Chapter 1

PREFERENCES AND UTILITY

1. Fundamental Assumptions

We suppose that there is a set of states, or alternatives, or bundles of goods, or "things" in the world. At various times we'll use various symbols to denote those things, but for now, we use the letters x, y, z, ... Later on we will be more explicit about the nature of our set of things.

The first fundamental assumption that we make about people is that they know that they like: they know their preferences among the set of things. If a person is given a choice between x and y, he can say (one and only one sentence is true):

1. He prefers x to y
2. He prefers y to x
3. He is indifferent between the two.

This is the axiom of *completeness*. It seems reasonable enough.

But some objections could be made to it. For a variety of reasons, a reasonable person might not be able to choose. If you are given the choice between shooting your dog and shooting your cat, you will balk. If your don't know what x and y really are; if, for example, both are complicated machines like cars and you don't know much about them, you may be unwilling to choose. If you are used to having your choices made for you; if you are dependent on your parents, your doctor, your religious guide, your government, you may be incapable of making choices yourself. Moreover, it may be painful, time consuming, distasteful, and nerve-wracking to make choices, and we will more or less ignore these costs of decision making. In spite of these objections, we make the assumption.

The second fundamental assumption is the axiom of *transitivity*. The assumption has four parts:

1. If a person prefers x to y and prefers y to z, then he prefers x to z.
2. If a person prefers x to y, and is indifferent between y and z, then he prefers x to z.
3. If a person is indifferent between x and y and prefers y to z, then he prefers x to z.
4. If a person is indifferent between x and y, and is indifferent between y and z, then he is indifferent between x and z.

There are several possible objections to the transitivity assumption. Parts (1), (2), and (3) may simply not be true for some people under some circumstances. It might be the case that you prefer apple to cherry pie, and cherry to peach pie, while you prefer peach to apple. In fact, experiments with real subjects sometimes do reveal intransitivities of this sort, although when they are brought to the subjects' attention, they typically change their minds. Part (4) is the least realistic, since it can be applied repeatedly to get nonsense results: Let x_1 be a cup of coffee with one grain of sugar in it; let x_2 be a cup of coffee with two grains of sugar in it; and so on. Now it's almost certainly the case that you can't taste the difference between x_k ad x_{k+1}, for any whole number k, and so you must be indifferent between them. Therefore, by repeated applications of (4), you must be indifferent between x_0 and $x_{1,000,000}$, which is probably false. The problem here is evidently the existence of psychological thresholds. It can be escaped by assuming those thresholds away, or by assuming away the existence of finely divisible states of the world.

It is possible for some purposes to do without parts (2)-(4) of the transitivity assumption, in which case we say preferences are *quasi-transitive*. And quasi-transitivity itself can be further weakened, by assuming:

> If a person prefers x_1 to x_2, and prefers x_2 to x_3, ..., and prefers x_{k-1} to x_k, then he does *not* prefer x_k to x_1.

If preferences satisfy this assumption we say they are *acyclic*. In most of what follows, however, we assume all of transitivity for individuals' preferences.

The third and last fundamental assumption is that people always choose an alternative which is preferred or indifferent to every alternative available to them. They choose "best" alternatives for themselves. In short, they are rational.

2. Best Alternatives and Utility Functions

In the middle and late nineteenth century it was popular in some philosophical circles to assume that pleasure and pain could be numerically measured. The measurement was in terms of *utils* or *utility units*, which were considered as scientifically real as units of length, mass, or temperature. Now a unit of length is scientifically real for several reasons: first, there is a standard object which everyone (at least everyone outside the U.S.) agrees represents one unit (e.g., a platinum rod in a vault in Paris); second, there is a natural zero for length; third, units of length can be added, subtracted, and multiplied by numbers according to the rules of arithmetic, and the results make sense: 2 meters + 2 meters = 4 meters.

Some of the nineteenth century advocates of utility calculus thought utility could be standardized and measured, like length; they thought the units could be used to measure everyone's happiness; they thought there was a natural zero between pleasure and pain; and they thought units of utility could be added and subtracted in a reasonable way.

But no one has yet succeeded in defining an objective unit of utility. Is it a level of electrical activity somewhere in the brain? Is it an index constructed from pulse, blood pressure, glandular activity data? Is it a rate of salivation, a degree of pupil dilation, or perspiration? We don't know. There is no way of comparing levels of satisfaction among different people. For that matter, there is no objective way of measuring utility at two different times for the same individual. This remains so despite the interesting developments in experimental psychology and neuroeconomics, although future research in these fields may shed important light on these issues.

But there is a subjective way: Ask him. (If you don't believe what a person says, you might choose instead to observe him. See what he chooses when he has what opportunities. If he chooses x when he might have chosen y, he reveals his preference for x.)

The problem with asking about utility is this. If you ask "How many units of happiness would you now get if I gave you a banana?" you will be laughed at. The question must be more subtly put. Ask instead, "Would you prefer a banana to an apple?" This is our fundamental question.

Asking "Would you prefer x to y" will never get you a measure of utility with well defined units, a zero, and other nice mathematical properties. But it will allow you to find alternatives that are at least as good as all others, and, remarkably, it will allow you to construct a numerical measure to reflect tastes. The determination of best alternatives and the construction of a measure of satisfaction are both made possible by the completeness and transitivity assumptions on preferences. Therefore, the theory of preferences, with those two assumptions, is connected to, and is a generalization of, the old-fashioned nineteenth century theory of utility.

3. The Formal Model of Preferences

Before we can proceed, we need to introduce some notation. Let x and y be two alternatives. We consider a group of people who are numbered 1, 2, 3, and so on. To symbolize the preferences of the i^{th} person we write xR_iy for "i thinks x is at least as good as y"; xP_iy for "i prefers x to y"; and xI_iy for "i is indifferent between x and y."

