Preference, Choice and Utility

Eric Pacuit*

February 2, 2016

1 Relations

Suppose that X is a non-empty set. The set $X \times X$ is the **cross-product** of X with itself. That is, it is the set of all pairs of elements (called **ordered pairs**) from X. For example, if $X = \{a, b, c\}$, then

$$X \times X = \{(a, a), (a, b), (a, c), (b, a), (b, b), (b, c), (c, a), (c, b), (c, c)\}$$

A **relation** R on a set X is a subset of $X \times X$ (the set of pairs of elements from X). Formally, R is a relation on X means that $R \subseteq X \times X$. It is often convenient to write $a\ R\ b$ for $(a, b) \in R$. To help appreciate this definition, consider the following example. Suppose that X is the set of people in a room. Further, suppose that everyone in the room is pointing at some person in the room. A relation can be used to describe who is pointing at whom, where for $a, b \in X$, $a\ R\ b$ means that person a is pointing at person b. A second example of a relation is "taller-than", denoted $T \subseteq X \times X$, where $a\ T\ b$ means that person a is taller than person b. Typically, we are interested in relations satisfying special properties.

Definition 1.1 (Key Properties of relations) Suppose that $R \subseteq X \times X$ is a relation.

- R is **reflexive** provided for all $a \in X$, $a\ R\ a$
- R is **irreflexive** provided for all $a \in X$, it is not the case that $a\ R\ a$
- R is **complete** provided for all $a, b \in X$, $a\ R\ b$ or $b\ R\ a$ (or both)
- R is **symmetric** provided for all $a, b \in X$, if $a\ R\ b$ then $b\ R\ a$
- R is **asymmetric** provided for all $a, b \in X$, if aRb then not-$b\ R\ a$ $\triangleleft$
- R is **transitive** provided for all $a, b, c \in X$, if $a\ R\ b$ and $b\ R\ c$ then $a\ R\ c$

*Department of Philosophy, University of Maryland, College Park. *Web*: `pacuit.org`, *Email*: `epacuit@umd.edu`

Remark 1.2 *As stated, completeness implies reflexivity (let $a = b$ in the above statement). Often, one states completeness as follows: for all* distinct *$a, b \in X$, aRb or bRa. In what follows, we will use the above stronger definition of completeness where completeness implies reflexivity.*

Recall the example of a relation R that describes people pointing at other people in the room. If R is reflexive, then this means everyone is pointing at themselves. If R is irreflexive, then this means that no-one is pointing at themselves. This example illustrates the fact that irrelexivity is *not* the negation of reflexivity. That is, there are examples of relations that are neither reflexive nor irreflexive. If R is complete, then this means that every person in the room is either pointing at somebody or being pointed at. Symmetry of R means that every person that is being pointed at is pointing back at the person pointing at them. Asymmetry of R means that nobody is pointing back at the person pointing at them. Similar to the relationship between reflexivity and irreflexivity, asymmetry is *not* the negation of symmetry. Finally, picturing transitivity of the relation R is a bit more complicated. If the relation R is transitive, then everyone is pointing at the person that is being pointed to by the person that they are pointing at.

Exercise 1 Suppose that there are 5 people in a room. Draw a picture of a situation where the people are pointing at each other and the relation that describes the situation is transitive.

Exercise 2 What properties does the "better-than" relations satisfy?

Notation 1.3 (Describing Relations) Suppose that $R \subseteq X \times X$ is a relation. We will often use the following shorthand to denote elements in the relation: If $x_1, \ldots, x_n \in X$, then

$$x_1 \; R \; x_2 \; R \cdots x_{n-1} \; R \; x_n$$

means that for all $i = 1, \ldots, n-1$, $(x_i, x_{i+1}) \in R$ or $(x_i, x_j) \in R$ for all $j < i$ if R is assumed to be transitive (or $j \leq i$ if R is assumed to also be reflexive). For example, if R is transitive and reflexive, then $a \; R \; b \; R \; c$ means that $(a, a), (a, b), (b, b), (a, c), (b, c), (c, c) \in R$. $\triangleleft$

The following two definitions will play an important role in this course.

Definition 1.4 (Cycle) A **cycle** in a relation $R \subseteq X \times X$ is a set of distinct elements $x_1, x_2, \ldots, x_n \in X$ such that for all $i = 1, \ldots, n-1$, $x_i \; R \; x_{i+1}$, and $x_n \; R \; x_1$. A relation R is said to be **acyclic** if there is no cycles. $\triangleleft$

Definition 1.5 (Maximal Elements) Suppose that X is a set and $S \subset X$. An element $a \in S$ is **maximal** provided there is no $b \in S$ such that $b \; R \; a$. Let $\max_R(S)$ be the set of maximal elements of Y. $\triangleleft$

Exercise 3 Suppose that X has three elements (i.e., $X = \{a, b, c\}$. How many cycles can be formed from elements in X?

Exercise 4 Is it possible to find a relation that has a cycle and a non-empty set of maximal elements? What about a relation that has a cycle, a non-empty set of maximal elements, and is complete and transitive?

Exercise 5 Prove that if R is acyclic, then $\max_R(Y) \neq \emptyset$. Is the converse true? (Why or why not?)

Relations are an important mathematical tool used throughout Economics, Logic and Philosophy. We will use relations to describe a decision maker's *preferences* over a set of objects X. This means, among other things, that a decision maker's preferences are *comparative*. So, if we say that the decision maker "prefers red wine", then this means that the decision maker prefers red wine to the other available alternatives (e.g., red wine more than white wine). In most cases, we will assume that the relations representing a decision maker's preference satisfies certain properties:

Definition 1.6 (Rational Preference) A relation $R \subseteq X \times X$ is called a (**rational) preference ordering** provided R is complete and transitive. $\triangleleft$

Unless stated otherwise, we always assume that a decision maker's preference on a set X is a complete and transitive ordering on X.

Exercise 6 The assumption that a decision maker's preferences are complete and transitive is not uncontroversial. Find arguments for and against these assumptions.

In order to simplify our discussion, we will make use of the following notation.

