Expected Qualitative Utility Maximization*

Daniel Lehmann

Institute of Computer Science, Hebrew University, Jerusalem 91904, Israel
E-mail: lehmann@cs.huji.ac.il

A model for decision making that generalizes Expected Utility Maximization is presented. This model, Expected Qualitative Utility Maximization, encompasses the Maximin criterion. It relaxes both the Independence and the Continuity postulates. Its main ingredient is the definition of a qualitative order on nonstandard models of the real numbers and the consideration of nonstandard utilities. Expected Qualitative Utility Maximization is characterized by an original weakening of von Neumann-Morgenstern's postulates. Subjective probabilities may be defined from those weakened postulates, as Anscombe and Aumann did from the original postulates. Subjective probabilities are numbers, not matrices as in the Subjective Expected Lexicographic Utility approach.

Key Words: Utility Theory, Non-Standard Utilities, Qualitative Decision Theory

1. INTRODUCTION

This paper proposes a deviation from the von Neumann-Morgenstern (vNM) postulates paradigm in decision theory. It suggests that it may, in certain circumstances, be rational to judge alternatives according to the issues in their main focus. If these issues are deemed as equal, the alternatives will continue to be judged as equal even if they are mixed with different side-effects - side-effects which would not be judged as equal were they to be in the main focus themselves. This contradicts the vNM Independence axiom. Moreover, outcomes of issues in the main focus may completely overshadow the outcomes of the side-issues, so alternatives of the former may be judged as infinitely more (or infinitely less) preferable than those of the latter. This contradicts the vNM Continuity axiom.

*This work was partially supported by the Jean and Helene Alfassa fund for research in Artificial Intelligence and by grant 136/94-1 of the Israel Science Foundation on "New Perspectives on Nonmonotonic Reasoning".

Suppose you have to choose between two lotteries. In the first lottery you may win, with probability p, a week's vacation in Hawaii. With probability $1-p$ you get nothing. In the second lottery you may win, with the *same* probability p, the same vacation in Hawaii, but, with probability $1-p$ you get a consolation prize: a free copy of your favorite magazine. Since the free copy is preferred to nothing, vNM's independence postulate implies that lottery two is preferred to lottery one. But couldn't a rational decision maker be indifferent between the two lotteries. One, often, I think, buys a lottery ticket in a frame of mind focused on the big prize and not on the consolation prize. This behavior is by no means general, as attested by the fact lotteries often offer consolation prizes, but should a decision maker indifferent between the two lotteries be considered *irrational* in all situations? I think not.

A variation on this example considers also a third lottery, in which one wins a trip to Paris with the same probability p as above, and nothing with probability $1-p$. Suppose you try to compare lotteries one and three. You ponder at length the advantages and disadvantages of the two vacation spots, and decide you are indifferent between the trip to Hawaii and the trip to Paris, all relevant considerations taken into account. You conclude that you are also indifferent between the lotteries one and three. The independence axiom of von Neumann and Morgenstern implies that lottery two is preferred to lottery three. But it has been argued that it is quite unreasonable to expect the very slight improvement that lottery two presents over lottery one to overcome the lengthy and delicate deliberation that made you conclude that the trips to Hawaii and Paris are equivalent for you.

Similar examples have been put forward to argue that the indifference relation is not always transitive: lottery three is equivalent to lottery one and to lottery two, but lottery two is preferred to lottery one. The system presented in this paper endorses the transitivity of indifference, but allows a decision maker to be indifferent between lotteries one and two.

Let us consider now a second situation. A patient has to choose between two options:

1. (option p) do nothing and die in a matter of weeks,
2. (option q) undergo surgery, the result of which, depending on some objective probabilities, may be a long and happy life (denoted by l) or an immediate death on the operation table (denoted by d). We shall denote by λ the probability of the surgery being successful, i.e., the probability of l. The probability of death on the operation table is, therefore, $1-\lambda$.

In other terms, one has to compare *mixtures* p and $\lambda\, l + (1-\lambda)\, d$. Assuming one prefers l to p and prefers p to d, vNM's postulates imply what we shall call property **P**: there is a unique $\lambda \in]0,1[$ for which one is indifferent between p and $\lambda\, l + (1 - \lambda)\, d$. For any $\mu > \lambda$, one prefers $\mu\, l + (1 - \mu)\, d$ to p, and for all $\mu < \lambda$, one prefers p to $\mu\, l + (1 - \mu)\, d$. If one thinks that a long and happy life is overwhelmingly preferable to p so as to make the distinction between p and d insignificant, i.e., that $\mu\, l + (1 - \mu)\, d$ is preferable to p for any $\mu \in]0, 1[$, property **P** fails and one deviates from vNM's rationality postulates.

But would we really dismiss as *irrational* such a behavior, or such a preference? After considering the question, in this section, I will propose a weakening of vNM's postulates that allows for the preference of $\mu\, l + (1 - \mu)\, d$ over p for any $\mu \in]0, 1[$.

Von Neumann-Morgenstern's point of view is perfectly acceptable and I have no criticism for someone who adheres to it and decides there is indeed some μ, very close to zero, perhaps, such that p is equivalent to $\mu\, l + (1 - \mu)\, d$. My only claim is that someone who thinks $\mu\, l + (1 - \mu)\, d$ is preferable to p for any $\mu \in]0, 1[$ cannot be considered *irrational* outright.

An argument, very similar to the one just presented, for preferring a mixture $\mu\, l + (1 - \mu)\, d$ to some p, for any $\mu \in]0, 1[$, even though p is preferred to d, appears in Pascal's [23]. There l denotes eternal bliss (the reward of the believer if God exists), d denotes a life spent in error by a believer in a God that does not exist, and p denotes a life spent by a non-believer. This argument, known as Pascal's wager, is very well-known and the reader may find in a previous version of this work [22], a fuller discussion.

Before entering a technical description, I will, in short, describe intuitively the reason for the failure of property **P** in the system presented in this work, Expected Qualitative Utility Maximization. In the examples above one may assume that consequence l is *infinitely* preferable to p and d. This will be technically translated in giving an infinite value to the utility of l, whereas p and d have finite utilities. For von Neumann and Morgenstern, there are no infinite utilities. This work develops a framework in which there are infinite utilities, by using nonstandard (in the sense of A. Robinson [25]) utilities.

A number of papers [5, 28, 2] discussed, in the setting of the St. Petersburg's paradox, the existence of unbounded utilities. In the last of these papers, Aumann argues very convincingly that utilities should be bounded. At first sight, one may think that infinite utilities imply unbounded utilities, and therefore Aumann argues also against infinite utilities, but this is not the case. His argument may be summarized in the following way: if utilities were unbounded, for any $\lambda \in]0, 1[$ there would be a consequence (c) such that a lottery $\lambda\, c + (1 - \lambda)\, d$ is preferred to a long and happy life (l). But this seems very unreasonable to Aumann. His argument is directed against *unbounded* utilities, but ineffective against *infinite* utilities. Certainly, no consequence is infinitely preferable to l and therefore, if there are infinite utilities, the utility of l is infinite. But there is absolutely no problem if one assumes that the utility of l is infinite and maybe even maximal (nothing is preferred to l). In this case, Aumann's argument disappears. The fact that l is infinitely preferred to some other consequence, d for example, or a sum of money, will influence the preferences of a decider between lotteries involving l and consequences such as d.

2. BACKGROUND

Utility theory is discussed in the framework of [30, Chapter 3], see also [11]. Let $\mathcal{H}$ be a boolean algebra of subsets of X, and P a convex set of probability measures on $\mathcal{H}$. Convex means here that:

$$\forall p, q \in P,\ \forall \lambda \in]0,1[,\ \lambda p + (1-\lambda) q \in P.$$

Here $]0,1[$ denotes the standard open real interval. Von Neumann and Morgenstern have characterized the binary relations $>$ on P that can be defined by a linear functional u on P in the following way:

$$\forall p, q \in P,\ p > q \ \Leftrightarrow\ u(p) > u(q). \tag{1}$$

In Equation 1, the functional u is a function from P to the standard set of real numbers $\mathbf{R}$ and the relation $>$ in the right hand side is the usual strict ordering on $\mathbf{R}$.

Their characterization is equivalent to the following, due to Jensen [13] (see [11, p. 1408]) three conditions, for all $p, q, r \in P$ and all $\lambda \in]0,1[$ (a weak order is an asymmetric and negatively transitive binary relation):

$$\mathbf{A1} \quad > \text{ on } P \text{ is a weak order},$$

$$\mathbf{A2} \quad p > q \ \Rightarrow\ \lambda p + (1-\lambda) r > \lambda q + (1-\lambda) r,$$

$$\mathbf{A3} \quad (p > q, q > r) \ \Rightarrow\ \exists \alpha, \beta \in]0,1[, \text{ such that}$$

$$\alpha p + (1-\alpha) r > q > \beta p + (1-\beta) r.$$

The three conditions above are not the original postulates of von Neumann and Morgenstern, they are equivalent to them. They will be referred to, nevertheless, in this work, as vNM's postulates. **A2** is usually denoted Independence and **A3** Continuity. The purpose of this work is to generalize von Neumann and Morgenstern's characterization to deal with *qualitative* utilities. In the sequel, $p \geq q$ will denote $q \not> p$ and $p \sim q$ will denote the conjunction of $p \geq q$ and $q \geq p$.