The relation R_i should be viewed as the logical primitive, the "given." The relations P_i and I_i can be derived from R_i with these definitions:

$$xP_iy \text{ if } xR_iy \text{ and not } yR_ix$$

$$xI_iy \text{ if } xR_iy \text{ and } yR_ix$$

In words: Person i prefers x to y if he thinks x is at least as good as y but he does not think y is at least as good as x. And i is indifferent between x and y if he thinks x is at least as good as y and he thinks y is at least as good as x.

Now our fundamental axioms of completeness and transitivity are formally put this way:

> *Completeness.* For any pair of alternatives x and y, either xR_iy or yR_ix.

> *Transitivity.* For any three alternatives x,y, and z, if xR_iy and yR_iz, then xR_iz.

Notice that these definitions are in terms of the primary relation R_i, rather than in terms of the derived relations P_i and I_i. The verbal definitions in the section above were in terms of P_i and I_i. The reader can check that the verbal and the formal definitions are in fact logically equivalent. That is, if R_i is transitive in the sense that, for all x, y, and z, xR_iy and yR_iz implies xR_iz, then the following must also be true:

1. xP_iy and yP_iz implies xP_iz. (See Proposition 1 below.)

2. xP_iy and yI_iz implies xP_iz.

3. xI_iy and yP_iz implies xP_iz.

4. xI_iy and yI_iz implies xI_iz.

The less fundamental (and weaker) assumptions of quasi-transitivity and acyclicity are formally put this way:

Quasi-transitivity. For any three alternatives x, y, and z, if xP_iy and yP_iz, then xP_iz.

Acyclicity. For any list of alternatives x_1, $x_2, \ldots, x_k$, if $x_1P_ix_2$, $x_2P_ix_3, \ldots$, and $x_{k-1}P_ix_k$, then not $x_kP_ix_1$.

Let us now prove that if a preference relation R_i is transitive, it must be quasi-transitive, and if it is quasi-transitive, it must be acyclic:

Proposition 1. If R_i is transitive, then it is quasi-transitive. If R_i is quasi-transitive, then it is acyclic.

Proof. Suppose first that R_i is transitive. We want to show it is quasitransitive. Suppose xP_iy and yP_iz. We need to show xP_iz, that is, xR_iz and not zR_ix. Now xP_iy means xR_iy and not yR_ix and yP_iz means yR_iz and not zR_iy. Since xR_iy and yR_iz, xR_iz follows by R_i's transitivity. If zR_ix were also true, then we would have zR_ix, xR_iy and, by R_i's transitivity, zR_iy, which contradicts not zR_iy. Consequently, zR_ix cannot be true; that is, not zR_ix. But xR_iz and not zR_ix means xP_iz, and R_i is quasi-transitive.

Next suppose R_i is quasi-transitive. We want to show it is acyclic. Suppose $x_1P_ix_2, x_2P_ix_3 \ldots, x_{k-1}P_ix_k$. We need to show not $x_kP_ix_1$. Since $x_1P_ix_2$ and $x_2P_ix_3$, $x_1P_ix_3$ by quasi-transitivity. Similarly, since $x_1P_ix_3$ and $x_3P_ix_4$, $x_1P_ix_4$ by quasi-transitivity. Repeated applications of this argument gives $x_1P_ix_k$, and not $x_kP_ix_1$ follows immediately. Q.E.D

We have already noted that preferences can be quasi-transitive without being transitive: the grains-of-sugar-in-coffee example shows this. Preferences can also be acyclic without being quasi-transitive or transitive. Suppose someone likes apples (A) better than bananas (B), and bananas better than cherries (C), but is indifferent between apples and cherries. Then his preferences relation is AP_iB, BP_iC, and AI_iC. This doesn't violate acyclicity since there is no preferences cycle. (If CP_iA holds, there is a cycle.) But the preference relation is not quasi-transitive, since quasi-transitivity would require AP_iC.

With the necessary tools in hand, we proceed to define what is meant by an individual's "best" choices. Suppose S is some collection of alternatives. Let x be an element of S. Then x is said to be best for person i if i thinks it is at least as good as every other element of S.

Formally, i's *best set in* S or i's *choice set in* S, denoted $C(R_i, S)$, is defined as follows:

$$C(R_i, S) = \{x \text{ in } S | x\, R_i y \text{ for all } y \text{ in } S\}.$$

This is read: "$C(R_i, S)$ is the set of all x's in S, such that xR_iy for all y's in S." (Note that braces $\{\quad\}$ means "the set" and a slash $|$ means "such that.")

Now to the next result. Proposition 2 answers the question "When can we be sure best things exist?" One answer is: Whenever a preference relation (defined on a finite set) is complete and transitive.

Proposition 2. Let S be a finite set of alternatives available to person i. Suppose R_i is complete and transitive. The $C(R_i, S)$ is nonempty. That is, best choices exist.

Proof. Choose one alternative, say x_1, from S. If it is best, we are done. If not, there is an alternative, say x_2, for which

$$x_1 R_i x_2$$

does not hold. By completeness $x_2 R_i x_1$ must hold, and therefore, by definition

$$x_2 P_i x_1.$$

If x_2 is best, we are done. If not, we can choose an x_3 such that

$$x_3 P_i x_2$$

by the same argument as above.

This process can either terminate at a best choice (in which case we are done), or it can go on indefinitely. Since S has only a finite number of elements, if the choice process goes on forever, it must repeat. Therefore, there must be a cycle:

$$x_1 P_i x_k P_i x_{k-1} P_i \ldots x_3 P_i x_2 P_i x_1.$$

Repeated applications of the transitivity assumption implies $x_k P_i x_1$. But this contradicts $x_1 P_i x_k$. Hence,the process cannot continue indefinitely and the choice set is nonempty. Q.E.D.