Notation 1.7 (Strict Preference, Indifference) If R is a preference on X. Let $P_R \subseteq X \times X$ denote the strict sub-relation of R and $I_R \subseteq X \times X$ the indifference sub-relation of R defined as follows:

- for all $a, b \in X$, $a \; P_R \; b$ iff $a \; R \; b$ and it is not the case that $b \; R \; a$
- for all $a, b \in X$, $a \; I_R \; b$ iff $a \; R \; b$ and $b \; R \; a$ $\triangleleft$

Notation 1.8 (Preferences) We will use $\succeq$ to denote a rational preference ordering with $\succ$ denoting the associated strict preference ordering and $\sim$ the associated indifference ordering. So, $a \succeq b$ means "the decision maker (weakly) prefers a to b", $a \succ b$ means "the decision maker strictly prefers a to b" and $a \sim b$ means "the decision maker is indifferent between a and b". $\triangleleft$

Exercise 7 Give an example of something about which you have a weak preference, something about which you have a strict preference and something about which you are indifferent.

The following Lemma gathers some important facts about rational preferences. These facts will be used quite often without explicit reference. The proofs of these facts are immediate from the definitions.

Lemma 1.9 *Suppose that $\succeq$ is a rational preference ordering on X. Then,*

- *$\succ$ is transitive and irreflexive*
- *$\sim$ is an equivalence relation (reflexive, transitive and symmetric)*
- *For all $a, b, c \in X$, if $a \succeq b$ and $b \succ c$, then $a \succ c$*
- *For all $a, b, c \in X$, if $a \sim b$ and $b \succ c$, then $a \succ c$*
- *$\succ$ is acyclic*
- *For all $a, b \in X$, either $a \succ b$, $b \succ a$ or $a \sim b$*
- *For all $a, b \in X$, if $a \not\succ b$, then $b \succeq a$*

There is an alternative way to characterize rational preferences which is used in some texts. I conclude this brief introduction to relations by giving the details of this equivalent approach to defining a rational preference.

Definition 1.10 (Negative Transitivity) A relation $R \subseteq X \times X$ is **negatively transitive** provided for all $a, b, c \in X$, if $a \not R\, b$ (this means that it is not the case that $a\ R\ b$, i.e., $(a, b) \notin R$) and $b \not R\, c$, then $a \not R\, c$. $\triangleleft$

Fact 1.11 *A relation R is negatively transitive if, and only if, for all $a, b \in X$, if $a\ R\ b$, then for all $x \in X$, either $a\ R\ x$ or $x\ R\ b$.*

Proof. Suppose that R is negatively transitive. We will show that for all $a, b \in X$, if $a\ R\ b$, then for all $x \in X$, either $a\ R\ x$ or $x\ R\ b$. Let $a, b \in X$ and suppose that $a\ R\ b$. Let $x \in X$ and suppose that $a \not R\, x$ and $x \not R\, b$. Then, by the negative transitive property, $a \not R\, b$, contradicting our assumption. Hence, $a\ R\ x$ or $x\ R\ b$. Suppose that for all $a, b \in X$, if $a\ R\ b$, then for all $x \in X$, either $a\ R\ x$ or $x\ R\ b$. We must show that R is negatively transitive. Suppose that $a \not R\, b$ and $b \not R\, c$. Suppose that, contrary to the negative transitive property, $a\ R\ c$. Then since, $b \in X$, we have either $a\ R\ b$ or $b\ R\ c$, which contradicts our assumption. Thus, $a \not R\, c$, as desired. QED

Theorem 1.12 *Suppose that R' is negatively transitive and irreflexive. Then, define R by aRb iff it is not the case that $bR'a$ and aIb iff neither $aR'b$ nor $bR'a$. Then, R is complete and transitive. Furthermore, $P_R = R'$ and $I_R = I$.*

2 Choices

In this section, we will introduce some notation to describe a decision maker's choices. Suppose that X is a finite set and $\mathcal{P}(X) = \{Y \mid Y \neq \emptyset \text{ and } Y \subseteq X\}$ is the set of non-empty subsets of X. Elements of $\mathcal{P}(X)$ are called **menus**. A choice function identifies the elements form a menu (i.e., a finite set of objects) chosen by a decision maker.

Definition 2.1 (Choice Function) Suppose that X is a finite set. A **choice function** on X is a function $c : \mathcal{P}(X) \to \mathcal{P}(X)$ such that for all $S \in \mathcal{P}(X)$, $c(S) \subseteq S$. $\triangleleft$

Remark 2.2 (Actual vs. Hypothetical Choice) *The mathematical formalism does not specify whether a choice function c represents a decision maker's* actual *or* hypothetical *choices. If it is the actual choices, then c is a record of the decision maker's observed choice behavior. If it is the hypothetical choices, then c represents what the decision maker* would *chose if given the opportunity to select an element from a given menu.*

Typically, it is assumed that the decision maker's choices are guided by some underlying (subjective) preference relation (together with the decision maker's beliefs).

Definition 2.3 (Derived Choice Function) Let R be a relation on a finite set X. The choice function derived from the relation R is $c_R : \mathcal{P}(X) \to \mathcal{P}(X)$ is defined as follows: for all $S \in \mathcal{P}(X)$,

$$c_R(S) = \{y \mid y \in S \text{ and there is no } x \in S \text{ such that } x \; R \; y\}$$

That is, $c_R(S) = \max_R(S)$. $\triangleleft$

The above definition works for *any* relation on a set X. In general, given an arbitrary relation R on X, c_R many not necessarily be a *choice* function. This would happen when there is a finite subset Y such that $\max_R(Y) = \emptyset$. The following Lemma states precisely when a function derived from a relation is a choice function.

Lemma 2.4 *Suppose that X is finite. A binary relation $R \subseteq X \times X$ is acyclic iff c_R is a choice function.*

Proof. Suppose that $R \subseteq X \times X$ is acyclic. By definition, for any nonempty set $S \in \mathcal{P}(X)$, $c_R(S) \subseteq S$. We must show $c_R(S) \neq \emptyset$. Suppose that $c_R(S) = \emptyset$. Choose an element $x_0 \in S$. Since $c_R(S) = \emptyset$, there is an element $x_1 \in S$ such that $x_1 \; R \; x$. Again, since $c_R(S) = \emptyset$ there must be some element $x_2 \in S$ such that $x_2 \; R \; x_1$. Since R is acyclic, we must have $x_2 \neq x$ (otherwise, $x \; R \; x_1 \; R \; x$ is a cycle). Continue in this manner selecting elements of S. Since S is finite, eventually all elements of S are selected. That is, we have $S = \{x_0, x_1, x_2, \ldots, x_n\}$ and

$$x_n \; R \; x_{n-1} \; R \cdots x_2 \; R \; x_1 \; R \; x_0$$

Since $c_R(S) = \emptyset$ there must be some element $x \in S$ such that $x \; R \; x_n$. Thus, $x = x_i$ for some $i = 0, \ldots, n$, which implies R has a cycle. This contradicts the assumption that $c_R(S) = \emptyset$. Hence $c_R(S) \neq \emptyset$.

