors consider that further work is required to ascertain when their use is admissible.

King's College, London.

BIBLIOGRAPHY

[Bayes, 1764] Thomas Bayes. An essay towards solving a problem in the doctrine of chances. *Philosophical Transactions of the Royal Society of London*, **53**, 370–418, 1764.

[Berger, 2000] James O. Berger. Bayesian analysis: a look at today and thoughts of tomorrow. *Journal of the American Statistical Association*, **95** (December), 2000.

[Bishop, 1995] Christopher M. Bishop. *Neural networks for pattern recognition*. Oxford University Press, 1995.

[Box & Tiao, 1973] G. Box & G. Tiao. *Bayesian Inference in Statistical Analysis*. Addison-Wesley, 1973.

[Carnap, 1950] Rudolf Carnap. *Logical foundations of probability*. Routledge & Kegan Paul Ltd, 1950. Second edition 1962.

[Carnap & Jeffrey, 1971] Rudolf Carnap & Richard C. Jeffrey(eds.). *Studies in inductive logic and probability*, Volume 1, University of California Press, 1971.

[Dawid, 1982] A.P. Dawid. The well-calibrated Bayesian. With discussion, *Journal of the American Statistical Association*, **77**, 604–613, 1982.

[Earman, 1992] John Earman. *Bayes or bust?*. M.I.T. Press, 1992.

[de Finetti, 1937] Bruno de Finetti. Foresight. Its logical laws, its subjective sources. In [Kyburg & Smokler, 1964], 53–118, 1937.

[Gaifman & Snir, 1982] H. Gaifman & M. Snir. Probabilities over rich languages. *Journal of Symbolic Logic*, **47**, 495–548, 1982.

[Gelman *et al.*, 1995] Andrew B. Gelman, John S. Carlin, Hal S. Stern & Donald B. Rubin. *Bayesian data analysis*. Chapman & Hall/CRC, 1995.

[Gillies, 2000] Donald Gillies. *Philosophical theries of probability*. Routledge, 2000.

[42] [Kadane *et al.*, 1999] §§3.7 & 3.8.

[Glymour & Cooper, 1999] Clark Glymour & Gregory F. Cooper(eds.). *Computation, causation, and discovery*. M.I.T. Press, 1999.

[Hausman, 1999] Daniel M. Hausman. The mathematical theory of causation. Review of [McKim & Turner, 1997], *British Journal for the Philosophy of Science*, **50**, 151–162, 1999.

[Hausman & Woodward, 1999] Daniel M. Hausman & James Woodward. Independence, invariance and the causal Markov condition. *British Journal for the Philosophy of Science*, **50**, 521–583, 1999.

[Horvitz *et al.*, 1998] Eric Horvitz, Jack Breese, David Heckerman, David Hovel & Koos Rommelse. The Lumière Project: Bayesian user modeling for inferring the goals and needs of software users. In *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence*, Morgan Kaufmann, pages 256–265, 1998.

[Howson & Urbach, 1989] Colin Howson & Peter Urbach. *Scientific reasoning: the Bayesian approach*. Open Court, 1989. Second edition, 1993.

[Jaynes, 1998] E.T. Jaynes. *Probability theory: the logic of science*. http://bayes.wustl.edu/etj/prob.html.

[Jordan, 1998] Michael I. Jordan(ed.). *Learning in Graphical Models*. MIT Press, Cambridge, MA, 1998.

[Kadane *et al.*, 1999] J. Kadane, M. Schervish & T. Seidenfeld(eds.). *Rethinking the Foundations of Statistics*. Cambridge University Press, 1999.

[Keynes, 1921] John Maynard Keynes. *A treatise on probability*. Macmillan, 1948.

[Kyburg & Smokler, 1964] H.E. Kyburg & H.E. Smokler(eds.). *Studies in subjective probability*. 1964. Second edition, Robert E. Krieger Publishing Company, 1980.

[Laplace, 1814] Pierre Simon - Marquis de Laplace. *A philosophical essay on probabilities*. Dover, 1951.

[Lewis, 1980] David K. Lewis. A subjectivist's guide to objective chance. In [Lewis, 1986], 83–132, 1980.

[Lewis, 1986] David K. Lewis. *Philosophical papers II*. Oxford University Press, 1986.

[McKim & Turner, 1997] Vaughn R. McKim & Stephen Turner. *Causality in crisis? Statistical methods and the search for causal knowledge in the social sciences*. University of Notre Dame Press, 1997.

[Muggleton & de Raedt, 1994] Stephen Muggleton & Luc de Raedt. Inductive logic programming: theory and methods. In *Journal of Logic Programming*, **19, 20**, 629–679, 1994.

[Paris & Vencovská, 1990] J.B. Paris & A. Vencovská. A note on the inevitability of maximum entropy. *International Journal of Approximate Reasoning*, **4**, 181–223, 1990.

[Pearl, 1988] Judea Pearl. *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Morgan Kaufmann, 1988.

[Pearl, 2000] Judea Pearl. *Causality: models, reasoning, and inference*. Cambridge University Press, 2000.

[Ramsey, 1926] Frank Plumpton Ramsey. Truth and probability. In [Kyburg & Smokler, 1964], 23–52, 1926.

[Rubin, 1987] H. Rubin. A Weak System of Axioms for "Rational" Behavior and the Nonseparability of Utility from Prior. *Statistics and Decisions*, **5**, 47–58, 1987.

[Spiegelhalter *et al.*, 1993] David J. Spiegelhalter, A. Philip Dawid, Steffen L. Lauritzen & Robert G. Cowell. Bayesian analysis in expert systems. *Statistical Science*, **8**(3), 219–283, with discussion, 1993.

[Spiegelhalter *et al.*, 2000] D.J. Spiegelhalter, J.P. Myles, D.R. Jones & K.R. Abrams. Bayesian methods in health technology assessment: a review. *Health Technology Assessment*, **4**(38), 2000.

[Spirtes *et al.*, 1993] Peter Spirtes, Clark Glymour & Richard Scheines. Causation, Prediction, and Search. *Lecture Notes in Statistics*, **81**, Springer-Verlag, 1993.

[Vapnik, 1995] Vladimir N. Vapnik. *The nature of statistical learning theory*. Springer-Verlag, 1995. Second edition 2000.

[Williamson, 1999] Jon Williamson. Countable additivity and subjective probability. *British Journal for the Philosophy of Science*, **50**(3), 401–416, 1999.

[Williamson, 2000] Jon Williamson. Probability logic. In Dov Gabbay, Ralph Johnson, Hans Juergen Ohlbach & John Woods (eds.), *Handbook of the Logic of Inference and Argument: The Turn Toward the Practical*, Elsevier, 393–419, 2000.

[Williamson, 2001] Jon Williamson. Bayesian networks for logical reasoning. *Proceedings of the 8th Workshop on Automated Reasoning*, A. Voronkov (ed.), 55–56, 2001.