But Proposition 2 could clearly be strengthened by substituting the assumption of quasi-transitivity, or of acyclicity, for our fundamental assumption of transitivity, since the key to the proof is the possible existence of a cycle in the individual's preferences. In fact, the following proposition is also true. The proof is virtually the same as for Proposition 2, and is left to the reader.

Proposition 3. Let S be a finite set of alternatives available to person i. Suppose R_i is complete and acyclic. Then $C(R_i, S)$ is nonempty. That is, best choices exist.

Proposition 3 can itself be strengthened to more clearly indicate the connection between the existence of best or choice sets, and acyclicity of the preference relation. The following proposition says that when R_i is complete, best sets are always nonempty if *and only if* R_i is acyclic:

Proposition 4. Suppose R_i is complete. Then $C(R_i, S)$ is nonempty for every finite set of alternatives S available to person i, if and only if R_i is acyclic.

Proof: The "if" part of the proof follows from Proposition 3. To prove the "only if" part, we assume $C(R_i, S)$ is nonempty for every finite set of alternatives S. We want to show R_i is acyclic.

Suppose to the contrary that R_i is not acyclic. Then there exist alternatives $x_1, x_2 \ldots, x_k$ such that $x_1 P_i x_2, x_2 P_i x_3, \ldots, x_{k-1} P_i x_k$, and $x_k P_i x_1$. Let $S = \{x_1, x_2, x_3, \ldots, x_k\}$. Then $C(R_i, S)$ is empty, since every alternative in S is inferior to some other alternative in S. But this is a contradiction. Consequently R_i must be acyclic. Q.E.D.

The propositions above answer this question: Given particular assumptions about a person's preferences, can he always identify best alternatives? The next proposition answers a different question: Is there a numerical function, a utility function, which represents a person's preferences? If the answer is yes, then familiar mathematical tools can be applied to the problem of identifying best alternatives, since the search for a best alternative reduces to the problem of maximizing a utility function. If the answer is no, the use of utility functions, indifference curves, and all the other common tools of economics, is very likely illegitimate.

It turns out that the answer is yes if preferences are complete and transitive. (And in this case, acyclicity cannot substitute for transitivity.)

Proposition 5. Let S be a finite set of alternatives available to person i. Suppose R_i is complete and transitive.

Then we can assign numerical values $u_i(x), u_i(y), u_i(z)$, etc., to the alternatives in S so that

$$u_i(x) \geq u_i(y) \text{ and only if } xR_iy.$$

In other words, there is a utility function u_i, which places values on the alternatives that exactly reflect i's preferences. The proof is in the appendix to this chapter.

We should note that u_i could be transformed without altering its preference representation property. For instance, if we define $v_i = u_i + C$, where C is any constant, then $v_i(x) \geq v_i(y)$ if and only if $u_i(x) \geq u_i(y)$, if and only if xR_iy. Therefore, v_i represents R_i as well as u_i does. And if $u_i(x) \geq 0$ for all x's, u_i^2 would represent R_i as well as u_i. In fact, any transformation of u_i that does not change relative values leaves the representation property intact. These are called monotone transformations. If a utility function represents a person's preferences, any monotone transformation of that utility function is another utility function that represents the same preferences.

For this reason, u_i is called an ordinal utility function and, unlike the hypothesized utility functions of nineteenth century philosophers, it does not behave like a cardinal measure such as length: For our utility function, there exist no standard units, there are no natural zeros, and it makes no sense to add $u_i(x)$ to $u_i(y)$. Nor does it make any sense to add $u_i(x) + u_j(y)$, if u_j is another person's utility function.

What then is the use of an ordinal utility function? In fact, it transmits exactly the same information as the preference relation it represents: neither more, nor less. But a utility function allows us to analyze, in a compact and easy way, the behavior of an individual in an economic environment. It is quite correct to say that a consumer chooses a bundle of goods to maximize his utility, and the utility approach is mathematically and graphically convenient. It allows us to use the standard tools of the economist's trade.

To be able to represent preferences by means of utility functions, Proposition 5 has dealt with the case of finite sets of alternatives. However, in many applications in this book an individual will be choosing from infinite sets of alternatives. For example, a consumer will choose bundles of goods where the amount of each good is measured by a real number. For such settings, if one wishes to represent preferences by a utility function the assumption of *continuous preferences* is important. Intuitively, continuity means that the preference relation has "no jumps." Here's the definition.

Continuity. For any bundle of goods x, the upper contour set of R_i at x and the lower contour set of R_i at x are closed, i.e., they contain their boundaries. (The upper contour set of R_i at x is the set of bundles $\{y|yR_ix\}$. The lower contour set of R_i at x is the set of bundles $\{y|xR_iy\}$).

With the aid of continuity, Proposition 5 can be extended as follows:

Proposition 6. Let S be a (possibly infinite) set of bundles of goods. Suppose R_i is complete, transitive and continuous over S. Then there exists a utility function u_i defined on S which exactly reflects i's preference relation R_i.

This proposition will be used extensively in the following chapters.

4. Decisions under Uncertainty and Expected Utility

In this section we present an important special case of decision theory. It concerns problems involving *uncertainty.* Uncertainty has come to be viewed in recent decades as an important factor in many economic decisions. For example, an individual making investment decisions is uncertain about the returns he will obtain. A sports team making players' hiring decisions does not know for sure how these hires will translate into victories. The government of a country, when implementing a policy change, may not know exactly its consequences for society. For these cases and many more, the decision makers are facing a problem in which uncertainty and risk are essential components. It turns out that the theory developed for these decision problems has a very interesting mathematical structure, which we shall outline in this section.

Suppose that the set of pure alternatives (i.e., those not involving uncertainty) is $\{x_1, \ldots, x_k\}$. Each of these pure alternatives could be anything, but for simplicity and to fix ideas, let's think of each of them as a prize, a different amount of money that the individual could win. Thus, for example, the individual could end up with a prize of $x_1 = \$0$, $x_2 = \$10$ or $x_3 = \$100$.