Suppose that c_R is a choice function. This means that for all $S \in \mathcal{P}(X)$, $c_R(S) \neq \emptyset$. Suppose that R is not acyclic. Then, there is a set of distinct elements $x_1, x_2, \ldots, x_n \in S$ such that

$$x_1 \; R \; x_2 \; R \cdots x_{n-1} \; R \; x_n \; R \; x_1.$$

But this means that $c_R(\{x_1, \ldots, x_n\}) = \emptyset$. (The above cycle means that there is no maximal element of $\{x_1, \ldots, x_n\}$.) This contradicts the assumption that c_R is a choice function. Thus, R is acyclic. QED

An immediate corollary is that if $\succeq$ is a rational preference ordering on X, then $c_{\succeq}$ is a choice function.

Definition 2.5 (Rationalizable Choice Functions) A choice function $c : \mathcal{P}(X) \to \mathcal{P}(X)$ is **rationalizable** if there is a rational preference relation $\succeq$ such that $c = c_{\succ}$. $\triangleleft$

Exercise 8 Give an example of a choice function that is not rationalizable.

The question we are interested in is *when is a choice function rationalizable*? The following properties of the choice function provide an answer to this question. Suppose that $c : \mathcal{P}(X) \to \mathcal{P}(X)$ is a choice function. We say c satisfies:

- **Houthakker's Axiom (WARP)** provided for all $x, y \in X$, if x and y are in both A and B and if $x \in c(A)$ and $y \in c(B)$ then $x \in c(B)$.
- **Sen's property** α provided for all $x \in X$ if $x \in B \subseteq A$ and $x \in c(A)$, then $x \in c(B)$
- **Sen's property** β provide for all $x, y \in X$ if $x, y \in c(A)$, $A \subseteq B$ and $y \in c(B)$, then $x \in c(B)$.

Exercise 9 1. Give an example of a choice function satisfying Sen's α but not Sen's β.

2. Give an example of a choice function satisfying Sen's β but not Sen's α.
3. Give and example of a choice function that does not satisfy Sen's α and does not satisfy Sen's β.
4. Give and example of a choice function satisfying both Sen's α and β.

The first result is that Sen's α and β are together equivalent to WARP.

Lemma 2.6 *A choice function c satisfies WARP iff c satisfies Sen's properties α and β.*

Proof. Suppose that c satisfies WARP. We must show c satisfies Sen's α and β:

Sen's α: Suppose that $x \in X$ with $x \in B \subseteq A$ and $x \in c(A)$. Suppose that $x \notin c(B)$. Then there is some $y \in B$ such that $y \in c(B)$ and $y \neq x$. Since $y \in B$ and $B \subseteq A$, we have $y \in A$. Hence, x and y are in both A and B. By the WARP axiom, since $x \in c(A)$ and $y \in c(B)$, we must have $x \in c(A)$. This contradicts the assumption that $x \notin c(B)$. Thus, $x \in c(B)$.

Sen's β: Suppose that $x, y \in X$ with $x, y \in c(A)$, $A \subseteq B$ and $y \in c(B)$. Since $c(A) \subseteq A$, we have $x, y \in A$; and since $A \subseteq B$, we have $x, y \in B$. Thus, x and y are in both A and B. By the WARP axiom, since $x \in c(A)$ and $y \in c(B)$, we must have $x \in c(B)$, as desired.

Suppose that c satisfies Sen's α and β. We must show that c satisfies WARP. Suppose that $x, y \in A \cap B$, $x \in c(A)$ and $y \in c(B)$. We must show that $x \in c(B)$. Since, $A \cap B \subseteq B$ and $y \in c(B)$, by Sen's α, $y \in c(A \cap B)$. Similarly, since $A \cap B \subseteq A$ and $x \in c(A)$, by Sen's α, $x \in c(A \cap B)$. Finally, Since $x, y \in c(A \cap B)$, $A \cap B \subseteq B$ and $y \in c(B)$, by Sen's β, we have $x \in c(B)$. QED

The main result of this section is the revelation theorem showing that WARP is equivalent to rationalizability.

Theorem 2.7 (Revelation Theorem) *c satisfies WARP iff c is rationalizable.*

Proof. Suppose that c is rationalizable. Then there is a rational preference ordering $\succeq$ such that $c = c_{\succ}$. We must show that c satisfies WARP. Suppose that $x, y \in A \cap B$, $x \in c(A)$ and $y \in c(B)$. We must show that $x \in c(B)$. Since $c = c_{\succ}$, we have $x \in \max_{\succ}(A)$ and $y \in \max_{\succ}(B)$. This means that there is no $z \in A$ such that $z \succ x$ and there is no $z \in B$ such that $z \succ y$. Suppose that $w \in B$. We will show that $w \not\succ x$. Since $w \in B$ and $y \in \max_{\succ}(B)$, we have $w \not\succ y$. Since $\succeq$ is complete, this means that $y \succeq w$. Furthermore, since $y \in A$ and $x \in \max_{\succ}(A)$, we have $y \not\succ x$. Since $\succeq$ is complete, this means that $x \succeq y$. Since $\succeq$ is transitive and $x \succeq y$ and $y \succeq w$, we have $x \succeq w$. This implies that $w \not\succ x$. That is, $x \in \max_{\succ}(B) = c(B)$.

Suppose that c satisfies WARP. Then by Lemma 2.6, c satisfies Sen's α and β. Define a relation $\succeq_c \subseteq X \times X$ as follows: for all $x, y \in X$,

$$x \succeq_c y \text{ iff } x \in c(\{x, y\})$$

We must show that 1. $\succeq_c$ is a preference relation and 2. for all $S \in \mathcal{P}(X)$, $c(S) = c_{\succ_c}(S)$. To see that 1. holds:

$\succeq_c$ is complete: For any $x, y \in X$, since $c(\{x, y\})$ is non-empty we have $c(\{x, y\}) = \{x\}$, $c(\{x, y\}) = \{y\}$ or $c(\{x, y\}) = \{x, y\}$. Thus, either $x \succeq_c y$ or $y \succeq_c x$ (or both).

