Knowledge Representation Issues in Default Reasoning

J. Terry Nutter

TR 89-21

KNOWLEDGE REPRESENTATION ISSUES IN DEFAULT REASONING

J. TERRY NUTTER
Department of Computer Science
Virginia Tech
Blacksburg, Virginia 24061

ABSTRACT

Most existing approaches to reasoning in uncertainty and with incomplete information appeal to formal theories, with relatively little attention to the phenomena they are intended to capture. This has had two major consequences. First, it has led to spurious disputes, in which participants criticize alternative approaches in the belief that they are competing, when in fact they are investigating different aspects of related phenomena, and should ultimately be viewed as cooperative efforts. Second, it has led to wasted effort of models which fail to reflect important aspects of kinds of reasoning which they are trying to capture, because the representational requirements have not been adequately spelled out. This paper delineates several different kinds of reasoning in uncertainty, establishes some distinctions within the field, and attempts to begin setting some ground rules for representational adequacy.

Section 1: INTRODUCTION

Most proposed systems for default reasoning rest their appeal on formal developments in logic, statistics, decision theory, or some other technical field. These discussions usually place little emphasis on analyzing the phenomena to motivate their approaches, but instead plunge directly into technical developments, and argue for one view over another on the basis of formal results. When we compare these discussions to those in other fields of artificial intelligence, the lack of attention to the phenomena which the model is trying to capture springs sharply into focus.

The most difficult part of developing an approach to default reasoning may lie not in coming up with an appropriate formalism, but in understanding precisely what distinctions and intuitions we would like our systems to capture. There are many different kinds of reasoning in uncertainty. It follows that there are many possible models, which differ without conflicting, once their appropriate domains of application are understood.

But this does not mean that the game is wide open, or that any coherently developed formal system has as much claim to validity as any other. In the heat of theorem proving and argument pressing, it is easy to forget the significance of the theorems — what they are intended to establish. Almost every formalism for reasoning these days is presented with soundness and completeness theorems, as if the existence of such a theorem by itself validated using inference engines based on them. Soundness and completeness are not properties of inference engines. They reflect a relationship between an inference engine and a formal semantics. Soundness at its most essential level says that all inferences are warranted by the semantics; completeness says that all entailments that the semantics warrants can be inferred. There are two different ways that either of these could fail: the inference engine could be defective (perhaps necessarily so, if the logic is second or higher order, for instance), *or the semantics could be inappropriate.* Because we are computer scientists, concerned with writing correct inference engines, we naturally focus on the first. But when developing models, it is at least equally important to look at the second.

A model of reasoning in uncertainty is useless unless it captures at least the most basic properties of some form of actual uncertainty reasoning. It follows that the conditions of truth which the formal semantics specifies must reflect important aspects of the real conditions under which we would want to claim that generalizations of the kind we are trying to model are true or false. Likewise, unless the definition of logical implication provided by the semantics reflects the actual criteria which determine whether one thing follows from other partial or uncertain information of the intended kind, soundness proofs become irrelevant.

In other words, formal systems must be analyzed in terms of the phenomena they model, and that they are representationally adequate or inadequate insofar as they do or do not reflect the structure of those phenomena. To be an adequate model of some aspect of reasoning in uncertainty, a formal model must provide a semantics which preserves both truth conditions and implications when we go from the motivating phenomena to their correlates in the model. In particular, the interpretation of non-universal generalizations in the formalism must preserve the truth and implication conditions of some interesting set of actual non-universal generalizations, or the formalism has no clear realm of application. It is this fact that makes distinctions among kinds of generalization so important: formalisms for reasoning in with them will be expressively adequate only insofar as they preserve those distinctions. This paper establishes some distinctions within the realm of reasoning in uncertainty, and then describes requirements on systems of default reasoning as knowledge representations, based on those distinctions and on the natures of the different tasks involved.

The paper is structured as follows. Section two discusses the difference between decision theory and logic, and argues that trying to locate decision theory in logic, whether for default reasoning tasks or for any other ones, is a mistake. It establishes what tasks belong to each or these portions of a default reasoning system, and what criteria apply to each. Section three establishes distinctions among four kinds of generalization, which will be used to structure the discussion of different kinds of default reasoning systems. Section four discusses other distinctions which at least some knowledge representation systems for different default reasoning tasks will want to retain. Section five discusses additional requirements for various kinds of logics of uncertainty and incompleteness, and section six provides conclusions and some preliminary remarks toward unifying the different kinds of default reasoning.

Section 2: *DECISION THEORY VERSUS LOGIC*

In contexts of incomplete information or uncertainty, we are often faced less with answers than with choices. There are at least two tasks which a system reasoning in contexts of uncertainty could be expected to perform. First, we want it to identify the reasonable candidates. Second, whenever possible, we would like an adjudication among the various options. This section argues that these are distinct tasks, to such an extent that they actually belong to different parts of the system altogether. Logic gives us alternatives; decision theory selects among them. We argue here that although many hold that the decision theory is the heart of any logic for reasoning in uncertainty, there are sound reasons for keeping it strictly separate from the logic.

2.1 How they are different

A logic is a formal system for telling what follows from what, and how to establish that. A decision theory is a substantive theory giving grounds for choices among options, when none can be shown to follow. The difference may be seen in either of two ways. On one view, decision theory is to uncertain situations as logic is to certain ones. That is, in situations allowing certainty, logic "selects" the options that follow (certainly); it is not necessary to ask *which* it selects, because all the possibilities that actually follow will be consistent with one another, if the initial description of the situation was coherent (that is, if the premises were consistent). In situations which do not allow certainty, logic is baffled, and decision theory picks up where it leaves off.

The problem with this view is that logic is *not* baffled in situations of uncertainty. Logic can tell us *exactly* what follows from our given situation. It can even demonstrate what the alternatives are, which of them are compatible with one another, and which are not. The difference lies not in what logic can do, but in *what we want from it*. In contexts of uncertainty or incomplete information, we will often have to choose among options none of what are entailed, and which often conflict. What we really need, in these cases, is an adjudication not of what follows from what we already believe (have already committed ourselves to), but of what further commitments we ought to make. Decision theory has "ought" built into it at the lowest level: it is a preferential evaluator. Logic works rather from "is".

2.2 *Why to keep them different*

It could be argued that the above distinction is not really important. That is, in contexts of incomplete or uncertain information, we want to know what "reasonable conclusions" can be drawn from the partial or uncertain information on hand; and this is close enough to what logic does that the same engine responsible for drawing entailments in contexts where that is possible should be doing this work. What is wrong with this, essentially, is that the engine by its nature is a single-criterion decider. When dealing with more familiar inferential contexts, that criterion is entailment. The natural impulse, then, is to extend the concept of entailment so that it will make our decisions for us in uncertain or partial contexts as well. But there are strong reasons not to do this.

2.2.1 *Multiple decision criteria depending on context*

The single most important reason not to leave a single-criterion engine in charge of decision theory is that decision-making is not a single-criterion procedure. There are many possible criteria for making decisions, and which criterion should take precedence depends very much on the circumstances. Suppose, for instance, we are working with a medical consulting system. Do we necessarily want it to choose the most probable diagnosis when faced with a new case?

Suppose that a patient has a sore throat, white spots on the tonsils, nausea, diarrhea, and vomiting. Consider two hypotheses: (A) the patient has a strep infection; (B) the patient has a strep infection and a gastrointestinal virus. Hypothesis A is necessarily the more probable of the two: the probability of a conjunction is always smaller than the probability of either conjunct (unless one entails the other or the probability of one of the conjuncts is zero, neither of which holds in this case). But surely hypothesis B is the better diagnosis.

From this example, we might conclude that what we want is really the best explanation, which we might define as the highest probability hypothesis which covers all the symptoms. But is this really true? Suppose a patient has a set of symptoms which could simply reflect a minor infection, but which occasionally indicate an early cancer. Would you want your doctor always to diagnose the infection, without checking for the cancer, just because it explains as many symptoms and has a higher probability?

In cases where being wrong can have serious costs attached, we probably want our decision theory to accommodate that. That is, we want cost of error to be a criterion, as well as probability and coverage of information. It could be argued that such a decision theory could be built into an inference engine, so that its "single criterion" is in fact a balanced sum of factors. But things are more complicated than that.

Suppose that instead of making decisions about "real life" cases, our system is trying to do something like scientific discovery. The inferences we are interested in are now things like arriving at a scientifically interesting hypothesis to explain data. Cost of being wrong is not a factor here. We don't want the system to make infer the safest hypotheses; we want it to infer the most fruitful ones. Measures of fruitfulness include such things as range of new phenomena explained, range of experimental investigations suggested, and so on. Some of these issues can be formalized more easily than others; but whatever the formalization winds up looking like, the criteria will be nearly disjoint from those of the medical consultant system above.

Now consider an economic forecaster/decision maker for a business organization. They aren't looking for fruitfulness, but neither do they simply want to minimize risk. Minimum risk schemes rarely make money. They are interested in minimizing risk *subject to* maximizing profit.

In other words, the criteria for making decisions vary from application to application. Worse yet, they vary *within* applications. There are times when doctors don't bother to get precise diagnoses at all: most frequently, when they are confident that the problem is an essentially harmless virus about which they can't do anything anyhow. In fact, the decision *whether to decide* is itself subject to all the considerations of decision theory: sometimes, it is better to wait until more information is available; other times, no decision may ever be worth the cost of getting it, and the best thing to do is never to decide (or to toss a coin, and then ignore the outcome).

Decision theory is a large and complex substantive field. Even as applied to the tasks of a single application area, the decision theoretic issues may greatly exceed anything that could reasonably be tied to a single concept of uncertain entailment, however complexly defined. The idea of trying to model such an entailment concept formally and use that model as a basis for soundness and completeness proofs is nightmarish. This just is not what logic does.

2.2.2 Knowledge-based decision criterion choice

Furthermore, there are positive advantages to be obtained by separating out the decision theoretic issues from the logic. In particular, suppose that the decision theory principles are themselves present in the knowledge base, in a form accessible to the inference engine. Then the system can use its inference engine and its representation of decision theory to *infer* whether this is a case that needs deciding, and if so, what criteria should be applied with what priorities. To put it differently, people faced with decision making tasks often reflect consciously on the decision making principles, and may decide among them on the kinds of grounds sketched above. To get systems to do the same things, we would have to provide them both with information about the kinds of situations they face and with explicit information about what matters to decision making in different situations.

