Lecture notes for Decision Theory

David Schmeidler

October 2004

von Neumann-Morgenstern utility.

Let X be a nonempty finite set, $\Delta(X) = \{\ p : X \to [0,1] \mid \sum_{x\in X} p(x) = 1\}$, and let $\succsim$ denote a binary relation on $\Delta(X)$. As usual, $\succ$ and $\sim$ denote the asymmetric and symmetric parts of $\succsim$. In our nomenclature elements of X are *outcomes* (or consequences or prizes), elements of $\Delta(X)$ are *lotteries*, and $\succsim$ is the *preference relation.*

Axioms: **A1** the relation $\succsim$ is complete and transitive. **A2** (continuity) $\forall p, q, r \in \Delta(X)$ with $p \succ q \succ r$, there are $\alpha, \beta \in (0,1)$ such that $\alpha p + (1-\alpha)r \succ q \succ \beta p + (1-\beta)r$. **A3** (independence or substitution) $\forall p, q, r \in \Delta(X)$, and $\alpha \in (0,1] : p \succ q \Rightarrow \alpha p + (1-\alpha)r \succ \alpha q + (1-\alpha)r$. (Note that for all $p, q \in \Delta(X)$, $\alpha p + (1-\alpha)q \in \Delta(X)$, and $[\alpha p + (1-\alpha)q](x) = \alpha p(x) + (1-\alpha)q(x)$. Obviously $\sum_{x\in X}[\alpha p(x) + (1-\alpha)q(x)] = 1$.)

Theorem (for lotteries). *A binary relation $\succsim$ on $\Delta(X)$ satisfies A1, A2, and A3 iff there exists a (utility) function $u : X \to R$ such that, $p \succsim q \Leftrightarrow \sum_{x\in X} p(x)u(x) \geq \sum_{x\in X} q(x)u(x)$. Moreover, a function $v : X \to R$ satisfies $p \succsim q \Leftrightarrow \sum_{x\in X} p(x)u(x) \geq \sum_{x\in X} q(x)u(x)$ iff $v(x) = \alpha u(x) + \beta$ on X for $\alpha, \beta \in R$, $\alpha > 0$.*

Proof. First a **Notational** comment: no quantifier means the quantifier "for all"with the appropriate range. Next, assuming A1, A2, and A3 the following five claims are proved:

(i) $p \succ q$, $0 < \alpha < 1$ imply $p \succ \alpha p + (1-\alpha)q \succ q$.

(ii) $p \succ q$, $0 \leq \alpha < \beta \leq 1$ imply $\beta p + (1-\beta)q \succ \alpha p + (1-\alpha)q$.

(iii) $p \sim q$, $\alpha \in (0,1)$ imply $p \sim \alpha p + (1-\alpha)q(\sim q)$.

(iv) $p \sim q$, $\alpha \in (0,1)$ imply $\alpha p + (1-\alpha)r \sim \alpha q + (1-\alpha)r$.

(v) $p \succsim q \succsim r$, $p \succ r$ imply the existence of a unique $\alpha^* \in [0,1]$ such that, $\alpha^* p + (1-\alpha^*)r \sim q$.

Proof of (i). $p = (1-\alpha)p + \alpha p \succ (1-\alpha)q + \alpha p = \alpha p + (1-\alpha)q \succ \alpha q + (1-\alpha)q = q$, where the preferences are implied by A3 with the appropriate substitutions, and transitivity.

Proof of (ii). If one only of $\alpha = 0$ or $\beta = 1$ hold, (ii) reduces to (i). If both hold the conclusion coincides with the assumption. In case of strict inequalities, by (i), $p \succ \beta p + (1-\beta)q \succ q$. Since $\alpha p + (1-\alpha)q = \gamma[\beta p + (1-\beta)q] + (1-\gamma)q$ with $0 < \gamma = \alpha/\beta < 1$, the conclusion follows again by (i). Note that (ii) and completeness yield the converse implication of (ii).

Proof of (iii). By way of negation suppose that $p \succ \alpha p + (1-\alpha)q =: s$. (**Notation**: $x =: y$ means that y is defined to be x. Similarly, $x := y$ means that x is defined to be y.) Let $t := (1/2)p + (1/2)s$. By (i) $p \succ t \succ s$. On the other hand $q \sim p \succ t$ and s is between q and t, .which in turn implies that for some $\delta \in (0,1)$, $s = \delta q + (1-\delta)t$. Since $q \sim p \succ t$, applying (i) again we get $s \succ t$, a contradiction. The case $s \succ p$ similarly leads to a contradiction.

Proof of (iv). If $r \sim p \sim q$ then (iii) yields the required conclusion. Let us assume that $r \succ p \sim q$.

Let $s := \alpha p + (1-\alpha)r$, and $t := \alpha q + (1-\alpha)r$, and suppose by way of negation that $s \succ t$. By A3 $r \succ s$. Hence, $r \succ s \succ t$. By A2 there exists $\beta \in (0,1)$ such that, $s \succ \beta r + (1-\beta)t = \beta r + (1-\beta)[\alpha q + (1-\alpha)r] = (\alpha - \alpha\beta)q + (1-\alpha+\alpha\beta)r =: v$.

In order to follow the arguments it may help to draw a triangle with $[p,q]$ as basis, and r the vertex. Then s lies on the $[r,p]$ side, and t and v on the $[r,q]$ side with v above t.

Let $w = \beta r + (1-\beta)q$. Again by A3 or (i), $r \succ w \succ q \sim p$. So by A3, $\alpha w + (1-\alpha)r \succ \alpha p + (1-\alpha)r = s$. But $\alpha w + (1-\alpha)r = v$, a contradiction to $s \succ v$ above.

Case 2, $p \sim q \succ r$, is proved similarly.

Proof of (v). Uniqueness is implied by (ii). Obviously if $p \sim q$ then $\alpha^* = 1$ and if $q \sim r$ then $\alpha^* = 0$. So we are left with the case $p \succ q \succ r$. Define $\alpha^* = \inf\{\ \alpha \in [0,1] \mid \alpha p + (1-\alpha)r \succ q\ \}$. By A2 (continuity) $\alpha^* < 1$. We have to show that $\alpha^* p + (1-\alpha^*)r \sim q$. Suppose by way of negation that $\alpha^* p + (1-\alpha^*)r \succ q$. So, $\alpha^* > 0$. Applying A2 to $\alpha^* p + (1-\alpha^*)r \succ q \succ r$ we get $\beta \in (0,1)$ such that $\beta[\alpha^* p + (1-\alpha^*)r] + (1-\beta)r \succ q$. As $\beta\alpha^* < \alpha^*$, it is a contradiction to the definition of α^*. The other negation assumption (case 2) is: $q \succ \alpha^* p + (1-\alpha^*)r$. Here we apply A2 to $p \succ q \succ \alpha^* p + (1-\alpha^*)r$ and get $\gamma \in (0,1)$ such that $q \succ \gamma p + (1-\gamma)[\alpha^* p + (1-\alpha^*)r] =$

$(\gamma + \alpha^* - \gamma\alpha^*)p + (1-\gamma-\alpha^*+\gamma\alpha^*)r = (\gamma(1-\alpha^*)+\alpha^*)p + (1-\gamma-\alpha^*+\gamma\alpha^*)r$. So, q is preferred to a convex combination of p and r where the coefficient of p is larger than α^*. This contradicts (ii) and the definition of α^*.

