Fifth Int'l Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems, July 1994.

Some Varieties of Qualitative Probability

Michael P. WELLMAN

University of Michigan
Dept of Electrical Engineering and Computer Science
Ann Arbor, MI 48109 USA
wellman@engin.umich.edu

ABSTRACT

In this essay I present a general characterization of qualitative probability, including a partial taxonomy of possible approaches. I discuss some of these in further depth, identify central issues, and suggest some general comparisons.

1. Introduction

In the standard theory of probability, degrees of belief for events or *propositions* take values in the real interval [0,1]. From degrees of belief on the primitive propositions, the theory dictates degrees of belief for various compound and conditional propositions, and vice versa. Computational schemes for *probabilistic reasoning* apply this theory to the automated derivation of degrees of belief for designated propositions of interest given prespecified degrees of belief over some other propositions and some particular conditioning propositions observed or hypothesized. This approach has, among other advantages, those accruing to a well understood and powerful underlying theory.

Despite these virtues, many have objected to the straightforward application of probability theory to uncertain reasoning because it appears to require undue precision in the specified degrees of belief. This and other supposed drawbacks has led researchers to invent innumerable variations on probability, as well as many other "non-probabilistic" uncertainty calculi. Among other distinctions, the variant approaches typically support the specification of degrees of belief that are less precise in some way or another. Motivations for these features include reduction of specification burden in representing uncertain knowledge, robustness of inference with respect to slight changes in degree of belief, and sometimes, the attempt to capture "ignorance" in a state of uncertain belief.

It is not my intent here to evaluate or survey the numerous mechanisms for incomplete or imprecise degree-of-belief specification in variant uncertainty calculi. Rather, I aim to identify and delineate some general classes of approaches that relax precision in representation and reasoning *within* the probabilistic framework. In doing so, I maintain a Bayesian perspective on probability, interpeting the imprecise specifications as *partial descriptions* of an uncertain belief state, not the belief state itself. (Nevertheless, most of the technical points should be ideologically neutral enough to avoid putting off non-Bayesian readers, if I still have any.) The various approaches have (legitimately) been termed "qualitative probability", though they are distinct enough to warrant an examination of their relative merits and limitations.

2. Qualitative Probability

Generally stated, *qualitative reasoning* is reasoning directly in terms of the qualities of interest for a particular problem or class of problems. The idea of qualitative reasoning is present (often implicitly) in many technical disciplines, where the aim of research is to identify exactly those qualities that are specific enough to characterize the phenomena of interest, yet general enough to comprise all situations where the phenomena apply. As a general methodology for representation and computation, qualitative reasoning has been investigated most explicitly and thoroughly within the field of Artificial Intelligence, especially by researchers interested in capturing commonsense or expert knowledge about the physical world (Weld and de Kleer 1989). These researchers have tended to focus on particular sorts of qualitative properties: signs of physical quantities and rates of change, monotone relations among parameters, and topological descriptions of space. But this emphasis can be attributed to the particular problems they have addressed, and we can consider qualitative reasoning more broadly as the approach to knowledge representation that shuns excess precision wherever possible in the effort to match the levels of description and objects of inference to those of the problem being addressed.

So how does the idea of qualitative reasoning apply to probabilistic or uncertain reasoning? Since few if any real problems in uncertain

reasoning (specifically, decisions under uncertainty) depend on the full precision of real-valued degrees of belief, the principle of qualitative reasoning suggests that we should not require that degrees of belief be specified to this full precision, nor that the reasoning method manipulate degrees of belief at this precision. The usual (Bayesian, at least) objection at this point is that we cannot generally tell in advance which level of precision will be required, so it will be irrational to accept anything but the fullest precision. But it really does not follow from the lack of an assured precision bound that we need apply unbounded precision *in all cases*. Rather the qualitative dictum just calls for our reasoner to be flexible in its use of precision, and says that in those cases where precision can be avoided, it should be.

