# Reference Classes and Multiple Inheritances

**Pei Wang**
Center for Research on Concepts and Cognition
Indiana University
*pwang@cogsci.indiana.edu*

July 25, 1994

**Abstract**

The reference class problem in probability theory and the multiple inheritances (extensions) problem in non-monotonic logics can be referred to as special cases of conflicting beliefs. The current solution accepted in the two domains is the specificity priority principle. By analyzing an example, several factors (ignored by the principle) are found to be relevant to the priority of a reference class. A new approach, Non-Axiomatic Reasoning System (NARS), is discussed, where these factors are all taken into account. It is argued that the solution provided by NARS is better than the solutions provided by probability theory and non-monotonic logics.

## 1 Introduction

How do we predict whether an individual has a certain property, if direct observation is impossible? A useful method is to look for a "reference class". The class should include the individual as an instance, and we should know something about how often the instances of the class have the desired property, or whether its typical instances have it. Then, the prediction can be done by letting the instance "inherit" the information from the class.

In the field of reasoning under uncertainty, there are (at least) two paradigms that use this type of inference: non-monotonic logics (for example, Touretzky's inheritance network in [15]), and probabilistic reasoning systems (for example, Pearl's Bayesian network in [9]).

In non-monotonic logics, if the only relevant knowledge is "$A$ is an instance of $R$" and "Normally, $R$'s instances have the property $Q$", a defeasible conclusion is "$A$ has the property $Q$".

In probabilistic reasoning systems, under the subjective interpretation of probability, if the only relevant knowledge is "$A$ is an instance of $R$" and "The probability for $R$'s instances to have the property $Q$ is $p$", a plausible conclusion would be "The probability for $A$ to have the property $Q$ is $p$".

Now a problem appears: if $A$ belongs to two classes $R_1$ and $R_2$ at the same time, and the two classes lead to different predictions about whether (or how probable) $A$ has the property $Q$, what conclusion can we reach? In different context, the problem is referred to as "multiple inheritance problem", "multiple extension problem", or "reference class problem" [3, 6, 7, 9, 10, 11, 15].

Though the above theories treat the problem differently, they have something in common: None of them suggest a general solution to the problem, though they agree on a special case: if $R_2$ is a (proper) subset of $R_1$, $R_2$ is the correct reference class to be used.

Let us see two examples.

1. In [15], Touretzky said: "Since Clyde is a royal elephant, and royal elephants are not gray, Clyde is not gray. On the other hand, we could argue that Clyde is a royal elephant, royal elephants are elephants, and elephants are gray, so Clyde is gray. Apparently there is a contradiction here. But intuitively we feel that Clyde is not gray, even though he is an elephant, because he is a special type of elephant: a royal elephant."

2. In [6], Kyburg said: "If you know the survival rate for 40-year old American male to be 0.990, and also that the survival rate for 40-year old American male white-collar workers to be 0.995, then, other things being equal, it is the latter that should constrain your beliefs and enter your utility calculations concerning the particular 40 year old male white-collar worker John Smith."

Let us call this principle "specificity priority principle". It looks quite reasonable, and it is not hard to find many examples to show that we do apply such a principle in common sense reasoning. However, the following questions are still open:

1. Why is the principle correct? Can it be justified by more basic axioms or assumptions?

2. Beside specificity, what are the "other things" that influence the priority of a reference class?

3. When neither reference class is more specific than the other, what should be done?

For the first question, Reichenbach made it a matter of definition by "regarding the individual case as the limit of classes becoming gradually narrower and narrower" [11]; Pearl said it is because "the influence of the remote ancestors is summarized by the direct parents" [9].

For the second question, Reichenbach said we need to have complete statistical knowledge on the reference class, that is, the probability for $R$ to be $Q$ should be supported by good statistical data [11]. In non-monotonic logics, this corresponds to sufficient evidence which can determine what properties a *normal* instance of the class has.

For the third question, few word is said, except Reichenbach's suggestion to "look for a larger number of cases in the narrowest common class at your disposal" [11].

