

# Problems in Decision Theory

Itzhak Gilboa

July 25, 2011

## 1 Introduction

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, the field seems to be asking very basic questions regarding its goal and purpose, its main questions and their answers, as well as its methods of research. This soul searching is partly due to empirical and experimental failures of the classical theory. Partly, 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 now.

In terms of goals and purposes, decision theory may be viewed as a mostly descriptive field aiming to help understanding economic phenomena, or 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 the application one has in mind. For example, having a more general theory, involving more parameters, may be a blessing if the theory is used for theoretical applications in economics. It may be a nightmare if these parameters should be estimated for aiding a patient with making a medical decision. Similarly, axioms might be plausible for individual decision making but not for group decisions; a theory may be a reasonable approximation for routine choices but a daunting challenge for deliberative processes; and so on. 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 the behavior of voters is governed by the same regularities as the behavior of consumers, or that the decision to get married obeys the same axioms as the decision to start a new business. Many of us 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 method have also become pertinent in a way they have never been before. Experimental economics has become a respectable field, offering a new method of research 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 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 and electrodes and other devices that can be dismissed as anything but nonscientific. Decision theory thus ponders, what would be a right mix of axioms and theorems, questionnaires and experiments, electrodes and fMRIs? In this note 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. It goes without saying that the present note 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 in this note. 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 very partial and subjective list.

The questions discussed here are the following:

1. Rationality: 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 vs. normative applications?

2. Probability: What do we mean by “probability” when we use the term in various applications? What are the limits, if such exist, of useful applications of the term? How should probabilities be specified, or modeled, when they are applied, and what should replace them when they are not?

3. Utility: 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: How do these two very 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: Do groups make decisions that are fundamentally different than individuals? Are the standard models better descriptions of the decision taken by organizations than those taken by individuals? Do groups make better decisions than the individuals constituting them? The body of this paper consists of five sections, each devoted to one of the questions above, attempting to give some background, as well as some personal biases.<sup>1</sup> The final section concludes.

## 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, I think it is a useful term, provided that we agree that rationality is a subjective notion. The definition of “rationality” used in Gilboa and 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.<sup>2</sup><sup>1</sup> 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 a case of a resolution of an unsuccessful bet, may still be rational. But rationality will be undermined if the decision maker regrets her choice 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. There are several features of this definition that are not standard

---

<sup>1</sup> In 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

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, say, to the Savage axioms (Savage, 1954), 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. 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 up a questionnaire and report preferences as in the Allais paradox (Allais, 1953). 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, while B was rational.

What are the merits of this definition, then? Consider a decision theorist equipped with the phenomenally elegant classical decision theory, and faced with the outpour of experimental evidence a la Kahneman and Tversky, showing that each and every axiom fails in carefully designed laboratory experiments.<sup>4</sup> 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 highschool, future generations might make less mistakes in probability judgments. That is, 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 we find that, when we explain the theory to decision makers, they are convinced and wish to change their choices, that is, if their choices were irrational to them, we

may declare a success of the classical theory as a normative one. It would indeed be reasonable to preach the classical theory and help decision makers make better decisions as judged by them. 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. 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. In Gilboa, Maccheroni, Marinacci, and Schmeidler (2008) we 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 a wrong thing to do. Typically, objective rationality would leave many choices open. However, since 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.<sup>5</sup> We typically believe that some axioms are more basic than others. For instance, transitivity is often considered more fundamental than 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. But whatever the answer, it would tell 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 are these failures, 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 — I believe that 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 US president six years from now will be a Democrat? Clearly, we do not know the answer. But according 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 president six years hence will be a Democrat with probability 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 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 statistics, machine learning, computer science, artificial intelligence, and so forth.

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 remains 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.