What does it mean to be less than fully precise about probability? There are several possible interpretation approaches, but the most straightforward is to regard an imprecise statement involving degrees of belief as a statement that the degrees of belief satisfy some constraint. Thus, whereas a fully precise description of an uncertain belief state will assign a unique number to the probability of each proposition, an imprecise description will merely constrain the probability to belong to some set. Or more generally, it will constrain the set of *joint* probability assignments to a subset of the joint space of probabilities.

Now, there are many potential ways that we can relax precision in degrees of belief. For example, we could simply replace our scalar representation of degrees of belief with a pair, where the two numbers are interpreted as lower and upper bounds on degrees of belief. This is, of course, standard in several of the variant calculi for uncertain reasoning.[1] But it is only the most obvious of the possible relaxations of precision, and not obviously the most advantageous for purposes of representation and reasoning. In order to make this judgment, we first need be more explicit about our desiderata for representation language and inference calculus. These will depend on the task we have in mind for qualitative probabilistic reasoning, and (once again) on the qualitative reasoning principle that reasoning be directly in terms of the qualities with which we are concerned.

[1] One representative attempt to apply probability bounds in uncertain reasoning was Quinlan's INFERNO system (Quinlan 1983). There have been many others, before and since. Concepts of lower probability functions are commonplace in the statistical literature, and the notion of intervals of degrees of belief (not necessarily interpreted as probability) is prominent in both Dempster-Shafer theory and Possibilistic Logic.

3. Qualitative Probabilistic Reasoning

3.1. Semantics

Given our constraint-based interpretation of qualitative probability, we specify the reasoning process in the conventional way. Let L be a qualitative probability language, and Q a sentence of L. The meaning of Q is defined by the set of probability distributions that satisfy it. We write $\mu \models Q$ to say that the probability distribution μ satisfies the constraints Q. The set

$$\mathrm{M}(Q) = \{\mu \mid \mu \models Q\}$$

directly characterizes the extension of Q. From this we can define entailment in the qualitative probability language:

$$Q_1 \models Q_2 \text{ iff } \mathrm{M}(Q_1) \subseteq \mathrm{M}(Q_2).$$

Typically, systems of qualitative probability specify some inference regime, denoted $\vdash$, such that $Q_1 \vdash Q_2$ means that Q_2 is derivable from Q_1. This inference regime is *sound* if $Q_1 \vdash Q_2$ implies $Q_1 \models Q_2$, and *complete* if the converse holds. Most of the qualitative probability schemes we are concerned with are sound but not complete.

The expressive power of a qualitative probability language is defined by the sets of probability distributions that can be expressed. A set M is *expressible* in L if $\mathrm{M} = \mathrm{M}(Q)$ for some Q in L. A qualitative probability language is *closed under conjunction* if, for any two sentences Q_1 and Q_2, $\mathrm{M}(Q_1) \cap \mathrm{M}(Q_2)$ is expressible. Closure under disjunction and negation can be defined similarly. A perhaps more interesting property is closure under marginalization, which requires that the set $\mathrm{M}(Q)$ remain expressible after averaging over some proposition or random variable.

3.2. Tasks

Qualitative probabilistic reasoning is the process of deriving some qualitative properties of a probability distribution from other ones. Which qualitative properties are of interest

depends on the objectives of the reasoner. Two qualitative reasoning tasks have been featured in the literature: *decision making* and *acceptance*.

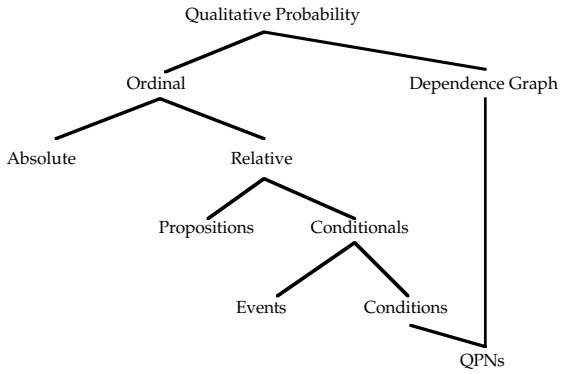
In decision making, the objective is to determine which action or strategy maximizes expected utility. This task requires that we have information about the available actions and preferences in addition to probability. The imprecision of qualitative probability means that we cannot in general determine a unique optimal decision, but the constraints may allow us to eliminate many of the possibilities based on qualitative dominance (Wellman 1990).

