

Heuristics and Normative Models of Judgment under Uncertainty

Pei Wang

*Center for Research on Concepts and Cognition
Indiana University **

ABSTRACT

Psychological evidence shows that probability theory is not a proper descriptive model of intuitive human judgment. Instead, some heuristics have been proposed as such a descriptive model. This paper argues that probability theory has limitations even as a normative model. A new normative model of judgment under uncertainty is designed under the assumption that the system's knowledge and resources are insufficient with respect to the questions that the system needs to answer. The proposed heuristics in human reasoning can also be observed in this new model, and can be justified according to the assumption.

KEYWORDS: *Subjective probability, normative and descriptive models, heuristics and bias, insufficient knowledge and resources, non-axiomatic reasoning system.*

1. Introduction

The study of human judgment under uncertainty reveals systematic discrepancy between actual human behaviors and conclusions of probability

*This work is supported by a research assistantship from the Center for Research on Concepts and Cognition, Indiana University.

*Address correspondence to 510 North Fess, Bloomington, IN 47408, USA
E-mail: pwang@cogsci.indiana.edu*

International Journal of Approximate Reasoning 1994 11:1-158

© 1994 Elsevier Science Inc.

655 Avenue of the Americas, New York, NY 10010 0888-613X/94/\$7.00

1

theory [18], that is, between what we should do (according to probability theory) and what we do (according to psychological experiments). Therefore, probability theory is not a good *descriptive* theory for human reasoning under uncertainty, though it is still referred to as a good *normative* theory.

When a normative model and a descriptive model conflict with each other, which one should be blamed? In this field, the dominant opinion is to explain the inconsistency as fallacies, errors, or illusions that happen in human thinking, for the following reasons:

1. Probability theory has a solid foundation. Its conclusions are derived deductively from a set of intuitive, or even self-evident axioms [5].
2. Most of the people who make the fallacy are disposed, after explanation, to accept that they made a mistake [19].

As a result, the research activities in this domain often consist of the following steps [9, 12]:

1. To identify the problem by carrying out psychological experiments, and compare the results with the conclusions of probability theory;
2. To explain the result by looking for the *heuristics* that are used by humans and the factors that affect their usage, and to suggest and verify methods to correct the errors.

Heuristics, as methods to assess subjective probability, “are highly economical and usually effective, but they lead to systematic and predictable errors” [18]. Compared with normative theories, such as probability theory, heuristics are not optimal, not formal, not systematic, and not always correct.

According to this opinion, the fact that probability theory cannot match actual human reasoning is not a problem of the theory. Though the discrepancy is well-known, probability theory, especially Bayesian approach, is becoming more popular as a normative model of reasoning under uncertainty.

“Bayesian approach” usually means the following in the current context:

1. Probability is a subjective measurement on uncertain belief, based on available evidence and background knowledge.
2. The beliefs of a idealized person about a domain can be represented by a (consistent) probability distribution function on a proposition space.
3. Bayes’ theorem is applied to revise one’s beliefs with new evidence.

However, besides the mainstream opinion expressed above, there are opinions to explain the discrepancy as a challenge to Bayesian approach:

- Probability theory can be interpreted differently, such as in the “frequentist” [9] or “propensity” [4] interpretation.

- There are alternative normative models that compete with Pascal probability theory, such as Baconian probability [4] and belief function [16].
- Some formal descriptive models are proposed, such as information integration theory [1].

Similarly, this paper attempts to address the following questions: Is Bayesian approach always the *correct* model to use? If not, when and why? Are there other normative models for reasoning under uncertainty? What is wrong with the heuristics?

In section 2, the assumptions and limitations of Bayesian approach are analyzed. In section 3, a new normative theory for judgment under uncertainty is briefly described. In section 4, the relationship between the heuristics and the new theory is discussed. Finally, there are conclusions arguing that the Bayesian approach is not always the appropriate normative model for the related problems.

2. Bayesian Approach as a Normative Model

Like other normative theories, Bayesian approach is based on certain assumptions, therefore it is applicable only when the assumptions are satisfied. Though such a statement sounds trivial when put in this way, the analysis about exactly when Bayesian approach can be applied is far from sufficient.