The idea here is that decision making should ideally be a knowledge-based activity, and selecting decision criteria should be done on the basis of knowledge about criteria and situations. To make this possible, though, it is necessary to separate the decision theory from the inference engine. In other words, I am arguing here that logics of uncertainty and incomplete information *should not* choose among conflicting alternatives. Rather they should provide an architecture for reasoning about the alternatives which *when combined with knowledge-based application of an explicit decision theory* can let the knowledge about decision making and about the situation guide the selection among alternatives.

Logic is not about making commitments. It is not about premise selection. It is only about what follows from commitments *if made*. Decision theory is about making commitments. Its considerations are different from those of logic, and it operates at an entirely different level. We should not be confusing the one with the other.

2.2.3 Decision criterion = knowledge; application = inference

The model which this section argues for, then, is a stratified model, in which logic lies below decision theory, and applies it to provide choices in uncertain situations. The distinction between the two is sharp. The logic lies in the inference engine. It is not explicit knowledge on the part of the system; it is rather the system's competence in dealing with knowledge. The decision theory, on the other hand, is knowledge, in the most straightforward sense. It is represented as explicit rules in a knowledge base, which give the criteria on which decisions are to be based. To infer the appropriate decision in a given situation, the inference engine works on the domain knowledge base, the given information about the current situation, and the decision theory, to result in an application of the decision theory's principles to the current case. That is the inference that we are looking for. The rest of this paper discusses aspects of reasoning in uncertainty which pertain to the structure of the logic portion of this reasoning apparatus. While many existing approaches fold the decision theory also into the inference engine, this paper will largely ignore those aspects of the approaches.

Section 3: DIFFERENT KINDS OF GENERALIZATION

In English, "all" rarely means "every single thing without exception". When it doesn't mean that, there are several different things it can mean, and these different meanings can lead to very different patterns of reasoning. In this section, we present four different kinds of generalizations, pointing out the salient distinctions from the standpoint of reasoning with them. We make no claims to completeness: there almost certainly are numerous possibilities that this discussion misses. This analysis also makes no prior judgment as to the desirability of modeling any of the four kinds of generalization discussed here. Our point here is that the four are different, and sufficiently different to require distinct handling in inference systems. Hence it should not be expected that a technique adequate for handling one kind of generalization to handle the others as well.

3.1 Universal generalizations

Universal generalizations (UGs) are the kind Aristotle handles in his logic, and which modern logic deals with using universal quantifiers. These are claims which, if true at all, are true of every individual in their domains. In developing artificially intelligent systems, we can no longer ignore the issue of how the domain of a particular UG is identified. However, assuming that it has been identified by the time information is represented in a data base, the representation of UGs presents no problem for systems capable of storing and manipulating the information contained in predicate calculus style propositions.

Brachman [Brachman 1985] argues persuasively that A.I. systems must represent genuine UGs — that is, UGs which cannot be overridden by "exception" clauses — if they are to be able to encode definitions. That is, if we always treat "all" as permitting exceptions, we can wind up storing nonsense like non-elephant elephants. The point here is that at least some aspects of definition require "real live" universals. This is even clearer if we want to provide a system with a definition of subset on the basis of membership, for instance, and allow the system to infer things like, if $x \in A$ and $A \subseteq B$ then $x \in B$, then for the definition of subset had better not allow exceptions.

In addition, if all generalizations allow exceptions, then finding $\forall x P(x)$ and $\sim P(a)$ only indicates an exception. It does not indicate a contradiction. The problem with this is that consistency maintenance can be used as a constructive constraint, but only if inconsistencies are detected as such. Hence it seems necessary to encode UGs in knowledge representation systems. However, this necessity at least prima facie presents no special problems.

3.2 Statistical generalizations

Statistical generalizations (SGs) are generalizations based on some interpretation of probability. In these generalizations, "all" means most, the majority, on the average, in the expected case, or some similar claim. Probabilities can be reported with varying precision, depending on the level of knowledge and interest. Measuring them forms the subject matter for an entire discipline, and so clearly goes beyond the scope of the present report. But because these form a very important class of generalizations which has deservedly received a great deal of attention in AI, it seems appropriate to go into some depth on their nature and a few important aspects of their behavior.

The philosophical nature of probabilities matters less for AI purposes than what kinds of phenomena classical and Bayesian probability analyses model. However, given the vehement disputes on this issue, a few observations may be useful. Statistics begins investigating probabilities in any particular instance by defining (at least loosely) a space of outcomes, that is, mutually exclusive observations of test results. Events are sets of outcomes from that space. When probability theorists refer to probabilities, they typically mean event probabilities, that is, the likelihood that the outcome of a particular test will belong to the set which defines the event. This likelihood is traditionally defined in terms of frequency: given a "sufficiently large" number of tests, what proportion of all outcomes fall in the event set?

The frequency view has been attacked for centuries; a recent criticism can be found in [Cheeseman 1985]. Probably the most persuasive argument against the frequency view from an AI standpoint is that on that view, each event has exactly one correct probability. But for AI purposes, such a probability is neither attainable nor in some cases even interesting. Rather, we are interested in the probability of an hypothesis *given the current evidence.* Critics further object that the frequency theory "restricts probability to domains where repeated experiments (e.g. sampling) are possible, or at least conceivable" [Cheeseman 1985]. In addition, the concept of "long run frequency" has bothered people for centuries. How long? How do you know? Why should "large numbers" (how large?) have special properties?

These objections can be met without deserting a frequency-based approach. The probability of any hypothesis on the basis of the current evidence can be — and in normal statistical practice is — interpreted as the conditional probability of the hypothesis given the conjunction of events which that evidence reflects. In other words, in addition to a single, well-defined probability for every event over the space, the frequency view also provides a way to represent precisely the relativized probabilities we are most interested in (and these are exactly the probabilities that statisticians investigate).

Classical statistics texts also contain chapters on game theory and decision theory which describe techniques for estimating probabilities on the basis of very small samples (see e.g. [Freund and Walpole 1980] Chapter 9, or almost any other freshman text). So not only does classical statistics recognize that this can be done, the theory instructs the interested in how to do it; only, it also warns not to place great faith in the accuracy of such estimates.

The hardest question to meet is the philosophical question of the significance of the Law of Large Numbers: what does it mean to talk about "long run" frequencies? Classical statistics provides some tests for whether an actual sample is large enough; but that cannot answer the philosophical question. The best that can be said here is that other approaches have their own philosophical questions that they cannot answer, but none of these philosophical questions seem to affect AI.

The classical alternative to the frequency view is the subjective probabilities view, which derives from the views of the 18th century English clergyman Thomas Bayes. On this approach, probabilities measure certainty levels. Two options here should be distinguished. The first is well-defined, and clearly subjective (as philosophers use the term): the probability of an event given the current evidence is the measure of the degree to which a particular specific "real live" individual believes that the event will occur on the basis of that evidence. The problem here is evident: people will believe all sorts of things, and different things at different times, for different reasons or none at all. There is no reason to suppose that one person's "probability" in this sense will match another's, and no grounds for a *science* of probability at all.

It is unlikely that many supporters of subjective probabilities ever meant that, though they often seem to say it:

> ... the following definition is put forward as one that withstands all previous criticisms: *The (conditional) probability of a proposition given particular evidence is a real number between zero and one, that is a measure of an entity's belief in that proposition, given the evidence.* ([Cheeseman 1985]; emphasis in original)

The alternative, and the view that is actually held, is that probabilities measure how much *an ideal rational subject ought* to believe that an event will occur, given the evidence. This option makes probabilities relative (to evidence), but not really subjective: no actual subjects are involved any more. This approach has two difficulties, both as obvious and as pressing as the problem the frequency theory has with understanding the long run. First, what makes someone an ideal rational subject? Probability cannot be considered well-defined on this view until that is spelled out. Second, how other than by measured frequencies can we establish the degree to which such a subject ought to believe that a given event will occur?

The mathematics for measuring probabilities is the same on both these competing definitions: Bayes's Theorem is a theorem of classical statistics, for example. The significant differences come in questions of when it is legitimate to apply the formulas, and what they can be taken as establishing. In this regard, it seems that the frequency analysis has an advantage: designers of AI systems generally care less whether their systems "ought" to believe their answers than how often those answers are right. For systems whose judgments have practical consequences, we should measure and maximize that if we measure anything. But whatever philosophical view of probabilities we take, the mathematics always agrees with long run frequency expectations in all situations in which we can make sense of them.

Finally, some simple properties of probabilities should be noted. Events are independent provided that whether an outcome belongs to one does not affect how likely it is to belong to another. The joint probability (probability that all events will occur) for independent events is the product of the probabilities of the events. Since all probabilities lie between zero and one, the joint probability of several independent events is always smaller than the probability of any one of them, unless all but one have probability one or at least one has probability zero. For dependent events, the joint probability is at most the minimum of the individual event probabilities, and it reaches that level only if the corresponding event entails all the others. The joint probability for dependent events may be zero even though none of the individual probabilities is (it will always be so if at least two of the events are mutually exclusive). More subtly, the joint probability of, say, six events may be zero even though no two of them are mutually exclusive, if, say, five of them together exclude the sixth.

Similarly, the probability that an outcome will fall into at least one of several independent events (the probability of their disjunction) is the sum of the probabilities of the events in question. If they are dependent, it is at least the maximum of the individual probabilities, and at most their sum (or one, whichever is smaller).

It is common in AI contexts to assume independence in the absence of information showing otherwise. A false assumption of independence never underestimates the probability of disjunctions, and usually overestimates it. For conjunctions, the situation is more complicated. If the events involved are dependent and mutually supportive (the probability of any increases the probabilities of the others), a false assumption of independence always underestimates the probability of the conjunction. If, on the other hand, the events are mutually antagonistic (the probability of any reduces the probabilities of the others) a false assumption of independence overestimates the probability of the conjunction. If some are mutually antagonistic and others mutually supportive, how the actual joint probability is related to the one derived from a false assumption of independence is anybody's guess. In a long chain of reasoning involving disjunctions and all three kinds of conjunctions, these offsetting errors may prove very hard to detect and isolate.

In addition to these problems, combining statistical inference with ordinary logical inference involves some nasty complications. Statistical measures do not combine truth functionally, and hence do not follow the rules of standard first order logic. That is, if $P(A)$ is the probability of A, $P(B)$ is the probability of B, and $P(A \wedge B)$ is the probability of $A \wedge B$, there is no function f such that $f(P(A), P(B)) = P(A \wedge B)$. In English, the probability of a conjunction is not a function of the probabilities of the conjuncts. We can see this immediately from the remarks above on the behavior of probabilities of conjunctions. Different functions are needed to compute the probability of $A \wedge B$ depending on whether they are dependent or independent, and the functions for deriving the probability in the case of dependent events require not only the probabilities of A and of B but the conditional probabilities of A given B and of B given A also.

