PETER J. HAMMOND

CONSEQUENTIALIST FOUNDATIONS FOR EXPECTED UTILITY

April 1987, revised October 1987

To appear in *Theory and Decision*

ABSTRACT. Behaviour norms are considered for decision trees which allow both objective probabilities and uncertain states of the world with unknown probabilities. Terminal nodes have consequences in a given domain. Behaviour is required to be consistent in subtrees. Consequentialist behaviour, by definition, reveals a consequence choice function independent of the structure of the decision tree. It implies that behaviour reveals a revealed preference ordering satisfying both the independence axiom and a novel form of sure-thing principle. Continuous consequentialist behaviour must be expected utility maximizing. Other familiar assumptions then imply additive utilities, subjective probabilities, and Bayes' rule.

1. INTRODUCTION

An almost unquestioned hypothesis of modern normative decision theory is that acts are valued by their consequences.[1] Indeed, Savage (1954) *defines* an act as a function mapping uncertain of the world into a domain of conceivable consequences, thus identifying an act with the state-contingent consequence function which it generates.

Normative decision theory then erects a superstructure of various possible axiom systems upon this basic "consequentialist" hypothesis. The axioms and their implications are discussed at great length, without general agreement as to their acceptability or not as principles of rational behaviour. This discussion is usually conducted in a "normal form" decision problem where, relying upon von Neumann and Morgenstern's (1944) reduction of extensive games, the agent is viewed as having to choose a decision strategy specifying what decision to make in every conceivable set of circumstances. In a normal form decision problem, consequentialism means that behaviour is judged to be acceptable if its consequences (or, more generally, its risky and uncertain consequences) lie in the choice set corresponding to the feasible set of consequences resulting from all possible decisions. In other words, consequentialism implies that behaviour should reveal a consequence choice function. No other rationality postulates or axioms are implied by consequentialism alone.

In decision trees, however, consequentialism has much more powerful implications. A decision tree (Raiffa, 1968) amounts to a one person game in extensive form, in which there is perfect information because uncertainty is deemed to be resolved only when the agent knows how it is resolved, and because of perfect recall. A decision strategy in the normal form implies using consistent decision strategies in all possible "continuation trees" — or "subtrees" which continue from any node of the decision tree. These consistent decision strategies are just continuations of the original decision strategy. Consequentialism applies to these continuation strategies in the continuation decision trees no less than to complete strategies in complete decision trees. That, at least, is the fundamental hypothesis of this and several related papers.[2] It would be false if missed opportunities, regrets, sunk costs, etc. affected behaviour and yet were excluded from the domain of consequences. As a normative principle, however, consequentialism requires everything which should be allowed to affect decisions to count as a relevant consequence — behaviour is evaluated by its consequences, and nothing else. If regrets, sunk costs, even the structure of the decision tree itself, are relevant to normative behaviour, they are therefore already in the consequence domain. Indeed, the content of a normative theory of behaviour is then largely a matter of what counts in practice as a relevant consequence, rather than whether consequentialism and other abstract axioms are satisfied. For example, the standard economists' injunction to ignore sunk costs is a *practical* normative principle. Whereas expected utility maximization is a principle which has no practical content (beyond continuity of behaviour with respect to changes in probabilities) until the consequences which are the arguments of the utility function have been specified.[3] Another practical normative principle in economics, which I happen to find ethically unacceptable, is to consider only aggregate consumption of each good and aggregate income, rather than the distribution between rich and poor consumers.

Here, however, my subject is the implications of consequentialism for the structure of normative behaviour — in particular, the extent to which consequentialism implies that behaviour must maximize expected utility. Consequentialism *per se* is not about practical normative theories. Indeed, consequentialist behaviour is very impractical, insofar as it requires consideration of the whole of a decision tree which may be unmanageably complex. It may also result in very bad consequences even in rather simple trees, as long as behaviour is explicable by those bad consequences.

Since the content of the consequence domain is really a subject for practical normative theory, I shall avoid it by taking as fixed the state contingent consequence domains Y_s ($s \in E$) for a fixed finite set E of possible states of the world. For most of the paper, the domain Y_s is allowed to depend on the state of the world s so that decision problems involving life or death issues can be treated. Where at least one uncertain state involves the agent's accidental death, or the loss of a limb, that is incompatible with consequences such as those which include playing cricket normally.

The other key hypothesis, apart from that consequentialism applies both to complete and to continuation decision trees, is that it applies to all logically possible finite decision trees whose terminal nodes have consequences in the given domain. This is the assumption that there is an *unrestricted domain of consequential decision trees*.[4] The formal part of the paper commences with Section 2, in which consequential decision trees are defined. In addition to decision, chance and terminal nodes as considered, for instance, by Raiffa (1968), the definition allows "natural" nodes n at which "nature" refines the set $S(n) \subset E$ of possible states of the world into a partition $\{ S(n') \mid n' \in N_{+1}(n) \}$, where $N_{+1}(n)$ denotes the nodes which immediately succeed the natural node n. Natural nodes differ from chance nodes in that chance nodes have *positive* probabilities $\pi(n'|n)$ $(n' \in N_{+1}(n))$ attached to their immediate successors. The exclusion of zero probabilities is important for reasons which will become clear in Section 6. The other difference from Raiffa (1968) is that terminal nodes have attached to them consequences rather than payoffs. Indeed, the whole aim of the paper is to give sufficient conditions for the existence of a payoff or utility function. At a terminal node n where the set of possible states of the world is $S(n)$, the attached consequence is taken to be a simple (i.e. finitely supported) probability distribution $\tilde{y}^{S(n)}$ in the set $\tilde{Y}^{S(n)}$ of all simple distributions on the product set $Y^{S(n)} := \prod_{s \in S(n)} Y_s$ of "contingent consequence functions" $y^S(n) := (y_s)_{s \in S(n)}$.

Section 3 considers consistent behaviour norms for the set of consequential decision trees. A behaviour norm is a correspondence specifying a non-empty subset of $N_{+1}(n)$ at every decision node n of every consequential decision tree. Consistency is a natural requirement given that the norm should apply to continuation trees as well as complete trees. Essentially, the norm must prescribe behaviour in any continuation tree which is a restriction of the behaviour prescribed in the original tree. That is, the behaviour norms in the original tree and in the continuation tree must specify the same subset of $N_{+1}(n)$ at any decision node of the continuation tree. The potential addict example of Hammond (1976) is used to elucidate this consistency condition.

In order to specify that the consequences of behaviour prescribed by the norm must depend only on set of all consequences of possible behaviour, the consequences of behaviour must be derived. Section 4 begins by giving a method of doing this, starting at all the terminal nodes of the tree and then applying backward recursion. In any tree T the sets $\Phi_\beta(T,n)$, $F(T,n)$ are constructed at all nodes n in turn, where $\Phi_\beta(T,n)$ denotes the set of consequences of behaviour described by β, and $F(T,n)$ denotes the feasible set of consequences. One has $\emptyset \neq \Phi_\beta(T,n) \subset F(T,n)$. Section 4 spells out the crucial consequentialist hypothesis, which is that $F(T,n) = F(T',n)$ implies $\Phi_\beta(T,n) = \Phi_\beta(T',n)$. Thus for every non-empty $S \subset E$, there must exist a "revealed" consequence choice function C^S_β such that, in every tree T and at every node n of T, one has

$$\emptyset \neq \Phi_\beta(T,n) = C_\beta^{S(n)}(F(T,n)) \subset F(T,n) \subset \tilde{Y}^{S(n)}.$$

In fact, because behaviour is consistent in continuation trees, it is enough to have the above hold at the initial node of any decision tree.

After the necessary preliminary definitions, Sections 5, 6, 7 and 8 give a complete characterization of a consequentialist behaviour norm. Using the fact that such a norm must be consistent and consequentialist in all continuous trees, Section 5 proves the existence, for every non-empty $S \subset E$, of a revealed preference ordering R^S_β (i.e. a complete, transitive, binary relation which the consequences of behaviour must maximize at any decision node n of any consequential decision tree, provided that $S(n) = S$). Then Section 6 proves that all these preference orderings must satisfy Samuelson's (1952) independence axiom.[5] Section 7 proves that they must satisfy a new version of Savage's (1954) sure-thing principle extended to allow *independent* probabilities. *Only* independent probabilities, however — in particular, other extended versions of the sure-thing principle, such as that due to Anscombe and Aumann (1963), are not implied. Finally, this part of the paper concludes in Section 8 with a proof that the three necessary conditions just derived are actually a complete characterization of consequentialist behaviour, because any collection R^S $(\emptyset \neq S \subset E)$ of preference orderings satisfying the independence axiom and the sure-thing principle for independent probabilities permits the construction of a corresponding consequentialist norm.

This complete characterization of consequentialist behaviour norms leaves open the possibility of lexicographic preferences which violate expected utility maximization. For expected utility to be satisfied, an additional continuity or “Archimedean” axiom is clearly required. Section 9 states one for the behaviour norm itself. As probabilities vary at the chance nodes of a consequential decision tree there is a correspondence expressing the dependence of behaviour on probabilities. This correspondence is required to have a closed graph. This implies that the preference orderings R^S ($\emptyset \neq S \subset E$) all satisfy Herstein and Milnor's (1953) continuity (or Archimedean) axiom, and that each corresponds to the expected value of each von Neumann-Morgenstern utility function (NMUF) v^S in a unique cardinal equivalence class.

Section 9 makes no use of the sure thing principle for independent probabilities. In Section 10 this other implication of consequentialism is invoked to characterize further the structure of the NMUF's v^S for different $S \subset E$. One possibility is that each v^S can be expressed in the form of an additive evaluation function $\sum_{s \in S} v_s$, as in Wilson (1968) and Myerson (1979). But two other possibilities are that v^S be expressible as a product of either the form $\prod_{s \in S} v_s$ or the form $-\prod_{s \in S} (-v_s)$ where, in the first case each v_s is positive, and in the second case each v_s is negative. These new possibilities arise because consequentialism only implies the sure-thing principle for independent probabilities. Section 10 presents a complete characterization of the set of continuous consequentialist behaviour norms — they must result in consequences which maximize expected utility, with each NMUF v^S being built up from the NMUF's v_s ($s \in S$) as in one of the three cases mentioned above. Section 10 concludes by showing how which of these three cases occurs depends on whether there is a preference to have uncertainty about which of the two events S_1 or S_2 holds resolved before chance moves unravel, and that this preference depends upon whether the random consequences y^{S_1} and y^{S_2} have positively or negatively correlated utilities.

To this point the consequence domains $Y_s (s \in E)$ have been unrestricted. Section 11 finally introduces the standard assumption of Savage (1954), Anscombe and Aumann (1963), Harsanyi (1978), Myerson (1979), etc. that Y_s is independent of s. As already suggested, the assumption is undesirable and the further implications of relaxing it will be explored elsewhere. It is used to here to see how justified are standard axiomatic formulations of subjective probabilities.

Section 11 also adapts another Savage (1954) assumption regarding “constant acts”. Here, it is required that when decision trees have only constant (state-independent) consequences, the constant consequences of behaviour should not depend upon the set of possible states of the world — that is, behaviour itself should be “state-independent”. This is actually rather stronger than Savage's assumption because it encompasses probability distributions over constant consequences. These two new assumptions exclude the multiplicative utility functions v^S when the constant consequence domain Y has at least three distinct indifference classes. Then $v^S \equiv \sum_{s \in S} v_s$. But if Y has at most two distinct indifference classes, the two multiplicative cases remain possible. Indeed one cannot then exclude the preferences associated with Ellsberg's (1961) paradox, as is shown in Section 12.

Reverting to the customary additive case, Section 13 shows how a consequentialist behaviour norm satisfying all the assumptions so far mentioned must reveal unique conditional subjective probabilities $p(s|S)$ ($s \in S \subset E$), as well as unique cardinal equivalence class of state-independent NMUF's defined on the constant consequence domain Y. Subjective probabilities must satisfy Bayes' rule, for reasons similar to those considered by Weller (1978). Thus Section 13 justifies the Anscombe and Aumann (1963) approach. It also contains a small surprise — null events are not allowed, and all subjective probabilities must be positive, just like their objective counterparts.

Section 14 contains a brief summary and some concluding remarks.

2. CONSEQUENTIAL DECISION TREES

As discussed in the introduction, a fixed non-empty finite set E of possible states of the world is assumed. For each state s in E, assume too that there is a fixed non-empty state contingent *consequence domain* Y_s whose members are the consequences which, if state s should occur, may result from behaviour in the domain of decision problems being considered.

For any non-empty subset S of E, S is an *event*, and $Y^S := \prod_{s \in S} Y_s$ denotes the set of *contingent* *consequence functions* given that event. Each member y^S of Y^S is a mapping from the states in event S to the appropriate state contingent consequence domain.

A *simple probability distribution of contingent consequences* — which will be called, for short, a "risky consequence" or even just a "consequence" — is a function $\mu(y^S)$, taking non-negative real values on its domain Y^S, for which there exists a finite *support* $K \subset Y^S$ satisfying both $\mu(y^S) > 0 \iff y^S \in K$ and also $\sum_{y^S \in K} \mu(y^S) = 1$. For every event S, the set of all such risky consequences will be denoted by $\Delta(Y^S)$ or just by $\tilde{Y}^S$.

In the absence of uncertain states of the world, decision trees with chance, decision and terminal nodes should be familiar from Raiffa's (1968) lucid introductory lectures. Here, therefore, I merely introduce notation and also adapt the definition to accommodate uncertain states of the world, by allowing "natural nodes" at which a "move of nature" partitions the set of possible states. Also, every terminal node x of a decision tree T will have attached to it a consequence $\gamma(x)$ in the space $\tilde{Y}^{S(x)}$ of risky contingent consequence functions given the event $S(x)$.

Formally, a *consequential decision tree* is a collection

$$T = \langle N, N^*, N^0, N^1, X, N_{+1}(\cdot), n_0, \pi(\cdot|\cdot), S(\cdot), \gamma(\cdot) \rangle$$

whose ten component parts are described and interpreted as follows:

(i) N is a non-empty finite set of *nodes* of the tree T, which is partitioned into the four disjoint sets N^*, N^0, N^1, and X described below;

(ii) N^* is the (possibly empty) set of *decision nodes*;

(iii) N^0 is the (possibly empty) set of *chance nodes*;

(iv) N^1 is the (possibly empty) set of *natural nodes*;

(v) X is the non-empty set of *terminal nodes*;

(vi) $N_{+1} : N \twoheadrightarrow N$ is the *immediate successor* correspondence satisfying:
  (a) $\forall n \in N : \; n \notin N_{+1}(n)$;
  (b) $\forall n \in N : \; N_{+1}(n) = \emptyset \iff n \in X$;
  (c) $\forall n, n' \in N : \; N_{+1}(n) \cap N_{+1}(n') \neq \emptyset \iff n = n'$;

(vii) n_0 is the unique *initial node* in N satisfying $\forall n \in N : \; n_0 \notin N_{+1}(n)$;

(viii) for each pair $n \in N^0$ and $n' \in N_{+1}(n)$, the *positive* real number $\pi(n'|n)$ is the *probability* of the chance move from n to n', with

$$\sum_{n' \in N_{+1}(n)} \pi(n'|n) = 1 \; (\forall\, n \in N^0);$$

(ix) at each node n of N, $S(n)$ is the *event at* n — the non-empty set of states of the world s that remain possible after reaching n, with
  (a) $\forall n \notin N^1; \; \forall n' \in N_{+1}(n) : \; S(n') = S(n)$;
  (b) $\forall n \in N^1 : \; \{S(n') \mid n' \in N_{+1}(n)\}$ is a partition of $S(n)$;

(x) γ is the *consequence mapping* with domain X, satisfying $\gamma(x) \in \tilde{Y}^{S(x)}$ for all $x \in X$, where $\gamma(x)$ is the consequence of reaching terminal node x.

In the above definition, (x) does not restrict consequences to "elementary" sure consequences in the set $Y := \cup_{s \in E} Y_s$. Of course, if the consequence $\gamma(x)$ at the terminal node x is not already an elementary consequence, the tree could be extended so that all terminal nodes do have elementary consequences. For one could replace x with a chance node $\hat{n}(x)$ which is succeeded immediately by the set

$$N_{+1}(\hat{n}(x)) = \{\, \hat{n}(y^{S(x)}) \mid \exists y^{S(x)} \in Y^{S(x)} : \; \gamma(x, y^{S(x)}) > 0 \,\}$$

of natural nodes corresponding to each contingent consequence function $y^{S(x)}$ which occurs with positive probability. At the chance node $\hat{n}(x)$ there should be transition probabilities given by $\pi(\hat{n}(y^{S(x)})|\hat{n}(x)) = \gamma(x, y^{S(x)})$ which correspond to the relevant probability of each contingent consequence function $y^{S(x)}$. Finally, each natural node $\hat{n}(y^{S(x)})$ of the set $N_{+1}(\hat{n}(x))$ should be succeeded by the terminal nodes in the set $\{\hat{x}(s, y^{S(x)}) \mid s \in S(x)\}$, each of which has an elementary consequence $\hat{\gamma}(\hat{x}(s, y^{S(x)})) = y_s^{S(x)}$, in an obvious notation. These extra nodes leading to elementary consequences will usually be omitted in order to make it easier both to describe and visualize the decision trees used in proofs later on.