A typical opinion on this issue can be found in the following statements:

- “The subjective assessment of probability resembles the subjective assessment of physical quantities such as distance or size.” [18]
- “Although the language of probability can be used to express any form of uncertainty, the laws of probability do not apply to all variants of uncertainty with equal force.” [12]

Some authors even take such a radical position by claiming that “the world operates according to Bayes’ Theorem” [14].

According to this opinion, Bayesian approach is *the* normative model for judgment under uncertainty, and it always gives the correct or optimal answer, although sometimes it is not easy to apply the model. This opinion is also advocated by some authors in the study of uncertainty reasoning in artificial intelligence [3, 15, 17].

It is well known that the axioms of probability theory can be derived from several assumptions about the relationships between evidence and belief [5]. These assumptions, though reasonable for many situations, set up limitations for Bayesian approach at the same time.

2.1. Consistency

All applications of Bayesian approach begin with a consistent prior probability distribution on a predefined proposition (or event) space. The requirement for *consistency*, though looks reasonable, is not always satisfiable, for the following reasons:

1. When a system is open to new evidence, that is, the system works in a continuous, incremental, or adaptive manner [11], it is always possible for new knowledge to conflict with previous knowledge.
2. Under a time pressure, it is often impossible for the system to locate and consider all relevant knowledge when a judgment is made, so the judgments based on different knowledge may conflict with each other.

In these situations, a belief system cannot be abstracted as a (consistent) probability distribution on a proposition space.

2.2. Ignorance and revision

Though a probability distribution is a useful way to express one's uncertainty about some events or propositions, it does not contain the information about the amount of evidence that supports the probability distribution [13]. This type of information is referred to by various authors as "ignorance", "confidence", "reliability", and so on [16, 21].

Some people argue that this information can be derived from a probability distribution [15, 17], but this argument is invalid, because it is actually based on a confusion between the background knowledge that supports a probability assignment and the proposition that appears within a conditional probability assignment as the condition. A detailed discussion on this issue can be found in [20]. In the following, we will summarize the argument briefly.

Suppose we are talking about the uncertainty of the propositions in a space S . For this purpose, we collect some background knowledge C , and accordingly set up a prior probability distribution on S . We refer to the distribution as $P_C(x)$ ($x \in S$), and use the subscript C to indicate the fact that the probability distribution is based on the background knowledge, or context consideration, C . Now a piece of new knowledge E comes. If $E \in S$, the updated beliefs should be $P_C(x|E)$, and can be calculated according to Bayesian theorem, when $P_C(E) > 0$.

However, the above procedure cannot be applied when E is not in S or $P_C(E) = 0$. These situations happen typically when what need to be changed is the background knowledge C . Intuitively, we know that some probability distributions are established according to huge statistical data or careful theoretical analysis, but some others are based on shaky guesses. However, this difference, what we usually call the ignorance about the domain, cannot be reflected *within* the probability distribution $P_C(x)$. When

all our concern is about decision making in S , and C is remain unchanged during the process, the above difference does not matter. However, if new evidence suggests a revision of the distribution by changing C , Bayesian theorem cannot help.

Sometimes an extension of Bayesian theorem, Jeffrey's rule, can be used to modify a probability distribution. If a proposition T 's previously estimated probability is $P(T) = v$, and there is a piece of new knowledge saying that T 's probability should be v' , T 's probability is changed to v' , and for every judgment x in the space, its probability is changed from $P(x)$ to $P'(x)$, where $P'(x) = P(x|T) \times v' + P(x|\neg T) \times (1 - v')$. This is Jeffrey's rule. When $v' = 1$, we get Bayesian conditionalization. Obviously, this rule can be used to change C , however, it is an *updating rule* (for replacing old knowledge by new knowledge), rather than a *revision rule* (for combining knowledge from distinct bodies of evidence). In updating, a probability distribution is modified according to a (new) single probability assignment on a proposition, whereas the previous probability assignment on the proposition is completely ignored. As a result, updating is asymmetric, but revision (or evidence combination [16]) is symmetric [6]. Although updating is a valid operation, it cannot be used to replace revision.