Basic comment. If we replace the set $\Delta(X)$ with an arbitrary convex subset, say K, of a linear space, the axioms are meaningfully stated and so are the five claims. The proofs use only convexity and don't require the spacial structure of $\Delta(X)$.Thus the claims (i),...,(v) hold with K instead of $\Delta(X)$. We can now state a von Neumannn-Morgenstern utility theorem for a more general domain. But first let us recall that a function $f : K \to \mathbb{R}$ is said to be **affine** if for all $p, q \in K$ and $\alpha \in [0,1]$, $f(\alpha p + (1-\alpha)q) = \alpha f(p) + (1-\alpha)f(q)$. Note also,

(vi) if $F : K \to R$ is affine than for all m, $f_1, f_2, ..., f_m$ in K, and $\theta_1, \theta_2, ..., \theta_m$ in $[0,1]$ with $\sum_{i=1}^m \theta_i = 1 : F(\sum_{i=1}^m \theta_i f_i) = \sum_{i=1}^m \theta_i F(f_i)$. The simple proof by induction is omitted.

vNM Utility Theorem. *A binary relation $\succsim$ on a convex set K satisfies A1, A2, and A3 iff there exists an affine (utility) function $U : K \to R$ such that, $p \succsim q \Leftrightarrow U(p) \geq U(q)$. Moreover, an affine function $V : K \to R$ satisfies $p \succsim q \Leftrightarrow V(p) \geq V(q)$ iff $V(\cdot) = \alpha U(\cdot) + \beta$ for some $\alpha, \beta \in R$, $\alpha > 0$.*

We continue the proof in several steps. First, U as above is constructed on a subset of K. In step 2 the derivation of the utility function, u, in the Theorem for lotteries is concluded. In step 3 the relative uniqueness of the utility functions in both theorems is proved. In step 4 U of step 1 is extended to all of K. Finally, step 5 deals with the direction from representation to axioms in both framework.

But first note that if $p \sim q$ for all $p, q \in K$ (or $\in \Delta(X)$), then U constant (or u constant) represent $\succsim$ and is affine (and unique as required). From now on we assume that there are $p, q \in K$ such that $p \succ q$.

Step 1. Let $a, b \in K$, with $a \succ b$. Denote by $[a, b]$ the order interval $\{ p \in K \mid a \succsim p \succsim b \}$. By (v), for each $p \in [a, b]$ there is a unique α^* such that $p \sim \alpha^* a + (1 - \alpha^*)b$. Define $U(p) = \alpha^*$. Obviously, $U(a) = 1 = U(p)$ whenever $p \sim a$, and $U(b) = 0 = U(p)$ whenever $p \sim b$. We have to prove that for all $p, q \in [a, b]$ and $\lambda \in [0, 1]$, $U(\lambda p + (1 - \lambda)q) = \lambda U(p) + (1 - \lambda)U(q)$. By our definition of U, $p \sim U(p)a + (1 - U(p)b$, and $q \sim U(q)a + (1 - U(q)b$. These two equivalences and (iv) imply

$$\lambda p + (1 - \lambda)q \sim \lambda(U(p)a + (1 - U(p))b) + (1 - \lambda)q$$

$$\sim \lambda(U(p)a + (1 - U(p))b) + (1 - \lambda)(U(q)a + (1 - U(q)b) \sim$$

$$[\lambda U(p) + (1 - \lambda)(U(q)]a + [\lambda(1 - U(p)) + (1 - \lambda)(1 - U(q))]b.$$

Once again, the definition of U implies $U(\lambda p+(1-\lambda)q) = \lambda U(p)+(1-\lambda)U(q)$. (Recall that $U(a) = 1$ and $U(b) = 0$.) So, affinity of U on $[a, b]$ is proved. Finally, (ii) implies that for all $p, q \in [a, b]$, $p \succ q$ iff $U(p) > U(q)$. I.e., U represents $\succsim$ on $[a, b]$.

Step 2. We will now complete the proof that the axioms imply representation for the case where $K = \Delta(X)$. First a **Notation**: we embed X in $\Delta(X)$. Thus we will identify $x \in X$ with $\delta_x \in \Delta(X)$ where $\delta_x(x) = 1$, and $\delta_x(z) = 0$ for $z \neq x$, $z \in X$. In these notations any $p \in \Delta(X)$ can be written as $p = \sum_{x \in X} p(x)x$. Equivalently, we can write $p = \sum_{i=1}^{k} p(x_i)x_i = \sum_{i=1}^{k} \lambda_i x_i$, where $p(x_i) := \lambda_i > 0$ for all i

Let a and b in X be such that for all $x \in X$, $a \succsim x \succsim b$. Such a and b exist because X is finite. Applying A3 or (iv), as the case may be, we get for $p = \sum_{i=1}^{k} \lambda_i x_i$ as above;

$$\lambda_1 a + (1 - \lambda_1) \textstyle\sum_{i=2}^{k} \frac{\lambda_i}{1-\lambda_1} x_i \succsim \sum_{i=1}^{k} \lambda_i x_i \succsim \lambda_1 b + (1 - \lambda_1) \sum_{i=2}^{k} \frac{\lambda_i}{1-\lambda_1} x_i.$$

Repeating the argument,

$$(\lambda_1 + \lambda_2)a + (1 - \lambda_1 - \lambda_2) \textstyle\sum_{i=3}^{k} \frac{\lambda_i}{1-\lambda_1-\lambda_2} x_i \succsim \sum_{i=1}^{k} \lambda_i x_i \succsim$$

$$(\lambda_1 + \lambda_2)b + (1 - \lambda_1 - \lambda_2) \textstyle\sum_{i=3}^{k} \frac{\lambda_i}{1-\lambda_1-\lambda_2} x_i.$$

Repeating the argument less than $|X|$ times we get $a \succsim p \succsim b$.

So, $\Delta(X) = [a, b]$. Hence, by step 1, there is an affine $U : \Delta(X) \to R$ representing $\succsim$. For each $x \in X$ define $u(x) = U(x)$. We have to show for any $p \in \Delta(X)$ that $U(p) = \sum_{x \in X} p(x)u(x)$. This is easily implied by (vi) and representation of p as $\sum_{x \in X} p(x)x$.

Step 3. First we show that for affine U and V on K,

$$U(p) \geq U(q) \Leftrightarrow V(p) \geq V(q) \text{ iff } V(\cdot) = \alpha U(\cdot) + \beta \text{ for some } \alpha, \beta \in \mathbb{R}\ ,\ \alpha > 0.$$

It is obvious that the right side implies the left side. Assume the left side of the iff implication. If U is constant, so is V and the implication holds. Suppose that for some a and b in K, $U(a) > U(b)$. Thus $V(a) > V(b)$. Define $W : K \to \mathbb{R}$ by:

$$W(p) = \tfrac{V(a)-V(b)}{U(a)-U(b)}[U(p) - U(a)] + V(a) =$$
$$\tfrac{V(a)-V(b)}{U(a)-U(b)}U(p) + [V(a) - \tfrac{V(a)-V(b)}{U(a)-U(b)}U(a)],$$

for $p \in K$. Thus $W(\cdot)$ is of the form, $\alpha U(\cdot) + \beta$, $\alpha > 0$. So it is affine and by the obvious implication, $U(p) \geq U(q) \Leftrightarrow W(p) \geq W(q)$. So, $V(p) \geq V(q) \Leftrightarrow W(p) \geq W(q)$. To complete the proof we show that $W(\cdot) = V(\cdot)$. Clearly, $W(a) = V(a)$ and $W(b) = V(b)$. For $p \in K$ one of the following three holds: $p \sim \lambda a + (1-\lambda)b$, $a \sim \xi p + (1-\xi)b$, $b \sim \varsigma a + (1-\varsigma)p$, with $\lambda \in [0, 1]$. and $\xi, \varsigma \in (0, 1)$. Thus $W(p) - V(p)$ for all p.

For the case $K = \Delta(X)$ the result follows because $\sum_{x \in X} p(x)u(x)$ is affine.

Step 4. Let us choose some c and d in K with $c \succ d$, (to be fixed throughout this step). We apply the result of step 1 to the $\succsim$-interval $[c, d]$. Thus there is an affine function $U : [c, d] \to \mathbb{R}$ representing $\succsim$ on its domain. We will extend U to an affine function, say V, on all of K such that V represents $\succsim$ on K.

