Questions in Decision Theory

Itzhak Gilboa

Eitan Berglas School of Economics, Tel-Aviv University, Tel Aviv 69978, Israel, and HEC, Paris 78351 Jouy-en-Josas, France; email: tzachigilboa@gmail.com

Key Words
rationality, probability, utility, reasoning, group decision

Abstract
This review surveys a few major questions in the field of decision theory. It is argued that a re-examination of some of the fundamental concepts in decision theory may have important implications to theoretical and even empirical research in economics and related fields.
1. INTRODUCTION

The field of decision theory appears to be at a crossroad, in more sense than one. Over half a century after the defining and, for many years, definitive works of the founding fathers, investigators in the field seem to be asking basic questions regarding the field’s goals and purpose, its main questions and their answers, as well as its methods of research. This soul searching has arisen partly because of empirical and experimental failures of the classical theory. In some ways, it is the natural development of a successful discipline, which attempts to refine its answers and finds that it also has to define its questions better. In any event, there is no denying that since the 1950s, decision theory has not been as active as it is presently.

In terms of goals and purposes, decision theory may be viewed as a mostly descriptive field aiming to help understand economic phenomena, a descriptive field in the service of normative economics, or perhaps a normative enterprise whose goal is to help individual decision makers pursue their own goals. Whereas subjective expected utility maximization may be a good starting point for all these (and other) goals, one may find that refinements of the theory depend on specific applications. For example, a more general theory, involving more parameters, may be beneficial if the theory is used for theoretical applications in economics. It may be an extreme disadvantage if these parameters should be estimated for aiding a patient in making a medical decision. Similarly, axioms might be plausible for individual decision making but not for group decisions, and a theory may be a reasonable approximation for routine choices but a daunting challenge for deliberative processes. The growing popularity of rational choice models also makes it harder to develop theories that fit all applications. It is not a priori clear that voter behavior is governed by the same regularities as consumer behavior or that the decision to marry obeys the same axioms as the decision to start a new business. Many may find it exhilarating that the same abstract theory can be applied to all these decisions and many more. Yet we should also be prepared to accept the conclusion that, when details are concerned, different decision theories might be needed for different goals.

Questions of methodology have also become pertinent in a way they have never been before. Experimental economics has become a respectable field, offering a new research method to a discipline that used to recognize little between pure theory and empirical work. The works of Daniel Kahneman and Amos Tversky have had an impact, and after several decades in which economics ignored psychological evidence, behavioral economics has come to the forefront, embracing laboratory findings with an enthusiasm that reminds one of the passion with which they were dismissed until the 1990s. At the same time, neuroscience has developed and become one of the most exciting fields in science at large. It has also provided a new way of addressing questions in decision theory. As opposed to mere questionnaires that have been used in cognitive and social psychology for ages, neuroeconomics uses brain scans, electrodes, and other devices that can be dismissed as anything but nonscientific. Decision theory thus ponders, what would be the right mix of axioms and theorems, questionnaires and experiments, electrodes and fMRIs?

In this review I attempt to focus on some questions that I find both exciting and important. I would have liked to present only questions, without also expressing personal biases regarding the applications of decision theory, the answers to the questions, or the methods used to address them. But these problems are interrelated, and personal biases are revealed through the choice of questions. I should therefore declare that the type of
applications I have in mind are mostly theoretical: the generation and examination of qualitative insights using theoretical models in economics and related fields.

This review does not attempt to cover all important or exciting questions faced by decision theory. There are many issues I do not mention for brevity’s sake, and many others I may not mention due to my ignorance. Problems that I have worked on in the past obviously get over-represented here. As often happens, I wonder whether I worked on these problems because I thought they were interesting or whether the causality is reversed. Be that as it may, the questions raised here constitute a partial and subjective list:

1. Rationality (Section 2): What is meant by this term, and how should we use it in a way that would minimize confusion and facilitate scientific discussions? How should the term be used in the context of descriptive versus normative applications?
2. Probability (Section 3): What do we mean when we use this term in various applications? What are the limits, if such exist, of its useful applications? How should probabilities be specified, or modeled, when they are applicable and what should replace them when they are not?
3. Utility (Section 4): What does the utility function measure, and specifically, does it have anything to do with notions such as well-being or even happiness? Should economics be concerned with such notions at all?
4. Rules and analogies (Section 5): How do these two fundamental reasoning techniques feature into the beliefs that agents form? How do they affect probability as well as moral judgments? Should these questions be of interest to economists?
5. Group decisions (Section 6): Do groups make decisions that are fundamentally different than individuals? Are the standard models better descriptions of decisions taken by organizations than those taken by individuals? Do groups make better decisions than their individual members?

2. RATIONALITY

The term rationality has been used in such a variety of ways as to render it more confusing than helpful. It might have been wise to avoid it altogether and use only new, clearly defined terms. Nevertheless, it is a useful term, provided that we agree that rationality is a subjective notion.

The definition of rationality used in Gilboa & Schmeidler (2001) is that a mode of behavior is rational for a given decision maker if, when confronted with the analysis of her behavior, the decision maker does not wish to change it. This test of rationality does not involve new factual information. That is, a decision maker who regrets her decision in light of new information, as in the case of a resolution of an unsuccessful bet, may still be rational. But rationality will be undermined if the decision maker regrets her decision only as a result of a theoretical argument, as might be the case if she exhibits cyclical preferences, and a decision theorist points this out to her.

---

1Sections 2 and 3 summarize views that have been expressed in recent publications, coauthored with Antoine Billot, Gabrielle Gayer, Offer Lieberman, Fabio Maccheroni, Massimo Marinacci, Andy Postlewaite, Dov Samet, and David Schmeidler.