2.3. Extensional interpretation

Probability theory is traditional interpreted in an extensional way [19], which means the following:

1. All sets are well-defined, that is, whether an object belongs to a set has a (maybe unknown) "Yes/No" answer.
2. The probability of " $A \subset B$ ", where both A and B are sets, is usually closely related to $|A \cap B|/|A|$. For instance, this ratio is often used as an estimation of the probability, and its limit, if known, is often taken as the probability [18].
3. The probability of " $a \in B$ ", where a is an object and B is a set, is often determined via another set R , the "reference class". When " $a \in R$ " is true and the probability of " $R \subset B$ " is known, this value can be inherited by " $a \in B$ " [24].

Though this is a very useful and reasonable way to apply probability theory to everyday life, we should keep the following points in mind.

First, this is a way to interpret probability, but not necessarily the *only* way to do it. There are several concepts that are often confused with one another: "probability" as a mathematical notion, "probability" under an extensional interpretation, and "probability" as used in ordinary language to express our (informal) degree of belief or uncertainty. These concepts are closely related, but not identical.

Second, by interpreting probability extensionally, some simplifications are introduced. We assume that the extension of all concepts are well-defined, and only extensional inclusion relations are related to probability evaluation. Again, these assumptions are reasonable for some purposes, but may be rejected for some other purposes.

There is a historical reason for why all mathematical and logical theories (including probability theory) are more closely related to the extension of concepts than to their intensions, but it does not mean that there is no way (even in the future) to formally process the intension of concept, or this should not be done.

In summary, Bayesian approach is useful, but also has limitations. In some situations, it cannot be applied or should not be applied.

3. NARS as a Normative Model

Non-Axiomatic Reasoning System (NARS) is an intelligent reasoning system which adapts to its environment under insufficient knowledge and resources. A formal and complete description about the system's logical kernel has been published, also in the *International Journal of Approximate Reasoning* [21]. It is assumed that the readers of the current paper have access to that paper, therefore in the following we only introduce the aspects of the system that are most closely related to our current issue.

3.1. Theoretical assumption

NARS is designed under the assumption that the knowledge and resources of the system are usually insufficient with respect to the questions that it needs to answer.

More concretely, the system's computing facilities (such as processor time and memory space) are usually in short supply; the questions asked by the environment have various time requirements attached; the system is always open to new knowledge (which is not necessarily consistent with the current knowledge of the system) and new questions (which may go beyond the current knowledge scope of the system).

Being adaptive, the system accommodates itself to new knowledge, makes judgments under the current knowledge–resource constrains, and adjusts its memory structure and the distribution of its resources to improve its time–space efficiency, under the assumption that future situations will be similar to past situations.

Because all the judgments are usually based on insufficient evidence, the system needs to measure how each of them is supported or refused by

available evidence. The system also needs rules to make plausible inference from given knowledge, and to revise previous beliefs in the light of new knowledge. Therefore, among other things, NARS attempts to provide a *normative* model for reasoning with uncertainty.

3.2. Uncertainty measurement

In the simplest version of NARS [21], each judgment has the form “ $S \subset P [t]$ ”, where “ S ” is the *subject term* of the judgment, “ P ” is the *predicate term*, “ \subset ” can be intuitively understood as “is a kind of” and “has the properties of” (see [21] for its formal definition), and “ t ” measures the uncertainty of the judgment.

Because all judgments in NARS are based on the system’s experience, the uncertainty of a judgment is actually represented by the *weights* of its (positive and negative) evidence. If the system knows (from its experience) a term M such that M is a kind of S and also a kind of P , or that both S and P have the property of M , then M is counted as a piece of positive evidence for $S \subset P$. If the system knows that M is a kind of S but not a kind of P , or P has the property of M but S has not, then M is counted as a piece of negative evidence for “ $S \subset P$ ”. Therefore, the uncertainty of a judgment can be represented by a pair $\langle w^+, w^- \rangle$, where w^+ is the total weight of positive evidence, and w^- is the total weight of negative evidence. w , the weight of all relevant evidence, is simply $w^+ + w^-$.

