1 Feasibility and Desirability

1.1 Examples

Aesop's Fox One afternoon a fox was walking through the forest and spotted a bunch of grapes hanging from a high branch.

"Just the thing to quench my thirst," said he.

Taking a few steps back, the fox jumped and just missed the hanging grapes. Again the fox took a few paces back, jumped, and tried to reach them but still failed. Finally, giving up, the fox turned up his nose and said, "They're probably sour anyway," and walked away.

Groucho Marx's Club "I don't care to belong to a club that accepts people like me as members."

Wishful Thinking "If P is a cause for Q, and Q is enjoyable, then P is true."

1.2 Separating *Can* from *Want*

These examples should make you smile. The first is a fable dating back to the sixth century B.C.E. It's intended to be more ironic than funny. The other two examples were meant as jokes but also to convey particular messages. These examples have one basic thing in common—they are silly because they involve the confounding of feasibility and desirability, of *can* and *want*.

In the first two examples, what the protagonist wishes depends on what he can achieve. Aesop's fox evidently wanted the grapes. Only when the grapes proved unattainable did he find that he actually had not wanted them, that is, that they were sour and not worth having.

Groucho Marx probably wanted to belong to clubs to be respected and accepted. But then he found he only liked those he couldn't get into. Once a club would accept him, he no longer valued it.

From a psychological point of view, Aesop's fox is much healthier than Groucho Marx. The fox declares that he doesn't want something because he *cannot* have it, whereas Groucho Marx, because he *can*. Thus, the fox brings closer to each other what he wants and what he has, whereas Groucho Marx keeps them apart. The fox may be a caricature of people who are willing to be intellectually dishonest in order to deal with frustration, disappointment, and envy.[1] Groucho Marx makes fun of people who suffer from self-hatred to a degree that does not allow them to be happy.

However, the two examples share the following feature: the feasibility of an option affects its desirability. An option is *feasible* if it can be chosen, if it is possible for the decision maker. The *desirability* of an option is the degree to which the decision maker wants it. Thus, feasibility has to do with beliefs about the world, and desirability with wishes. It appears irrational to mix the two. For example, if you think the grapes are tasty, then they are probably still tasty even if they are hanging higher than expected. If you think that a club is respectable and would be fun to join, then it should remain so after it admitted you. Rationality, we argue, requires that desirability be independent of feasibility.

Wishful thinking refers to considering a state of affairs true only because it is desirable. Assuming that a choice is feasible because we would like it to be is a type of wishful thinking. The sentence, "If P is a cause for Q, and Q is enjoyable, then P is true," adds a humorous twist, by giving the statement the general form of a principle of logic such as *modus ponens* ("If P implies Q, and P is true, then Q is true"), but it could also be read, "If Q is enjoyable, then Q is true." Again, it seems irrational to judge the feasibility of Q (or P) based on how much we like it (or its implications). When we analyze a problem, we should be able to judge what is feasible (possible for us) independently of our goals and desires. Doing otherwise would mean failing to face reality and deluding ourselves.

We are therefore led to suggest that one of the cornerstones of rational choice is a sharp distinction between desirability and feasibility. By sharp distinction we mean not only that the two can be told apart but also that they are causally independent; one does not affect the other.

1.3 What Is Meant by *Rational*?

We identified one pillar of rational choice: the dichotomy between feasibility and desirability. This does not imply that examples that violate it, like the ones shown, cannot be found in everyday reasoning. Indeed, these examples are funny mostly because they do remind us of real cases. Moreover, we should be content that there are some real life phenomena that we do not consider rational; otherwise *rationality* would be a vacuous term because everything would qualify as rational.

What precisely is meant by *rationality*? The answer is not obvious. Often rationality is taken to imply the collection of models of individual choice developed in economics. This definition is accepted by most economists, who believe that economic agents can, for the most part, be modeled as rational according to this definition. It is also accepted by most psychologists and behavioral decision theorists, who tend to believe that these models are at odds with the data, and that people are therefore not rational. These two camps disagree on the empirical question of how close economic behavior is to the rational model, but they often agree on the definition of rationality.