While 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 be due to 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 pre-suppose 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 very 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 up-dating. 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 fact that much of economic theory is not subjected to concrete empirical tests might not have been helpful in allowing a beautiful but unrealistic paradigm to dominate the field. To remove any doubt, let me declare that I do not think that any piece of theoretical work should be directly tested. On the

contrary: I happen to believe that 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 forth. It would be wrong to limit economic theory to that which can be directly tested and verified. But, when dealing with metaphors and illustrations, we enjoy a freedom of imagination that might sometimes becloud some important points. The fact that many uncertainties cannot be quantified is one of them.

Most of the early works 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 and 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 and Tversky (1979), dealt with decision making under risk.<sup>2</sup> Other works by Kahneman and Tversky exposed blatant irrationalities such as mistakes in Bayesian updating and framing effects (Tversky and Kahneman, 1974, 1984). 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 very 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 give it up as a desirable standard of decision making.

---

<sup>2</sup> Only in Tversky and Kahneman (1992), suggesting an improvement of Prospect Theory under risk, did the authors also extend it to situations of uncertainty. This followed Schmeidler's contribution. whose minimal expected utility (over all prior probabilities in the set) is the highest.



By contrast, the claim that for many events of interest we simply do not have enough data to generate probabilities is a normative challenge to the theory. In a seminal paper, Schmeidler (1989) made the general point and suggested an alternative model for decision making under uncertainty, involving non-additive probabilities and a notion of integration due to Choquet (1953-54). Gilboa and Schmeidler (1989) offered another model, involving a set of prior probabilities, coupled with a decision rule that chooses an act. Some of these authors make the 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 doesn't.

The meaning of probability and the scope of applicability of the concept remain central questions of decision theory. There is little theoretical work in decision theory on the 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.<sup>3</sup> 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 “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, God forbid, happiness. And 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

---

<sup>3</sup> See Gilboa, Postlewaite, and Schmeidler (2008, 2009a, 2009b).

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 don't 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.

Viewed thus, 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, since they compare their lot to 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 to a certain reference point, and notice only changes from that reference point. Helson modeled these phenomena by "adaptation level theory" (Helson, 1947, 1948, 1964). Brickman and Campbell (1971) applied this theory to the generation of well-being as a result of increase in income or material consumption, concluding that in order 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" but "getting off the Hedonic Treadmill."

The insight that material well-being does not guarantee happiness was not discovered in the 20th century. Almost all religions and ancient cultures have parables and sayings to that effect. King Midas and Buddhist monks, the teachings of Christ and the saying of 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 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, Coates, and Janoff-Bulman (1978) showed that people who underwent very dramatic positive and negative life changing events (winning a lottery vs. becoming paraplegic) reported major changes in well-being immediately following the event, but, after a year, went back to their normal level.

Again, the conclusion seemed to be that there is no point in pursuing more material well-being. And 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 have been?

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, while 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 will the lottery winners be willing to switch fates? And if the answer is negative, isn't this a proof that reported wellbeing may miss some important factors?

Indeed, the reported well-being studies are relatively easy to dismiss. But 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 Daniel Kahneman in recent years. In Kahneman, Krueger, Schkade, Schwarz, and Stone (2004), the Day Reconstruction Method (DRM) is introduced as a measure of wellbeing. 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 to a rat-race 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 very 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 for ever. She may perform her tasks and go around her business as if she has never competed. But knowing that she "has made it" and fulfilled her dream may make her happier. Similarly, spiritual aspects, serenity and acceptance or belief in the afterlife can also affect well-being and happiness without being 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 and trust sages and writers, religious thinkers and philosophers, and suggest that we seek happiness in the love of God or of people, in self-determination or in Being, but surely not in material wealth. This, however, is a very dangerous conclusion. First, the mere richness of the above list, coupled with the absence of measurement, suggest that we do not propose anything 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 children or without? 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 the material rat race; but convincing others to do so is harder to justify. Third, when it comes to the provision of food and medication, or relief to victims of natural disasters, it turns out that the countries with higher GDP can help more than others. While we do not know what may promise happiness, we have gathered sufficient data to know what guarantees misery. And material wealth is needed to cope with universally painful phenomena such as famine and diseases. Wealth may not maximize the happiness of the rich, but it may minimize the misery of the poor. And since 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 recipes for happiness. It is even possible that many people are incapable of being happy by birth, 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, I think it would be hard to explain why economics should not be interested in this question. I therefore believe that, at least at 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 largely neglected all the rest. For the most part, decision makers were assumed to be perfect logical reasoners, to know all mathematical theorems and be aware of anything that the modeler might be aware of, and so forth. Similarly, they were assumed to have a prior probability over anything of import, and to perform perfect Bayesian updating. As mentioned above, the theory did not address the question of the origin of probabilistic beliefs. Hence, it had no room for additional modes of reasoning: logical proofs and Bayesian updating leave no room for any other ways of thinking, and no room for originality or imagination either. Luckily, other ways of thinking can be embedded in 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 in certain situations there may not exist meaningful probabilities at all, we are led to ask, how do people reason? Gladly, decision theory need not address this question from scratch. There are several millennia of thought in philosophy and in 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 simple enough to incorporate into decision theory: analogies and rules. “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.