3. QUALITATIVE UTILITIES

Qualitative decision theory has been developed mostly in opposition to quantitative decision theory, stressing decision methods that do not satisfy vNM's or Savage's [26] postulates, the postulates generally accepted for quantitative decision theory. The focus in qualitative decision theory has always been on methods and algorithms, more than on an axiomatic treatment ([4] is an exception).

A different approach is proposed here: qualitative and quantitative decision theories can be viewed as special cases of a unified general theory of decision that contains both. This unified theory is a generalization of the quantitative theory. The power of the generalization lies in the consideration of nonstandard models of the set of real numbers for utilities and a definition of indifference that neglects infinitesimally small differences. In this paper probabilities will always be standard. The case of nonstandard probabilities and standard utilities has been treated, after the elaboration of this paper, in [21]. Preliminary ideas already appeared in [19] and [20].

A well-established tradition in Decision Theory considers Expected Utility Maximization as the only rational policy. Following this view, an act f is strictly preferred to an act g iff the utility expected from f is strictly larger than that expected for g. Since expected utilities are real numbers, *strictly larger* has its usual, *quantitative* meaning. The main claim of this paper is that the qualitative point of view may be subsumed by a slightly different definition of *strictly larger*. Suppose we consider any model elementarily equivalent to the real numbers, more precisely, any (standard or nonstandard) model of the real numbers, $\mathcal{R}$ (for the standard model, we shall use $\mathbf{R}$). Let x and y be elements of $\mathcal{R}$. To make matters simpler, suppose that both x and y are positive. The number x is quantitatively larger than y iff $x - y > 0$. What could be a reasonable definition of *qualitatively larger*? Clearly, if x is qualitatively larger than y then it must be quantitatively larger: in a sense *qualitatively larger* means *definitely larger*. A first idea that may be considered is to use a notion that proved fundamental for nonstandard analysis (the monads of [25], or see [16]): the notion of two numbers being *infinitely close*, and consider that a number x is *qualitatively* larger than a number y iff x is larger than y and *not infinitely close* to y, i.e., iff $x - y$ is strictly larger than some positive standard number. At first this idea looks appealing: if ϵ is strictly positive and infinitesimally close to zero, and x is a standard, strictly positive, real, we do *not* want $x + \epsilon$ to be qualitatively larger than x. At a second look, one realizes that the size of $x - y$ should not be judged absolutely, but relatively to the size of x: for example $\epsilon^2 + \epsilon$ should be qualitatively larger than ϵ^2. Therefore I propose the following definition:

Definition 3.1. Let x and y be positive. We shall say that x is qualitatively larger than y and write $x \succ y$ iff $\frac{x-y}{x}$ is strictly positive and *not* infinitesimally close to zero: in other terms, iff there is a strictly positive *standard* number r such that $\frac{x-y}{x} \geq r$.

To the best of my knowledge, the idea of comparing nonstandard numbers qualitatively, i.e., by $\succ$ and not by the usual ordering $>$ is original. Notice also that Definition 3.1 relies on the notion of a *nonstandard* number, and that this notion is not first-order definable. We may therefore expect for the qualitative order properties that are quite different than the ones enjoyed by the usual order. The qualitative order is *not* a lexicographic ordering. It will be shown in Section 7 that its properties are very different from those of lexicographic orderings.

The definition may be extended to arbitrary numbers in the obvious way:

1. if $x \geq 0$ and $0 > y$, then $x \succ y$, and
2. $x \succ y$ iff $-y \succ -x$,

but we shall limit ourselves to the consideration of positive utilities in this paper. $x \preceq y$ shall denote $x \not\succ y$ and $x \sim y$ shall denote that $x \preceq y$ and $y \preceq x$. Properties of the qualitative ordering $\succ$ are described and proved in Appendix A.

Note that $x \preceq y$ does not imply $x - c \preceq y - c$. If ϵ is positive and infinitesimally close to zero, $1 + \epsilon \preceq 1$ but $1 + \epsilon - (1 - \epsilon^2) \not\preceq \epsilon^2$. Note also that $\succ$ is not complete and that $\preceq$ is not antisymmetric. Notice that, if we choose, for $\mathcal{R}$, the standard model of the reals, $\mathbf{R}$, then $x \succ y$ iff $x > y$. Therefore our treatment would include the classical approach, if we allowed also negative utilities. As said above, in this paper, we concentrate on the case all utilities are positive. Is our framework, with positive nonstandard utilities, a generalization of the classical theory, with standard positive and negative utilities? Since, in the classical setting, utilities are defined only up to an additive constant, bounded utilities may always be considered to be positive, by adding a positive large enough constant. In view of Aumann's [2] critique of unbounded utilities, we feel that the present framework encompasses the most important part of classical theory.

Expected Qualitative Utility Maximization is the version of Expected Utility Maximization that obtains when, for utilities,

- the model chosen for the real numbers *may* be nonstandard, and
- real numbers are compared *qualitatively*, i.e., by $\succ$.

We want to characterize the binary relations $>$ on P that can be defined by a linear functional $u : P \rightarrow \mathcal{R}_+$ in the following way:

$$\forall p, q \in P, p > q \Leftrightarrow u(p) \succ u(q). \tag{2}$$

One could, as well, consider functionals of the type: $u : P \rightarrow \mathcal{R}$, but the characterization of relations obtained in such a general way is more intricate than the task proposed in this paper and is left for future work. We shall note the main difference between the two characterizations in the sequel. One should immediately notice that, by Lemma A.1 part 9., if $c > 0$, the utility function $c\, u(p)$ defines the same ordering as $u(p)$. But, contrary to what happens both in the classical and in the lexicographic ordering settings, if $d \in \mathcal{R}$ the function $d + u(x)$ does not, in general, define the same ordering as $u(p)$: consider Lemma A.1 part 11.. Such an instability under an additive constant, and in particular an asymmetry between gains and losses has been found in the behavior of decision makers in many instances [10, 14, 27]. The question of whether Expected Qualitative Utility Maximization, in the present form, or in the generalized version that allows for negative utilities, is a realistic model for explaining such behavior cannot be discussed in this paper.

4. MAXIMIN AS EXPECTED QUALITATIVE UTILITY MAXIMIZATION

As noticed above, Expected Qualitative Utility Maximization generalizes Expected Utility Maximization: if one chooses the standard model for real numbers then Expected Qualitative Utility Maximization boils down exactly to Expected Utility Maximization, at least when utilities are bounded. We shall show now that considering nonstandard utilities enables us to obtain decision criteria that do not satisfy vNM's postulates and were so far considered as part of the realm of *qualitative* decision theory. A version of the Maximin criterion will be presented. The Maximin criterion has been proposed by A. Wald [31], in a different framework. The criterion to be presented is a variation on this theme.

Notice that any bona fide Maximin criterion violates Independence, and is therefore not obtainable by a lexicographic ordering, since all such methods satisfy Independence. Suppose $p \succ q \succ r$. Then, under the maximin criterion $\lambda r + (1 - \lambda)p$ must be considered indifferent to $\lambda r + (1 - \lambda)q$ since in both the worst possible outcome is the same and obtains with the same probability, violating Independence.

Assume the set X is finite and $\mathcal{H}$ contains all subsets of X. Let the elements of X be $x_0, \ldots, x_{n-1}$. Let ϵ be a number that is positive and infinitesimally close to zero and let our utility function u be the linear function defined by: $u(x_i) = \epsilon^{n-i-1}$, for $i = 0, \ldots, n-1$. Notice that $x_i < x_j$ iff $i < j$. The utility of a mixture $\lambda x_i + (1 - \lambda)x_j$ is $\lambda\epsilon^{n-i-1}$ if $x_i < x_j$ and ϵ^{n-i-1} if $x_i \sim x_j$. Suppose $i < j$ and $i' < j'$. Then, $\lambda x_i + (1 - \lambda)x_j < \mu x_{i'} + (1 - \mu)x_{j'}$ holds iff $x_i < x_{i'}$ or $x_i \sim x_{i'}$ and $\lambda > \mu$. The decision maker therefore compares different mixtures by comparing the worst possible outcomes and, if they are the same, their respective probabilities. This is some form of Maximin criterion and does not satisfy **A2** or **A3**, but it has been considered a rational way of deciding by many authors, and it is amenable to Expected Qualitative Utility Maximization.

5. POSTULATES

Suppose $>$ is defined by some functional u as in Equation 2. First, one easily sees that the relation $>$ defined by Equation 2 is a weak order since $\succ$ is a weak order. Therefore **A1** holds.