Let $Q = \{\ [a, b] \subset K \mid a \succsim c \text{ and } d \succsim b\ \}$. For $[a, b] \in Q$, let $W^{ab} : [a, b] \to \mathbb{R}$ be affine and representing $\succsim$ on $[a, b]$ from step 1. Define $V^{ab} : [a, b] \to \mathbb{R}$ by $V^{ab}(e) = \frac{V(e)}{V(c)-V(d)} - \frac{V(d)}{V(c)-V(d)}$. Clearly, V^{ab} also is affine and represents $\succsim$ on $[a, b]$. Moreover, $V^{ab}(c) = 1$ and $V^{ab}(d) = 0$. thus V^{ab} extends U from $[c, d]$ to $[a, b]$.

If $e \in [a, b]$, then one of three holds: $e \in [c, d]$, $e \succ c$, $d \succ c$. In the first case $V^{ab}(e) = U(e)$. In the second case there is a unique $\rho \in (0, 1)$ such that, $c \sim \rho e + (1-\rho)d$. So, $V^{ab}(c) = \rho V^{ab}(e) + (1-\rho)V^{ab}(d)$, or equivalently $1 = \rho V^{ab}(e)$, or $V^{ab}(e) = \frac{1}{\rho}$. In the third case there is a unique $\tau \in (0, 1)$ such that, $d \sim \tau c + (1-\tau)e$. in this case $V^{ab}(e) = \frac{\tau}{\tau-1}$. In conclusion, if $e \in [a, b] \cap [g, h]$ for some $[a, b], [g, h] \in Q$ then $V^{ab}(e) = V^{gh}(e)$. So we define $V : K \to \mathbb{R}$ by, $V(e) = V^{ab}(e)$ for some, or equivalently all, $[a, b] \in Q$ such that $e \in [a, b]$.

For e and f in K there is $[a, b] \in Q$ with $e, f \in [a, b]$. (Choose a to be a first in $\succsim$ among $\{c, e, f\}$ and similarly let b be a last among $\{d, e, f\}$ in $\succsim$.) Then $V^{ab}(e) = V(e)$ and $V^{ab}(f) = V(f)$. Since V^{ab} is affine and represents $\succsim$ on $[a, b]$. Thus $V(e) \geq V(f)$ iff $e \succsim f$ and $V(\lambda e + (1-\lambda)f) = \lambda V(e) + (1-\lambda)V(f)$.

Thus step 4 and the direction from axioms to representation in the vNM Utility theorem have been proved.

Step 5. Obvious, and the proof of the theorems is complete.

In the sequel we will use a special form of the vNM theorem for products of lotteries. For a finite and nonempty set S we define $A = \Delta(X)^S$. Clearly A is a convex set and the vNM utility theorem applies to $K = A$. **Notation**: To avoid confusion, from now on, elements of A (and K) are denoted by $a, b, c, ..$; elements of $\Delta(X)$ are denoted by p, q; elements of S are denoted by s, t, r : and elements of X are denoted by x, y, z, w. So, if $a \in A$ then $a(s) \in \Delta(X)$, and $a(s)(x)$ is the probability that a assigns to x given s.

We can represent A equivalently as follows:

$$A = \{\ a : S \times X \to [0,1] \mid \forall s \in S,\ \sum_{x \in X} a(s,x) = 1\ \}.$$

In our notations $a(s)(x) = a(s,x)$.

Lemma. *Let an affine function $U : A \to R$ be given. Then there exist a function $u : S \times X \to R$ such that for all $a \in A$,*

$$U(a) = \sum_{s \in S} \sum_{x \in X} a(s,x) u(s,x).$$

Moreover, suppose that for some $v : S \times X \to R$, $V : A \to R$ is defined by,

$$V(a) = \sum_{s \in S} \sum_{x \in X} a(s,x) v(s,x).$$

Then for all a, b in A : $[U(a) \geq U(b) \Leftrightarrow V(a) \geq V(b)]$ iff there are numbers, $\alpha > 0$, and β_s for $s \in S$, such that for all $a(s,x) \in S \times X$, $v(s,x) = \alpha u(s,x) + \beta_s$.

Note that u of the lemma is covariant under a larger class of transformations than u and U of the previous theorems.

Comment. If $\succsim$ is a binary relation on A satisfying A1,A2, and A3 then, by the vNM utility theorem, there is an affine U representing it. The opposite direction holds too: an affine U represents a binary relation which satisfies A1,A2, and A3.

Proof. Let $S = \{s_1, s_2, ..., s_n\}$ where $|S| = n$. For $a \in A$ we will write $a = (p_1, p_2, ..., p_n)$ where $p_i \in \Delta(X)$ for $i = 1, ..., n$. Let $b = ((q_1, q_2, ..., q_n)$ be an arbitrary element of A, fixed for the rest of the proof,. For any $a = (p_1, p_2, ..., p_n) \in A$, and any i, let $a_i = (q_1, ..., q_{i-1}, p_i, q_{i+1}, ..., q_n)$.

Than $\frac{1}{n}a + \frac{n-1}{n}b = \sum_{i=1}^{n} \frac{1}{n} a_i$. By affinity of U and (vi),$\frac{1}{n}U(a) + \frac{n-1}{n}U(b) = \sum_{i=1}^{n} \frac{1}{n} U(a_i)$ $(\star)$. For all i define $U_i : \Delta(X) \to \mathbb{R}$ by

$$U_i(p) = U(q_1, ..., q_{i-1}, p, q_{i+1}, ..., q_n) - \frac{n-1}{n} U(b).$$

Thus for any $a \in A$, $U_i(p_i) = U(a_i) - \frac{n-1}{n}U(b)$. Summing over i and dividing by n we get, $\frac{1}{n}\sum_{i=1}^{n} U_i(p_i) = \frac{1}{n}\sum_{i=1}^{n} U(a_i) - \frac{n-1}{n}U(b)$. In view of $(\star)$ above we get, $U(a) = \sum_{i=1}^{n} U_i(p_i)$ $(\ddagger)$.
Since U_i coincides with U on a convex subset of A, it is affine. (Note that,

$$\Delta(X) \approx \{\ a \in A \mid a = (q_1, ..., q_{i-1}, p, q_{i+1}, ..., q_n),\ p \in \Delta(X)\ \}.)$$

For $x \in X$ and $i = 1, ..., n$, define, $u_i(x) = U_i(x)$. (Recall that x is same as δ_x.)
So, by (vi), $U_i(p) = \sum_{x \in X} p(x) u_i(x)$, and by (‡) above, $U(a) = \sum_{i=1}^{n} \sum_{x \in X} p(x) u_i(x)$.
QED

Comment: Consider the set,

$$Ex(A) = \{\ a \in A \mid a(s, x) \in \{0, 1\}\ \}.$$

The elements of $Ex(A)$ are the extreme points of A. For any n-list, $x_1, x_2, ..., x_n$ of elements of X, let $e^{x_1 x_2 \dots x_n}$ denote the element of $Ex(A)$ where $e^{x_1 x_2 \dots x_n}(s_i, x_i) = 1$ for $i = 1, ..., n$. Any $a \in A$ can be represented as a convex combination of elements of $Ex(A)$:

$$a = \sum_{x_1 x_2 \dots x_n} \Pi_{i=1}^{n} a(s_i, x_i) e^{x_1 x_2 \dots x_n},$$

where the sum runs over all $|X|^n$ distinct n-lists of elements of X.

November 2004

States of nature and uncertainty

The primitives of decisions under uncertainty consist of two nonempty sets, A' and X, and a correspondence φ which assigns a nonempty subset $\varphi(a')$ of X to every $a' \in A'$. We assume that $\cup_{a' \in A'} \varphi(a') = X$. The elements of A' are termed *feasible* or available *alternatives*. Next define $S = \{\ s : A' \to X \mid s(a') \in \varphi(a')\ \}$, and $A'' = \{\ a : S \to X\ \}$. The elements of S are termed *states* of Nature, or just states, and the elements of A'' are termed conceivable alternatives or just alternatives, (acts in Savage's terminology). Note that we can embed any element $a' \in A'$ in A by defining $a(s) = s(a')$ for all $s \in S$. Anscombe and Aumann, AA for short, whose model as restated by Fishburn we introduce here included additional conceivable alternatives. Thus the set of alternatives we have here is as in the lemma above is:

$$A = \{\ a : S \to \Delta(X)\ \}.$$

Finally, we assume existence of a binary relation, $\succsim$, on A. We need two additional axioms.

