A Symmetric View of Utilities and Probabilities

Yoav Shoham*
Computer Science Department
Stanford University
Stanford, CA 94305
shoham@cs.stanford.edu
http://robotics.stanford.edu shoham

Abstract

Motivated by the need to reason about utilities, and inspired by the success of bayesian networks in representing and reasoning about probabilities, we introduce the notion of utility distributions, in which utilities have the structure of probabilities. We furthermore define the notion of a bi-distribution, a structure that includes in a symmetric fashion both a probability distribution and a utility distribution. We give several examples of bi-distributions. We also show that every state space with standard probability distribution and utility function can be embedded in a bi-distribution, and provide bounds on the size requirements of this bi-distribution. Finally, we suggest a re-interpretation of the von-Neumann and Morgenstern theorem in light of this new model.

1 introduction

AI enjoys a rich arsenal of tools with which to represent shades of certainty. One can formally represent varieties of certainty using logics of knowledge, nonmonotonic logics, and, importantly, probability theory. One reason the latter is particularly significant is that one has tools for efficiently representing and reasoning about probabilities, notably in the form of bayesian networks.

Of course, AI is interested in uncertainty in general, and in probability in particular, only to the extent that it provides help with decision making. While wildly successful, bayesian networks constitute a mechanism to reason purely about probabilities. In contrast, AI has been fairly impoverished when it comes to mechanisms for reasoning about the motivational components of decision making, such as preferences, goals, and, importantly, utilities.

*This work was supported in part by NSF grant IRI-9503109.

I don't mean that AI hasn't designed computational mechanisms to deal with some of these notions; obviously, the notions of goals and plans lie at the heart of AI planning research. What I mean is that there is not the analog of bayesian networks, that is, a simple and clear computational mechanism to effectively reason about preferences or utilities that rests on crisp, well-understood mathematical foundations.

Influence diagrams, and the related dynamic bayesian networks, are the closest we get to mechanisms for reasoning about utilities. However, while these mechanisms are undisputedly important and do contain a utility component, they provide very little power to reason about the utility component. I will discuss this further in the comparison section.

Why might we want to reason purely about utilities? The answer might be self evident, but here are a few examples:

(Personalization) A recipe-generation program can use information about the user's gastronomical preferences to devise tailored recipes.

(Software agency) A software agent needs to know its owner's preferences so as to act in the owner's best interests.

(Electronic commerce) A software agent engaged in a strategic interaction with another (human or computer) agent can exploit an understanding of the other agent's preferences to maximize its gain. For example, a software agent negotiating over access fees to a database might benefit from knowing that the database owner pays royalties to a third party on certain items and not on others.

It would be quite convenient if we had a mechanism analogous to bayesian networks to reason purely about utilities, and this paper makes a contribution in that direction. At the heart of bayesian networks lie three concepts: probability distribution, conditional probability, and probability independence. If we manage to mirror those notions in the case of utilities, we will have availed ourselves of a ready-made mechanism for reasoning about utilities. This paper is devoted to the first task, namely, defining a notion of distribution for utilities. A companion paper [Sho97] completes the story by discussing conditional utility and utility independence.

Perhaps the best starting point for explaining the approach advocated here is to note the striking asymmetry between probabilities and utilities in the traditional view (e.g., [Kre88; Sav72; Fis69]). A probability distribution has a rich structure, which allows you among other things to compute the probability of an event (that is, a set of states) and to meaningfully add the probabilities of disjoint events, whereas a utility function allows neither. The crux of my argument will be that this is an arbitrary choice, and that in fact one can define a coherent notion of utility distributions in which, for example, in makes perfect sense to add the utilities of disjoint events. In fact, the same argument will suggest that in a symmetric fashion the notion of probability can be coherently weakened to a notion of graded certainty that has only the properties of traditional utilities.

The structure of the rest of the paper is as follows.

- In Section 2 I review some of the importance manifestations of the asymmetry between probabilities and utilities. Here I discuss only the more familiar quantitative case; in the long version of this paper I discuss also the (arguably more fundamental) qualitative case.
- In Section 3 I give an overview of an alternative, more symmetric model by way of a simple example, and accompanying intuition.
- In Section 4 I provide the mathematical definition of this alternative model, and explore some of its rudimentary properties.
- In Section 5 I discuss the possible impact on the foundations of decision theory; specifically, 1 speculate on how one might re-interpet von Neumann and Morgenstern's seminal representation theorem.
- Finally, in Section 6 I discuss related work.

