Numeric and Symbolic Reasoning About Uncertainty in Expert Systems

Paul R. Cohen
Department of Computer and Information Science
University of Massachusetts, Amherst
COINS Technical Report 85-25

EKSL Report 4

This research is funded by a grant from the National Science Foundation IST 8409623 and DARPA-RADC Contract F30602-85-C-0014.

1.1 Introduction

This paper is about *qualification* of reasoning due to ambiguity, lack of evidence, poor quality data, and the many other factors that are usually associated with uncertainty. A broad distinction is often made between unqualified, categorical, or definite reasoning; and qualified, or uncertain reasoning. We start this paper from the position that all reasoning is subject to qualifications, though they may seem insignificant. Indeed, we are concerned with how people come to view qualifications as insignificant — how people can act as if certain under uncertainty. This paper discusses ways to represent and reason about qualifications. Our goal is to provide these abilities for expert systems and other artificial intelligence (AI) programs, so that they can reason intelligently under uncertainty. **

In overview, we will discuss the sources of uncertainty in reasoning, and the responses to uncertainty that, in people, we call intelligent. Then we survey current AI approaches to uncertainty. We find that few of our criteria for intelligent reasoning under uncertainty are manifested by AI programs. We assess the reasons for these deficits, and discuss three AI programs that, collectively, lay the groundwork for a new technology for reasoning under uncertainty. In the course of the paper, we find that reasoning under uncertainty is closely related to two other persistent problems for expert systems. One is the problem of controlling the behavior of large, knowledge-based systems; the other is the issue of explanation. The control problem, as we see it, is how to select a course of action that is responsive to one's uncertainty. Should the program pursue one hypothesis at a time, or all together,

** Our use of the word "qualification" is not incongruent with that of McCarthy (1980). His "qualification problem" refers to the need to act despite the fact that the conditions for action can not be stated completely. McCarthy asks, as we do, how we decide that we know enough to act.

or postpone this decision and search, instead, for more discriminating evidence? Which of several evidence-gathering plans is best? To answer these questions properly, we need to know the qualifications on reasoning — the reasons for uncertainty. Since these qualifications are the impetus for control decisions, they are also the basis for explanations of reasoning.

Our emphasis is on representing knowledge about uncertainty to facilitate reasoning under uncertainty. This is a common perspective in AI, where representing knowledge adequately is understood to be a prerequisite for intelligent reasoning. Thus, the current reliance on inadequate numeric representations is puzzling. One explanation is that probability and uncertainty are so closely associated that the one is mistaken for the other. The situation is analogous to mistaking a reproduction of a painting for the painting itself. A reproduction allows us to make *some* inferences about the painting, maybe enough inferences to tempt us to say, incorrectly, that we know what the painting looks like. And as the distinction between the original and the reproduction fades, we loose sight of the fact that different *kinds* of reproductions support different kinds of inferences about the original: any representation supports some kinds of inferences at the expense of others. Probabilities support inferences about the *degree* of uncertainty at the expense of inferences about the *reasons* for uncertainty. This paper suggests reversing these priorities.

1.2 Sources of Uncertainty

Uncertainty is a state of mind that arises during reasoning. By asking what aspects of reasoning give rise to uncertainty, we focus on its causes and consequences, not on the mental phenomenon itself. The sources of uncertainty are many but they can be discussed under three headings (see Cohen and Gruber, 1984, for more detail). First, uncertainty is introduced by evidence that is errorful, irrelevant, insufficient, and so on. Second, reasoning about evidence depends on heuristic knowledge, which can sometimes lead to a wrong conclusion. Third, the organization of knowledge, and the methods by which it is accessed, can introduce uncertainty.

Uncertainty due to evidence is especially problematic for systems that rely on sensory input. These include vision, robotic, and speech understanding systems. Evidence is typically *noisy*, meaning that parts of the evidence have been deleted or are obscured. The transducers that make evidence available to the system that interprets it can also introduce uncertainty. Most transducers have a limited bandwidth — they reproduce only some evidence faithfully; for example, sound transducers limit the frequencies they pass. Even if evidence is not noisy, and is not degraded by its transducers, its *relevance* may be uncertain. Most of the sensory data available to humans and other organisms is filtered by attentional processes. AI programs require procedures to select evidence from masses of information; these procedures introduce uncertainty. Finally, relevant, noise-free evidence may still be *inadequate*. Many tasks are uncertain not because the quality of evidence is poor but because there isn't enough evidence to complete the task. Sometimes the needed evidence is too expensive, sometimes it just isn't available.

Uncertainty in evidence can be managed if one knows its source. For example, high spatial frequency noise is common in vision systems, and the common remedy is to run the noisy image through a bandpass filter that cuts off the high frequencies. This eliminates the noise but introduces another kind of uncertainty: sharp intensity gradients (edges) become blurred. The remedy here is often edge enhancement of some kind. If the source of uncertainty is known, it can be managed. This argues for explicit, informative knowledge about uncertainty; it argues against limiting our knowledge of uncertainty to our degree of belief.