A4. Nondegeneracy: there are a^* and a_* in A such that $a^* \succ a_*$.

Notation: an event $E \subset S$ is said to be **null** if for any two alternatives, a and b which agree outside E, $a \sim b$. It is easy to see that if E and F are null events, then so is $E \cup F$. A4 implies that S is not null.

A5. State independence: Let $a = (p_1, p_2, ..., p_n) \in A$, and for some $s_i \in S$, and $p, q \in \Delta(X)$, $(p_1, p_2, ..., p_{i-1}, p, p_{i+1}, ..., p_n) \succ (p_1, p_2, ..., p_{i-1}, q, p_{i+1}, ..., p_n)$.
Then for any non-null $s_j \in S$,

$$(p_1, p_2, ..., p_{j-1}, p, p_{j+1}, ..., p_n) \succ (p_1, p_2, ..., p_{j-1}, q, p_{j+1}, ..., p_n).$$

Note that A5 together with A1 imply:

$$(p_1, p_2, ..., p_{i-1}, p, p_{i+1}, ..., p_n) \sim (p_1, p_2, ..., p_{i-1}, q, p_{i+1}, ..., p_n) \Rightarrow$$
$$(p_1, p_2, ..., p_{j-1}, p, p_{j+1}, ..., p_n) \sim (p_1, p_2, ..., p_{j-1}, q, p_{j+1}, ..., p_n).$$

Thus A5 means that preferences over consequences do not depend on the occurring state of nature.

AA Theorem. *A binary relation $\succsim$ on A satisfies A1, A2, A3, A4, and A5 iff there exist a unique probability, π on S and a (utility) function $u : X \to R$ such that,*

$$a \succsim b \Leftrightarrow \sum_{s \in S} \pi(s) \sum_{x \in X} a(s, x)u(x) \geq \sum_{s \in S} \pi(s) \sum_{x \in X} b(s, x)u(x).$$

Moreover, a function $v : X \to R$ satisfies

$$a \succsim b \Leftrightarrow \sum_{s \in S} \pi(s) \sum_{x \in X} a(s, x)v(x) \geq \sum_{s \in S} \pi(s) \sum_{x \in X} b(s, x)v(x)$$

iff $v(x) = \alpha u(x) + \beta$ on X for some $\alpha, \beta \in R$, $\alpha > 0$.

Proof. By the lemma of the previous section there exists a function $u : S \times X \to \mathbb{R}$ such that.

$$a \succsim b \Leftrightarrow \sum_{s \in S} \sum_{x \in X} a(s, x)u(s, x) \geq \sum_{s \in S} \sum_{x \in X} b(s, x)u(s, x).$$

It is obvious that if a state s is null than $u(s, \cdot)$ is constant, and vice-versa, if $u(s, \cdot)$ is constant then s is null. By the uniqueness part of the lemma we may assume that this constant is zero. Let us fix a non-null state, say r. Then for any other non-null state t, A5 with A1 imply that $u(t, \cdot)$ represent the same preferences as $u(r, \cdot)$. Hence, by the uniqueness part of the vNM theorem for lotteries, $u(t, \cdot) = \alpha(t)u(r, \cdot) + \beta(t)$ with $\alpha(t) > 0$. By the uniqueness part of the lemma we can choose $\beta(t) = 0$. If t is null then the same equality holds with $\alpha(t) = 0$. Defining $\alpha(r) = 1$ we have,

$$a \succsim b \Leftrightarrow \sum_{s \in S} \sum_{x \in X} a(s, x)\alpha(s)u(r, x) \geq \sum_{s \in S} \sum_{x \in X} a(s, x)\alpha(s)u(r, x).$$

To get the theorem's representation we define $u(\cdot) = u(r, \cdot)$, and $\pi(s) = \frac{\alpha(s)}{\sum_{t \in S} \alpha(t)}$. for all $s \in S$.

Uniqueness of π is the result of nondegeneracy. For a null state, say s, it must be that $\pi(s) = 0$. If s is non-null, and the representation holds with both, π and μ where, say, $\pi(s) > \mu(s)$, then for some other non-null state t, $\pi(t) < \mu(t)$. In this case it is easy to show that the two representation contradict one another. The other direction of the theorem and the relative uniqueness of u are obvious.

The state independence axiom, A5, in the AA model can be replaced by either one of the two axioms we will introduce now; monotonicity and strict monotonicity. To state these axioms we first have to **define** a new binary relation on $\Delta(X)$ which also will be denoted by $\succsim$. For p and q in $\Delta(X)$ we write $p \succsim q$ iff $a = (p,p,p,...,p) \succsim b = (q,q,q,...,q)$. Obviously, if $\succsim$ on A satisfies Ai, so does $\succsim$ on $\Delta(X)$ for i=1,2,3.

A6. Monotonicity: if $a(s) \succsim b(s)$ for all $s \in S$, then $a \succsim b$.

Note that applying A6 twice we get: if $a(s) \sim b(s)$ for all $s \in S$, then $a \sim b$.

A7. Strict monotonicity: if $a \succ b$ and for some $E \subset S$, $a(s) = p$ and $b(s) = q$ for $s \in E$, and $a(s) = b(s)$ for $s \notin E$, then $p \succ q$.

We will now state a claim connecting the two axioms.

(vii) If A1 holds, then A6 and A7 are equivalent.

Proof: Assume A6 and the antecedent of A7. We have to prove that $p \succ q$. Otherwise, by A1 (completeness), $q \succsim p$. So by A6, $b \succsim q$, a contradiction. Hence the consequent of A7 holds.

Assume now A7 and the antecedent of A6. We have to prove that $a \succsim b$. Define c^i, $i = 0,1,...,n$ by: $c^i(s_j) = b(s_j)$ when $j \leq i$, and $c^i(s_j) = a(s_j)$ otherwise. Clearly $c^0 = a$, and $c^n = b$. We will show that for $i \geq 1$, $c^{i-1} \succsim c^i$; A1(transitivity) will then complete the proof. If, by way of negation, $c^i \succ c^{i-1}$, than by A7, $b(s_i) \succ a(s_i)$, a contradiction. This completes our proof.

Comment: The AA theorem holds when A5 is substituted with A6 or A7.

Proof: We restrict attention to the main part of the theorem, that is, axioms imply representation of $\succsim$: Recall that by the lemma A1, A2, A3 already imply that:

$$a \succsim b \Leftrightarrow \sum_{s\in S}\sum_{x\in X} a(s,x)u(s,x) \geq \sum_{s\in S}\sum_{x\in X} a(s,x)u(s,x), \qquad (\heartsuit)$$

and moreover, for any null state s, set $u(s,\cdot) = 0$. A5 is used to show that for any non-null s and t, $u(s,\cdot)$ and $u(t,\cdot)$ represent the same preferences. Here we will show, using A6 and/or A7, that for any non-null state t, $u(t,\cdot)$ represents the relation $\succsim$ on $\Delta(X)$. Suppose that for some p and q in $\Delta(X)$, $\sum_{x\in X} p(x)u(t,x) > \sum_{x\in X} q(x)u(t,x)$. Let $a(s) = b(s)$ for $s \neq t$, and let $a(t) = p$ and $b(t) = q$. Then

$$\sum_{s\in S}\sum_{x\in X} a(s,x)u(s,x) > \sum_{s\in S}\sum_{x\in X} b(s,x)u(s,x),$$

which implies, by $(\heartsuit)$, $a \succ b$. The latter implies by A7 that $p \succ q$.