2In Gilboa (1991) rationality is related to a notion of an ascriptive theory, defined as a theory that describes a decision maker's behavior, but that can also be ascribed to her without refuting itself. Roughly, a theory is ascriptive if it is robust to its own publication.
There are several features of this definition that are not standard and may require justification. First, the definition is subjective. A mode of behavior that is rational for one person may not be rational for another. Asking whether a certain behavior, axiom, or model is rational becomes a quantitative and empirical issue, and the answer might vary with the population in question.

Second, the definition is not based on behavior alone. As opposed to the Savage (1954) axioms, for example, which are taken by many to be the definition of a rational individual, the definition above cannot be tested simply by observing choice data. One would have to be able to talk to the decision maker, expose her to analysis, and measure her degree of regret or mental discomfort.\(^3\) It may be hard to judge whether organizations, animals, or computer programs behave in a rational way, and this restricts the realm of application of the definition.

Third, this definition of rationality allows less intelligent people to be more rational than more intelligent ones. Suppose that A and B fill out a questionnaire and report preferences as in the Allais (1953) paradox. A decision theorist walks into the room and explains the von Neumann–Morgenstern (vNM, 1944) theory to them. A understands the independence axiom and wishes to change her choice. B gazes at the decision theorist but fails to understand the axiom. The decision theorist walks out and concludes that A behaved in an irrational way, whereas B was rational.

What are the merits of this definition then? We are equipped with the phenomenally elegant classical decision theory and faced with the outpour of experimental evidence à la Kahneman and Tversky, showing that each and every axiom fails in carefully designed laboratory experiments.\(^4\) What should we do in face of these violations? One approach is to incorporate them into our descriptive theories, to make the latter more accurate. This is, to a large extent, the road taken by behavioral economics. Another approach is to go out and preach our classical theories, that is, to use them as normative ones. For example, if we teach more probability calculus in high school, future generations might make less mistakes in probability judgments. In other words, we can either bring the theory closer to reality (making the theory a better descriptive one) or bring reality closer to the theory (preaching the theory as a normative one). Which should we choose?

The definition of rationality quoted here is precisely the test we need for this decision: If decision makers become convinced once the theory has been explained to them and they then wish to change their choices (that is, if their choices were irrational to them), we may declare the classical theory successful as a normative one. It would indeed be reasonable to preach the classical theory and help decision makers make better (as judged by themselves) decisions. If, however, the decision makers shrug their shoulders, consider us to be crazy mathematicians, or simply find our suggestions impractical, they will not regret their choices and will probably behave in the same way in the future. That is, they behaved in a rational way. In this case our theory failed not only descriptively but also normatively, and we have to change it to make it useful. In short, if a decision is rational for a decision maker, it is here to stay, and we can neither ignore it nor drive it away by education. If, however, it is irrational according to our definition, education may decrease its prevalence.

\(^3\)In certain situations one may translate regret to behavior, by repeating the same decision problem and observing choices before and after exposure to analysis.

\(^4\)Tversky used to say, “Give me an axiom and I’ll design the experiment that refutes it.”
The definition is rather liberal. The final judge of rationality is the decision maker herself. Rationality is not a medal of honor bestowed upon particularly smart decision makers by the university dons. Rather, it is a term that facilitates the discourse between the decision maker and the theorist.

Gilboa et al. (2010) refine this definition to distinguish between objective and subjective rationality. Both terms are defined in the context of a discussion between the decision maker and others—experts, friends, or colleagues. A decision is objectively rational if one can convince any reasonable person that it is the right thing to do. It is subjectively rational if the decision maker cannot be convinced that it is the wrong thing to do. Typically, objective rationality would leave many choices open. However, as decisions have to be made, there is an interest in completing the objectively justified choices by subjective ones, and the latter are expected not to be blatantly silly.

Accepting the subjective and therefore empirical nature of rationality has two related implications. First, it calls for a project of experimentally delineating the scope of rationality.5 We typically believe that some axioms are more basic than others. For instance, transitivity is often considered more fundamental than the vNM independence axiom. But whether this is the case should not be resolved by hard thinking by the blackboard. It is an empirical question that should be settled by gathering evidence. It is possible that the answer would vary with years of education, profession, or culture. Regardless of the answer, it would inform us in which population we may expect a certain axiom or theory to be accepted.

Second, the definition of rationality and the experimental findings we may hopefully have about rationality may serve as guides to future developments of the theory. Beyond the fact that axioms fail in the laboratory, we wish to know how colossal these failures are and, importantly, whether they can be reversed by education. If we can, theoretically and practically, improve the decisions made by people—as judged by these people themselves—it behooves upon us to try to do so. The alternative of modeling their irrationalities would do them little good (and may even help others exploit these irrationalities). If, by contrast, we are faced with a mode of behavior that is considered rational by most relevant decision makers, it will be a waste of time to try to change it, and we will do wisely to improve our models by incorporating this mode of behavior.

3. PROBABILITY

Will the U.S. president six years from now be a democrat? Clearly, we do not know the answer. However, according to the Bayesian approach, which is the dominant approach in economics, we should be able to state our subjective probability for this event. That is, we should be able to state that the U.S. president six years from now will be a democrat with a probability of 37%, or 62%, or some other number between 0 and 1. We are not allowed to say that this probability lies in a range such as [40%,60%], and we are not allowed to say that we simply do not know and have no subjective probability for the event in question.

The popularity of the Bayesian approach in economics is an interesting sociological phenomenon. I do not think that there exists any other field in which Bayesianism is so overwhelmingly dominant, or in which it is practiced with such generality, as in economic

---

5I owe this observation to John Kagel, who asked how would one experimentally test these definitions.
theory. There are two distinct assumptions that are often referred to as Bayesian. The first is that, given prior beliefs, they should be updated to posterior beliefs based on Bayes’s rule. This also means that evidence may help us form beliefs on the probability rules that generated them, taking into account the conditional probability of the observations gathered given different rules. This notion dates back to Thomas Bayes, and it is a powerful inference tool used in fields such as statistics, machine learning, computer science, and artificial intelligence.