Let $l = (q_1, \ldots, q_k)$ be a *lottery* over the pure alternatives. That is, l is a probability distribution, whereby alternative x_j occurs with probability q_j. Of course, $q_j \geq 0$ for $j = 1, \ldots, k$ and $\sum_{j=1}^{k} q_j = 1$. Continuing with the example of three monetary prizes, one could think of several lotteries: lottery $l_1 = (0.5, 0, 0.5)$ is a fair coin toss that pays \$100 if heads, and nothing if tails. Lottery $l_2 = (1/3, 1/3, 1/3)$ is a fair die toss that pays \$0 if faces 1 or 2 turn up, \$10 if 3 or 4 do, and \$100 if 5 or 6 do. Lottery $l_3 = (0, 1, 0)$ is also a lottery, but it is called a *degenerate lottery*, because it pays one of the prizes for sure (in this case, \$10).

Suppose that now the individual is asked to choose among the lotteries. Which should he choose? Note that two rational individuals may choose differently. For instance, presented with the choice between l_1 and l_3, one individual may choose l_3 because he is afraid of the high probability (one half) of getting nothing in l_1, while another person may choose l_1 because its expected prize (weighted average of prizes) is so much higher than that in l_3.

In any event, since individuals will be making decisions involving uncertainty, we model these situations as individuals choosing over the set of possible lotteries. Therefore, we assume that individuals have preferences over lotteries.

Given a set of pure alternatives $\{x_1, \ldots, x_k\}$, the set of lotteries over it is the set of all possible probability distributions. This is called the *probability simplex*:

$$\{(q_1, \ldots, q_k) | q_j \geq 0 \text{ for all } j, \ \sum_{j=1}^{k} q_j = 1\}.$$

The preference relation R_i over the probability simplex describes the preferences of the decision maker. The statement "$l_1 R_i l_2$" is read "lottery l_1 is at least as good as lottery l_2 according to person i." The preference relation R_i is used to define both the strict preference relation P_i and the indifference relation I_i, as before.

We shall assume that person i's preference relation R_i over the set of lotteries satisfies completeness, transitivity and continuity. Before we proceed, it is worth noting an important property of the set of lotteries: for any pair of lotteries l_1 and l_2 and any nonnegative constant α no greater than 1 ($\alpha \in [0, 1]$), the convex combination of the two lotteries, that is, $[\alpha l_1 + (1 - \alpha) l_2]$, is also a lottery. This is interpreted as first playing a lottery over lotteries, leading to l_1 with probability α and to l_2 with probability $1 - \alpha$, and then playing either l_1 or l_2, depending on which was chosen in the first stage. We refer to this property as the *linearity* of the set of lotteries.

Because of linearity, the assumption of continuity of preferences reduces to the following simple form:

> *Continuity.* For any three lotteries l_1 , l_2 and l_3, if $l_1 P_i l_2 P_i l_3$, there exists a number $\alpha \in (0, 1)$ such that $[\alpha l_1 + (1 - \alpha) l_3] I_i l_2$.

That is, if an individual has a strict ranking among three lotteries, so that he judges one "best" among the three, the second one "in the middle" and the third one "worst," continuity of preferences means that there must be a way to combine the best and the worst lotteries to get something that is indifferent to the one that was judged in the middle. Preference jumps are excluded.

Finally, we shall require another assumption on preferences over lotteries, also driven by the linearity of this set.

> *Independence.* For any lotteries l_1, l_2 and l_3, $l_1 R_i l_2$ if and only if $[\alpha l_1 + (1-\alpha)l_3] R_i [\alpha l_2 + (1-\alpha)l_3]$ for every number $\alpha \in [0,1]$.

Although one can construct violations of the independence assumption, its content is very intuitive. Suppose an individual judges lottery l_1 at least as good as l_2. Then, this preference should persist, should be independent, of mixing these lotteries with the same third lottery: if the choices now are that: (a) with probability α lottery l_1 will be played, and lottery l_3 will happen with probability $1-\alpha$, or (b) with probability α lottery l_2 will be played, and lottery l_3 will happen with probability $1-\alpha$, the same individual should prefer (a) over (b) or be indifferent between the two. This is simply because with probability α he is facing the choice between l_1 and l_2 (and $l_1 R_i l_2$), while with the rest of probability he is offered the same thing, i.e., l_3.

These assumptions characterize the so-called *von Neumann-Morgenstern* or *expected utility* preferences. The four axioms on preferences over lotteries lead to the von Neumann-Morgenstern expected utility theorem, named after the great mathematician and physicist John von Neumann and the economist Oskar Morgenstern:

> *von Neumann-Morgenstern Expected Utility Theorem.* The preference relation R_i over lotteries satisfies completeness, transitivity, continuity and independence if and only if it can be represented by a function that has the expected utility form. That is, there exist numbers $u_1, \ldots, u_k$ such that for any pair of lotteries $l = (q_1, \ldots, q_k)$ and $l' = (q'_1, \ldots, q'_k)$, lR_il' if and only if $\sum_{j=1}^{k} q_j u_j \geq \sum_{j=1}^{k} q'_j u_j$.
>
> *Proof*: It is easy to see that, if preferences are representable by a utility function that has the expected utility form, those preferences must satisfy the four axioms required.
>
> For the other direction, we provide a graphic proof for the case of three pure alternatives x_1, x_2 and x_3, which correspond to the degenerate lotteries l_1, l_2 and l_3, respectively. We deal with the nontrivial case in which the individual has a strict preference among these three. Let's say that $l_1 P_i l_3 P_i l_2$. The probability simplex is depicted in Figure 1.1.