Restrictions (a), (b) and (c) of part (vi) are imposed so that T is indeed a tree. Then the set of nodes N can in fact be constructed recursively from the correspondence N_{+1}, starting with the initial node n_0, then proceeding to $N_{+1}(n_0)$, then to $\bigcup_{n \in N_{+1}(n_0)} N_{+1}(n)$, etc. until terminal nodes are reached (after a finite number of iterations, because N is assumed to be finite).

In part (viii), the transition probabilities $\pi(n'|n)$ are restricted to be strictly *positive* for reasons which will become apparent later in Section 6.

For each event S, write $\mathcal{T}(S)$ for the set of all possible consequential decision trees satisfying $S(n_0) = S$ at the initial node, and let $\mathcal{T}$ denote $\cup \mathcal{T}(S)$ as S varies over all the non-empty subsets of E. Then $\mathcal{T}$ is the domain of all possible finite consequential decision trees — or *trees* for short.

Where the tree T is variable, it will sometimes be included as an argument in expressions such as $N^*(T), N_{+1}(T, n), n_0(T)$, etc.

3. CONSISTENT BEHAVIOUR NORMS

A *behaviour norm* is a correspondence β with a domain consisting of all pairs $T \in \mathcal{T}$ and $n \in N^*(T)$. Its value $\beta(T, n)$ is a non-empty subset of the appropriate set $N_{+1}(T, n)$. Here, of course, $N_{+1}(T, n)$ is the set of all decisions which are feasible at the decision node n of the tree T in the domain $\mathcal{T}$ of all consequential decision trees. The value $\beta(T, n)$ should be interpreted as the set of acceptable or recommended decisions at this decision node. As usual in decision theory, the possibility of more than one decision being acceptable is specifically allowed. But *stochastic behaviour* norms, in which the value $\beta(T, n)$ consists of probability distributions over $N_{+1}(T, n)$, are excluded — I plan to discuss them in later work.

Given any tree $T = \langle N, N^*, N^0, N^1, X, N_{+1}(\cdot), n_0, \pi(\cdot|\cdot), S(\cdot), \gamma(\cdot)\rangle$ of $\mathcal{T}$ and any node $\bar{n}$ of N, there is a "subtree" or *continuation tree*

$$\bar{T} = \langle \bar{N}, \bar{N}^*, \bar{N}^0, \bar{N}^1, \bar{X}, \bar{N}_{+1}(\cdot), \bar{n}_0, \bar{\pi}(\cdot|\cdot), \bar{S}(\cdot), \bar{\gamma}(\cdot)\rangle$$

starting with initial node $\bar{n}_0 = \bar{n}$. To define it explicitly, first let $>$ be the binary *successor* relation on N with the property that $n' > n$ if and only if there exists a chain $n_1, n_2, ..., n_k$ in N such that $n_1 = n$, $n_k = n'$ and $n_{j+1} \in N_{+1}(n_j)$ for $j = 1$ to $k - 1$. Second, let $N(n) := \{ n' \in N \mid n' > n \text{ or } n' = n \}$ be the set of nodes in N which either succeed or coincide with n. Now define :

(i) $\bar{N} := N(\bar{n})$;
(ii) $\bar{N}^* := N^* \cap \bar{N}$;
(iii) $\bar{N}^0 := N^0 \cap \bar{N}$;
(iv) $\bar{N}^1 := N^1 \cap \bar{N}$;
(v) $\bar{X} := X \cap \bar{N}$;
(vi) $\bar{N}_{+1} : \bar{N} \twoheadrightarrow \bar{N}$ as the restriction to $\bar{N}$ of $N_{+1} : N \twoheadrightarrow N$ which, because $\bar{N}$ contains all the successors of any of its members, must satisfy $\bar{N}_{+1}(n) = N_{+1}(n) \subset \bar{N}$ for all $n \in \bar{N}$;
(vii) $\bar{n}_0 = \bar{n}$ as the new initial node;
(viii) $\bar{\pi}(\cdot|\cdot)$ as the restriction to all pairs $n \in \bar{N}, n' \in \bar{N}_{+1}(n) = N_{+1}(n)$ of the probabilities $\pi(n'|n)$;
(ix) $\bar{S}(\cdot)$ as the restriction to $\bar{N}$ of the correspondence $S : N \twoheadrightarrow E$;
(x) $\bar{\gamma}(\cdot)$ as the restriction to $\bar{X}$ of the mapping γ with domain X.

From this definition it is obvious that $\bar{T}$ is itself a consequential decision tree in $\mathcal{T}$.

Henceforth, let $T(n)$ denote the continuation from node $n \in N(T)$ of any tree $T \in \mathcal{T}$.

Let n be any decision node of a tree T in $\mathcal{T}$. When the agent comes to make a decision at node n, the decision problem is really described by the tree $T(n)$ since all other nodes, branches, and consequences are no longer possible. Whether the agent foresees it or not, behaviour at node n is described by $\beta(T(n), n)$. Thus, to be an accurate description of behaviour in the tree T, the behaviour norm β must be *consistent* in the sense that $\beta(T, n) = \beta(T(n), n)$ at all decision nodes n of every tree T of $\mathcal{T}$. From now on only consistent behaviour norms will be considered, and they will be called simply *norms*.

Many opponents of consequentialism have chosen to attack this consistency condition, arguing that the agent does not just face tree $T(n)$ at node n of the tree T. Instead, other aspects of T are regarded as relevant, even though they are not included in the consequences. McClennen (1986, 1987), for example, puts forward an interesting theory of "resoluteness", in which the agent is expected to take account at n of resolutions made at previous nodes of T. If such resolutions are indeed relevant, and are not to be counted as consequences, this is actually a violation of consequentialism rather than of consistency because behaviour has to be explained by resolutions as well as by its consequences.

In fact, to appreciate better the weakness of consistency and the strength of consequentialism, it may be helpful to reconsider the "potential addict" of Hammond (1976). The agent faces a consequential decision tree T with two decision nodes n_0, n_1, and three terminal nodes x_y ($y \in \{a, b, c\}$) as in Figure 3.1. There are no chance or natural nodes. The immediate successor correspondence N_{+1} satisfies $N_{+1}(n_0) = \{x_c, n_1\}, N_{+1}(n_1) = \{x_a, x_b\}$. The consequence mapping satisfies $\gamma(x_y) = y$ for $y \in \{a, b, c\}$. The three possible consequences a, b, c are given the following interpretations:

(i) a is enjoying some addictive activity for a while but then giving it up before any permanent damage to health is caused;
(ii) b is becoming addicted and suffering permanent damage;
(iii) c is never exposing oneself to the possibility of addiction.

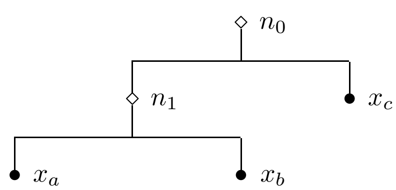

Figure 3.1

The potential addict decides at n_0 whether to embark on the addictive activity, and at n_1 whether to give up the activity after addiction has set in but before any other permanent damage has occurred. A *naïve* potential addict plans to enjoy the consequence a and moves to n_1, but moves on from there to x_b because addition has set in. A *sophisticated* potential addict predicts this behaviour and moves first to x_c in order to avoid addiction and the consequence b. Both types of behaviour are consistent in T because they both satisfy

$$\beta(T, n_1) = \beta(T(n_1), n_1) = \{x_b\}.$$

In addition naïve behaviour has $\beta(T, n_0) = \{n_1\}$ and sophisticated behaviour has $\beta(T, n_0) = \{x_c\}$. Thus even naïve behaviour is actually consistent, though only unintentionally so. Naïve behaviour has consequence b and sophisticated behaviour has consequence c.

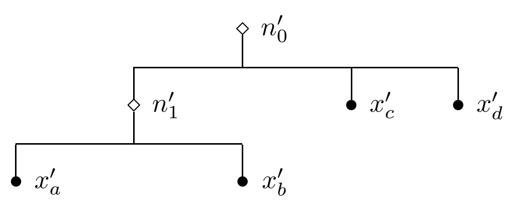

Figure 3.2

Though they are both consistent, neither naïve nor sophisticated behaviour is consequentialist. For consider the decision tree T' of Figure 3.2 in which an extra option of commitment is available at n'_0, leading to terminal node x'_d whose consequence is taken to be a. Now the naïve agent, who cannot see the need to commit, has $\beta(T', n'_0) = \{n'_1, x'_d\}$ and the ultimate consequence of naïve behaviour is either a or b depending on whether the first move is to x'_d or to n'_1. The sophisticated agent, however, takes full advantage of the opportunity to commit by behaving according to $\beta(T', n'_0) = \{x'_d\}$ with consequence a. Thus, although the decision trees T, T' both offer the same range $\{a, b, c\}$ of feasible consequences, the actual consequences of naïve and sophisticated behaviour change as one changes the structure of the tree. This is a violation of the consequentialist principle to be defined below.

In the face of the potential addict's decision problem, sophisticated behaviour seems clearly the best, despite its violation of consequentialism. This does not imply, however, that consequentialism is irrational. Rather, the potential addict is really two (potential) persons, before and after addiction, and the decision problem has to be analysed as a "game" between two "rational players". This paper will consider only single person decision theory and defer all questions regarding multi-person decisions for later analysis.

4. CONSEQUENTIALIST BEHAVIOUR

As explained in the introduction, *consequentialism* entails the behaviour norm β revealing, for every event S, a *a consequence choice function* C^S_β with the property that, for all trees T of $\mathcal{T}$, and at all nodes $n \in N(T)$,

$$\Phi_\beta(T, n) = C^{S(n)}_\beta(F(T, n)).$$

Here $F(T, n)$ is the non-empty subset of $\tilde{Y}^{S(n)}$ consisting of all the consequences of behaviour that are still possible after reaching node n. And $\Phi_\beta(T, n)$ is the non-empty subset of $F(T, n)$ consisting of all the consequences which could possibly result from following the behaviour norm β after reaching node n. Both $\Phi_\beta(T, n)$ and $F(T, n)$ are easily calculated by backward recursion, and the calculation permits a formal proof that indeed

$$\emptyset \neq \Phi_\beta(T, n) \subset F(T, n) \subset \tilde{Y}^{S(n)}$$

for all nodes n in all trees $T \in \mathcal{T}$. Obviously the backward recursion starts at the terminal nodes $x \in X$, where one has

$$\emptyset \neq \Phi_\beta(T, x) = F(T, x) = \{\gamma(x)\} \subset \tilde{Y}^{S(x)}$$

because only the one consequence $\gamma(x)$ is possible. For the other nodes in N, there are three obvious cases to consider

Case 1. Decision nodes $n \in N^*$. Here:

$$F(T, n) = \bigcup\nolimits_{n' \in N_{+1}(n)} F(T, n'); \quad \Phi_\beta(T, n) = \bigcup\nolimits_{n' \in \beta(T, n)} \Phi_\beta(T, n').$$

Case 2. Chance nodes $n \in N^0$. Here, given the probabilities $\pi(n'|n)$ of the immediately succeeding nodes $n' \in N_{+1}(n)$, one has:

$$F(T, n) = \sum\nolimits_{n' \in N_{+1}(n)} \pi(n'|n)\, F(T, n');$$
$$\Phi_\beta(T, n) = \sum\nolimits_{n' \in N_{+1}(n)} \pi(n'|n)\, \Phi_\beta(T, n').$$

Case 3. Natural nodes $n \in N^1$. Here, given the partition $\cup_{n' \in N_{+1}(n)} S(T, n')$ of $S(T, n)$, one has:

$$F(T, n) = \prod\nolimits_{n' \in N_{+1}(n)} F(T, n'); \quad \Phi_\beta(T, n) = \prod\nolimits_{n' \in N_{+1}(n)} \Phi_\beta(T, n').$$

Thus $F(T, n)$ is the set of all probability distributions $\mu \in \tilde{Y}^{S(n)}$ which:

(i) in Case 1 are equal to a probability distribution $\mu(n') \in F(T, n')$ for at least one immediately succeeding node $n' \in N_{+1}(n)$;
(ii) in Case 2 can be expressed as the probability mixture $\sum_{n' \in N_{+1}(n)} \pi(n'|n)\, \mu(n')$ of probability distributions satisfying $\mu(n') \in F(T, n')$ for all immediately succeeding nodes $n' \in N_{+1}(n)$;
(iii) in Case 3 can be expressed as the product joint distribution $\prod_{n' \in N_{+1}(n)} \mu(n')$ of the independent probability distributions $\mu(n')$ satisfying $\mu(n') \in F(T, n')$ for all immediately succeeding nodes $n' \in N_{+1}(n)$.

The construction of $\Phi_\beta(T, n)$ in each case is similar. The product set arises in Case 3 because the probabilities at all the chance nodes of any decision tree are independent.

This backward recursion can be applied throughout the decision tree T until the initial node n_0 is reached, at which the feasible set is $F(T) := F(T, n_0)$ and the revealed choice set is $\Phi_\beta(T) := \Phi_\beta(T, n_0)$. The existence of a consequence choice function C^S_β for every event S evidently implies that whenever two decision trees T, T' are *consequentially equivalent*[6] in the sense that $F(T) = F(T')$, then behaviour in the two trees must also be *consequentially equivalent*, in the sense that $\Phi_\beta(T) = \Phi_\beta(T')$. Thus the structure of the decision tree must be irrelevant to the consequences of acceptable or recommended behaviour. Of course, such consequential equivalence of the two trees T, T' requires that $S(n_0) = S'(n'_0)$ because $F(T) \subset \tilde{Y}^{S(n_0)}$ and $F(T') \subset \tilde{Y}^{S'(n'_0)}$. This explains why there is a different revealed consequence choice function C^S_β for each different event S.

The domain of C^S_β has not yet been specified. It is the complete collection of all non-empty finite subsets of $\tilde{Y}^S$ because of

THEOREM 4: *For all non-empty $S \subset E$ and non-empty finite $Z \subset \tilde{Y}^S$, there exists a consequential decision tree $T \in \mathcal{T}(S)$ such that $F(T) = Z$ and $\emptyset \neq C^S_\beta(Z) = \Phi_\beta(T) \subset Z$.*

Proof. Given any non-empty finite subsets S and Z as in the hypothesis of the theorem, there is a *trivial* decision tree T with just one initial decision node n_0 at which the set of immediate successors $N_{+1}(n_0)$ is equal to the set of terminal nodes, with one terminal node for each consequence in the finite set Z. Thus

$$N_{+1}(n_0) = X = \{\, n(\zeta) \mid \zeta \in Z \,\}, \quad S(n_0) = S(n(\zeta)) = S \quad (\forall \zeta \in Z),$$

and the consequence mapping $\gamma : X \mapsto \tilde{Y}^S$ has $\gamma(n(\zeta)) = \zeta$ (all $\zeta \in Z$). So $F(T, n_0) = F(T) = Z$ in this tree T. Therefore:

$$C^S_\beta(Z) = C^S_\beta(F(T)) = \Phi_\beta(T) = \gamma(\beta(T, n_0)).$$

But then it must be true that $C^S_\beta(Z)$ is a non-empty subset of Z because $\beta(T, n_0)$ is a non-empty subset of X. ∎

Were the domain of the behaviour norm restricted to trivial trees such as that used to prove Theorem 4, then *any* consequence choice function would be consistent with consequentialism. With an unrestricted (or less restricted) domain, however, the consistency property of Section 3 becomes important, and severely restricts the set of consequentialist behaviour norms which are logically possible.

5. CONSEQUENTIALISM IMPLIES ORDINALITY

In this section it is shown that a consequentialist behaviour norm β reveals not only a consequence choice function C^S_β for each event S, but also a consequence preference ordering R^S_β, assuming that the domain of consequential decision trees is unrestricted. Recall that a preference ordering is a complete transitive binary relation. Let R^S_β be the binary preference relation revealed by choices among pairs — namely, the relation defined, for all $\lambda, \mu \in \tilde{Y}^S$, by

$$\lambda \, R^S_\beta \, \mu \iff \lambda \in C^S_\beta(\{\lambda, \mu\}).$$

Because of Theorem 4, $C^S_\beta(\{\lambda,\mu\})$ is non-empty for every pair $\lambda,\mu \in \tilde{Y}^S$, so the relation R^S_β is indeed complete. It will be shown that C^S_β maximizes R^S_β, for all non-empty finite sets $Z \subset \tilde{Y}^S$, because

$$C^S_\beta(Z) = \{\lambda \in Z \mid \mu \in Z \Rightarrow \lambda R^S_\beta \mu\}$$

and also that R^S_β is transitive. The proof uses a series of lemmas, of which the first is based upon considering, for any non-empty $S \subset E$, non-empty finite $Z \subset \tilde{Y}^S$, and pair $\lambda,\mu \in Z$, the decision tree

$$T(\lambda,\mu,Z) := \langle N, N^*, N^0, N^1, X, N_{+1}(\cdot), n_0, \pi(\cdot|\cdot), S(\cdot), \gamma(\cdot)\rangle$$

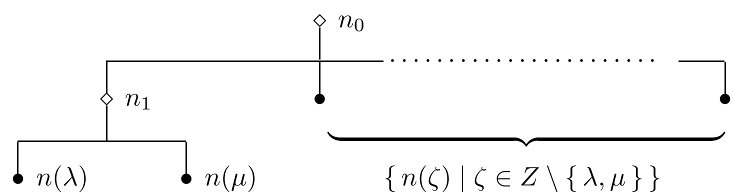

Figure 5.1. The Decision Tree $T(\lambda,\mu,Z)$

illustrated in Figure 5.1 above. This is constructed as follows:

(i) $N := \{ n_0, n_1 \} \cup \{n(\zeta) \mid \zeta \in Z\}$;
(ii) $N^* := \{n_0, n_1\}$;
(iii) $N^0 := \emptyset$;
(iv) $N^1 := \emptyset$;
(v) $X := \{n(\zeta) \mid \zeta \in Z\}$;
(vi) $N_{+1}(n_0) := \{n_1\} \cup \{n(\zeta) | \zeta \in Z \setminus \{\lambda,\mu\}\}$, and $N_{+1}(n_1) := \{n(\lambda), n(\mu)\}$;
(vii) n_0 is the initial node;
(viii) conditional probabilities are undefined because $N^0 = \emptyset$;
(ix) $S(n) := S$ (all $n \in N$);
(x) $\gamma(n(\zeta)) := \zeta$ (all $\zeta \in Z$).