If for some p and q in $\Delta(X)$, $\sum_{x\in X} p(x)u(t,x) = \sum_{x\in X} q(x)u(t,x)$, we define $a(s) = p$, and $b(s) = q$ for all s. Then,

$$\sum_{s\in S}\sum_{x\in X} a(s,x)u(s,x) > \sum_{s\in S}\sum_{x\in X} a(s,x)u(s,x),$$

which implies, by $(\heartsuit)$, $a \sim b$. The latter implies by the definition of $\succsim$ on $\Delta(X)$, that $p \sim q$. The rest of the proof is as in AA theorem above.

Two alternatives, say a and b, are said to be **comonotonic** if for all $s,t \in$: $a(s) \succ a(t) \Rightarrow b(s) \succsim b(t)$, and $a(t) \succ a(s) \Rightarrow b(t) \succsim b(s)$.

A8. Comonotonic independence: Suppose that $a \succ b$, a and c, and b and c are comonotonic, and $0 < \lambda \leq 1$. Then $\lambda a + (1-\lambda)c \succ \lambda b + (1-\lambda)c$.

A function $v : S \to [0,1]$ is termed **nonadditive probability** if $v(\varnothing) = 0$, $v(S) = 1$, and $E \subset F \Rightarrow v(E) \leq v(F)$. For any $f : S \to R$ we define,

$$\int_S f dv = \sum_{i=1}^n [f(s_i) - f(s_{i+1})] v(\{s_1, s_2, ..., s_i\}),$$

where $s_1, s_2, ..., s_n$ is an ordering of the states in S, such that $f(s_1) \geq f(s_2) \geq ... \geq f(s_n)$, and $f(s_{n+1}) = 0$

DS Theorem. *A binary relation $\succsim$ on A satisfies A1, A2, A4, A6, and A8 iff there exist a unique nonadditive probability, v on S and a (utility) function $u : X \to \mathbb{R}$ such that,*

$$a \succsim b \Leftrightarrow \int_S [\sum_{x \in X} a(s,x)u(x)] dv \geq \int_S [\sum_{x \in X} b(s,x)u(x)] dv. \qquad (\blacklozenge)$$

Moreover, u is unique as in vNM or AA theorems.

For the next result we need another axiom on preferences.

A9. Uncertainty aversion: $a \succsim b$ implies $\lambda a + (1-\lambda) b \succsim b$ for all $\lambda \in [0,1]$.

Corollary. *If axiom 9 is added to the other axioms in DS theorem, then the resulting unique nonadditive probability v satisfies convexity (or supper linearity), i.e.,*

$$v(E) + v(F) \leq v(E \cup F) + v(E \cap F).$$

The opposite direction holds too, i.e., if v is convex then the relation $\succsim$ defined by ($\blacklozenge$) above satisfies A9.

Outline of the proof DS theorem. (Only the part: axioms imply representation.) Note first that any two constant (over S) alternatives are comonotonic. (Indeed, any alternative and a constant alternative are comonotonic.) So A1, A2, and A8 restricted to constant alternatives satisfy the vNM conditions. Thus it is also true for $\succsim$ on $\Delta(X)$, and by the vNM theorem for lotteries, there is a utility function, $u : X \to \mathbb{R}$ representing $\succsim$ on $\Delta(X)$.

Notations: (i) For $a \in A$, and $s \in S$ we write $U(a(s))$ instead of $\sum_{x \in X} a(s,x)u(x)$. (ii) For $p \in \Delta(X)$ we write p^* for the constant alternative $a : S \to \Delta(X)$, where for all s, $a(s) = p$. Similarly, for $\alpha \in \mathbb{R}$ we write α^* for the constant function $f : S \to \mathbb{R}$, where for all s, $f(s) = \alpha$. (iii) We denote by $x^\#$ a best element in X in the $\succsim$ ranking. If there is more than one, $x^\#$ is chosen arbitrarily, but remains fixed from now on. Similarly, $x_\#$ denotes a worst element of X in the $\succsim$ ranking. (Note that for all $p \in \Delta(X)$, $x^\# \succsim p \succsim x_\#$.)

We assume, without loss of generality, that $u(x^\#) = 1$, and $u(x_\#) = 0$. For any $a \in A$ let $V(a)$ denote the function from S to $\mathbb{R}$, where $V(a)(s) = U(a(s))$. Thus $V : A \to \mathbb{R}^S$. Note that if $a(s) \sim b(s)$, for some s, then $U(a(s)) = U(b(s))$. If $a(s) \sim b(s)$ for all s, then by monotonicity, (A6), $a \sim b$.

For any $a \in A$, there is a unique $\lambda \in [0,1]$, such that $a \sim \lambda x^{\#*} + (1-\lambda) x_\#^*$. This is a special case of claim (v) from the first theorem. The proof of (v) there uses A3. Proof of (i) uses A3 and the claims are proved consecutively where each proof uses the previous claims. But in our case we only deal with alternatives of the form $\theta x^{\#*} + (1-\theta)x^*_{\#}$, $\theta \in [0,1]$, and a. All are pairwise comonotonic, and A8 restricted to them coincides with A3.

Let us define a function $J : A \to \mathbb{R}$ by $J(a) = \lambda u(x^{\#}) + (1-\lambda)u(x_{\#}) = \lambda$, where $a \sim \lambda x^{\#*} + (1-\lambda)x^*_{\#}$. To close the triangle we define a function I from the codomain of V to R as follows: For $f : S \to \mathbb{R}$ with $1 \geq f(s) \geq 0$ for all s, there is an alternative a with $V(a)(s) = U(a(s)) = f(s)$; then $I(f) = J(a)$. The function I is well defined: If for some other alternative b, $V(b)(s) = U(b(s)) = f(s)$ for all s, then by monotonicity $a \sim b$, which, in turn, implies $J(b) = J(a)$.

By the definition of I, and by the axioms A6, A8, it is straightforward that I has the following properties: (1) $I(\alpha^*) = \alpha$ for any $\alpha \in [0,1]$. (2) If f, g, h are pairwise comonotonic, and $I(f) > I(g)$ for $0 < \lambda < 1$, then $I(\lambda f + (1-\lambda)h) > I(\lambda g + (1-\lambda)h)$. (3) If $f(s) \geq g(s)$ for all s then $I(f) \geq I(g)$. We show that these 3 properties imply that if f, g are comonotonic, $0 < \lambda < 1$, then $I(\lambda f + (1-\lambda)g) = \lambda I(f) + (1-\lambda)I(g)$.
Denote $I(f) = \alpha$, $I(g) = \beta$ and let $\delta > 0$ be small. (3) and (1) imply that $I(f) < I((\alpha+\delta)^*)$, $I(g) < I((\beta+\delta)^*)$. Therefore,

$$\begin{aligned}\lambda I(f) + (1-\lambda)I(g) + \delta = I(\lambda(\alpha+\delta)^* + (1-\lambda)(\beta+\delta)^*) > \\ > I(\lambda f + (1-\lambda)(\beta+\delta)^*) > I(\lambda f + (1-\lambda)g),\end{aligned}$$

where we used the comonotonicity of f, g and of f and g with the constant functions. Taking the limit as δ goes to 0 yields $\lambda I(f) + (1-\lambda)I(g) \geq I(\lambda f + (1-\lambda)g)$. Repeating the same argument, this time with $\delta < 0$, gives the other inequality and the claim is proved.

Next, let θ be a non-negative number. We show that I is homogeneous, that is $I(\theta f) = \theta I(f)$ (of course we only consider cases where $0 \leq \theta f \leq 1$). If $0 \leq \theta \leq 1$ then $I(\theta f) = I(\theta f + (1-\theta)0^*) = \theta I(f) + (1-\theta)0 = \theta I(f)$. If $\theta > 1$ then $I(f) = I(\frac{1}{\theta}\theta f + (1-\frac{1}{\theta})0^*) = \frac{1}{\theta} I(\theta f)$ so again $I(\theta f) = \theta I(f)$. Note, that this implies that I is additive on comonotonic functions. Indeed, if f, g are comonotonic then $I(f+g) = I(2\frac{f+g}{2}) = 2I(\frac{1}{2}f + \frac{1}{2}g) = 2\frac{I(f)+I(g)}{2} = I(f) + I(g)$.