When a relative measurement is preferred, the same information can be represented by a pair of real numbers in $[0, 1]$, $\langle f, c \rangle$, where $f = w^+/w$, the *frequency* (or proportion) of positive evidence among all relevant evidence, and $c = w/(w + 1)$, a monotonically increasing function of the total weight of relevant evidence. c is referred to as the *confidence* of the judgment, because of the familiar phenomenon: the more evidence one has collected, the more confident one feels when making a judgment on the issue, though it does not follow that the judgment become “truer” or “more accurate” in an objective sense [7, 8].

For a detailed discussion of the related semantics issues, see [22].

3.3. Inference rules

In NARS there are two types of rules, one is for new judgments derivation (including deduction, induction, abduction, and so on), and the other is for conflict management. For our current purpose, we will concentrate on the latter.

By a conflict between two judgments, we mean that the two judgments are about the same “ $S \subset P$ ” relation, but they are based on different bodies of evidence, so they may attach different uncertainty to the relation.

As mentioned previously, this kind of conflict is a normal phenomenon

in NARS. With insufficient knowledge, it is always possible for new knowledge to conflict with previous knowledge. With insufficient resources, the system cannot afford the time to consider all of its knowledge to make a judgment, so a judgment is usually based on part of the system's knowledge. Therefore, even without new evidence, conflicting judgments may co-exist.

Though a normal phenomenon, the system does not let a conflicting pair of judgments stay in that way when it is found. When the inference engine is fed two judgments " $S \subset P < f_1, c_1 >$ " and " $S \subset P < f_2, c_2 >$ ", two different cases are distinguished: if the two judgments are based on *correlated* evidence, then the *updating* rule is applied, otherwise the *revision* rule is applied.

By "correlated evidence", we mean that some evidence is used to evaluate the uncertainty of both judgments (for an exact definition and how the system can recognize its happening, see [21]). The correlation may be either full (i.e., the evidence of one judgment is included by that of the other) or partial. For partially correlated evidence, an ideal solution is to merge the evidence without repeatedly counting the shared part. However, under insufficient resources, it is simply impossible to distinguish the contribution of each piece of evidence to the uncertainty of the judgment. Therefore in both situations (full and partial correlations) NARS chooses the judgment with a higher confidence (that is, based on more evidence) as the result, and ignores the other one.

When the evidence is not correlated, the revision rule is applied to get a judgment based on the *merged* evidence. From the definition of f and c (in terms of w^+ and w^-), and the convention that the weight of evidence is additive during revision, we can directly get the conclusion " $S \subset P < f, c >$ " where $f = (w_1 f_1 + w_2 f_2) / (w_1 + w_2)$, and $c = (w_1 + w_2) / (w_1 + w_2 + 1)$.

We can see from the function that after a revision the conflicting frequency evaluations are "averaged" with a (monotonically increasing) function of confidence as weight, and the confidence is increased due to the accumulation of evidence from different sources. Therefore, confidence indicates the stability of a frequency assignment in the face of confliction judgments.

For how the uncertainty measurement is used to predict future situations, see [21].

3.4. Compared with Bayesian approach

NARS and Bayesian approach are based on different assumptions. In Bayesian model, whether an event will happen, or whether a proposition is true, is uncertain, but their probability, or degree of uncertainty, is usually certain. The resources expenses of the rules (for example, Bayes' theorem)

are ignored. On the contrary, in NARS the insufficiency of knowledge and resources is consistently and completely assumed. From here, some concrete differences follow:

1. In the traditional interpretation of probability theory, only extensional evidence is considered when the probability of a statement is evaluated. In NARS, as defined above, extensional evidence (“shared instances”) and intensional evidence (“shared properties”) are equally treated when the uncertainty of a judgment is determined.
2. All the operations in Bayesian approach are *within* the same distribution function (with updating as an exception), therefore all of the probability evaluations involved are based on the same chunk of background knowledge, which can be omitted in formulae. In NARS, each judgment is evaluated individually, so it is necessary to somehow indicate the amount of its evidence. This is why a *confidence* measurement is introduced.
3. In NARS, all rules are “local”, in the sense that the uncertainty of the conclusion only depends on the premises. Therefore, the application of a rule only involves several judgments. On the contrary, Bayesian approach uses “global” rules. For example, when Bayes’ theorem (or Jeffrey’s rule) is used to update a distribution function, most probability assignments in the whole proposition space need to be re-calculated. Pearl correctly argues in [15] that local rules cause incorrect conclusions by neglect relevant information. For a system with insufficient resources, however, local rules become the only choice. The incorrect conclusions can be revised when the relevant information is located in a later time [21].
4. As a result, NARS may contain (explicitly or implicitly) conflicting judgments. To handle them, NARS has both updating rule and revision rule, whereas the latter is not available in a Bayesian model, because the information about confidence is absent there [20].