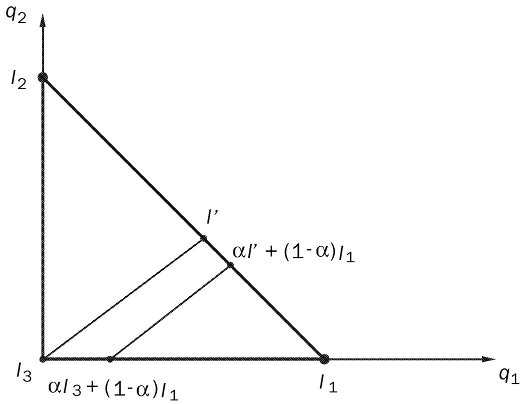

Figure 1.1.

A point in this triangle represents a lottery over the three pure alternatives (which are the degenerate lotteries l_1, l_2 and l_3). Note how the coordinates (q_1, q_2) of any point, measured from the usual origin, tell us the probabilities that the given lottery assigns to the best and to the worst alternatives (obviously, the probability that this lottery assigns to the middle alternative l_3 is simply $1 - q_1 - q_2$).

Now, completeness, transitivity and continuity of R_i guarantee the existence of a utility function representing those preferences (Proposition 6). Given such a utility function u, let $u_1 = u(l_1)$, $u_2 = u(l_2)$ and $u_3 = u(l_3)$, with $u_1 > u_3 > u_2$. What we shall show now is that this function is linear in probabilities: for any lottery $l = (q_1, q_2, q_3)$, the utility of lottery l is $u(l) = q_1u_1 + q_2u_2 + q_3u_3$.

Since $l_1P_il_3P_il_2$, by continuity, there exists $\alpha \in (0, 1)$ such that $l' = [\alpha l_1 + (1-\alpha)l_2]$ is indifferent to l_3, i.e., $l'I_il_3$, which implies that these two lotteries, l' and l_3, lie on the same *indifference curve* (a locus of points among which the individual is indifferent). Furthermore, by independence, one has that for any $\alpha \in [0, 1]$:

$$l_3 = [\alpha l_3 + (1-\alpha)l_3]I_i[\alpha l' + (1-\alpha)l_3],$$

which implies that the indifference curve passing through l' and l_3 is a straight line (recall that the locus of points that are convex combinations of two extreme points is the straight line segment connecting them). See Figure 1.1.

Finally, also from independence, since $l'I_il_3$, one also has that for any $\alpha \in [0, 1]$, $[\alpha l' + (1-\alpha)l_1]I_i[\alpha l_3 + (1-\alpha)l_1]$, and applying the previous step, we construct a new indifference curve for each value of α that is parallel to the one through l' and l_3. Next, taking combinations of l' and l_3 with l_2, one concludes that the indifference map is one of parallel straight lines. This corresponds to a function that is linear in probabilities. See Figure 1.1 again. Q.E.D.

Thus, in the problems involving uncertainty that we shall cover, we shall assume that agents have von Neumann-Morgenstern or expected utility preferences.

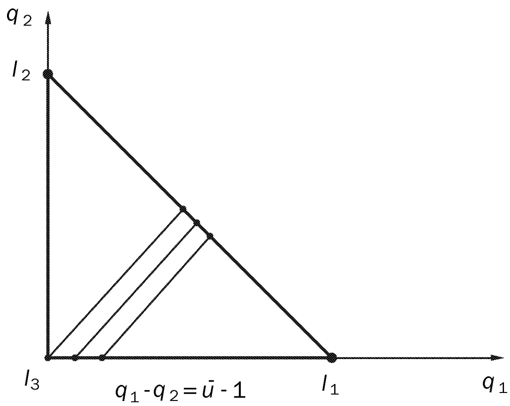

Figure 1.2.

As an illustration, Figures 1.2 and 1.3 depict two different preferences over the probability simplex, where the three degenerate lotteries are l_1, l_2, and l_3. In Figure 1.2, let the corresponding utilities $u_i(l_1) = 2$, $u_i(l_3) = 1$ and $u_i(l_2) = 0$ according to preferences R_i. For these preferences, the indifference curve of level $\bar{u}$ is the locus of points in the simplex whose equation is $2q_1 + (1 - q_1 - q_2) = \bar{u}$ or $q_1 - q_2 = \bar{u} - 1$. Not surprisingly, the top ranked point in the simplex is the degenerate lottery l_1, while the worst lottery is l_2. Figure 1.3 shows an indifference map with different expected utility preferences over lotteries. In it, $u_i'(l_1) = 4$, $u_i'(l_3) = 3$ and $u_i'(l_2) = 0$, and we call these preferences R_i'. For them, the indifference curve of level $\bar{u}$ has the equation $4q_1 + 3(1 - q_1 - q_2) = \bar{u}$ or $q_1 - 3q_2 = \bar{u} - 3$.

Note that, despite the fact that u_i' is a *monotone transformation* of u_i, both utility functions do not represent the same preferences over lotteries. This is true because, to preserve the expected utility feature, preferences can be represented only by functions that are *positive affine transformations* of one another. That is, if u_i and u_i' are two expected utility functions representing the same preferences over lotteries,

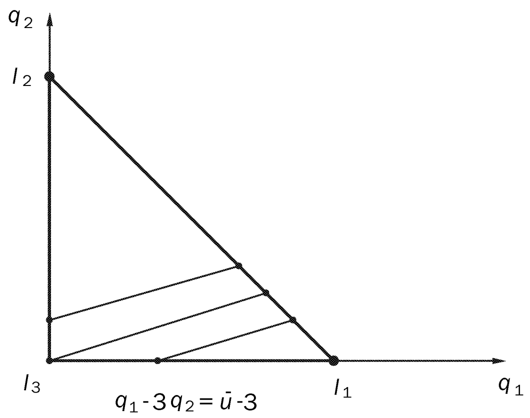

Figure 1.3.

there must exist a positive constant α and another constant β such that $u_i'(l_j) = \alpha u_i(l_j) + \beta$ for each degenerate lottery l_j.