For any $E \subset S$, define $v(E) = I(\mathbf{1}_E)$, where $\mathbf{1}_E(s) = 1$ if $s \in E$, and 0 otherwise. We will show that if v is defined as above, then $I(f) = \int_S f dv$ for any function f. The proof is by induction on k - the number of different values (and different from 0) that f attains. If $k = 1$ then there is some non-empty set $E \subset S$ and a number $0 < \alpha \leq 1$ such that $f(s) = \alpha \mathbf{1}_E(s)$. Therefore, $I(f) = I(\alpha \mathbf{1}_E) = \alpha I(\mathbf{1}_E) = \alpha v(E) = \int f dv$ (recall that $\int f dv = \sum_{i=1}^n [f(s_i) - f(s_{i+1})] v(\{s_1, ..., s_i\})$ where the states are in descending order, $f(s_1) \geq \ldots \geq f(s_n) \geq f(s_{n+1}) = 0$). Assume that the equality holds for functions with $k-1$ different positive values and let f have k different positive values. This means that there are k non-empty and disjoint in pairs subsets $E_1, \ldots, E_k$, and positive numbers $1 \geq \alpha_1 > \alpha_2 > \ldots > \alpha_k > 0$ such that $f = \sum_{i=1}^k \alpha_i \mathbf{1}_{E_i}$. Define $g = \sum_{i=1}^{k-1} (\alpha_i - \alpha_k) \mathbf{1}_{E_i}$ and $h = \alpha_k \sum_{i=1}^k \mathbf{1}_{E_i}$. Obviously $f = g + h$ and g, h are comonotonic, and so $I(f) = I(g) + I(h)$. Notice that g has only $k-1$ different positive values, so by the induction hypothesis,

$$I(g) = \int g dv = \sum_{i=1}^{k-1}[(\alpha_i - \alpha_k) - (\alpha_{i+1} - \alpha_k)]v(E_1 \cup \ldots \cup E_i) =$$
$$= \sum_{i=1}^{k-1}[\alpha_i - \alpha_{i+1}]v(E_1 \cup \ldots \cup E_i).$$

Also, $I(h) = \alpha_k v(E_1 \cup \ldots \cup E_k)$. Summing up, we get,

$$I(f) = I(g) + I(h) = \sum_{i=1}^{k-1}[\alpha_i - \alpha_{i+1}]v(E_1 \cup \ldots \cup E_i) + \alpha_k v(E_1 \cup \ldots \cup E_k) =$$
$$\sum_{i=1}^{k}[\alpha_i - \alpha_{i+1}]v(E_1 \cup \ldots \cup E_i) = \int f dv.$$

QED

Proof of the Corollary. (only the direction A9 $\Rightarrow$ convexity of v.) Since I represents $\succsim$, $a \succsim b$ is equivalent to $I(V(a)) \geq I(V(b))$. It follows that $I(f) \geq I(g) \Rightarrow I(\alpha f + (1-\alpha)g) \geq I(g)$ for all $\alpha \in [0,1]$. In particular, if $I(f) = I(g)$ then $I(f+g) = 2I(0.5f + 0.5g) \geq 2I(g) = I(f) + I(g)$.

Let E, F be two subsets of S, and assume without loss of generality that $v(E) \geq v(F)$. Then there is $\theta \geq 1$ such that $v(E) = \theta v(F)$. Therefore, $I(\mathbf{1}_E) = v(E) = \theta v(F) = I(\theta \mathbf{1}_F)$, so by the previous paragraph, $I(\mathbf{1}_E + \theta \mathbf{1}_F) \geq v(E) + \theta v(F)$. However, we can rewrite $\mathbf{1}_E + \theta \mathbf{1}_F$ as $\mathbf{1}_{E \cap F} + (\theta - 1)\mathbf{1}_F + \mathbf{1}_{E \cup F}$. Notice that each pair of functions in the sum are comonotonic, so we can use additivity to obtain,

$$v(E) + \theta v(F) \leq I(\mathbf{1}_{E \cap F} + (\theta-1)\mathbf{1}_F + \mathbf{1}_{E \cup F}) = v(E \cap F) + (\theta - 1)v(F) + v(E \cup F),$$

so v is convex. QED

A10. Certainty Independence: Suppose that $a \succ b$, c is a constant alternative, and $0 < \lambda \leq 1$. Then $\lambda a + (1-\lambda)c \succ \lambda b + (1-\lambda)c$.

Note that any alternative and a constant alternative are comonotonic. Thus A8 implies A10.

GS Theorem. *A binary relation $\succsim$ on A satisfies A1, A2, A4, A6, A9, and A10 iff there exist a unique convex and compact subset of* $\Delta(S)$, say K, *and a (utility) function $u : X \to \mathbb{R}$ such that,*

$$a \succsim b \Leftrightarrow \min_{P \in K} \sum_{s \in S} P(s)[\sum_{x \in X} a(s,x)u(x)]dv \geq$$
$$\min_{P \in K} \sum_{s \in S} P(s)[\sum_{x \in X} b(s,x)u(x)]dv. \qquad (\spadesuit)$$

Moreover, u is unique as in vNM or AA theorems.

Outline of the proof GS theorem. (Only the part: axioms imply representation.) The first part of the proof of the DS theorem, including the construction of the functionals J and I applies here. It is so because the use of the independence axiom, A3, in the proof is restricted to constant (over S) functions. This restricted form of independence follows from A10 (the certainty independence). The functionals J and I just constructed hold somewhat different properties than J and I in the proof of the DS theorem. Unlike in the previous proof we normalize the utility function u so that $u(x^{\#}) > 1$, and $u(x_{\#}) < -1$. Thus

$J(a) = \lambda u(x^{\#}) + (1-\lambda)u(x_{\#})$, where $a \sim \lambda x^{\#*} + (1-\lambda)x^*_{\#}$. I is defined through J as previously.

By the definition of I, and by axioms A6, A9, and A10, it is easy to see that I has the following properties: (1) $I(\alpha \mathbf{1}_S) = \alpha$ for any $\alpha \in [0,1]$. (2) If $f(s) \geq g(s)$ for all s then $I(f) \geq I(g)$. (3) $I(\lambda f) = \lambda I(f)$ for $\lambda > 0$. (4) $I(f+g) = I(f) + I(g)$. (5) $I(f + \lambda \mathbf{1}_S) = I(f) + \lambda$. These properties hold in the domain of definition of I. However, because of homogeneity (3) and monotonicity (2), I can be extended, through homogeneity, to all of $\mathbb{R}^S$.