Once an expert system acquires evidence, the next step is to interpret it. Expert systems, more than other kinds of AI programs, rely on heuristic knowledge to interpret evidence; knowledge acquired from experts who will not always vouch for its accuracy. Expertise is experiential and pragmatic, and is sometimes unsupported by theory. Most tools for building expert systems allow the expert to qualify his or her knowledge with a degree of belief; but these rarely express qualifications satisfactorily (e.g., Gadsden, 1984). One kind of qualification has to do with *exceptions*. Since expert heuristics are *compilations* of expert experience, some aspect of the experience will be left out. A heuristic will "work" most of the time; uncertainty is introduced because a situation could arise in which the heuristic won't work. Doyle (1983) has suggested making these exception cases explicit when they are known. Then heuristics could be used with certainty in the standard cases and with caution at other times. Again, we see that uncertainty can be managed if its source is known.

But one cannot know, ahead of time, *all* the situations in which an expert heuristic should *not* be used. This source of uncertainty is unavoidable, but not necessarily unmanageable. Heuristics have applicability conditions which, in rule-based expert systems are the clauses in the *condition-part* of an inference rule. If the condition-part is satisfied, the action-part is asserted. Yet most rule-based systems do *not* execute all applicable rules, but select among them according to a *control strategy*. If the control strategy exploits some uncertainty-reducing aspect of a domain, such as redundancy, then the rules selected for execution are more apt to be those that *should* be accepted. This technique for managing uncertainty is discussed further in the section on control approaches, below.

Uncertainty is introduced in evidence and the knowledge that interprets evidence, and also in the *strategies* that control the use of the knowledge. For example, the question often arises, how long should one wait, or how much effort should one expend, to find some evidence? Some strategies cut off the waiting or search for evidence, thus introducing the uncertainty that a little more time or effort would have provided it. Many strategies are based on assumptions about the organization or extent of our knowledge. For example, the *closed world assumption* supports the conclusion that a fact is false if an exhaustive search of a knowledge base fails to turn it up (Reiter, 1980). The idea of a closed world is that we know all relevant facts; this is usually false, so inferences based on the assumption are uncertain. A similar assumption underlies *lack-of-knowledge* inferences, described by Collins (1978). Asked, "Is the Mekong River very long?" I reason that if it were, I would know it, and since I don't, it isn't. Similar knowledge is used to assess subjective probabilities. One method, called *availability*, is used by humans to estimate probability based on the ease of calling something to mind. Concepts that are "available" in memory are judged relatively probable; unavailable concepts are judged improbable (Tversky and Kahneman, 1982). We overestimate the probability of publicized events, such as winning lotteries; and students, for example, underestimate the probability of dying of heart disease, since few instances come to mind. Availability introduces uncertainty about the accuracy of our assessments of probability.

These heuristic methods for controlling access to our knowledge, like other heuristics, introduce uncertainty. But, as we noted above, if the source of uncertainty is known, it can be managed. The source of uncertainty in the Mekong River example is the assumption, "If the Mekong was long, I would know it." If I mistrust the assumption, then I can consult an authority – a person for whom the assumption is true. The credibility of a lack-of-knowledge inference is directly proportional to the amount one knows about the topic. Once one knows the source of uncertainty – in this case an assumption – and the factors that affect credibility, then the uncertainty is manageable. Assumptions play a major role in managing uncertainty, since they are explicit records of uncertain "stepping stones" in lines of argument. The assumption above is *needed* to answer the question, "Is the Mekong a long river?" Doyle (1983b) has developed *reason maintenance* mechanisms for managing the uncertainty represented by assumptions. This work, and the endorsement-based methods described below, recognise the need for explicit knowledge about uncertainty. For Doyle, assumptions are explicit records of the deliberate introduction of uncertainty, and, as such, pinpoint the source of uncertainty and provide a basis for its management.

1.3 Desiderata for Intelligent Reasoning About Uncertainty

This section asks what behaviors we should require of expert systems that reason intelligently about uncertainty. The requirements are of two kinds: first, we discuss what an expert system ought to *do* about uncertainty, then we focus on the representation of knowledge required to reason as we desire. It is striking that contemporary expert systems do very little about uncertainty besides measuring it. Some expert systems assess degrees of belief for hypotheses, but they do not use these numbers except to rank hypotheses and for some rudimentary control decisions. What more should an expert system do? We focus on two behaviors: planning (or control) and explanation.

Intelligent behavior under uncertainty requires a plan for the management of the uncertainty. Here are some examples of plans:

1. Confronted with uncertainty about which of two diseases afflict a patient, try to rule out the most serious one. Specifically, order relatively inexpensive, noninvasive tests before more costly ones, and give the patient a therapeutic trial of medication for the more serious disease. See the patient again after the test results are known and after the therapeutic trial has an opportunity

to alleviate symptoms.

2. Since I am uncertain whether my weekday bus runs on the weekend, I decide to drive my car.

3. I am going to visit my parents, who say they have a birthday present for me. They won't tell me what it is, so just to be safe, I put the roof-rack on my car.