Let us know consider **A2**. We shall see that **A2** does not hold. Let ϵ be positive and infinitesimally close to zero. Let $u(p) = 2 \times \epsilon$, $u(q) = \epsilon$ and $u(r) = 1$. Then $0.5 \times 2 \times \epsilon + 0.5 \sim 0.5 \times \epsilon + 0.5$ and $0.5p + 0.5r \sim 0.5q + 0.5r$. Let us find a weakening of **A2** that holds. Assume now that $p > q$, i.e., $u(p) \succ u(q)$. Let $\lambda \in \mathbf{R}$, be a standard real number, $0 < \lambda < 1$. Let us compare $\lambda p + (1 - \lambda)r$ and $\lambda q + (1 - \lambda)r$. Since u is linear,

$$u(\lambda p + (1 - \lambda)r) = \lambda u(p) + (1 - \lambda)u(r)$$

and

$$u(\lambda q + (1 - \lambda)r) = \lambda u(q) + (1 - \lambda)u(r).$$

By Lemma A.1, parts 10. and 11., two cases must be considered. If $(1 - \lambda)u(r)/\lambda u(p)$ is finite, i.e., if $u(r)/u(p)$ is finite, then,

$$\lambda u(p) + (1 - \lambda)u(r) \succ \lambda u(q) + (1 - \lambda)u(r),$$

and therefore $\lambda p + (1 - \lambda)r > \lambda q + (1 - \lambda)r$, which corresponds exactly to **A2** above. If, on the contrary, $u(r)/u(p)$ is infinite, then, for any $\lambda' \in]0, 1[$ (the *standard* unit interval),

$$(1 - \lambda')u(r) \sim \lambda' u(p) + (1 - \lambda')u(r) \sim \lambda' u(q) + (1 - \lambda')u(r) \succ u(p).$$

To formulate our independence property, it is best to set the following definition.

Definition 5.1. We shall say that p overrides q and write $p \gg q$ iff $p > q$ and for any q' such that $q > q'$ and for any $\lambda \in]0, 1[$, $\lambda q + (1 - \lambda)p \sim \lambda q' + (1 - \lambda)p$.

The intuitive meaning of $p \gg q$ is that p is so much preferred to q that, in any lottery in which p and q are the prizes, if one does not win p, one does not even care to cash q, but would as well get any lesser prize q'. Notice that, since $>$ is asymmetric, the relation $\gg$ is also asymmetric and therefore irreflexive. The following lemma describes the tight relationship between the properties $p \gg q$ and $u(p)/u(q)$ is infinite.

LEMMA 5.1. *If the relation* $>$ *is defined by a linear functional as in Equation 2, then* $p \gg q$ *iff either* $u(p)/u(q)$ *is infinite, or* $p > q$ *and* $q \leq r$ *for any* $r \in P$*.*

Proof. Let us proof the *if* part. If $q \leq r$ for any $r \in P$ and $p > q$, then obviously $p \gg q$. If $u(p)/u(q)$ is infinite, then $u(p) \succ u(q)$ and, if $u(q) > u(q')$,

$$\lambda u(q) + (1 - \lambda)u(p) \sim (1 - \lambda)u(p) \sim \lambda u(q') + (1 - \lambda)u(p).$$

Let us prove now the *only if* part. If $u(q) \succ u(q')$ and

$$\lambda u(q) + (1 - \lambda)u(p) \sim \lambda u(q') + (1 - \lambda)u(p),$$

then $\lambda(u(q) - u(q'))/(\lambda u(q) + (1 - \lambda)u(p)$ is infinitesimally close to zero and therefore

$$\frac{u(q)}{u(q) - u(q')} + \frac{1 - \lambda}{\lambda} \frac{u(p)}{u(q) - u(q')}$$

is infinite. But $u(q) \succ u(q')$ implies the first term is not infinite. Therefore the second one is. Then $u(p)/(u(q) - u(q'))$ is infinite. But $u(q) - u(q') > \gamma u(q)$, for some *standard* positive γ, $u(p)/(u(q) - u(q')) < u(p)/\gamma u(q)$ and $u(p)/u(q)$ is infinite. ▮

Our independence property may now be formulated as:

$$\mathbf{A'2} \quad p > q, r \not\gg p \ \Rightarrow \ \forall \lambda \in]0, 1[\ \lambda p + (1 - \lambda)r > \lambda q + (1 - \lambda)r.$$

The intuitive meaning of **A'2** is that any lottery is sensitive to both its prizes, unless one of the prizes overrides the other one.

Let us study now property **A3**. Let us show it does not hold either. Let ϵ be positive and infinitesimally close to zero. Let $u(p) = 1$, $u(q) = 2\epsilon$, $u(r) = \epsilon$. We have $p > q > r$, but for any $\beta \in]0, 1[$, $u(q) \leq \beta u(p) + (1 - \beta)u(r)$.

Let us find a sound weakening of **A3**. Assume $p > q > r$. We have $u(p) \succ u(q) \succ u(r)$. Consider the values $\alpha u(p) + (1 - \alpha)u(r)$ for the different values of $\alpha \in]0, 1[$. Since $u(p) \succ u(q)$, there is some standard $\gamma > 0$ such that $(u(p) - u(q))/u(p) > \gamma$. We have $(1 - \gamma)u(p) > u(q)$. Let $\alpha = 1 - \gamma$, $\alpha \in]0, 1[$. We have

$$\alpha u(p) + (1 - \alpha)u(r) = (1 - \gamma)u(p) + \gamma u(r) > u(q) + \gamma u(r) > u(q).$$

Take any $\alpha' > \alpha$, $\alpha' < 1$. We have $\alpha' u(p) + (1 - \alpha')u(r) \succ u(q)$. We have shown that $\alpha' p + (1 - \alpha')r > q$.

Let us now consider the existence of a $\beta \in]0,1[$ such that $q > \beta p + (1 - \beta)r$. We have $u(q)/u(p) \succ u(r)/u(p)$. If $u(q)/u(p)$ is not infinitesimal, there is some standard $\gamma < 1$ such that $u(q)/u(p) \succ \gamma \succ u(r)/u(p)$. If $\beta > 0$ approaches zero, $(\gamma - \beta)/(1 - \beta)$ approaches γ from below. There is therefore some standard $\beta \in]0,1[$ such that $\gamma \succ (\gamma - \beta)/(1 - \beta) \succ u(r)/u(p)$. Therefore, we have $(\gamma - \beta)u(p) \succ (1 - \beta)u(r)$, and $\gamma u(p) \succ \beta) u(p) + (1 - \beta)u(r)$. But $u(q) \succ \gamma u(p)$. We conclude that $q > \beta p + (1 - \beta)r$. If, on the contrary, $u(q)/u(p)$ is infinitesimal, for any $\beta \in]0,1[$, $\beta u(p) \succ u(q)$ and $\beta p + (1 - \beta)r > q$. In this case, note that, for any $\beta \in]0,1[$,

$$\beta u(p) + (1 - \beta)u(r) \sim \beta u(p) \sim \beta u(p) + (1 - \beta)u(q).$$

We conclude that the following properties hold:

$$\mathbf{A'3} \quad p > q > r \;\Rightarrow\; \exists \alpha \in]0,1[\text{ such that } \alpha p + (1 - \alpha)r > q.$$

$$\mathbf{A''3} \quad p > q > r, p \not\gg q \;\Rightarrow\; \exists \beta \in]0,1[\text{ such that } q > \beta p + (1 - \beta)r.$$

6. REPRESENTATION RESULT

All the technical work needed to prove our theorem may be found in Appendix B. The general structure of the proof is very similar to that of the classical proof of the completeness of Jensen's postulates for representation by a standard linear functional. Nevertheless, there are a few subtle points, since one has to carefully avoid using those properties that hold in the standard case but do not hold here. The main result of this work may now be presented.

THEOREM 6.1. *If the set P is finitely generated, then the relation $>$ satisfies* **A1**, **A'2**, **A'3** *and* **A"3** *iff there is some nonstandard model $\mathcal{R}$ and some linear functional $u : P \longrightarrow \mathcal{R}_+$ such that $\forall p, q \in P$, $p > q$ iff $u(p) \succ u(q)$.*

The proof of Theorem 6.1 is contained in Appendix B. One may certainly weaken the restriction to a finitely generated P a bit, but some restriction is needed here. The following example will show that Theorem 6.1 does not hold if one removes the condition that P be finitely generated. Consider, for example, any nonstandard model of the real numbers and let P be the subset including all positive numbers that are neither infinite nor infinitesimally close to zero. Let us consider the usual ordering $>$ (*not* $\succ$) on P.

This ordering is easily seen to satisfy vNM's postulates. Suppose there is a nonstandard model $\mathcal{R}$ and a linear function $u : P \longrightarrow \mathcal{R}_+$ such that $\forall p, q \in P$, $p > q$ iff $u(p) \succ u(q)$. Since postulate **A2** (independence) holds, the image of u has to be contained in one single equivalence class under $\asymp$. All such equivalence classes are isomorphic (by multiplication). We may therefore assume that the image of u is contained in the class of numbers that are neither infinite nor infinitesimally close to zero. Any member of this class is $\sim$-equivalent to a unique standard real. We may therefore take u into the standard real numbers. But this is impossible if the nonstandard model chosen has a cardinality that is larger than the continuum. In this case, P has the same cardinality and, since u must be one-to-one, its image has a cardinality larger than that of $\mathbf{R}$.

7. COMPARISON WITH PREVIOUS WORK

Numerous works during the fifties and the sixties considered weakenings of vNM's postulates, see [11] for a survey. Some of these weakenings try to describe the actual preferences of human deciders and reject the Independence property **A2**. Often they also reject the requirement that preferences be transitive, or that indifference be transitive. Others, based on lexicographic orderings, keep the Independence postulate and reject only the Continuity postulate **A3**: [12, 6, 7, 8, 9, 3, 17, 18, 24]. Nonstandard analysis [25] appeared late on the scene. Few authors applied it to decision theory: [29, 15].