The tree $T(\lambda,\mu,Z)$ has $F(T,n_1) = \{\lambda,\mu\}$ and $F(T,n_0) = Z$.

LEMMA 5.1: *For all non-empty $S \subset E$, non-empty finite $Z \subset \tilde{Y}^S$, and pairs $\lambda,\mu \in Z$:*

(a) $\lambda \in C^S_\beta(Z) \Rightarrow \lambda\ R^S_\beta\ \mu$;

(b) $\mu \in C^S_\beta(Z)\ \&\ \lambda\ R^S_\beta\ \mu \Rightarrow \lambda \in C^S_\beta(Z)$.

Proof.

(1) Construct $T := T(\lambda,\mu,Z)$ as above. Then, as $F(T) = Z$, so

$\lambda \in C^S_\beta(Z) \iff \lambda \in \Phi_\beta(T)$ (by definition of C^S_β)
$\iff n_1 \in \beta(T,n_0)\ \&\ \lambda \in \Phi_\beta(T,n_1)$ (by backward recursion)
$\iff n_1 \in \beta(T,n_0)\ \&\ n(\lambda) \in \beta(T,n_1)$ (because only $n(\lambda)$ has $\gamma(n(\lambda)) = \lambda$).

(2) Consistency implies that $\beta(T(n_1),n_1) = \beta(T,n_1)$ and so

$$n(\lambda) \in \beta(T,n_1) \iff n(\lambda) \in \beta(T(n_1),n_1)) \iff \lambda \in \Phi_\beta(T(n_1)).$$

(3) Obviously $F(T(n_1)) = \{\lambda,\mu\}$, so consequentialism entails

$$\lambda \in \Phi_\beta(T(n_1)) \iff \lambda\ R^S_\beta\ \mu$$

by definition of the binary relation R^S_β.

(4) To complete the proof of part (a), observe that

$$\begin{aligned} \lambda \in C^S_\beta(Z) &\Rightarrow n(\lambda) \in \beta(T, n_1) && \text{(by (1))} \\ &\Rightarrow \lambda \in \Phi_\beta(T(n_1)) && \text{(by (2))} \\ &\Rightarrow \lambda \; R^S_\beta \; \mu && \text{(by (3)).} \end{aligned}$$

(5) Replacing λ by μ in (1) above gives, in particular,

$$\mu \in C^S_\beta(Z) \Rightarrow n_1 \in \beta(T, n_0).$$

(6) Also

$$\begin{aligned} \lambda \; R^S_\beta \; \mu &\Rightarrow \lambda \in \Phi_\beta(T(n_1)) && \text{(by (3))} \\ &\Rightarrow n(\lambda) \in \beta(T, n_1) && \text{(by (2)).} \end{aligned}$$

(7) To complete the proof of part (b), observe finally that, by combining (5) and (6) and then using (1),

$$\begin{aligned} \mu \in C^S_\beta(Z) \;\&\; \lambda \; R^S_\beta \; \mu &\Rightarrow n_1 \in \beta(T, n_0) \;\&\; n(\lambda) \in \beta(T, n_1) \\ &\Rightarrow \lambda \in C^S_\beta(Z). \end{aligned}$$

■

LEMMA 5.2: *For all non-empty $S \subset E$ and non-empty finite $Z \subset \tilde{Y}^S$:*

$$C^S_\beta(Z) = \{ \lambda \in Z \mid \mu \in Z \Rightarrow \lambda R^S_\beta \mu \}.$$

Proof.

(1) If $\lambda \in C^S_\beta(Z)$, then obviously $\lambda \in Z$. In addition, by part (a) of Lemma 5.1, $\lambda \; R^S_\beta \; \mu$ for all $\mu \in Z$.

(2) Conversely, by Theorem 4, $C^S_\beta(Z)$ must be non-empty, so suppose $\lambda^* \in C^S_\beta(Z)$. Then, if $\lambda \in Z$ and if $\lambda \; R^S_\beta \; \mu$ for all $\mu \in Z$, it follows that $\lambda \; R^S_\beta \; \lambda^*$ in particular, because $\lambda^* \in Z$. So, by part (b) of Lemma 5.1, $\lambda \in C^S_\beta(Z)$. ■

LEMMA 5.3: *For all non-empty $S \subset E$, the binary relation R^S_β is transitive.*

Proof. Suppose $\lambda, \mu, \nu \in \tilde{Y}^S$ are such that $\lambda \; R^S_\beta \; \mu$ and $\mu \; R^S_\beta \; \nu$. Define Z as the set $\{ \lambda, \mu, \nu \}$. By Theorem 4 $C^S_\beta(Z)$ is not empty. Three cases are then possible:

(1) If $\lambda \in C^S_\beta(Z)$ then, by part (a) of Lemma 5.1, $\lambda \; R^S_\beta \; \nu$.

(2) If $\mu \in C^S_\beta(Z)$ then, because $\lambda \; R^S_\beta \; \mu$ by hypothesis, it follows from part (b) of Lemma 5.1 that $\lambda \in C^S_\beta(Z)$. So case (1) applies.

(3) If $\nu \in C^S_\beta(Z)$ then, because $\mu \; R^S_\beta \; \nu$ by hypothesis, it follows from part (b) of Lemma 5.1 that $\mu \in C^S_\beta(Z)$. So case (2) applies.

Thus, in all three cases, $\lambda \; R^S_\beta \; \nu$. ■

So we have proved

THEOREM 5.4: *If the consistent behaviour norm β is consequentialist for the unrestricted domain of all consequential decision trees $T \in \mathcal{T}$, then it reveals an ordinal choice function C^S_β for every non-empty $S \subset E$, with a corresponding revealed preference ordering R^S_β.*

Next I shall turn to properties of the orderings R^S_β ($S \subset E$). From now on, the subscript β will be omitted.

6. CONSEQUENTIALISM IMPLIES INDEPENDENCE

The following *independence axiom* is due to Samuelson (1952). For all non-empty $S \subset E$, all $\lambda, \mu, \nu \in \tilde{Y}^S$, and all real numbers α with $0 < \alpha < 1$, one has

$$\lambda \; R^S \; \mu \iff [\alpha\lambda + (1-\alpha)\nu] \; R^S \; [\alpha\mu + (1-\alpha)\nu].$$

This is another implication of consequentialism, as can be seen from the decision tree $T^S(\alpha, \lambda, \mu, \nu)$ shown in Figure 6.1, with $Z := \{\lambda, \mu, \nu\}$ and:

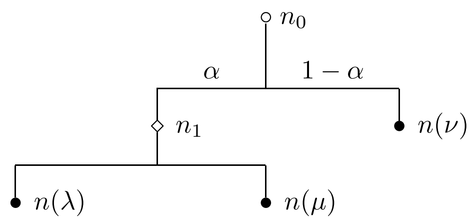

Figure 6.1. The Decision Tree $T^S(\alpha, \lambda, \mu, \nu)$

(i) $N := \{ n_0, n_1 \} \cup \{ n(\zeta) \mid \zeta \in Z \}$;
(ii) $N^* := \{n_1\}$;
(iii) $N^0 := \{n_0\}$;
(iv) $N^1 := \emptyset$;
(v) $X := \{ n(\zeta) \mid \zeta \in Z \}$;
(vi) $N_{+1}(n_0) := \{n_1, n(\nu)\}$ and $N_{+1}(n_1) := \{ n(\lambda), n(\mu) \}$;
(vii) n_0 the initial node;
(viii) $\pi(n_1|n_0) := \alpha, \quad \pi(n(\nu)|n_0) := 1-\alpha$;
(ix) $S(n) := S$ (all $n \in N$);
(x) $\gamma(n(\zeta)) := \zeta$ (all $\zeta \in Z$).

THEOREM 6.1. *If the consistent behaviour norm β is consequentialist for the unrestricted domain of all consequential decision trees $T \in \mathcal{T}$, then, for all non-empty $S \subset E$, the revealed preference ordering R^S must satisfy the independence axiom.*

Proof. For any event S, construct the tree $T := T^S(\alpha, \lambda, \mu, \nu)$ as in Figure 6.1. Then backward recursion gives $F(T, n_1) = \{\lambda, \mu\} = F(T(n_1))$ and:

$$F(T) = \alpha\{ \lambda, \mu \} + (1-\alpha)\{\nu\} = \{ \alpha\lambda + (1-\alpha)\nu \; , \; \alpha\mu + (1-\alpha)\nu \};$$
$$\Phi(T) = \alpha\Phi(T, n_1) + (1-\alpha)\{\nu\}.$$

Now

$$\begin{aligned} & [\alpha\lambda + (1-\alpha)\nu] \;\; R^S \;\; [\alpha\mu + (1-\alpha)\nu] \\ \iff & [\alpha\lambda + (1-\alpha)\nu] \;\; \in \;\; \Phi(T) = \alpha\Phi(T, n_1) + (1-\alpha)\{\nu\} \\ \iff & \;\; \lambda \in \Phi(T, n_1) \;\; \iff n(\lambda) \in \beta(T, n_1) \;\; \iff \;\; n(\lambda) \in \beta(T(n_1), n_1) \\ \iff & \;\; \lambda \in \Phi(T(n_1)) \;\; \iff \lambda \; R^S \; \mu. \end{aligned}$$

■

Note that the probabilities $\pi(n'|n)$ ($n \in N^0$, $n' \in N_{+1}(n)$) in any finite consequential decision tree must indeed be strictly positive, in general. For if not, the argument used to prove Theorem 6.1 remains valid even when $\alpha = 0$, and implies that $\nu \; R^S \; \nu \iff \lambda \; R^S \; \mu$ for all $\lambda, \mu, \nu \in \tilde{Y}^S$. This would lead to the absurd conclusion that all the random consequences in $\tilde{Y}^S$ are indifferent, according to the preference ordering R^S.

7. THE SURE-THING PRINCIPLE FOR INDEPENDENT PROBABILITIES

Originally, Savage (1954) stated his sure-thing principle regarding preferences for contingent consequence functions in Y^S ($\emptyset \neq S \subset E$) rather than for probability distributions in $\tilde{Y}^S$. Savage's sure-thing principle requires that, for all non-empty disjoint $S_1, S_2 \subset E$, all $y^{S_1}, z^{S_1} \in Y^{S_1}$, and all $\bar{y}^{S_2} \in Y^{S_2}$,

$$(y^{S_1}, \bar{y}^{S_2})\ R^{S_1 \cup S_2}\ (z^{S_1}, \bar{y}^{S_2}) \iff y^{S_1}\ R^{S_1}\ z^{S_1}.$$

Since Anscombe and Aumann's (1963) fundamental contribution, it has been customary to extend this principle to all simple probability distributions $\lambda, \mu \in \tilde{Y}^{S_1}$ and $\nu \in \tilde{Y}^{S_2}$ so that

$$(\lambda, \nu)\ R^{S_1 \cup S_2}\ (\mu, \nu) \iff \lambda\ R^{S_1}\ \nu.$$

In the above statement, however, the joint distributions of the random vectors (y^{S_1}, y^{S_2}) are left unspecified — only the marginal distributions λ, μ, ν are given. Of course, Anscombe and Aumann (1963) assume in effect that only these marginal distributions are relevant. As Drèze (1985, 1986, 1987) points out, this means that the agent is indifferent whether the lotteries are resolved before or after the state of the world is known. Rather than make the Anscombe and Aumann assumption, it seems preferable to derive it, if possible, from consequentialism, and this will be done in due course under an extra structural condition. Meanwhile, the precise form of the joint distribution of the random variables of the vector y^S will be treated as a relevant consequence for normative behaviour.

In general, consequentialism implies only a weakened form of Anscombe and Aumann's version of the sure-thing principle, in which the random vectors y^{S_1}, y^{S_2} must be independent. This weakened from, called the *sure-thing principle for independent probabilities*, requires that for all non-empty disjoint $S_1, S_2 \subset E$ and all simple probability distributions $\lambda, \mu \in \tilde{Y}^{S_1}$, $\nu \in \tilde{Y}^{S_2}$:

$$(\lambda \times \nu)\ R^{S_1 \cup S_2}\ (\mu \times \nu) \iff \lambda\ R^{S_1}\ \mu.$$

The difference from Anscombe and Aumann is evident — only product distributions $\lambda \times \nu$, $\mu \times \nu$ satisfy the condition. On the other hand, Savage's original sure-thing principle is indeed an implication, when λ, μ, ν are degenerate distributions attaching probability one to the consequences y^{S_1}, z^{S_1} and $\bar{y}^{S_2}$ respectively.

The proof that consequentialism implies this principle relies upon constructing the decision tree $T(S_1, S_2, \lambda, \mu, \nu)$ as in Figure 7.1, for any pair of non-empty disjoint sets $S_1, S_2 \subset E$ and any $\lambda, \mu \in \tilde{Y}^{S_1}, \nu \in \tilde{Y}^{S_2}$, with $Z := \{\lambda, \mu, \nu\}$ and:

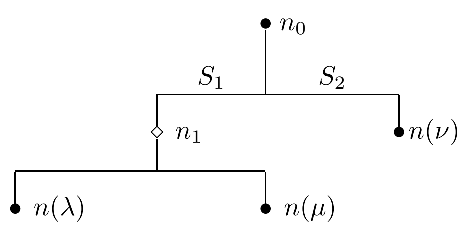

Figure 7.1. The Decision Tree $T(S_1, S_2, \lambda, \mu, \nu)$

(i) $N := \{ n_0, n_1 \} \cup \{ n(\zeta) \mid \zeta \in Z \}$;
(ii) $N^* := \{n_1\}$;
(iii) $N^0 := \emptyset$;
(iv) $N^1 := \{n_0\}$;
(v) $X := \{ n(\zeta) \mid \zeta \in Z \}$;
(vi) $N_{+1}(n_0) := \{ n_1, n(\nu) \}$ and $N_{+1}(n_1) := \{ n(\lambda), n(\mu) \}$;
(vii) n_0 the initial node;

(viii) $\pi(\cdot \mid \cdot)$ undefined;
(ix) $S(n_0) := S_1 \cup S_2$, $S(n_1) := S(n(\lambda)) := S(n(\mu)) := S_1$ and $S(n(\nu)) := S_2$;
(x) $\gamma(n(\zeta)) := \zeta$ (all $\zeta \in Z$).

THEOREM 7. *If the consistent behaviour norm β is consequentialist for the unrestricted domain of all consequential decision trees $T \in \mathcal{T}$, then for all pairs of non-empty disjoint events $S_1, S_2 \subset E$, the revealed preference orderings R^{S_1} and $R^{S_1 \cup S_2}$ must satisfy the sure-thing principle for independent probabilities.*

Proof. Construct $T := T(S_1, S_2, \lambda, \mu, \nu)$ as in Figure 7.1. Then

$$\begin{aligned} F(T, n_1) &= \{\lambda, \mu\} = F(T(n_1)) \\ F(T) &= \{\, \lambda, \mu \,\} \times \{\nu\} = \{\, \lambda \times \nu, \mu \times \nu \,\} \\ \Phi(T) &= \Phi(T, n_1) \times \{\nu\} \end{aligned}$$

by backward recursion. Let $S := S_1 \cup S_2$. Then

$$\begin{aligned} (\lambda \times \nu)\ R^S\ (\mu \times \nu) &\iff \lambda \times \nu \in \Phi(T) \iff \lambda \in \Phi(T, n_1) \\ &\iff n(\lambda) \in \beta(T, n_1) \iff n(\lambda) \in \beta(T(n_1), n_1) \\ &\iff \lambda \in \Phi(T(n_1)) \iff \lambda\ R^{S_1}\ \mu. \end{aligned}$$

∎

8. SUFFICIENCY OF ORDINALITY, INDEPENDENCE AND THE SURE-THING PRINCIPLE

This section shows that the three necessary conditions for consequentialism which were derived in Sections 5, 6 and 7 are also sufficient for the existence of a consequentialist norm.

THEOREM 8. *Suppose there are preference orderings R^S for all non-empty $S \subset E$ satisfying the independence axiom and the sure-thing principle for independent probabilities. Then there exists a consequentialist consistent behaviour norm β, defined on the unrestricted domain $\mathcal{T}$ of all finite consequential decision trees, for which the revealed preference orderings satisfy $R^S_\beta = R^S$ for all non-empty $S \subset E$.*

Proof.