To see that the preferences depicted in Figures 1.2 and 1.3 differ, note that the indifference curves have different slopes, and so the indifference maps are not the same. More clearly, let's exhibit two lotteries l_1 and l_2 such that l_1 is preferred to l_2 according to preferences R_i $(l_1 P_i l_2)$, while l_2 is preferred to l_1 according to R_i' $(l_2 P_i' l_1)$. Such lotteries could be, for example, $l_1 = (1/3, 1/3, 1/3)$ and $l_2 = (0.1, 0.7, 0.2)$. Indeed, for this pair of lotteries, $u_i(l_1) = 1 > 0.9 = u_i(l_2)$, but $u_i'(l_1) = 7/3 < 2.5 = u_i'(l_2)$.

5. Introduction to Social Preferences

Interest in quasi-transitivity and acyclicity arises largely from the analysis of social preferences, rather than of individual preferences. It is hard to imagine, for instance, that a person could have preferences which are acyclic but not quasi-transitive. But society's preferences are not, as we shall explain at length in later chapters, nearly so sensible as a person's.

A few examples will clarify the idea of social preferences, and the possibilities of nontransitivities for them. Suppose a group is making choices between alternatives, by using some voting rule. If x defeats y in a vote, let us say x is socially preferred to y, which we now write xPy. If x and y tie, let us say x and y are socially indifferent, which we now write xIy. If x is socially preferred to y or socially indifferent to y, we now write xRy. Where we had R_i, P_i, and I_i for individual i's preference, strict preference, and indifference relations, we now have

R, P, and I for society's preference, strict preference, and indifference relations.

Let us be more specific about the voting rules. Assume for simplicity that there are only three people in the group that is making the choices, and assume there are only three alternatives, x, y, and z.

Our first example is an instance of Condorcet's voting paradox, to which we shall return in Chapter 9 below. The voting rule is simple majority rule: a vote is taken between a pair of alternatives, and if alternative A gets more votes than alternative B, then A wins. Suppose the individuals' preferences are as follows: Person 1 prefers x to y to z. Person 2 prefers y to z to x. Person 3 prefers z to x to y. Each individual has sensible transitive preferences, but they evidently disagree on the relative merits of the three alternatives. We can indicate these preferences diagrammatically by listing the alternatives from top to bottom in the order of each person's preferences:

$$\begin{array}{ccc} \underline{1} & \underline{2} & \underline{3} \\ x & y & z \\ y & z & x \\ z & x & y. \end{array}$$

Consider a vote between x and y. Evidently, if the individuals vote according to their preferences, which we assume they do, person 1 votes for x; person 2 votes for y; and person 3 votes for x. Consequently, xPy. Next, consider a vote between y and z. Now person 1 votes for y; person 2 votes for y; and person 3 votes for z. Consequently, yPz. Finally, consider a vote between x and z. Now person 1 votes for x; person 2 votes for z; and person 3 votes for z. Consequently, zPx. We have a cycle here, since xPy, yPz, and zPx. These social preferences are not even acyclic.

The moral is social preferences might be very odd indeed — they need not share the sensible rational qualities of individual preferences. What about best sets in this example? We do have $C(R, \{x, y\}) = \{x\}$: x is best if the choice is limited to x and y. Similarly, $C(R, \{y, z\}) = \{y\}$, and $C(R, \{x, z\}) = \{z\}$. But R has a cycle. So Proposition 4 warns us that there is some set of available alternatives S for which $C(R, S)$ is empty. And, in fact, $C(R, \{x, y, z\})$ is empty: if all three alternatives are available, none is best according to majority rule. Each alternative is worse than one of the others.

Now we turn to a slightly different example. Suppose the people, alternatives, and preferences are as above, but the majority rule mechanism is modified as follows: A vote is taken between a pair of alternatives, and if alternative A gets more votes than alternative B, then A wins — unless person 1 prefers B to A. If 1 prefers B to A, and A wins a majority over B, then A and B are declared tied, or socially indifferent. We call this rule simple majority rule with a vetoer. Person 1 has a veto, in the sense that he can prevent any alternative from actually beating another alternative he prefers. What are the voting results for this rule? Consider a vote between x and y. Alternative x gets two votes to one for y, and person 1, who prefers x anyway, does not exercise his veto. Consequently, xPy. Next, consider a vote between y and z. Alternative y gets two votes to one for z, and person 1 again does not exercise his veto. Consequently, yPz. Finally, consider a vote between x and z. Alternative z gets two votes to one for x, but now person 1 does exercise his veto, since he prefers x to z. Consequently, xIz. In sum, xPy, yPz and xIz. These social preferences are acyclic, although they are not quasi-transitive. Since they are acyclic, Proposition 4 tells us that best sets are always nonempty. In fact, $C(R\{x, y, z\}) = \{x\}$ in this case; the alternative x is socially best. (It is no accident, of course, that x is also person 1's favorite.)

For the third example, we again continue with the people, alternatives and preferences above, but majority rule is now discarded. The new rule is an oligarchy of persons 1 and 2, and it works like this: A is socially preferred to B if and only if both persons 1 and 2 prefer A to B. Otherwise, A and B are socially indifferent. Now consider a "vote" between x and y. Person 1 prefers x to y, but 2 prefers y to x. Consequently, xIy. Next, consider a vote between y and z. Person 1 prefers y to z and person 2 prefers y to z. Consequently, yPz. Finally, consider a vote between x and z. Person 1 prefers x to z but person 2 prefers z to x. Consequently, xIz. In sum, xIy, yPz, and xIz. Here there are no cycles, so the social preference relation is acyclic. Moreover, the definition of quasi-transitivity is (vacuously) satisfied. (It would not be satisfied if xPy and yPz, and xIz, as in the former example.) But the social preference relation is not transitive, because transitivity requires that if xIy and yPz, then xPz must follow. So this is an example of a quasi-transitive, but not transitive, social preference relation. Note that $C(R, \{x, y, z\}) = \{x, y\}$, the favorite alternatives of the two oligarchs.