This work proposes an original weakening based on nonstandard analysis. We shall describe its relation to lexicographic orderings first, and then to other uses of nonstandard analysis in decision theory.

The motivation for this work is, in fact, opposite to the basic motivation for considering lexicographic orderings. The motivation for the consideration of lexicographic orderings on probabilities is the consideration of those perfection refinements of Nash Equilibria which required optimization also off the equilibrium path, on zero-probability events. Replacing the zero probabilities by infinitesimal probabilities and requiring optimization in the usual sense (i.e. with the usual linear ordering on (non-standard) numbers) gives the desired result. The current paper is trying to achieve exactly the opposite - "side issues" should not matter, and hence the special ordering proposed.

For a more technical comparison, notice that our postulates are very close to the original postulates of von Neumann and Morgenstern. In particular the ordering $<$ is modular (weak total) and the indifference relation $\sim$ is transitive. But both **A2** and **A3** are weakened, in a closely linked manner. Notice that lexicographic orderings define a preference relation that satisfies **A2**, and that **A2**, **A'3** and **A"3** together imply **A3**, since **A2** says that $p \gg q$ implies that for any w, $w \geq q$. The weakening offered by lexicographic orderings is therefore orthogonal to ours: any lexicographic ordering that satisfies our postulates satisfies the full set of vNM's postulates.

To explain better the difference between lexicographic orderings and our qualitative ordering, assume P is the real plane $\mathbf{R}^2$ ordered by the lexicographic ordering:

$$(x, y) < (x', y') \text{ iff either } x < x' \text{ or } x = x' \text{ and } y < y'.$$

For $\lambda \in [0,1]$, define $\lambda(x,y) + (1-\lambda)(x',y')$ to be $(\lambda x + (1-\lambda)x', \lambda y + (1-\lambda)y')$. Notice that $(0,0) < (1,10) < (2,0)$, but there is no $\lambda \in]0,1[$ such that $(1,10) = \lambda(0,0) + (1-\lambda)(2,0)$ since $\lambda(0,0) + (1-\lambda)(2,0) = (2(1-\lambda), 0)$. Both lexicographic and qualitative orderings imply the failure of property **P** of section 1. But notice that lexicographic ordering also implies the failure of Lemma B.4 since there exists some $\beta \in]0,1[$ such that:

$$(1,10) > \beta(2,0) + (1-\beta)(0,0) = \beta(2,0),$$

for example $\beta = 0.4$, and nevertheless there is no λ as above. Lexicographic and qualitative orderings stem from different concerns and have very different characteristics.

The previous uses of nonstandard numbers mentioned above consider the usual ordering on those nonstandard models, *not* the qualitative ordering we consider. Most of the considerations of [29] concerning strategic games in which some pay-offs are infinite are valid in the present framework too, and therefore very relevant for further work.

8. SUBJECTIVE PROBABILITY

In [1], Anscombe and Aumann showed that, in a slightly modified framework in which the decider compares not lotteries but acts the outcomes of which depend on the state of nature, an additional very natural postulate is sufficient to ensure the existence of subjective probabilities, in the decider's mind. This result is described in Section 8.1. In 8.2 and 8.3, it is shown that the same holds, with minimal changes, for Expected Qualitative Utility Maximization. This result must be contradistincted with LaValle and Fishburn's [17], showing that, in the Lexicographic Subjective Expected Utility framework, subjective probabilities have to be matrices rather than numbers.

8.1. Anscombe-Aumann's framework

Let P be a convex set of probability measures on $\mathcal{H}$, as in Section 2. In [1], Anscombe and Aumann consider a finite set S (of states) and the set **F** (of acts) of mappings: $S \longrightarrow P$. The set **F** is a mixture set when, for $a, b \in \mathbf{F}$, $\lambda \in [0,1]$, one defines $[\lambda a + (1-\lambda)b](s)$ as $\lambda a(s) + (1-\lambda)b(s)$. If f is an element of P we shall identify it with the constant function $h \in \mathbf{F}$ defined by $h(s) = f$ for any $s \in S$. They consider binary relations ($>$) on **F** that may be defined a probability measure p on S and a linear function $u : P \longrightarrow \mathbf{R}$ by:

$$a > b \Leftrightarrow \sum_{s \in S} p(s)u(a(s)) > \sum_{s \in S} p(s)u(b(s)).$$

Assuming P is finitely generated (in fact they have a slightly weaker condition), they show that a relation $>$ on **F** may be defined by a probability measure p on S and a linear u in the way above iff $>$ satisfies **A1**, **A2**, **A3** (where $>$ is defined on **F**) and one additional postulate. For any $a, b \in \mathbf{F}$

$\mathbf{A'4}$ If $\forall s \in S, s \neq s_0 \Rightarrow a(s) = b(s)$, then $a > b \;\Rightarrow\; a(s_0) > b(s_0)$.

In the last part of **A'4**, $a(s_0)$ and $b(s_0)$ stand for the corresponding constant functions.

8.2. Additional Postulate

We shall need an additional postulate, specific to our consideration of nonstandard utilities. A definition will be handy. Roughly, an element t of S is null iff the singleton $\{t\}$ is Savage-null, i.e., if the value taken by an act on t is insignificant.

Definition 8.1. Let $t \in S$. The element t is said to be *null* iff for any $a, b \in \mathbf{F}$ we have $a \sim b$ if a and b agree everywhere except possibly on t, i.e., for any $s \neq t$, $a(s) = b(s)$.

Our last postulate ensures that probabilities are *standard*. Assume $t \in S$, $a \in \mathbf{F}$ and one of the possible values of a, $a(t)$, overrides a, i.e., $a(t) \gg a$. It must be the case that t is null.

$\mathbf{A'5}$ $\quad t \in S,\ a \in \mathbf{F},\ a(t) \gg a \;\Rightarrow\; t$ is null.

A first consequence of **A'4** shall be proved now.

Lemma 8.1. *Assume* **A1** *and* **A'4** *hold. If* $a(s) \preceq b(s)$ *for any* $s \in S$, *then* $a \preceq b$.

Proof. By induction on the number of elements s of S for which $a(s) \neq b(s)$, using **A'4**. ▌

8.3. Characterization Theorem

A theorem similar to Anscombe and Aumann's holds for Expected Qualitative Utility Maximization.

THEOREM 8.1. *Assume P is finitely generated and $>$ is a binary relation on* $\mathbf{F}$*. There is a probability measure p on S, a nonstandard model $\mathcal{R}$ and a linear function $u : P \longrightarrow \mathcal{R}_+$ such that $\forall a, b \in \mathbf{F}$,*

$$a > b \Leftrightarrow \sum_{s \in S} p(s)u(a(s)) \succ \sum_{s \in S} p(s)u(b(s)) \tag{3}$$

iff the relation $>$ satisfies **A1**, **A'2**, **A'3**, **A"3**, **A'4** *and* **A'5**. *If the relation $>$ on P is not trivial, the probability measure p is unique.*

The proof may be found in Appendix C. Its general structure is similar to Anscombe and Aumann's, but there are a number of steps that have to be managed in a different manner.

9. CONCLUSION

Nonstandard models of the real numbers provide for a natural notion of qualitative equivalence and a principle of Qualitative Utility Maximization. This work characterizes in full the situation in which one allows nonstandard utilities but insists on standard probabilities. In this framework one may consider consequences that are infinitely preferable to others and criteria of the Maximin family. Subjective probabilities may be shown to exist as in the classical framework. They are numbers, not matrices, as in the Lexicographic Subjective Expected Utility framework. This represents probably a clear advantage for Qualitative Utility Maximization over lexicographic methods. But both are really orthogonal extensions of the classical Expected Utility Maximization paradigm.

The study of games with nonstandard utilities seems appealing. The dual case of nonstandard probabilities and standard utilities and the most general case of nonstandard probabilities and utilities are yet to be characterized. They will include consideration of subjective probabilities infinitesimally close to zero.

APPENDIX A
Properties of the Qualitative Ordering

The following lemma summarizes the main properties of the relations just defined. Let $\mathcal{R}_+$ be the subset of $\mathcal{R}$ consisting of all positive numbers (including zero).

LEMMA A.1. *For any $x, y, z \in \mathcal{R}_+$*

1. $x \succ y \Rightarrow x > y$,

2. $x > 0 \Rightarrow x \succ 0$,

3.the relation $\succ$ *is asymmetric: i.e.,* $x \succ y \Rightarrow y \preceq x$,

4. $x \succ y$ *and* $y \geq z$ *imply* $x \succ z$,

5. $x \geq y$ *and* $y \succ z$ *imply* $x \succ z$,

6.the relation $\succ$ *is transitive,*

7.the relation $\succ$ *is negatively transitive, i.e., modular: i.e.,* $x \preceq y$ *and* $y \preceq z$ *imply* $x \preceq z$,

8.the relation $\sim$ *is an equivalence relation,*

9.if $c > 0$, *then* $x \succ y$ *iff* $cx \succ cy$,

10.if $x \succ y$, $c > 0$ *and* c/x *is finite, then* $x + c \succ y + c$.