Dissatisfied by the above answers, this paper is an attempt to discuss the issue of reference class in more detail. An example, with its variations, will be discussed as a starting point, then, after analyzing the factors that influence the result, the solution provided by Non-Axiomatic Reasoning System (NARS) [16, 17, 18] is discussed and compared with the specificity priority principle.

## 2 A thought experiment

Let us reconstruct Kyburg's example in the following way: Imaging that you are working for a life insurance company, and you need to predict whether John Smith can live to 40. You have John's personal information, and for some special reasons (such as you just woke up from a 200-year-long sleep or you are actually an extraterrestrial spy), you have no background knowledge about the survival rates at 40 for various groups of people. Fortunately, you have access to personal files of some Americans, who are alive or died in recent years, and you decide to make the prediction by the "reference class method" defined above.

At first, knowing that John is a male, you begin to build the first reference class $R_1$ by picking up some files randomly. $R_1$ consists of two subsets: $P_1$ includes the positive evidence for John's survival, that is, American males who are more than 40 year old (including those who are already deceased), and $N_1$ includes the negative evidence, that is, those who died before 40. You should keep in mind that American males who are alive and younger than 40 (including John himself) are neither positive evidence nor negative evidence for the prediction, so they do not belong to $R_1$.

If you weight everyone as equal (why do not you?), your prediction should be determined by the relative size of $P_1$ and $N_1$. Let us say $|P_1| > |N_1|$. Therefore you predict that John Smith can live to 40.

After returning the files, you have a new idea: why not consider the fact that John is, among other things, a white-collar worker? So you build another reference class $R_2$ similarly. Let us assume, unfortunately, this time you find that $|P_2| < |N_2|$. Here you meet the reference class problem: to see John as a "male" and a "male white-collar worker" will lead to different predictions.

If we apply the *specificity priority principle* here, the result should be dominated by $R_2$, since "male white-collar worker" is a proper subset of "male". However, it is easy to find a situation to show that sometimes the result is counter-intuitive. If you have looked through 1000 files, and all of them are males and live to 40, and after that you find 1 male white-collar worker who died at 35, will you predict that John will die before 40? It seems very unlikely.

Does this mean that the specificity priority principle is wrong? Of course not. Sample size is obviously one of the "other things" that influence the priority of a reference class. One sample is far from enough to tell us about how a "typical" or "normal" instance looks like, or to support a statistical assertion on the instances. In such a case, the principle is inapplicable, since there is another relevant difference between the two reference classes, beside their specificities.

If you have to make predictions in such an environment, what will you do? Let us consider a simple psychological experiment. Assuming $R_1$ includes positive evidence only (that is, $R_1 = P_1$, no male is found to be died before 40), but $R_2$ includes negative evidence only (that is, $R_2 = N_2$, no male white-collar worker is found to be alive at 40). Even before really carrying out such an experiment on human subjects, We are confident to make the following prediction: If $|P_1|$ is fixed at a big number (say 1000), and $|N_2|$ is increased one by one, starting from 1, the predictions made by subjects will be positive before $|N_2|$ reaching a certain point, and negative after reaching that point. That critical point may vary from person to person, but is always smaller than $|P_1|$.

The "sample size effect" can also be used to answer the following question: If a more specific reference class is always better, why do not we simply use the *most specific reference class*, defined by all available properties of John Smith? The reason is simple: in most situations such a class is *empty* — nobody is similar to John to such an extent. With more and more properties used to define a reference class, the extension of the class becomes narrower and narrower. As a result, fewer and fewer samples can be found to support the prediction or to against it. From this point of view, specificity is not preferred.

Previously, we talked about the reference classes $R_1$ and $R_2$, as if they are accurately defined. Obviously it is a simplification. Though we can ignore the boundary cases for "male", the fuzziness in "white-collar worker" cannot be neglected so easily. As argued by fuzzy set theory [19] and prototype theory [13], whether an instance belongs to a concept is usually a matter of degree. This membership function is also related to the current issue: if John can be referred to as a "white-collar worker", but not a typical one, the influence of $R_2$ will be reduced.