The aim of the acceptance task is to establish that the likelihood of particular propositions is sufficiently high that they should be accepted as beliefs. To accept a proposition is essentially to act as though it were certainly the case. Although a strict Bayesian would deny that it is ever rational to accept uncertain propositions,[2] some have advocated it on the grounds that it simplifies reasoning. Indeed, this is one of the central motivations for nonmonotonic logics. See Kyburg and respondents (Kyburg 1994) for a discussion of acceptance based on probabilistic thresholds. In any case, once an acceptance criterion is stated, the problem becomes a well-posed exercise in qualitative probability.

Irrespective of the overall task, another aim of qualitative probability has been to capture common inference patterns. That is, qualitative probabilistic reasoning schemes are often judged by whether they sanction particular inference schemas or how well they match the expected results on canonical examples. Examples include probabilistic theories of causality (Suppes 1970), and accounts of the phenomenon of "explaining away" in evidential reasoning (Pearl 1988; Wellman and Henrion 1993).

4. A Taxonomy of Qualitative Probability

Having described the general idea and formal framework, we are ready to classify some different forms of qualitative probability. The various distinctions considered in this paper are laid out in the taxonomy below.



The two main categories of qualitative probability concepts, ordinal and dependence graph, are described in the sections below. Ordinal comparisons constitute the most common form of qualitative probability, and can be distinguished along several dimensions. First, we can compare probabilities against an absolute scale, or relative to each other. Second, the comparisons can involve unconditional propositions or conditional expressions (this distinction makes sense for absolute as well as relative comparisons, but is shown in the taxonomy only for the latter). Finally, in comparing conditional expressions, we can vary the base event part of the expression, or the conditioning events (or both). The qualitative probability language I have developed, QPNs (Wellman 1990), consists of relative ordinal comparisons among conditional expressions varying only the conditions, embedded within a dependence graph formalism.

5. Ordinal Qualitative Probability

Several forms of qualitative probability are defined by introducing an ordering relation $\preccurlyeq$ over events signifying "not less likely". Its negation, $\succ$, denotes "less likely", and its symmetric restriction, $\sim$, "equally likely". In some foundational accounts of probability (Koopman 1940; Savage 1972), the numeric representation is derived from axioms on $\preccurlyeq$. However, even if we accept the axioms, we are not *obliged* to use the numeric representation, and can design schemes that reason in terms of $\preccurlyeq$ directly.

Specifying partial rankings among degrees of belief rather than assessing their magnitudes on a cardinal scale is perhaps the most prevalent form of qualitative probability. Many languages based on this idea have been proposed; the following sections outline some of the major distinctions.

[2] Or at least the Bayesian would argue against separating acceptance of beliefs from decision making based on them. Nevertheless, the separation between so-called *theoretical* and *practical* reasoning is widely adopted in philosophy, and much research in uncertain reasoning presumes that the tasks are indeed separable. The case for categorical acceptance of uncertain beliefs can be made (in some cases) on computational grounds, but this is rarely explicit in treatments of the acceptance task.

5.1. Absolute vs. Relative

Often, ordinal qualitative probability schemes are augmented with a set of reference events measured on a cardinal scale. When all comparisons in the language are between a propositional expression and one of these reference events, we say that the language supports statements of *absolute* qualitative probability. All of the various interval calculi fall in this category.