(1) For the given collection of orderings R^S ($\emptyset \neq S \subset E$), construct the consequence choice function

$$C^S(Z) = \{\, \lambda^* \in Z \mid \lambda \in Z \Rightarrow \lambda^*\ R^S\ \lambda \,\}$$

which has a non-empty value for all non-empty finite $Z \subset \tilde{Y}^S$. Then, for all trees $T \in \mathcal{T}$ and for all nodes $n \in N(T)$, construct the non-empty set

$$F^*(T, n) := C^S(F(T, n)) = \{\, \lambda^* \in F(T, n) \mid \lambda \in F(T, n) \Rightarrow \lambda^*\ R^S\ \lambda \,\}.$$

(2) The construction by backward recursion of the sets $F(T, n)$ in Section 4 obviously implies that, whenever $\bar{n} \in N$ and $n \in N(\bar{n})$, then $F(T, n) = F(T(\bar{n}), n)$. So the above construction (1) implies also that

$$F^*(T, n) = F^*(T(\bar{n}), n).$$

(3) Now, given any tree $T \in \mathcal{T}$ and any decision node $n \in N^*(T)$, define the behaviour set

$$\beta(T, n) := \{n' \in N_{+1}(n) \mid F^*(T, n) \cap F(T, n') \neq \emptyset\}.$$

Because $F(T,n) = \bigcup_{n' \in N_{+1}(n)} F(T,n')$ and $F^*(T,n)$ is a non-empty subset of $F(T,n)$, the set $\beta(T,n)$ is a non-empty subset of $N_{+1}(n)$. Thus β is a behaviour norm.

(4) When $\bar{n} \in N$ and $n \in N^*(\bar{n})$, then by (2) and (3) above

$$\begin{aligned}\beta(T,n) &= \{n' \in N_{+1}(n) \mid F^*(T,n) \cap F(T,n') \neq \emptyset\} \\ &= \{n' \in N_{+1}(n) \mid F^*(T(\bar{n}),n) \cap F(T(\bar{n}),n') \neq \emptyset\} \\ &= \beta(T(\bar{n}),n)\end{aligned}$$

so that β is a consistent behaviour norm.

(5) It remains only to prove that

$$\Phi_\beta(T,n) = F^*(T,n) \tag{H}$$

at all nodes n of all trees T in $\mathcal{T}$. For then the construction of F^* in (1) will imply that the consistent behaviour norm β is indeed consequentialist, revealing the consequence choice functions $C^S_\beta = C^S$ and preference orderings $R^S_\beta = R^S$ for all non-empty $S \subset E$. The proof that (H) is true will be by backward induction.

(6) At every terminal node n of X one has

$$\Phi_\beta(T,n) = F^*(T,n) = \{\gamma(n)\}$$

so that (H) is clearly satisfied.

(7) Suppose then, as the backward induction hypothesis, that (H) is satisfied at every node of $N_{+1}(n)$, for some node n of a tree T in $\mathcal{T}$. The proof requires considering three obvious cases:

Case A. n is a decision node in N^*.

(8) At any $n \in N^*$ the definition of Φ_β, together with (3) and the induction hypothesis (7), implies that

$$\begin{aligned}\Phi_\beta(T,n) &= \bigcup_{n' \in \beta(T,n)} \Phi_\beta(T,n') \\ &= \{\, \lambda \mid \exists n' \in N_{+1}(n) : \; F^*(T,n) \cap F(T,n') \neq \emptyset \;\&\; \lambda \in F^*(T,n') \,\}.\end{aligned}$$

(9) If $\lambda^* \in \Phi_\beta(T,n)$ then (8) implies that there exists $n^* \in N_{+1}(n)$ for which both $\lambda^* \in F^*(T,n^*)$ and $F^*(T,n) \cap F(T,n^*) \neq \emptyset$. Let $\hat{\lambda}$ be a member of $F^*(T,n) \cap F(T,n^*)$. Then $\lambda^* \; R^{S(n^*)} \; \hat{\lambda}$ by definition of $F^*(T,n^*)$. But $S(n^*) = S(n)$ because n is a decision node and $n^* \in N_{+1}(n)$. Therefore $\lambda^* \; R^{S(n)} \; \hat{\lambda}$. But $\hat{\lambda} \in F^*(T,n)$ and so, for all $\lambda \in F(T,n)$, $\hat{\lambda} \; R^{S(n)} \; \lambda$ which implies that $\lambda^* \; R^{S(n)} \; \lambda$ because the preference relation $R^{S(n)}$ is transitive, by hypothesis. So it follows that $\lambda^* \in F^*(T,n)$ also.

(10) Conversely, suppose that $\lambda^* \in F^*(T,n)$. Then there must exist $\hat{n} \in N_{+1}(n)$ for which $\lambda^* \in F(T,\hat{n})$ because $F^*(T,n) \subset F(T,n) = \bigcup_{n' \in N_{+1}(n)} F(T,n')$.

(11) For all $\lambda \in F(T,\hat{n})$ one has $\lambda \in F(T,n)$ because $\hat{n} \in N_{+1}(n)$ and n is a decision node. By the hypothesis (10) and the definition of F^*, it follows that $\lambda^* \; R^{S(n)} \; \lambda$. Again, because n is a decision node, $S(\hat{n}) = S(n)$ and so $\lambda^* \; R^{S(\hat{n})} \; \lambda$ for all $\lambda \in F(T,\hat{n})$. Thus $\lambda^* \in F^*(T,\hat{n})$.

(12) From (10) and (11), $\lambda^* \in F^*(T,n) \cap F(T,\hat{n})$, so (8) implies $F^*(T,\hat{n}) \subset \Phi_\beta(T,n)$. But $\lambda^* \in F^*(T,\hat{n})$ by (11), so $\lambda^* \in \Phi_\beta(T,n)$.

(13) From (9), (10), (11) and (12) it follows that (H) is satisfied at n.

Case B. n is a chance node in N^0.

(14) At any $n \in N^0$ the definition of Φ_β, together with the induction hypothesis (7), implies

$$\Phi_\beta(T,n) = \sum_{n' \in N_{+1}(n)} \pi(n'|n) \Phi_\beta(T,n') = \sum_{n' \in N_{+1}(n)} \pi(n'|n) F^*(T,n').$$

(15) Given $\lambda, \lambda^* \in F(T, n)$ write $\lambda := \sum_{n' \in N_{+1}(n)} \pi(n'|n)\lambda(n')$ where $\lambda(n') \in F(T, n')$ (all $n' \in N_{+1}(n)$) and similarly for λ^*.

(16) If $\lambda^* \in \Phi_\beta(T, n)$ then, for all $\lambda \in F(T, n)$ and all $n' \in N_{+1}(n)$, one has $\lambda^*(n') \in F^*(T, n')$ and $\lambda(n') \in F(T, n')$, so $\lambda^*(n')\ R^{S(n')}\ \lambda(n')$. Because n is a chance node, $S(n') = S(n)$ for all $n' \in N_{+1}(n)$. Repeated application of the independence axiom implies that, for all $\lambda \in F(T, n)$,

$$\lambda^* = \sum_{n' \in N_{+1}(n)} \pi(n'|n)\lambda^*(n') \quad R^{S(n)} \quad \sum_{n' \in N_{+1}(n)} \pi(n'|n)\lambda(n) = \lambda.$$

Therefore $\lambda^* \in F^*(T, n)$ by definition of F^*.

(17) Conversely, suppose $\lambda^* \in F^*(T, n)$. Take any $\hat{n} \in N_{+1}(n)$ and any $\lambda(\hat{n}) \in F(T, \hat{n})$. Define $\hat{\lambda} \in F(T, n)$ by

$$\hat{\lambda} := \pi(\hat{n}|n)\lambda(\hat{n}) + \sum_{n' \in N_{+1}(n) \setminus \{\hat{n}\}} \pi(n'|n)\lambda^*(n').$$

Then $\lambda^*\ R^{S(n)}\ \hat{\lambda}$ by definition of F^*. Since $\pi(\hat{n}|n) > 0$ because of the assumption that all probabilities are positive at all chance nodes, the independence axiom implies that $\lambda^*(\hat{n})\ R^{S(n)}\ \lambda(\hat{n})$. This is true for any $\lambda(\hat{n}) \in F(T, \hat{n})$ and so $\lambda^*(\hat{n}) \in F^*(T, \hat{n})$. Because this is true for all $\hat{n} \in N_{+1}(n)$, it follows from (14) and (15) that $\lambda^* \in \Phi_\beta(T, n)$.

(18) From (16) and (17) it follows that (H) is satisfied at n.

Case C. n is a natural node in N^1.

(19) At any $n \in N^1$ the definition of Φ_β, together with the induction hypothesis (7), implies

$$\Phi_\beta(T, n) = \prod_{n' \in N_{+1}(n)} \Phi_\beta(T, n') = \prod_{n' \in N_{+1}(n)} F^*(T, n').$$

(20) Given $\lambda, \lambda^* \in F(T, n)$, write λ in the product form $\prod_{n' \in N_{+1}(n)} \lambda(n')$ where $\lambda(n') \in F(T, n')$ (all $n' \in N_{+1}(n)$), and similarly for λ^*.

(21) If $\lambda^* \in \Phi_\beta(T, n)$ then, for all $\lambda \in F(T, n)$ and all $n' \in N_{+1}(n)$, one has $\lambda^*(n') \in F^*(T, n')$ and $\lambda(n') \in F(T, n')$, so $\lambda^*(n')\ R^{S(n')}\ \lambda(n')$. Because the collection $\{\, S(n') \mid n' \in N_{+1}(n) \,\}$ is a partition of $S(n)$, repeated application of the sure-thing principle for independent probabilities implies that, for all $\lambda \in F(T, n)$,

$$\lambda^* = \prod_{n' \in N_{+1}(n)} \lambda^*(n') \quad R^{S(n)} \quad \prod_{n' \in N_{+1}(n)} \lambda(n') = \lambda.$$

Therefore $\lambda^* \in F^*(T, n)$.

(22) Conversely, suppose $\lambda^* \in F^*(T, n)$. Take any $\hat{n} \in N_{+1}(n)$ and any $\lambda(\hat{n}) \in F(T, \hat{n})$. Now define $\hat{\lambda} \in F(T, n)$ by

$$\hat{\lambda} := \lambda(\hat{n}) \times \left\{ \prod_{n' \in N_{+1}(n) \setminus \{\hat{n}\}} \lambda^*(n') \right\}.$$

Then $\lambda^*\ R^{S(n)}\ \hat{\lambda}$ by definition of F^*. By the sure-thing principle for independent probabilities, this implies that $\lambda^*(\hat{n})\ R^{S(\hat{n})}\ \lambda(\hat{n})$. This is true for any $\lambda(\hat{n}) \in F(T, \hat{n})$ and so $\lambda^*(\hat{n}) \in F^*(T, \hat{n})$. Since this is true for all $\hat{n} \in N_{+1}(n)$, it follows from (19) and (20) that $\lambda^* \in \Phi_\beta(T, n)$.

(23) From (21) and (22) it follows that (H) is satisfied at n.

(24) Thus (H) is satisfied at all n, by backward induction. ∎

In combination with the earlier results of Sections 5, 6 and 7, Theorem 8 gives a complete characterization of consequentialist behaviour in finite consequential decision trees with all probabilities positive. The proof relies on constructing β as in step (3). Obviously a form of dynamic programming is being used in which each node n of the tree T is given a "value" in the sense of a set $\Phi_\beta(T, n) = F^*(T, n)$ of consequences which are all indifferent according to the preference ordering $R^{S(n)}$.

9. CONTINUITY AND CONDITIONAL EXPECTED UTILITY

The characterization of consequentialist behaviour in Theorem 8 allows general preference orderings satisfying independence and the sure-thing principle for independent lotteries. Discontinuous preferences, which only admit lexicographic expected utility representations, are not excluded — see Hausner (1954), Chipman (1960), Skala (1975) and Fishburn (1982). In the spaces $\tilde{Y}^S (S \subset E)$, such representations may also generate lexicographic hierarchies of subjective probabilities, as considered recently by Blume (1986) and Brandenburger and Dekel (1986). Here, however, such discontinuous preferences will be excluded by an additional assumption. As in Herstein and Milnor's (1953) simplification of the original von Neumann and Morgenstern (1944) axioms, this could be the following

CONTINUITY ASSUMPTION: *If $\emptyset \neq S \subset E$ and $\lambda, \mu, \nu \in \tilde{Y}^S$ with $\lambda \; P^S \; \mu$ and $\mu \; P^S \; \nu$, then there exist α and β with $0 < \alpha < \beta < 1$ such that*

$$[(1-\alpha)\lambda + \alpha\nu] \; P^S \; \mu \quad \textit{and} \quad \mu \; P^S \; [(1-\beta)\lambda + \beta\nu].$$

As Herstein and Milnor show, this assumption, together with the fact that R^S is a preference ordering satisfying the independence axiom, implies that there is a real-valued von Neumann-Morgenstern utility function (NMUF) v^S on Y^S such that, for all $\lambda, \mu \in \tilde{Y}^S$,

$$\lambda \; R^S \; \mu \iff \mathbb{E}_\lambda \, v^S \geq \mathbb{E}_\mu \, v^S$$

where $\mathbb{E}_\lambda \, v^S := \sum_{y^S \in Y^S} \lambda(y^S) v^S(y^S)$ denotes the expected value of v^S with respect to λ, and similarly $\mathbb{E}_\mu \, v^S$.

The two NMUF's v^S, $\tilde{v}^S$ on Y^S are said to be *cardinally equivalent* if there exist real numbers $\rho > 0$ and α such that $\tilde{v}(y^S) \equiv \alpha + \rho v(y^S)$ on Y^S. Then, as is well known, there is a unique cardinal equivalence class of NMUF's whose expected values represent R^S.

Rather than assume continuity of preferences, it is more in the spirit of this paper to postulate continuity of behaviour, whether or not it is consequentialist. Consider a family of decision trees T^π in which only the probabilities $\pi(n'|n)$ vary, for all $n \in N^0$ and $n' \in N_{+1}(n)$ — the other features of the decision tree are all independent of π. Consider then each behaviour set $\beta(T^\pi, n)$ as π varies with $n \in N^*$ fixed. This gives a correspondence whose graph, for each $n \in N^*$, is

$$G_\beta(n) := \{ \, (\pi, n') \mid n' \in \beta(T^\pi, n) \, \}.$$

Since N^0 is finite, the product set $\prod_{n \in N^0} \; \Delta(N_{+1}(n))$ of independent probability distributions $\pi(\cdot|n)$ $(n \in N^0)$ is evidently compact. Since the set $N_{+1}(n)$ is finite, the above correspondence will have a compact graph, and so be upper hemi-continuous, provided $G_\beta(n)$ is a closed set for all $n \in N^*$. So the behaviour norm β is said to be *continuous* provided that the graph $G_\beta(n)$ defined above is indeed closed, for every $T \in \mathcal{T}$ and $n \in N^*(T)$.

This definition applies to general behaviour norms. There is a technical problem on the boundary of the set $\prod_{n \in N^0} \; \Delta(N_{+1}(n))$ for consequentialist behaviour norms. If $\pi(n'|n)$ becomes zero for some $n \in N^0$ and $n' \in N_{+1}(n)$, then T^π is not properly in the domain of a consequentialist behaviour norm. But then, in defining $\beta(T^\pi, n^*)$ at any decision node $n^* \in N^*$ which is succeeded some time later by n and then immediately thereafter by n', one can replace T^π by a decision tree with set of nodes $N \setminus N(n')$ — i.e., all nodes which must occur with probability zero are simply removed from the tree.

A consequentialist continuous behaviour norm does indeed give rise to revealed preference orderings R^S satisfying the above continuity assumption. To see this, for any given non-empty $E \subset S$, and $\lambda, \mu, \nu \in \tilde{Y}^S$ with $\lambda \, P^S \, \mu$ and $\mu \, P^S \, \nu$, construct the family $T^\delta(\lambda, \mu, \nu)$ $(0 < \delta < 1)$ of decision trees defined by

$$T^\delta := T^\delta(\lambda, \mu, \nu) := \langle N, N^*, N^0, N^1, X, N_{+1}(\cdot), n_0, \pi^\delta(\cdot|\cdot), S(\cdot), \gamma(\cdot) \rangle$$

as shown in Figure 9.1, where $Z := \{\lambda, \mu, \nu\}$ and with:

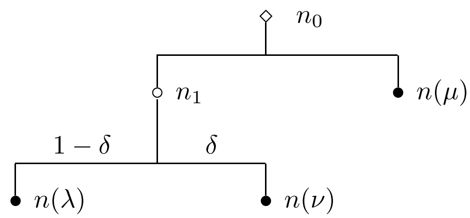

Figure 9.1. The Decision Tree $T^\delta(\lambda, \mu, \nu)$

(i) $N := \{n_0, n_1\} \cup \{\, n(\zeta) \mid \zeta \in Z \,\}$;
(ii) $N^* := \{n_0\}$;
(iii) $N^0 := \{n_1\}$;
(iv) $N^1 := \emptyset$;
(v) $X := \{\, n(\zeta) \mid \zeta \in Z \,\}$;
(vi) $N_{+1}(n_0) := \{n_1, n(\mu)\}$ and $N_{+1}(n_1) := \{n(\lambda), n(\nu)\}$;
(vii) n_0 the initial node;
(viii) $\pi^\delta(n(\lambda)|n_1) := 1 - \delta$ and $\pi^\delta(n(\nu)|n_1) := \delta$;
(ix) $S(n) := S$ for all $n \in N$;
(x) $\gamma(n(\zeta)) := \zeta$ for all $\zeta \in Z$.