11.if $x \succ y$, $c > 0$ *and* c/x *is infinite, then, for any* standard *positive* λ, μ, ν, $\lambda c \sim \mu x + \lambda c \sim \mu y + \lambda c \succ \nu x$,

12.if $x \succ y$, $c > 0$ *and* $y - c > 0$, *then* $x - c \succ y - c$.

Proof. Parts 1. and 2. are obvious from the definition of $\succ$.

Part 3. follows from part 1..

For part 4., notice that $x - z \geq x - y$ and $(x - z)/x \geq (x - y)/x$.

For part 5., note that $x - z \geq x - y$ and $(x - z)/x \geq (x - y)/x$.

Part 6. follows from parts 1., 4. and 5..

For part 7., assume $x \succ y$. We shall show that either $x \succ z$ or $z \succ y$. If $y \leq z$ or $z \leq x$ we conclude by parts 4. or 5.. We may therefore assume that $x > z > y$. Notice that

$$(x - y)/x = (x - z)/x + (z - y)/x < (x - z)/x + (z - y)/z.$$

Since the left hand side is not infinitesimally close to zero, at least one of the terms in the sum in the right hand side must be larger than some standard number and we have shown that the relation $\succ$ is negatively transitive. It follows that the relation $\preceq$ is transitive. It is reflexive by part 1.. The relation $\sim$ is therefore reflexive, transitive and symmetric, and part 8. is proved.

Part 9. is obvious.

For part 10., assume $x \succ y$. We have $x > y$. Therefore $x + c > y + c$. Consider

$$\frac{(x + c) - (y + c)}{x + c} = \frac{x - y}{x + c} = \frac{x - y}{x} \times \frac{x}{x + c}.$$

If c/x is finite, so is $(x + c)/x$ and $x/(x + c)$ is not infinitesimally close to zero. But $x \succ y$ implies that $(x - y)/x$ is not infinitesimally close to zero. We conclude that $x + c \succ y + c$.

For part 11., assume c/x is infinite. Both $\mu x/\lambda c$ and $\mu y/\lambda c$ are infinitesimally close to zero. Therefore $\lambda c \sim \mu x + \lambda c \sim \mu y + \lambda c$ and $\lambda c \succ \nu x$.

For part 12., notice that $x - c > y - c$ and $[(x-c)-(y-c)]/(x-c) > (x-y)/x$. ▮

So far, the reader may get the feeling that the qualitative order is just like the usual one. This would be a mistaken impression. Note that $x \preceq y$ does not imply $x - c \preceq y - c$. If ϵ is positive and infinitesimally close to zero, $1 + \epsilon \preceq 1$ but $1 + \epsilon - (1 - \epsilon^2) \not\preceq \epsilon^2$. Note also that $\succ$ is not complete and that $\preceq$ is not antisymmetric. A sequence of technical lemmas will be stated now and used in Section 8.3. Had we used the usual order $>$, they would have been obvious, for $\prec$ they are not.

Lemma A.2. *Let $\lambda \in]0, 1[$ be a standard real number. Assume $a, b, c \in \mathcal{R}_+$ such that $a = \lambda b + (1 - \lambda)c$, $a \preceq b$, $a \preceq c$. Then, $a \sim b$ and $a \sim c$.*

Proof. Suppose, without loss of generality, that $b < c$. Then $b < a < c$. Clearly $b \preceq a$ and $a \sim b$. But $c - a = \frac{\lambda}{1-\lambda}(a - b)$. If $\frac{a-b}{a}$ is infinitesimally close to zero, so is $\frac{c-a}{c}$ and $a \sim c$. ▮

Lemma A.3. *Let $n > 0$, $\lambda_i > 0$, $i = 0, \ldots, n-1$, standard real numbers such that $\sum_{i=0}^{n-1} \lambda_i = 1$. Let a, a_i, $i = 0, \ldots, n-1$ be elements of a nonstandard model $\mathcal{R}_+$ such that $a = \sum_{i=0}^{n-1} \lambda_i a_i$ and $a \preceq a_i$ for any $i = 0, \ldots, n-1$. Then, for any $i = 0, \ldots, n-1$, $a \sim a_i$.*

Proof. By induction on n. The case $n = 1$ is obvious. The case $n = 2$ is dealt with by Lemma A.2. For the general case, consider that

$$a = \lambda_0 a_0 + (1 - \lambda_0) \sum_{i=1}^{n-1} \frac{\lambda_i}{1 - \lambda_0} a_i,$$

and use Lemma A.2 to conclude that $a \sim a_0$. Similarly for a_i. ▮

Lemma A.4. *Let $n > 0$, $\lambda_i \geq 0$, $i = 0, \ldots, n-1$, standard real numbers such that $\sum_{i=0}^{n-1} \lambda_i = 1$. Let a, a_i, $i = 0, \ldots, n-1$ be elements of a nonstandard model $\mathcal{R}_+$ such that $a = \sum_{i=0}^{n-1} \lambda_i a_i$ and $a \preceq a_i$ for any $i = 0, \ldots, n-1$. Then, for any $i = 0, \ldots, n-1$ such that $\lambda_i > 0$, $a \sim a_i$.*

Proof. By induction on n. The case $n = 1$ is obvious. For the general case, if, for some i $\lambda_i = 0$, use a straightforward induction argument. If, for any i, $\lambda_i > 0$, use Lemma A.3. ▮

APPENDIX B
Derivation of the completeness result

We assume that the postulates **A1**, **A'2**, **A'3** and **A"3** hold and derive some consequences. In most cases, the derivation is very similar to the classical case, but we have to do with the weaker assumptions. First, notice that, since $p \gg q$ implies $p > q$ and, by **A1**, the relation $>$ is asymmetric, so is $\gg$. In particular, $\gg$ is irreflexive.

Lemma B.1. *Assume $p > q$, $\lambda \in]0, 1[$. Then, $p > \lambda p + (1 - \lambda)q > q$.*

Proof. Assume $p > q$. We know that $p \not\gg p$, therefore, by **A'2**, for any $\mu \in]0, 1[$, $\mu p + (1 - \mu)p > \mu q + (1 - \mu)p$. By taking, $\mu = 1 - \lambda$, for any $\lambda \in]0, 1[$, $p > (1 - \lambda)q + \lambda p$. But we know also that $q \not\gg p$ and therefore we have $\lambda p + (1 - \lambda)q > \lambda q + (1 - \lambda)q$. ∎

The next lemma is most important.

Lemma B.2. *Assume $p > q$, $\lambda, \mu \in [0, 1]$, and $\lambda > \mu$. Then, $\lambda p + (1 - \lambda)q > \mu p + (1 - \mu)q$.*

Proof. Note that $\mu < 1$.

$$\lambda p + (1 - \lambda)q = \frac{\lambda - \mu}{1 - \mu} p + \frac{1 - \lambda}{1 - \mu}[\mu p + (1 - \mu)q]$$

By Lemma B.1, $p > \mu p + (1 - \mu)q$. Therefore $\mu p + (1 - \mu)q \not\gg p$ and, by **A'2**,

$$\frac{\lambda - \mu}{1 - \mu} p + \frac{1 - \lambda}{1 - \mu}[\mu p + (1 - \mu)q] >$$

$$\frac{\lambda - \mu}{1 - \mu}[\mu p + (1 - \mu)q] + \frac{1 - \lambda}{1 - \mu}[\mu p + (1 - \mu)q] = \mu p + (1 - \mu)q.$$

∎

Corollary B.1. *Let $\lambda, \mu \in [0, 1]$. If $\lambda p + (1 - \lambda)q \sim \mu p + (1 - \mu)q$, then either $\lambda = \mu$ or $p \sim q$.*

Proof. Let $\lambda p + (1 - \lambda)q \sim \mu p + (1 - \mu)q$. Assume, without loss of generality, that $p > q$. By lemma B.2, if $\lambda > \mu$, we have $\lambda p + (1 - \lambda)q > \mu p + (1 - \mu)q$. We conclude that $\mu \leq \lambda$, and, similarly, $\lambda \leq \mu$. ∎

Corollary B.2. *For any $p, q, r \in P$, if $p > q$, there is at most one $\lambda \in [0, 1]$ such that $r \sim \lambda p + (1 - \lambda)q$.*

Proof. Obvious from Corollary B.1. ▮

LEMMA B.3. *If $p \gg q$ and $q > r$, then $p \gg r$.*

Proof. Let $p \gg q$ and $q > r$. Let r' be such that $r > r'$. We shall show that, for any $\lambda \in]0,1[$, $\lambda r + (1-\lambda)p \sim \lambda r' + (1-\lambda)p$. We have, $q > r$, $q > r'$ and $p \gg q$. Therefore $\lambda q + (1-\lambda)p \sim \lambda r + (1-\lambda)p$ and $\lambda q + (1-\lambda)p \sim \lambda r' + (1-\lambda)p$. ▮

LEMMA B.4. *Let $p > q > r$. If one assumes that there exists some $\beta \in]0,1[$ such that $q > \beta p + (1-\beta)r$, then there exists some $\gamma \in]0,1[$ such that $q \sim \gamma p + (1-\gamma)r$.*