Main Lemma. *Suppose that $I : \mathbb{R}^S \to \mathbb{R}$ satisfies properties (1),...,(5) above. Then there exist a closed and convex subset, say K, of $\Delta(S)$ such that $I(f) = \min_{P \in K} \sum_{s \in S} P(s)f(s)$ for all f .*

Proof. Let $f \in \mathbb{R}^S$ with $I(f) > 0$ be given. We will construct a probability measure P_f such that $I(f) = P_f \cdot f$ and $I(g) \leqq g \cdot P_f$ for all $g \in \mathbb{R}^S$. ("·" denotes the inner product in $\mathbb{R}^S$, i.e., $P_f \cdot f. = \sum_{s \in S} P_f(s)f(s)$.). To this end we define

$$D_1 = \{\ g \in \mathbb{R}^S \mid I(g) > 1\ \}\ ,$$

$$D_2 = \mathrm{conv}(\{\ g \in \mathbb{R}^S \mid g \leqq \mathbf{1}_S\ \} \cup \{\ g \in \mathbb{R}^S \mid g \leqq f/I(f)\ \})\ .$$

We now show that $D_1 \cap D_2 = \emptyset$. Let $d_2 \in D_2$ satisfy $d_2 = \alpha g_1 + (1-\alpha)g_2$ where $g_1 \leq \mathbf{1}_S$, $g_2 \leqq (f/I(f))$ and $\alpha \in [0,1]$. By monotonicity, homogeneity and C-independence of I,

$$I(d_2) \leqq \alpha + (1-\alpha)I(g_2) \leqq 1\ .$$

Note that each of the sets D_1, D_2 has an interior point and that they are both convex. Thus, by a separation theorem there exists a non-zero vector P_f and an $\alpha \in \mathbb{R}$ such that:

$$\text{for all } d_1 \in D_1 \text{ and } d_2 \in D_2, \quad P_f \cdot d_1 \leqq \alpha \leqq P_f \cdot d_2\ .$$

Since the unit ball of $\mathbb{R}^S$ is included in D_2, $\alpha > 0$. (Otherwise P_f would have been identically zero). We may therefore assume without loss of generality that $\alpha = 1$.

By (1), $P_f \cdot \mathbf{1}_S \leqq 1$. Since $\mathbf{1}_S$ is a limit point of D_1, $P_f \cdot \mathbf{1}_S \geqq 1$ is also true, hence $P_f \cdot \mathbf{1}_S = 1$. We now show that P_f is non-negative, or more specifically, that $P_f \cdot \mathbf{1}_E \geqq 0$. Since

$$P_f \cdot \mathbf{1}_E + P_f \cdot (\mathbf{1}_S - \mathbf{1}_E) = P_f \cdot \mathbf{1}_S = 1\ ,$$

and $\mathbf{1}_S - \mathbf{1}_E \in D_2$, the inequality follows.

We will now show that $g \cdot P_f \geqq I(g)$ for all $g \in \mathbb{R}^S$, with equality for $g = f$. First assume $I(g) > 0$. It is easily seen that $g/I(g) + (1/n)\mathbf{1}_S \in D_1$, so in the limit we get, $g \cdot P_f \geqq I(g)$. For the case $I(g) \leqq 0$ the inequality follows from C-independence. Since $f/I(f) \in D_2$, we obtain the converse inequality for f, thus $f \cdot P_f = I(f)$.

We now define the set C as the closure of the convex hull of $\{P_f \mid I(f) > 0\}$ (which, of course, is convex). It is easy to see that $I(g) \leqq \min\{\int gdP \mid P \in C\}$. For a such that $I(g) > 0$ we have shown the converse inequality holds as well. For a such that $I(g) \leqq 0$, it is again a simple implication of C-independence. □

The conclusion of the proof of Theorem is obvious. For a and b we have, $a \succsim b \Leftrightarrow J(a) \geq J(b)$. Setting $f(s) = U(a(s))$ and $g(s) = U(b(s))$ for all s, we have by the definition of I, $a \succsim b \Leftrightarrow I(f) \geq J(g)$. Thus the main lemma yields the required result.

Mid December 2004

Case-Based approach.

The primitives of our model consist of two non-empty and finite sets X and $\mathbb{T}$. We interpret X as the set of all conceivable *eventualities* in a given prediction problem, p, whereas $\mathbb{T}$ represents the set of *types* of *cases*. To simplify notation, we suppress the prediction problem p whenever possible. The predictor is equipped with a finite set of cases, her *memory* (or database), and her task is to rank the eventualities by a binary relation, "at least as likely as".

Memories are represented by vectors of non-negative integers, counting how many cases of each type appear in memory. Formally, $\mathbb{Z}_+^{\mathbb{T}} = \{I : T \to \mathbb{Z}_+ \mid I \neq 0\}$ where $\mathbb{Z}_+$ stands for the non-negative integers. $I \in \mathbb{Z}_+^{\mathbb{T}}$ is interpreted as a counter vector, where $I(t)$ counts how many cases of type t appear in the memory represented by I.

For $I \in \mathbb{Z}_+^{\mathbb{T}}$, let $\succsim_I \subset X \times X$ denote the predictors binary relation, "at least as likely as"given the memory I.

We will now present the axioms on $\{\succsim_I\}_{I \in \mathbb{Z}_+^{\mathbb{T}}}$.

A1* Order: For every $I \in \mathbb{Z}_+^{\mathbb{T}}$, $\succsim_I$ is complete and transitive on X.

A2* Combination: For every $I, J \in \mathbb{Z}_+^{\mathbb{T}}$ and every $x, y \in X$, if $x \succsim_I y$ ($x \succ_I y$) and $x \succsim_J y$, then $x \succsim_{I+J} y$ ($x \succ_{I+J} y$).

A3* Archimedeanity: For every $I, J \in \mathbb{Z}_+^{\mathbb{T}}$ and every $x, y \in X$, if $x \succ_I y$, then there exists $l \in N$ such that $x \succ_{lI+J} y$.

Observe that in the presence of Axiom 2, Axiom 3 also implies that for every $I, J \in \mathbb{Z}_+^{\mathbb{T}}$ and every $x, y \in X$, if $x \succ_I y$, then there exists $l \in \mathbb{N}$ such that for all $k \geq l$, $x \succ_{kI+J} y$.

A4* Diversity: For every list (x, y, z, w) of distinct elements of X there exists $I \in \mathbb{Z}_+^{\mathbb{T}}$ such that $x \succ_I y \succ_I z \succ_I w$. If $|X| < 4$, then for any strict ordering of the elements of X there exists $I \in \mathbb{Z}_+^{\mathbb{T}}$ such that $\succ_I$ is that ordering.

Definition: A matrix $v : X \times \mathbb{T} \to \mathbb{R}$, where $|X| \geq 4$, is *diversified* if there are no distinct four elements $x, y, z, w \in X$ and $\lambda, \mu, \theta \in \mathbb{R}$ with $\lambda + \mu + \theta = 1$ such that $v(x, \cdot) \leq \lambda v(y, \cdot)) + \mu v(z, \cdot) + \theta v(w, \cdot)$. If $|X| < 4$, v is diversified if no row in v is dominated by an affine combination of the others. (Linear combination where the sum of the coefficients is 1, is termed *affine combination*.)

Theorem 1 *: Let there be given X, $\mathbb{T}$, and a family of binary relations,* $\{\succsim_I\}_{I\in\mathbb{Z}_+^{\mathbb{T}}}$. *Then the following two statements are equivalent:*

(i) $\{\succsim_I\}_{I\in\mathbb{Z}_+^{\mathbb{T}}}$ *satisfy A1*-A4*;*

(ii) There is a diversified matrix $v : X \times \mathbb{T} \to \mathbb{R}$ *such that:*

$$(\circledast) \quad \begin{cases} \text{for every } I \in \mathbb{Z}_+^{\mathbb{T}} \text{ and every } x, y \in X, \\ x \succsim_I y \quad \text{iff} \quad \sum_{t\in\mathbb{T}} I(t)v(x,t) \geq \sum_{t\in\mathbb{T}} I(t)v(y,t) \ , \end{cases}$$

Furthermore*, in this case the matrix* v *is unique in the following sense:* v *and* u *both satisfy* $(\circledast)$ *iff there are a scalar* $\lambda > 0$ *and a matrix* $w : X \times \mathbb{T} \to \mathbb{R}$ *with identical rows (i.e., with constant columns) such that* $u = \lambda v + w$.

Derivation of Utility in the Context of a Game

Assume that a decision maker is facing a decision problem with a non-empty set of *acts* A and a finite, non-empty set of *states of the world* Ω. Such problems are often represented by a "decision matrix", or a "game against nature", attaching an outcome to each act-state pair (a, ω). We do not assume any knowledge about this set of outcomes or about the structure of the matrix, and hence suppress it completely. (Equivalently, one introduces a formal set of abstract outcomes that is simply the set of pairs $A \times \Omega$.) Let $\Delta = \Delta(\Omega)$ be the set of *probability distributions* on Ω. We assume that, for every probability vector $p \in \Delta$, the decision maker has a binary preference relation $\succsim_p$ over A. We now formulate axioms on $\{\succsim_p\}_{p\in\Delta}$:

A1 Order:** For every $p \in \Delta$, $\succsim_p$ is complete and transitive on A.