The next examples are not hypothetical as the three preceding ones. They were first discussed, in the 1970s, by Donald Brown:

We now consider two voting rules used by the United Nations Security Council. The first was in force prior to August 31, 1965. At that time there were five permanent and six nonpermanent members of the Security Council. To be passed, a motion needed seven affirmative votes, and the concurrence of all five permanent members. That is, each permanent member had to vote aye on a motion, or to abstain, or that motion would be defeated. Each permanent member had a veto. Now assuming that each nation's Ambassador had transitive (i.e., sensible) preferences, the procedure could not cycle. To see this, suppose there were a series of motions, or amendments to motions, or amendments to amendments, such that x_1 defeated x_2, x_2 defeated x_3, x_3 defeated $x_4, \ldots,$ and x_{k-1} defeated x_k. Since x_1 defeated x_2, x_1 got seven affirmative votes from the eleven members of the Council. Consequently, one of the permanent members must have voted affirmatively for x_1 over x_2. Say the United States voted affirmatively for x_1. Then the United States presumably preferred x_1 to x_2. Now x_2 was passed over x_3. Consequently, x_2 had seven affirmative votes over x_3, and the concurrence of all five permanent members. That means every permanent member either preferred x_2 to x_3, or was indifferent between the two In particular, the United States either preferred x_2 to x_3, or was indifferent between the two. Similar reasoning shows the United States either preferred x_n to x_{n+1}, or was indifferent between the two, for $n = 3, 4, \ldots, k-1$. Consequently, by repeated applications of transitivity, the United States preferred x_1 to x_k. Therefore, the United States would have used its veto power to prevent x_k's winning over x_1: so x_k could not possibly defeat x_1. A cycle could not occur: the voting rule was acyclic. From Proposition 4 we know that no matter what set of alternatives was available, the voting procedure would sensibly identify at least one best alternative.

The second United Nations Security Council voting rule was put in force on September 1, 1965. At that time, the nonpermanent membership of the Council was increased from six to ten. The permanent membership remained at five. To be passed, a motion now needs nine affirmative votes, and the concurrence of all five permanent members. (This rule remains in effect in 2005.) This procedure can cycle. To see this, we construct an example. There are ten alternatives, labeled $x_1, x_2, \ldots, x_{10}$. Assume for the sake of argument that the five permanent members are all indifferent about all these alternatives: None feels strongly enough about any of the alternatives to veto it. Assume that the preferences of the nonpermanent members are as follows: (Under member 1, we list the alternatives, from top to bottom, in that Ambassador's order of preference; similarly for 2, 3, and so on.)

The table is formidable, but the analysis is perfectly simple: Consider a vote between x_1 and x_2. Everyone except the Ambassador from Country 10 prefers x_1 to x_2. (The permanent members are indifferent.) Consequently, x_1 defeats x_2. Consider a vote between x_2 and x_3. Everyone except the Ambassador from Country 9 prefers x_2 to x_3. (The permanent members are indifferent.) Consequently, x_2 defeats x_3. Sim-

1	2	3	4	5	6	7	8	9	10
x_1	x_{10}	x_9	x_8	x_7	x_6	x_5	x_4	x_3	x_2
x_2	x_1	x_{10}	x_9	x_8	x_7	x_6	x_5	x_4	x_3
x_3	x_2	x_1	x_{10}	x_9	x_8	x_7	x_6	x_5	x_4
x_4	x_3	x_2	x_1	x_{10}	x_9	x_8	x_7	x_6	x_5
x_5	x_4	x_3	x_2	x_1	x_{10}	x_9	x_8	x_7	x_6
x_6	x_5	x_4	x_3	x_2	x_1	x_{10}	x_9	x_8	x_7
x_7	x_6	x_5	x_4	x_3	x_2	x_1	x_{10}	x_9	x_8
x_8	x_7	x_6	x_5	x_4	x_3	x_2	x_1	x_{10}	x_9
x_9	x_8	x_7	x_6	x_5	x_4	x_3	x_2	x_1	x_{10}
x_{10}	x_9	x_8	x_7	x_6	x_5	x_4	x_3	x_2	x_1

ilarly, x_3 defeats x_4, x_4 defeats x_5, and so on, until x_9 defeats x_{10}. Now consider a vote between x_1 and x_{10}. Everyone except the Ambassador from Country 1 prefers x_{10} to x_1. (The permanent members are indifferent.) Consequently, x_{10} defeats x_1, and there is a voting cycle!

To briefly summarize the observations of this section, the question of transitivity for a preference ordering, which hardly arises for an individual's ordering, does arise with a vengeance for a social preference ordering. In our discussion of individuals, where it is comfortable to assume completeness and transitivity for preferences, we shall largely use Proposition 5 and the utility functions that proposition guarantees exist. But when we return to social preferences, we shall have to return to the concepts of this chapter, and pay careful attention to ideas like completeness, transitivity, and transitivity's weaker cousins, quasi-transitivity and acyclicity.

6. Exercises

1 Show that if a preference relation R_i is transitive in the sense that xR_iy and yR_iz implies xR_iz for all x, y, and z, then (i) xP_iy and yI_iz implies xP_iz, and (ii) xI_iy and yI_iz implies xI_iz.

2 Hockey team A defeats hockey team B. Hockey team B defeats hockey team C. Hockey team A ties hockey team C.

(a) Is this preference order complete? Is it transitive? Quasi-transitive? Acyclic?

(b) Can you identify a best hockey team?

(c) Can you construct a "quality" function u for hockey teams, with the property that $u(x) > u(y)$ if and only if x defeats y? Show with numbers why you can or cannot do this.