Proof. By **A'3**, there is some $\alpha \in]0,1[$ such that $\alpha p + (1-\alpha)r > q$. Let γ_1 be the infimum of the set $\{\alpha \mid \alpha p + (1-\alpha)r > q\}$ and γ_2 be the supremum of the set $\{\beta \mid q > \beta p + (1-\beta)r\}$ (by assumption this set is non-empty). By Lemma B.2, $\gamma_1 \geq \gamma_2$. Take any γ such that $\gamma_1 \geq \gamma \geq \gamma_2$. The choice of γ_1 and γ_2, together with **A1** imply that $q \sim \lambda p + (1-\lambda)r$. ▮

COROLLARY B.3. *Let $p \geq q \geq r$, $p > r$ and $p \not\gg r$. Then, there exists a unique $\lambda \in [0,1]$ such that $q \sim \lambda p + (1-\lambda)r$.*

Proof. Uniqueness has been shown in Corollary B.2. Let us show existence. If $p \sim q$ or $q \sim r$, one can take λ to be equal to one or zero respectively. We may therefore assume that $p > q > r$ and $p \not\gg r$. By Lemma B.3, $p \not\gg q$. By **A"3** the assumptions of Lemma B.4 hold and the existence of a suitable λ is ensured. ▮

LEMMA B.5. *If $p > p'$ and $q > q'$, then, for any $\lambda \in [0,1]$, $\lambda p + (1-\lambda)q > \lambda p' + (1-\lambda)q'$.*

Proof. The cases $\lambda = 0$ and $\lambda = 1$ are easily treated. Assume $\lambda \in]0,1[$. By **A'2**, either,

1. for any $\lambda \in]0,1[$, $\lambda p + (1-\lambda)q > \lambda p' + (1-\lambda)q$, or
2. $q \gg p$. Similarly, either
3. for any $\lambda \in]0,1[$, $\lambda p' + (1-\lambda)q > \lambda p' + (1-\lambda)q'$, or
4. $p' \gg q$.

The only non-trivial case is when both cases 2. and 4. occur. But then, we have both $q > p$ and $p' > q$, contradicting $p > p'$. ▮

LEMMA B.6. *If $p \sim q$, $\lambda \in [0,1]$, then $p \sim \lambda p + (1-\lambda)q$.*

Proof. Suppose $p \sim q > \lambda p + (1-\lambda)q$. Then, by Lemma B.5,

$$\lambda p + (1-\lambda)q > \lambda[\lambda p + (1-\lambda)q] + (1-\lambda)[\lambda p + (1-\lambda)q] = \lambda p + (1-\lambda)q,$$

a contradiction. Similarly $\lambda p + (1-\lambda)q > p \sim q$ leads to a contradiction. ▌

LEMMA B.7. *Let $\lambda \in]0,1[$. For any p, q, $p \not\gg \lambda q + (1-\lambda)p$.*

Proof. Assume $p \gg \lambda q + (1-\lambda)p$. We have $p > \lambda q + (1-\lambda)p$. By Lemmas B.1 and B.6, $p > q$ and $\lambda q + (1-\lambda)p > q$. Since $p \gg \lambda q + (1-\lambda)p$, we know that $\alpha[\lambda q + (1-\lambda)p] + (1-\alpha)p \sim \alpha q + (1-\alpha)p$, for any $\alpha \in]0,1[$. Therefore, for any such α, $\alpha\lambda q + (1-\alpha\lambda)p \sim \alpha q + (1-\alpha)p$. By Corollary B.1, $q \sim p$, a contradiction. ▌

The next lemma does not mention the relation $\gg$, but it does not hold in the wider framework in which negative utilities are allowed: $1 + \epsilon \sim 1$, but $0.5(1+\epsilon) + 0.5(-1) > 0.5(1) + 0.5(-1)$. Its proof is more intricate than the corresponding proof in the classical framework, or any framework in which Independence holds.

LEMMA B.8. *Let $\lambda \in [0,1]$. If $p \sim q$, then $\lambda p + (1-\lambda)r \sim \lambda q + (1-\lambda)r$.*

Proof. Assume $p \sim q$. The result is obvious for λ equal to one or zero. We may assume $\lambda \in]0,1[$. If $r \sim p$, then, by Lemma B.6, $r \sim \lambda p + (1-\lambda)r$ and, since $r \sim q$ $r \sim$ $\lambda q + (1-\lambda)r$, and the result is proved. Assume therefore that $p \not\sim r$. Without loss of generality, we may assume that $\lambda q + (1-\lambda)r > \lambda p + (1-\lambda)r$, We shall derive a contradiction. Assume, first, that $p > r$. Since $p \sim q$ we have, $q > r$ and, by Lemma B.1, $p \sim q > \alpha q + (1-\alpha)r > r$, for any $\alpha \in]0,1[$. Since $r \not\gg p$, we have, by **A'2**, for any $\alpha \in]0,1[$,

$$\lambda p + (1-\lambda)r > \lambda[\alpha q + (1-\alpha)r] + (1-\lambda)r.$$

But

$$\lambda[\alpha q + (1-\alpha)r] + (1-\lambda)r = \lambda\alpha q + (1-\lambda\alpha)r =$$

$$\alpha[\lambda q + (1-\lambda)r] + (1-\alpha)r.$$

We see that $\lambda q + (1-\lambda)r > \lambda p + (1-\lambda)r > r$. But

$$\lambda p + (1-\lambda)r > \alpha[\lambda q + (1-\lambda)r] + (1-\alpha)r,$$

for any $\alpha \in]0,1[$, in contradiction with **A'3**.

Let us assume now that $r > p$. By Lemma B.1, $r > \alpha p + (1-\alpha) r > p \sim q$, for any $\alpha \in]0,1[$. By Lemma B.7 $r \not\gg \alpha p + (1-\alpha) r$, and therefore, by **A'2**, we have $\lambda[\alpha p + (1-\alpha) r] + (1-\lambda) r > \lambda q + (1-\lambda) r$, for any $\alpha \in]0,1[$. But we have $r > \lambda q + (1-\lambda) r > \lambda p + (1-\lambda) r$ and,

$$\lambda[\alpha p + (1-\alpha) r] + (1-\lambda) r = \lambda \alpha p + (1-\lambda\alpha) r = \tag{B.1}$$

$$\alpha[\lambda p + (1-\lambda) r] + (1-\alpha) r.$$

We see that, for any $\alpha \in]0,1[$, $\alpha[\lambda p + (1-\lambda) r] + (1-\alpha) r > \lambda q + (1-\lambda) r$. By **A"3**, $r \gg \lambda q + (1-\lambda) r$, contradicting Lemma B.7. ▮

Our goal is now to show that the relation $\gg$ is a weak order.

LEMMA B.9. *If $p \gg q$ and $q \geq r$, then $p \gg r$.*

Proof. If $q > r$, the result follows from Lemma B.3. If $q \sim r$, let r' be such that $r > r'$. We shall show that, for any $\lambda \in]0,1[$, $\lambda r + (1-\lambda) p \sim \lambda r' + (1-\lambda) p$. Indeed, by Lemma B.8, $\lambda r + (1-\lambda) p \sim \lambda q + (1-\lambda) p$, and, since $q > r'$ and $p \gg q$, $\lambda r' + (1-\lambda) p \sim \lambda q + (1-\lambda) p$. ▮

LEMMA B.10. *If $p \geq q$ and $q \gg r$, then $p \gg r$.*

Proof. Let $p \geq q$ and $q \gg r$. Let r' be such that $r > r'$. We shall show that, for any $\lambda \in]0,1[$, $(1-\lambda) r + \lambda p \sim (1-\lambda) r' + \lambda p$. We know that $(1-\lambda) r + \lambda q \sim (1-\lambda) r' + \lambda q$. If $p \sim q$, Lemma B.8 yields the desired conclusion. We may therefore assume that $p > q > r$. We may assume $p \not\gg r$. By Lemma B.3, there is some $\alpha \in]0,1[$ such that $q \sim \alpha p + (1-\alpha) r$. Therefore, by Lemma B.8, $(1-\lambda) r + \lambda q \sim (1-\lambda) r + \lambda[\alpha p + (1-\alpha) r] = (1-\lambda\alpha) r + \lambda\alpha p$. Similarly, $(1-\lambda) r' + \lambda q \sim (1-\lambda\alpha) r' + \lambda\alpha p$. We have $r > r'$ and $(1-\lambda\alpha) r + \lambda\alpha p \sim (1-\lambda\alpha) r' + \lambda\alpha p \sim$. **A'2** implies $p \gg r$. ▮

LEMMA B.11. *The relation $\gg$ is negatively transitive.*

Proof. Assume $p \not\gg q$ and $q \not\gg r$. We shall show that $p \not\gg r$. If $r \geq p$, the result holds. We may therefore assume that $p > r$. If $q \geq p$, $q \not\gg r$ implies $p \not\gg r$, by Lemma B.10. We may therefore assume that $p > q$. If $r \geq q$, $p \not\gg q$ implies $p \not\gg r$, by Lemma B.9. We may therefore assume that $q > r$. We have $p > q > r$.