The second Bayesian assumption is that any uncertainty can and should be quantified probabilistically. This ideological and almost religious belief is probably due to Bruno de Finetti. It has had a following within statistics and philosophy of science, but in neither field did it become the dominant approach. Importantly, the bread and butter of statistical inference remain classical statistics, using confidence intervals and hypotheses tests. Aiming to establish objective statistical proofs, classical statistics tries to avoid subjective beliefs and leaves much uncertainty unquantified (such as the uncertainty about an unknown parameter of a distribution). Clearly, classical statistics remains to this day the main workhorse of scientific research in all fields, economics included.

Whereas economists such as Keynes (1921) and Knight (1921) discussed uncertainty that could not be quantified by probabilities, de Finetti’s approach became dominant in economics. Some of its success must result from the compelling axiomatizations of subjective expected utility maximization. This idea was first suggested by Ramsey (1931), sketched (for expected value) by de Finetti (1937), and culminated in the monumental work of Savage (1954). Savage’s axioms are particularly enticing as they do not presuppose any numerical notion such as probability or utility. Yet both probability and utility, coupled with the expected utility formula, are derived from the axioms. It is hard to exaggerate the mathematical and conceptual beauty of this result.

It is also possible that the theoretical coherence of the Bayesian approach would have sufficed to popularize it among economists. The method is theoretically clean: There is but one type of uncertainty and one way to model it. There is only one way to learn, namely, to perform Bayesian updating. Moreover, various paradoxes of statistics and philosophy of science turn out to be resolved by this approach.

Over the decades, the state space describing uncertainty has expanded. A turning point was the works of Harsanyi (1967/1968), incorporating incomplete information in game theory. Harsanyi’s idea was simple and brilliant: to model any source of uncertainty explicitly in the game. Coupled with the Bayesian dogma, rational agents were now assumed to have probabilistic beliefs over the state of nature, but also over other agent’s beliefs, and so on. Moreover, all these beliefs were supposedly derived, by Bayesian updating, from a prior that these agents presumably entertained as embryos before they were born and before they found out their identity.

There is no denying that formulating problems as a game of incomplete information and computing Bayesian-Nash equilibria may be insightful. At the same time, it is remarkable that this approach became the dominant one. One cannot help wondering if the lack of concrete empirical testing in much of economic theory may have helped a beautiful but unrealistic paradigm to dominate the field. To remove any doubt, I do not think that every piece of theoretical work should be tested directly. On the contrary, some of the most important applications of economic research are general ideas, metaphors, illustrations, and so forth. These are not concrete theories that can or should be tested directly; rather,
they shape our thinking, making us aware of various effects and so on. It would be wrong to limit economic theory to that which can be directly tested and verified. However, when dealing with metaphors and illustrations, we enjoy a freedom of imagination that might sometimes becloud some important points. That many uncertainties cannot be quantified is one of them.

Most early work on violations of expected utility theory did not address the question of probability. Allais (1953) showed a violation of expected utility theory under risk, that is, with known probabilities. Ellsberg (1961), by contrast, provided examples in which many decision makers violate Savage’s axioms and do not behave as if they had a probability measure that describes their choices. In the language of Machina & Schmeidler (1992), these decision makers are not probabilistically sophisticated. Importantly, the subject of Ellsberg’s critique is not the way probabilities are used for decision making, but the very notion that there exist probabilities that summarize the relevant information for decision making.

However, Ellsberg’s experiments and the meaning of probability were largely neglected for a long while. The most famous attack on expected utility theory, namely prospect theory, proposed by Kahneman & Tversky (1979), dealt with decision making under risk. Other works by Kahneman and Tversky exposed blatant irrationalities, such as mistakes in Bayesian updating and framing effects (Tversky & Kahneman 1974, 1981). These are behaviors that are irrational in the sense above—most people who fall prey to these mistakes can be convinced that they made wrong decisions. Descriptively, Kahneman and Tversky made a powerful point: They showed that people violate even the most basic and widely accepted axioms of the classical theory. These findings, for the most part, did not challenge decision theory from a normative point of view. The more basic the axiom, the more striking it is that people violate it. At the same time, the less likely it is that we no longer consider it a desirable standard of decision making.

By contrast, the claim that we simply do not have enough data to generate probabilities for many events of interest is a normative challenge to the theory. In a seminal paper, Schmeidler (1989) suggested an alternative model for decision making under uncertainty, involving nonadditive probabilities and a notion of integration due to Choquet (1953/1954). Gilboa & Schmeidler (1989) offered another model, involving a set of prior probabilities, coupled with a decision rule that chooses an act whose minimal expected utility (over all prior probabilities in the set) is the highest. There are today several alternatives and extensions of these models, notably Klibanoff et al. (2005), Maccheroni et al. (2006), Gajdos et al. (2007), and Seo (2007). Some of these authors claim that behaving in a Bayesian way is not the highest standard of rationality: In the absence of information needed to generate a prior, it is less rational to behave as if one had such information than to admit one does not.

The meaning of probability and the scope of the concept’s applicability remain central questions of decision theory. There is little theoretical work in decision theory on

---

6 Only in Tversky & Kahneman (1992) did the authors also extend it to situations of uncertainty, suggesting an improvement of prospect theory under risk. This followed Schmeidler’s (1989) contribution.

7 Schmeidler (1989) and Gilboa & Schmeidler (1989), as well as other works mentioned below, are axiomatic models. That is, they describe a set of conditions on presumably observed behavior that are shown to be equivalent to the respective representations.