A2 Combination:** For every $p, q \in \gneqq$ and every $a, b \in A$, if $a \succsim_p b$ $(a \succ_p b)$ and $a \succsim_q b$, then $a \succsim_{\alpha p+(1-\alpha)q} b$ $(a \succ_{\alpha p+(1-\alpha)q} b)$ for every $\alpha \in (0, 1)$.

A3 Archimedean Axiom:** For every $a, b \in A$ and $p \in \Delta$, if $a \succ_p b$, then for every $q \in \Delta$ there exists $\alpha \in (0, 1)$ such that, $a \succ_{\alpha p+(1-\alpha)q} b$.

A4 Diversity:** For every list (a, b, c, d) of distinct elements of A there exists $p \in \Delta$ such that $a \succ_p b \succ_p c \succ_p d$. If $|A| < 4$, then for any strict ordering of the elements of A there exists $p \in \Delta$ such that $\succ_p$ is that ordering.

We need the following definition: a matrix of real numbers is called *diversified* if no row in it is dominated by an affine combination of three (or less) other rows in it. Formally:

Definition: A matrix $u : A \times Y \to \mathbb{R}$, where $|A| \geq 4$, is *diversified* if there are no distinct four elements $a, b, c, d \in A$ and $\lambda, \mu, \theta \in \mathbb{R}$ with $\lambda + \mu + \theta = 1$ such that $u(a, \cdot) \leq \lambda u(b, \cdot) + \mu u(c, \cdot) + \theta u(d, \cdot)$. If $|A| < 4$, u is diversified if no row in u is dominated by an affine combination of the others.

Theorem 2 *: Given a family of binary relations $\{\succsim_p\}_{p\in\Delta}$, the following two statements are equivalent:*

*(i) $\{\succsim_p\}_{p\in\Delta}$ satisfies A1** - A4**;*

(ii) There is a diversified matrix $u : A \times \Omega \to R$ such that:

$$(\spadesuit) \qquad \begin{array}{c} \text{for every } p \in \Delta \text{ and every } a, b \in A, \\ a \succsim_p b \quad \text{iff} \quad \sum_{\omega\in\Omega} p(\omega)u(a,\omega) \geq \sum_{\omega\in\Omega} p(\omega)u(b,\omega) \ , \end{array}$$

***Furthermore**, in this case the matrix u is unique in the following sense: u and w both satisfy ($\spadesuit$) iff there are a scalar $\lambda > 0$ and a matrix $v : A \times \Omega \to R$ with identical rows (i.e., with constant columns) such that $w = \lambda u + v$.*

Subjective Distributions

Assume that a decision maker is facing a decision problem with a finite and non-empty set of *acts, A*. Each act will result in one (and only one) physical *outcome* from the set $N = \{1, ..., n\}$ for $n \geq 1$. A *context* is a real-valued function on N. The space of all contexts, $\mathbb{R}^N$, is identified with $\mathbb{R}^n$, endowed with the natural topology and the standard algebraic operations. Given a context $x \in \mathbb{R}^n$, $\succsim_x \subset A \times A$ is a binary relation over acts.

The interpretation is as follows. The physical outcomes are abstract, and they do not determine the decision maker's utility. Rather, it is the context x which associates a utility value to each possible outcome. Put differently, the set of contexts is the set of possible utility functions on the abstract set of outcomes N. As in the introduction, an outcome might be, for instance, "a ball drawn from urn 1 is red". But, as opposed to the example used above, we do not assume that different acts have disjoint sets of possible outcomes. Rather, every physical outcome might, a priori, result from any act. It is assumed that we can observe the decision maker's preferences over acts given *any* utility function.

We now formulate axioms on $\{\succsim_x\}_{x\in\mathbb{R}^n}$:

A1* Order:** For every $x \in \mathbb{R}^n$, $\succsim_x$ is complete and transitive on A.

A2* Additivity:** For every $x, y \in \mathbb{R}^n$ and every $a, b \in A$, if $a \succsim_x b$ and $a \succ_y b$, then $a \succ_{x+y} b$.

A3* Continuity:** For every $a, b \in A$ the sets $\{x \mid a \succ_x b\}$ and $\{x \mid b \succ_x a\}$ are open.

A4* Diversity:** For every list (a, b, c, d) of distinct elements of A there exists $x \in \mathbb{R}^n$ such that $a \succ_x b \succ_x c \succ_x d$. If $|A| < 4$, then for any strict ordering of the elements of A there exists $x \in \mathbb{R}^n$ such that $\succ_x$ is that ordering.

A5* Neutrality:** For every constant $c \in \mathbb{R}^n$ (i.e., $c_i = c_j$ for all $i, j \in N$), and every $a, b \in A$, $a \sim_c b$.

Axiom 1 is standard. Axiom 2 is the most crucial axiom, as it guarantees that the set of contexts (utility functions) for which act a is preferred to act b is convex. Axiom 3 states that this set is also open. The diversity axiom (A4) rules out certain preferences. For instance, it does not allow one lottery to be always preferred to another. Finally, A5 is a weak consequentialism axiom. It holds whenever the decision maker cares only about the final utility derived from outcomes: if this utility function happens to be constant, no act should be preferred to any other.

The statement of the theorem requires two additional definitions. A function $P : N \to [0,1]$ with $\sum_{i \in N} P(i) = 1$ is called a *lottery*. Algebraic operations on lotteries are performed pointwise. In particular, the α-mixture of lotteries P and Q, $\alpha P + (1-\alpha)Q$ is also a lottery. A collections of lotteries $\{P_a\}_{a \in A}$ is called *4-independent* if every four (or fewer) lotteries in it are linearly independent.(That is, if $|A| < 4$, $\{P_a\}_{a \in A}$ is 4-independent if it is linearly independent.)

Theorem 3 *: Given a family of binary relations $\{\succsim_x\}_{x \in \mathbb{R}^n}$, the following two statements are equivalent:*

*(i) $\{\succsim_x\}_{x \in \mathbb{R}^n}$ satisfies A1*** - A4***;*

(ii) There is a collection of 4-independent lotteries $P = \{P_a\}_{a \in A}$ such that:

$$(\blacktriangleright) \quad \begin{array}{c} \text{for every } x \in \mathbb{R}^n \text{ and every } a, b \in A, \\ a \succsim_x b \quad \text{iff} \quad \sum_{i \leq n} P_a(i)x(i) \geq \sum_{i \leq n} P_b(i)x(i) \ , \end{array}$$

To what extent are the lotteries $\{P_a\}_{a \in A}$ unique? Clearly, if $\{P_a\}_{a \in A}$ satisfy $(*)$, then, for any lottery R and any $\alpha \in (0,1]$, the collection $\{\alpha P_a + (1-\alpha)R\}_{a \in A}$ also satisfies $(*)$. Observe that in $\{P_a\}_{a \in A}$ differences between lotteries are more pronounced than in $\{\alpha P_a + (1-\alpha)R\}_{a \in A}$. This gives rise to the following definition.

For two collections of lotteries, $P = \{P_a\}_{a \in A}$ and $Q = \{Q_a\}_{a \in A}$, we say that P is *more extreme than* Q if there exists a lottery R and $\alpha \in (0,1)$ such that $\alpha P_a + (1-\alpha)R = Q_a$ for all $a \in A$. We can now state the uniqueness result.

Proposition 4 *There exists a unique collection of 4-independent lotteries $P = \{P_a\}_{a \in A}$ that satisfies $(\blacktriangleright)$ and that is more extreme than any other collection $Q = \{Q_a\}_{a \in A}$ that satisfies $(\blacktriangleright)$.*