$\succeq_c$ is transitive: Suppose that $x \succeq_c y$ and $y \succeq_c z$. Then, $x \in c(\{x, y\})$ and $y \in c(\{y, z\})$. We must show that $x \succeq_c z$; that is, $x \in c(\{x, z\})$. By Sen's α, if $x \in c(\{x, y, z\})$, then $x \in c(\{x, z\})$. Thus, if we show that $x \in c(\{x, y, z\})$, then we are done. There are three cases to consider. Suppose that $c(\{x, y, z\}) = \{y\}$. By Sen's α, since $\{x, y\} \subseteq \{x, y, z\}$ and $y \in c(\{x, y, z\})$ we must have $y \in c(\{x, y\})$. Thus, $c(\{x, y\}) = \{x, y\}$. By Sen's β, this implies that $x \in c(\{x, y, z\})$ (this follows since $\{x, y\} \subseteq \{x, y, z\}$, $x, y \in c(\{x, y\})$ and $y \in c(\{x, y, z\}))$. This contradicts the assumption that $c(\{x, y, z\}) = \{y\}$. Thus, $c(\{x, y, z\} \neq \{y\}$. A similar argument shows that $c(\{x, y, z\}) \neq \{z\}$. Suppose that $c(\{x, y, z\}) = \{y, z\}$. Then, $y \in c(\{x, y, z\})$, and, as above, by Sen's α, we have $c(\{x, y\}) = \{x, y\}$. This implies, by Sen's β, that $x \in c(\{x, y, z\})$, which contradicts that assumption that $c(\{x, y, z\}) = \{y, z\}$.

Hence, $x \in c(\{x, y, z\}$. By Sen's α, since $\{x, z\} \subseteq \{x, y, z\}$, we have $x \in c(\{x, z\})$. That is, $x \succeq_c z$. This completes the proof that $\succeq_c$ is transitive.

Suppose that $S \in \mathcal{P}(X)$. First of all, if S is a singleton (i.e., $S = \{x\}$ for some $x \in X$), then by definition $c(S) = S = c_{\succ_c}(S)$. Thus, in what follows we assume that S has at least two elements. We must show that $c(S) = c_{\succ_c}(S)$. We first show that $c(S) \subseteq c_{\succ_c}(S)$. Suppose that $x \in c(S)$. We must show that $x \in c_{\succ_c}(S)$. Let $y \in S$. We must show that $y \not\succ_c x$. Since $\succ_c$ is complete, this is equivalent to showing that $x \succeq_c y$. Since $\{x, y\} \subseteq S$ and $x \in c(S)$, by Sen's α, we have $x \in c(\{x, y\})$. Thus, $x \succeq_c y$; and so, $y \not\succ_c x$, which implies that $x \in \max_{\succ_c}(S) = c_{\succ_c}(S)$. Next, we show that $c_{\succ_c}(S) \subseteq c(S)$. Suppose that $x \in c_{\succ_c}(S)$. If $x \notin c(S)$. Then there is some $y \neq x$ such that $y \in c(S)$. By Sen's α, this implies that $y \in c(\{x, y\})$. Furthermore, if $c(\{x, y\}) = \{x, y\}$, then, by Sen's β, $x \in c(S)$. This contradicts the assumption that $x \notin c(S)$. Thus, $c(\{x, y\}) = \{y\}$. By definition, this means that $y \succeq_c x$ but $x \not\succeq_c y$; i.e., $y \succ_c x$. QED

We end this section by pointing out that there are alternative ways to define a preference ordering from a choice function.

Fact 2.8 *Suppose that X is a finite set and $c : \mathcal{P}(X) \to \mathcal{P}(X)$ is a choice function. Consider the following two alternative ways to define a preference ordering from the choice function c:*

- *For all $x, y \in X$, $x \succeq_c^1 y$ iff there is a set $S \in \mathcal{P}(X)$ such that $x, y \in S$ and $x \in c(S)$.*
- *For all $x, y \in X$, set $x \succ_c^2 y$ iff $x = c(\{x, y\})$ and $x \sim_c^2 y$ iff $c(\{x, y\}) = \{x, y\}$. Then, define $x \succeq_c^2 y$ iff $x \succ_c^2 y$ or $x \sim_c^2 y$.*

If c satisfies WARP, then for all $x, y \in X$, the following are equivalent:

- $x \succeq_c y$
- $x \succeq_c^1 y$
- $x \succeq_c^2 y$

The proof of this follows from the definitions and is left to the reader.

3 Utility

A **utility function** on X is a function $u : X \to \Re$, where $\Re$ is the set of real numbers.

Definition 3.1 (Representing a Preference Ordering) Suppose that $\succeq \subseteq X \times X$ is a preference ordering. We say that $\succeq$ is **representable by a utility function** provided there is a $u : X \to \Re$ such that for all $x, y \in X$, $x \succeq y$ iff $u(x) \geq u(y)$. $\triangleleft$

Lemma 3.2 *Suppose that X is a finite set. A relation $R \subseteq X \times X$ is a rational preference ordering iff R is representable by a utility function.*

Proof. We leave it to the reader to show that if R is representable by a utility function, then R is transitive and complete.

We prove the following: For all $n \in \mathbb{N}$, any preference relation $\succeq$ on a set of size n is representable by a utility function $u_{\succeq} : X \to \Re$. The proof is by induction on the size of the set of objects X. The base case is when $|X| = 1$. In this case, $X = \{a\}$ for some object a. If $\succeq$ is a transitive and complete ordering on X, then $\succeq = \{(a,a)\}$. Then, $u_{\succeq}(a) = 0$ (any real number would work here) clearly represents $\succeq$. The induction hypothesis is: if $|X| = n$, then any preference ordering on X is representable. Suppose that $|X| = n+1$ and $\succeq$ is a preference ordering on X. Then, $X = X' \cup \{a\}$ for some object a, where $|X'| = n$. Note that the restriction[1] of $\succeq$ to X', denoted $\succeq'$, is a preference ordering on X'. By the induction hypothesis, $\succeq'$ is representable by a utility function $u_{\succeq'} : X' \to \Re$. We will show how to extend $u_{\succeq'}$ to a utility function $u_{\succeq} : X \to \Re$ that represents $\succeq$. For all $b \in X'$, let $u_{\succeq}(b) = u_{\succeq'}(b)$. For the object a (the unique object in X but not in X'), there are four cases:

1. $a \succ b$ for all $b \in X'$. Let $u_{\succeq}(a) = \max\{u_{\succeq'}(b) \mid b \in X'\} + 1$.
2. $b \succ a$ for all $b \in X'$. Let $u_{\succeq}(a) = \min\{u_{\succeq'}(b) \mid b \in X'\} - 1$.
3. $a \sim b$ for some $b \in X'$. Let $u_{\succeq}(a) = u_{\succeq'}(b)$.
4. There are $b_1, b_2 \in X'$ such that $b_{\succ} a \succ b_2$. Let $u_{\succeq}(a) = \frac{u_{\succeq'}(b_1) + u_{\succeq'}(b_2)}{2}$.