When $\delta = 0$ or 1, the node $n(\nu)$ or $n(\lambda)$, respectively, is omitted from T^δ. Because $\lambda\ P^S\ \mu$ and $\mu\ P^S\ \nu$, the behaviour norm satisfies $\beta(T^0, n_0) = \{n_1\}$ when $\delta = 0$ and $\beta(T^1, n_0) = \{n(\mu)\}$ when $\delta = 1$. But, by continuity, the behaviour correspondence $\beta(T^\delta, n_0)$ has a closed graph as δ varies. Thus the two sets :

$$D_1 := \{\, \delta \in [0,1] \mid n_1 \quad \in \beta(T^\delta, n_0) \,\}$$
$$D_2 := \{\, \delta \in [0,1] \mid n(\mu) \in \beta(T^\delta, n_0) \,\}$$

are both closed subsets of $[0,1]$ with $0 \in D_1 \setminus D_2$, $1 \in D_2 \setminus D_1$, and $D_1 \cup D_2 = [0,1]$. So the two points $\delta_1 := \max\{\, \delta \mid \delta \in D_1 \,\}$ and $\delta_2 := \min\{\, \delta \mid \delta \in D_2 \,\}$ are both well defined, with $0 < \delta_2 \leq \delta_1 < 1$. It follows that there must exist α, β satisfying $0 < \alpha < \delta_2 \leq \delta_1 < \beta < 1$. But then $\alpha \in D_1 \setminus D_2$ and $\beta \in D_2 \setminus D_1$, which implies that

$$[(1-\alpha)\lambda + \alpha\nu]\ P^S\ \mu \text{ and } \mu\ P^S\ [(1-\beta)\lambda + \beta\nu],$$

so verifying that R^S is continuous. To summarize

THEOREM 9. *If a consistent behaviour norm β is both continuous and consequentialist for the domain of all consequential decision trees $T \in \mathcal{T}$, then for each non-empty $S \subset E$ there is a unique cardinal equivalence class of NMUF's $v^S(\cdot) : Y^S \to \mathbb{R}$ such that behaviour in any tree $T \in \mathcal{T}(S)$ maximizes conditional expected utility $\mathbb{E}v^S$.*

10. EXPECTED UTILITY SUMS AND PRODUCTS

Theorem 9 just uses continuity, the existence of a preference ordering, and the independence axiom. Another implication of consequentialism, the sure-thing principle for independent probabilities, relates the NMUF's v^S for different events S to each other. Indeed, under the hypotheses of Theorem 9 it will be shown that when $v_s : Y_s \to \mathbb{R}$ is defined as $v^{\{s\}}$ for each $s \in E$, then, for every non-empty $S \subset E$, v^S can be expressed in one of the following three forms, after suitable normalization:

(i) $v^S(y^S) \equiv \sum_{s \in S} v_s(y_s)$;
(ii) $v^S(y^S) \equiv \prod_{s \in S} v_s(y_s)$ (with $v_s(y_s) > 0$ everywhere);
(iii) $v^S(y^S) \equiv -\prod_{s \in S} [-v_s(y_s)]$ (with $v_s(y_s) < 0$ everywhere).

Form (i) is familiar from existing work, as mentioned in the introduction, but forms (ii) and (iii) are novel. Notice that all three forms are consistent with having $v^{\{s\}} = v_s$ for all $s \in E$. Notice too that the products

(ii) and (iii) allow the sure-thing principle to hold for independent probabilities, since then the expected value of the utility product is equal to the product of expected utilities.

Some new notation will facilitate the argument. Write μ^S for the product $\prod_{s\in S}\mu_s$, a typical member of the product space $\prod_{s\in S}\tilde{Y}_s$ of independent probability distributions. Write $U^E(\mu^E)$ for the expected utility expression $\mathbb{E}_{\mu^E} v^E(y^E)$. Choose any fixed $\bar{\mu}^E$ in $\prod_{s\in E}\tilde{Y}_s$ and then normalize v^E, and so U^E also, in order to satisfy $U^E(\bar{\mu}^E)=0$. Now define the function U^S, for any non-empty $S\subset E$, by $U^S(\mu^S) := U^E(\mu^S\times\bar{\mu}^{E\setminus S})$. Of course $U^S(\bar{\mu}^S)=U^E(\bar{\mu}^S\times\bar{\mu}^{E\setminus S})=0$, by the normalization.

The proof will involve considering non-empty sets $S\subset E$ which can be partitioned into m non-empty disjoint sets S_j ($j\in J:=\{1,2,\ldots,m\}$). For each $j\in J$, write μ_j for μ^{S_j} and $U_j(\mu_j)$ for $U^{S_j}(\mu^{S_j})$. Also write $U(\prod_{j\in J}\mu_j)$ for $U^S(\prod_{j\in J}\mu^{S_j})$ and write $U_{-j}(\prod_{k\in J\setminus\{j\}}\mu_k)$ for $U^{S\setminus S_j}(\prod_{k\in J\setminus\{j\}}\mu^{S_k})$. Often, too, the arguments μ_j etc. of the expected utility functions will be omitted.

LEMMA 10.1. *When $m=2$ there exists a constant ρ (positive, negative, or zero) such that $U\equiv U_1+U_2+\rho U_1U_2$.*

Proof. Suppose $\{j,k\}=\{1,2\}$. By Theorem 7 the sure-thing principle for independent probabilities is satisfied. So both $U(\mu_j\times\mu_k)$ and $U_j(\mu_j)$ represent the same preference ordering R^{S_j} on the set $\prod_{s\in S_j}\tilde{Y}_s$ of independent probability distributions in $\tilde{Y}^{S_j}$, for any fixed $\mu_k\in\prod_{s\in S_k}\tilde{Y}_s$. So the two utility functions $U(\mu_j\times\mu_k)$ and $U_j(\mu_j)$ are cardinally equivalent, for each such fixed μ_k. That implies the existence of $\rho_j(\mu_k)>0$ and $\alpha_j(\mu_k)$, both independent of μ_j, for which

$$U(\mu_j\times\mu_k)\equiv\alpha_j(\mu_k)+\rho_j(\mu_k)U_j(\mu_j).$$

In particular, when $\mu_j=\bar{\mu}_j$, then $U_j(\bar{\mu}_j)=0$ by our normalization, and so $\alpha_j(\mu_k)\equiv U(\bar{\mu}_j\times\mu_k)\equiv U_k(\mu_k)$. It follows that

$$\begin{aligned}U(\mu_j\times\mu_k)&\equiv U_k(\mu_k)+\rho_j(\mu_k)U_j(\mu_j)\\&\equiv U_j(\mu_j)+\rho_k(\mu_j)U_k(\mu_k),\end{aligned}$$

because j and k can be interchanged throughout the above argument. But then

$$[\rho_j(\mu_k)-1]U_j(\mu_j)\equiv[\rho_k(\mu_j)-1]U_k(\mu_k).$$

So, when both $U_j(\mu_j)$ and $U_k(\mu_k)$ are not zero, then

$$\frac{[\rho_k(\mu_j)-1]}{U_j(\mu_j)}\equiv\frac{[\rho_j(\mu_k)-1]}{U_k(\mu_k)}\equiv\rho$$

for some number ρ which is independent of both μ_j and μ_k, and which can be positive, negative, or zero. This leads to

$$U(\mu_j\times\mu_k)\equiv U_j(\mu_j)+U_k(\mu_k)+\rho U_j(\mu_j)U_k(\mu_k)$$

which is obviously true even when $U_j(\mu_j)=0$ (or $U_k(\mu_k)=0$), because then

$$U(\mu_j\times\mu_k)=U(\bar{\mu}_j\times\mu_k)=U_k(\mu_k).$$

This completes the proof. ∎

LEMMA 10.2. *For all m there exists a constant ρ (positive, negative, or zero) such that*

$$U\equiv\sum_{\{K|\emptyset\neq K\subset J\}}\rho^{\#K-1}\prod_{k\in K}U_k.$$

Proof. The proof will be by induction on m, the number of members of the set J. When $m = 1$, the equation becomes $U \equiv U_1$ and is satisfied trivially. When $m = 2$ it becomes $U = U_1 + U_2 + \rho U_1 U_2$ which was proved in Lemma 10.1. So, as an induction hypothesis, suppose it is true for $m - 1$. Then, for any $j \in J$, one has

$$U_{-j} \equiv \sum\nolimits_{\{ K | \emptyset \neq K \subset J \setminus \{j\}\}} (\rho_{-j})^{\#K-1} \prod\nolimits_{k \in K} U_k.$$

But also, applying Lemma 10.1 to the partition $S_j \cup S_{-j}$ of S leads to

$$U \equiv U_j + U_{-j} + \rho_j U_j U_{-j}.$$

Substituting for U_{-j} from the previous formula gives

$$U \equiv \sum\nolimits_{\{ K | \emptyset \neq K \subset J\}} \rho^*_K \prod\nolimits_{k \in K} U_k$$

where

$$\rho^*_K := \begin{cases} (\rho_{-j})^{\#K-1} & \text{if } K \subset J \setminus \{j\} \\ \rho_j (\rho_{-j})^{\#K-2} & \text{if } j \in K \subset J \text{ and } \#K \geq 2 \\ 1 & \text{if } K = \{j\}. \end{cases}$$

But the choice of j was arbitrary, so the above expression for ρ^*_K must be independent of j. When $K = \{j, k\}$ this implies that $\rho^*_{\{j,k\}} = \rho_j = \rho_k$ for all pairs $j, k \in J$, and also that $\rho^*_{\{j,k\}} = \rho_{-\ell}$ for all $\ell \in J \setminus \{j, k\}$. Thus for all $j \in J$, $\rho_j = \rho_{-j} = \rho$, independent of j. So $\rho^*_K = \rho^{\#K-1}$ and the lemma is also true for m. This completes the proof by induction on m. ∎

A more convenient expression of U in terms of $U_j (j \in J)$ is as follows. When $\rho = 0$ one has

$$U \equiv \sum\nolimits_{j \in J} U_j$$

because $0^0 = 1$ and $0^m = 0$ for $m = 1, 2, \ldots$. When $\rho \neq 0$ one has

$$1 + \rho U \equiv \prod\nolimits_{j \in J} (1 + \rho U_j),$$

as can be verified by direct expansion, or proved by induction on $m = \#J$. Since E can be partitioned into the subsets $\{\{s\} \mid s \in E\}$, one has:

$$\begin{aligned} U^E(\mu^E) &\equiv \textstyle\sum_{s \in E} U_s(\mu_s) && \text{if } \rho = 0; \\ 1 + \rho\, U^E(\mu^E) &\equiv \textstyle\prod_{s \in E} [1 + \rho\, U_s(\mu_s)] && \text{if } \rho \neq 0. \end{aligned}$$

When $\mu^{E \setminus S} = \bar{\mu}^{E \setminus S}$ and so $U_s(\mu_s) = 0$ (all $s \in E \setminus S$), one has:

$$\begin{aligned} U^S(\mu^S) &\equiv \textstyle\sum_{s \in S} U_s(\mu_s) && \text{if } \rho = 0; \\ 1 + \rho\, U^S(\mu^S) &\equiv \textstyle\prod_{s \in S} [1 + \rho\, U_s(\mu_s)] && \text{if } \rho \neq 0. \end{aligned}$$

The forms stated at the head of this section can now be derived by renormalization of the utilities v^S and expected utilities U^S. Indeed, if $\rho = 0$, no renormalization is necessary, because $U^S \equiv \sum_{s \in S} U_s$ leads immediately to case (i), which is

$$v^S(y^S) \equiv \sum\nolimits_{s \in S} v_s(y_s).$$

If $\rho > 0$, renormalize by taking, for all non-empty $S \subset E$:

$$\tilde{U}^S \equiv 1 + \rho U^S; \quad \tilde{v}^S \equiv 1 + \rho v^S.$$

Then, with a slight abuse of notation, the renormalized utilities satisfy case (ii), which is

$$v^S(y^S) \equiv \prod\nolimits_{s \in S} v_s(y_s).$$

Finally, if $\rho < 0$, renormalize by taking, for all non-empty $S \subset E$:

$$\tilde{U}^S \equiv -1 - \rho U^S; \quad \tilde{v}^S \equiv -1 - \rho v^S.$$

Then, again with a slight abuse of notation, the renormalized utilities satisfy case (iii), which is

$$v^S(y^S) \equiv -\prod_{s \in S} [-v_s(y_s)].$$

By the sure-thing principle for independent lotteries, the product $\prod_{s \in S} v_s$ must be increasing in each v_s in case (ii), and the negative product $-\prod_{s \in S}(-v_s)$ must be increasing in each v_s in case (iii). This requires that $v_s(y_s) > 0$ everywhere in case (ii), and that $v_s(y_s) < 0$ everywhere in case (iii).

The above sums and products depend upon appropriate normalizations of each utility function v^S. If each v^S is replaced by the cardinally equivalent

$$\tilde{v}^S(y^S) \equiv \alpha^S + \rho^S v^S(y^S)$$

with $\rho^S > 0$, one has

$$\tilde{v}^S(y^S) \equiv \alpha^S + \rho^S \sum_{s \in S} v_s(y_s) \equiv \alpha^S + \rho^S \sum_{s \in S} [\tilde{v}_s(y_s) - \alpha_s]/\rho_s$$

in case (i), so the simple summation $\tilde{v}^S(y^S) \equiv \sum_{s \in S} \tilde{v}_s(y_s)$ remains valid only if both $\rho^S = \rho$ and $\alpha^S = \sum_{s \in S} \alpha_s$ for all non-empty $S \subset E$. In case (ii) one has

$$\tilde{v}^S(y^S) \equiv \alpha^S + \rho^S \prod_{s \in S} [\tilde{v}_s(y_s) - \alpha_s]/\rho_s \ ,$$

so the simple product $\tilde{v}^S(y^S) \equiv \prod_{s \in S} \tilde{v}_s(y_s)$ remains valid only if both $\alpha^S = 0$ and $\rho^S = \prod_{s \in S} \rho_s$ for all non-empty $S \subset E$. Case (iii) is like case (ii).

THEOREM 10.3. *Suppose that a consistent behaviour norm β is continuous and consequentialist for the domain of all consequential decision trees $T \in \mathcal{T}$. Then, for each non-empty $S \subset E$ and each $s \in E$, there are NMUF's $v^S : Y^S \to I\!R$ and $v_s : Y_s \to I\!R$ such that behaviour in any $T \in \mathcal{T}(S)$ maximizes conditional expected utility $I\!E v^S$, and such that one of the following three cases is true:*

(i) $v^S(y^S) \equiv \sum_{s \in S} v_s(y_s)$ *and the NMUF's $v_s(\cdot)$ in all states $s \in E$ are unique up to co-cardinal transformations of the form*

$$\tilde{v}_s(y_s) \equiv \alpha_s + \rho v_s(y_s)$$

with $\rho > 0$ independent of s;

(ii) $v^S(y^S) \equiv \prod_{s \in S} v_s(y_s)$*, where each $v_s(\cdot)$ takes only positive values, and the NMUF's $v_s(\cdot)$ in all states $s \in E$ are unique up to independent unit transformations of the form*

$$\tilde{v}_s(y_s) \equiv \rho_s v_s(y_s)$$

with $\rho_s > 0$ for all $s \in S$;

(iii) $v^S(y^S) \equiv -\prod_{s \in S} [-v_s(y_s)]$*, where each $v_s(\cdot)$ takes only negative values, and the NMUF's $v_s(\cdot)$ in all states $s \in E$ are unique up to the same equivalence class as in (ii).*

This gives a complete characterization of continuous consequentialist behaviour norms because an analogy to Theorem 8 can be proved. The dynamic programming rules used to prove that result become very much simpler because each node n of a consequential decision tree T can be given an expected utility value $w(T, n)$, using backward induction. Indeed, at any terminal node $n \in X$,

$$w(T, n) = I\!E_{\gamma(n)} \ v^{S(n)}(y^{S(n)}).$$

Then, at any decision node $n \in N^*$,

$$\left.\begin{array}{r} w(T,n) = \\ \beta(T,n) = \arg \end{array}\right\} \max_{n'} \{ \, w(T,n') \mid n' \in N_{+1}(n) \, \}.$$

At any chance node $n \in N^0$,

$$w(T,n) = \sum\nolimits_{n' \in N_{+1}(n)} \pi(n'|n) \, w(T,n').$$

At any natural node $n \in N^1$, the rule for calculating the value $w(T,n)$ from the list $\langle w(T,n') \rangle_{n' \in N_{+1}(n)}$ depends upon which of the three possible cases applies:

$$w(T,n) = \begin{cases} \sum_{n' \in N_{+1}(n)} w(T,n') & \text{in case (i);} \\ \prod_{n' \in N_{+1}(n)} w(T,n') & \text{in case (ii);} \\ -\prod_{n' \in N_{+1}(n)} [-w(T,n')] & \text{in case (iii).} \end{cases}$$

Now one has

THEOREM 10.4. *Suppose there are preference orderings R^S for all non-empty $S \subset E$ which are represented by expected utilities $\mathbb{E}v^S$ satisfying one of the three cases of Theorem 10.3. Then there exists a consistent consequentialist continuous behaviour norm β, defined on the domain $\mathcal{T}$ of all consequential decision trees, for which $R^S_\beta = R^S$ (all $S \subset E$).*