This absence of truth-functionality becomes even clearer when a causal relationship is sought or presumed. In this case, probabilities become inextricably linked to the theoretical context, and in some sense take on a different meaning. Given one set of results R, the probability of R will differ depending on the hypothesis relative to which it is computed. More importantly, what changes tends to be not the probabilities of individual occurrences, but precisely the probabilities of coöccurrences. That is, the probability of the conjunctions changes, without that of the conjuncts changing.

Saying that the probability of $A \wedge B$ is not a function of the probabilities of A and of B, does not mean that it is not a function of A and B themselves. Given A and B and sufficient information about them, we can calculate the probability of $A \wedge B$. To calculate probabilities of complex propositions on the basis of their logical form, though, we would need a well-defined function reflecting the relationship between the *probabilities* of A and of B and that of their conjunction. But none exists. It remains only to note that this problem affects not only conjunction, but all connectives except negation.

The problem strikes even deeper. Traditional statistics does not even measure or combine probabilities for *propositions*. It measures probabilities of *events*, where an event is defined as a set of sample points in a sample space, whose points are taken to represent mutually exclusive outcomes of tests. By the conventions of ordinary propositional and predicate logic which form the basis for both the deductive and the formal semantic rules, no two atomic propositions are mutually exclusive. Hence atomic propositions do not reflect outcomes, so do not correspond straightforwardly to points in a sample space. Nor can non-atomic propositions be viewed as events over a sample space of propositions, unless we are extremely careful to restrict those atomic propositions to outcomes in the statistical sense, say by including the necessary axioms.

This is not to say that no systematic rules for determining complex event probabilities exist. On the contrary, the discipline of statistics has formulated such rules very precisely. The claim here is that to do so in an A.I. system, statistical laws must be implemented directly. Not only does logic provide no short cut here, it cannot even be combined with statistical operations in the obvious ways.

3.3 Generalizations arising from degree-of-applicability claims

A different kind of uncertainty centers on the extent to which a given property applies to an individual. This is the issue of vagueness, and the kind of inferences justified on the basis of degree-of-applicability are different from those based on other kinds of uncertainty. These are claims that involve "mushy quantifiers" like "sort of", "rather", "not very", "-ish", and the like, as for instance in the proposition "Fred is a youngish man".

The most common models for these degree-of-applicability claims (DACs), fuzzy set theory and fuzzy logic, look superficially a great deal like probabilities. In fact, however, DACs do not work like probability. Consider the following two claims about Oscar the Ostrich:

(i) Oscar is a (typical) bird at 0.6
(ii) Oscar is male at 0.5

Claim (i) says that Oscar is not very birdlike (ostriches aren't), although he is more birdlike than a lot of other non-birdish things (Oscar is more than 0.5 birdlike, because, for instance, he is more like a bird than, say, reptiles are.) This is the sort of claim fuzzy set theory was originally developed to handle; it tries to measure the extent to which an individual falls in the bounds established by a fuzzy concept. Claim (ii) is a probability claim, reflecting that the system doesn't know whether Oscar is male but does know that Oscar is a bird, and that half of all birds are male, making the chances that Oscar is male 50-50. That is *not* to say that Oscar is half male: the system can consistently hold (ii) and also hold that any given bird is either completely male or not at all.

The claims look superficially alike, but they cannot be taken the same way: (i) says that Oscar is not a very typical bird; (ii) does not say that Oscar is not a very typical male. The claim which (i) embodies does not really reflect incomplete information at all: it reflects a fundamental fact about how Oscar relates to a vague concept. In case (ii), the information is incomplete and can be completed by a single experiment (look at Oscar and see). If no difference in representation reflects this basic difference in content, the system will reason incorrectly a good part of the time. Translations of DACs into probabilities do not preserve inferences.

3.4 Generalizations as abbreviations

Generalizations as abbreviations (GaAs) are generalizations whose intuitive meaning is "If x is of kind K, then x has P unless I know of some problem preventing that." A GaA is a short cut, in place of actually stating a (perhaps cumbersome) list of known exceptions. Any GaA can be transformed into a correct UG as follows. Let Γ represent a "quasi-quantifier" for "in general". Suppose the GaA has the form $\Gamma x\phi(x)$, and the known exception conditions are $\phi_1(x),\ldots, \phi_n(x)$. Then $\forall x(\sim[\phi_1(x) \vee \ldots \vee \phi_n(x)] \rightarrow \phi(x))$is logically equivalent to $\Gamma x\phi(x)$.

The Closed World Assumption (CWA) is a global assumption about a data or knowledge base to the effect that all true positive ground facts are known [Reiter 1978]. GaAs are naturally related to the CWA and hence have many useful applications in fields like intelligent data bases. If exceptions to generalizations all took the form of positive ground facts (or could all be proved from them), then the CWA would entail that all exceptions to a given generalization are known. GaAs go well beyond the bounds of the CWA, since they do not restrict the form of possible exception-causing propositions (the ϕ_i need not be either positive ground facts or provable from them); but the spirit is similar. GaAs also have a great deal in common with the spirit of PROLOG, which does not include true negation, but does have negation by failure. That is, in PROLOG, there is a sense in which you cannot say "If $\sim\phi$ then ψ", but you can say "If I don't know ϕ, then ψ".

Since GaAs are assumed to have equivalent formulations in standard first order logic, and since these formulations can *ex hypothesi* be recovered from the knowledge base, their existence and use do not require us to formulate a new logic. We can simply treat them as the abbreviations they are, within a standard first order predicate logic. That is, they may be implemented in ways that are very different from implementing their equivalent universals. But insofar as that implementation can be shown to allow the same inferences as the formulation in genuine universals would, we can use first order logic in determining issues like soundness and completeness. This is not to claim that within A.I. systems, adding GaAs introduces no new behavior. But when we go to justify the conclusions of the systems in logical terms, we need nothing beyond standard logic.

There is an important sense in which GaAs are closer to UGs than to SGs. Suppose our knowledge base contains a generalization which says "In general, things of kind K have property P," and suppose it can show that x is of kind K and none of the exception clauses holds of x, but it also contains the proposition that x does not have property P. If the generalization were statistical in nature, this would simply say that an unusual thing has come to pass; but if it is a GaA, then the knowledge base contains an inconsistency. GaAs do not allow unknown (or unprovable) exceptions. Discovering something which cannot be proven to be an exception but which nonetheless is constitutes reason to worry. Something is wrong. If any exceptions exist which are not demonstrable by means of information in the data base, then the GaA is false.

3.5 Typicality-based generalizations

The generalizations of the fourth kind, which we will call typicality-based generalizations (TBGs), are subtly different from any of the foregoing kinds. These generalizations are based on something like a prototype notion (see for instance [Rosch 1975], [Rosch and Mervis 1975], and [Rosch et al. 1976]). They report facts which typically hold for members of a given class (or objects answering a given description, or so on).

TBGs encode things it is reasonable to presume: they warrant not inferences, but presumptions. Inferences show what follows from what. TBGs do not tell you that something follows; they tell you that it might be reasonable to suppose something. They do not add conclusions; instead they suggest presumptions. This warrant of presumption, however, transmits across propositions the same way that inferences do: presumability is inherited truth-functionally. If there is reason to suppose that A, and there is reason to suppose that B, then there is reason to suppose that $A \wedge B$ — although there may be better reasons not to suppose the conjunction, as for example when it proves logically false.

TBGs are neither statistical claims nor abbreviations for UGs with known deviations. Rather they reflect situations of genuinely partial knowledge, in which the information present warrants making certain presumptions, without warranting their actual assertion. Because TBGs are often hard to tell from other kinds of generalization, we contrast them with each class individually.

3.5.1 TBGs versus UGs

[Brachman 1985] explains this distinction clearly. To use his example, if we say that all elephants are mammals which are gray and which have four legs, and if we treat "all" as indicating genuine universality, then we have no way to talk about Clyde the unfortunate amputee elephant with only three legs (or about more natural examples like albino elephants). On the other hand, suppose we always treat "all" as indicating a non-universal generalization. Then we can talk about Clyde the three-legged elephant, but unfortunately we can talk with equal ease about Clyde the non-mammalian elephant, or even about Clyde the non-elephant elephant.

In other words, if both non-universal generalizations and UGs are to be dealt with, they must be clearly distinguished. A.I. systems must either implement genuine UGs or abandon definitions altogether. At the same time, if we want to be able to deal with situations which are atypical *in unpredictable ways*, we must incorporate in our systems some form of TBG in addition to and distinct from UGs, so as to provide for atypicality without ruling out lawfulness.

3.5.2 TBGs versus SGs

TBGs cannot be treated like statistical claims either, although the difference here is more subtle, and the need to avoid confusion harder to argue. Most people realize that over half the population is female. Yet in the absence of information concerning a person's sex, one does not usually presume that the person in question is female (indeed, the presumption tends to go the other way). On the other hand, the number of flightless birds (emus, ostriches, kiwis, penguins, and so on) is hardly negligible. Yet we feel justified in presuming of birds in general that they fly.

Pace Cheeseman [Cheeseman 1985] and many, many others, not all that is not universal is probabilistic. For instance: if, as Cheeseman claims, the by now tormented example "Birds fly" really means "Most birds fly", then birds don't fly in the spring. In nesting season, baby birds outnumber adults. Baby birds don't fly. Hence in nesting season, "Most birds fly" is false. (By the way, we can do even better with "Birds lay eggs," which is out-and-out false year round of at least half the population: none of the males do.) So if Cheeseman is right, anyone who says in the spring that birds fly or at any time that birds lay eggs is mistaken. This is nonsense.

"Birds fly" must be decoded with respect to typicality. If typicality can be modeled by any statistical concept, it is category cue validity, not probability [Rosch 1975; Rosch and Mervis 1975;

[Rosch et al. 1976]. "Birds lay eggs," on the other hand, is not statistical at all. It is shorthand for a genuine, accept-no-substitutes universal — but not for "For all x, if x is a bird, then x lays eggs". Instead, it is in a class with the non-universal generalizations "Mammals bear young alive" (duck-billed platypi lay eggs) and "Reptiles and fish lay eggs" (garter snakes and sharks bear live young). By the way, these generalizations cannot be translated straightforwardly into probability claims counting over species instead of individuals: *no* species *either* bears live young *or* lays eggs; only (female) individuals belonging to species do.