Then, it is straightforward to show that $u_{\succeq} : X \to \Re$ represents $\succeq$ (the details are left to the reader). QED

Exercise 10 Suppose that $X = \{a, b, c, d\}$ and $\succeq$ is the following preference relation:

$$b \succ a \sim c \succ d$$

Find two utility functions that represent this preference relation.

Remark 3.3 *The above proof can be extended to relations on infinite sets X. However, additional technical assumptions are needed. It is beyond the scope of this article to discuss these technicalities here.*

It is not hard to see that if a preference relation is representable by a utility function, then it is representable by *infinitely* many different utility functions. To make this more precise, say that a function $f : \Re \to \Re$ is **monotone** provided $r \geq r'$ implies $f(r) \geq f(r')$.

Lemma 3.4 *Suppose that $\succeq$ is representable by a utility function $u_{\succeq}$ and $f : \Re \to \Re$ is a monotone function. Then, $f \circ u_{\succeq}$ also represents $\succeq$.*

Proof. The proof is immediate from the definitions. Suppose that $a, b \in X$. Then, $a \succeq b$ iff $u_{\succeq}(a) \leq u_{\succeq}(b)$ (since u represents $\succeq$) iff $f(u_{\succeq}(a)) \geq f(u_{\succeq}(b))$ (since f is monotone). QED

[1]If R is a relation on X and $Y \subseteq X$, then $R_Y \subseteq Y \times Y$ is the **restriction of R to Y** provided for all $a, b \in Y$, $a \, R_Y \, b$ iff $a \, R \, b$.

The main result of this section is the von Neumann-Morgenstern Theorem. This results characterizes when a preference relation on *lotteries* is representable. Suppose that X is a finite set. A probability function on X is a function $p : X \to [0, 1]$ such that $\sum_{x \in X} p(x) = 1$. If $S \subseteq X$, then $p(S) = \sum_{x \in S} p(x)$.[2] In the remainder of this section, elements of X are called *prizes*.

Definition 3.5 (Lottery) Suppose that $Y = \{x_1, \ldots, x_n\}$ is a set of n elements from X. A lottery on Y is denoted

$$[x_1 : p_1, x_2 : p_2, \ldots, x_n : p_n]$$

where $\sum_{i=1}^{n} p_i = 1$. ◁

Remark 3.6 *We have defined lotteries for any subset of a fixed set X. Without loss of generality, we can restrict attention to all lotteries on X. For instance, suppose that $X = \{x_1, \ldots, x_n, y_1, \ldots, y_m\}$ and L is a lottery on $Y = \{x_1, \ldots, x_n\}$. That is, $L = [x_1 : p_1, \ldots, x_n : p_n]$. This lottery can be trivially extended to a lottery L' over X as follows:*

$$L' = [x_1 : p_1, \ldots, x_n : p_n, y_1 : 0, \ldots, y_m : 0].$$

Suppose that X is a finite set and let $\mathcal{L}$ be the set of lotteries on X. There are two technical issues that need to be addressed. First of all, we can identify elements $x \in X$ with lotteries $[x : 1]$. Thus, we may abuse notation and say that "X is contained in $\mathcal{L}$". Second, we will need the notion of a **compound lottery**.

Definition 3.7 (Compound Lottery) Suppose that $L_1, \ldots, L_n$ are lotteries. Then, $[L_1 : p_1, \ldots, L_n : p_n]$ is **compound lottery**, where $\sum_{i=1}^{n} p_i = 1$. ◁

We are interested in decision makers that have preferences over the set $\mathcal{L}$ of lotteries. Suppose that $\succeq$ is a preference relation on $\mathcal{L}$. The goal is to show that any preference relation can be represented by a linear utility function:

Definition 3.8 (Linear Utility Function) A utility function $u : \mathcal{L} \to \Re$ is **linear** provided for all $L = [L_1 : p_1, \ldots, L_n : p_n] \in \mathcal{L}$,

$$u(L) = \sum_{i=1}^{n} p_i u(L_i)$$

If $u : \mathcal{L} \to \Re$ is a linear utility function, then for $L \in \mathcal{L}$, $u(L)$ is called the **expected utility of** L. ◁

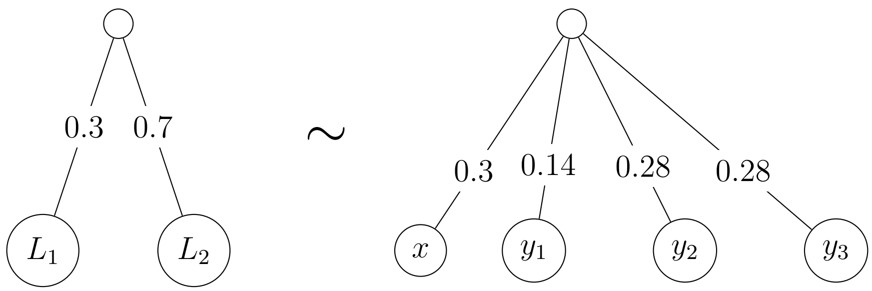
The first axiom is that compound lotteries can always be "reduced" to simple lotteries.

[2] There are a number of mathematical details about probability measures that we are glossing over here. Our discussion in this section is greatly simplified since we assume that the set of objects X is finite.

Simplifying lotteries Suppose that $[L_1 : p_1, \ldots, L_n, p_n]$ is a compound lottery, where for each $i = 1, \ldots, n$, we have $L_i = [x_1 : p_1^i, \ldots, x_n : p_n^i]$. Then,

$$[L_1 : p_1, \ldots, L_n, p_n] \sim [x_1 : (p_1 p_1^1 + p_2 p_1^2 + \cdots p_n p_1^n), \ldots, x_1 : (p_1 p_n^1 + p_2 p_n^2 + \cdots p_n p_n^n)]$$

The axiom means that decision makers do not get any utility from the "thrill of gambling". That is, what matters to the decision maker is how likely she is to receive prizes that she prefers. For example, suppose that $L_1 = [x : 1]$ and $L_2 = [y_1 : 0.2, y_2 : 0.4, y_3 : 0.4]$. Then, the decision maker is assumed to be indifferent between the compound lottery $L_3 = [L_1 : 0.3, L_2 : 0.7]$ and the simple lottery $L_4 = [x : 0.3, y_1 : 0.14, y_2 : 0.28, y_3 : 0.28]$. This can be pictured as follows:



The next two axioms play a central role in the von Neumann-Morgenstern theorem.