I have a personal preference for a different definition of rationality, which is much more subjective. According to this definition, a mode of behavior is rational for a given person if this person feels comfortable with it, and is not embarrassed by it, even when it is analyzed for him. For example, if you don't care for clubs that are willing to accept you, I could point out, "Notice that you wanted this club until they admitted you. You don't care for them *because* they are feasible. Why would you aspire to be admitted by the next club, knowing that you will despise it, too, as soon as you're admitted to it?" I would expect most people to feel uncomfortable with Groucho Marx's choices. That is, I would expect that the separation of desirability from feasibility will be rational for most people. But if someone insisted that they felt perfectly happy with this mode of behavior, I would prefer to think of this mode as rational for them rather than dub them irrational.

The reason I like this peculiar definition of rationality is that I find it useful. An irrational mode of behavior is one that I can hope to change by talking to the decision maker, by explaining the theory to him, and so forth. A rational mode of behavior is one that is likely to remain in the data despite my teaching and preaching. I prefer to think of rationality as a notion of stability, or coherence of the decision with the

decision maker's personal standards, rather than as a medal of honor bestowed upon certain decision makers by decision theorists.

According to this view, I present in the next few chapters various ingredients of so-called rational choice, and readers are free to choose and decide which ingredients fit their notions of ideal decision making. It is likely to be the case that a principle of rational choice will be acceptable in some contexts but not in others. My goal in this exercise is not to be convinced that you should make decisions in a certain way, or that most people make decisions in this way, but to enrich your understanding of the choices made by yourself as well as by others.

1.4 Uncertainty

Often you do not know whether an option is feasible for you or whether an outcome is desirable. Do these cases result in violations of the separation of feasibility from desirability? The answer is no. Let us start with uncertainty about the feasible options. If I do not know whether I can do something, I can at least *try* to do it, and then the absence of information will be reflected in uncertainty about the outcome of this attempt. For example, I may not know if I can solve a difficult problem, but then I can think of the act "try to solve the problem for two hours," which I can (presumably) choose, and then I have uncertainty about the outcome of this act but not about its feasibility. Thus, there is no difficulty in not knowing whether something is feasible as long as our beliefs about its feasibility are determined independently of its desirability.

Next consider uncertainty about desirability. Suppose that I come to the market at the end of the day. I see only one box of strawberries left for sale. Do I want it? Well, I might suspect that if this is the only box left unsold, there might be something wrong with it. Maybe other buyers have examined it and decided to leave it for a good reason. Of course, I cannot be sure that this is the reason the box is still for sale. But the fact that it is still on the market is a signal about its quality. Taking this into account a priori, I may decide to forgo the trip to the market; if I find anything for sale, it's probably not worth having.

This sounds similar to the Groucho Marx's line. In both cases the decision makers decide not to choose an option because it is feasible. But the similarity is only superficial. In the market example, my preferences about strawberries are inherently independent of the feasible

options. In the presence of uncertainty, if I make some plausible assumptions about the behavior of other consumers, I can infer something about the quality of the good from the fact that it is feasible. That is, the link between feasibility and desirability is not a direct causal link; it is mediated by information. Had I known the quality of the strawberries, the fact that they are available for sale would not change their desirability.

In this and the following two chapters, I discuss alternatives whose outcomes are known with certainty. Later, I discuss decisions in the presence of uncertainty. We look first at alternatives that are available to the decision maker but whose outcomes are not necessarily known at the time the decision has to be taken. Then we have to refine the dichotomy between the feasible and the desirable to distinguish among three concepts: feasible, possible, and desirable. The term *feasible* will still refer to what the decision maker can decide to do, whereas *possible* will mean "can happen but not as a result of the decision maker's choice." The term *acts* is often used to refer to the feasible choices of the decision maker, and *states* ("states of nature" or "states of the world") to designate possible scenarios, the choice among which is not under the decision maker's control. This choice will be made by other decision makers or by "nature"—a nickname for randomness or chance—but not by the decision maker herself.

Under conditions of certainty, the emphasis is on the importance of the distinction between feasibility and desirability. Under uncertainty, it will be equally important to distinguish between acts and states, or between feasibility and possibility. Often people arrive at erroneous conclusions when they mistakenly assume that they have control over choices that are not actually theirs to make, or vice versa.