**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.”

while Wittgenstein (1922, 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, philosophy of science, as well as machine learning and artificial intelligence can be viewed as dealing with case-to-rule induction: finding which is the appropriate generalization of the data, which theory 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 and Hodges, 1951, 1952), and “case-based reasoning” in artificial intelligence (Schank, 1986).<sup>4</sup>

---

<sup>4</sup> For axiomatic approach to the problem, see Gilboa and Schmeidler (2003), Billot,

Behavior Facing a decision problem under uncertainty, it makes sense to try to reason about the potential outcomes of various acts and make a decision based on these predictions. But it may be the case that it is hard to imagine all possible outcomes and to judge their probabilities.

Correspondingly, a decision that is 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 it might lead to.

Reasoning about acts may also be rule-based or case-based. Rule-based behavior such as “admit a student if their GPA is 3.8 or higher” 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, a case-based behavior might be “admit this student because he is similar to other cases in which ‘admit’ yielded a desirable outcome”.

Gilboa and Schmeidler (1995, 2001) develop a theory of case-based decision making. The theory assumes that the only criteria 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 for determining the act in a given problem. In the example above, one can imagine a potential candidate for whom certain rules suggest acceptance, and others — rejection. In this case, the degree of accuracy of the rules, their specificity and degree of relevance to the case at hand may all be factored into their weight in the final decision.<sup>11</sup> Moral judgment A different domain in which rule-based and case-based reasoning appear is moral judgment. Asked to judge what is the “right”, or “just” thing to do, people resort both to general rules and to analogies. The legal code is basically a collection of rules. Its application often involves casebased reasoning, especially when precedents are discussed and compared.<sup>12</sup> Similarly, rules and analogies guide our moral judgments of political acts, taxation policy, and so on.

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

---

Gilboa, Samet, and Schmeidler (2005), and Gilboa, Lieberman, and Schmeidler (2006).

that would result from their generalization, namely, the state of affairs in which everyone is “doing the same”. Since this mental exercise involves case-to-rule induction, it is not always well defined. Still, in many situations the generalization is obvious, and the Categorical Imperative offers a clear moral judgment. Similarly, the Golden Rule, suggesting to treat others as we would have liked to be treated by them, employs hypothetical analogy for moral judgment.

To sum, 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:

1. How do and how should we generalize cases to rules?
2. How should we resolve conflicts between the predictions or advice of different rules?
3. How should we aggregate cases? How do we judge similarity of cases and how do we use it?
4. When do (and when should) people use case-based reasoning, and when — rule-based reasoning?