Proof. As in the proof of Theorem 8 it is easily shown that, when β is constructed as above to maximize $w(T,n')$ over $n' \in N_{+1}(n)$ at each decision node $n \in N^*$, then, for every $n \in N$,

$$\left.\begin{array}{r} w(T,n) = \\ \Phi_\beta(T,n) = \arg \end{array}\right\} \max_{\lambda} \{ \, \mathbb{E}_\lambda \, v^{S(n)} \mid \lambda \in F(T,n) \, \}.$$

The proof does, of course, rely on the fact that the sure-thing principle for independent lotteries is satisfied. This is true, however, because if $\lambda, \mu \in \tilde{Y}^{S_1}$, $\nu \in \tilde{Y}^{S_2}$, $S_1 \cap S_2 = \emptyset$, and $S_1 \cup S_2 = S$, then

$$\begin{aligned} \lambda \; R^{S_1} \; \mu \; &\Longleftrightarrow \; \mathbb{E}_\lambda \, v^{S_1} \geq \mathbb{E}_\mu \, v^{S_1} \\ (\lambda \times \nu) \; R^S \; (\mu \times \nu) \; &\Longleftrightarrow \; \mathbb{E}_{\lambda \times \nu} \, v^S \geq \mathbb{E}_{\mu \times \nu} \, v^S \end{aligned}$$

Also, in case (i), when $v^S = v^{S_1} + v^{S_2}$, then

$$\begin{aligned} \mathbb{E}_\lambda \, v^{S_1} \geq \mathbb{E}_\mu \, v^{S_1} \; &\Longleftrightarrow \; \mathbb{E}_\lambda \, v^{S_1} + \mathbb{E}_\nu \, v^{S_2} \geq \mathbb{E}_\mu \, v^{S_1} + \mathbb{E}_\nu \, v^{S_2} \\ &\Longleftrightarrow \; \mathbb{E}_{\lambda \times \nu} \, (v^{S_1} + v^{S_2}) \geq \mathbb{E}_{\mu \times \nu} \, (v^{S_1} + v^{S_2}) \\ &\Longleftrightarrow \; \mathbb{E}_{\lambda \times \nu} \, v^S \geq \mathbb{E}_{\mu \times \nu} \, v^S. \end{aligned}$$

In case (ii), when $v^S = v^{S_1} v^{S_2}$ with v^{S_1} and v^{S_2} both positive, then

$$\begin{aligned} \mathbb{E}_\lambda \, v^{S_1} \geq \mathbb{E}_\mu \, v^{S_1} \; &\Longleftrightarrow \; [\mathbb{E}_\lambda \, v^{S_1}][\mathbb{E}_\nu \, v^{S_2}] \geq [\mathbb{E}_\mu \, v^{S_1}][\mathbb{E}_\nu \, v^{S_2}] \\ &\Longleftrightarrow \; \mathbb{E}_{\lambda \times \nu} \, v^{S_1} v^{S_2} \geq \mathbb{E}_{\mu \times \nu} \, v^{S_1} v^{S_2} \\ &\Longleftrightarrow \; \mathbb{E}_{\lambda \times \nu} \, v^S \geq \mathbb{E}_{\mu \times \nu} \, v^S. \end{aligned}$$

In case (iii), when $v^S = -[-v^{S_1}][-v^{S_2}]$ with v^{S_1} and v^{S_2} both negative, then

$$\begin{aligned} \mathbb{E}_\lambda \, v^{S_1} \geq \mathbb{E}_\mu \, v^{S_1} \; &\Longleftrightarrow \; -[\mathbb{E}_\lambda \, (-v^{S_1})][\mathbb{E}_\nu \, (-v^{S_2})] \geq -[\mathbb{E}_\mu \, (-v^{S_1})][\mathbb{E}_\nu \, (-v^{S_2})] \\ &\Longleftrightarrow \; -\mathbb{E}_{\lambda \times \nu} \, (-v^{S_1}) \, (-v^{S_2}) \geq -\mathbb{E}_{\mu \times \nu} \, (-v^{S_1}) \, (-v^{S_2}) \\ &\Longleftrightarrow \; \mathbb{E}_{\lambda \times \nu} \, v^S \geq \mathbb{E}_{\mu \times \nu} \, v^S. \end{aligned}$$

Thus in all three cases $\lambda\ R^{S_1}\ \mu \iff (\lambda \times \nu)\ R^S\ (\mu \times \nu)$ as in the sure-thing principle for independent probabilities. As for continuity, it is easy to check by backward induction that the utilities $w(T, n)$ and the behaviour sets $\beta(T, n)$ $(n \in N^*)$ all vary continuously with probabilities. ∎

Finally, Anscombe and Aumann (1963) derived additive utility from the assumption that it did not matter whether states of the world were determined before chance moves, or *vice versa*. One might therefore surmise that multiplicative probabilities arise when the agent definitely prefers that states be resolved before chance moves, or *vice versa*, and that the two multiplicative cases can be distinguished by their preferences regarding when uncertainty is resolved. This is indeed true when uncertainty simply concerns which of the two events S_1, S_2 occurs, where $\{S_1, S_2\}$ is a partition of the given non-empty $S \subset E$. For if chance moves are resolved first, the consequence can be any $\lambda \in \tilde{Y}^S$. But if uncertainty regarding S_1 or S_2 is resolved first, λ must be replaced by the corresponding consequence $\lambda^* = \lambda_1^* \times \lambda_2^*$ where, whenever the set $\{j, k\} = \{1, 2\}$, then $\lambda_j^* \in \tilde{Y}^{S_j}$ is the marginal probability distribution given by

$$\lambda_j^*(y^{S_j}) = \lambda\left(\{\, (y^{S_j}, \bar{y}^{S_k}) \mid \bar{y}^{S_k} \in Y^{S_k} \,\}\right).$$

Now consider the three possible forms for v^S in terms of v^{S_1} and v^{S_2}. In the additive case (i), one has

$$\mathbb{E}_{\lambda^*} v^S = \mathbb{E}_{\lambda_1^*} v^{S_1} + \mathbb{E}_{\lambda_2^*} v^{S_2} = \mathbb{E}_\lambda (v^{S_1} + v^{S_2}) = \mathbb{E}_\lambda v^S$$

so that λ^* and λ are indeed indifferent. In the multiplicative case (ii) with positive utilities, independence of the two distributions λ_j^* $(j = 1, 2)$ implies

$$\begin{aligned} \mathbb{E}_{\lambda^*} v^S &= \mathbb{E}_{\lambda^*} (v^{S_1} \times v^{S_2}) = [\mathbb{E}_{\lambda_1^*} v^{S_1}]\, [\mathbb{E}_{\lambda_2^*} v^{S_2}] \\ &= [\mathbb{E}_\lambda v^{S_1}]\, [\mathbb{E}_\lambda v^{S_2}] \quad \gtrless \quad \mathbb{E}_\lambda (v^{S_1} \times v^{S_2}) = \mathbb{E}_\lambda v^S \end{aligned}$$

according as the random variables y^{S_1}, y^{S_2} have negatively or positively correlated utilities. In particular, if y^{S_1} and y^{S_2} are perfectly correlated, then λ is preferred to λ^*, and the agent would rather not have uncertainty about S_1 or S_2 resolved before any chance moves occur. In the multiplicative case (iii), with negative utilities, the products all carry a minus sign, so the inequalities are reversed, and $\mathbb{E}_{\lambda^*}\ v^S \gtrless \mathbb{E}_\lambda\ v^S$ according as the random variables y^{S_1}, y^{S_2} have positively or negatively correlated utilities. In particular, if y^{S_1} and y^{S_2} are perfectly correlated, then λ^* is preferred to λ, and the agent would like to have uncertainty about S_1 or S_2 resolved before any chance moves occur. Later, in Section 12, the three cases will also be distinguished by the preferences they imply in Ellsberg's paradox.

11. CONSTANT CONSEQUENCES AND STATE INDEPENDENCE

Up to here the consequence domains Y_s $(s \in S)$ have been entirely general. From now on, following Savage (1954), it will be assumed that there is a *constant consequence domain* Y with $Y_s = Y$ (all $s \in E$). Then y^S is said to be a *constant consequence* if there exists $y \in Y$ with $y_s^S = y$ (all $s \in S$). In this case y^S is written as $y1^S$. The set of all constant consequences is written as $Y1^S$.

Let T be any consequential decision tree in $\mathcal{T}$ which has no chance nodes. Suppose that every consequence in $F(T)$ is a constant consequence. Then all possible behaviour also has constant consequences. It is as if there were no uncertainty at all because each constant consequence $y1^S$ is effectively the same, no matter what the state of the world s may be. The same is true if T has chance nodes but no natural nodes, and if $\gamma(n)$ is a constant consequence for every terminal node $n \in X$. Only $F(T)$ will then consist of probability distributions $\lambda \in \tilde{Y}^S$ with the property that constant consequences occur with probability one — i.e., consequences in different states are perfectly correlated. Thus $F(T)$ will be a subset of $\tilde{Y}1^S$, where $\lambda \in \tilde{Y}1^S$ if and only if there exists a probability distribution $\lambda^* \in \tilde{Y} = \Delta(Y)$ for which $\lambda = \lambda^* 1^S$ in the sense that

$$\lambda(y^S) = \begin{cases} \lambda^*(y) & \text{if } y^S = y1^S \text{ for some } y \in Y \\ 0 & \text{if } y^S \notin Y1^S. \end{cases}$$

These informal arguments motivate extending the definition of consequential equivalence in Section 4 as follows. When the tree T and the behaviour norm β are such that $F(T) \subset \Phi_\beta(T) \subset \tilde{Y}1^S$, then let

$$F^*(T) := \{ \lambda^* \in \tilde{Y} \mid \lambda^* 1^S \in F(T) \}$$
$$\Phi^*_\beta(T) := \{ \lambda^* \in \tilde{Y} \mid \lambda^* 1^S \in \Phi_\beta(T) \}$$

be respectively the set of feasible consequences and the set of consequences of behaviour, where now states of the world are ignored as having no bearing on the distribution of consequences. Notice especially that consequences are required to be *perfectly correlated*: it is not enough to have the same marginal distribution in all states of the world. An extended notion of consequential equivalence then requires the two trees $T \in \mathcal{T}(S)$ and $T' \in \mathcal{T}(S')$ to be regarded as equivalent whenever $F^*(T) = F^*(T')$, even if $S \neq S'$ so that they are not equivalent under the earlier definition. And requires behaviour in the two trees to be regarded as consequentially equivalent whenever $\Phi^*_\beta(T) = \Phi^*_\beta(T')$. An extended notion of consequentialism then demands that $\Phi^*_\beta(T) = \Phi^*_\beta(T')$ whenever$F^*(T) = F^*(T')$. Such consequentialist behaviour will be called *state independent* to reflect the fact that, when only constant consequences are possible, the states of the world have no influence on behaviour.

When behaviour is state independent as well as consequentialist, consider any decision tree $T \in \mathcal{T}(S)$ of the following trivial form, for any non-empty $S \subset E$. There is an initial decision node $n_0 \in N^*$ followed by two terminal nodes n_1, n_2 with $\gamma(n_1) = \lambda 1^S$, $\gamma(n_2) = \mu 1^S$. Then $n_1 \in \beta(T, n_0) \iff \lambda 1^S R^S \mu 1^S$, by definition of R^S. By state independence, this must be true for all non-empty $S \subset E$. So there exists a single state-independent binary relation R^* on $\tilde{Y}$ such that, for all $\lambda, \mu \in \tilde{Y}$ and all non-empty $S \subset E$, one has

$$\lambda \; R^* \; \mu \iff \lambda 1^S \; R^S \; \mu 1^S.$$

Because each R^S is a complete and transitive preference ordering, so is R^*, and continuity of behaviour implies that R^* inherits continuity from each R^S. The relation R^* also satisfies the independence axiom because each R^S satisfies that axiom. So there exists a unique cardinal equivalence class of state-independent NMUF's $v^* : Y \to \mathbb{R}$ such that, for all $\lambda, \mu \in \tilde{Y}$, one has

$$\lambda R^* \mu \iff \mathbb{E}_\lambda \; v^* \geq \mathbb{E}_\mu \; v^*.$$

But R^S is represented on $\tilde{Y}1^S$ by the expected value of $v^S(y1^S)$ mapping Y into $\mathbb{R}$. So, for every non-empty $S \subset E$ and every $\lambda, \mu \in \tilde{Y}$, one has

$$\begin{aligned} \mathbb{E}_\lambda \; v^*(y) \geq \mathbb{E}_\mu \; v^*(y) &\iff \lambda \; R^* \; \mu \iff \lambda 1^S \; R^S \; \mu 1^S \\ &\iff \mathbb{E}_\lambda \; v^S(y1^S) \geq \mathbb{E}_\mu \; v^S(y1^S). \end{aligned}$$

This implies that the utility functions $v^S(y1^S)$ on $Y1^S$ and $v^*(y)$ on Y must be cardinally equivalent functions of y. So, for all non-empty $S \subset E$, there exist constants $\rho^S > 0$ and α^S such that

$$v^S(y1^S) \equiv \alpha^S + \rho^S v^*(y).$$

In particular, when $S = \{s\}$ with $s \in E$, then

$$v_s(y) \equiv \alpha_s + \rho_s v^*(y).$$

The implications now depend on how each v^S depends on $v_s (s \in S)$, as in the three cases set out in Section 10. First, in the additive case (i),

$$v^S(y1^S) \equiv \alpha^S + \rho^S v^*(y) \equiv \sum\nolimits_{s \in S} \; v_s(y) \equiv \sum\nolimits_{s \in S} \; [\alpha_s + \rho_s v^*(y)]$$

and so:

$$\alpha^S = \sum\nolimits_{s \in S} \; \alpha_s; \quad \rho^S = \sum\nolimits_{s \in S} \; \rho_s.$$

Then, for general $y^S \in Y^S$, one has

$$v^S(y^S) \equiv \sum\nolimits_{s\in S} \; v_s(y_s) \equiv \sum\nolimits_{s\in S} \; [\alpha_s + \rho_s v^*(y_s)].$$

Further implications of this case are discussed in Section 13.

Next, in case (ii) with multiplicative positive utilities,

$$v^S(y1^S) \equiv \alpha^S + \rho^S v^*(y) \equiv \prod\nolimits_{s\in S} \; v_s(y) \equiv \prod\nolimits_{s\in S} \; [\alpha_s + \rho_s v^*(y)].$$

To explore this further, take the simple case when $S = \{s_1, s_2\}$ and write α_j, ρ_j for α_{s_j}, ρ_{s_j} $(j = 1, 2)$. Then, for all $y \in Y$, one has

$$\alpha^S + \rho^S v^*(y) = [\alpha_1 + \rho_1 v^*(y)][\alpha_2 + \rho_2 v^*(y)].$$

For any pair $y', y'' \in Y$, subtracting the version of the above equation for $y = y''$ from that for $y = y'$ gives

$$\rho^S[v^*(y') - v^*(y'')] = (\alpha_1\rho_2 + \rho_1\alpha_2)[v^*(y') - v^*(y'')] + \rho_1\rho_2\{[v^*(y')]^2 - [v^*(y'')]^2\}.$$

Thus, for any pair $y', y'' \in Y$ with $v^*(y') \neq v^*(y'')$, one has

$$\rho^S = \alpha_1\rho_2 + \rho_1\alpha_2 + \rho_1\rho_2[v^*(y') + v^*(y'')].$$

Therefore $v^*(y') + v^*(y'')$ must be a constant, independent of the choice of the pair $y', y'' \in Y$ with $v^*(y') \neq v^*(y'')$. This implies that the range $v^*(Y)$ of the NMUF can have at most two different values. But if indeed $v^*(Y) = \{v_1^*, v_2^*\}$ with $v_1^* > v_2^* > 0$, then one can satisfy the two equations

$$\alpha^S + \rho^S v_j^* = \prod\nolimits_{s\in S} \; (\alpha_s + \rho_s v_j^*) =: v_j^S \quad (j = 1, 2)$$

for any non-empty $S \subset E$ by taking:

$$\rho^S = (v_1^S - v_2^S)/(v_1^* - v_2^*) > 0; \quad \alpha^S = (v_1^* v_2^S - v_2^* v_1^S)/(v_1^* - v_2^*).$$

Then $v^S \equiv \alpha^S + \rho^S v^*$ on Y, as required for state-independent behaviour.