1.5 Zen and the Absurd

Is it so obvious that desirability should be independent of feasibility? There seem to be situations in which we wish certain things precisely because they are attainable, or unattainable, and these situations are not as funny as a Groucho Marx line. For example, consider a mathematician who attempts to solve hard problems. She dismisses trivial problems as uninteresting and "not fun" and seeks to solve precisely those problems that have so far eluded her. In this sense, the mathematician would be similar to a mountain climber who seeks to conquer a summit *because* he has not yet done it; or to an imperialist who wishes

to add another country to his list of conquests; or to an athlete who attempts to break her own record once more. In fact, we seem to be surrounded by people who seek goals precisely because they may not be attainable and who lose interest in them as soon as they are proven feasible. All of the characters Camus thinks of as "absurd" are of this type.

You may also find reasonable people who tell you that the goal doesn't really matter, it is the road that matters. Zen philosophy might be a source of inspiration for this line of thinking. And if you're interested in the way to a goal rather than in the goal itself, you may prefer a goal that is unattainable. That is, it will be desirable *because* it is not feasible.

Do these examples confound desirability and feasibility? Not necessarily. There are several distinct issues in these examples, and some are simple to incorporate in the standard model of rationality, provided the alternatives are defined appropriately. Suppose, first, that you observe me devouring peanuts. Are you going to conclude that I enjoy having many peanuts in my stomach? Probably not. It will be more reasonable to assume that I derive pleasure from the taste of peanuts rather than from their weight in my stomach. That is, I enjoy the act of consuming peanuts rather than the state of having them. Similarly, I can enjoy swimming in the pool or strolling in the woods without trying to get anywhere.

Next consider a traveler who wishes to visit as many places as possible. He enjoys traveling but derives no pleasure from a daily stroll in the woods. He finds a known place less desirable than a new one. However, he does not seek a new place *because* it may not be feasible to get there; he simply enjoys the discovery, being somewhere for the first time. This phenomenon is also within the scope of rational choice as previously described. As in the case of consuming peanuts, the carrier of utility is the act rather than the final state. Also, in this case the pleasure derived from an act is history-dependent.

The mathematician's example is a little more complicated. As in the case of devouring peanuts, the mathematician enjoys the act more than the state. As in the case of the traveler, the mathematician also seeks the pleasure of a discovery and enjoys the act only the first time. But, as opposed to the previous examples, the mathematician enjoys a solution more, the harder is the problem. That is, she desires a conquest more, the less it appears feasible at first sight. What distinguishes her from Groucho Marx, then?

The answer is not obvious. One may argue that mathematicians, like athletes, enjoy a certain type of exercise and cannot derive pleasure from exercise that requires no effort. According to this account, they do not desire an achievement because it may not be feasible; they simply need to feel their muscles flexed, as it were, to enjoy the solution. Alternatively, you may decide that a mathematician's or an athlete's career is not rational enough for you. As will always be the case, you will make the final decision about what is rational for you.

1.6 On Theories and Paradigms

The previous two sections may seem like mental acrobatics. Rather than admitting that the definition of *rationality* involving separation of desirability from feasibility is very restricted, we come up with redefinitions of concepts to save the principle we were trying to promote. Is this honest? And is there anything that could not be classified as rational by some appropriate redefinition of terms?

Theories are supposed to be refutable, and when they are refuted, we should be honest enough to admit that. However, part of the merchandise we are trying to sell is not a specific theory, but a paradigm, a system of thought, a way of organizing the world in our minds. A paradigm consists of certain more or less formal, idealized terms, but, as opposed to a specific theory, it leaves some freedom in the way these terms are mapped onto real life phenomena. Thus, what gives pleasure to the mathematician is flexible enough to be changed from "being able to prove a theorem" to "finding a proof for a theorem that has not been known before."

Throughout this book there are examples of such redefinitions. The rational choice paradigm will often be useful and insightful even when particular theories of rational choice may fail. This is, in fact, why the book is called *Rational Choice* rather than the more common "Rational Choice Theory": in the social sciences it is often hard to come up with theories that are both useful and accurate. But there are many insights and organizing principles that change the way we think about the world. The focus in this book is on the latter.

2 Utility Maximization

2.1 Example

[*Ann is sitting by a table. Barbara, her sister, enters.*]

Barbara: Hey, what's up?

Ann: Nothing.

Barbara: But you're depressed.

Ann: No, I'm not.

Barbara: C'mon, I know you better than that. You are obviously, positively, definitely depressed.