The first case is taken from a series of interviews with a physician on the problem of diagnosing chest pain. Two causes of pain, angina and esophageal spasm, can have identical manifestations, but one is more serious than the other. Thus, physicians will try to *rule out* angina first, and may prescribe therapy for angina on a trial basis. The angina/esophageal spasm differential is not usually resolved by ruling *in* esophageal spasm, since it is difficult to get direct, physical evidence of spasm. However, this plan is appropriate if less costly tests fail to resolve between the disease hypotheses. In contrast, one can sometimes quickly rule out angina by demonstrating that the pain is due to damage to the muscles of the chest. This "rule-out by ruling-in" plan may not be appropriate, however, if the patient is at risk for heart disease because of smoking, age, family history, and so on, since this patient may have *both* heart disease and some other cause of chest pain.

Thus, intelligent reasoning under uncertainty involves selecting a plan appropriate to the nature of the uncertainty. The "rule-out by ruling in" plan may be appropriate in some cases but not to the angina/esophageal spasm differential if the patient is at risk for heart disease and if less difficult tests have not yet been tried.

If one knows enough about the nature of one's uncertainty to intelligently select a plan, then this knowledge can be used to explain one's behavior:

- Why did you try to rule out angina before esophageal spasm?

- Because the consequences of my uncertainty about angina are more serious; and because it is difficult to find direct evidence for or against esophageal spasm; and because there is evidence that the patient is at risk for heart disease, so ruling in esophageal spasm would not rule out heart disease.

Many plans for managing uncertainty are much simpler than this one. The second example, above, is a case of sidestepping uncertainty. Instead of facing the uncertainty of whether a bus is running, the question is made irrelevant by deciding to drive a car. The third case is similar: it involves anticipating possible outcomes and preparing for the most extreme. When uncertain about the size of a birthday present, one prepares for the worst (best?) case by arranging transportation for the biggest possible object.

One characteristic of these examples is that the *probability* of the various uncertain outcomes is both insufficient to determine a response to the uncertainty, and furthermore, it is largely irrelevant. In the medical example, provided there is "enough" evidence for angina, the physician pursues the angina hypothesis not because it is more likely than esophageal spasm but because it is more dangerous. In the second case, if there is "not enough" evidence that the bus is running, the commuter decides to drive. The extent of the uncertainty in these cases, and the third case, is not the salient factor in deciding on a plan to manage the uncertainty.

Yet, the probability of outcomes plays a *small* role in these examples, and a greater role in other cases, such as this one:

An airplane has crashed in dense jungle. Searchers superimpose a grid on a map of the area and calculate, for each square in the grid, the probability that the plane crashed in that square. They search the high-probability areas first.

Here, the appropriate plan for managing uncertainty depends on knowing the likelihood of outcomes. Thus, in addition to planning and explanation, we need the ability to believe one proposition more than another. This, in turn, requires the ability to update degrees of belief in light of evidence.

In summary, the behaviors that make for intelligent reasoning about uncertainty are: the ability to plan a course of action appropriate to one's uncertainty, the ability to explain one's actions, and the ability to determine degrees of belief in alternatives given evidence. We now consider the conceptual tools required to build expert systems with these abilities.

An expert system requires a representation of knowledge about its uncertainty and methods for manipulating this knowledge to plan and explain actions, and to modify its belief in propositions. A good representation supports all the concepts one wishes to reason about, and all the methods one uses to reason about them. A good representation makes important distinctions explicit. One should not have to struggle to represent a situation — the representational techniques should make the "translation" between a situation and its representation easy. If these representational criteria are met, then we will be able to represent the knowledge required to achieve the three performance criteria outlined above. Table 1 summarizes the performance and representational criteria. We now survey current AI approaches to reasoning under uncertainty from the perspective of these criteria.

TABLE 1.

Performance Criteria

Planning:	Plan actions that are appropriate to uncertainty
Explanation:	Explain plans for managing uncertainty
Measurement:	Modify degree of belief in light of evidence

Representational criteria

Adequacy:	Support all interesting concepts and methods for reasoning about them
Explicitness:	Make important distinctions explicit
Ease-of-use:	Make the "translation" between situation and representation easy

1.4 AI Approaches to Uncertainty

Many techniques for reasoning under uncertainty have been adopted or invented for AI programs. We group them according to how they represent uncertainty, and, thus, by the extent to which uncertainty is actively managed.

Parallel Certainty Inferences. The *parallel certainty inference* approach divides reasoning under uncertainty into parallel streams: one is a stream of domain inferences; the other, a stream of calculations of the credibilities of the domain inferences (Cohen, 1983). This is shown in Figure 1. Along the top of the figure is a chain of domain inferences – if a person is on fixed income then he or she has low risk tolerance, and if a person has low risk tolerance, then he or she ought to buy