2 Asymmetries in quantitative probabilities and utilities

Here are some obvious asymmetries between the quantitative models of probabilities and utilities, in the traditional view.

1. Probability of each state lies in [0,1], and the probability of all states sum to 1. Utilities of states (or outcomes, or prizes) do not have these constraints. This on the surface is no big deal; certainly in the finite case one can normalize the utilities to comply with these constraints.

2. Under some conditions we can meaningfully add or subtract probabilities of different states; there appears to be no sense in performing similar arithmetic on utilities of states. This is perhaps the most telling asymmetry between current quantitative theories of probability and utility.

3. The fundamental notion in the case of probability is a probability distribution (or lottery), a function applied to sets (specifically, a function defined over a σ-algebra), whose values on different sets are constrained by the set-theoretic relationships between these sets. In the case of utility, the fundamental construct is a function applied to a single state (and, as it happens, yielding a real value). There is no principled way to lift this function to sets of states without appealing to the added notion of probability, i.e., via the notion of expected utility (but see section 6 for discussion on some proposals to lift an ordering on points to an ordering on sets of points).

In short, there are fairly interesting things we can do with probabilities alone, and almost nothing we can do with utilities alone.

3 Two examples

In this section I will explain all the important ingredients of the construction through two examples and informal discussion; the formal definitions and results will be given in the next section.

Consider the possibility of owning any of three cars: a Rolls Royce (R), a Maserati (M), and a Ford (F). This gives rise to eight different events, corresponding to whether each of these cars is owned. Suppose furthermore that we define a probability distribution over these events, and also attach utilities to each of the states, as follows:

state	–	R	M	F	RM	RF	FM	RFM
prob.	1/8	1/8	1/8	1/8	1/8	1/8	1/8	1/8
util.	0	0.1	0.2	0.7	0.3	0.8	0.9	1.0

(The uniform distribution is chosen here for concreteness; any other would do as well.) We can sum any of these probabilities and get something meaningful; for example, we can compute the probability of owning a Ford by summing the probabilities of the four states that include a Ford, getting the value 1/2. We cannot meaningfully sum any utilities we wish. However, it is easy to see that the utility function in this example has a special structure. Specifically, it *can* be interpreted as arising from assigning the three cars the values 0.1, 0.2, and 0.7, respectively, and defining the utility of any subset of the three cars as the sum of the individual utilities of the cars in the subset.[1]

*The reader familiar with multi-attribute utility theory will recognize that this is a special case of MAUT; more on

Given this observation, let us construct another space as follows:

factor	r	m	f
"utility"	0.1	0.2	0.7
"probability"	1/2	1/2	1/2

The first two rows of this structure form an instance of what I will call a *utility distribution.* The states in this distributions will be called *factors.* (From now on I'll reserve the term 'states' to denote elements on which a probability distribution is defined, and 'factors' to denote elements on which a utility distribution is defined.) Factors are informally thought of as the various independent contributions to one's sense of satisfaction or well being. The utilities associated with each factor determine how much of one's sense of well being is supplied by that factor. A *factor set* is simply a set of factors, for example {r,f}, and it plays a role analogous to an event in a probability distribution; its contribution to one's satisfaction is simply the sum of the contributions of its members. One can additionally attach to factors "probabilities" and on the basis of those compute expected utilities, but these are mere numbers that cannot be meaningfully added up. In other words, this structure is the exact dual of the structure we started out with.

Finally, note that these two structures are intimately connected, in the sense that certain factors "co-occur" with certain states. Specifically, r co-occurs with all the states that contain R, and ditto for m and f. So in fact, one need not explicitly list the utilities in the first structure, nor the "probabilities" in the second structure. These can be inferred: The utility of a state is the sum of the "utilities" of all factors that co-occur with it, and the "probability" of a factor is the sum of the probabilities of all states that co-occur with it.