8 This view has been stated explicitly in Gilboa et al. (2008b, 2009b) (see also Shafer 1986).
formation of beliefs. Knowing more about where beliefs come from and how they are generated might help us understand which beliefs economic agents are likely to entertain in various situations, as well as which type of economic situations lend themselves to probabilistic modeling in the first place (see Gilboa et al. 2008b; 2009a,b). The Bayesian model is likely to remain the favorite benchmark of economic theory, for many good reasons. Moreover, there are many economic situations in which Bayesian analysis yields perfectly valid insights that need not be cluttered with more complicated models. But there are also many problems in which the Bayesian analysis might be misleading, suggesting insights that are a bit too simple.

4. UTILITY

What does the utility function mean? Most economists would say not much. The utility function is a mathematical device that helps us represent preferences and choices. People are typically not aware of their utility functions, and the process of utility maximization need not correspond to any actual mental process. It is a way to describe the behavior of a so-called black box, and nothing more should be read into it. In particular, the utility function is not a measure of one's well-being or happiness.

Yet we use the utility function, or equivalent constructs, to define Pareto optimality, and we treat the latter concept as an important goal for economic policies and institutions. We tend to feel that increasing the utility of people is a worthy goal. Typically, the justification for this notion is liberal: We wish for them what they would have wished for themselves, as evidenced by revealed preferences. But we do not feel the same about the supply of addictive drugs. In this case many of us feel that the utility used for describing choices is too far from our notion of well-being to warrant policies that take into account only the utility function. In other words, we can view economic analysis as interested in individuals’ well-being, but accepting that, apart from some extreme cases, revealed preferences are the best measure of well-being.

Thus viewed, one may wonder whether other goods are similar to addictive drugs in terms of the gap they introduce between utility, as measured by choice, and well-being, otherwise conceptualized. For example, Duesenberry’s (1949) relative income hypothesis suggests that well-being is determined by one’s relative standing in society. This implies that, whenever sampled, individuals would prefer more money to less. Yet obtaining more money would not necessarily make them happier, as they compare their lot with the others around them. In this case, it is true that each individual (apart, perhaps, from the richest) will be happier with higher income given the income of others, but the pursuit of material well-being by the society in its entirety is a futile enterprise.

There are many psychological phenomena in which people compare a perceptual input with a certain reference point and notice only changes from that reference point. Helson (1947, 1948, 1964) modeled these phenomena by adaptation level theory. Brickman & Campbell (1971) applied this theory to the generation of well-being as a result of an increase in income or material consumption, concluding that to derive happiness from income, one would need ever increasing levels of the latter. They therefore argued that “there is no true solution to the problem of happiness” other than “getting off the Hedonic Treadmill.”

The insight that material well-being does not guarantee happiness is not a twentieth-century discovery. Almost all religions and ancient cultures have parables and sayings to that effect. King Midas, Buddhist monks, Jesus Christ, and Jewish sages seem to
be united in making the point that money does not guarantee happiness. In modern times, the same theme is rather popular in Hollywood movies and in modern spiritualism. Happiness, we are told, may reside in love or religion, in meditation or righteousness—but not in money.

In the 1970s there were several influential studies that measured well-being and its relationship to income or to life events. Easterlin (1973, 1974) found almost no correlation between income and well-being when measured over a person’s lifetime but found a positive correlation across a cohort at a given time. He explained the findings by an adjustable aspiration level: Over one’s life, one may get richer but also raise one’s aspirations, resulting in no gain in well-being. By contrast, within a cohort, aspirations were assumed more or less constant, resulting in a positive correlation between income and reported well-being. A famous study by Brickman et al. (1978) showed that people who underwent dramatic positive and negative life-changing events (winning a lottery versus becoming paraplegic) reported major changes in well-being immediately following the event, but, after a year, returned to their normal level.

Again, the conclusion seems to be that there is no point is pursuing more material well-being. If this is the case, one wonders whether economics should indeed focus on GDP and growth as the criteria for economic success. Should we encourage people to work long hours, move in pursuit of job opportunities, and try to produce as much as they can sell? Or should we encourage them to minimize competition, spend more time with family and friends, and generally work less? Is it possible that our economic policies and institutions, designed with classical economic criteria in mind, make entire societies less happy than they could be?

The measurement of well-being by subjective reports has been criticized on several grounds. It has been shown that such reports are highly manipulable and that drawing the attention of the respondents to various aspects of their lives might have a significant effect on the reported well-being. Moreover, although reported well-being may be relative to aspiration levels, it is not obvious that these relative quantities are a valid measure of well-being. Lottery winners and paraplegics may indeed adjust their aspirations and report a similar level of well-being. But would the lottery winners be willing to switch fates? And if the answer is negative, is this not proof that reported well-being may miss some important factors?

Indeed, the reported well-being studies are relatively easy to dismiss. However, this does not mean that money does indeed buy happiness. It only implies that one needs to seek yet another measure of well-being, distinct from income and self-report. This is the goal of a project led by Kahneman in recent years. Kahneman et al. (2004) introduced the day reconstruction method (DRM) as a measure of well-being. The method assumes that well-being is the integral over time of the quality of one’s experiences and that this quality can be judged in a more or less objective way. Thus, individuals are asked to report only factual information about themselves, namely, how much time they spent engaged in various activities during their day. These activities are ranked based on pleasantness. The resulting measure is much more robust than subjective self-reports of well-being, and it does not depend on aspirations or any other subjective factors. At the same time, this measure goes beyond income or material well-being, and it may well favor less material consumption coupled with active social life over a stressful, competitive career accompanied by high income.