How should we empirically determine the membership function for a concept like "white-collar worker"? Psychologists suggest that it can be determined by the degree of *similarity* of the instance to a *prototype* [13] or an *exemplifier* [8] of the concept. A common way to determine the similarity between two concepts is to compare their properties. Like the situation of probability prediction, similarity evaluation is also influenced by two factors: the proportion of shared properties and the amount of properties that has been checked during the evaluation. As a result, we can predict that the more properties an instance and a reference class shares, the higher the priority of the class is.

## 3 Analysis of the factors

In the previous section, some factors are listed which influence the priority of a reference class. Let us summarize them formally as the following: To predict whether $A$ has property $Q$, two reference classes $R_1$ and $R_2$ are taken into consideration. If we write "$A$ has property $Q$" as $A \in Q$, "$A$ is a member of $R$" as $A \in R$, and "$R$'s members have property $Q$" as $R \subset Q$, the problem becomes: to evaluate (to predict its truth value or probability)

$$T_0 : \; A \in Q$$

from the given evaluations of

$$T_1 : A \in R_1, \quad T_2 : A \in R_2, \quad T_3 : R_1 \subset Q, \quad T_4 : R_2 \subset Q.$$

How should we evaluate each of the above statements? As discussed in the previous section, we can see two major factors that need to be measured: one is the frequency or proportion of the positive evidence among all relevant evidence, and the other is the total amount of evidence that have been considered. In the following discussion, we will refer to the weight of positive, negative, and total evidence for $T_i$ as $w_i^+$, $w_i^-$, and $w_i$, respectively, where $w_i^+ + w_i^- = w_i$, and $i = 0, 1, 2, 3, 4$.

In traditional binary logic, as used in set theory, the result is straightforward. Under the assumption of complete knowledge, there are only two types of statements: affirmative and negative, where either $w_i^+ = w_i$ or $w_i^- = w_i$. Such a system is monotonic in the sense that the evaluation of statements are insensitive to new knowledge, that is, $w_i$ is constant for all statements. In this situation, the reference class problem cannot appear: given the consistency of the premises, the conclusions derived along different paths cannot conflict with each other.

However, problems emerge as soon as we allow general rules (as $T_3$ and $T_4$) to have *exceptions*, and to allow the accommodation of new evidence. If $T_3$ is such a rule, from the fact that $T_1$ is true, we can no longer guarantee the truthfulness of $T_0$, since $A$ may turn out to be an exception. On the other hand, since $A$ is not necessarily included in $w_3$, $T_0$ is not necessarily true even if $T_3$ has no available negative evidence. Returning to the previous example: if the current survival rate is 1 for males at the age of 40, it does not follow that John Smith (currently 20 years old) cannot die before 40, because the "rule" may have *potential* negative evidence which is not available at present.

In such a situation, how can the "reference class method" be justified? It can be derived from a more fundamental principle: the *extension* (instances) and *intension* (properties) of a concept is *co-ordinated* during the development of human classification behavior [4]. As a result, we can predict extensional relations from intensional relations (e.g., to determine whether a concept includes an instance by checking the properties of the concept), and to predict intensional relations from extensional relations (e.g., to determine whether a concept has a property by checking the instances of the concept).

However, with insufficient knowledge and resources, the co-ordination between extension and intension is imperfect, that means (1) the statements usually have both positive and negative evidence, and (2) the current evaluations should be revised according to future evidence.

When two (or more) competing evaluations (about the same statement, but based on different sources of evidence) are taken into consideration at the same time, there are two possibilities:

1. The evidence that support them is *uncorrelated*, that is, they come from independent sources. Therefore, the two pieces of evidence should be *combined*, since the weight of evidence is an additive measurement. As a result, the frequency of the result is a weighted sum of the

frequencies of the competing evaluations, and the weight of the result is the sum of the weights of the competing evaluations (see the following section). Obviously, between the two competing evaluations, the one with a larger *weight* should have a stronger influence on the result.