These problems have received varying degrees of attention. For example, in the context of prediction, problem (1) is the subject of a vast literature in philosophy, statistics, and machine learning. By contrast, very little seems to be known about problem (4). Yet, such a problem may have implications regarding economic behavior. For example, casual observation suggests that traders sometimes believe in certain rules, such as “the market soars at a rate of 5% a year” and sometimes they do not engage in any general theorizing, and based 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 being at odds with the data, people do not only change 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 (1-4) 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 smart to have groups make decisions rather than individuals? 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 a part of a proof for the entire group to agree and add that part to its toolbox. Unless there are serious personality problems or particularly poor groups 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 majority vote might be cyclical, and Arrow's impossibility theorem (Arrow, 1950) shows that the problem is not intrinsic to majority vote. Indeed, we often find groups in a deadlock, unable to reach consensus, or making compromise decisions that are not quite coherent.

The mathematical problem is an example of what experimentalist call "Truth Wins".<sup>15</sup> In such examples there is a correct answer, and when it is shown, everyone can verify that it is indeed correct. That is, 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 case of decision under certainty with differing tastes, putting several decision makers together causes problems more than solves them.

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 up to us to decide. For instance, students who are allowed to work in groups may decide to do so, or to split to 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 the 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", which 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  $r \leftrightarrow (p \wedge q)$ . 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  $2/3$  majority for  $p$ , a  $2/3$  majority for  $q$ , but a  $2/3$  majority against  $r$ . In other words, all three judges individually accept the doctrine, but the majority vote among them does not. Moreover, List and Pettit (2002) proved an impossibility result a 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 8 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  $(r \leftrightarrow (p \wedge q))$ , 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 8 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 of the premises,  $p$  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 society is to be rational as each of its members is (in the sense of satisfying vNM's axioms), and to follow unanimity preferences when these exist, society had to aggregate preferences in a utilitarian way.

However, the vNM set-up 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 set-up, 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 and Zeckhauser (1979). Mongin (1995) provided 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, Samet, and Schmeidler (GSS, 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 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 than that it doesn't. Should society also have the same preferences, as implied by the Pareto condition?

GSS argue that the answer should be negative. There are no beliefs that, if shared, would make both gentlemen risk their lives in the duel. The fact 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, since it is obvious that at least one of them is wrong.

The same type of reasoning casts a doubt on the concept of Pareto domination in speculative markets. (See Gilboa and Schmeidler, 2008.) Consider, for example, a financial 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 sub-optimal 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 and when they also differ in beliefs. Tastes cannot be wrong in the same sense that beliefs can.<sup>17</sup> 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-20th 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 be used in different problems. Similarly, different methods may enrich each other 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.

## 8 References

- Akaike, H. (1954), "An Approximation to the Density Function", *Annals of the Institute of Statistical Mathematics*, 6: 127-132.
- Allais, M. (1953), "Le Comportement de L'Homme Rationnel devant le Risque: critique des Postulats et Axiomes de l'Ecole Americaine", *Econometrica*, 21: 503-546.
- Arrow, K.J., (1950), "A Difficulty in the Concept of Social Welfare", *Journal of Political Economy* 58, 4, pp. 328—346.
- Billot, A., I. Gilboa, D. Samet, and D. Schmeidler (2005), "Probabilities as Similarity-Weighted Frequencies", *Econometrica*, 73, 1125-1136.
- Brickman, P., & Campbell, D. T. (1971). Hedonic relativism and planning the good society. In M. H. Appley (Ed.), *Adaptation level theory: A symposium*. New York: Academic Press.
- Brickman, P., Coates, D., & Janoff-Bulman, R. (1978). "Lottery Winners and Accident Victims: Is Happiness Relative?" *Journal of Personality and Social Psychology*, 36, 917-927.
- Choquet, G. (1953-4), "Theory of Capacities," *Annales de l'Institut Fourier* 5: (Grenoble), 131-295.
- Davis, J. H. (1992) "Some Compelling Intuitions About Group Consensus Decisions, Theoretical and Empirical Research, and Interpersonal Aggregation Phenomena: Selected Examples, 1950-1990", *Organizational Behavior and Human Decision Processes*, 52: 3-38.
- de Finetti, B. (1937), "La Prevision: Ses Lois Logiques, Ses Sources Subjectives", *Annales de l'Institut Henri Poincare*, 7, 1-68.
- Duesenberry, J. S. (1949), *Income, Saving and the Theory of Consumer Behavior*. Cambridge, MA: Harvard University Press.
- Easterlin, R. (1973). "Does money buy happiness?" *The Public Interest*, 30, 3-10.
- Easterlin, R. A. (1974). *Does Economic Growth Improve the Human Lot? Some Empirical Evidence. Nations and Households in Economic Growth*. P. David and M. W. Reder, Academic Press: 89-125.