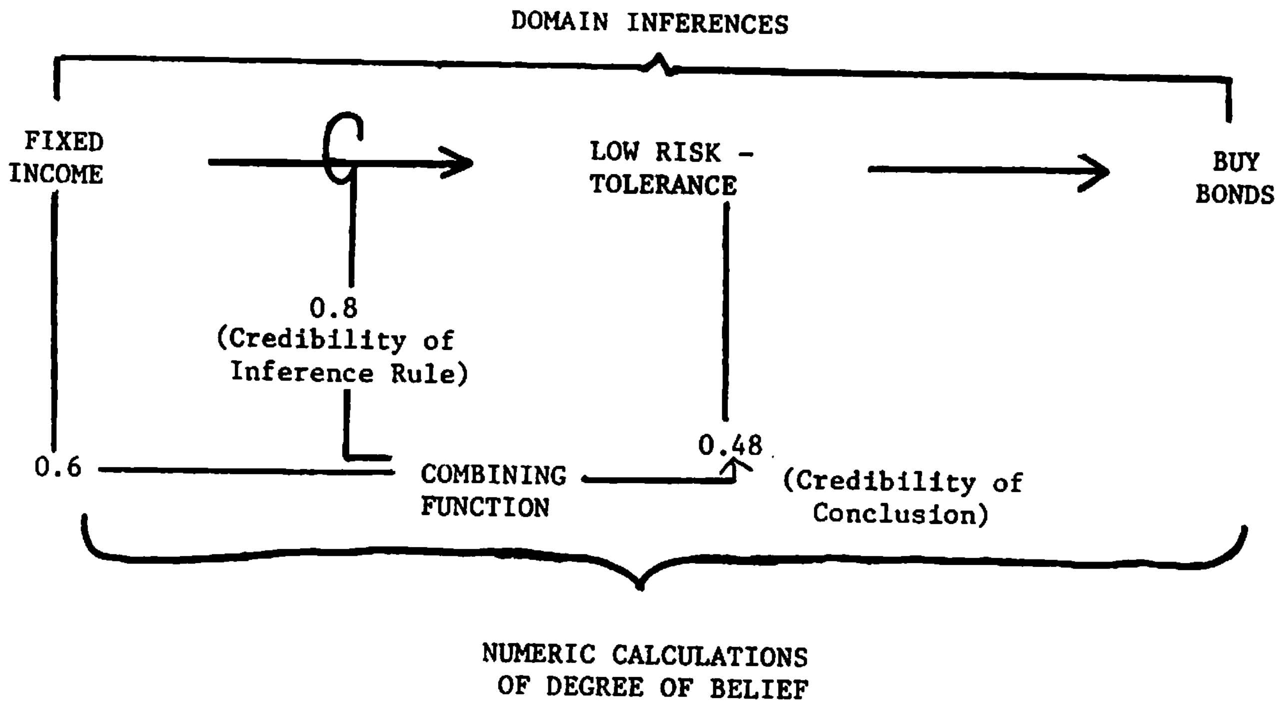

FIGURE 1: PARALLEL CERTAINTY INFERENCE APPROACH.

bonds. Along the bottom of the figure is a series of calculations of the credibilities of the data and conclusions. The first inference rule (fixed income implies low risk tolerance) is not entirely credible; its degree of belief is only 0.8. Moreover, the finding that this client is on fixed income is not entirely credible; its degree of belief is just 0.6. How credible is the conclusion of low risk tolerance, given the credibilities of the inference rule and the data? A *combining function* calculates, by multiplication, the degree of belief in the conclusion to be 0.48.

Among the parallel certainty inference methods we find strict probabilistic methods (e.g., Pearl, 1982), subjective probability techniques that are more or less distantly related to Bayesian updating (Shortliffe and Buchanan, 1975; Duda, Hart, and Nilsson, 1976), Dempster-Shafer calculi (Ginsberg, 1984; Strat, 1984; Gordon and Shortliffe, 1984; Lowrance and Garvey, 1982), and fuzzy logic (e.g., Zaheh, 1975) Though the proponents of the individual methods argue about their relative merits, for our purposes they may be grouped as the techniques that keep domain inferences and credibility calculations in separate compartments, using numbers to represent credibility.

The parallel certainty inference approach, though common, is unsatisfactory in terms of our performance and representation criteria. The good news is that degrees of belief are easily adjusted in light of evidence. But this advantage is not unqualified, since, in practice, we cannot assume that the numbers evoked from experts and propagated through chains of inferences are accurate. Nor can we guarantee that the combining functions used by the more subjective methods preserve the meanings of the numbers they combine. The bad news is that this approach denies to expert systems the ability to *plan* actions to manage their uncertainty; degrees of belief serve no purpose other than to find the highest-ranking conclusion, and sometimes to throw away conclusions with very low degrees of belief. The chains of logical and numerical inferences in Figure 1 are parallel in the sense that degree of belief has little or no effect on which domain inferences are made, when they are made, how they are corroborated, and so on. Since degrees of belief have no role in planning actions, they cannot be used to *explain* behavior. Furthermore, it is difficult to explain what a degree of belief means, since it is a poor represention of the complex mental processes that evoke it. Degrees of belief fail the *representational adequacy* criterion because they represent only the extent of one's belief, not the reasons for believing and disbelieving. They fail the criterion of making important distinctions *explicit*, because the degree of belief is a summary the many factors that contribute to uncertainty, including probability and utility. Finally, experts and others generally dislike the process of trying to quantify all aspects of their uncertainty, so degrees of belief fail the *ease-of-use* criterion.