TBGs usually represent causal claims, albeit masked and incomplete ones. Most birds fly, not by accident or coincidence, but because the features which distinguish something as a bird evolved to facilitate flight. There are exceptions to the rule (birds that don't fly) because evolution did not always stop at that point. But in general the rule holds, and for a good reason.

Statistical claims are frequently (though not always) assumed to be related to some causal fact. The difference is that statistical claims are evidence for, rather than embody, a causal claim. Furthermore, many statistical "facts" result from accident: we accept as supporting causal claims only certain statistical evidence. In particular, the evidence must show a persistent trend among data concerning phenomena which there is independent reason to suppose might be relevant to one another.

For example, I recall reading somewhere that for many years, the membership rolls of a bakers' union in New York City precisely paralleled the births and deaths in a town in India. Whether this actually happened is not important here; my point is, it could happen, and if it did, no reasonable person would take it as anything more than a striking (and somewhat humorous) coincidence.

A particularly dramatic difference between the logic behind TBGs and the logic behind SGs lies in the transitivity of truth-functional inferences based on generalizations. Presumability can be inherited through truth-functional inferences; but statistical relationships are far more complex, and statistical inferences follow utterly different rules (see section 3.2 above).

This distinction between judgments based on prototypicality and SGs is not new. Rosch's group [Rosch et al. 1976] relates prototypicality to category cue validity, and then points out (page 384) that "category cue validity is not a probability ... it does not have the same set theoretical properties as a probability." The following passage is a long footnote to that remark.

> Were category cue validity a true probability, the most inclusive category would always have the highest validity. This follows from the fact that if category A includes category B, the probability that x belongs to category A always exceeds the probability at [*sic*] x belongs to category B. Category cue validity refers to a psychological factor — the extent to which cues to category membership are available at all (attributes common to the category) and the extent to which those cues are not misleading (attributes which do not belong to other categories). This measure disregards the base rate probabilities of membership in categories — as do most people.... [p. 384-385]

I belabor this point, because there is a strong temptation to view all non-universal generalizations, and especially TBGs, as a sort of SG. Generalizations are naturally associated with degrees of certainty or uncertainty, bounded below by known-false and above by known-true. Any property of propositions which can reasonably be measured on a continuous scale bounded below and above has a tendency to be assimilated to probability. This tendency is dangerous: the laws of probability distribution cannot be arbitrarily extended to other domains.

3.5.3 TBGs versus DACs

The difference between DACs and typicality is subtle but real. Typical birds fly. But how typical a bird Tweety is does not measure how well Tweety flies, or how even how likely Tweety is to fly (hummingbirds are atypical in many ways, but spectacularly good fliers). DACs are based on vagueness: they measure the extent to which an individual falls under an inherently vague concept. Typicality can be viewed as a sort of vague concept, and the idea of measuring typicality itself by a DAC is within limits tempting. But typicality-based generalizations are not simply estimates of how typical an individual is. Rather they are generalizations which indicate what other properties — vague or otherwise — things of a certain kind tend to have. The degree of typicality does not translate to a degree of having these other properties, or even to a likelihood of having any particular one of them (the second would bring us back to SGs).

3.5.4 TBGs versus GaAs

GaAs can be rephrased equivalently as universal implications whose antecedent says that no exceptions occur. The question, then, is whether TBGs can also be rephrased in this way. How could that fail? There are two reasons why such a list of exceptions may not be possible. First, it may not be possible to predict exceptions: this is a failure in our knowledge. Second, it may be that no such list is possible, regardless of the state of our knowledge, because of the nature of the concepts involved. We take up these two problems separately.

If we knew everything there was to know about every individual in every domain which our system would ever deal with, then clearly we would not need any kind of generalizations — including universal ones — although they might prove convenient as abbreviations. The CWA is not the only assumption under which TBGs collapse into GaAs.

Consider the following "quasi-closed-world" principle: the system knows all rules relevant to its domain, and has procedures which let it identify, obtain, and verify all currently unknown individual facts. Under this assumption, TBGs would collapse into GaAs. However, in many domains, this assumption is very nearly as unrealistic as the "pure" closed world assumption, since it requires us to identify all possible exceptions to any generalization and determine in fact whether or not they are exceptions before we make any inferences using the generalization. But in that case, two things follow. First, before we use any generalization, we must be able to determine the truth value of its unquantified version for every legitimate instantiation. This is not practical. Second, if we must in fact make all those determinations before we may use the generalization, what good is it? Why not simply determine the case one is interested in and be done with it?

Let me call the "open world assumption" the view that the system's knowledge is essentially incomplete, in the sense that there may always be relevant information which the system lacks and cannot obtain (indeed, may not even be able to identify). If the open world assumption holds, then no version of GaAs will handle Clyde-like cases, and so TBGs constitute a separate class of generalization.

The foregoing difficulty is essentially practical: it relates to the difficulty of acquiring adequate information. There is a further problem, which is not practical but theoretical in nature. It may seem that generalizations result from imprecision, and that to eliminate them we need only obtain sufficient knowledge and then state our actual knowledge precisely. But even ignoring all practical difficulties, this is not always true.

In some cases the range — not the identity, but the range of possible identities — of "pathological" examples (flightless birds, white ravens, three-legged elephants, etc.) cannot be wholly determined *a priori*. Second, and more importantly, we may need to deal with cluster concepts (see e.g. [Wittgenstein 1953]; Rosch [Rosch 1975] cites this as a forerunner to her view of prototypicality), for which persuasive arguments have been made that precise definitions are *in principle* impossible.

The classical example of such a concept is "game". Many games involve more than one player, but some do not. Many games involve teams, but some do not. Many games involve boards, or balls, or cards, but some do not. It has been argued that in fact, there is nothing that all games have in common, beyond simply being games. What holds the class together is a sort of "family resemblance", not specific shared traits. If this view is correct, and if we cannot express TBGs as opposed to GaAs, we could hardly say anything about games that was true.

It remains only to note that many concepts seem to act like cluster concepts. While precise definitions in terms of UGs may be possible, they are surely very difficult to find, and we should not count on being able to find adequate ones. A system which includes no TBGs, and which does not want to make frequent false assertions, could say very little in such cases.

In summary, there are at least four kinds of propositions involving "all": genuine UGs, SGs, GaAs, and TBGs. To capture ordinary reasoning appropriately (and to avoid conclusions which are not justified by the data on which they are based), these must be distinguished.

3.5.5 Problems TBGs raise

Perhaps the most commonly approach to implementing TBGs takes the form, "in the absence of evidence that ~A, you may infer A" (see e.g. [Reiter 1980], [McDermott and Doyle 1980] and [McDermott 1982]). This kind of default is implemented as follows. When the system is asked "A?" and finds the default rule, it attempts to derive ~A. If it fails to do so, it then returns A as the answer (depending on the details of the system in question, it may also build A into the data base). Hence systems augmented by this kind of rule can take advantage of a version of non-universal generalizations. So far, so good.

But this procedure only looks reasonable for what I call TBGs so long as we deal with questions like "Can Roger the bird fly?". Then, saying "Of course, he's a bird" seems unobjectionable, but only because nothing depends on the answer. If we don't care what the answers to our questions are, there is little motivation to implement any form of reasoning with TBGs. If we don't care, we just say "I don't know".

But suppose that we do care what answer we get. For instance, consider a medical consulting system. Suppose that for a particular syndrome S, treatment x is generally very beneficial, but that in exceptional cases treatment x kills. Now if Smith has syndrome S, we do not want to recommend treatment x just because we don't know that Smith is exceptional. On the other hand, if syndrome S can itself prove fatal, neither do we want to say that we don't know anything about what to do for Smith.

What we would like the system to do in this kind of case is give a guarded response: that is, say something like, "Treatment x usually helps in cases like this," or "Presumably treatment x helps." A better answer would tell the user directly what the counterindications are; but at the very least, a responsible system should warn the user that the information results from a presumption, and not an inference. Once the system has issued the warning, the user can pursue it with further questions.

Another way to put the point is this: the implementation described here captures GaAs. These it handles adequately, because exceptions are already known and demonstrable. But it cannot handle TBGs, precisely because of the differences between them and generalizations as abbreviations. Hence this implementation models a real sense of "all", but not TBGs.

Much of the unhappiness over the existence of multiple extensions in systems such as Reiter's [Reiter 1980] or McDermott and Doyle's [McDermott and Doyle 1980] [McDermott 1982] can be traced to a combination of this strategy with equivocation between GaAs and TBGs. The strategy for producing inferences owes much to GaAs; the recognition of inconsistent extensions arises from dealing with TBGs. Insofar as these logics model GaAs, inconsistent extensions are inappropriate; but in that case, the entire formalism becomes unmotivated, since standard first order predicate logic can accommodate GaAs (see section 3.3 above). Insofar as they model TBGs, inconsistent extensions are not only well motivated but necessary; but the strategy, which uses logic to choose among them, now seems inappropriate, forcing decision theory into the wrong domain (see section 2).

A further difficulty with TBGs lies in deciding what it means for them to be true or false. "If Roger is a bird, then presumably Roger can fly" can be true even if Roger is a bird, but Roger can not fly. Indeed, "Presumably Roger can fly" can be true even though "Roger can fly" is false. That is the whole point of saying "presumably": it protects the speaker from saying something false when the facts go the "wrong" way. On the other hand, the generalizations do mean something: it follows that it must be possible for them to be false.

In other words, the truth value of TBGs cannot be a simple function of the truth values of the component propositions: TBG operators are not truth functional. Furthermore, there is good reason to despair of logic ever giving an adequate account of how the truth values of propositions involving TBGs depend on their components. TBGs make sense because they reflect non-logical connections among their constituents. The missing information guarantees that their content cannot be a simple function of the contents of the components. But then we should not expect to be able to give a purely logical account of TBGs (see [Israel 1980] for a discussion of this point).

Section 4: *INFORMATION TO BE RETAINED*

Of course, the distinctions between different senses of "all" do not exhaust the distinctions important to reasoning in uncertainty. Not only are there many other important issues, there are important distinctions which do not fall neatly into the boundaries of any one kind of generalization, but affect several or even all. Some of these place constraints on acceptable knowledge representations, in that they indicate differences which representation schemes must reflect.

One important distinction is grounds of doubt. While all of the substantial information in any system may be open to some degree of doubt, it will not as a rule all be subject to the same *kind* of doubt. There may be differences among kinds of doubt which go beyond what can effectively be reflective by numeric measures, however complex. We take this up in section 4.1. Section 4.2 concerns the difference between results of default reasoning (presumptions) and results of more standard logical reasoning (which for the moment I will call inferences), and argues that this difference and the manner in which it is represented have effects not only on what is believed, but on what is expressible. Section 4.3 makes a further distinction between making inferences and presumptions on the one hand, and committing to premises on the other. Section 4.4 looks at the distinction between uncertainty and incompleteness, and section 4.5 discusses the distinction between uncertainty which results from the state of our knowledge and uncertainty which results from inherent properties of the state of affairs.