Independence For all $L_1, L_2, L_3 \in \mathcal{L}$ and $a \in (0, 1]$,

$$L_1 \succ L_2 \text{ if, and only if, } [L_1 : a, L_3 : (1-a)] \succ [L_2 : a, L_3 : (1-a)].$$

Continuity For all $L_1, L_2, L_3 \in \mathcal{L}$ and $a \in (0, 1]$,

$$\text{if } L_1 \succ L_2 \succ L_3, \text{ then there exists } a, b \in (0,1) \text{ such that}$$
$$[L_1 : a, L_3 : (1-a)] \succ L_2 \succ [L_1 : b, L_3 : (1-b)].$$

Both axioms have been criticized as *rationality* principles. For example, consider the continuity axiom. Consider the prizes $x =$ "win \$1000", $y =$ "win \$100" and $z =$"get hit by a car". Clearly, it is natural to assume that a decision make would have the preference $x \succ y \succ z$. Now, Continuity implies that there is some number $a \in (0, 1)$ such that $[x : a, z : (1-a)] \succ [y : 1]$. Thus, the decision maker would strictly prefer a lottery in which there is some non-zero chance of getting hit by a car to a lottery in which gives a guaranteed payoff of \$100. Arguably, many people would not hold such a preference no matter how small the chance is of getting hit by a car. Here we bracket this are related philosophical discussions about the above axioms and focus on what follows from the axioms. The first observation is a straightforward consequence of Independence and the assumption that the preference ordering is complete.

Observation 3.9 *Suppose that $\succeq$ is a preference relation on $\mathcal{L}$ satisfying the Independence axiom. For all lotteries $L_1, L_2, L_3 \in \mathcal{L}$ and real numbers $a \in [0,1]$, if $L_1 \sim L_2$, then $[L_1 : a, L_3 : (1-a)] \sim [L_2 : a, L_3 : (1-a)]$.*

The second observation is that decision makers prefers lotteries in which there is a better chance of winning a more preferred prize.

Lemma 3.10 *If $\succeq$ is a preference relation on $\mathcal{L}$ satisfying Simplifying lotteries and Independence, then for all lotteries $L_1, L_2 \in \mathcal{L}$, if $L_1 \succ L_2$, and $1 \geq a > b \geq 0$, then $[L_1 : a, L_2 : (1-a)] \succ [L_1 : b, L_2 : (1-b)]$.*

Proof. Suppose that $b = 0$. Then, $1 \geq a > 0$. Apply the Independence axiom with $L_3 = L_2$. We have $[L_1 : a, L_2 : (1-a)] \succ [L_2 : a, L_2 : (1-a)] = L_2$. Since $b = 0$, we have $L_2 = [L_1 : b, L_2 : (1-b)]$ (this follows from the simplifying lotteries axiom). Thus, $[L_1 : a, L_2 : (1-a)] \succ [L_1 : b, (1-b) : L_2]$, as desired.

Suppose that $1 \geq a > b > 0$. Then, $\frac{b}{a} < 1$ and $1 - \frac{b}{a} < 1$. Let $L' = [L_1 : a, L_2 : (1-a)]$. By the previous argument, we have $L' \succ L_2$. Since $(1 - \frac{b}{a}) \in (0,1]$ and $L' \succ L_2$, using the Independence axiom (let $L_3 = L'$), we have

$$\left[L' : (1 - \frac{b}{a}), L' : \frac{b}{a}\right] \succ \left[L_2 : (1 - \frac{b}{a}), L' : \frac{b}{a}\right]$$

Now, the left-hand side simplifies to the following lottery: $[L' : (1 - \frac{b}{a}), L' : \frac{b}{a}] = L' = [L_1 : a, L_2 : (1-a)]$. The right-hand side of the above equation simplifies as follows:

$$\begin{array}{rcl} [L_2 : (1-\frac{b}{a}), L' : \frac{b}{a}] & = & [L_2 : (1-\frac{b}{a}), [L_1 : a, L_2 : (1-a)] : \frac{b}{a}] \\ & = & [L_1 : a \times \frac{b}{a}, L_2 : ((1-a) \times \frac{b}{a} + (1 - \frac{b}{a}))] \\ & = & [L_1 : b, L_2 : (\frac{b}{a} - b + (1 - \frac{b}{a}))] \\ & = & [L_1 : b, L_2 : (1-b)] \end{array}$$

Thus, $[L_1 : a, L_2 : (1-a)] \succ [L_1 : b, L_2 : (1-b)]$, as desired. QED

Exercise 11 Prove that if a relation $\succeq$ on $\mathcal{L}$ is representable by a linear utility function, then $\succeq$ satisfies Simplifying lotteries, Independence, and Continuity.

The von Neumann-Morgentern Theorem proves the converse of the above exercise.

Theorem 3.11 (von Neumann-Morgenstern Representation Theorem) *A binary relation $\succeq$ on $\mathcal{L}$ satisfies Simplifying lotteries, Independence and Continuity iff $\succeq$ is representable by a linear utility function $u : \mathcal{L} \to \Re$.*

Moreover, $u' : \mathcal{L} \to \Re$ represents $\succeq$ iff there exists real numbers $c > 0$ and d such that $u'(\cdot) = cu(\cdot) + d$. ("u is unique up to linear transformations.")

3.1 "Counterexamples" to Expected Utility Theory

In this section, we present two purported "counterexamples" to expected utility theory. We start with an example that illustrates that utility should not necessarily be identified with monetary payoff. Suppose that there are three prizes: M_1 ="win \$1,000,000", M_0 = "win \$0" and M_3 = "win \$3,000,000". Consider the following two lotteries:

- $L_1 = [M_1 : 0.5, M_1 : 0.5] = [M_1 : 1]$, and
- $L_2 = [M_3 : 0.5, M_0 : 0.5]$

Which of the above two lotteries do you prefer? We can represent the payoff matrix for this decision as follows:

Options	1/2	1/2
L_1	M_1	M_1
L_2	M_3	M_0

The expected monetary payout for the two lotteries is:

$$\begin{aligned} \mathsf{EM}(L_1) &= 1/2 \cdot 1,000,000 + 1/2 \cdot 1,000,000 = 1,000,000 \\ \mathsf{EM}(L_2) &= 1/2 \cdot 3,000,000 + 1/2 \cdot 0 \qquad\qquad\quad = 1,500,000 \end{aligned}$$

Thus, if monetary payout is identified with utility (and the decision maker maximizes expected utility), then any rational decision maker should prefer L_2 strictly over L_1. Arguably, we should not classify a decision maker that prefers L_1 over L_2 as *irrational*. However, this is not a problem for expected utility theory. It is consistent with the von Neumann-Morgenstern Theorem that a decision maker assigns an expected utility of, say, 2 to L_1 since it is a guaranteed payoff of \$1,000,000 while simultaneously assigning an expected utility of 1.5 to L_2 (since there is a chance of not winning any money). The next two examples are more serious challenges to expected utility theory.