2. The two competing evaluations are based on *correlated* evidence, when some evidence has been used to supporting both evaluations. In such a case, the two pieces of evidence should not be combined by addition. If the system has to make a decision without the help of other information about *how* the evidence is correlated, the evaluation with a larger *weight* should be chosen, since it is supported by more evidence, and the other is ignored.

Since in both cases the evaluation with a larger weight has a priority over the other one, let us see how the weight is determined. Intuitively, to get a large weight for $T_0$ ("$A$ has property $Q$"), using $R_1$ as the reference class, both $T_1$ and $T_3$ must have large weights, that means, there should be enough evidence to support the evaluation for "$A$ is a member of $R$" and "$R$'s members have property $Q$". As discussed in the previous section, for $T_1$ it means that $R_1$ is intensionally described by a large set of properties (shared by $A$), and for $T_3$ it means that $R_1$ is extensionally exemplified by a large set of instances (shared by $Q$).

Since a more specific concept is always described by more properties than a more general concept, the former has a priority, other things being equal. However, when there is a sample size difference or typicality difference which favors the more general concept, the specificity priority can be more or less cancelled out. This conclusion from above analysis is consistent with the result of the thought experiment in the previous section.

## 4 Another approach

Non-Axiomatic Reasoning System (NARS) [16, 17, 18] is an intelligent system which works and adapts to its environment under the assumption of insufficient knowledge and resources. As a result of the assumption, it also meets the reference class problem.

For our current purpose, let's say that each statement in NARS has the form of $S \subset P$ or $S \in P$, where $S$ is the *subject term*, and $P$ the *predicate term*. Since $S \in P$ can be identically rewritten as $\{S\} \subset P$, we will talk only about the first form in the following.

Intuitively speaking, either a common *instance* or a common *property* of $S$ and $P$ is counted as a piece of positive evidence for $S \subset P$, with a *unit weight*, that is, $w = w^+ = 1$. On the contrary, either an instance of $S$ which is not shared by $P$ or a property of $P$ which is not shared by $S$ is counted as a piece of negative evidence for $S \subset P$, with a *unit weight*, that is, $w = w^- = 1$. For a formal definition of evidence and its weight, see [17].

In NARS, the *truth value* of a statement, indicating how the statement is supported by available evidence, is represented by a pair of real numbers in [0, 1], $< f, c >$, where $f$ is $\frac{w^+}{w}$, the relative *frequency* of positive evidence, and $c$ is the *confidence*. Confidence is defined in NARS as $\frac{w}{w+k}$, a monotonic increasing function of total weight, indicating the stability of the current frequency evaluation. $k$, a positive constant, is assumed to be 2 in this paper. It is easy to see that given the definition of truth value in terms of weight of evidence, we can also calculate the latter from the former.

NARS can carry out several types of inference on *judgments* (i.e., statements with truth value). In this paper, only three of them are mentioned: *updating, revision* and *deduction*.

The updating rule and the revision rule are used to deal with conflicting evidence. We assume that the system has a pair of judgments which are about the same statement:

$$J_1 : \ S \subset P \ \ <f_1, c_1> \quad \text{and} \quad J_2 : \ S \subset P \ \ <f_2, c_2>$$

At first, the system needs to check whether the judgments are based on correlated evidence. This can be done somehow in NARS (see [17]), and a "Yes/No" result is reported. As discussed before, if the evidence of the two judgments is not correlated, an evidence combination, or *revision*, should be done. For the revised judgment, its weight of total, positive, and negative evidence should be the sum of the corresponding weights of $J_1$ and $J_2$. Given the relations between weight and truth value, it is easy to get the truth value of the revised judgment $S \subset P \ <f_0, c_0>$:

$$f_0 = \frac{w_1 f_1 + w_2 f_2}{w_1 + w_2} \ , \quad c_0 = \frac{w_1 + w_2}{w_1 + w_2 + k}$$

We can see from the function that the revised frequency is a weighted sum of the frequencies of the competing judgments, and the confidence of the revised judgment is higher than those of $J_1$ and $J_2$, since evidence is accumulated through revision.