This is a convenient situation; not only have we managed to separate the representation of probabilities from that of utilities, but the representation of utilities is identical to that of probabilities, providing hope that we can use, e.g., the mechanism of bayesian networks to reason about utilities.

On the face of it, our example was contrived so as to make this separation possible. Can we view *any* given utility function as having this particular form? We will see in the next section that the answer is 'yes,' but that the set of factors might not be as small as this particular example suggests. The next example illustrates this point, and also the fact that the notion of 'factor' is quite broad.

Suppose we have seven envelopes; one contains one $1 bill, another contains two $1 bills, and so on up to seven $1 bills. A subject receives one of these envelopes drawn from a given distribution. For concreteness, consider uniform distribution, and a sigmoid-like utility function; the smallest-valued envelopes have little value since there's little you can buy with them, then they start to pick up steam, and at some point their value starts to level off because there's just that many things you can buy. This is captured in the following table:

state	$1	$2	$3	$4	$5	$6	$7
prob.	1/7	1/7	1/7	1/7	1/7	1/7	1/7
util.	.03	.1	.3	.7	.9	.97	1.0

As usual, it would make no intuitive sense to ask what the numerical utility was of the set consisting of the second and fifth envelopes, without appealing to probabilities. However, let us now construct the dual space where we can ask these sorts of questions. This time the factors will consist of the seven pairs *(k, k—* i) for 1 < *k* < 7; the informal interpretation of the pair will be the Ar-th "order statistic", that is, the *k'th* $ bill. The utility of the k-th $ bill will be defined as the difference between the utilities (in the first, "state" space) of the $k-envelope and the *$(k —* l)-envelope, that is, the marginal utility of the k-th $1 bill (we take the utility of a $0-envelope to be 0, to cover the case of *k* = 1). In our concrete example, these marginal utilities form a bell-shaped curve; the marginal utility drops the closer *k* is to either 1 or 7. This is captured in the following table:

factor	1st	2nd	3rd	4th	5th	6th	7th
util.	.03	.07	.2	.4	.2	.07	.03
prob.	1	6/7	5/7	4/7	3/7	2/7	1/7

In this factor space it makes perfect sense to add the different contributions to joy: the second $1 bill adds an independent value from that of the fifth $1 bill.

Here's an important point. You might worry that "it doesn't make sense to add up the utilities of only the second and fifth $1 bills, since we can't get the two bills without also getting all first five." But that's confusing states with factors; the fact that one cannot experience a particular combination of factors doesn't mean that this combination has no meaning or value.

Of course, we would like to capture the fact that only certain factor combinations are possible, and here is where the co-occurrence relation between states and factors comes in. In the first example, for each combination of factors there was a state whose co-occurring factors were exactly that combination. This is not the case in this second example. Here, the k -th-$ factor co-occurs with the $/-envelope iff *k* < *I.*

A final word about the numerical range of probabilities and utilities, before we move on to the formal treatment. Why do we insist that utilities sum to 1? In fact we don't have to, but nor do we in the case of probabilities. The question is whether we wish to model relative notions of "chunk of reality" or "chunk of satisfaction," or absolute

this in the comparison section.

notions. Relative notions model totality of the quantity as 1, and portions of totality as fractions. Absolute models have no notion of totality of the quantity, and measure portions in some arbitrary units. Traditionally, probability has been modeled as relative and utility as absolute, but that (I believe) is an arbitrary choice. Here I will continue to model both as relative.

4 The formal model

In the formal treatment I'll only discuss finite, fully measurable sets. The first definition is standard:

Definition 1 *A* distribution *is a pair* (5, *F) such that S is a finite set (whose members are called* components/, *and F is a function* $F : S \rightarrow [0,1]$ *such that* $\Sigma_{s\in S}F(s) = 1$. *F is lifted to* 2^S *(the set of all subsets of S) by* $F(x) = \Sigma_{s\in x}F(s)$.

The previous definition intentionally avoids a particular interpretation of the notions, so that they can be applied to both probabilities and utilities. The following definitions could also be presented in neutral terms, but that generality is not needed in this paper.