3.1.1 Allais Paradox

Suppose that there are three prizes: M_1 ="win \$1,000,000", M_0 = "win \$0" and M_5 = "win \$5,000,000". Consider the following 4 lotteries:

- $L_1 = [M_1 : 0.01, M_1 : 0.89, M_1 : 0.10]$
- $L_2 = [M_0 : 0.01, M_1 : 0.89, M_5 : 0.10]$
- $L_3 = [M_1 : 0.01, M_0 : 0.89, M_1 : 0.10]$
- $L_4 = [M_0 : 0.01, M_0 : 0.89, M_5 : 0.10]$

To help visualize these four lotteries, suppose that there is an urn with 100 balls. Each of the balls are either red, white or blue. Further, suppose that there is 1 red ball, 89 white balls and 10 blue balls. Then, lottery L_3 can be executed as follows: a single ball is drawn from the urn and the decision maker is paid $1,000,000 if it is red or blue. Similarly for the other lotteries. These 4 lotteries are listed in the table below:

Lotteries	Red (1)	White (89)	Blue (10)
L_1	M_1	M_1	M_1
L_2	M_0	M_1	M_5
L_3	M_1	M_0	M_1
L_4	M_0	M_0	M_5

Exercise 12 How would you rank lotteries L_1 and L_2? What about the lotteries L_3 and L_4?

Many people report that they prefer L_1 over L_2 while simultaneously preferring L_4 over L_3. The question is is there anything "wrong" with this preference ordering?

Observation 3.12 *The preference ordering $L_1 \succ L_2$ and $L_4 \succ L_3$ is inconsistent with the von Neumann-Morgenstern Theorem.*

Proof. Suppose that $\succeq$ satisfies the assumption of the von Neumann-Morgenstern Theorem and that $L_1 \succ L_2$ and $L_4 \succ L_3$. By the von Neumann-Morgenstern Theorem, there is a linear utility function u that represents $\succeq$. This means that $u(L_1) > u(L_2)$ and $u(L_4) > u(L_3)$. Since u is a linear utility function we have:

$$u(L_1) = 0.01u(M_1) + 0.89u(M_1) + 0.1u(M_1) > u(L_2) = 0.01u(M_0) + 0.89u(M_1) + 0.1u(M_5)$$

Simplifying, gives us

$$(*) \qquad 0.11u(M_1) > 0.01u(M_0) + 0.1u(M_5)$$

We also have:

$$u(L_4) = 0.01u(M_0) + 0.89u(M_0) + 0.1u(M_5) > u(L_3) = 0.01u(M_1) + 0.89u(M_0) + 0.1u(M_1)$$

Simplifying the above inequality, gives us

$$(**) \qquad 0.01u(M_0) + 0.1u(M_5) > 0.01u(M_1) + 0.1u(M_1) = 0.11u(M_1)$$

Inequalities $(*)$ and $(**)$ are contradictory. Thus, the preferences $L_1 \succ L_2$ and $L_3 \succ L_4$ are inconsistent with the von Neumann-Morgenstern Theorem. QED

Observation 3.13 *If $\succeq$ is a preference relation that satisfies the Independence axiom, then if $L_1 \succ L_2$, then $L_3 \succ L_4$.*

Proof. We first show that if $L_1 = [M_1 : 0.01, M_1 : 0.89, M_1 : 0.10] \succ L_2 = [M_0 : 0.01, M_1 : 0.89, M_5 : 0.10]$, then $[M_1 : \frac{1}{11}, M_1 : \frac{10}{11}] \succ [M_0 : \frac{1}{11}, M_5 : \frac{10}{11}]$. Suppose not. Then, $[M_1 : 0.01, M_1 : 0.89, M_1 : 0.10] \succ [M_0 : 0.01, M_1 : 0.89, M_5 : 0.10]$ and $[M_1 : \frac{1}{11}, M_1 : \frac{10}{11}] \not\succ [M_0 : \frac{1}{11}, M_5 : \frac{10}{11}]$. By completeness of $\succeq$, the second conjunct implies that $[M_0 : \frac{1}{11}, M_5 : \frac{10}{11}] \succeq [M_1 : \frac{1}{11}, M_1 : \frac{10}{11}]$. There are two cases

- Case 1. $[M_0 : \frac{1}{11}, M_5 : \frac{10}{11}] \succ [M_1 : \frac{1}{11}, M_1 : \frac{10}{11}]$. By the Independence axiom (with $L_3 = [M_1 : 1]$), we have

  $$\left[\left[M_0 : \frac{1}{11}, M_5 : \frac{10}{11}\right] : \frac{11}{100}, [M_1 : 1] : \frac{89}{100}\right] \succ \left[\left[M_1 : \frac{1}{11}, M_1 : \frac{10}{11}\right] : \frac{11}{100}, [M_1 : 1] : \frac{89}{100}\right]$$

  Simplifying the above compound lotteries gives:

  $$[M_0 : 0.01, M_1 : 0.89, M_5 : 0.1] \succ [M_1 : 0.01, M_1 : 0.89, M_1 : 0.1].$$

  This contradicts the assumption that $L_1 \succ L_2$.

- Case 2. $[M_0 : \frac{1}{11}, M_5 : \frac{10}{11}] \sim [M_1 : \frac{1}{11}, M_1 : \frac{10}{11}]$. By Observation 3.9,

  $$\left[\left[M_0 : \frac{1}{11}, M_5 : \frac{10}{11}\right] : \frac{11}{100}, [M_1 : 1] : \frac{89}{100}\right] \sim \left[\left[M_1 : \frac{1}{11}, M_1 : \frac{10}{11}\right] : \frac{11}{100}, [M_1 : 1] : \frac{89}{100}\right]$$

  Simplifying the above compound lottery gives:

  $$[M_0 : 0.01, M_1 : 0.89, M_5 : 0.1] \sim [M_1 : 0.01, M_1 : 0.89, M_1 : 0.1].$$

  This contradicts the assumption that $L_1 \succ L_2$.