In spite of the differences, there are still many similar properties in the two models. Both of them are normative models for judgment under uncertainty, but they are based on different assumptions about the environment where the model is applied.

4. Heuristics and NARS

Though designed as a normative model, NARS shows some behaviors that are usually explained in term of “heuristics and biases”, when these phenomena happen in human judgments [18].

4.1. Availability

Availability, “the ease with which instances or occurrences can be brought to mind”, is a common heuristics in intuitive judgment of probability. It is “affected by factors other than frequency and probability”, therefore “leads to predictable biases” [18].

The same phenomenon happens in NARS. Because NARS is built under the assumption of insufficient knowledge and resources, the following properties are implied:

1. The system has to base its judgments on the *available*, though usually incomplete, knowledge. Therefore, the estimation of the frequency of an event is actually about the *experienced* frequency, rather than the *objective* frequency.
2. Judgments must be made with the *available* resources. Therefore, the system often cannot consider all of its knowledge, but only part of it.
3. Which part of the system’s knowledge is consulted is determined by several factors, such as relevance, importance, usefulness, and so on. Therefore, it is not surprising that certain events, like priming and association, influence the availability distribution [2].

Because which piece of knowledge to use at each step of reasoning is determined by the current context (by priming) and past experience (by association), it is inevitable that some knowledge, necessary for the assessment of uncertainty of a proposition, may be either unknown to the system or cannot be recalled at the time. As a result, the system will have *expectation errors* — i.e., the conflicts between the system’s expectations and the system’s future actual experience, but this type of error is not caused by mis-designing or malfunction of the system. Under the knowledge and resources constraints, the system has done its best. As long as it can revise its beliefs according to new evidence, there is no error in the system’s *operations*, though there may be errors in the *results* of these operations.

4.2. Representativeness

Representativeness, or degree of similarity, is often used as probability by human beings. “This approach to the judgment of probability leads to serious errors, because similarity, or representativeness, is not influenced by several factors that should affect judgments of probability” [18]. The basic difference between them is that “the laws of probability derive from extensional considerations” [19], but similarity judgments are based on the sharing of properties, so they are *intensional*.

As mentioned previously, here we need to distinguish three different meanings of “probability”:

1. As a pure mathematical concept, probability is neither extensional nor intensional.

2. Probability theory is usually interpreted extensionally when applied to a practical domain.
3. In everyday language and intuitive thinking, both extensional and intensional interpretations of probability happen.

Why is only the extensional interpretation referred to as “correct”? There is a historical reason: the normative theories about extension are well-developed, but the theories about intension are not. Actually there is no commonly accepted theory about how to define and process the intension of a concept. However, it does not imply that intensional factors should not be taken into consideration when we make predictions about uncertain events.

NARS is an attempt to equally treat extension and intension. When the uncertainty of a judgment is determined, both the extensional factor (shared instances) and intensional factor (shared properties) are considered [21, 22]. By doing this, it does not mean that they are not different, but that their *effects* are the same in the judgment. It is valid to build normative theories to process extension or intension separately, but it is also valid, and maybe more useful, to have theories that process both of them in a unified manner. In the latter case, it is valid to use representativeness and probability indiscriminately for certain purposes.

4.3. Adjustment and anchoring

For any system that accepts new knowledge or makes judgments by incrementally considering available knowledge, there must be a rule by which a previous probability judgment is adjusted in light of new evidence or further consideration [1].

The anchoring phenomenon, or insufficient adjustment from the initial point, is observed in human thinking [18]. By calling the observed adjustments “insufficient”, it is assumed that the correct adjustment rule is Bayes’ theorem, or its extension, Jeffrey’s rule.