(d) Can you construct a pseudo quality function v for hockey teams, which only satisfies this property: if x defeats y then $v(x) > v(y)$?

3 Show that if preferences over lotteries satisfy independence, then for all lotteries l_1, l_2 and l_3, one has that $l_1 I_i l_2$ if and only if $[\alpha l_1 + (1-\alpha)l_3] I_i [\alpha l_2 + (1-\alpha)l_3]$ for every $\alpha \in [0,1]$.

4 Show that if preferences over lotteries are representable by an expected utility function, they must satisfy completeness, transitivity, continuity and independence.

5 Suppose a committee has five rational members, and, for motion x to defeat motion y, x needs four affirmative votes out of the five.

(a) Show that if there are five alternatives available, there can be a voting cycle.

(b) Show that if there are only four alternatives available, there cannot be a voting cycle.

7. Appendix

Proof of Proposition 5. For notational convenience in this proof, we will drop the subscript i wherever it appears.

Suppose S is finite and R is complete and transitive. We want to show that there exists a utility function u such that

$u(x) \geq u(y)$ if and only if xRy.

First, we subdivide S into "indifference classes."

Let $C_1 = C(R, S)$. C_1 is nonempty by Proposition 2.

The alternatives in S which are not in C_1 we call $S - C_1$.

Let $C_2 = C(R, S - C_1)$. C_2 is nonempty by Proposition 2.

The alternatives in S which are not in C_1 or in C_2 we call $S - C_1 - C_2$.

Let $C_3 = C(R, S - C_1 - C_2)$. C_3 is nonempty by Proposition 2.

We continue in this fashion until we have exhausted S. This we must be able to do because S is finite. Let C_h be the last class so constructed.

$$\text{Now define } u(x) = \begin{cases} h \text{ if } x \text{ is in } C_1 \\ h-1 \text{ if } x \text{ is in } C_2 \\ \cdot \\ \cdot \\ \cdot \\ 1 \text{ if } x \text{ is in } C_h \end{cases}$$

Next we show that $u(x) \geq u(y)$ implies xRy. Suppose $u(x) \geq u(y)$. Then x is in the same class as y, or in a class constructed before the class containing y. Let C_k be the class containing x. Then x is in $C(R, S - C_1 - C_2 - \ldots - C_{k-1})$ while y is in $S - C_1 - C_2 - \ldots - C_{k-1}$. Therefore, xRy.

Finally, we will establish that xRy implies $u(x) \geq u(y)$. We will argue that $u(x) < u(y)$ implies not xRy. Suppose $u(x) < u(y)$. Let C_k be the indifference class containing x, and C_j be the indifference class containing y.

Since $u(x) < u(y)$, x's class C_k was constructed after y's class C_j. Therefore, y is in $C(R, S - C_1 - \ldots - C_{j-1})$, x is in $S - C_1 - \ldots - C_{j-1}$, but x is not in $C(R, S - C_1 - \ldots - C_{j-1})$. Therefore, yRx and there is some alternative z in $S - C_1 - \ldots - C_{j-1}$ such that yRz (because y is in the best set $C(R, S - C_1 - \ldots - C_{j-1})$) but not xRz (because x is not). By completeness, if not xRz, then zPx.

Now by transitivity, if yRz and zPx, then yPx. Hence, not xRy, which is what we wanted to establish. Q.E.D.

8. Selected References

(Items marked with an asterisk (*) are mathematically difficult.)

1. K. Arrow, *Social Choice and Individual Values*, 2nd Edition, John Wiley and Sons, Inc., New York, 1963, Chapter II.

   This is an easy to read chapter of the classic monograph by Kenneth Arrow. It has short but useful observations on older literature. Arrow's notation and formalization of preferences and best or choice sets are the ones followed in this book.

2. D.J. Brown, "Aggregation of Preferences," *Quarterly Journal of Economics*, V. 89, 1975, pp. 456-469.

   This relatively nontechnical piece by Donald Brown is meant to introduce the nonspecialist to modern variants of Arrow's Impossibility Theorem. In these variants *oligarchies* and what Brown calls *collegial polities* take the place of dictators in Arrow's original theorem. Our example above of a majority rule mechanism with a vetoer is a Brown collegial polity. Our observations about the Security Council of the United Nations are taken from this Brown article.

*3. G. Debreu, *Theory of Value*, John Wiley and Sons, Inc., New York, 1959, Chapter 4.

   Chapter 4 of Gerard Debreu's classic monograph has a rigorous proof for existence of continuous utility functions. The mathematics is rather sophisticated.

*4. R.D. Luce, "Semiorders and a Theory of Utility Discrimination," *Econometrica*, V. 24, 1956, pp. 178-191.

This article deals with individual preferences that are complete, but not transitive. Instead of transitivity, quasi-transitivity, or acyclicity, Luce assumes the following intuitive property: It is possible to "string out all the elements of S [the set of available alternatives] on a line in such a fashion that an indifference interval never spans a preference interval." For such preferences, there is a theorem similar to, but slightly different than our Proposition 5.

5. A.K. Sen, *Collective Choice and Social Welfare*, Holden-Day, Inc., San Francisco, 1970, Chapter 1*.

Amartya Sen's book provides an extremely clear treatment of most topics in social choice. Chapter 1*, on preferences relations, is formal but not difficult. Sen's Lemma 1*1 is the original version of our Proposition 4 above.

6. J. von Neumann and O. Morgenstern *Theory of Games and Economic Behavior*, Princeton University Press, 1st edition, 1944.

In its first part, this fundamental book provides the model of decision making under uncertainty, and derives expected utility from assumptions on preferences over lotteries. The rest of the book studies problems involving more than one individual, in which the decisions made by each one influence the others. These problems are called games. Although one could cite some previous contributions, this book constitutes the formal birth of game theory.