Ann: I'm not depressed, it's just that....

Barbara: ...yes?

Ann: Well, you won't tell anyone, will you?

Barbara: Of course not, you can trust me; this is what you have big sisters for.

Ann: The same way I could trust you then with the chocolate?

Barbara: Oh, don't be silly, we were kids then. [*Both smile.*]

Ann: Well, the thing is that I have three guys who want to date me, and I can't make up my mind.

Barbara: I see. Well, I have some experience in this matter. Do you like them?

Ann: Uh-huh.

Barbara: All three of them?

Ann: Uh-huh.

Barbara: You're not very selective, are you?

Ann: Thank you very much. Why not say, my little sister is so wonderful that she attracts the best guys around?

Barbara: Sure, sure, that's exactly what I meant. Anyway, you like all three?

Ann: Yes, sort of, you know, there are pluses and minuses, no one is perfect.

Barbara: Do you love any of them?

Ann: I don't know, I *think* so, I mean I sort of love each of them in some way.

Barbara: That means you're not *in* love with any of them.

Ann: Maybe. But I still don't want to be all alone. What happens if I'm never in love?

Barbara: Okay, here's my idea: you sit down, and attach to each one of them a number. The better the guy is, the higher the number. Then you select the one with the highest number.

Ann: That sounds crazy. Did you learn that at school?

Barbara: Yes, we called it utility maximization.

Ann: Sounds just like the kind of thing that you would study in a business school. How to maximize your utility. Great. Was the course titled "How to use and abuse your boyfriend"?

Barbara: Why abuse? What are you working yourself up about?

Ann: Just listen to your words: utility, maximization—this sounds so cold, so heartless! Do they also teach you to choose the boy who's richest or whose father is best connected?

Barbara: No....

Ann: This is love we're talking about, not money! This is about people, and relationships, and emotions, not about stocks and, and.... [*She begins crying.*]

Barbara: Wait a minute, cool down, okay? First, they do not teach us how to choose boyfriends there; it's a business school, not a summer camp. I was just thinking about this idea because of how we make decisions. Second, I think you're carried away with rhetoric.

Ann: Yes, sure, if I don't think you're the greatest genius on earth, I'm carried away with rhetoric.

Barbara: No, I mean it, could you give me a chance to explain?

[*Ann is silent, but it's clear she's willing to listen.*]

Barbara: And please, without getting overexcited and without attaching meaning to the particular words—that's what I meant by rhetoric: forget about the terms, think about their contents.

Ann: OK, I'm listening. But do me a favor, and don't make it as long as last time with the derivatives. I understood nothing.

Barbara: Don't worry, this is purely about concepts. And it's short.

Ann: Okay, go ahead!

Barbara: Think of your choice between any pair of these candidates.

Ann: "Candidate"! This isn't politics!

Barbara: You see, you get all hung up on the words. What do you care if I call them *candidates* or *choices* or *guys* or *alternatives*?

Ann: It's important how you refer to people. Language has an impact on the way we think. You think of them as alternatives, and immediately I start thinking that each of them is dispensable.

Barbara: I see your point. In fact, I may even agree with you, for a change. Seriously, I think that what you just said is quite deep. I wonder if economists don't get a lot of unnecessary criticism because of a poor choice of words.

Ann: It's not unnecessary. You just agreed that language has its power.

Barbara: I meant, unnecessary in the sense that what these economists have to say is actually quite sensible, but because they often choose words that turn people off, people don't listen to what they have to say.

Ann: Okay, but I'm mature and open-minded and I'm listening.

Barbara: So: consider your choice between any pair of guys.

Ann: *Any* pair?

Barbara: With three guys you have exactly three pairs. With four guys you would have six pairs, with five, ten pairs, and so on.

Ann: You promised no derivatives.

Barbara: Derivatives? Derivatives have to do with calculus. This is combinatorics.

Ann: You know what I mean.

Barbara: Okay, so take these three pairs—think of a–b, b–c, a–c.

Barbara: Would you like to be able to choose between any two?

Ann: Yes, of course, that's what I'm trying to do.

Barbara: We call this *completeness*. It means that you can always make a choice, that your preferences are *complete*.

Ann: And if I find two of them just as good?