Ellsberg, D. (1961), "Risk, Ambiguity and the Savage Axioms", *Quarterly Journal of Economics*, 75: 643-669.

Fix, E. and J. Hodges (1951), "Discriminatory Analysis. Nonparametric Discrimination: Consistency Properties". Technical Report 4, Project Number 21-49-004, USAF School of Aviation Medicine, Randolph Field, TX.

Fix, E. and J. Hodges (1952), "Discriminatory Analysis: Small Sample Performance". Technical Report 21-49-004, USAF School of Aviation Medicine, Randolph Field, TX.

Gajdos, T., T. Hayashi, J.-M. Tallon, and J.-C. Vergnaud (2007), "Attitude toward imprecise information", mimeo.

Gayer, G., I. Gilboa, and O. Lieberman (2007) "Rule-Based and Case-Based Reasoning in Housing Prices", *BE Journals in Economics*, 7.

Gick, M. L. and K. J. Holyoak (1980), "Analogical Problem Solving", *Cognitive Psychology*, 12: 306-355.

——— (1983), "Schema Induction and Analogical Transfer", *Cognitive Psychology*, 15: 1-38.

Gilboa, I., (1991), "Rationality and Ascriptive Science", unpublished manuscript.

Gilboa, I., O. Lieberman, and D. Schmeidler (2006), "Empirical Similarity", *Review of Economics and Statistics*, , 88, 433-444.

——— (2008), "On the Definition of Objective Probabilities by Empirical Similarity", *Synthese*, forthcoming.

Gilboa, I., F. Maccheroni, M. Marinacci, and D. Schmeidler (2008), "Objective and Subjective Rationality in a Multiple Prior Model", mimeo.

Gilboa, I., A. Postlewaite, and D. Schmeidler (2008), "Probabilities in Economic Modeling", *Journal of Economic Perspectives*, 22: 173-188.

——— (2009a), "Rationality of Belief", *Synthese*, forthcoming.

——— (2009b), “Is It Always Rational to Satisfy Savage’s Axioms?”, *Economics and Philosophy*, forthcoming.

Gilboa, I., D. Samet, and D. Schmeidler (2004), “Utilitarian Aggregation of Beliefs and Tastes”, *Journal of Political Economy*, 112, 932-938.

Gilboa, I. and D. Schmeidler (1989), “Maxmin Expected Utility with a Non- Unique Prior”, *Journal of Mathematical Economics*, 18: 141-153.

——— (1995), “Case-Based Decision Theory”, *The Quarterly Journal of Economics*, 110: 605-639.

——— (2001), *A Theory of Case-Based Decisions*, Cambridge: Cambridge University Press.

——— (2003), “Inductive Inference: An Axiomatic Approach”, *Econometrica*, 71, 1-26.

——— (2008), “A Difficulty with Pareto Domination”, mimeo.

Harsanyi, J.C. (1955), “Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparison of Utility”, *Journal of Political Economics*, 63, 309-321.

Harsanyi, J. (1967), “Games of Incomplete Information: Part I”, *International Journal of Game Theory*, .

——— (1968), “Games of Incomplete Information: Part II”, *International Journal of Game Theory*, .

Helson, H. (1947). Adaptation-level as frame of reference for prediction of psychophysical data. *American Journal of Psychology*, 60, 1-29.