However, the DRM seems to miss some important determinants of happiness as well. In particular, it ignores completely the meaning and emotional value that people attach to their experiences. Getting a hug from one’s baby at the beginning of the day may make a person happy beyond the duration of the hug. It may also make the commute to work
much easier. The self-reported well-being as well as the DRM-measured well-being might indicate that having children is detrimental to one’s well-being. Yet only a minority of parents would accept this proposition, even though these are the same parents who report stress, worries, and hours spent in less than pleasurable duties. To consider another example, having won a gold medal in the Olympic games may change an athlete’s well-being forever. She may perform her tasks and go about her business as if she has never competed. Yet knowing that she has fulfilled her dream may make her happier. Similarly, spiritual aspects, serenity, and acceptance or belief in an afterlife can also affect well-being and happiness, and these are not captured by the DRM.

It seems that well-being and happiness are not satisfactorily measured by income, self-report, or even the DRM. Given these difficulties, one is tempted to do away with measurement; trust sages, writers, religious thinkers, and philosophers; and suggest that we seek happiness in the love of God or of people, in self-determination, or in the act of simply existing, but surely not in material wealth. This, however, is a dangerous conclusion. First, the mere richness of the above list, coupled with the absence of measurement, suggests that these propositions are not practicable. For how should individuals decide if their happiness lies in religious faith or in existentialism? And how would they know if they are going to be happier with or without children?

Second, the proposition that people forego material well-being for a richer spiritual life is all too familiar. It brings to mind the Marxist critique of religion, enslaving the masses for the benefit of the elite. Needless to say, Communist ideology was later subject to precisely the same critique. And the critique remains valid: It may be a laudable decision for one to drop out of modern life and its materialism, but convincing others to do so is harder to justify.

Third, with regard to the provision of food and medication, or relief for victims of natural disasters, countries with higher GDP can help more than others. Although we do not know what may promise happiness, we have gathered sufficient data to know what guarantees misery. Material wealth is needed to cope with universally painful phenomena such as famine and disease. Wealth may not maximize the happiness of the rich, but it may minimize the misery of the poor. And because we may not be able to do better, we should be careful to dismiss the pursuit of material well-being.

It is possible that the social sciences and philosophy will not be able to find prescriptions for happiness. It is even possible that many people are intrinsically incapable of being happy and that the only legitimate goal for the social sciences is the reduction of suffering. But it appears too early to reach this conclusion. If there could be a meaningful way to measure well-being and thereby to rank economic policies according to the degree of well-being they bring about, it would be hard to explain why economics should not be interested in this question. Therefore, given the present state of knowledge, we should treat the measurement of well-being as a valid and respectable research problem.

5. REASONING

Of the various modes of human reasoning, decision theory has fully embraced two—logical and Bayesian—and has largely neglected the rest. For the most part, decision makers are assumed to be perfect logical reasoners, to know all mathematical theorems, to be aware of anything that the modeler might assume or infer, and so forth. Similarly, they are assumed to have a prior probability over anything of import and to perform perfect Bayesian updating. As mentioned above, the theory does not address the question
of the origin of probabilistic beliefs. Hence, it has no room for additional modes of reasoning: Logical proofs and Bayesian updating do not allow for any other ways of thinking, or for originality or imagination. Luckily, other ways of thinking can be embedded into the Bayesian model by assuming a large-enough state space, with a prior probability that reflects the conclusions that can be arrived at by other modes of reasoning.

However, if we are interested in the formation of beliefs (probabilistic or otherwise), and if we suspect that meaningful probabilities may not exist in certain situations, we are led to question how people reason. Gladly, decision theory need not address this question from scratch. There are several millennia of thought in philosophy and psychology, and several decades of developments in statistics, machine learning, artificial intelligence, and neuroscience to draw upon in coping with this question.

Two modes of reasoning appear easily accessible to introspection and are simple enough to incorporate into decision theory: analogies and rules. The term analogy refers to a similarity that is found between two cases; a rule refers to a generalization of many cases. Both types of reasoning are sometimes referred to as inductive—one can perform case-to-case induction, or case-to-rule induction. It is worthwhile to note that both modes of reasoning appear in at least three distinct types of applications.

5.1. Prediction

The most common instance of reasoning is prediction, that is, learning from the past regarding the future. Hume (1748, section IV) pointed out the role of analogical reasoning in this task: “In reality, all arguments from experience are founded on the similarity. . . . From causes which appear similar we expect similar effects. This is the sum of all our experimental conclusions.” Wittgenstein (1922, section 6.363) attempted to define case-to-rule induction: “The procedure of induction consists in accepting as true the simplest law that can be reconciled with our experiences.” Much of statistical inference and philosophy of science, as well as machine learning and artificial intelligence, can be viewed as dealing with case-to-rule induction: finding the appropriate generalization of the data, the theory that best explains the observations, and so forth. Case-to-case induction is typically less popular, but it has also appeared in many domains, in the guise of kernel estimation in statistics (Akaike 1954), nearest-neighbor techniques in machine learning (Fix & Hodges 1951, 1952), and case-based reasoning in artificial intelligence (Schank 1986) (for an axiomatic approach to the problem, see Gilboa & Schmeidler 2003, Billot et al. 2005, and Gilboa et al. 2006; see Gilboa et al. 2008a for the definition of objective probabilities based on these techniques).

5.2. Behavior

Facing a decision problem under uncertainty, one could try to reason about the potential outcomes of various acts and make a decision based on these predictions. However, it may be hard to imagine all possible outcomes and to judge their probabilities. Correspondingly, a decision based on explicit prediction might be sensitive to misspecification errors, and one might reach a better decision by reasoning directly in terms of the act chosen rather than in terms of the outcomes to which it might lead.