Barbara: Ties are allowed. You can say that you are indifferent between the two; each is as good as the other. Then you may choose the first that comes to mind, but you won't need to change your choice later on. By the way, it's good for your guys, too.

Ann: Huh?

Barbara: Otherwise you'd drive them nuts. You'd say yes and no, first you and then him, and then maybe. Do you know, for instance, that Franz Kafka was twice engaged to marry the same woman, and he canceled the marriage both times?

Ann: Really?

Barbara: Yes, it didn't really make her happy.

Ann: Why did he do that?

Barbara: Well, he was just incapable of making a decision. The point is there's nothing very romantic about this.

Ann: Okay, I get it.

Barbara: Good. Now: would your like your choices between pairs to be transitive?

Ann: What's that?

Barbara: Transitive. This means that if you think that a is at least as good as b, and b is at least as good as c, you also think that a is at least as good as c.

Ann: I guess so.

Barbara: Sure, you want to make such decisions!

Ann: Here we go again. Big wise sister telling Ann what she wants.

Barbara: No, no, no, not because I'm your big sister, and not because I'm wise, though both are true.

[*Ann rolls her eyes.*]

Barbara: You want to be transitive because otherwise you'll be dating c and leaving him for b, then dating b and leaving him for a, and then you'll send a away and go back to c, and so on, until they're all fed up with you. If you are not transitive, you will be cruel to all the guys involved, and if they have any backbone, you'll be cruel to yourself, too.

Ann: Oh, I thought that being faithful to one means being cruel to all the others.

Barbara: Did I ever say that?

Ann: Da Ponte did, giving this line to Don Giovanni.

Barbara: Oh, good. I was afraid I might have been too honest.

Ann: Very funny.

Barbara: But you get the point—if you want to be neither as indecisive as Kafka nor as fickle as Don Giovanni, you have to be complete and transitive.

Ann: Okay, suppose I am. What wouldn't one do for one's sister!

Barbara: The point is that if you are complete and transitive in your preferences, then it is as if you are maximizing a utility function.

Ann [*suspiciously*]: Function? This is something with a derivative, isn't it?

Barbara [*smiling*]: It might have a derivative in calculus. But all I mean is a rule, a way to assign numbers to alternatives.

Ann: What's a way? What is not a way?

Barbara: Just think of a table, where in one column you have the name of the alternative, and in another, the numerical value you attach to it.

Ann: If you mean a table, why do you call it a function? Sometimes I feel you really don't want me to understand what you're saying.

Barbara: I'm sorry. Don't give me this look, I really mean it. The reason it's called a function is that sometimes it will not be given by a table but by a formula. You know, like writing $2x$ instead of listing the value for each and every value of x.

Ann: Okay. But I can think of a function as a table?

Barbara: Yes, you can think of it as a table of values that is sometimes more succinctly described by a formula.

Ann: Great. But what did you want a function for?

Barbara: You're so argumentative, I nearly forgot why I mentioned a function in the first place. But I think it's coming back to me. I said that if your preferences are complete and transitive, then I can think of you as if you were maximizing a utility function.

Ann: As if? But I'm not.

Barbara: Well, this is up to you. But let's start by agreeing that this is now only a matter of representation. One can say, "Ann is choosing among her alternatives by maximizing a utility function" and one can also say, "Ann is choosing whom to date in a complete and transitive way, or a decisive and faithful way," and these two statements mean exactly the same thing. It's a mathematical theorem.

Ann: What is?

Barbara: That if you have a preference—a way to compare pairs of alternatives—that is complete and transitive, then it can be represented by a utility function, so that between any two alternatives the one with the higher utility is chosen.

Ann: Always?

Barbara: Well, at least if you have finitely many alternatives. And, pretty as you are, I think that even you don't have infinitely many suitors.

Ann: You're so clever.

Barbara: More than you'd believe. There's even more: not only can I look at you and say, "Ann is maximizing a utility function," without thinking anything bad about you, I can even tell you that finding a utility function and maximizing it is the only method I know that can guarantee that you will indeed be complete and transitive in your preferences.

Ann: So you seriously suggest that I assign a number—call it *utility* if this makes you happy—to each guy and choose the one with the highest number.

Barbara: Yes, that is precisely what I suggest.

Ann: But I really hate the word *utility*. It makes me think of gas, electricity, and cable TV, not of love.