$q \not\gg r$ implies there exists some r', $r > r'$ and some $\lambda \in]0,1[$, such that $\lambda r + (1-\lambda) q > \lambda r' + (1-\lambda) q$. By **A'2**, since $q \not\gg r$, for any $\alpha \in]0,1[$, we have $(1-\alpha) r + \alpha q > (1-\alpha) r' + \alpha q$.

By **A"3**, since $p \not\gg q$, there is some $\gamma \in]0,1[$ such that $q > \gamma p + (1-\gamma)r$. By Lemma B.4, there is some $\beta \in]0,1[$ such that $q \sim \beta p + (1-\beta)r$. By Lemma B.8,

$$(1-\alpha)r + \alpha[\beta p + (1-\beta)r] > (1-\alpha)r' + \alpha[\beta p + (1-\beta)r]$$

and

$$\alpha\beta p + (1-\alpha\beta)r > \alpha\beta p + (1-\alpha\beta)r'$$

and we conclude that $p \not\gg r$. ▮

Let $p \asymp q$ denote that $p \not\gg q$ and $q \not\gg p$.

Corollary B.4. *The relation $\gg$ is a weak order and therefore the relation $\asymp$ is an equivalence relation.*

Proof. We notice that $\gg$ is asymmetric, by **A1**. Lemma B.11 shows that it is negatively transitive. ▮

Lemma B.12. *Let $A \subseteq P$ be an equivalence class of P under $\asymp$. The relation $>$ on the set A satisfies* **A1**, **A2** *and* **A3**. *It may therefore be represented by a linear functional $u : A \longrightarrow \mathbf{R}$, the standard set of real numbers. If the set P is finitely generated, then, the functional u may be taken into the positive real numbers.*

Proof. Since $>$ is a weak order on P, it is a weak order on $A \subseteq P$. For any $p, q \in A$, $p \not\gg q$, therefore **A'2** implies **A2** and **A'3** and **A"3** imply **A3**. By Theorem 4.1 of [11], the relation $>$ on A may be represented by a linear functional on A. This functional is uniquely defined up to a positive affine transformation. If the set P is finitely generated, by Lemma B.2, the range of u is bounded and therefore the addition of a large enough constant to u ensures its range contains only positive numbers. ▮

Lemma B.13. *For any $p, q \in P$, $\lambda \in [0,1]$, one of the following two cases obtains:*

- $\lambda p + (1-\lambda)q \asymp p$, *or*
- $\lambda p + (1-\lambda)q \asymp q$.

Proof. If λ is equal to zero or to one, the result is obvious. Assume $\lambda \in]0,1[$. If $p \sim q$, Lemma B.6 implies $\lambda p + (1-\lambda)q \sim p$ and therefore $\lambda p + (1-\lambda)q \asymp p$. Without loss of generality, assume $p > q$. Then, by Lemma B.1, $p > \lambda p + (1-\lambda)q$ and therefore $\lambda p + (1-\lambda)q \not\gg p$. By Lemma B.7, $p \not\gg \lambda q + (1-\lambda)p$. ▮

The proof of Theorem 6.1 will now be presented. .1

THEOREM B.6. *If the set P is finitely generated, then the relation $>$ satisfies* **A1**, **A'2**, **A'3** *and* **A"3** *iff there is some nonstandard model $\mathcal{R}$ and some linear functional $u : P \longrightarrow \mathcal{R}_+$ such that $\forall p, q \in P$, $p > q$ iff $u(p) \succ u(q)$.*

Proof. The *if* part has been proven in Section 5. Let us deal with the *only if* part. By Lemma B.13 any element of P is $\asymp$-equivalent to one of the generators. Therefore there is a finite number of $\asymp$-equivalence classes. Let them be ordered by $\gg$: $A_0 \gg A_1 \gg \ldots A_{n-1}$. Let us choose any non-standard model $\mathcal{R}$and any $\epsilon \in \mathcal{R}$, positive and infinitesimally close to zero. By Lemma B.12, there are linear functions: $u_i : A_i \longrightarrow \mathbf{R}$ such that $\forall p, q \in A_i, p > q$ iff $u_i(p) > u_i(q)$. We shall define, for every $p \in P$: $u(p) \stackrel{\text{def}}{=} \epsilon^i u_i(p)$.

We shall show that, if $p > q$, then $u(p) \succ u(q)$. Suppose $p > q$. If $p \gg q$, $u(p)/u(q)$ is infinite and $u(p) \succ u(q)$. If $p \not\gg q$, then $p \asymp q$. Suppose $p, q \in A_i$. Since $u_i(p)$ and $u_i(q)$ are standard numbers such that $u_i(p) > u_i(q)$, we have $u_i(p) \succ u_i(q)$ and $\epsilon^i u_i(p) \succ \epsilon^i u_i(q)$.

We shall show now that, if $u(p) \succ u(q)$, then $p > q$. Assume $u(p) \succ u(q)$. If $p \asymp q$, then $\epsilon^i u_i(p) \succ \epsilon^i u_i(q)$ and therefore $u_i(p) > u_i(q)$ and $p > q$. If $p \not\asymp q$, then there are i, j, $i < j$ such that $p \in A_i$ and $q \in A_j$ and therefore $p \gg q$ and $p > q$. ▮

APPENDIX C
Proof of Theorem 8.1

THEOREM 8.1. Assume P is finitely generated and $>$ is a binary relation on $\mathbf{F}$. There is a probability measure p on S, a nonstandard model $\mathcal{R}$ and a linear function $u : P \longrightarrow \mathcal{R}_+$ such that $\forall a, b \in \mathbf{F}$,

$$a > b \Leftrightarrow \sum_{s \in S} p(s)u(a(s)) \succ \sum_{s \in S} p(s)u(b(s)) \tag{C.1}$$

iff the relation $>$ satisfies **A1**, **A'2**, **A'3**, **A"3**, **A'4** and **A'5**. If the relation $>$ on P is not trivial, the probability measure p is unique.

Proof. Let us prove first the *only if* part. Assume Equation C.1 holds. We notice that $\mathbf{F}$ is a mixture set and that, by letting $u'(a) \stackrel{\text{def}}{=} \sum_{s \in S} p(s)u(a(s))$, we have

$$a > b \Leftrightarrow u'(a) \succ u'(b).$$

By the results of Section 5, **A1**, **A'2**, **A'3**, and **A"3** hold.

Let us check that **A'4** holds. Assume that $a(s) = b(s)$ for any $s \in S$, $s \neq s_0$ and that $a > b$. If

$$\sum_{s \in S} p(s)u(a(s)) \succ \sum_{s \in S} p(s)u(b(s)),$$

then, by Lemma A.1, part 12., $p(s_0)u(a(s_0)) \succ p(s_0)u(b(s_0))$. By part 9., we have $u(a(s_0)) \succ u(b(s_0))$. But $\sum_{s \in S} p(s)u(a(s_0)) = u(a(s_0))$ and similarly $\sum_{s \in S} p(s)u(b(s_0)) = u(b(s_0))$. Therefore $\sum_{s \in S} p(s)u(a(s_0)) \succ \sum_{s \in S} p(s)u(b(s_0))$ and $a(s_0) > b(s_0)$.

Let us check now that **A'5** holds. By Lemma 5.1, $a(t) \gg a$ implies that either $u(a(t))/u(a)$ is infinite, or $a(t) > a$ and $a \leq b$ for any $b \in \mathbf{F}$. Since $u(a) \geq p(t)u(a(t))$, if $u(a(t))/u(a)$ is infinite, $p(t)$ must be equal to zero and t is null. If $a(t) > a$ and $a \leq b$ for any $b \in \mathbf{F}$, then, letting $x = \sum_{s \in S} p(s)u(a(s))$, we have $x \preceq u(a(s))$ for any $s \in S$. By Lemma A.4, $x \not\prec u(a(t))$ implies that $p(t) = 0$, and t is null.

Let us deal now with the more challenging *if* part. Assume **A1**, **A'2**, **A'3**, **A"3**, **A'4** and **A'5** are satisfied. By Theorem 6.1, since **A1**, **A'2**, **A'3** and **A"3** hold, there is a nonstandard model $\mathcal{R}$ and a linear function $u^\star : \mathbf{F} \longrightarrow \mathcal{R}_+$ such that $\forall a, b \in \mathbf{F}$,

$$a > b \Leftrightarrow u^\star(a) \succ u^\star(b). \tag{C.2}$$

Since P is finitely generated, by Lemma B.1, there are minimal (l) and maximal (h) elements of P. Both l and h are constant functions in $\mathbf{F}$. By Lemma 8.1, for any $a \in \mathbf{F}$, $l \preceq a \preceq h$. If $l \sim h$, all acts are equivalent and the Theorem is easily proved. Suppose, therefore, that $h \succ l$. By Lemma A.1, parts 9. and 12.,

$$a > b \Leftrightarrow \frac{u^\star(a) - u^\star(l)}{u^\star(h)} \succ \frac{u^\star(b) - u^\star(l)}{u^\star(h)}.$$

This means we may assume, from now on, that $u^\star(l) = 0$ and $u^\star(h) = 1$. For $t \in S$, let us define $c_t \in \mathbf{F}$ by $c_t(t) = h$ and, for $s \neq t$, $c_t(s) = l$. If t is not null, by **A'5**, $h \not\gg c_t$ and, by **A"3** and Lemma B.4 there is some $\lambda \in [0, 1]$ such that $c_t \sim \lambda h + (1 - \lambda)l$; in this case we shall define $p(t) = \lambda$ and notice that $u^\star(c_t) \sim p(t)$. If t is null we take $p(t) = 0$. In this case, notice that $c_t \sim l$ and therefore $u^\star(c_t) = 0$.