4.1 Differing grounds of doubt

Consider the following argument. In any non-trivial field, the best information available always contains errors. The real difference between UGs, GaAs, and TBGs lies in the degree of confidence they reflect and command. Since nothing commands absolute confidence, why single these out? Why not just associate confidence levels with all statements, work out laws for their combination, and implement some version of belief revision to recover from detected anomalies?

TBGs need a separate logical treatment, because they are logically different from other sources of doubt. When I ask how confident I am of *A*, I could mean any of the following: (1) how certain am I that *A* is true? (2) how good is my evidence for *A*? (3) how likely is *A* on the basis of my evidence and theories? The first two questions, frequently confounded, are in fact distinct. My confidence in *A* reflects psychological phenomena which may be completely independent of my evidence for *A*. I can be highly confident of *A* despite strong evidence against *A*, or in the complete absence of evidence one way or the other, simply because for some reason, *A* appeals to my less rational nature. On the other hand, the strength of my knowledge as evidence for *A* is (at least arguably) an objective fact, subject at least in principle to some sort of determination. It is important to note, though, that all the evidence we have (including all the evidence for our evidence) is rarely compelling, since we are rarely certain of our premises.

However, it is possible to make certain that given our premises, our conclusions are not merely likely but certain. That is, we can make sure that our conclusions are no worse than our premises — but only if we either outlaw all generalizations or distinguish them carefully from UGs.

If a system contains false premises, we cannot protect it from drawing false conclusions, by logic or by any other means (except by preventing it from asserting anything at all). But we can protect our systems from drawing false conclusions from true premises. TBGs represent a very special kind of doubt, namely the doubt concerning a conclusion which arises when we know that even if all our current beliefs are true, they only suggest, and do not entail the conclusion.

4.2 Inferences versus presumptions

When we talk about systems that reason in contexts of uncertainty, we may have either of two very different things in mind. We may be talking in short range terms about systems dealing in very limited ways with very limited domains, and with "tame" uncertainties: tame in the sense that we understand at system design time the nature and range of the kinds of uncertainty the system will be called upon to deal with. This is the situation with systems which we might anticipate actually using (as opposed to doing research on) in the next five to ten years. So long as this is the case, we can with reasonable safety tailor special purpose algorithms which, while they may have severe shortcomings relative to the general problem of uncertainty, are safe enough within the context of the system's intended application.

On the other hand, we may be talking in long range terms about target systems of a more open-ended kind. Their domains may be undetermined as yet; and the kinds of uncertainty they need to deal with may be far from tame. But we do know some of the things we want them to be able to do, and these desired functions place constraints on the mechanisms they can use for reasoning in uncertainty that can be used to guide research.

We want these systems to receive new information, store representations of it, perform inferences using new information, and report the results of all these operations in some reasonable fashion. They should be able to reason with generalizations in situations of incomplete information. At least in principle, they should be able to represent any information in their domains in which a human might be interested, and form propositions expressing that information. What "in principle" means to current research is that while our current prototypes lack these abilities, we know that we want them, and we anticipate adding them to our systems, although it will generally take research to find out how. One thing such systems will need is a distinction between presumptions and inferences, because systems without such a distinction cannot support the capacities listed above, even in principle.

Suppose that we start out in situation *S*, with information (including both unqualified beliefs and generalizations) which supports conclusion *C* without strictly entailing it. For instance, we know what birds are, and we know that Roger is a living unplucked bird. This information supports the conclusion that Roger flies, but we know that this might fail. This is situation *S*, and "Roger flies" is *C*.

Now suppose that we learn something new which contradicts our previously supported conclusion. In our example, we might learn that Roger is a kiwi. Call the new situation S^*.

Supporters of most nonmonotonic† approaches would describe the situation as follows. S entails C, S does not entail $\neg C$, S^* entails $\neg C$, and S^* does not entail C. This set of entailment relations provides a consistent basis for stating whether or not, to the best of our current knowledge, Roger flies. But it is not a sufficient basis for all of our relevant knowledge in either S or S^*.

In situation S, a careful speaker would not say "Roger flies," but rather something like "It is reasonable to suppose that Roger flies," or "There is reason to believe that Roger flies." These qualifications would not be placed on the front of all conclusions about Roger: for instance, "Roger has feathers" is justified absolutely by our beliefs in S. We may be mistaken that Roger is a living unplucked bird, but if we are not mistaken about that, we cannot be mistaken about his having feathers, since by virtue of the biological definition of birds, they all without exception have a genetic disposition to produce feathers.

The relationship between our knowledge and the conclusion that Roger has feathers is fundamentally different from the relationship between our knowledge and the conclusion that Roger flies. This has to do with the difference between UGs and TBGs (see section 3.5.1). People are frequently interested in this kind of difference: questions like "Are you sure?" would not otherwise be so common. There is an important distinction between expectations and knowledge.

But nonmonotonic approaches treat the relationship in both cases as entailment. Worse, they produce exactly the same conclusion in situation S as they would in situation S^{**}, where we add the information that Roger is a mature parrot, uncrippled and in normal health, and that he had enough liberty when young to learn how to fly. But we know that Roger the parrot flies, whereas we only suspect that Roger the bird flies. And we already know in situation S that this uncertainty exists. This is the first shortcoming of nonmonotonic approaches: they lose the distinction, present in the "real life" situation, between justified beliefs and justified presumptions. (For more on this point, see section 3.5.)

---

† In all standard logics (and most non-standard ones), if a proposition C follows from a set of propositions S, and if S is a subset of S', then C also follows from S'. That is, adding information without eliminating any cannot eliminate entailments. This property of a deductive system has been called monotonicity.

In situation S^*, there is a further loss of information. When we have found out that Roger is a kiwi, we don't just know that he doesn't fly. We know that

(a) Roger is a bird; so
(b) there is reason to believe that Roger flies; but in fact
(c) Roger is a kiwi, so
(d) he doesn't fly after all.

Notice that (b) and (d) not only do not contradict one another, they contain substantial information: ***they tell us that an expectation has failed.*** A nonmonotonic setting forces the difference between justified belief and justified presumption into the logic instead of leaving it in the sentences themselves (so that "presumably Roger flies" becomes the same proposition as "Roger flies"). In that setting, (b) and (d) represent an out-and-out contradiction. To maintain consistency in the logic, one of them is rejected. As stated above, S^* entails $\sim C$ ("Roger does not fly") and does not entail C ("Roger flies" — which under a nonmonotonic approach is the same as "Presumably Roger flies").

So in situation S^*, nonmonotonic approaches can not support inferring that presumably-C and $\sim C$. But this is certainly information a human might be interested in. This problem arises explicitly from nonmonotonicity. The information lost in S^* is precisely what is known in S. If that information is not lost, then adding premises can not cause conclusions to be rejected, and the logic is by definition monotonic.

Hence we have two kinds of information which nonmonotonic approaches fail to support. First, they fail to support the distinction at the propositional level between knowledge and supposition. Associating a measure instead of a truth value with propositions cannot do the work needed here, especially if the metric is taken as measuring probability or confidence: the probability that a single fair toss of an unbiased coin will land heads is 0.5; and of that fact we are very confident. Knowledge of probabilities *is itself knowledge* which may be either justified completely by other knowledge or only suggested. In any case, TBGs are not statistically based. Second, because nonmonotonic approaches fail to support the distinction between knowledge and supposition, they also fail to support reports of failed expectations.

Consider the following two descriptions of birds.

(a) Felix is a bird who lives in North America. He is under four feet tall, he flies, and he travels slowly on the ground.

(b) Oscar is a bird who lives in Africa. He is over four feet tall, he can't fly, but he travels rapidly on the ground.

Which of these birds do we know more about? Felix could be almost any North American bird (except a road runner). Oscar is an ostrich, and couldn't be anything else. Yet both descriptions give only the continent they live on, their height in vague terms, whether they fly, and their speed on the ground. How do we come to have so much more information about Oscar than about Felix?

When generalizations hold, that fact doesn't tell us much. But when they fail, their failure conveys a lot of information. Oscar's height, flightlessness, and speed on the ground are all atypical features; together, they pin down his species. (If he were Australian, he'd be an emu.)

Salience has been described as the key to a major problem of natural language generation: what and how much to say (see e.g [Conklin and McDonald 1982]). Most techniques for determining salience depend on either marking particular properties for a class of objects or determining differences between a pair of objects when that is specifically asked for (see the above, [McCoy 1982], and [McKeown 1982]). Neither of these techniques will let us get a system to produce all of paragraph (b) when asked to describe Oscar but produce only the first sentence of paragraph (a) when asked to describe Felix. However, the following is a simple rule for determining salience. If an object or event x belongs to kind K, and members of kind K typically have property P but this one does not, that is interesting and should be reported. A representation and underlying logic which distinguish "Presumably A" from "A" and allow deducing "Presumably A, but not A" should extend to support this rule for determining salience (depending how easy it is to generalize over things having properties). But the logic of such a system will be monotonic.

4.3 Performing inferences versus adopting premises

There are two separate motivations for trying to formulate alternatives to monotonic logic. First, we may want to warrant conjectures in situations of incomplete information. Second, we may want to model belief spaces — as opposed to knowledge spaces — in which some believed propositions later become disbelieved.

Notice that these are different in several ways. In the first case, we may or may not be trying to model a system of truths (as opposed to beliefs; that is, a system which tries, successfully or otherwise, to model some domain other than a human belief system), but in either case, the conjectures which we want to add are known to be conjectures, which may or may not be in error. If, as I have proposed, they are actually guarded assertions, then they are different assertions from their unguarded versions. Hence unlike the second case, in which a single proposition is at one point believed and at another disbelieved, in the first case, if the conjecture is in fact overruled, the unguarded form of the denial is in the end asserted alongside the guarded claim. The result is something like, there is reason to believe A, but in fact not A. Notice that this contains no contradiction.

Systems which try to simulate human belief may be very different from systems which try to capture known truths about some domain. For instance, people can and frequently do hold inconsistent beliefs, even after the inconsistency has been pointed out. (This is one form of irrational behavior, which people manifest with deplorable frequency.) So inconsistency may be tolerable in a system which tries to model human behavior. But it is surely intolerable in a system which tries to capture truths about some domain, since it infallibly indicates that at least one domain assumption embodies an error.