Definition 2 *A* bi-distribution, *also called a* pu-distribution, *is a triple* $((S_p, P), (S_u, U), L)$ *such that*

- *(S_P, P) is a distribution called the* probability distribution,
- *(S_U, U) is a distribution called the* utility distribution,
- *S_p and* S_u are *disjoint, and*
- *L is a nonempty subset of S_p x S_u.*

Next, we use the structure to compute expected utilities:

Definition 3 *Given a bi-distribution as above, the expected utility of a set $x \subseteq S_p$, denoted $EU(x)$, is defined by*

$$EU(x) = \Sigma_{s\in x,(s,t)\in L}P(s)U(t)$$

The expected utility of a non-empty set $x \subset S_u$, also denoted $EU(x)$, is defined symmetrically:

$$EU(x) = \Sigma_{t\in x,(s,t)\in L}P(s)U(t)$$

Proposition 1 *Given a bi-distribution as above,*

$$EU(S_p) = EU(S_u) = \Sigma_{(s,t)\in L}P(s)U(t)$$

The next proposition delivers on the promise that one need not get lucky to enjoy the bi-distribution representation: it states that given any probability distribution with an associated utility function on its states, we can embed the probability distribution in a bi-distribution that induces back the original utilities on the components of the probability distribution.

Proposition 2 *Given a probability distribution (S_p, P) and a utility function $U' : S_p \rightarrow [0,1]$, there exists a bi-distribution $((S_p, P), (S_u, U), L)$ such for any $s \in S_p$ it is the case that*

$$U'(s) = \Sigma_{(s,t)\in L}U(t)$$

and therefore also for any set $x \subseteq S_P$ it is the case that

$$\Sigma_{s\in x}P(s)U'(s) = EU(x)$$

Proof: By construction similar to the envelopes example from the previous section. Order the states in S_p in ascending utility according to U', so that $U'(s_1) \leq U'(s_2) \leq \ldots \leq U'(s_n)$. Now create $n+1$ factors $t_1, t_2, \ldots, t_{n+1}$, and define the following utility distribution on them: $U(t_1) = U'(s_1)$, $U(t_2) = U'(s_2) - U'(s_1)$, $\ldots$, $U(t_n) = U'(s_n) - U'(s_{n-1})$, $U(t_{n+1}) = 1 - U'(s_n)$. Finally, let L consists of all the links (s_i, t_j) such that $i \leq j$.

This above construction appears to be wasteful in the number of factors created in order to represent a given utility function. We might ask how small that set of factors might be.

Proposition 3

(lower bound) Given any set S of n distinct natural (real) numbers, there does not exist a set T of less than $\log_2 n$ *natural (real) numbers such that each number in S is the sum of some numbers in T.*

(upper bound) There exist infinitely many sets S of n distinct natural (real) numbers (including infinitely many in the interval [0,l]), for which there does not exist a set T of less than n natural (real) numbers such that each number in S is the sum of some numbers in T.

Proof: (lower bound) k numbers yield at most 2^k distinct sums, *(upper bound)* Take S to be any finite set of distinct powers of 2, for example {1,2,4,8}, or any set of distinct negative powers of 2, for example {1,1/2,1/4,1/8}.

These facts still leave open several important questions. Perhaps the most important mathematical question is what is the minimal set of factors required for a *given* set of utilities. And most importantly, these mathematical constructions do not address the question of how natural the factors are.

Finally, a short and somewhat whimsical comment on expected utilities. So far, we've discussed probability and utility as rich notions, but expected utility as little more than the result of numerical calculations on given probabilities and utilities. It's possible to invert the picture, and view the links of a bi-distribution as the basic elements of our world model. The link *(s,t)* will be assigned the weight *P(s)U(t),* and the weight of sets of links will be simply the sum of their weights. What do those links represent? Well, if the state *s* describes a piece of reality, and the factor *t* describes a piece of joy, then surely the pair *(s,t)* describe a piece of real(ized) joy. While it's hard to object to a theory of real joy, this line of reasoning does not play a role in the paper and is not developed further.