Reasoning about acts may also be rule based or case based. Rule-based behavior (such as college admission based on student grade point averages) may be arrived at as a generalization of many cases in which this strategy yielded a desirable outcome. It does not
require the elaboration of the beliefs over the outcomes that might result from the act. Similarly, an example of a case-based behavior would be one where college admission is based on a student’s similarities to other cases in which admittance yielded a desirable outcome.

Gilboa & Schmeidler (1995, 2001) developed a theory of case-based decision making. The theory assumes that the only criterion used to judge the desirability of an act is how well it (or similar acts) fared in similar problems in the past. The theory ignores beliefs or predictions. Along similar lines, one can imagine a rule-based decision theory in which different rules are generalized from cases and compete to determine the act in a given problem. In the example above, one can imagine a potential candidate for whom certain rules suggest acceptance, with others suggesting rejection. In this case, the degree of accuracy of the rules and their specificity and degree of relevance to the case at hand may all be factored into their weight in the final decision.9

5.3. Moral Judgment

A different domain in which rule-based and case-based reasoning appear is moral judgment. When asked to judge what is the right, or just thing to do, people resort to both general rules and analogies. The legal code is basically a collection of rules. Its application often involves case-based reasoning, especially when precedents are discussed and compared.10 Similarly, rules and analogies guide our moral judgments of political acts and taxation policy, for example.

Rule-based and case-based reasoning are also used hypothetically to judge the moral acceptability of acts. Kant’s categorical imperative suggests that we judge acts by the outcome that would result from their generalization, namely, the state of affairs in which everyone is doing the same. Because this mental exercise involves case-to-rule induction, it is not always well-defined.11 Still, in many situations the generalization is obvious, and the categorical imperative offers a clear moral judgment. Similarly, the Golden Rule, to treat others as we would like to be treated by them, employs hypothetical analogy for moral judgment.

To summarize, we reason about what is likely to occur, what is a wise choice, and what is a just choice in terms of rules as well as analogies. In all three domains of applications, we are faced with the following problems: (a) How do we, and how should we, generalize cases to rules? (b) How should we resolve conflicts between the predictions or advice of different rules? (c) How should we aggregate cases? How do we judge the similarity of cases, and how do we use it? (d) When do people, and when should people, use case-based and rule-based reasoning?

These problems have received varying degrees of attention. For example, in the context of prediction, problem (a) is the subject of a vast literature in philosophy, statistics, and machine learning. By contrast, little seems to be known about problem (d).12 Yet such a

---

9Holland’s (1975) genetic algorithms are an example of such a system for a classification problem.
10In many systems, however, legal cases that are designed to be precedents typically come with an explanation of their scope of application as precedents. That is, they are partly generalized to rules. This differs from the way cases present themselves in history or in medical studies.
11To complicate matters, every argument for or against an act might be incorporated into the description of the problem, thereby changing the generalization that results.
12Gayer et al. (2007) study this problem empirically in an economic setup.
problem may have implications regarding economic behavior. For example, casual observation suggests that traders sometimes believe in certain rules (e.g., the market soars at a rate of 5% a year) and sometimes they do not engage in any general theorizing and base predictions on similarity to past cases. Indeed, the switch between these reasoning modes may be a contributing factor to fluctuations in stock markets: When a theory is at odds with the data, people do not change only a particular parameter of the theory; rather, they may abandon this mode of reasoning in favor of the more conservative case-based one. More generally, a better understanding of problems \((a–d)\) might provide new insights into economic problems.

6. GROUP DECISIONS AND GROUP BELIEFS

Social choice is a vast and active field that offers formal, often axiomatic treatment of aggregation of preferences, voting schemes, social welfare functions, and so forth. However, it appears that many issues having to do with the beliefs that can be ascribed to a group are far from resolved. A few examples follow.

6.1. Are Groups Better than Individuals?

Is it smarter to have groups, rather than individuals, make decisions? Will groups solve problems better? Invest more wisely? Make more coherent decisions?

Suppose first that a group of students tries to cope with a homework assignment in mathematics. Casual observation as well as experimental data suggest that the group will do better than the individuals in it—often, better than each individual in isolation. The reason appears obvious: Mathematical proofs may be hard to find, but they tend to be obvious once explicitly stated. It suffices that one individual conceives of part of a proof for the entire group to agree on and to add that part to its toolbox. Unless there are serious personality problems or particularly poor group dynamics, the group will perform better than the individuals.

By contrast, if a group has to make a choice under certainty, different tastes might complicate matters. Condorcet’s paradox famously shows that a majority vote might be cyclical, and Arrow’s (1950) impossibility theorem shows that the problem is not intrinsic to a majority vote. Indeed, we often find groups in a deadlock, unable to reach consensus, or making compromises that are not quite coherent.

The mathematical problem is an example of what experimentalists call “truth wins” (see Lorge & Solomon 1955 and Davis 1992). In such examples there is a correct answer, and when it is shown, everyone can verify that it is indeed correct. In other words, there is a choice that is objectively rational—every reasonable person will be convinced by it. In this case the group can be likened to a parallel processor computer, where each individual helps in searching the solution space, and every finding is shared with all. By contrast, in the case of a decision under certainty with differing tastes, putting several decision makers together causes more problems than it solves.

The investment problem is an intermediate case. On the one hand, some reasoning about possible investments may be acceptable by all, as in the case of a mathematical proof. On the other hand, there are aspects of taste, such as degrees of risk aversion, that make the aggregation problem more difficult and may result in choices that are less coherent than those of the individuals involved.
It is important to know when groups make better decisions than do individuals because sometimes the size of the group may be the individuals’ decision. For instance, students allowed to work in groups may decide to form a group or to do individual work. Organizations may decide to decentralize decisions or rely on a joint policy determined by a larger group. It would be desirable to be able to say more about the size and composition of optimal groups, as a function of the type of problem they face.