As discussed previously, in NARS, two different cases are distinguished when judgments conflict with each other. If the evidence supporting the two judgments are correlated, the updating rule is applied, otherwise the revision rule is applied.

In updating, there are also two possibilities: if the confidence of the previous estimation is no lower than the confidence of the new estimation, then nothing is changed, otherwise the former is replaced by the latter. Though the second possibility is the same with Jeffrey’s rule, what follows is different: NARS usually cannot afford the resources to update all related judgments, therefore only some of them are updated accordingly, by applying the inference rules and the updating rule of NARS.

In revision, the new frequency is a weighted sum of those of the premises, as discussed in the previously.

Therefore, in all situations, the adjustment of frequency in NARS is no more than what required by probability theory. If conditionalization (Bayes' theorem and Jeffrey's rule) is *the* correct way of adjustment, NARS shows the anchoring bias, too. However, as argued above and in [20], it is not always valid to use updating as revision, or to assume sufficient resources for global updating. Again, there is nothing wrong in NARS.

5. Conclusions

This paper is a follow-up of [21], and its purpose is to show some implications of the formal model defined in the previous paper. For a more recent and complete description of the NARS project, see [23].

Though the above discussions only address some, but not all, aspects of the system, we can still get some conclusions about the models of judgment under uncertainty.

Despite the fact that NARS is designed as a normative model, the system shows some behaviors similar to those happens in human thinking, which are usually explained in terms of heuristics.

NARS is no less normative than probability theory in the sense that it is developed from some basic principles and assumptions about what *should* a system (human or computer) do with incomplete and inaccurate knowledge [21]. It is true that when applied into a practical domain, NARS may produce wrong expectations, but so does probability theory.

NARS is not proposed to replace Bayesian models. In Good's terms [10], Bayesian approach is toward a "Type I" rationality by maximizing the expected utility, while NARS is toward a "Type II" rationality where the cost of computing must be taken into account. If Bayesian approach can be applied in a situation (i.e., the computational cost and the revision of background knowledge can be ignored there), it is still better than NARS. It is in situations where Bayesian approach cannot or should not be applied that approaches like NARS will take over.

NARS is not proposed as a descriptive model for actual human thinking, such as Anderson's model [1]. Its behavior is still different from that of a human being. The approach is not justified by psychological data, but by logical analysis. Therefore there is no psychological experiment conducted to verify the theory.

However, psychological observations, as those reported in [18], do have a strong relation to the study of normative models. From the above discussion we conclude that there is no unique normative model for judgment under uncertainty — different models can be established according to different theoretical assumptions. NARS is "less idealized" than Bayesian approach, because it assumes stronger knowledge–resource constraints. The

behavior of NARS is more similar to those of people, therefore we have reason to believe that its assumptions are more “realistic” — that is, more similar to the human cognitive mechanism. This result can be explained by the observation that human mind was evolved, and still works, in an environment where the knowledge and resources are usually insufficient to solve its problems.

On the other hand, we see that it is possible to find a normative interpretation for the “heuristics”. They are not necessarily “efficient but biased”. Sometimes they indicate the right thing to do, though they do not always succeed.

As for the “biases” and “fallacies” discussed in the psychological literature, the situation is complex. NARS cannot explain all of them, but it does suggest a distinction: some violations of probability theory happen in the situations where probability theory cannot or should not be applied, and they may be explained by other normative theories, therefore they are not necessarily errors. The real errors happen when probability theory should be applied, but the person fails to do so.

Even for the latter case, an explanation is suggested from the study of NARS: because the human mind usually works under some assumptions about knowledge and resources that is quite different from what probability theory assumes, it needs some special effort (which does not always succeed) to suppress the “natural law of thinking”, and to learn, to remember, and to follow probability theory.

Now we can say that by analyzing the so called “heuristics and biases”, we not only find limitations in human reasoning, but also find limitations in probability theory, especially in Bayesian approach. Just like nobody is born with a digital calculator embedded in brain, a brain does not include a Bayesian network, and even for a good reason — in the environment for a human to survive, the assumptions made by Bayesian approach are not always correct, or, usually incorrect.