In what cases will an absolute scheme achieve our goal of capturing the central qualities that we care about in uncertain reasoning? For example, suppose we fix a set of reference events corresponding to degrees of belief, such as 0.01, 0.05, 0.5, etc. These points may correspond to intuitive landmarks on the probability interval, but in reality it will be only by coincidence that they line up with significant thresholds in any decision or acceptance problem. (Kyburg 1994) proposes that in limited domains only a few threshold acceptance values will be required, but this claim is not really supported. The reason to be skeptical is that such thresholds are generally functions of a complex of other degrees of belief and preferences, and there is no a priori reason they should cluster around particular values. Except for the endpoints, no values of the probability interval seem to have general significance outside of some particular context.

Even if there exist some qualitatively meaningful reference probabilities, it is not clear that typical operations will respect the qualitative boundaries. For example, Bacchus (Bacchus 1989) has investigated an inheritance scheme based on properties holding with probability $\geq 1/2$, finding that it did not support chaining of inferences about typicality. Analysis of the idea of using fixed totally-ordered (though not cardinally scaled) reference events (Xiang et al. 1990) suggests that the problem is quite general, although Dubois et al. (1992) have found that a scheme with parametric thresholds can be surprisingly robust.

An alternate approach is to invoke the ordering relation among the domain events rather than special reference events. We refer to such approaches as *relative* qualitative probability. The advantage of relative approaches is that the objects of comparison are known to be meaningful in the domain, and hence their relative likelihood is more apt to be a qualitatively significant property.

5.2. Conditional vs. Unconditional

Another significant distinction among ordinal approaches is whether the objects of comparison are event probabilities or conditional event probabilities. The use of conditionals is possible whether the comparison is relative or absolute. For example, the scheme of (Dubois et al. 1992) is based on classifying within parametric interval ranges conditional expressions corresponding to terms in a syllogism. Order-of-magnitude schemes (e.g., e-semantics (Pearl 1988) or System-Z$^+$ (Goldszmidt and Pearl 1992)) rank conditional probabilities corresponding to default rules against an infinitesimal scale.

When comparing conditional expressions in a relative scheme, it also makes sense to distinguish whether the comparison varies the base proposition, the conditioning events in the expression, or both. For example, the logic of (Gärdenfors 1975) seems to allow arbitrary comparisons, whereas Neufeld's qualitative account of defaults (Neufeld 1989) makes use only of comparisons where the conditioning event is negated from one term to the other.

6. Dependence Graphs

Probabilistic independence, conditional or unconditional, is a basic qualitative concept exploited implicitly or explicitly in many probabilistic reasoning schemes. Independence imposes a constraint on probability distributions, and can be directly incorporated in languages of qualitative probability.

The logic of probabilistic independence has been studied in depth by Pearl and colleagues (1989), among others. In particular, they have proposed a set of axioms for conditional independence. Although these axioms could be used directly in a logic for reasoning about independence, much more common are representations based on *dependence graphs*, where dependence among propositions is encoded via connectivity in a graph. The most widely applied representation is that of directed acyclic graphs, variously called Bayesian, belief, causal, or probabilistic networks.

The independence relations encoded in a probabilistic network are that each node x is independent of its non-descendants given its immediate predecessors in the graph. That is,

$$Y \cap succ^*(x) = \emptyset \rightarrow I(x, pred(x), Y),$$

where $pred(x)$ denotes the direct predecessors of x, $succ^*(x)$ denotes the direct and indirect successors (i.e., nodes with a directed path from x, including x), and $I(X,Z,Y)$ means that the nodes in X are conditionally independent of those in Y given those in Z. From these explicit I relations, all that follow can be determined via *d-separation*, a sound, complete, and efficiently testable graphical connectivity criterion for conditional independence (Geiger et al. 1990).

How can we view dependence graphs as a qualitative probability logic? Although the language is not sentential in form, we can cast it sententially in several ways. First, we can treat the whole graph as a sentence, asserting the conditional independence statements above. This can be decomposed into a set of sentences $I(x, pred(x), \overline{succ^*(x)})$, one for each node x. However, in this decomposition the conditions are not local, as the I relation at each node depends on the structure of the rest of the graph.