Clearly, for any $s \in S$, $0 \leq p(s) \leq 1$. We claim that, for any $a \in \mathbf{F}$,

$$u^\star(a) \sim \sum_{s \in S} p(s)\, u^\star(a(s)).$$

Notice that we do not claim equality, only *similarity*, i.e., $\sim$, between the left hand side and the right hand side. Let us justify our claim now. Let $a \in \mathbf{F}$ and $c = \sum_{s \in S} u^\star(a(s))$. If $c = 0$, then, for any $s \in S$, $u^\star(a(s)) = 0$ and $a(s) \sim l$, and by Lemma 8.1, $a \sim l$, $u^\star(a) = 0$, by Lemma A.1, part 2. and the result holds. If $c > 0$, let

$$b \stackrel{\text{def}}{=} \sum_{s \in S} \frac{u^\star(a(s))}{c}\, c_s.$$

Let us, first, compare a and b. For any $t \in S$,

$$b(t) = \frac{u^\star(a(t))}{c}\, h + (1 - \frac{u^\star(a(t))}{c})\, l.$$

We see that, for any $t \in S$,

$$u^\star(b(t)) = \frac{1}{c}\, u^\star(a(t)). \tag{C.3}$$

Let us assume, first, that $c \leq 1$. Define the act a' by:

$$a' \stackrel{\text{def}}{=} c\, b + (1 - c)\, l.$$

For any $s \in S$, $a'(s) = c\, b(s) + (1 - c)l$, $u^\star(a'(s)) = c\, u^\star(b(s))$ and therefore $u^\star(a'(s)) = u^\star(a(s))$. We conclude, by Lemma 8.1, that $a' \sim a$, and therefore $u^\star(a) \sim u^\star(a')$. But, by linearity of $u^\star$, $u^\star(a') = c\, u^\star(b)$. We conclude that $u^\star(a) \sim c\, u^\star(b)$. The case $c > 1$ is treated similarly by defining

$$b' \stackrel{\text{def}}{=} (1/c)\, a + (1 - 1/c)\, l.$$

But, now, by linearity of $u^\star$,

$$u^\star(b) = \frac{1}{c} \sum_{s \in S} u^\star(a(s))\, u^\star(c_s).$$

Therefore

$$u^\star(a) \sim \sum_{s \in S} u^\star(a(s))\, u^\star(c_s).$$

We have seen that $u^\star(c_s) \sim p(s)$ when $p(s) > 0$ and $u^\star(c_s) = p(s)$ when $p(s) = 0$. Therefore

$$u^\star(a) \sim \sum_{s \in S} p(s)\, u^\star(a(s)),$$

as claimed. The first conclusion we draw is that:

$$a > b \Leftrightarrow \sum_{s \in S} p(s)\ u^{\star}(a(s)) \succ \sum_{s \in S} p(s)\ u^{\star}(b(s)).$$

Our second conclusion is that $1 = u^{\star}(h) \sim \sum_{s \in S} p(s) u^{\star}(h) = \sum_{s \in S} p(s)$. Therefore $\sum_{s \in S} p(s) \sim 1$. But, for any $s \in S$, $p(s)$ is *standard*. It follows that $\sum_{s \in S} p(s) = 1$.

For the uniqueness claim, notice that $\sum_{s \in S} p(s) u(c_t(s)) = p(t)u(h) + (1 - p(t))u(l)$. By Equation C.1, then, $c_t \sim p(t)h + (1 - p(t))l$. By Corollary B.3, there is only one such $p(t)$. ▮

ACKNOWLEDGMENT

I want to thank Israel Aumann for his encouragements and suggestions to improve the presentation of this work and Aviad Heifetz who helped correct an argument. An anonymous referee provided unusually helpful insights, two of which have been directly incorporated into the paper, thanks a lot!

REFERENCES

1. Anscombe, F. J., and Aumann, R. J. (1963). A definition of subjective probability. *Annals of Mathematical Statistics*, 34:199–205.
2. Aumann R. J. (1977). The St. Petersburg paradox: A discussion of some recent comments. *Journal of Economic Theory*, 14(2):443–445.
3. Blume, L., Brandenburger, A, and Dekel E. (1991). Lexicographic probabilities and choice under uncertainty. *Econometrica*, 59(1):61–79.
4. Brafman R. I., and Tennenholtz, M. (1996). On the foundations of qualitative decision theory. In *Proceedings of the National Conference on AI (AAAI'96)*, pages 1291–1296.
5. Brito, D. L. (1975). Becker's theory of the allocation of time and the St. Petersburg paradox. *Journal of Economic Theory*, 10:123–126.
6. Chipman, J. S. (1960). The foundations of utility. *Econometrica*, 28:198–224.
7. Chipman, J. S. (1971). Non-archimedean behavior under risk: An elementary analysis - with applications to the theorey of assets. In J. S. Chipman, L. Hurwicz, M. K. Richter, and H. F. Sonnenschein, Eds., *Utility and Demand*. Harcourt Brace Jovanovich, New York.
8. P.C. Fishburn. A study of lexicographic expected utility. *Management Science*, 17:672–678, 1971.
9. P.C. Fishburn. Lexicographic orders, utilities, and decision rules: a survey. *Management Science*, 20:1442–1471, 1974.
10. P.C. Fishburn and G.A. Kochenberger. Two-piece von neumann-morgenstern utility functions. *Decision Sciences*, 10:503–518, 1979.
11. Peter C. Fishburn. Utility and subjective probability. In Robert J. Aumann and Sergiu Hart, editors, *Handbook of Game Theory with Economic Applications*, volume 2 of *Handbooks in Economics 11*, chapter 39, pages 1397–1435. North-Holland, 1994.

12. Melvin W. Hausner. Multidimensional utilities. In Robert M. Thrall, Clyde H. Coombs, and Robert L. Davis, editors, *Decision Processes.* John Wiley & sons, New York, 1954.

13. N. E. Jensen. An introduction to bernoullian utility theory, I: Utility functions. *Swedish Journal of Economics*, 69:163–183, 1967.

14. D. Kahneman and A. Tversky. Prospect theory: an analysis of decision under risk. *Econometrica*, 47:263–291, 1979.

15. Yakar Kannai. Non-standard concave utility functions. *Journal of Mathematical Economics*, 21:51–58, 1992.

16. H. J. Keisler. *Foundations of Infinitesimal Calculus.* Prindle, Weber&Schmidt Inc., Boston, 1976.

17. I. H. LaValle and P. C. Fishburn. Lexicographic state-dependent subjective expected utility. *Journal of Risk and Uncertainty*, 4:251–269, 1991.

18. I. H. LaValle and P. C. Fishburn. State-independent subjective expected lexicographic utility. *Journal of Risk and Uncertainty*, 5:217–240, 1992.

19. Daniel Lehmann. Generalized qualitative probability: Savage revisited. In Eric Horvitz and Finn Jensen, editors, *Proceedings of the Twelfth Conference on Uncertainty in Artificial Intelligence*, pages 381–388, Portland, Oregon, August 1996. Morgan Kaufmann.

20. Daniel Lehmann. What is qualitative: A framework for qantitative and qualitative decision theory. In *Qualitative Preferences in Deliberation and Practical Reasoning*, Stanford, CA, March 1997. AAAI Spring Symposium.

21. Daniel Lehmann. Nonstandard numbers for qualitative decision making. In Itzhak Gilboa, editor, *Proceedings of the Seventh Conference on Theoretical Aspects of Rationality and Knowledge (TARK 98)*, pages 161–174, Evanston, Ill., July 1998. Morgan Kaufmann.

22. Daniel Lehmann. Nonstandard utilities. Technical Report TR-98-3, Leibniz Center for Computer Science, Computer Science Institute, Hebrew University, Jerusalem, Israel, January 1998. revised version.

23. Blaise Pascal. *Pensées*, volume 28 of *Les Cents Chefs d'Œuvre de la Langue Française.* Robert Laffont, Paris, France, 1960. (first published in 1670).

24. Uday Rajan. Trembles in bayesian foundations. *Journal of Economic Theory*, 1998.

25. Abraham Robinson. *Non-standard Analysis.* North-Holland, Amsterdam, 1966.

26. Leonard J. Savage. *The Foundations of Statistics.* John Wiley, 1954.

27. P.J.H. Schoemaker. *Experiments On Decision Under Risk.* Martinus Nijhoff, Boston, 1980.

28. L.S. Shapley. The St. Petersburg paradox: A con game? *Journal of Economic Theory*, 14:439–442, 1977.

29. H. J. Skala. Nonstandard utilities and the foundations of game theory. *International Journal of Game Theory*, 3:67–81, 1974.

30. John von Neumann and Oskar Morgenstern. *Theory of Games and Economic Behavior.* Princeton University Press, second edition edition, 1947.

31. Abraham Wald. *Statistical Decision Functions.* John Wiley & Sons, New York, 1950.