5 The scriptures revisited

The primary motivation for this paper has been the search for a mechanism for reasoning about utilities. In service of this goal, the particular focus of this paper has been to level the playing field between probabilities and utilities. However, the notion of utility distributions calls into question foundational assumptions in choice theory regarding the asymmetric roles and structure of probabilities and utilities. Indeed, it suggests a reinterpretation of some of the most influential developments in choice theory. Here we will discuss one of them the representation theorem of von Neumann and Morgenstern.[2]

In the foundations of decision theory, mental notions such as utilities are usually presented as convenient auxiliary constructs, to be justified based on other, observable phenomena such as choices made. The representation theorem of von Neumann and Morgenstern, and especially that of Savage (the "crowning achievement of decision theory," to quote one mathematical economist), are among the deepest embodiments of this "revealed preference" doctrine. Some economists have argued to me in private that the fact that the traditional properties of probabilities and utilities enable these deep theorems is in itself justification for accepting these properties.

Recall von Neumann and Morgenstern's theorem (here presented in its finite version). The setting is a finite set of prizes (or outcomes) Z, the set *P* of all (!) probability distributions over Z, and a binary relation > on *P*. The intended interpretation of $p > q$ is "p is preferred to q."

vNM introduce the following three postulates:

(Preference) $>$ is asymmetric (if $a > b$ then not $b > a$) and negatively transitive (if $a > b$ then either $a > c$ or $c > b$).

(Independence) For all $p, q, r \in P$ and $a \in (0,1]$, if $p > q$ then $ap + (1-a)r > aq + (1-a)r$.

(Continuity) For all $p, q, r \in P$, if $p > q > r$ then there exist $a, b \in (0,1)$ such that $ap + (1-a)r > q > bp + (1-b)r$.

and prove their famous representation theorem:

vNM Theorem:
Given Z, P and $>$ as above, $>$ satisfies the three postulates iff there exists a function $u : Z \rightarrow \mathcal{R}$ such that

$$p > q \text{ iff } \Sigma_{z \in Z} p(z) u(z) > \Sigma_{z \in Z} q(z) u(z)$$

(with the usual overloading of the > symbol). Furthermore, *u* is unique up to positive affine transformation.

I will now present a new representation theorem, which I'll call the vMN theorem. The setting for the vMN theorem consists of a set of factors Z, the set *P* of all utility distributions over Z, and an ordering > on P. The intended interpretation of $p > q$ is "*p* is preferred to *q*."[3]

We are now ready to present the vMN theorem, but we need not; it is identical to the **vNM** theorem.

This is of course tongue-in-cheek; I've presented no new mathematics, and have merely re-interpreted the vNM theorem. However, I have hopefully made the point that the conceptual foundations of the vNM theorem are open to debate. While the theorem was presented with a certain interpretation in mind, it admits at least one other interpretation as well, in which the roles of probabilities and utilities are reversed. Specifically, the original interpretation suggests a picture of a selfish person attempting to select among lotteries presented to him so as to maximize his own expected payoff. The new interpretation suggests a picture of a benevolent person attempting to select among multiple other people with varying tastes, in order to maximize the payoff to the selected person. (Note that the properties represented by the three postulates seem to make as much sense under the reverse interpretation as under the intended one).

Note, by the way, that this re-interpretation of the vNM theorem flies in the face of its reputation as the most extreme embodiment of objectivist view on probabilities. In the re-interpretation, the utilities are exogenous, or objective, and the "probabilities" (quoted here since they no longer have the rich structure of a distribution) are imputed, or subjective.

I believe similar discussion is possible in the context of the more complex representation theorem of Savage, but that is beyond the scope of this paper.

6 Related work

The advertized motivation for the work described here is a dearth of mechanisms to reason about utilities. We should mention one weak exception to this dearth, namely influence diagrams [Sha90]. These have (in addition to *chance* and *decision* nodes) a special node called a *value* node. This node, which can have no successors and cannot be part of the evidence set, is merely used to compute a given expected utility function as a result of evidence propagation. The usual way this utility is used is to search the space of values for decision nodes so as to maximize this utility. Thus, while technically speaking

[2] All the non-heretical material in this section here can be found in any standard textbook, such as [Kre88].

[3] Yet a different version would substitute here "p is mroe likely than g," but we do not pursue this further here.

influence diagrams reason with utilities, in fact all their smarts is in how they represent probabilities.