Alternately, we could treat each link as a sentence, yet here we cannot associate a constraint with sentences, because the presence of a link by itself does not disallow any probability distributions. This suggests the final approach, taking the *absence* of links as sentences. The absence of a link does indeed express independence (if there is no link from y to x then $I(x, pred(x), y)$, assuming $y \notin succ^*(x)$), however, this too is a non-local decomposition.

Therefore, if we desire a logic with a modular semantics, that is, one where we can give a meaning to each sentence separately, it appears that we need to treat the entire graph as one sentence, or accept sentences with non-local scope. (Alternately, we could give up modularity and describe dependence graphs in a *nonmonotonic* logic (Grosof 1988).) The resulting logic is well characterized, and has a sound, complete, and efficient inference procedure based on d-separation (i.e., we can compute efficiently whether the independence conditions expressed in one graph are entailed by those of another). We could incorporate conjunction by allowing collections of graphs represent the union of associated independence conditions. This would strictly increase the expressive power of the language, though I am not sure if it would permit expression of arbitrary patterns of conditional independence.

Marginalization of a dependence graph can be accomplished via the graphical operations of node reduction and arc reversal (Shachter 1988). Although this preserves as many independencies as possible, the result may still lose information about conditional independence among the original variables. Just as adding auxiliary variables can enhance the expressive power of dependence graphs (even with respect to the original nodes) (Pearl 1988), eliminating variables can decrease it. Thus, dependence graphs are not closed under marginalization.

7. Qualitative Probabilistic Networks

Qualitative probabilistic networks (QPNs) are dependence graphs augmented with a specific form of ordinal, conditional comparison where the conditioning events vary *ceteris paribus*. Space does not permit a full account of QPNs here, but happily, there exist several available descriptions (Wellman 1990; Wellman and Henrion 1993). The main constructs in QPNs are *qualitative relations*, which specify the direction of interactions among random variables. For instance, a positive qualitative influence from node a to binary node b means that

$$a' > a'' \rightarrow \Pr(B \mid a'x) \geq \Pr(B \mid a''x),$$

for any assignment x to the nodes in $pred(b) - \{a\}$. There are also two varieties of qualitative *synergy*, which express ternary constraints capturing the interactions among qualitative influences.

As we can see above, QPNs impose constraints on probability distributions beyond those inherent in the dependence graph. However, just as for dependence graphs, it is difficult to interpret the qualitative relations sententially, as their meaning depends on the rest of the graph (in the case above, the other predecessors of node b). However, we can locally interpret all of the incoming links on a given node as a single sentence constraining the conditional distribution of a node given its predecessors. These, in conjunction with the sentence(s) describing the entire graph structure, constitute a complete description of the constraints imposed by a QPN.

Given this sketch, we can note some of the technical properties of QPNs as a qualitative probability logic. QPNs are trivially more expressive than dependence graphs, as any dependence graph can be represented by a QPN

where all links are of sign **?**. There are also sound inference algorithms, based on graph transformations (Wellman 1990) or qualitative sign propagation (Druzdzel and Henrion 1993). Both are efficient, and the latter is conjectured to be complete, at least with respect to queries about the qualitative influence relations among pairs of nodes. The transformation operations (node reduction and arc reversal) lose information; this is reflected in a lack of closure under marginalization and Bayesian revision.

8. This is not a Comprehensive Survey

I have described a general framework for characterizing schemes for qualitative probability, and identified some of the major distinctions. The analysis is not nearly exhaustive, having omitted discussion of several lines of work that are rightfully considered qualitative. One significant omission is that of order-of-magnitude systems based on infinitesimal probabilities (Goldszmidt and Pearl 1992). Also slighted are non-probabilistic approaches for qualitative uncertain reasoning. See (Darwiche 1993) for a treatment of qualitative dependence using abstract degrees of belief, and (Parsons 1993) for qualitative reasoning methods using various uncertainty calculi.