Second, we show that $L_1 \succ L_2$ implies that $L_3 \succ L_4$. Suppose that $L_1 \succ L_2$. Then, by the above argument, $[M_1 : \frac{1}{11}, M_1 : \frac{10}{11}] \succ [M_0 : \frac{1}{11}, M_5 : \frac{10}{11}]$. By independence, we have

$$\left[\left[M_1 : \frac{1}{11}, M_1 : \frac{10}{11}\right] : \frac{11}{100}, [M_0 : 1] : \frac{89}{100}\right] \succ \left[\left[M_0 : \frac{1}{11}, M_5 : \frac{10}{11}\right] : \frac{11}{100}, [M_0 : 1] : \frac{89}{100}\right]$$

Simplifying the above compound lottery gives:

$$L_3 = [M_1 : 0.01, M_0 : 0.89, M_1 : 10] \succ [M_0 : 0.01, M_0 : 0.89, M_5 : 0.10] = L_4.$$

QED

Exercise 13 Prove the converse of the above observation: if $L_3 \succ L_4$, then $L_1 \succ L_2$.

The Allais paradox has generated an extensive literature in Economics and Philosophy. The discussions in the literature fall into three general categories:

1. The axioms of von Neumann and Morgenstern fail to adequately capture our understanding of rational choice.
2. Decision makers that rank L_1 above L_2 *and* L_4 above L_3 are *irrational*.
3. The model of rational choice must represent a decision maker's attitudes towards *risk*.

3.1.2 Ellsberg Paradox

Suppose that there is an urn with 100 balls. Each ball is either blue, yellow or green. Suppose that there are 30 blue balls and the remaining balls are either yellow or green, though the precise distribution of yellow and green balls is unknown. Consider the following four lotteries:

L_1 : win $\$1,000$ if a blue ball is drawn

L_2 : win $\$1,000$ if a yellow ball is drawn.

L_3 : win $\$1,000$ if a blue or green ball is drawn.

L_4 : win $\$1,000$ if a yellow or green ball is drawn.

Exercise 14 How would you rank lotteries L_1 and L_2? What about the lotteries L_3 and L_4?

Many people report a preference of L_1 over L_2 while simultaneously preferring L_4 over L_3. To simplify the discussion, the lotteries are pictured in the following table:

	30	60	
Lotteries	Blue	Yellow	Green
L_1	M_1	M_0	M_0
L_2	M_0	M_1	M_0
L_3	M_1	M_0	M_1
L_4	M_0	M_1	M_1

Observation 3.14 *The preference ordering $L_1 \succ L_2$ and $L_4 \succ L_3$ is inconsistent with the von Neumann-Morgenstern Theorem.*

Proof. Suppose that $\succeq$ satisfies the assumption of the von Neumann-Morgenstern Theorem and that $L_1 \succ L_2$ and $L_4 \succ L_3$. By the von Neumann-Morgenstern Theorem, there is a linear utility function u that represents $\succeq$. Let Y denote the number of yellow balls in the urn and G denote the number of green balls in the urn (so the possible values of Y and G range from 0 to 60 under the constraint that $Y+G = 60$). Then, the probability of choosing a yellow ball is $Y/90$ and of choosing a green ball is $G/90$. We will show that $u(L_1)-u(L_2) = u(L_3)-u(L_4)$.

$$\begin{aligned}
u(L_1) - u(L_2) &= (\tfrac{60}{90}u(M_0) + \tfrac{30}{90}u(M_1)) - (\tfrac{30+G}{90}u(M_0) + \tfrac{Y}{90}u(M_1)) \\
&= \tfrac{60}{90}u(M_0) - \tfrac{30+G}{90}u(M_0) + \tfrac{30}{90}u(M_1) - \tfrac{Y}{90}u(M_1) \\
&= \tfrac{30-G}{90}u(M_0) + \tfrac{30-Y}{90}u(M_1)
\end{aligned}$$

$$
\begin{aligned}
u(L_3) - u(L_4) &= \left(\tfrac{Y}{90}u(M_0) + \tfrac{30+G}{90}u(M_1)\right) - \left(\tfrac{30}{90}u(M_0) + \tfrac{60}{90}u(M_1)\right) \\
&= \tfrac{Y}{90}u(M_0) - \tfrac{30}{90}u(M_0) + \tfrac{30+G}{90}u(M_1) - \tfrac{60}{90}u(M_1) \\
&= \tfrac{Y-30}{90}u(M_0) + \tfrac{G-30}{90}u(M_1) \\
&= \tfrac{60-G-30}{90}u(M_0) + \tfrac{60-Y-30}{90}u(M_1) \quad \text{(since } Y = 60 - G \text{ and } G = 60 - Y\text{)} \\
&= \tfrac{30-G}{90}u(M_0) + \tfrac{30-Y}{90}u(M_1)
\end{aligned}
$$

Thus, since $u(L_1) - u(L_2) = u(L_3) - u(L_4)$, we must have $u(L_1) \geq u(L_2)$ iff $u(L_3) \geq u(L_4)$, which implies that $L_1 \succeq L_2$ iff $L_3 \succeq L_4$. QED

As in the previous section, we can show that the rankings $L_1 \succ L_2$ and $L_4 \succ L_3$ is inconsistent with the Independence axiom.

Exercise 15 The four lotteries in the Ellsberg case are:

$L_1 = [M_1 : \frac{30}{90}, M_0 : \frac{Y}{90}, M_0 : \frac{G}{90}]$
$L_2 = [M_0 : \frac{30}{90}, M_1 : \frac{Y}{90}, M_0 : \frac{G}{90}]$
$L_3 = [M_1 : \frac{30}{90}, M_0 : \frac{Y}{90}, M_1 : \frac{G}{90}]$
$L_4 = [M_0 : \frac{30}{90}, M_1 : \frac{Y}{90}, M_1 : \frac{G}{90}]$

Prove that if $\succeq$ satisfies the Independence axiom and $L_1 \succ L_2$, then $L_3 \succ L_4$. (*Hint*: the proof is similar to the proof of Observation 3.9).

The Ellsberg paradox has generated a large literature on *imprecise probabilities* and *ambiguity aversion*. The details of this literature are beyond the scope of these notes.