For a proposition to be believed, to be true, and to be justified are three different things (albeit they are generally held to cooccur in the case of knowledge). In developing a logic, it is important to keep straight what it is that our semantics is supposed to model. This is normally taken to be truth conditions, not belief states, and not the presence, absence, or degree of justification. Should a logic attempt to use its semantics for some other purpose, its formulators will have to rethink the entire purpose of the logic, and the burden will be on them to show what the logic does, since it can no longer characterize entailments or provide a mechanism for deducing valid arguments.

The point here is that it is one thing to build in error recovery, and quite another to claim that no error was made. Standard logic is monotonic, because additional information cannot disrupt entailments. It can show that something which before was reasonable to presume is in fact false; but that is quite different from saying that the original premise set entailed the conclusion while the expanded premise set does not. If nonmonotonic logics are trying to model truth, then to say that truth changes with evidence simply represents a confusion. It amounts to trying to recover from a mistake without admitting that it ever took place.

4.4 Incomplete versus uncertain knowledge

Authors frequently use the expressions "reasoning in contexts of incomplete information" and "reasoning in uncertainty" virtually interchangeably. When we design systems for reasoning in these kinds of contexts, it is important to notice that the expressions mean different things. Not all uncertain knowledge is incomplete; and while incomplete knowledge may generally lead to uncertainty, it is not the same thing as uncertainty.

To say that a situation provides incomplete information is to say that information is missing. The usual "Does Tweety fly?" problem is one of incomplete information. There is a specific piece of information about Tweety that is missing; if that piece of information were supplied, the uncertainty in the situation would vanish. The problem is not that our context is uncertain; it's just that there's something we don't know.

But the "Does Tweety fly?" case can be looked at from another angle, from which genuine considerations of uncertainty arise. This is the case when we move from asking what we know about Tweety to asking what we know about birds. It has long been argued that knowledge often resides not with individuals, but with the classes to which they belong. But when we talk about knowledge about classes. we very early encounter things like "Birds fly" — that is, much knowledge about classes takes the form of TBGs.

The status of TBGs is not really a question either of incomplete knowledge or of uncertainty in the most obvious everyday sense. Even quite young children have complete enough information to know that some birds don't fly; on the other hand, not only do we readily agree to the TBG that birds fly, we are quite certain of it — as a TBG. The uncertainty lies not in the generalization itself, but in the way in which we understand it as applying to individual cases.

TBGs (as opposed to GaAs) can themselves result from either of two cases. First, we may have a situation in which we know all conditions which would cause a TBG to fail, although we may not know for a specific case whether any of those conditions hold. This is like analyses of "Birds fly" which hold that we know what could keep a bird from flying (belonging to species like penguins that don't; being too young; being injured; having clipped wings; being tethered to something too heavy to lift; etc.), although we may not know of a given bird whether its wings are clipped, or so on. On the other hand, we may have a situation like the one described in section 3.5.4 regarding the definition of "game", the classic example of a cluster concept.

If we are to represent knowledge about individual games at the level of "game" at all, such knowledge must reside in TBGs. This is true, not because we don't know all the facts about individual games that would let us tell whether they are exceptional, but because the fundamental principles themselves necessarily apply to games only uncertainly. To put this another way: when we run into a flightless bird, it is reasonable to ask for an explanation why it doesn't fly. But when we run into a game which is untypical in some regard, say for instance in having only one player, it is not necessarily reasonable to ask for an explanation. It simply is atypical in that way. For cluster concepts, the assumption of a causal connection between TBGs and their success or failure when applied to individuals becomes far more tenuous, and may vanish entirely.

Arguably the most extreme cases of uncertainty are those that are inherently statistical. If modern physics is anything like on the right track, no amount of information we could ever get would let us predict exactly when a radioactive atom will emit an alpha particle. We can give probability estimates describing the overall behavior of lumps of radioactive matter in the process of decay, but we will never be able to make more than a probabilistic estimate of the behavior of any particular atom, because the only laws governing the atom's behavior — known or otherwise — are probabilistic. Similarly, it will never be possible to know the precise location and velocity of any elementary particle. Not because we don't know how to measure that precisely, or because we lack a piece of information which could be filled in, but because the problem is inherently uncertain.

In less extreme cases, while the phenomena involved may not be inherently uncertain, our models may be. Any model that is essentially statistical has uncertainty built into its deepest levels, and that uncertainty results not from incomplete information, but from the *nature* of the information, *even when it is complete*. Incompleteness is, relatively speaking, a simple matter. All systems have always dealt with incomplete information (either by confessing ignorance or by ignoring it). Standard logic has no problem with incompleteness. It is excellent at detecting at. It just doesn't tell us what to do with it. By contrast, there is no place in standard logic for uncertainty: there is no way to say, "We rather think that this might be so, but we aren't sure."

There are many kinds of uncertainty, many of which behave very differently from one another. Confusing different kinds of uncertainty can already lead to problems; adding in a confusion between uncertainty and incompleteness raises the noise level to a point where it is unclear that a useful signal could get through.

4.5 Uncertain knowledge versus inherently uncertain propositions

The foregoing discussion made use of a distinction which we may also be interested in, which has to do with the locus or origin of uncertainty. On the one hand, *our knowledge* may be uncertain, because (for instance) we do not have a deep enough understanding of the principles of a domain to state its laws as true UGs. On the other hand, a domain's laws, correctly and deeply understood, may themselves be uncertain, for instance by being essentially statistical. Modern physics claims this to be the case relative to phenomena at the quantum level.

If a system has knowledge involving both these kinds of uncertainty, but views them in the same way, its explanations will tend to be confused at best. Notice that explanations are needed most when the user is least confident of what the system is saying. One way to trigger low confidence in a user is for the system to express uncertainty. Hence when the system reports uncertain results as uncertain, we may expect a particularly great need for clear and precise explanation. Sources of uncertainty will form an important part of this explanation. One thing users may want to do is look (outside the system) for information which, when added, will eliminate the uncertainty. If the uncertainty is inherent and irreducible, it would certainly save time and effort if the system could explain that.

4.6 Errors versus atypicality

There is a difference between discovering that something has an atypical feature and discovering that you have been wrong. Much of the discussion about reasoning in uncertainty fails to distinguish clearly between the need to deal with unusual cases and the need to repair mistakes which arise from going beyond the state of the system's knowledge. It should be noted that standard logic can always deal with error recovery, by identifying the culprit assumption and noting that it fails. Recording that ostriches are unusual birds because they fly has nothing to do with error recovery. There is an essential difference between noting exceptions to TBGs on the one hand, and repairing bad conclusions on the other. Failure to make the distinction can once again lead to very confused explanations.

Section 5: ADDITIONAL REQUIREMENTS FOR LOGICS OF INCOMPLETENESS AND UNCERTAINTY

The distinctions discussed in the previous two sections do not exhaust the issues surrounding logics for reasoning in contexts of uncertainty or incomplete information. On the contrary, there are many more concepts a logic may to capture than just those discussed here. Each kind of reasoning pattern which we attempt to reproduce will place its own requirements on the fundamental representation underlying any proposed inference mechanism. This section goes through a few established classes of such logics and points out some of the requirements in each case.

5.1 For logics of evidence

One interesting potential class of logics of incomplete information is the class of logics of evidence. Most work in this domain to date has been more suggestive than anything else. The fundamental idea is that of a non-numerical scheme for reasoning about information which is both incomplete and uncertain (not in the inherent sense, but in the sense that it may contain both errors and inconsistencies). Cohen's early work falls into this class [Cohen and Grinberg 1983; Cohen 1985]. While there is relatively little substance in this area as yet, it is both attractive enough and suggestive enough to raise several specific issues.

5.1.1 Weights versus Cohen-style endorsements

It is common to speak of weighing evidence. In AI applications, concepts like "weigh" tend to be interpreted as numerical measures to mind. In his early papers (see e.g. [Cohen and Grinberg 1983]), Cohen argued strongly that representations of evidence can not be reduced to simple numerical measures. He gave essentially three reasons for this. First, using simple numerical measures as degrees of belief does not allow different sources of evidence to be treated differently. Second, a single source of information might be more reliable in some contexts than in others. If the system only has a number, it cannot know which numbers to change which it switches contexts, let alone how to change them. Third, Cohen argues that certainty is best judged relative to tasks. Numerical measures will not relativize in any obvious way.

Thus reasoning with endorsements places several constraints. The first and most obvious, pointed out clearly in [Cohen and Grinberg 1983], is the need to represent richer evidentiary information than numerical measures of degree-of-belief (Cohen calls this information *endorsements*). These endorsements must be able to capture everything we would want to know about the sources for "given" information in our knowledge base.

Second, there must be ways of ranking endorsements. In the simplest case, this means that the system must have enough information to let it determine the relative strength of its various sources of information. This permits ranking the information given explicitly in the knowledge base.

Third, there must be some representation for the source of derived material which lets the system also rank beliefs which arise from inferences. This means essentially that the system needs two things. It needs to be able to represent that a given fact was derived on the basis of information from several sources, all of which in part support this particular item. Second, it needs to be able to rank the belief level of the item in question relative to other items in the system (this corresponds to combination rules for propagating simple numerical measures in measure based systems). This combination is likely to be more complex for Cohen than for a normal measure-based system, for two reasons. The first is that if the weight assigned to any single source can change, then this ranking will change with it, in rule-governed ways. In some sense, we want a formula rather than an answer. Second, it is unclear how to reflect issues like how heavily a given piece of information relies on one source rather than another.

Fourth, if we are to reason about endorsements as Cohen proposes, the endorsements themselves must be objects of knowledge, and we must be able to represent about them the kind of information on the basis of which we might establish rankings among different endorsements or alter rankings based on context or task. To date, little detail on these issues has been forthcoming, perhaps because of their innate complexity.

5.2 For logics of likelihood

Logics based on some concept somehow resembling probability (likelihood, certainty, etc.) generally fall into two categories. First, there are those which draw their concepts at least primarily from the field of statistics. Second, there are those which work with various other kinds of measures, which are not claimed to be the same as probabilities. We take these up separately.

5.2.1 Statistical approaches

These again break into three classes. First, there are relatively straightforward applications of statistical theory (either classical or Bayesian) to a given domain. Second, there are the

Dempster-Shafer [Shafer 1976] [Dempster 1968] based approaches. Third, there are attempts at statistical logics. Each of these has its own requirements.