Acknowledgment

This work was supported in part by grant F49620-94-1-0027 from the US Air Force Office of Scientific Research.

References

Bacchus, F. (1989). A modest, but semantically well founded, inheritance reasoner. *Proceedings of the Eleventh International Joint Conference on Artificial Intelligence*, Detroit, MI, Morgan Kaufmann.

Darwiche, A. Y. (1993). *A Symbolic Generalization of Probability Theory*. PhD Thesis, Stanford University.

Druzdzel, M. J., and M. Henrion (1993). Efficient Reasoning in qualitative probabilistic networks. *Proceedings of the Eleventh National Conference on Artificial Intelligence*, Washington, DC, AAAI Press.

Dubois, D., H. Prade, L. Godo, et al. (1992). A symbolic approach to reasoning with linguistic quantifiers. *Proceedings of the Eighth Conference on Uncertainty in Artificial Intelligence*, Palo Alto, CA, Morgan Kaufmann.

Gärdenfors, P. (1975). Qualitative probability as an intensional logic. *Journal of Philosophical Logic* **4**: 171-185.

Geiger, D., T. Verma, and J. Pearl (1990). *d*-separation: From theorems to algorithms. *Uncertainty in Artificial Intelligence 5* Ed. M. Henrion et al. North-Holland.

Goldszmidt, M., and J. Pearl (1992). Reasoning with qualitative probabilities can be tractable. *Proceedings of the Eighth Conference on Uncertainty in Artificial Intelligence*, Palo Alto, CA, Morgan Kaufmann.

Grosof, B. N. (1988). Non-monotonicity in probabilistic reasoning. *Uncertainty in Artificial Intelligence 2* Ed. J. F. Lemmer, and L. N. Kanal. North-Holland. 237-249.

Koopman, B. O. (1940). The axioms and algebra of intuitive probability. *Annals of Mathematics* **42**: 269-292.

Kyburg, Jr., H. E. (1994). Believing on the basis of the evidence. *Computational Intelligence* **10**(1).

Neufeld, E. (1989). Defaults and probabilities; Extensions and coherence. *First International Conference on Principles of Knowledge Representation and Reasoning*, Toronto, Morgan Kaufmann.

Parsons, S. D. (1993). *Qualitative Methods for Reasoning under Uncertainty*. PhD Thesis, Queen Mary and Westfield College.

Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Mateo, CA, Morgan Kaufmann.

Pearl, J., D. Geiger, and T. Verma (1989). Conditional independence and its representations. *Kybernetika* **25**: 33-44.

Quinlan, J. R. (1983). Inferno: A cautious approach to uncertain inference. *Computer Journal* **26**: 255-269.

Savage, L. J. (1972). *The Foundations of Statistics*. New York, Dover Publications.

Shachter, R. D. (1988). Probabilistic inference and influence diagrams. *Operations Research* **36**: 589-604.

Suppes, P. (1970). *A Probabilistic Theory of Causality*. Amsterdam, North-Holland Publishing Co.

Weld, D. S., and J. de Kleer, Ed. (1989). Readings in Qualitative Reasoning About Physical Systems. Morgan Kaufmann.

Wellman, M. P. (1990). *Formulation of Tradeoffs in Planning Under Uncertainty*. London, Pitman.

Wellman, M. P. (1990). Fundamental Concepts of Qualitative Probabilistic Networks. *Artificial Intelligence* **44**: 257-303.

Wellman, M. P., and M. Henrion (1993). Explaining "explaining away". *IEEE Transactions on Pattern Analysis and Machine Intelligence* **15**: 287-292.

Xiang, Y., M. P. Beddoes, and D. Poole (1990). Can uncertainty management be realized in a finite totally ordered probability algebra? Uncertainty in Artificial Intelligence 5 Ed. M. Henrion et al. North-Holland. 41-57.