Barbara: Can we call it *payoff*?

Ann: Payoff is what you get when you gamble on horses. Or when you're killed by the mafia.

Barbara: Call it whatever you like. I thought we agreed not to attach too much importance to names. Just assign numbers to your alternatives.

Ann: But I really don't know how I would do that. How do I know if Bob should have a higher number than, say, Jim?

Barbara: Ask yourself, which one do you like better?

Ann: But that's precisely the point; I don't know which one I like better!

[*Barbara is silent.*]

Ann: I mean, this is what you were supposed to help me sort out in the first place, weren't you?

Barbara: You know what? I'll think about it.

2.2 Two Points

The example in the previous section illustrates two main points. The first is that terms like *utility* and *maximization* should not turn you off. They do not preclude emotional decision making, love and hate, lofty or base motives. To say that someone maximizes a utility function is merely to say that she is coherent in her choices. Mother Teresa could possibly be described as maximizing the number of healthy children in the world. That is, she maximized a certain function. Adolf Hitler tried to maximize the percentage of Aryan people in Germany. He also maximized a function. Thinking of Mother Teresa and Adolf Hitler as utility maximizers only says that each of them pursued a goal in a coherent way. It does not mean that they are equivalent in terms of ethical judgments, character, or anything of the sort. You are likely to admire Mother Teresa for her utility function and to loathe Adolf Hitler for his. The notion of utility maximization leaves room for all these attitudes.

The second important point, made at the end of the dialogue, is that it doesn't always help to want to maximize a utility function. In the absence of additional structure in the problem, the mathematical equivalence mentioned in the dialogue leaves us no better off than we were when we started.

The dialogue refers to a theorem stating that comparison between pairs is complete and transitive if and only if it can be described by maximization of a function (a utility function). Appendix B provides mathematical details and two formal versions of this theorem. I now turn to the theorem's interpretations.

2.3 Interpretations

The theorem in appendix B has three types of interpretations. One

namely, applications of the theory recommending modes of behavior to decision makers. The second deals with descriptive applications, that is, with situations in which the theory is interpreted as attempting to describe reality or to predict behavior. Finally, the theorem can be interpreted in a metascientific way, as a way of defining the theoretical terms.

2.3.1 Normative

Normative science refers to the activity of scientists, such as decision and game theorists, economists, and political scientists, who address members of an audience and recommend what they should be doing. The audience may be a single decision maker, as in the case of choosing a retirement plan, or a whole country, as in the case of writing a constitution. The main point about normative science is that it's not about describing reality but rather about changing it. Normative science does not try to say how things are but how they *should* be.

Wait a minute, one might think. How does the scientist know? Where does she derive her authority from? Isn't it a bit pretentious to tell people how they should run their lives or to preach to societies what laws they should abide by?

Indeed, a good question. Sometimes people forget what can and what cannot be expected of a social scientist. Let us agree that social scientists are not religious preachers, and they do not have access to any external source of authority. All the social scientist can do is to help decision makers think about what's best for them. Analyzing problems, using general rules as well as specific analogies, invoking mathematical results alongside empirical and experimental findings, the scientist can try to convince decision makers that they would like to make decisions differently than they do. But it is the decision maker who has to make the final choice—the worker who decides on a retirement plan or the country that votes on a constitutional amendment.

If we take the view that the role of the normative scientist is to convince decision makers that they would like to behave in a certain way, what tools does the scientist have? How can she convince others?

In principle, one can use all strategies of debate in order to convince others. But let us assume (perhaps unrealistically) that the scientist has no ulterior motives and that she really wants to do the best for the decision maker. She doesn't want to convince him to buy her software or to keep using her services. She wants the decision maker to be convinced of her teachings and to think, even years later, that he has learned a lot from her. Hence, the scientist does not want to resort to rhetorical tricks in order to win a particular debate; she only wants to use rhetorical tools that provide robust conclusions. Note that I use the term *rhetorical* in a slightly different way than is customary; in this usage rhetoric need not be negative. To be precise, we can distinguish between negative rhetoric and positive rhetoric. Negative rhetoric refers to tricks that may make one lose a debate but for which one has good replies the morning after the debate. Positive rhetoric refers to the type of arguments that make one view the issue differently. Roughly, positive rhetoric consists of devices that you can take from the debate and later use to convince others of what you were convinced of yourself.