I have discussed the several overall requirements for applying statistical analyses at some length elsewhere [Nutter 1987]. These include above all the knowledge to analyze the domain into an outcome space, define appropriate events, determine base distributions, determine event dependence and independence for at least the most important events, and of course come up with reliable numbers (especially for Bayesian approaches).

People find the Dempster-Shafer approach, as opposed to classical or Bayesian ones, attractive because it appears to palliate the last of these requirements. That is, the use of ranges instead of point probabilities appears to reflect not only uncertainty but so-called second order uncertainty, that is, our degree of confidence (or lack of it) in our probability estimates. It should be noted that this appearance only reflects some reality to the extent that we have good reason to believe that our estimates of our confidence in our guesses are any better than our guesses themselves. The evidence on the subject (see e.g. [Kahneman et al. 1982]) does not bear out any such optimism. It is also not clear that any of the usual propagation rules reflect a sensible interpretation of what happens to second order uncertainty as we reason with uncertain probabilities. At the very least, the propagation rules must insure that in the degenerate case (both limits the same), we actually compute probabilities, and that insofar as we can project ranges onto single-valued probabilities at all, the projection of the result of combining several ranges is the same as the result of combining their projections in normal statistics: that or admit that the system is not dealing with probabilities, and provide principled explanation of what it is up to.

Statistical logics like Kyburg's ([Kyburg 1987], [Kyburg 1988a], [Kyburg 1988b]) and Pearl's work on dependency graphs ([Pearl 1987], [Pearl and Verma 1987] are especially interesting because they attempt to combine logical with statistical reasoning in ways that respect the technical work in both fields. That is, their probabilities really act like probabilities, their (logical) inferences really act like (logical) inferences, and they are trying to approach the difficult problem of how to combine the two, given the difficulties outlined in section three above. The main requirement here is that the systems actually deliver on the promise of keeping the statistical aspects really statistical. Because of the complexity of the work, further details lie beyond the scope of this article.

5.2.2 Certainty factors, etc.

On the one hand, many problems plague traditionally statistical approaches, ranging from the difficulty of analyzing domains to the intractability of determining event dependencies to the difficulty of combining statistical with logic-like reasoning to the classic problem of getting the numbers. On the other hand, some find probabilities unsatisfying even supposing they could get them, for example because if A is evidence *against* B, it seems odd that we should express this by assigning a positive conditional probability to B given A. On the other hand, these same researchers nonetheless want numerical measures to use in ranking their system's beliefs. This constellation of issues has given rise to measures which are numerical, and which are said to reflect in one sense or another the plausibility of (or warrantedness of, or level of commitment to, or so on) propositions without modeling probabilities in the sense of statistics.

The best known such measures are probably the certainty factors used in MYCIN [Shortliffe and Buchanan 1975]. Rich [Rich 1983] has proposed a similar but less well articulated measure which she calls likelihoods. The certainty factor approach has numerous well-known difficulties. From the point of view of knowledge representation issues, two of these stand out sharply. First, while Shortliffe and Buchanan provide a set of formulas for manipulating their certainties, and in that sense have a well defined theory, the *interpretation* of that theory is far less evident. The origin of their numbers lies in the responses to questions like "On a scale of -10 to 10, how sure are you of that?" The elicitation process is more sophisticated than is shown here, but ultimately the initial certainty factors are people's estimates of their personal sense of conviction.

There are data on how people's senses of conviction behave (see e.g. [Kahneman et al. 1982]), and Shortliffe and Buchanan did not use them in formulating their theory. This may be just as well, since the data show (among other things) that the behavior is not consistent. Unfortunately, this raises two specters. The first is that the initial data may be meaningless. The second is the question, given that Buchanan and Shortliffe's formulas jointly compute some function, of exactly what that function models, of whether the initial values are reasonable inputs, and of whether given reasonable inputs, we should place credence or reliance on its outputs. That is, without some non-mathematical theory that links the mathematics of certainty factors (or any other nonstatistical numerical measure) with some actual property of propositions *that we care about and that should influence our willingness to accept them*, certainty factors are magic numbers. That is the first problem.

The second problem is one of representational inadequacy. On the certainty factor approach, positive numbers are intended to represent a measure of how much the system believes a proposition, while negative numbers are intended to represent a measure of disbelief (or belief in the negation). This would appear to leave zero to represent agnosticism. Unfortunately, agnosticism can take either of two forms. First, the system may be agnostic because is has no evidence on the subject at all. Second, it may be agnostic because it has conflicting evidence that balances out.

There are important differences between ignorance and conflict, especially in terms of deciding whether to resolve the situation or to let the sleeping dog lie. In addition, if the system has to resolve the question, what it should do may well differ depending on whether it is just ignorant or whether it already has mixed results. Also, if the system has to explain why it did not reach a given hypothesis, surely its explanation should be different if it had no relevant evidence from the explanation it would give if it had mixed evidence. For that matter, one would hope — especially in a system making diagnoses! — that the system would treat mixed but mild evidence differently from the way it would treat very strong indicators which are in conflict. Ignorance is not necessarily a motivation to further investigation. Given an arbitrarily chosen patient, a doctor may have no evidence as to whether that patient has, say, lung cancer. With no reason to suspect it (and assuming that the patient is having routine chest X-rays on the normal schedule), there is also no reason to go looking for it. But if there is strong evidence in favor, then even if there is also evidence against which cumulatively balances it, it would be well to make sure!

In computer science in general, and in AI in particular, there is a strong and justifiable tendency to be impressed with working programs. If a program computes a well-defined function from propositions into numerical measures, it is tempting to think it is doing something useful. In this regard, we should bear two points always in mind. First, suppose we take a program that computes a well-defined function, and change it as follows. If the leading digit of the original result was even, we add the number of flower species on the endangered species list divided by the number of buildings over ten stories tall in Manhattan. If it is odd, we subtract the same amount. The result is also a well-defined function, and so what? We need more than an internally consistent mathematics to guarantee meaningful results. Second, numerical measures summarize evidence. Summaries lose information. It is important to know what information we are losing, and to be sure that we can recover anything important, and more subtly, that we can tell when we need to initiate recovery.

5.3 For logics of vagueness

There are essentially two established approaches to vagueness, both due to the work of Lotfi Zadeh. Stemming from his early work [Zadeh 1965; Zadeh 1968] in the theory of fuzzy sets, one approach models properties in its domain as fuzzy sets, implements fuzzy set operations to model unions, intersections, relative complements, and the like, and hence works in a more or less immediate model of a vague domain. The second approach uses fuzzy concepts to formulate a semantics for a first order fuzzy logic, then uses fuzzy logic propositions and inference rules to reason in and about a domain. We take these up separately.

5.3.1 Fuzzy sets and their application

To apply fuzzy sets to a domain directly, the primary requirements are that the domain be finite and that measure functions be known for all predicates on the domain (i.e. for those sets which will be intersected, unioned, complemented, and so on to produce the results). The domain must be finite, because in this model there is no analog to quantifier reasoning. Measure functions are given extensionally. That is, for any "basic" set, the system needs a list of pairs, each consisting of an object in the domain and a value between 0 and 1, such that every object occurs in exactly one pair. The only effective way to give extensions is by listing them, and the lists must therefore be finite. They must also be complete, or the combining functions will not be well defined.

The obvious problem is getting reasonable data for the measure functions. This is not as severe a problem as it is for probabilities, since the granularity is not so fine, nor the system so sensitive. On the other hand, the results are no more sensitive, and reflect no higher granularity, than the system can represent. This means that this technique can not produce as many distinctions (or as reliable ones) as statistical techniques can when they can be applied. It should also be noted that the distinctions it does produce are not differences in event probabilities. They are levels of applicability of concepts. While fuzzy set theory is a reasonable model of vagueness, it is not only an unreasonable model of probability but also a mathematically inaccurate one.

5.3.2 Fuzzy logic

The fuzzy logic approach ([Mamdani and Gaines 1981]) has different and more subtle problems. Fuzzy logic admits of a continuum of truth values in the closed range [0,1], ultimately derived from measure functions on sets which serve as the extensions of predicates, but not in simple or obvious ways from fuzzy set theoretic combinations. Several questions immediately arise. First, every "assertion" in the data base must have an associated fuzzy truth value: where are we to get these from? For atomic sentences, this reduces to having measure functions; but what about quantified ones? For that matter, on this view, what is the natural interpretation of material implication?

Second, how are the truth values of propositions related to those of their components, and how are the truth values of conclusions related to those of the premises of the demonstration in question? So far as I know, no progress has been made on the "complex proposition vs. component" question. There are some preliminary results with regard to premises and conclusions (see [Aronson et al. 1980]), which boil down to the unsurprising claim that the conclusions are no better than the premises, but also on the whole no worse (where "better" is interpreted as numerical "greater than"). It is significant that this is already non-trivial to establish. Third, how do we deal with the apparent result that different demonstrations of the same proposition "establish" different truth values? For more on resolution and fuzzy logic, see [Lee 1972].

But the largest problem, in my opinion, lies in the irresistible temptation to view these fuzzy truth values as probabilities. Fuzzy logic is not a logic of probabilities. The mathematics of fuzzy truth values does not match that of probabilities. But researchers who advocate this approach do associate fuzzy truth values with probabilities. The tendency is encouraged by the need to assign what, in context, look much like Bayesian prior probabilities to the propositions in the data base. Zadeh has attempted to make fuzzy logic the basis of what he calls a theory of possibilities [Zadeh 1978; Zadeh 1981; Zadeh 1985]. While "possibilities" is not the same as "probabilities", the inherent temptation to confusion is powerful, and it may also be deliberate. As a nonstatistical numerical approach, Zadeh's possibilities in fact shares all the "magic numbers" problems that plague confidence factors, likelihoods, and the rest.

5.5 For logics of typicality

One of the more delicate tasks in assessing approaches to reasoning in uncertainty is distinguishing logics of abbreviation with error recovery from logics of typicality. Many of the schemes for reasoning in uncertainty seem to be representing GaAs, reasoning from them, and then allowing "fixes" in case the system in fact does not have the abbreviation right. The difficulty is that these seem to be *motivated* by TBGs, and further that the elaborateness of their "fixes" seems to go beyond what would be needed in a GaA-based system.