Control Strategies. A second category of techniques, the *control* methods, manage uncertainty actively by *ordering* problem solving actions or sequences of actions. For example, I want to buy my wife a birthday present and a colorful box in which to wrap it. Since I haven't bought the present yet (and I'm not sure what I will buy), I am uncertain how big the box should be. The "obvious" solution is to buy the box after the present. So obvious, in fact, that it obscures an important conclusion: uncertainty is often due to the *timing* of evidence, and it can therefore be minimized by ordering one's actions so that the timing of evidence is most facilitative. This principle underlies *least-commitment planning* and related techniques (Sacerdoti, 1977; Stefik, 1980).

Other control approaches exploit characteristics of a domain to order problem-solving actions. For example, when building a jigsaw puzzle, it is best to start with "border" pieces, and then extend in from the border. This is because the border pieces are easily recognized, and once placed, constrain the placement of the other pieces. Redundancy is an important characteristic of some domains, and is exploited by control approaches to problems such as speech understanding. This problem is uncertain due to the noise and ambiguity inherent in the speech signal, but because speech is redundant, it is possible to work on the relatively certain parts of a speech signal first, then use them to constrain work on the uncertain parts. This approach was used in the HEARSAY-II speech understanding system (Erman, Hayes-Roth, Lesser, and Reddy, 1980).

Control approaches satisfy several of our performance and representation criteria. First, systems like HEARSAY-II actively plan which of several uncertain hypotheses to work on next. Unfortunately, they typically use numeric *evaluation functions* to decide where to direct their attention. The terms of evaluation functions represent, as numbers, factors relevant to managing uncertainty. In HEARSAY-II these include measures of the validity of data, the cost-effectiveness of actions, the desirability of understanding a particular segment of the speech signal, and so on. These numbers are combined into summary measures that control focus of attention. HEARSAY-II is thus able to select from among the many tasks it *might* do those which reduce its uncertainty about the speech signal. Its representation of knowledge about uncertainty supports productive reasoning methods (the *representational adequacy* criterion). But the factors that determine focus of attention are summarized in a single measure of worth, violating the *explicitness* criterion; and, since the requisite knowledge is not explicit, the *explanation* criterion. Roughly, the system works but it doesn't know why. HEARSAY-II is typical of systems that use control strategies to manage uncertainty.

Endorsement-based Reasoning. We turn now to four efforts to reason symbolically about uncertainty that, collectively, represent stages in the development of *endorsement-based* reasoning. Endorsements are explicit records of reasons to believe and disbelieve propositions. Endorsement-based reasoning satisfies most of the criteria in Table 1. Since it relies on reasons for uncertainty, it can plan and explain its plans to manage uncertainty. But since reasons for uncertainty are not *quantities*, precise reasoning about degrees of belief is awkward. This tradeoff is acceptable if one's emphasis is actively managing uncertainty instead of just measuring it.

Our first endorsement-based program, called SOLOMON for the wisdom we wished it had, attached mnemonic endorsements to propositions in place of numeric degrees of belief (Cohen and Grinberg, 1983; Cohen, 1984?). Each endorsement was used to select a course of action appropriate to the kind of uncertainty it represented. For example, we endorsed the rule

IF age > 65 THEN risk-tolerance = low

with the mnemonic **overgeneralisation**, meaning that, for some individual, the conclusion *could* be false when the premise is true. Now, one can imagine adding clauses to the premise to pinpoint more certainly the criteria for low risk tolerance. The same effect can be had by finding another rule with a different premise but the same conclusion. Thus, given a conclusion endorsed as an **overgeneralisation**, SOLOMON searched for a corroborating conclusion, that is, a rule with the same conclusion but a different premise. If this succeeded, SOLOMON endorsed the conclusion as **corroborated.**

The theme of the SOLOMON program is familiar: a system must respond appropriately to its uncertainty. Endorsements characterize uncertainty and are the key to intelligent responses. But this early work considered relatively few endorsements and responses to uncertainty. Nor did we modify the endorsements associated with propositions in the light of evidence. This became the focus of our next study.

Numeric approaches to uncertainty modify the degrees of belief in propositions as evidence becomes available. The "running total" belief for a hypothesis is increased or decreased in response to evidence pro or con. Endorsements do not represent degrees of belief, but rather, reasons for belief. We explored how these reasons are adjusted by evidence in the context of a plan recognition program called HMMM (Cohen, 1984; Sullivan and Cohen, 1985). Imagine a simple device that can execute one of two plans, each composed of 3 steps:

plan		steps
plan 1	:	a b c
plan 2	:	b d e

If the device takes step **a**, what plan does it have "in mind"? Since step **a** is unique to plan 1, the device either has made a mistake, or it intends plan 1. Assume the device now takes step **b**. This provides no evidence to discriminate the interpretations of the first step: the device may be pursuing plan 1 or it may have recovered from its mistake and started plan 2. If the next step is **c**, then it looks as though plan 1 was intended all along; if it is **d**, then apparently plan 1 was started and abandoned for plan 2. The question we want to answer is, if the endorsement of the plan 1 interpretation of **a** is **may be a mistake**, what happens to this endorsement as more evidence — subsequent plan steps — becomes available? Answering this question is analogous to finding a combining function for numeric representations of uncertainty.