Mathematics is such a device. Consider the utility maximization issue again. If a scientist told the decision maker to maximize a utility function, her proposition might appear strange. But if she suggested that decisions be made in accordance with the completeness and transitivity axioms, her recommendation would seem much less controversial, perhaps even trivial. And then the theorem can be invoked to say and prove that whoever agrees with the axioms has to agree with the conclusion as well. It would be very embarrassing to accept the axioms but to reject their implications.

To conclude, the first type of interpretation of the theorem is normative; it can help convince decision makers, ourselves included, that we would actually like to behave in accordance with a particular decision model.

2.3.2 Descriptive

Theories in the social sciences are often intended to be descriptions of reality. They provide better understanding of phenomena, and enable predictions, without trying to change reality. This type of interpretation is called descriptive. If this is how we conceive of the theory of utility maximization, what does the theorem teach us? After all, it is an equivalence result. Hence, it does not say anything new about reality; it is just about (the equivalence between) different representations of the same mode of behavior.

Indeed, if a theory makes specific predictions, and it is judged by the accuracy of those predictions, then different mathematical representations of that theory will, by definition, have the same degree of accuracy. But even in the natural sciences, where one can find successful specific theories, theories are selected based not only on their accuracy but also on other criteria such as simplicity and generality. These

criteria, among others, do depend on representation. A theory may appear complex in one formulation and simple in another. Similarly, rephrasing a theory may show that it is much more general than previously suspected because in its new formulation it encompasses theories that were thought disparate.

Different representations of the same theory may be even more important when, as in the social sciences, theories often do not provide specific predictions but rather ways of thought and general insights. When we understand theories this way, as paradigms, we find that their degrees of relevance and applicability depend on our intuitive judgment of their plausibility. For example, I later discuss free markets and the reason that economists tend to like them. The argument for the optimality (or efficiency) of the market relies on the notion of utility maximization. If I told you that I believe most people maximize a utility function, you might think I was out of my mind. But if I redescribed the same theory by saying that I believed most people satisfy completeness and transitivity, my claim might appear more reasonable. Thus, the degree to which you believe in the accuracy of my claim depends on how I represent it. The more inaccurate our theories are, and the more we rely on intuition and qualitative arguments, the more important is mathematical analysis, which allows us to view theories in more than one way.

2.3.3 Metascientific

Finally, the theorem stated in section 2.2 can be viewed as relating the theoretical term *utility* to the observable term *choice*. This is in line with the logical positivist view in the philosophy of science, which held that the meaning of theoretical concepts is in their observable manifestations. While this view has been criticized within the philosophy of science, it remains a useful guideline in conducting scientific work as well as in everyday life and in political debates. Before we start arguing, it is always a good idea to ask what terms mean precisely. We may find that we are referring to the same thing by different names, or that we use the same word for completely different notions. In our case, the theorem says what is the meaning of *utility*: according to this *revealed preference* paradigm, utility is that function whose maximization is compatible with the choices of the decision maker. This means, in particular, that two utility functions that are equivalent in terms of the observed choices they predict should not be considered different, and we should not waste time and energy trying to decide which one is the correct one.

2.4 Measurement Issues

If we think of observable choice behavior as defining the theoretical concept of utility, we can ask, is the utility function that is compatible with the data unique? Or, given certain choice data, can there be different functions that deserve to be called the utility of the decision maker because each of them can provide a description of her choice by utility maximization?

The question of uniqueness arises whenever one attempts to measure a certain quantity. Typically, a measurement function cannot be unique because the unit of measurement matters. For instance, one can measure weight by grams, kilograms, or ounces. It is meaningless to say that the weight of an object is 5 unless one specifies a unit of measurement and gets a more meaningful measure, such as 5 grams or 5 ounces. If one has measurements of weight using grams, one can multiply all numbers by 0.001 and get equivalent measurements of weight using kilograms. Any reasonable theory that can be stated in terms of grams can be restated in terms of kilograms. The same applies to length, which can be measured by meters, centimeters, feet, and so on. Thus, all physical quantities are measured with at least one degree of freedom, namely, the choice of the unit of measurement.