On the other hand, if the two pieces of evidence are correlated, they cannot be accumulated. Instead of revision, the *updating* rule is applied, which picks up the judgment with a higher confidence, and ignores the other.

The *deduction* rule can infer $S \subset P \ <f_0, c_0>$ from $M \subset P \ <f_1, c_1>$ and $S \subset M \ <f_2, c_2>$. The truth value function is derived from *T-norm* and *T-conorm* [1, 2], and the derivation can be found in [17]. The resulting function is:

$$f_0 = \frac{f_1 f_2}{f_1 + f_2 - f_1 f_2} \ , \quad c_0 = (f_1 + f_2 - f_1 f_2)\, c_1 c_2$$

Now we are ready to see how NARS treats the reference class problem. Putting into the format of NARS, the premises are:

$$J_1 : A \in R_1 \ <f_1, c_1> \ , \quad J_2 : A \in R_2 \ <f_2, c_2> \ ,$$

$$J_3 : R_1 \subset Q \ <f_3, c_3> \ , \quad J_4 : R_2 \subset Q \ <f_4, c_4> \ .$$

Since John shares one property with $R_1$ ("male") and two properties with $R_2$ ("male" and "white-collar worker"), we have $w_1 = w_1^+ = 1$ and $w_2 = w_2^+ = 2$. It follows that (assuming $k = 2$) $f_1 = f_2 = 1$, $c_1 = \frac{1}{3}$, and $c_2 = \frac{1}{2}$ (the fuzziness in the concepts is temporally ignored). Under the assumption that $R_1$ consists of 1000 positive samples, we have $f_3 = 1$ and $c_3 = \frac{1000}{1002}$. Let us say that $R_2$ includes negative samples only, but leaves the number of samples, $n$, as a variable, to see how it influences the final evaluation of $A \in Q$, that is, "John Smith will live to 40". Therefore, we have $f_4 = 0$ and $c_4 = \frac{n}{n+2}$.

Applying the deduction rule, from $J_1, J_3$ and $J_2, J_4$, respectively, we get

$$J_5 : A \in Q \ \ <1, \ c_1 c_3> \ , \quad \text{and} \quad J_6 : A \in Q \ \ <0, \ c_2 c_4>$$

Since the knowledge that "John is male" is used to evaluate both $J_1$ and $J_2$, and they are used in the derivation of $J_5$ and $J_6$, respectively, the evidence for $J_5$ and $J_6$ is correlated. As a result, the updating rule is applied to pick up the judgment that has a higher confidence as the conclusion. Which reference class will win the competition?

By solving the inequality $c_1 c_3 > c_2 c_4$, we can see that

1. When $0 < n < 4$, $R_1$ is selected. The specificity priority of $R_2$ is undermined by the fact that the sample size of $R_2$ is too small.

2. When $n \geq 4$, $R_2$ is selected. The specificity priority can be established even by a pretty small sample size: with $|R_1| = 1000$ and $|R_2| = 4$, the prediction is still determined by $R_2$ due to its specificity.

If John is not a typical white-collar worker (i.e., $f_2 < 1$), $R_2$'s confidence is smaller than $c_2 c_4$, so it may need a bigger $n$ for $R_2$ to be dominant. Therefore, when NARS is selecting a reference class, several factors are balanced against one another, including specificity, typicality, sample size, and so on. It provides a generalization of the specificity priority principle, by taking more relevant factors into consideration.

NARS' approach is more general than the specificity priority principle in another way. The including of reference classes is only a special case for two judgments to be based on correlated evidence. It follows that the specific priority principle is a special case of NARS' updating rule.