There is a large class of theories that seem, at least *prima facie*, to have aimed at TBGs, but to have fallen between TBGs and GaAs, perhaps because they failed adequately to distinguish the two targets. Theories falling into this general category, at least for me, include Reiter's default logic [Reiter 1980], the nonmonotonic logic of McDermott and Doyle [McDermott and Doyle 1980; [McDermott 1982], and most work on inheritance hierarchies with exceptions (see e.g. [Etherington and Reiter 1983], [Etherington 1987] or [Horty et al. 1987]). I am even more hesitant in categorizing work on circumscription ([McCarthy 1980], [McCarthy 1986]), which at present looks more like this to me than like striking either target squarely, but which may in fact be aiming at another altogether, which I simply haven't seen. In any case, none of these theories explicitly preserve the distinctions discussed in section four (particularly inferences versus presumptions and performing inferences versus adopting premises) or adequately separate decision theory from logic (see section two), and so suffer from the shortcomings discussed earlier. For more complete remarks on the shortcomings of several of these as theories of TBGs, see [Nutter 1982], [Nutter 1983a], and [Nutter 1983b].

I have published elsewhere [Nutter 1983b; Nutter 1983c] a proposed monotonic logic for reasoning with TBGs, which distinguishes explicitly from propositions ("Tweety flies") and related warranted presumptions ("Presumably Tweety flies"), and which distinguishes between the logic of defaults and the decision theory which their use entails. This approach has the potential to meet all the requirements discussed here, but at present it has two grave shortcomings. First, while the semantics is sufficient to prove soundness, it is unappealing in several fundamental ways (for details, see the publications). A new semantic founding, based on a semantics for relevance logic, would be a major improvement. Work in this direction is ongoing, but too preliminary to report any results at present. Second, until some of the relevant decision theory has been worked out, it will not be clear that the architecture will actually support it. I continue to prefer this approach to the others I have seen, but that may be personal prejudice.

Section 6: Conclusions

The single strongest message of this paper is that reasoning in uncertainty and in contexts of incomplete information is not a monolithic phenomenon. There are lots of different kinds of uncertainty, lots of different kinds of incompleteness, and lots of different kinds of reasoning with each that would be desirable in our systems. There are three separate morals.

- Different kinds of uncertainty of incompleteness are associated with different reasoning patterns. Because many of these patterns of reasoning are desirable and useful, it follows that there is no such thing as "the correct approach" to reasoning in uncertainty. We should be cooperating, not fighting.

- Different kinds of reasoning place different requirements on the systems that model them. Before formulating a theory of reasoning in uncertainty or with incomplete information, it is critical to determine what kind of reasoning the system will try to model. Otherwise, we risk aiming at nothing and hitting what we aimed at.

- Ultimately, of course, we want systems that can do it all. That makes it particularly important not only to foster diversity, but to pay attention to the various diverse efforts with an eye toward what might allow unification; lest we find ourselves with a full set of working parts, no two of which can work together.

In this paper, I have tried to outline some of the different kinds of reasoning we will eventually want to model, and to sketch out a few of the ground rules associated with each. With time, it may prove as important to progress in this very difficult area that we have as many clear, specific, agreed ground rules for the various subareas as that we have potential solutions in any one.

REFERENCES

Aronson et al. 1980
Aronson, A.R., B.E. Jacobs, and J. Minker. A note on fuzzy deduction, *JACM* **21** (1980) 599-603.

Brachman 1985
Brachman, R. I lied about the trees, *AI Magazine* **6**, 1985, 80-93.

Cheeseman 1985
Cheeseman, P. In defense of probability, *Proc. IJCAI–85*, Los Angeles, California, August, 1985, pp. 1002–1009.

Cohen 1985
Cohen, P.R. *Heuristic Reasoning about Uncertainty: An Artificial Intelligence Approach*, Morgan Kaufmann Publishers (Palo Alto, California) 1985.

Cohen and Grinberg 1983
Cohen, P.R. and M.R. Grinberg. A framework for heuristic reasoning about uncertainty, *Proc. IJCAI–83*, Karlsruhe, Federal Republic of Germany, August, 1983, pp. 355–357.

Conklin and McDonald 1982
Conklin, E.J. and D.D. McDonald. Salience: the key to the selection problem in natural language generation, *Proc. 20th Ann. Meeting of the A.C.L.*, Toronto, June, 1982, 129-135.

Dempster 1968
Dempster, A.P. A generalization of Bayesian inference, *J. Royal Statistical Society* **30** (Series B), 1968, 205–247.

Etherington 1987
Etherington, D.W. More on inheritance hierarchies with exceptions: default theories and inferential distance, *Proc. AAAI-87*, Seattle, July, 1987, 352-357.

Etherington and Reiter 1983
Etherington, D.W. and R. Reiter. On inheritance hierarchies with exceptions, *Proc. AAAI-83*, Washington, August, 1983, 104-108.

Freund and Walpole 1980
Freund, J.E. and R.E. Walpole *Mathematical Statistics*, third edition, Prentice-Hall, Inc. (Englewood Cliffs, New Jersey) 1980.

Horty et al. 1987
Horty, J.F., R.H. Thomason, and D.S. Touretzky. A skeptical theory of inheritance in nonmonotonic semantic networks, *Proc. AAAI-87*, Seattle, July, 1987, 358-363.

Israel 1980
Israel, D.J. What's wrong with non-monotonic logic?, *Proc. First Ann. Natl. Conf. on A.I.* (AAAI-80), Stanford, California, 1980, 25-50.

Kahneman et al. 1982
Kahneman, D., P. Slovic and A. Tversky, editors. *Judgement under Uncertainty: Heuristics and Biases*. Cambridge University Press (New York) 1982.

Kyburg 1987
Kyburg, H.E. Objective probabilities, *Proc. IJCAI-87*, Milan, August, 1987, 902-904.

Kyburg 1988a
Kyburg, H.E. Probabilistic inference and probabilistic reasoning. Department of Computer Science Technical Report 248, University of Rochester, Rochester, NY 1988.

Kyburg 1988b
Kyburg, H.E. Probabilistic inference and non-monotonic inference. Department of Computer Science Technical Report 249, University of Rochester, Rochester, NY 1988.

Lee 1972
Lee, R.C.T. Fuzzy logic and the resolution principle, *JACM* **19**, 1972, 109-119.

Mamdani and Gaines 1981
Mamdani, E.H. and B.R. Gaines, *Fuzzy Reasoning and its Applications*, Academic Press, London, 1981.

McCarthy 1980
McCarthy, J. Circumscription -- a form of non-monotonic reasoning, *Artif. Intel.* **13**, 1980, 27-40.

McCarthy 1986
McCarthy, J. Applications of circumscription to formalizing common-sense knowledge, *Artif. Intel.* **28**, 1986, 89-116.

McCoy 1982
McCoy, K.F. Augmenting a database knowledge representation for natural language generation, *Proc. 20th Ann. Meeting of the A.C.L.*, Toronto, June, 1982, 121-128.

McDermott 1982
McDermott, D. Non-monotonic logic II, *JACM* **29**, 1982, 33-57.

McDermott and Doyle 1980
McDermott, D. and J. Doyle, Non-monotonic logic I, *Artif. Intel.* **13**, 1980, 41-72.

McKeown 1982
McKeown, K.R. *Text Generation: Using Discourse Strategies and Focus Constraints to Generate Natural Language Text*, Cambridge University Press (New York) 1985.

Nutter 1982
Nutter, J.T. Defaults revisited, or 'Tell me if you're guessing', *Proc. 4th Ann. Cog. Sci.*, Ann Arbor, June, 1982, pp. 67–69.

Nutter 1983a
Nutter, J.T. What else is wrong with non-monotonic logics? Representational and informational shortcomings, *Proc. 5th Ann. Conf. of the Cog. Sci. Soc.*, Rochester, New York, 1983.

Nutter 1983b
Nutter, J.T. Default reasoning in AI systems. M.S. Thesis, Department of Computer Science Technical Report 204, SUNY at Buffalo, Buffalo, NY, 1983.

Nutter 1983c
Nutter, J.T. Default reasoning using monotonic logic: a modest proposal, *Proc. AAAI-83*, Washington, August, 1983, 297-300.

Nutter 1987
Nutter, J.T. Probability and uncertainty, *Proc. IJCAI-87*, Milan, August, 1987, 373-379.

Pearl 1987
Pearl, J. Embracing causality in formal reasoning, *Proc. AAAI-87*, Seattle, July, 1987, 369-373.

Pearl and Verma 1987
Pearl, J. and T. Verma. The logic of represeting dependencies by directed graphs, *Proc. AAAI-87*, Seattle, July, 1987, 374-379.

Reiter 1978
Reiter, R. "On Closed World Data Bases", in H. Gallaire and J. Minker, eds, *Logic and Data Bases*, Plenum Press, New York, 1978, pp. 55-76.

Reiter 1980
Reiter, R. A logic for default reasoning, *Artif. Intel.* **13**, 1980, 81-132.

Rich 1983
Rich, E. Default reasoning as likelihood reasoning, *Proc. AAAI-83*, Washington, August, 1983, 348-351.

Rosch 1975
Rosch, E. Cognitive representations of semantic categories, *J. Exp. Psych.: General* **104**, 1975, 192–233.

Rosch and Mervis 1975
Rosch, E. and C. B. Mervis. Family resemblances: studies in the internal structure of categories, *Cog. Psych.* **7**, 1975, 573–605.

Rosch et al. 1976
Rosch, E., C.B. Mervis, W.D. Gray, D.B. Johnson and P. Boyes–Braem. Basic objects in naturalcategories, *Cog. Psych.* **8**, 1976, 382–439.

Shafer 1976
Shafer, G. *A Mathematical Theory of Evidence,* Princeton University Press (Princeton, New Jersey) 1976.

Shortliffe and Buchanan 1975
Shortliffe, E.H. and B.G. Buchanan. A model of inexact reasoning in medicine, *Math. Biosciences* **23**, 1975, 351–379.

Wittgenstein 1953
Wittgenstein, L. *Philosophical Investigations*, MacMillan (New York) 1953.

Zadeh 1965
L.A. Zadeh, Fuzzy sets, *Inf. Control* **8**, 1965, 338-353.

Zadeh 1968
Zadeh, L.A. Fuzzy algorithms, *Inf. Control* **12**, 1968, 92-102.

Zadeh 1978
Zadeh, L.A. Fuzzy sets as a basis for a theory of possibility, *Fuzzy Sets and Systems* **1**, 1978, 3-28.

Zadeh 1981
Zadeh, L.A. "Possibility Theory and Soft Data Analysis". In *Mathematical Frontiers of the Social and Policy Sciences*, L. Cobb and R.M. Thrall, eds., Westview Press (Boulder) 1981, 69-129.

Zadeh 1985
Zadeh, L. Syllogistic reasoning as a basis for combination of evidence in expert systems, *Proc. IJCAI-85*, Los Angeles, August, 1985, 417-419.