In some cases, we have even more freedom in the choice of the scale. We can choose not only the unit of measurement but also the location of zero. Consider temperature. Fahrenheit and Celsius measures differ not only in the unit of measurement, namely, the meaning of "one degree," but also in the temperature that is called zero. Height of the surface of the earth is another example. We have chosen to measure it relative to the height of the oceans, but we could have chosen a different zero.

When it comes to the measurement of utility, we can hardly expect to be able to have fewer degrees of freedom than in the measurement of physical quantities. We don't hope to be able to say, for example, "My utility from this movie is 6." If we have a function that measures utility, we can multiply it by any positive number and get another function that also measures utility. Moreover, as in the case of temperature, we can probably also determine where to set zero. In other words, given one utility function, we can add a certain number to all its values without changing anything of import. The two types of transformations—changing the unit of measurement (multiplying by a positive number) and shifting the scale (adding a number)—together allow us to take any increasing linear transformation of the utility

function in order to get another utility function for the same decision maker. Indeed, it seems natural that the same degrees of freedom that exist in the measurement of temperature will also be present in the measurement of utility.

But with the measurement of utility we have even more freedom than with temperature. If *utility* means "the function being maximized," then any increasing transformation, even if it is not linear, will have the same observable meaning. The utility function that is measured from choice behavior is said to be only *ordinal*: no meaning should be assigned to the particular values of the function; only their *ordering* matters. Thus, if a has a higher utility than b, and b has a higher utility than c, we will get the same observable implications if we set their respective values to (10, 1, 0), (10, 9, 0), or (90, 54, 2). The fact that the first alternative can be assigned the number 10 or 90, or that the range of the utility values can be 10 or 88, has no observable manifestations. Similarly, there are no observable implications to the comparison of the utility drop between a and b versus b and c. The utility values mean only that the first alternative is preferred to the second, and both are preferred to the third. Any decreasing triple of numbers describes this relation, and therefore any decreasing triple of numbers can be the utility function for this decision maker.

There are other sources of data to help identify a utility function with fewer degrees of freedom. For example, if we expect the function to be used in algebraic calculations, not only in binary comparisons, the family of utility functions that are observationally equivalent shrinks. In chapter 4, I discuss such a theory: the utility function will be used for calculations of expectation, and it will be as unique as the measure of temperature. Alternatively, if we have more data about the probability of choice, we can also pin down the utility function with fewer degrees of freedom. But it is important to recall that utility will typically not be defined in a unique way, and that, in particular, multiplying a function by a positive number yields another function that is, according to most models, observationally equivalent.

2.5 Utility and Disutility

It often seems more natural to think of the minimization of disutility rather than the maximization of utility. The two are not necessarily synonymous. Psychology distinguishes between pleasure-seeking and pain avoidance activities. For instance, when you decide which concert

to attend, it seems most natural to describe your behavior as utility maximization. By contrast, when you buy a headache medication, minimization of disutility appears to be a more intuitive description of your behavior. Moreover, there are certain patterns of decisions that differ between the two types of activities. It is tempting to think of pleasure-seeking activity as having utility values in the positive range and of pain avoidance as dealing with negative utility values. In such a model, utility will have a meaningful zero, and shifting the utility function by adding a constant will not describe the same behavior. However, it is not clear that the interaction between the two types of motivation can be fully captured by postulating different decision rules in the positive as compared to the negative ranges of the utility scale.

Importantly, data on choice behavior may not suffice to tell whether a problem is one of utility maximization or disutility minimization. When you prefer alternative a over b, it is possible that a gives you more pleasure than does b, or that a causes less pain than does b. Further, there are cases in which the distinction between pleasure seeking and pain avoidance is not obvious even to the individual involved. We eat to satisfy hunger, but we also enjoy the taste of food. We need clothes and shelter to avoid suffering, but we also derive pleasure from them when they are aesthetic and functional.

Thus, the distinction between the two types of motivation is not always sharp and is hard to draw based on choice data. Luckily, in many situations this distinction is not necessary for the description and prediction of choices. If, confronted with the choice between a and b, you consistently choose the former, I need not know what drives this choice in order to predict it on the next occasion. For these reasons, classical decision theory does not distinguish between utility maximization and disutility minimization. However, it is useful to bear in mind that in certain problems we may wish to delve into the decision maker's motivation and perhaps treat pleasure seeking differently than pain avoidance.