How about competing reference classes that do not involve correlated evidence? Let us say in the previous examples, $R_1$ is still for "male", but $R_2$ is changed for "smoker and white-collar worker". If the deduced judgments $J_5$ and $J_6$ are not based on correlated evidence in some other ways, the two judgments will be combined by the revision rule of NARS. Other things being equal, $R_2$ has a higher priority, since it matches better with John's properties. However, in this case a higher priority only means a higher *weight* in determining the frequency of the conclusion. The judgment from the other reference class is not ignored. In this situation, the reference class competing is solved not by *choosing one of them*, but by *combining the two.*

Let us see how NARS treats the famous "Nixon Diamond" [14]. Putting into the previous framework, in this problem we have "Nixon" as $A$, "Quaker" as $R_1$, "Republican" as $R_2$, and "Pacifist" as $Q$. It is also given that $J_1$, $J_2$, and $J_3$ are positive, but $J_4$ is negative. By deduction, two conflicting judgments $J_5$ ("Nixon is a pacifist") and $J_6$ ("Nixon is not a pacifist") can be derived as in the previous example.

Since we can assume the un-correlation of evidence of the judgments ($R_1$ and $R_2$ have no known relation), $J_5$ and $J_6$ will be combined by the revision rule, and the result depends on the truth value of the premises.

1. If $f_1 = f_2$, $c_1 = c_2$, $f_3 = 1 - f_4$, and $c_3 = c_4$, we will get $f_0 = 0.5$. That is, when the positive evidence and the negative evidence exactly balance with each other, the system is indifferent between a positive prediction and a negative prediction.

2. If $c_1 > c_2$, and the other conditions as in (1), we will get $f_0 > 0.5$. That is, when Nixon shares more property with Quaker, the system will put more weight on the conclusions suggested by the evidence about Quaker.

3. If $f_3 > 1 - f_4$ or $c_3 > c_4$, and the other conditions as in (1), we will get $f_0 > 0.5$. That is, when we have stronger statistical data about Quaker, the system will put more weight on the conclusions suggested by the evidence about Quaker, too.

In any situation, what NARS does is to combine the evidence from both sources. Even if "Quaker" is given a higher priority, the evidence provided by "Republican" still has its effect on the result. On the other hand, this kind of conflict does not always (though sometimes it does) cause complete indifference or ambiguity, as it does in non-monotonic logics [15].

## 5 Comparisons

Compared with non-monotonic logics and probability theory, the processing of the reference class problem in NARS has the following characteristics:

1. While still following the specificity priority principle, several factors, such as sample size and degree of membership, are taken into account to quantitatively determine the priority of a reference class, and all the factors are projected into a unique dimension, the weight of evidence.

2. The specificity priority principle has been generalized into a "confidence priority principle" which will pick up a judgment with the highest confidence among the competing ones. The principle is applied when the competing judgments are supported by correlated evidence. As discussed above, specificity is one way to get a high confidence, and inclusion relation between reference classes is one reason that causes evidence correlation.

3. When conflicting judgments come from different sources, the revision rule is applied to combine them by summarizing the evidence. This operation in unavailable in non-monotonic logics and probability theory.

Why cannot we do similar things in non-monotonic logics and probability theory? One of the major reasons is that the *confidence* (or identically, *weight of evidence*) measurement cannot be easily introduced there. From the view point of NARS, the confidence of all the default rules (in non-monotonic logics) and probability assignments (in probability theory) is 1, that is, they cannot be revised by accommodating its current evaluation to new evidence.

Because in [16] we have already argued that revision cannot be done in a first-order probability distribution, let us concentrate on non-monotonic logics here.

Non-monotonic logics are often referred as "defeasible logic", but actually what is defeasible are only the *conclusions* derived from the default rules, rather than the rules themselves. The rules are treated as *conventions* [12], which are immune from empirical revision. As long as "Birds fly" is a default rule, it remains to be valid no matter how many birds found later cannot fly.

To treat default rules as convention is possible and even desired in many situations. In communication between systems (human or computer), these conventions are often intentionally followed, and if we already have lots of evidence, or if we only study the judgment making of the system in a short period, the influence of new evidence can be ignored. In these situations, a binary logic is preferred for its simplicity and clarity. However, such assumptions about environment are not always valid. Another thing we need to keep in mind is: when treated as *conventions*, these rules become *a priori* to the system, and what the researchers concern about them is different from when they are treated as *generalized experience*. As conventions, their generation, acceptance, comparison, modification, and rejection are no longer determined, or even influenced, by the experience of the system. Though it is correct to say that "normality" or "typicality" should not be interpreted in a pure frequentist way as "in most cases", we still have reason to argue that for many purposes, it is better to see them as closely related to empirical evidences, and have different degrees [5, 13]. Therefore, it makes sense, and often necessary, to measure the relations between the default rules and available evidence, which cannot be done in the framework of binary logic.