One minor modification of the basic influence diagram is to introduce several value nodes, and to sum the utilities of all the value nodes. This is based on assumption that the utility function is decomposable in such as way. The theory governing such decomposable utility functions is multi-attribute utility theory (MAUT) [KR76]. MAUT has attracted some attention in AI in recent years, since it seems to offer a handle on complexity. Indeed, utility distributions are closely related to MAU functions. It is beyond the scope of this paper to discuss this connection in detail; this is precisely the topic of the companion paper [Sho97]. Here I will only remark that from the mathematical point of view utility distributions can be seen as a special case of additive multi-attribute utility functions, but that the difference is also conceptual in nature, and hinges on novels senses of conditional utility and utility independence.

The main "complaint" in the paper about the standard notion of utility function has been that it applies to states but not to events, or sets of states. There have been several proposals in AI to lift an ordering defined on points (whether the ordering reflects a degree of certainty or a preference) to an ordering on sets, for example by Doyle and Wellman [DW95] and Halpern [Hal96]. All these proposals are qualitative in nature; they usually boil down to quantifying over points in various sets, which reduces to using the Min and Max operations in some combination. None of these proposals have the quantitative flavor afforded by (probability or utility) distributions.

7 Summary and what's next

There is growing interest in AI in representing and reasoning about utilities. We have suggested that endowing utilities with the properties of utilities will get us closer to the goal of applying bayesian-network-like mechanisms to utilities. The primary contribution of this paper has been to introduce the notion of utility distribution, and the related notion of bi-distribution. A side effect of this development has been to call into question conventional wisdom from mathematical decision theory.

This is clearly only the beginning of the story. Now that we have a sense for utility distributions, we can revisit the familiar notions of conditional utility and utility independence. As was mentioned, this is the topic of [Sho97], where we show new senses of these notions that are isomorphic to their probabilistic counterparts. This means that, at least in principle, we can use the mechanism of bayesian networks (that would more aptly be called *utility networks* in this context) to reason about utilities. However, whether this promise can be realized, depends on whether one can take these formal ideas and apply them in practice. Can one identify natural factors in realistically large and natural domains? Can we in fact elicit preferences using the new model? It must be admitted that at this time factors seem more mysterious than states. It is not clear to me if this is a reflection of their novelty, the inherent elusiveness of mental state, or the fact that factors are not in general a natural category. While I'd like to believe that the framework described here, and further developed in [Sho97], is more than an idle exercise, only time will tell.

Acknowledgements. I have discussed the ideas described here informally with many people, who have made very useful suggestions and other comments. Among the AI/CS people are Xavier Boyen, Urszula Chajewska, Denise Draper, Moises Goldszmidt, Daphne Roller, Christos Papadimitriou, Stuart Russell, and Mike Wellman. Among the economists are Ken Arrow, Paul Milgrom, and especially Tzachi Gilboa. In addition the referees made useful comments on an earlier draft of the paper. This, however, does not imply that any of the above necessarily agree with the ideas expressed here.

References

[DW95] J. Doyle and M. P. Wellman. Defining preferences as ceteris paribus comparatives. In *Proc. AAAI Spring Symp. on Qualitative Decision Making,* pages 69-75, 1995.

[Fis69] P. C. Fishburn. *Utility Theory for Decision Making.* John Wiley &; Sons, Inc., 1969.

[Hal96] J. Y. Halpern. Defining relative likelihood in partially-ordered structures. In *Proc. Twelveth Conference on Uncertainty in Artificial Intelligence,* pages 299-306, 1996.

[KR76] R. H. Keeney and H. Raiffa. *Decision with Multiple Objectives: Preferences and Value Tradeoffs.* John Wiley & Sons, Inc., 1976.

[Kre88] D. M Kreps. *Notes on the Theory of Choice.* Westview Press, Boulder, Colorado, 1988.

[Sav72] L. J. Savage. *The Foundations of Statistics.* Dover Publications, Inc., 1972. (2nd edition).

[Sha90] R. D. Shachter. Evaluating influence diagrams. In G. Shafer and J. Pearl, editors, *Readings in uncertain reasoning,* pages 79-90. Morgan Kaufmann Publishers, 1990.

[Sho97] Y. Shoham. Two senses of conditional utility. In *Proc.of UAI-97,* page (to appear), 1997.