6.2. Should Groups Agree on Reasons?

Legal decision making gave rise to a version of Condorcet’s paradox, called the doctrinal paradox, that deals with opinions rather than with preferences. Opinions are not constrained by transitivity, but they might be constrained by logic. For instance, assume that there is a legal doctrine saying that a conclusion \( r \) can be reached if and only if both premises \( p \) and \( q \) are valid. In symbols, the doctrine is \( r \leftrightarrow (p \land q) \). Next assume that there are three judges, all of whom accept the doctrine. One believes that \( p \) is true but not \( q \). The other believes that \( q \) is true but not \( p \). Thus, they both reject the conclusion \( r \). The third judge believes that both \( p \) and \( q \) hold, and therefore she also believes that \( r \) should follow. Taking a majority vote, we find that there is a two-thirds majority for \( p \), a two-thirds majority for \( q \), but a two-thirds majority against \( r \). In other words, all three judges individually accept the doctrine, but the majority vote among them does not. Moreover, List & Pettit (2002) proved an impossibility result à la Arrow, showing that the only aggregation functions that will not be exposed to such paradoxes are dictatorial.

This impossibility result (as well as generalizations thereof) hinges on an independence axiom, stating that the aggregation of opinions on each issue should be independent of opinions on the other issues (this is akin to Arrow’s IIA axiom). One can imagine reasonable ways to aggregate opinions that do not satisfy the axiom and to which the impossibility theorem does not apply. For example, we may ask each judge to provide her subjective belief on the state space defined by \( p,q,r \) (that is, on the eight possible assignments of truth values to the three propositions) and average these beliefs to generate an aggregate belief. If each individual probability measure assigns 1 to the event \( [r \leftrightarrow (p \land q)] \), so will their average, and consistency is retained. However, it is not obvious that actual judges can be asked to specify a probability vector over eight states and to perform this task meaningfully. Casting a binary vote on each issue separately appears to be a much less demanding task.

How should inconsistency be avoided if we restrict attention to binary opinions? We may have a vote on each premise, \( p \) and \( q \), and then use the doctrine to determine the verdict on the conclusion \( r \), ignoring the individual opinions on the latter. By contrast, we may have a vote on the conclusion \( r \) and ignore the votes on the premises \( p, q \). Which method would result in better decisions?

As above, it appears that one might want to distinguish between situations that are inherently conflictual and situations that are supposedly consensual. For example, the formation of a government in a coalitional system is a result of negotiation among parties that do not even pretend to have identical interests. In such a situation an agreement on a joint action might be followed without delving into the reasoning that led to it. By contrast, consultation in a team of doctors, who are supposed to share a common goal, may reach better decisions if the doctors share their reasoning and attempt to convince each other on each of their premises. The averaging of probabilities offers a third alternative, which treats premises and
conclusions symmetrically. Finding optimal aggregation rules for various group decision situations is an interesting problem with potentially important applications.

6.3. Pareto Dominance with Subjective Beliefs

Harsanyi (1955) offered a celebrated result in support of utilitarianism. Assuming that all individuals in society, as well as society itself, are vNM expected utility maximizers, he showed that a mild Pareto condition is basically sufficient to conclude that the utility function attributed to society is an average of those of the individuals. Thus, if a society is to be rational as each of its members is (in the sense of satisfying vNM’s axioms) and is to follow unanimity preferences when these exist, the society had to aggregate preferences in a utilitarian way.

However, the vNM setup is restricted to decision problems under risk, that is, with known probabilities. Most real-life problems do not present themselves with given probabilities. Moreover, on many issues there are genuine differences of opinions. People often have different predictions regarding the results of welfare policies, the success of military operations, and even future natural phenomena such as global warming. It would have been reassuring to know that Harsanyi’s result extends to Savage’s setup, namely, to problems in which both utilities and probabilities may vary across individuals.

It turns out that this is not the case, as pointed out by Hylland & Zeckhauser (1979). Mongin (1995) provided an impossibility result, showing that one cannot have a society that is a subjective expected utility maximizer and that agrees with individual preferences whenever these agree among themselves. The obvious candidate, namely, a social utility function and a social probability measure that are averages of the individual ones, would fail to satisfy the Pareto condition in general.

These results might be disheartening. If there is no way to aggregate preferences coherently, the best intentions of political leaders cannot guarantee a desirable outcome, namely, decision making that is internally coherent (as are the individuals) and that respects unanimity vote. However, Gilboa et al. (2004) argue that the Pareto condition is not as compelling as it may seem. They suggest the following example. Two gentlemen are about to sort out a matter of honor in a duel. Each is experienced and skillful, and each believes that he is going to win the duel and come out unscathed with a probability of 90%. If one’s probability of a victory were 80% or less, the gentleman in question would rather flee town overnight. But, given their respective beliefs, each prefers that the duel takes place. Should society also have the same preferences, as implied by the Pareto condition?

Gilboa et al. (2004) argue that the answer should be negative. There are no beliefs that, if shared, would make both gentlemen risk their lives in the duel. That they agree on the preference is a result of vast disagreement over beliefs, as well as over tastes. These disagreements cancel out, as it were, and result in an agreement on the conclusion without any agreement on the premises. It seems inappropriate for society to adopt a preference for the duel based on the individuals’ preferences, as it is obvious that at least one of them is wrong. 13

The same type of reasoning casts a doubt on the concept of Pareto domination in speculative markets (see Gilboa & Schmeidler 2008). Consider, for example, a financial

---

13Gilboa et al. (2004) continue to restrict the Pareto condition to acts upon which distributions are agreed. They show that, under some conditions, this restricted Pareto condition suffices to conclude that both society’s utility and its probability are linear combinations of those of the individuals.
market. Two risk-averse (or risk-neutral) individuals trade with each other because their beliefs differ. Clearly, their tastes also differ: Each one prefers to have the other’s money. The differences in utilities and in probabilities suffice to generate a consensus that prefers trade to no trade. But, as in the duel example, this trade is no more than a bet. Should society endorse it? Alternatively, should we dismiss as suboptimal an equilibrium in which the two individuals cannot trade because the market is not complete?