A similar argument applies in case (iii), with multiplicative negative utility. Only the signs of certain expressions change. So

THEOREM 11.1. *Suppose that there is a constant consequence domain Y. Suppose that the consistent behaviour norm β is continuous and consequentialist for the domain $\mathcal{T}$ of all consequential decision trees, and that β also satisfies state-independence whenever there are only constant consequences. Then there exists a unique cardinal equivalence class of NMUF's $v^* : Y \to \mathbb{R}$ and, for each such NMUF, and each state $s \in E$, there exist constants α_s and ρ_s (with $\rho_s > 0$), such that, when $v_s : Y \to \mathbb{R}$ is defined by $v_s(y) \equiv \alpha_s + \rho_s v^*(y)$, then behaviour in every decision tree $T \in \mathcal{T}(S)$ (with $\emptyset \neq S \subset E$) maximizes the expected utility function $\mathbb{E}v^S$, where v^S is given by:*

(i) the utility sum

$$v^S(y^S) \equiv \sum\nolimits_{s\in S} \; v_s(y_s) \equiv \sum\nolimits_{s\in S} [\alpha_s + \rho_s v^*(y)],$$

where the constants α_s $(s \in E)$ are arbitrary, and the positive constants ρ_s $(s \in E)$ are unique up to a common unit transformation $\tilde{\rho}_s = \tilde{\rho}\rho_s$ for some $\tilde{\rho} > 0$ which is independent of s;

or, provided that the range $v^*(Y)$ of possible values of the NMUF v^* has at most two distinct members, one of the following two other possibilities:

(ii) the product of positive utilities

$$v^S(y^S) \equiv \prod_{s\in S} v_s(y_s) \equiv \prod_{s\in S} [\alpha_s + \rho_s v^*(y)],$$

where the sets of constants α_s and ρ_s ($s \in E$) are both arbitrary up to common state-dependent unit transformations of the form $\tilde{\alpha}_s = \xi_s \alpha_s$ and $\tilde{\rho}_s = \xi_s \rho_s$ for some $\xi_s > 0$, and where $v_s(y) > 0$ for all $y \in Y$ and all $s \in E$;

(iii) minus the product of negative utilities

$$v^S(y^S) \equiv -\prod_{s\in S} [-v_s(y_s)] \equiv -\prod_{s\in S} [-\alpha_s - \rho_s v^*(y)],$$

where the constants α_s and ρ_s ($s \in E$) can be transformed as in (ii), and where $v_s(y) < 0$ for all $y \in Y$ and all $s \in E$.

The conclusion of Theorem 11.1 gives a complete characterization of the set of consequentialist behaviour norms satisfying the hypotheses of the theorem, because of the following obvious counterpart to the earlier Theorems 8 and 10.4

THEOREM 11.2. *Suppose that a unique cardinal equivalence class V^* of NMUF's $v^* : Y \to I\!R$ is specified, together with constants α_s and ρ_s ($\rho_s > 0$) for all $s \in E$. Suppose that, for any NMUF $v^* \in V^*$, the equivalence class V^S of functions $v^S : Y^S \to I\!R$ is constructed, for all non-empty $S \subset E$, as in the conclusion of Theorem 11.1. Then the behaviour norm whose consequences maximize $I\!Ev^S$ over $F(T)$ in every decision tree T of $\mathcal{T}(S)$, for all non-empty $S \subset E$ and all $v^S \in V^S$, is a continuous consequentialist consistent behaviour norm which satisfies state-independence whenever $F(T)$ contains only constant consequences.*

12. ELLSBERG'S PARADOX

An important implication of Theorem 11.1 is that only the additive case is possible when $\#v^*(Y) \geq 3$ — i.e. when Y contains at least three separate indifferent classes of consequences. This certainly justifies the almost exclusive attention that has been paid in the past to the additive case. Nevertheless the multiplicative case helps shed some light on the following example, which Ellsberg (1961) claimed as a violation of Savage's sure-thing principle, although it is more exactly described as a violation of Anscombe and Aumann's (1963) extension of that principle to allow some objective probabilities.

Lottery	Red	Black	Yellow
a	1	0	0
b	0	1	0
c	1	0	1
d	0	1	1

Table 12.1. Ellsberg's Four Lotteries

An urn is known to contain 90 balls, of which exactly 30 are red and the remaining 60 are either black or yellow without the exact numbers of black or yellow balls being known. One of the balls is drawn at random; all 90 balls are equally likely to be drawn. There are four lotteries a, b, c, d having consequences 0 or 1 which depend on the colour of the ball drawn, as set out in Table 12.1, where 0 represents "no prize" and 1 represents a fixed "prize" (in Ellsberg's article, it was \$ 100 — at 1961 prices in the U.S., presumably). Typically, experimental subjects stated a preference for a over b, perhaps because they preferred a definite $\frac{1}{3}$ probability of winning to an unknown probability. But they also stated a preference for d over c, perhaps because they preferred a definite $\frac{2}{3}$ probability of winning to an unknown probability (cf. the discussion by Gärdenfors and Sahlin, 1982). Yet Anscombe and Aumann's extension of the Savage sure-thing principle requires that

$$a \; P \; b \iff c \; P \; d$$

because, for each pair, winning or losing a prize when a yellow ball is drawn is a "sure-thing". In addition, if only the marginal probabilities in each state of the world matter, then the probability mixtures $\frac{1}{2}a + \frac{1}{2}d$ and $\frac{1}{2}b + \frac{1}{2}c$ have identical consequences (as pointed out by Raiffa, 1961, p. 694) because there is a 50–50 chance of winning the prize no matter what colour ball is drawn from the urn. Thus state-independence requires $\frac{1}{2}a + \frac{1}{2}d$ and $\frac{1}{2}b + \frac{1}{2}c$ to be indifferent, which violates the Anscombe and Aumann sure-thing principle if $a \; P \; b$ and $d \; P \; c$.

Nevertheless the preferences $a \; P \; b$ and $d \; P \; c$ are entirely consistent with consequentialism when there are just the two consequences 0 or 1, because then consequentialism allows maximization of an expected utility product. Indeed, there are really 61 different states of the world, $q = 0$ to 60, where q is the unknown number of black balls in the urn and $60 - q$ is the number of yellow balls. Suppose that $S := \{ q \mid q = 0 \text{ to } 60 \}$ and that, in an obvious notation, $v^S(\langle y_q \rangle_{q=0}^{60})$ takes the positive multiplicative form $\prod_{q=0}^{60} \; v_q(y_q)$. Because all the 90 balls actually in the urn are equally likely to be drawn, the expected utilities of the four lotteries are respectively:

$$U(a) = \prod_{q=0}^{60} \; \{[30 v_q(1) + 60 v_q(0)] \;/\; 90\};$$
$$U(b) = \prod_{q=0}^{60} \; \{[q v_q(1) + (90 - q) v_q(0)] \;/\; 90\};$$
$$U(c) = \prod_{q=0}^{60} \; \{[(90 - q) v_q(1) + q v_q(0)] \;/\; 90\};$$
$$U(d) = \prod_{q=0}^{60} \; \{[60 v_q(1) + 30 v_q(0)] \;/\; 90\}.$$

Many specifications of $v_q(0), v_q(1)$ ($q = 0$ to 60) are consistent with $a \; P \; b$ and $d \; P \; c$. For example, suppose that for $q = 0$ to 60 there are positive constants k_q for which:

$$v_q(0) := k_q; \quad v_q(1) := 2k_q. \tag{V}$$

Let K be the positive constant $\prod_{q=0}^{60} \; (k_q/90)$. Then the expected utilities of the four lotteries become:

$$U(a) = K \, (120)^{61};$$
$$U(b) = K \prod_{q=0}^{60} \; (90 + q);$$
$$U(c) = K \prod_{q=0}^{60} \; (180 - q);$$
$$U(d) = K \, (150)^{61}.$$

Because the respective arithmetic means 120, 150 of the two sequences $90 + q$, $180 - q$ ($q = 0$ to 60) exceed the geometric means $[\prod_{q=0}^{60} \; (90 + q)]^{1/61}$ and $[\prod_{q=0}^{60} \; (180 - q)]^{1/61}$, it must be true that $a \; P \; b$ and $d \; P \; c$. On the other hand, when $v^S(\langle y_q \rangle_{q=0}^{60})$ takes the negative multiplicative form $-\prod_{q=0}^{60} \; [-v_q(y_q)]$ with $v_q(\cdot)$ given by (V) above, then preferences must be "anti-Ellsberg" in the sense that $b \; P \; a$ and $c \; P \; d$.

13. REVEALED SUBJECTIVE PROBABILITIES AND BAYES' RULE

Decision theorists have naturally concentrated on the additive case, which is the only possible one when there is a common consequence domain Y, when dynamically consistent behaviour is continuous, consequentialist, and state independent for constant consequences, and when $v^*(Y)$ has at least three distinct members. In the additive case, whenever $\emptyset \neq S \subset E$ behaviour in any tree of $\mathcal{T}(S)$ will have consequences which maximize $\mathbb{E} \sum_{s \in S} [\alpha_s + \rho_s v^*(y_s)]$, where the constants α_s ($s \in E$) are arbitrary and the positive constants ρ_s ($s \in E$) are unique up to a common unit transformation of the form $\tilde{\rho}_s = \rho \rho_s$ (all $s \in E$) with $\rho > 0$ independent of s. Then, given any non-empty set $S \subset E$, one can drop the additive constants, and normalize the multiplicative constants ρ_s ($s \in S$) in order to satisfy $\sum_{s \in S} \rho_s = 1$. This yields a *unique* collection $\langle p(s|S) \rangle_{s \in S \subset E}$ of positive constants with $\sum_{s \in S} p(s|S) = 1$ such that the agent maximizes $\mathbb{E} \sum_{s \in S} p(s|S) v^*(y_s)$ in any tree of the set $\mathcal{T}(S)$.

These positive constants $p(s|S)$ $(s \in S \subset E)$ are *conditional subjective probabilities* revealed by consequentialist behaviour. All the Anscombe and Aumann (1963) axioms for subjective probabilities have been derived from consequentialism together with the standard additional assumptions stated above. Actually, consequentialism implies even more because the probabilities $p(s|S)$ are all *positive*; null events must be excluded from decision trees, just as zero probability events were excluded in Section 6 by the strengthened form of independence which consequentialism implies.

These "revealed" subjective probabilities are not only probabilities in the technical sense. They are also the unique set of probabilities for which the maximand $\mathbb{E}_\lambda \sum_{s \in S} p(s|S)\, v^*(y_s)$ represents expected utility with respect to the objective probability distribution $\lambda \in \tilde{Y}^S$ combined with the subjective probabilities $p(s|S)$ $(s \in S)$. This reflects the fact that subjective probabilities also function as objective probabilities in the following sense. Suppose that every natural node $n \in N^1$ of every decision tree $T \in \mathcal{T}(S)$ were replaced with a chance node with probabilities specified by

$$\pi(n'|n) = \sum\nolimits_{s \in S(n')} p(s|S(n))$$

for every $n' \in N_{+1}(n)$. Suppose too that the terminal nodes $n \in X$ with consequences $\gamma(n) \in \tilde{Y}^{S(n)}$ were instead given consequences $\gamma^*(n) \in \tilde{Y}$ defined by

$$\gamma^*(n)(y) = \sum\nolimits_{s \in S(n)} p(s|S(n))\ \gamma(n)(\{\, y^S \mid y_s = y \,\}).$$

Then the backward recursions of Section 4 would replace $\Phi_\beta(T,n)$ and $F(T,n)$ with $\Phi^*_\beta(T,n)$ and $F^*(T,n)$ whose members $\lambda^* \in \tilde{Y}$ would be derived from $\lambda \in \tilde{Y}^{S(n)}$ just as $\gamma^*(n)$ was derived from $\gamma(n)$ above. This can be proved by backward induction. And the expected utility values derived in Section 10 would satisfy

$$\begin{aligned} w(T,n) &= \max_\lambda \{\, \mathbb{E}_\lambda\, v^{S(n)} \mid \lambda \in F(T,n) \,\} \\ &= \max_{\lambda^*} \{\, \mathbb{E}_{\lambda^*}\, v^* \mid \lambda^* \in F^*(T,n) \,\} \\ &= \max_\lambda \{\, \mathbb{E}_\lambda \sum\nolimits_{s \in S(n)} p(s|S(n)) v^* \mid \lambda \in F(T,n) \,\} \end{aligned}$$

of course, because of expected utility maximization. This shows that subjective probabilities indeed function as objective probabilities which can be attached to natural nodes in any consequential decision tree.

Another implication of consequentialism is that revealed subjective probabilities must satisfy Bayes' rule. Since Bayes' rule has recently come under critical scrutiny by philosophers such as Brown (1976), Hacking (1967) and Teller (1973) and by statisticians such as Shafer (1976) and Diaconis (1978), this will be reassuring to Bayesians.

First, if $S' \subset S$, write $p(S'|S) := \sum_{s \in S'} p(s|S)$. Now suppose $s \in S' \subset S$. Consider a decision tree T with a natural node $n \in N^1$ followed immediately by a succeeding natural node $n' \in N_{+1}(n) \cap N^1$ such that $S(n) = S$, $S(n') = S'$. Suppose that $n'' \in N_{+1}(n')$ and $S(n') = \{s\}$. Because the collection

$$\{\, S(\bar{n}') \mid \bar{n}' \in N_{+1}(n') \,\} \cup \{\, S(\bar{n}) \mid \bar{n} \in N_{+1}(n) \setminus \{n'\} \,\}$$

is a partition of $S(n)$, this decision tree is consequentially equivalent to a tree T' in which the only change is that node n' is omitted, and replaced in $N_{+1}(n)$ by all the nodes of $N_{+1}(n')$, so that $N'_{+1}(n) = N_{+1}(n) \cup N_{+1}(n') \setminus \{n'\}$. But the natural nodes n and n' in T can be replaced by chance nodes with probabilities:

$$\begin{aligned} \pi(n'|n) &= p(S(n')|S(n)) = p(S'|S); \\ \pi(n''|n') &= p(S(n'')|S(n')) = p(s|S'). \end{aligned}$$

The probability of reaching n'' from n is then $p(s|S')\ p(S'|S)$. When n is replaced by a chance node in the equivalent tree T', the corresponding probability is

$$\pi(n''|n') = p(s|S(n)) = p(s|S).$$

Because of consequential equivalence, the probabilities of reaching node n'' from n must be equal in these two trees, so

$$p(s|S) = p(s|S')\ p(S'|S).$$

Alternatively, and perhaps more simply, one can use Weller's (1978) argument regarding consistency of expected utility maximizing behaviour in decision trees.

When $S'' \subset S' \subset S$, summing the last equation over $s \in S''$ gives

$$p(S''|S) = p(S''|S')\ p(S'|S).$$

When S' is not a subset of S, one can define, obviously

$$p(S'|S) := p(S \cap S'|S).$$

Now consider any prior "hypothesis" $s \in H_0 \subset E$, together with an "observation" that $s \in G \subset E$ and any posterior "hypothesis" $s \in H \subset E$. Then one has

$$\begin{aligned} p(G \cap H \cap H_0|H_0) &= p(G \cap H \cap H_0|G \cap H_0)\ p(G \cap H_0|H_0) \\ &= p(G \cap H \cap H_0|H \cap H_0)\ p(H \cap H_0|H_0), \end{aligned}$$

so that $p(H|G \cap H_0) = p(G|H \cap H_0)\ p(H|H_0)/p(G|H_0)$. In particular, given any two posterior hypotheses H_1, H_2, the likelihood ratio is

$$\frac{p(H_1|G \cap H_0)}{p(H_2|G \cap H_0)} = \frac{p(G|H_1 \cap H_0)}{p(G|H_2 \cap H_0)}\ \frac{p(H_1|H_0)}{p(H_2|H_0)}.$$

That is, the posterior likelihood ratio is equal to the product of the conditional likelihood ratio and the prior likelihood ratio. This is the usual formula in Bayesian statistics, of course. So

THEOREM 13.1. *Suppose that the consistent behaviour norm β is continuous and consequentialist for the domain $\mathcal{T}$ of all consequential decision trees, and that β also satisfies state-independence whenever there are only constant consequences. Suppose also that there is a constant consequence domain Y with at least three distinct indifference classes. Then there exists a unique cardinal equivalence class of NMUF's $v^* : Y \to \mathbb{R}$ and there exist unique positive subjective probabilities $p(s|S)$ $(s \in S \subset E)$ satisfying Bayes' Rule $p(S''|S) = p(S''|S')\ p(S'|S)$ (all $S'' \subset S' \subset S$), such that the consequences of β in any decision tree $T \in \mathcal{T}(S)$ maximize subjectively expected utility $\mathbb{E} \sum_{s \in S} p(s|S)\, v^*(y_s)$.*

Yet again, this is an alternative complete characterization of those consequentialist behaviour norms which satisfy the hypotheses of Theorem 11.1, because of the following obvious counterpart to the earlier Theorems 8, 10.4 and 11.2

THEOREM 13.2. *Suppose that a cardinal equivalence class V^* of NMUF's $v^* : Y \to \mathbb{R}$ is specified, together with strictly positive subjective probabilities $p(s|E)$ $(s \in E)$. Suppose that for every non-empty $S \subset E$, the conditional probabilities $p(s|S) := p(s|E)/\sum_{s' \in S} p(s'|E)$ are constructed, together with the cardinal equivalence class V^S of NMUF's $v^S : Y^S \to \mathbb{R}$ which satisfy*

$$v^S(y^S) \equiv \sum\nolimits_{s \in S} p(s|S) v^*(y_s)$$

for some $v^* \in V^*$. Then there is a unique behaviour norm whose consequences maximize $\mathbb{E}v^S$ over $F(T)$ in every decision tree T of $\mathcal{T}(S)$, for all non-empty $S \subset E$ and all $v^S \in V^S$, and this behaviour norm is dynamically consistent, continuous, consequentialist, and satisfies state independence whenever $F(T)$ contains only constant consequences.