References

1. N. Anderson. A cognitive theory of judgment and decision. In B. Brehmer, H. Jungermann, P. Lourens, and G. Sevón, editors, *New Directions in Research on Decision Making*, pages 63–108. Elsevier Science Publishers, Amsterdam, 1986.
2. H. Arkes. Costs and benefits of judgment errors: implications for debiasing. *Psychological Bulletin*, 110:486–498, 1991.
3. P. Cheeseman. In defense of probability. In *Proceedings of the Eighth International Joint Conference on Artificial Intelligence*, pages 1002–1009, 1985.

4. L. Cohen. Can human irrationality be experimentally demonstrated? *The Behavioral and Brain Sciences*, 4:317–331, 1981.
5. R. Cox. Probability, frequency and reasonable expectation. *American Journal of Physics*, 14(1), 1946.
6. D. Dubois and H. Prade. Updating with belief functions, ordinal conditional functions and possibility measures. In P. Bonissone, M. Henrion, L. Kanal, and J. Lemmer, editors, *Uncertainty in Artificial Intelligence 6*, pages 311–329. North-Holland, Amsterdam, 1991.
7. H. Einhorn and R. Hogarth. Confidence in judgment: persistence of illusion of validity. *Psychological Review*, 35:395–416, 1978.
8. B. Fischhoff, P. Solvic, and S. Lichtenstein. Knowing with certainty: the appropriateness of extreme confidence. *Journal of Experimental Psychology: Human Perception and Performance*, 3:552–564, 1977.
9. G. Gigerenzer. How to make cognitive illusions disappear: beyond “heuristics and biases”. In W. Stroebe and M. Hewstone, editors, *European Review of Social Psychology, Volume 2*, chapter 4, pages 83–115. John Wiley & Sons Ltd, 1991.
10. I. Good. *Good Thinking: The Foundations of Probability and Its Applications*. University of Minnesota Press, Minneapolis, 1983.
11. R. Hogarth. Beyond discrete biases: Functional and dysfunctional aspects of judgmental heuristics. *Psychological Bulletin*, 90:197–217, 1981.
12. D. Kahneman and A. Tversky. On the study of statistical intuitions. In D. Kahneman, P. Slovic, and A. Tversky, editors, *Judgment under Uncertainty: Heuristics and Biases*, chapter 34, pages 493–508. Cambridge University Press, Cambridge, England, 1982.
13. J. Keynes. *A Treatise on Probability*. Macmillan, London, 1921.
14. D. Lyon and P. Slovic. Dominance of accuracy information and neglect of base rates in probability estimation. *Acta Psychologica*, 40:287–298, 1976.
15. J. Pearl. *Probabilistic Reasoning in Intelligent Systems*. Morgan Kaufmann Publishers, San Mateo, California, 1988.
16. G. Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, New Jersey, 1976.
17. D. Spiegelhalter. A statistical view of uncertainty in expert systems. In W. Gale, editor, *Artificial Intelligence and Statistics*, pages 17–56. Addison Wesley, Reading, 1986.
18. A. Tversky and D. Kahneman. Judgment under uncertainty: heuristics and biases. *Science*, 185:1124–1131, 1974.
19. A. Tversky and D. Kahneman. Extensional versus intuitive reasoning: the conjunction fallacy in probability judgment. *Psychological Review*, 90:293–315, 1983.
20. P. Wang. Belief revision in probability theory. In *Proceedings of the Ninth Conference on Uncertainty in Artificial Intelligence*, pages 519–526. Morgan Kaufmann Publishers, San Mateo, California, 1993.
21. P. Wang. From inheritance relation to non-axiomatic logic. *International Journal of Approximate Reasoning*, 11(4):281–319, November 1994.
22. P. Wang. Grounded on experience: Semantics for intelligence. Technical Report 96, Center for Research on Concepts and Cognition, In-

diana University, Bloomington, Indiana, 1995. Available via WWW at <http://www.cogsci.indiana.edu/farg/peiwang/papers.html>.

23. P. Wang. *Non-Axiomatic Reasoning System: Exploring the Essence of Intelligence*. PhD thesis, Indiana University, 1995.
24. P. Wang. Reference classes and multiple inheritances. *International Journal of Uncertainty, Fuzziness and and Knowledge-based Systems*, 3(1):79–91, 1995.