We do not believe that all kinds of evidence should be combined with the same combining function. That is one of our complaints against numeric approaches. We devised several *combining schemas* for endorsements that captured the flux of our reasons for uncertainty in the plan recognition problem. We noted above that the evidence **b** cannot reduce our concern that **a** was a mistake, since **b** is ambiguous with respect to plan 1 and plan 2. On the other hand, the input **c** is unique to plan 1 and "finishes off" the plan, and seems to reduce the concern that **a** was a mistake. This kind of reasoning is captured in the following combining schema, in which the endorsements — the reasons to believe and disbelieve intepretations of plan steps — are shown in uppercase.

IF step I IS-UNIQUE-TO plan N, and
    step J IS-UNIQUE-TO plan N, and
    step J FOLLOWS-IN-THE-PLAN step I, and
    the plan N interpretation of step I
    is endorsed by MAY-BE-A-MISTAKE
THEN erase the endorsement

By eliminating the second clause of this schema, the negative endorsement on the plan 1 interpretation of **a** is erased as soon as the evidence **b** becomes available. This seems premature, as we said, since **b** is ambiguous, but we give the example to raise a point: Our goal in this work was not to provide a *prescriptive* theory of how endorsements should combine, but rather, to give a framework for subjectively combining endorsements.

Clearly, erasing endorsements is a degenerate form of combining them, and loses valuable information about the interrelationships between pieces of evidence. A more realistic scheme would reduce the weight of the **may be mistake** endorsement as subsequent, consistent evidence becomes available; alternatively, one endorsement might be made dependent on another, as in Doyle's (1983b) work on reason maintenance.

Before implementing such a scheme, however, we were diverted by a difficult question: Where do endorsements come from, and what do they mean? The mnemonic value of endorsements like **may be a mistake** disguises the fact that endorsements are arbitrary symbols, whose meaning comes from the rules by which they are combined with other endorsements. We were concerned that, for complex domains, dozens of endorsements and combining schemes would have to be acquired. Although we had no objection in principle to acquiring this knowledge from an expert (much as other domain knowledge is acquired), we wondered whether the endorsements and combining schemas of a domain could be derived from other knowledge about the domain, such as inference rules. If so, we would worry less about whether we had the "right" endorsements and combining schemas.

We focused on the uncertainty inherent in a single problem-solving task, namely classification, to pinpoint the sources of uncertainty (and thus endorsements) of all classification tasks. Classification is the problem solved by many or most expert systems (Clancey, 1984): Given data, find the conclusion (or classification of the data) that fits the data best. Uncertainty in classification tasks is due, primarily, to mismatch between evidence and its various classifications. Degree of belief in a conclusion given evidence reflects the degree of fit between the evidence and the classification. For example, if the flu is characterized by fatigue, nausea, and aching limbs, then one's certainty in a diagnosis of flu depends on the degree of fit between symptoms and this characterization. Does a midafternoon nap constitute evidence for fatigue? Is skipping lunch evidence for nausea? Is a headache evidence for aching limbs? To the extent that these findings correspond to the symptoms of flu, the diagnosis of flu is credible. The chief source of uncertainty in classification tasks is partial matching between the evidence one needs and the evidence one has. Endorsements ought to describe these partial matches, and ideally should be derived from knowledge about the classification task.

Consider the evidence for flu: is a midafternoon nap evidence of fatigue? Fatigue is good cause for a nap, quite possibly the *only* cause. In contrast, there are many reasons to skip lunch, of which nausea is only one. Given this, it seems reasonable to suggest that the midafternoon nap is stronger evidence of fatigue than the skipped lunch is of nausea. Finally, headache seems to be very weak evidence for aching limbs because the head and the limbs are different parts of the body. Figure 2 shows how evidence and conclusions are associated for each of these cases. Figure 2a shows a causal relationship between fatigue and taking a nap; Figure 2b shows that nausea is one of several phenomena associated causally with skipping lunch; Figure 2c shows the head and limbs as *siblings* in a *part-of* hierarchy.