In summary, though non-monotonic logics and probability theory can be successfully used in many domains, their solution to the reference class (or multiple inheritance) problem is quite limited. Many related factors are ignored, and the problem is unsolvable when the involved reference classes do not include one another. For the more general problem, that is, how to revise beliefs, no solution is given. The problem is not one that can be solved by working harder in the two paradigms: as shown above, the solution involves factors that are ruled out by the fundamental definitions of the two paradigms. NARS is not always better than non-monotonic logics and probability theory in all situations, but it is better when the available knowledge and resources are insufficient. Though still a simplification, it does consider more factors than the other two.

## Acknowledgment

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

Thanks to Helga Keller for correcting my English.

## References

[1] P. Bonissone. Summarizing and propagating uncertain information with triangular norms. *International Journal of Approximate Reasoning*, 1:71–101, 1987.

[2] D. Dubois and H. Prade. A class of fuzzy measures based on triangular norms. *International Journal of General Systems*, 8:43–61, 1982.

[3] B. Grosof. Defeasible reasoning and uncertainty: comments. In M. Henrion, R. Shachter, L. Kanal, and J. Lemmer, editors, *Uncertainty in Artificial Intelligence 5*, pages 61–66. North-Holland, Amsterdam, 1990.

[4] B. Inhelder and J. Piaget. *The Early Growth of Logic in the Child.* W. W. Norton & Company, Inc., New York, 1969. Translated by E. Lunzer and D. Papert.

[5] D. Kahneman and D. Miller. Norm theory: comparing reality to its alternatives. *Psychological Review*, 93:136–153, 1986.

[6] H. Kyburg. The reference class. *Philosophy of Science*, 50:374–397, 1983.

[7] E. Neufeld. Defaults and probabilities: extensions and coherence. In *Proceedings of the First International Conference on Principles of Knowledge Representation and Reasoning*, 1989.

[8] R. Nosofsky. Typicality in logically defined categories: exemplar-similarity versus rule instantiation. *Memory and Cognition*, 17:444–458, 1991.

[9] J. Pearl. *Probabilistic Reasoning in Intelligent Systems.* Morgan Kaufmann Publishers, San Mateo, California, 1988.

[10] D. Poole. On the comparison of theories: preferring the most specific explanation. In *Proceedings of the Ninth International Joint Conference on Artificial Intelligence*, pages 144–147, 1985.

[11] H. Reichenbach. *The Theory of Probability.* University of California Press, Berkeley, California, 1949. Translated by E. Hutten and M. Reichenbach.

[12] R. Reiter. Nonmonotonic reasoning. *Annual Review of Computer Science*, 2:147–186, 1987.

[13] E. Rosch. On the internal structure of perceptual and semantic categories. In T. Moore, editor, *Cognitive Development and the Acquisition of Language*, pages 111–144. Academic Press, New York, 1973.

[14] D. Touretzky. Implicit ordering of defaults in inheritance systems. In *Proceedings of the National Conference of Artificial Intelligence*, pages 322–325, 1984.

[15] D. Touretzky. *The Mathematics of Inheritance Systems.* Pitman Publishing, London, 1986.

[16] 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.

[17] P. Wang. From inheritance relation to non-axiomatic logic. Technical Report 84, Center for Research on Concepts and Cognition, Indiana University, Bloomington, Indiana, 1993.

[18] P. Wang. Non-axiomatic reasoning system (version 2.2). Technical Report 75, Center for Research on Concepts and Cognition, Indiana University, Bloomington, Indiana, 1993.

[19] L. Zadeh. Fuzzy sets. *Information and Control*, 8:338–353, 1965.