It appears that the notion of Pareto dominance (and therefore also Pareto optimality) is quite different when individuals differ only in tastes as opposed to when they also differ in beliefs. Tastes cannot be wrong in the same sense that beliefs can (see Gilboa et al. 2009b). It may be an interesting challenge to understand what type of optimality criterion is appropriate for situations in which subjective beliefs might differ.

7. CONCLUSION

Decision theory touches upon fundamental questions such as rationality and reasoning, probability and uncertainty, learning and inference, justice and happiness. Correspondingly, it often overlaps with fields ranging from philosophy to machine learning, from psychology to statistics. Tracing its historical roots can be as fascinating as finding its contemporary allies or imagining its future applications.

It is quite amazing that a few thinkers in the early and mid-twentieth century could come up with simple principles that summarized a large body of philosophical thinking through the ages and charted the way for applications in decades to come. Their contributions are elegant and general, philosophically profound and mathematically brilliant. These contributions will most likely be taught centuries hence.

However, it should come as no surprise that such an elegant theory may need to be fine-tuned to accommodate specific applications. We cannot be sure that the same notion of rationality would meaningfully apply to all decision makers, individuals, or organizations, independently of culture, education, and context. We may not be able to use a single model to capture uncertainty about dice and wars, insurance and stock market behavior, product quality and global warming. We may also find that different ways of reasoning apply in different situations or that different notions of utility are relevant to different applications.

Decision theory should therefore retain a degree of open-mindedness, allowing for the possibility that different models and even different basic concepts can be used in different problems. Similarly, different methods may enrich others in addressing the same questions. The edifice we have inherited from our forefathers appears to be robust enough to support several new wings without risking collapse or disintegration.

DISCLOSURE STATEMENT

The author is not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

This review has been greatly influenced by discussions with many colleagues over the years, including Daron Acemoglu, Antoine Billot, Eddie Dekel, Gabrielle Gayer,
John Kagel, Offer Lieberman, Fabio Maccheroni, Massimo Marinacci, Andy Postlewaite, Dov Samet, Larry Samuelson, Peter Wakker, and, most of all, David Schmeidler. Daron Acemoglu has also provided many insightful comments on an earlier version. I cannot make any claim to the originality of the ideas presented here. At the same time, their content and style are sometimes different from my colleagues’, and the responsibility for unwarranted or silly claims remains with me.

LITERATURE CITED


18 Gilboa


Contents

Questions in Decision Theory
Itzhak Gilboa................................................1

Structural Estimation and Policy Evaluation in Developing Countries
Petra E. Todd and Kenneth I. Wolpin. ............................. 21

Currency Unions in Prospect and Retrospect
J.M.C. Santos Silva and Silvana Tenreyro.......................... 51

Hypothesis Testing in Econometrics
Joseph P. Romano, Azeem M. Shaikh, and Michael Wolf.............. 75

Recent Advances in the Empirics of Organizational Economics
Nicholas Bloom, Raffaella Sadun, and John Van Reenen ............. 105

Regional Trade Agreements
Caroline Freund and Emanuel Ornelas ............................... 139

Partial Identification in Econometrics
Elie Tamer .................................................................. 167

Intergenerational Equity
Geir B. Asheim ...................................................... 197

The World Trade Organization: Theory and Practice
Kyle Bagwell and Robert W. Staiger................................. 223

How (Not) to Do Decision Theory
Eddie Dekel and Barton L. Lipman ................................. 257

Health, Human Capital, and Development
Hoyt Bleakley ....................................................... 283

Beyond Testing: Empirical Models of Insurance Markets
Liran Einav, Amy Finkelstein, and Jonathan Levin. ............ 311

Inside Organizations: Pricing, Politics, and Path Dependence
Robert Gibbons ..................................................... 337
Identification of Dynamic Discrete Choice Models
    Jaap H. Abbring ............................................ 367

Microeconomics of Technological Adoption
    Andrew D. Foster and Mark R. Rosenzweig .................... 395

Heterogeneity, Selection, and Wealth Dynamics
    Lawrence Blume and David Easley ............................. 425

Social Interactions
    Steven N. Durlauf and Yannis M. Ioannides .................... 451

The Consumption Response to Income Changes
    Tullio Jappelli and Luigi Pistaferri .......................... 479

Financial Structure and Economic Welfare: Applied
    General Equilibrium Development Economics
    Robert Townsend ........................................... 507

Models of Growth and Firm Heterogeneity
    Erzo G.J. Luttmer ........................................... 547

Labor Market Models of Worker and Firm Heterogeneity
    Rasmus Lentz and Dale T. Mortensen .......................... 577

The Changing Nature of Financial Intermediation and the
    Financial Crisis of 2007–2009
    Tobias Adrian and Hyun Song Shin ............................ 603

Competition and Productivity: A Review of Evidence
    Thomas J. Holmes and James A. Schmitz, Jr. .................. 619

Persuasion: Empirical Evidence
    Stefano DellaVigna and Matthew Gentzkow ..................... 643

Commitment Devices
    Gharad Bryan, Dean Karlan, and Scott Nelson .................. 671

Errata

An online log of corrections to Annual Review of Economics
articles may be found at http://econ.annualreviews.org