14. SUMMARY AND CONCLUDING REMARKS

Normative decision theory has usually considered the "reduced form" of a decision problem, in which the problem is reduced to that of choosing a single decision strategy. Axioms are then introduced which concern behaviour in this normal form. Here a different approach has been taken, which retains the "extensive form" of the decision problem — that is, the decision tree — and requires behaviour in each decision tree to be *consistent*, as defined in Section 3. Now, in reduced form decision theory, an axiom which is so fundamental that it is often left implicit is that decision strategies are evaluated by their consequences. Savage (1954) indeed even *defines* an act as a mapping from the uncertain states of the world to consequences. Here this axiom is called *consequentialism* and is applied in an obvious way to dynamically consistent behaviour norms which are definied on the domain of all possible consequential finite decision trees with strictly positive probabilities at each chance node, as explained in Section 4. Indeed consequentialism is precisely the assumption which justifies considering only the reduced form of a decision problem.

The implications of this "consequentialist" approach to normative decision theory are quite striking. Many of the standard axioms, which before only had an "intuitive" justification that was often highly questionable and frequently questioned, now become logical implications of the consequentialist "pre-axiom". In particular, consequentialism implies the existence of a (complete and transitive) revealed preference ordering. It also implies that preferences must satisfy Samuelson's (1952) controversial independence axiom, and a new version of the sure-thing principle which is intermediate between Savage's (1954) original version — which applies in the absence of objective probabilities — and Anscombe and Aumann's (1963) extension — which applies to all probability distributions over state-contingent consequence functions. Consequentialism implies that the sure-thing principle must apply just to independent probability distributions of state-contingent consequences — a version of the principle which appears never to have been considered previously.

Consequentialism does fall somewhat short of being a complete justification of subjective expected utility maximization. For example, even in the complete absence of uncertain states of the world it does not imply expected utility because consequentialist behaviour could be discontinuous as objective probabilities vary. Imposing continuity, however, leads naturally to expected utility maximization. But not to subjective probability. For this, additional assumptions are required, even if one imposes the standard hypothesis of a constant (state-independent) consequence domain, as has been usual since Savage's and Anscombe and Aumann's derivations of revealed subjective probabilities. Following Savage (1954), Section 11 required that when behaviour can only have constant consequences, independent of the state, then the set of possible states of the world should be irrelevant. Even then, if there are no more than two indifference classes in the constant consequence domain, expected utility could be multiplicatively rather than additively separable across different states of the world. In this special case consequentialism is entirely consistent with the pattern of preferences most frequently observed by experimental subjects who are confronted with Ellsberg's (1961) example. When there is state independence this is really just a *curiosum* which is excluded by consequentialism as soon as there are at least three distinct indifference classes among the domain of constant consequences. When there are no consequences which are common to any pair of different states, however, as is implicitly the case when state-dependent utility functions are considered, then there is no justification which consequentialism can offer for restricting attention to additively rather than multiplicatively separable expected utility functions.

Consequentialism has another implication that goes beyond orthodox normative decision theory. To avoid inconsistencies of behaviour, there can be no zero probabilities at any chance node of a decision tree, nor any null states of the world whose subjective probability would have to be zero. In single person decision theory and for finite decision trees, this is not a problem because one can argue that zero probability or null events should be excluded from all decision trees. When the set of possible states of the world is a continuum, or when one considers games in which one cannot restrict attention to completely mixed strategies, this implication of consequentialism is serious and suggests that a rather richer space of probabilities needs to be considered.

From Section 11 on, this paper invoked the standard hypothesis that the domain of possible consequences is independent of the state. The work of Drèze (1958, 1961, 1962, 1985, 1986, 1987) — see also Dehez and Drèze (1982) — and of some notable successors shows that this hypothesis is completely unacceptable when states of the world may include such calamities as accidental death or injury. In later work I hope to characterize continuous consequentialist consistent behaviour norms which do satisfy state independence in those decision trees for which only constant consequences of the form $y1^S$ can result from behaviour, but without invoking the unacceptable hypothesis that such constant consequences are always possible for all simple consequences $y \in Y$ and all non-empty sets $S \subset E$. Any such characterization will be considerably more complicated than results such as Theorems 10.1, 10.2, 11.1, 11.2, 13.1, and 13.2 in this paper, and will require an approach such as that of Karni, Schmeidler and Vind (1983) if subjective probabilities are to be derived in more than rather special cases.

This paper has also considered just single person decision theory. The implications of consequentialism for behaviour in extensive games remain to be explored in further work. The implications for social choice theory have already been touched on in Hammond (1977, 1983, 1986, 1987a). Indeed, the last three of these papers build upon results presented here.

ENDNOTES

1. In ethics, where the term "consequentialism" was first used by Anscombe (1958), the hypothesis is not only questioned, but often regarded as untenable. See especially Williams (1973, 1985) and Sen and Williams (eds.) (1982). Some of the issues are incompletely discussed in Hammond (1986a).
2. See Hammond (1977, 1983, 1985, 1986, 1987a, b), Karni and Safra (1986) and McClennen (1986, 1987). Of these, the first paper discusses "metastatic choice", which is a less fortunate description of the same idea.
3. See Section 7 of Hammond (1987b) for a discussion of how even such apparent non-expected utility maximizing behaviour as that discussed by Allais (1953, 1979) and Machina (1982) is really expected utility maximizing when the utility function is allowed to include more arguments. For related discussion, see also Markovitz (1970), Morrison (1967), Mossin (1969) and Broome (1986).

In a recent interesting paper Sen (1985) has considered what he claims to be three counter-examples to the "strong" independence axiom which is an implication of expected utility theory. The first example, the "no-letter response", has an agent react enirely rationally to the anticipation of what a particular letter would have contained had it been sent - e.g., to the absence of good news on the one hand, and of bad news on the other. This instance of "psychological sensitivity", along with the closely related "Bergen paradox" of Drèze (1987, pp. 14–15), can surely be handled by including psychological mood in the description of each consequence in the relevant domain.

The second example, the "doctor's dilemma", may simply be an ethically irrelevant and possibly irrational instance of moral cowardice in the face of the agonizing decision as to which of two patients' lives to save when only one can be. The claim is that the example illustrates the relevance of "agency sensitivity": my counterclaim is that if such sensitivity to "agency" — in the sense of responsibility for a life-and-death decision of this kind — is indeed relevant, then it too can be captured with a suitable extension of the consequence domain.

Sen's third example has Ayesha choose a different career, as a civil rights lawyer, because in the event that she is *not* deported from the U.K., her experience of having faced possible deportation will serve as useful experience. At first this seems easy to treat consequentially too, since experience and information are clearly relevant consequences in such decision problems concerning the choice of a career. The difficulty is that Ayesha is apparently "boundedly rational" to the extent that she seems unable to imagine the complete experience of facing possible deportation unless she has faced it personally. Consequentialism is not yet well adapted to deal with bounded rationality, yet often it is rational to admit that one's rationality *is* bounded.

Thus, while I wholeheartedly agree with Sen (1985, p. 123) that "rationality deserves a less mechanical approach" than mere expected utility maximization, such an approach will have to come, in my view, from specifying what consequences *should* be relevant for rational behaviour — i.e., what should count as an argument of a rational agent's utility function — and what restrictions rationality puts on the form of that utility function.

4. Johnsen and Donaldson (1985) appear to deny the validity of some of my results. However, they effectively consider a restricted domain of decision trees so that the implications of consequentialism become muted — in particular, expected utility maximization is no longer implied. Karni and Safra (1986) also relax this unrestricted domain assumption, but show how, even so, consequentialism implies the independence axiom for a restricted domain of decision trees corresponding to English auctions. Pope's (1985) discussion also rests on their being a restricted domain. Indeed, the unrestricted domain assumption may be restrictive. Of course, a decision tree can hardly include, as part of a consequence, regret at missing an opportunity to have consequence y, unless there was an opportunity in the past to have had y. But the decision tree may be a continuation of an earlier tree in which this opportunity existed, but has now gone. So how restrictive the assumption is remains unclear to me at the time of writing.

5. A somewhat related justification of the independence axiom has recently been provided by Lavalle and Wapman (1986).

6. A thorough exploration of a related notion of consequential equivalence has recently been provided by Lavalle and Fishburn (1987).

ACKNOWLEDGEMENTS

The paper is based upon some notes with the same title which were circulated at the 3rd International Conference on the Foundations and Applications of Utility, Risk and Decision Theories, Aix-en-Provence, June 1986, together with material destined to appear in the conference proceedings as Hammond (1987b). Together with related work, it has been supported by the U.S. National Science Foundation, the John Simon Guggenheim Memorial Foundation, and C.O.R.E. at the Université Catholique de Louvain. I recall helpful discussions with Ken Arrow, Nick Baigent, Eddie Dekel, Jacques Drèze, Yossi Greenberg, Frank Hahn, Martin Hellwig, Mark Machina, Ned McClennen, Jean-François Mertens, Hervé Moulin, Prasanta Pattanaik, Amartya Sen and Bob Sugden. I owe much to all of these and no doubt several others too, especially some interesting opponents.

Peter J. Hammond
Department of Economics
Stanford University, CA 94305–6072
U.S.A.

REFERENCES

Allais, M.: 1953, "Le comportement de l'homme rationnel devant le risque: critique des postulats et des axiomes de l'école américaine", *Econometrica*, **21**, 503–546.

Allais, M.: 1979, "The foundations of a positive theory of choice involving risk and a criticism of the postulates and axioms of the American school", in *The Expected Utility Hypothesis and the Allais Paradox*, M. Allais and O. Hagen (eds.), 27–145. Dordrecht: D. Reidel.

Anscombe, F.J. and R.J. Aumann: 1963, "A definition of subjective probability", *Annals of Mathematical Statistics*, **34**, 199–205.

Anscombe, G.E.M.: 1958, "Modern moral philosophy", *Philosophy*, **33**, 1–19.

Blume, L.E.: 1986, "Lexicographic refinements of Nash equilibrium", Economics Discussion Paper, University of Michigan.

Brandenburger, A. and E. Dekel: 1986, "On an axiomatic approach to refinements of Nash equilibrium", University of Cambridge, Economic Theory Discussion Paper No. 104.

Broome, J.: 1986, "Should a rational person conform to the sure-thing principle?", presented at the 3rd International Conference on the Foundations and Applications of Utility, Risk and Decision Theories, Aix-en-Provence.

Brown, P.M.: 1976, "Conditionalization and expected utility", *Philosophy of Science*, **43**, 415–419.

Chipman, J.S.: 1960, "The foundations of utility", *Econometrica*, **28**, 193–224.

Dehez, P. and J.H. Drèze: 1982, "State-dependent utility, the demand for insurance and the value of safety", in *The Value of Life and Safety: Proceedings of a Conference Held by the 'Geneva Association'*, M.W. Jones-Lee (ed.), 41–65. Amsterdam: North-Holland.

Diaconis, P.: 1978, Review of Shafer (1976), *Journal of American Statistical Association*, **73**, 677–678.

Drèze, J.H.: 1958, "On the identifiability of subjective probabilities and cardinal utilities" (abstract), *Econometrica*, **26**, 319–320.

Drèze, J.H.: 1961, "Fondements logiques de la probabilité subjective et de l'utilité", in *La décision*, Colloques Internationaux du CNRS (Paris), pp. 73–87.

Drèze, J.H.: 1962, "L'utilité sociale d'une vie humaine", *Revue française de recherche opérationnelle*, **6**, 93–118.

Drèze, J.H.: 1985, "Decision theory with moral hazard and state-dependent preferences", in Drèze (1987), ch. 2, pp. 23–89.

Drèze, J.H.: 1986, "Moral expectation with moral hazard", in *Contributions to Mathematical Economics in Honor of Gerard Debreu*, W. Hildenbrand and A. Mas-Colell (eds.), ch. 11, 187–204. Amsterdam: North-Holland.

Drèze, J.H.: 1987, *Essays on Economic Decisions under Uncertainty.* Cambridge: Cambridge University Press.

Ellsberg, D.: 1961, "Risk, ambiguity, and the Savage axioms", *Quarterly Journal of Economics*, **75**, 643–669.

Fishburn, P.C.: 1982, *The Foundations of Expected Utility.* Dordrecht: D. Reidel.

Gärdenfors, P. and N.-E. Sahlin: 1982, "Unreliable probabilities, risk taking, and decision making", *Synthese*, **53**, 361–386.

Hacking, I.: 1967, "Slightly more realistic probability", *Philosophy of Science*, **34**, 311–325.

Hammond, P.J.: 1976, "Changing tastes and coherent dynamic choice", *Review of Economic Studies*, **43**, 159–173.

Hammond, P.J.: 1977, "Dynamic restrictions on metastatic choice", *Economica*, **44**, 337–350.

Hammond, P.J.: 1983, "Ex-post optimality as a dynamically consistent objective for collective choice under uncertainty", in *Social Choice and Welfare*, P.K. Pattanaik and M. Salles (eds.), ch. 10, 175–205. Amsterdam: North-Holland.

Hammond, P.J.: 1986, "Consequentialist social norms for public decisions", in *Social Choice and Public Decision Making: Essays in Honor of Kenneth J. Arrow, Vol. I*, W.P. Heller, R.M. Starr and D.A. Starrett (eds.) , ch. 1, 3–27. Cambridge: Cambridge University Press.

Hammond, P.J.: 1987a, "On reconciling Arrow's theory of social choice with Harsanyi's fundamental utilitarianism", in *Arrow and the Foundations of the Theory of Economic Policy*, G.R. Feiwel (ed.), ch. 4, 179–222. London: Macmillan.

Hammond, P.J.: 1987b, “Consequentialism and the independence axiom”, to appear in Munier (1987).

Harsanyi, J.C.: 1978, “Bayesian decision theory and utilitarian ethics”, *American Economic Review, Papers and Proceedings*, **68**, 223–228.

Hausner, M.: 1954, “Multidimensional utilities”, in R. M. Thrall, C. H. Coombs and R. L. Davis (eds.), *Decision Processes*, ch. 12, 167–180. New York: John Wiley.

Herstein, I.N. and J. Milnor: 1953, “An axiomatic approach to measurable utility”, *Econometrica*, **21**, 291–297.

Johnsen, T.H. and J.B. Donaldson: 1985, “The structure of intertemporal preferences under uncertainty and time consistent plans”, *Econometrica*, **46**, 186-200.

Karni, E. and Z. Safra: 1986, “Dynamic consistency in English auctions and expected utility theory”, mimeo.

Karni, E., Schmeidler, D. and K. Vind: 1983, “On state dependent preferences and subjective probabilities”, *Econometrica*, **51**, 1021–1031.

LaValle, I.H. and P.C. Fishburn: 1987, “Equivalent decision trees and their associated strategy sets”, *Theory and Decision*, **23**, 37–63.

LaValle, I.H. and K.R. Wapman: 1986, “Rolling back decision trees requires the independence axiom”, *Management Science*, **32**, 382–385.

McClennen, E.F.: 1986, “Dynamic choice and rationality”, to appear in Munier (1987).

McClennen, E.F.: 1987, *Rationality and Dynamic Choice: Foundational Explorations*.

Markovitz, H.: 1970, *Portfolio Selection: Efficient Diversification of Investment*. New Haven: Yale University Press.

Morrison, D.G.: 1967, “On the consistency of preference in Allais’ paradox”, *Behavioral Science*, **12**, 373–383.

Mossin, J.: 1969, “A note on uncertainty and preferences in a temporal context”, *American Economic Review*, **59**, 172–174.

Munier, B. (ed.): 1987, *Risk, Decision and Rationality: Proceedings of the 3rd International Conference on the Foundations and Applicatons of Utility, Risk and Decision Theories*. Dordrecht: D. Reidel, forthcoming.

Myerson, R.B.: 1979, “An axiomatic derivation of subjective probability, utility and evaluation functions”, *Theory and Decision*, **11**, 339–352.

Pope, R.: 1985, “Timing contradictions in von Neumann and Morgenstern’s axioms and in Savage’s ‘sure-thing’ proof”, *Theory and Decision*, **18**, 229–261

Raiffa, H.: 1961, “Risk, ambiguity, and the Savage axioms”, *Quarterly Journal of Economics*, **75**, 690–694.

Raiffa, H.: 1968, *Decision Analysis : Introductory Lectures on Choices under Uncertainty*. Reading, Mass.: Addison-Welsey.

Samuelson, P.A.: 1952, “Probability, utility, and the independence axiom”, *Econometrica*, **20**, 670–678.

Savage, L.J.: 1954, *The Foundations of Statistics*. New York: John Wiley.

Sen, A.K.: 1985, “Rationality and uncertainty”, *Theory and Decision*, **18**, 109–127.

Sen, A.K. and B. Williams (eds.): 1982, *Utilitarianism and Beyond*. Cambridge: Cambridge University Press.

Shafer, G.: 1976, *A Mathematical Theory of Evidence*. Princeton: Princeton University Press.

Skala, H.J.: 1975, *Non-Archimedean Utility Theory*. Dordrecht: D. Reidel.

Teller, P.: 1973, "Conditionalization and observation", *Synthese*, **26**, 218–258.

von Neumann, J. and O. Morgenstern: 1944, *Theory of Games and Economic Behavior*. Princeton: Princeton University Press.

Weller, P.A.: 1978, "Consistent intertemporal decision making under uncertainty", *Review of Economic Studies*, **45**, 263–266.

Williams, B.A.O.: 1973, 'A critique of utilitarianism", in *Utilitarianism for and against*, J.J.C. Smart and B.A.O. Williams (eds.). Cambridge: Cambridge University Press.

Williams, B.: 1985, *Ethics and the Limits of Philosophy*. London: Fontana Press/ Collins.

Wilson, R.B.: 1968, "The theory of syndicates", *Econometrica*, **36**, 119–132.