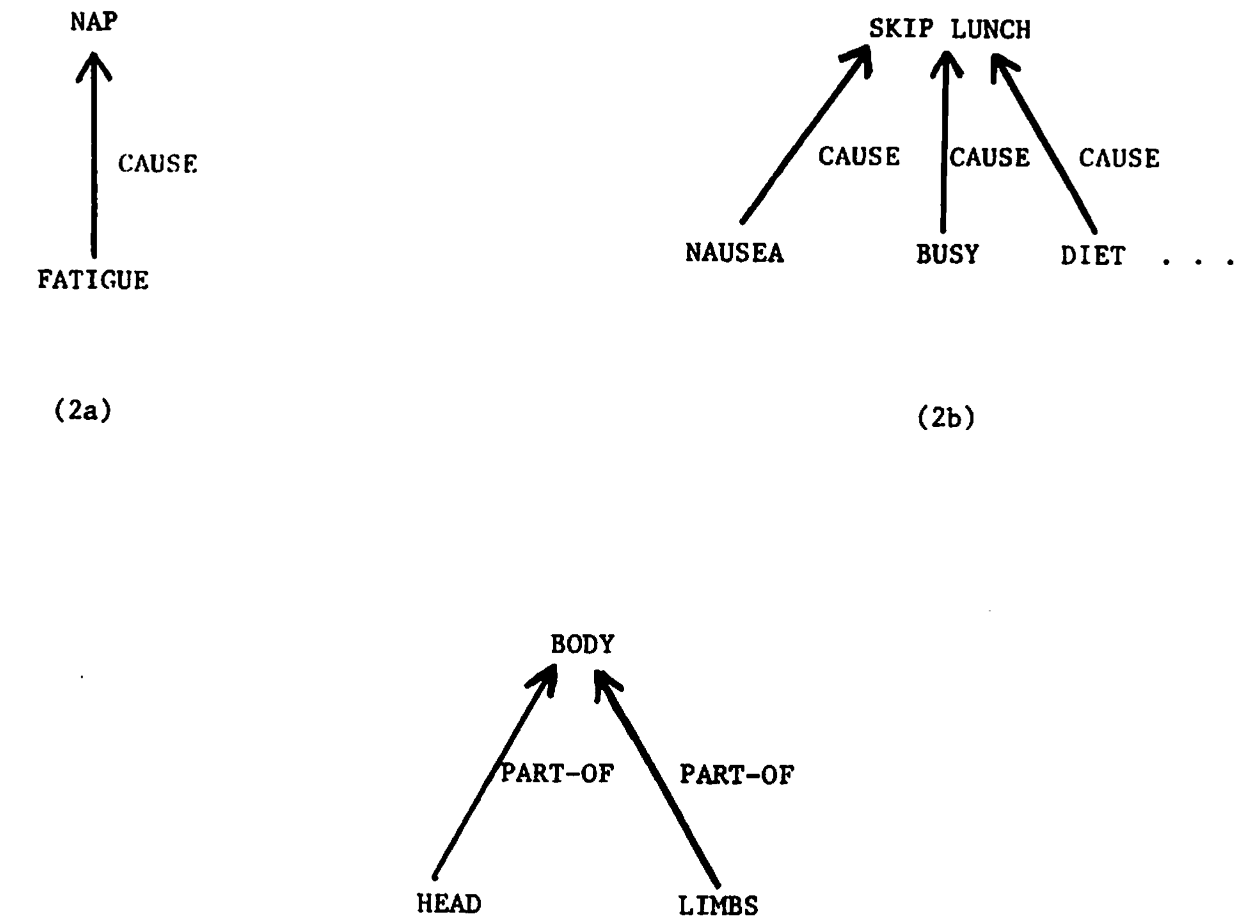

FIGURE 2: DERIVING "PATH ENDORSEMENTS" FROM ASSOCIATIONS BETWEEN EVIDENCE AND CONCLUSION.

Path endorsements reflect the associations between pieces of domain knowledge, such as those shown in Figure 2. The credibility of a conclusion such as "the patient has aching limbs" depends on the associations that form a path between the conclusion and the evidence. The *sibling* path between head and limbs is, in general, not the basis for credible inferences: something that is true of an object is not necessarily true of its siblings. A headache is not evidence that the rest of the body aches. On the other hand, the single causal association between fatigue and taking a nap is the basis for credible inferences; given that a person takes a nap, it is credible to infer fatigue. But it is not credible to infer nausea given that a person skips lunch, because of the many possible causes for skipping lunch. Path endorsements describe typical patterns of associations between evidence and conclusions. Inferences based on these associations are more or less credible, as we discussed, so path endorsements are the basis for judging the credibility of inferences. Note that path endorsements are derived from knowledge about how objects in a domain are associated. They are not "made up" by knowledge engineers to represent suspected sources of uncertainty.

We developed an expert system, called GRANT, based on path endorsements. Its task is to match researchers with funding agencies that are likely to support their work (Cohen, Davis, Day, Greenberg, Kjeldsen, Lander, and Loiselle, 1985). This is a classification problem in which the evidence is a research proposal, and the conclusions are the funding agencies that best fit the proposal. The chief source of uncertainty is partial matches between the interests and requirements of funding agencies and the interests and needs of researchers. To the extent that the match between an agency and a researcher is good, the agency is likely to support the researcher. Path endorsements are used to find matches between the respective research interests of the parties. For example, an agency interested in neurological diseases is unlikely to fund a researcher interested in osteopathic diseases, because the path between neurology and osteopathy includes the sibling relationship between the head and the limbs, shown in Figure 2c.

All endorsements in GRANT were derived *after* the knowledge for performance of the matching task was in place. The endorsements literally "come from" the associations that are needed to encode a large semantic network of research topics. The network contains over 2000 concepts that describe the research interests of about 250 funding agencies. The interests of researchers are described by the same concepts, and GRANT finds agencies to fund researchers by following well-endorsed associative pathways between concepts. So far, the path endorsements are discerning enough that less that 1/3rd of the agencies found by GRANT are judged, by our expert, unlikely to fund the researcher's proposal. Moreover, GRANT finds over 80% of the agencies judged acceptable by our expert.

Our fourth effort at reasoning with endorsements is currently in progress. We are developing an expert system for diagnosing the causes of chest pain. This problem was selected because it gives us an opportunity to study intelligent responses to uncertainty. Although we started to explore how endorsements could select responses in the SOLOMON project, the range of responses was small. The issue lay dormant in the HMMM and GRANT projects. But in diagnosing chest pain, a physician has access to a rich source of actions, and must select the appropriate ones based on his or her uncertainty. Many factors influence this choice. For example, the amount of time it takes to get evidence from tests must be weighed with the time course and seriousness of the disease, to decide whether to prescribe therapy or wait for evidence.