Helson, H. (1948). Adaptation-level as a basis for a quantitative theory of frames of reference. *Psychological Review*, 55, 297-313.

Helson, H. (1964), *Adaptation Level Theory: An Experimental and Systematic Approach to Behavior*. New York: Harper and Row.

Holland, J. H. (1975), *Adaptation in Natural and Artificial Systems*. Ann Arbor: University of Michigan Press.



- Hume, D. (1748), *An Enquiry Concerning Human Understanding*. Oxford: Clarendon Press.
- Hylland A. and R. Zeckhauser (1979), "The Impossibility of Bayesian Group Decision Making with Separate Aggregation of Beliefs and Values", *Econometrica*, 6, 1321-1336.
- Kahneman, D., Krueger A.B., Schkade, D.A., Schwarz, N. and Stone, A.A., (2004). "A Survey Method for Characterizing Daily Life Experience: The Day Reconstruction Method", *Science*, 306, 1776, pp. 1776—780.
- Kahneman, D. and A. Tversky (1979), "Prospect Theory: An Analysis of Decision Under Risk," *Econometrica*, 47: 263-291.
- Keynes, J. M. (1921), *A Treatise on Probability*. London: MacMillan and Co.
- Klibanoff, P., M. Marinacci, and S. Mukerji (2005), "A Smooth Model of Decision Making under Ambiguity," *Econometrica*, 73, 1849-1892.
- Knight, F. H. (1921), *Risk, Uncertainty, and Profit*. Boston, New York: Houghton Mifflin.
- List, C. and Pettit, P. (2002), "Aggregating Sets of Judgments: An Impossibility Result", *Economics and Philosophy*, 18: 89-110.
- Lorge, I. and H. Solomon, (1955) "Two Models of Group Behavior in the Solution of Eureka-Type Problems", *Psychometrika*, 20: 139-48.
- Machina, M. J. and D. Schmeidler (1992), "A More Robust Definition of Subjective Probability", *Econometrica*, 60: 745-780.
- Maccheroni, F., M. Marinacci, and A. Rustichini (2006), "Ambiguity Aversion, Robustness, and the Variational Representation of Preferences," *Econometrica*, 74: 1447-1498.
- Mongin, P. (1995), "Consistent Bayesian Aggregation", *Journal of Economic Theory*, 66, 313-351.
- Ramsey, F. P. (1931), "Truth and Probability", in *The Foundation of Mathematics and Other Logical Essays*. New York, Harcourt, Brace and Co.

Savage, L. J. (1954), *The Foundations of Statistics*. New York: JohnWiley and Sons. (Second addition in 1972, Dover)

Schmeidler, D. (1989), "Subjective Probability and Expected Utility without Additivity", *Econometrica*, 57: 571-587.

Seo, K. (2007), "Ambiguity and Second-Order Belief", mimeo.

Schank, R. C. (1986), *Explanation Patterns: Understanding Mechanically and Creatively*. Hillsdale, NJ, Lawrence Erlbaum Associates.

Shafer, G. (1986), "Savage Revisited", *Statistical Science*, 1: 463-486.

Simon, H. A. (1957), *Models of Man*. New York: JohnWiley and Sons.

Tversky, A. (1977), "Features of Similarity", *Psychological Review*, 84: 327- 352.

Tversky, A., and D. Kahneman (1974), "Judgment under uncertainty: Heuristics and biases", *Science*, 185: 1124-1131.

——— (1981), "The Framing of Decisions and the Psychology of Choice", *Science*, 211: 453-458.

——— (1992), "Advances in Prospect Theory: Cumulative Representation of Uncertainty", *Journal of Risk and Uncertainty*, 5: 297-323. 33

von Neumann, J. and O. Morgenstern (1944), *Theory of Games and Economic Behavior*. Princeton, N.J.: Princeton University Press.

Wittgenstein, L. (1922), *Tractatus Logico Philosophicus*. London, Routledge and Kegan Paul.