1.5 Conclusion

Our research in medical problem-solving is not at the stage that we have a running program, but the architecture of the system is guided by principles that, together, summarize the themes of this paper: Reasoning about uncertainty is knowledge-intensive, so one's representations of knowledge about uncertainty should be informative and explicit, not summary in nature. From these representations, plans to manage uncertainty can be formulated and explained. Uncertainty has many sources; intelligent management of uncertainty responds to them differently. But however responses are selected, uncertainty must be actively managed instead of passively measured.

References

Clancey, W.S., 1984 "Classification problem Solving," Proceedings of the AAAI, p. 49.

Cohen, P.R., 1983, "Heuristic reasoning about uncertainty: An artificial intelligence approach," Doctoral dissertation, Stanford University. STAN-CS-83-896, Also, Pitman Advanced Publishing Program. London, 1985.

Cohen, P.R. and Gruber, T., 1984 "Reasoning About Uncertainty: A Knowledge Representation Perspective," Pergamon State of the Art Report: Expert Systems.

Cohen, P.R., Davis, A., Day, D., Greenberg, M., Kjeldsen, R., Lander, S., and Loiselle, C., 1985, "Representativeness and Uncertainty in Classification Systems," *AI Magazine,* Fall, 1985.

Cohen, P.R., 1984 *A Progress on the theory of endorsements*, COINS-TR, University of Massachusetts, Amherst. Also, Proceedings of the IEEE Workshop on Principles of Knowledge-based Systems. Denver, Colorado. p. 139.

Collins, A.M., 1978, "Fragments of a theory of human plausible reasoning," *TINLAP-2,* 194-201.

Doyle, J., 1983a, "The ins and outs of reason maintenance," *Proceedings IJCAI-83*, August 1983, p. 349-351.

Doyle, J., 1983b, "Some theories of reasoned assumptions: Some essays in rational psychology," CMU-CS-83-125, Carnegie-Mellon University, 1983b.

Duda, R.O., Hart, P.E., and Nilsson, N., 1976, "Subjective Bayesian methods for rule-based inference systems," Technical Note 124, Artificial Intelligence Center, SRI International, Menlo Park, California.

Erman, L.D., Hayes-Roth, F., Lesser, V.R., and Reddy, D.R., "The HEARSAY-II speech understanding system: integrating knowledge to resolve uncertainty," *Computing Surveys,* 12, 1980, p. 213-253.

Gadsden, J.A., 1984, "An expert system for evaluating electronic warfare tasking plans for the royal navy," *Proceedings of the First Conference on Artificial Intelligence Applications.*, IEEE Computer Society Press.

Ginsberg, M.L., 1984, "Non-monotonic reasoning using Dempster's Rule," *Proceedings AAAI-84,* August, 1984.

Gordon, J., and Shortliffe, E.H., 1984, "The Dempster-Shafer Theory of Evidence," in Buchanan, B.G. and Shortliffe, E.H. (eds.) *Rule-Based Expert Systems,* pp.272-292.

McCarthy, J. 1980. Circumscription – A form of nonmonotonic reasoning. *Artificial Intelligence,* 13 (1,2) p. 27-39.

Pearl, J., 1982a, "Reverend Bayes on Inference Engines: a Distributed Hierarchical Approach," *Proceedings of the National Conference on Artificial Intelligence,* p. 133-136, Pittsburgh, PA.

Reiter, R., 1980, "A logic for default reasoning," *Artificial Intelligence,* 13, 1980, p. 81-132.

Sacerdoti, E.D, 1977, "A structure for plans and behavior," New York: American Elsevier.

Shortliffe, E.H. and Buchanan, B.G., 1975, "A Model of Inexact Reasoning in Medicine," *Mathematical Biosciences,*, 23 p. 351-379.

Stefik, M.J., 1980, "Planning with constraints" (Doctoral Dissertation), Department of Computer Science Report 80-784, Stanford University, 1980.

Strat, T.M., 1984, "Continuous Belief Functions for Evidential Reasoning",, *Proceedings of the National Conference on Artificial Intelligence,* Austin, Texas, August, 1984, p. 308-313.

Sullivan, M., and Cohen, P.R., 1985, "An endorsement-based plan recognition program," COINS T-R , University of Massachusetts, Amherst. Also *Proceedings of* *the International Joint Conference on Artificial Intelligence*. Forthcoming.

Tversky, A. and Kahneman, D., 1982, "Availability: A heuristic for judging frequency and probability." in Kahneman, Slovic, and Tversky (Eds.), *Judgment under uncertainty: Heuristics and Biases,* Cambridge.

Wesley, L.P. "Reasoning about control: the investigation of an evidential approach," *Proceedings IJCAI-83,* August 1983, p. 203-206.

Zadeh, L.A., 1975b, "Fuzzy logic and approximate reasoning," *Synthese,* Vol. 30, pp. 407-428.