Chapter 18
Introduction

Horacio Arló-Costa, Vincent F. Hendricks, and Johan van Benthem

The classical account of decision-making derives from seminal work done by Frank P. Ramsey (1926) and later on by Von Neumann and Morgenstern (1947). This work culminates later on with the influential account of Leonard Savage (1954), Ascombe and Aumann (1963) and de Finetti (1974). We can recapitulate here in a compact form the classical presentation by Von Neumann and Morgenstern. Define a lottery as follows: If $A_1, \ldots, A_m$ is a partition of the possible outcomes of an experiment with $\alpha_j = Pr(A_j)$ for each j, then the lottery $(\alpha_1, \ldots, \alpha_m)$ awards prize z_j if A_j occur. We can assume that the choice of the partition events does not affect the lottery. We can then introduce some central axioms for preferences among lotteries.

Axiom 18.1 (Weak Order) There is a weak order, $\geq$, among lotteries such that $L_1 \geq L_2$ iff L_1 is not strictly preferred to L_2.

Then we have a second crucial axiom:

Axiom 18.2 (Independence) For each L, L_1, L_2, and $0 < a < 1$, $L_1 \geq L_2$ iff $aL_1 + (1 - a)L \geq aL_2 + (1 - a)L$.

Horacio Arló-Costa was deceased at the time of publication.

H. Arló-Costa (deceased)
Carnegie Mellon University, Pittsburgh, PA, USA

V.F. Hendricks (✉)
Center for Information and Bubble Studies, University of Copenhagen, Copenhagen, Denmark
e-mail: vincent@hum.ku.dk

J. van Benthem
University of Amsterdam, Science Park 904, 1098 XH Amsterdam, The Netherlands

Stanford University, Stanford, United States
e-mail: johan@science.uva.nl

H. Arló-Costa et al. (eds.), *Readings in Formal Epistemology*, Springer Graduate Texts in Philosophy 1, DOI 10.1007/978-3-319-20451-2_18

A third Archimedean axiom is often introduced to guarantee that utilities are real valued. These axioms suffice to prove that there exists a utility over prizes U such that $(\alpha_1, \ldots, \alpha_m) > (\beta_1, \ldots, \beta_m)$ iff $\sum_{I=1,m} \alpha_i U(z_i) \leq \sum_{I=1,m} \beta_i U(z_i)$ This utility is unique up to positive affine transformation. Anscombe and Aumann (1963) introduced a fourth axiom designed to explicitly say that preferences among prizes did not vary with the state, or, as it is usually put that utilities are state-independent. Savage (1954) provided an alternative set of axioms for preferences among acts (functions from states to consequences) that do not rely on an auxiliary randomization. He then shows that these axioms lead to a unique probability and state-independent utility such that acts are ranked according to their expected utilities. Savage's postulates are consistent with VonNeumann and Morgenstern's three central axioms. The corresponding theory offers what we can consider as the standard received view on models of preference by maximizing expected utility.

Savage's account of decision-making has been challenged by a series of paradoxes and counterexamples. Two of these paradoxes occupy a central role in recent theorizing. The first was offered by Maurice Allais (1953), the second by Daniel Ellsberg (1961). We will discuss here as well two additional paradoxes: one initially suggested by the physicist William Newcomb (presented explicitly in a published form by Robert Nozick (1969)) and a paradox due to Teddy Seidenfeld which appears in the article reprinted here that he coauthored with Mark J. Schervish and Joseph B. Kadane.

We can focus first on Allais's conundrum. Consider three rewards, $r_1 = \$0$, $r_2 = \$1$ million, $r_3 = \$5$ million. Now consider the following lotteries: $L_1 = 1$ million for certain; $L_2 = (.01, .89, .10)$ with prizes $z_i = r_i$; $L_3 = (.90, .10)$ with prizes $z_1 = r_1$ and $z_2 = r_3$; $L_4 = (.89, .11)$ with prizes $z_1 = r_1$ and $z_2 = r_2$. Most people choose L_1 over L_2, and L_3 over L_4. If we assume that choices reveal the underlying preference it is easy to see that this violates the classical theory of expected utility. There are, nevertheless, many cognitive explanations for this behavior. For example, Ariel Rubinstein has suggested (Similarity and Decision Making under Risk (Is there a utility theory resolution to the Allais Paradox?) *Journal of Economic Theory*, 46, 145–153, 1988) that in the first choice subjects simply go for the sure thing and in the second choice the probabilities .11 and .10 are sufficiently similar but r_3 clearly dominates r_2. Although the use of the corresponding heuristic usually leads to choices that can be justified by the theory of expected utility in this case the use of the heuristics clashes with expected utility (see Seidenfeld's article for a detailed explanation). Many psychologists concluded that in situations of this type agents reliably commit certain cognitive errors and tried to construct a theory capable of predicting this type of behavior. Daniel Kahnemann and Amos Tversky proposed a theory of this sort in a paper initially published in 1979 in the journal *Econometrica* (see the reference in the article reprinted here). The theory in question is usually called Prospect Theory (PT). The axiomatic presentation of PT abandons the corresponding version of Independence. Philosophers had different types of reactions to Allais but in general they accepted that there are at least some versions of the paradox that constitute examples of systematic errors caused by the bias induced by the use of certain heuristics.

The paradox proposed by Daniel Ellsberg is quite different. We can present here the simplest version of the paradox. Urn A contains exactly 100 balls. 50 of these balls are solid black and the remaining 50 are solid white. Urn B contains exactly 100 balls. Each of these balls is either solid black or solid white, although the ratio of black balls to white balls is unknown. Consider now the following questions: How much would you be willing to pay for a ticket that pays $25 ($0) if the next random selection from Urn A results in black (white) ball? Repeat then the same question for Urn B. It is well known that subjects tend to offer higher maximum buying prices for urn A than for urn B. This indicates that subjects do not have identical probabilities for both urns (.5 for each color) as Savage's theory predicts. It is considerably less clear that this behavior has to be interpreted as some sort of error. Ellsberg himself saw this behavior as an indication that Savage's theory has to be amended to deal with situations where uncertainty and vague or imprecise probabilities are involved. One can perfectly think, for example, that probabilities remain indeterminate in the case of Urn B. There is a vast literature dealing with decisions under ambiguity that is reviewed in the article by Gilboa and Marinacci reprinted here. As Seidenfeld's article indicates there are two main choices: either embracing a theory that abandons Axiom 18.1 (Ordering) or alternatively embracing a theory that abandons Axiom 18.2 (Independence). Seidenfeld argues that abandoning Independence (a solution that is rather popular and that Ellsberg himself supported) has a costly price: it leads to a form of sequential incoherence. Seidenfeld's argument requires the use of axioms for sequential decision making that many have found controversial. Seidenfeld's article remains mainly concerned with normative solutions to the paradoxes. The article by Tversky and Kahnemann reprinted here intends to extend the initial version of prospect theory to the case of uncertainty as well. So, they think that the common choices elicited by Ellsberg constitute also an error. This implies having a conservative attitude regarding the normative status of standard decision theory that clearly clashes with the motivation and some of the central theoretical ideas that motivated Ellsberg's work.

Mark J. Schervish, Teddy Seidenfeld and Josheph B. Kadane question in their paper another central tenet of the standard theories of decision making: the assumption that utility has to be state-independent. They show via an ingenious example that the uniqueness of probability in standard representations is relative to the choice of what counts as a constant outcome. Moreover they prove an important result showing how to elicit a unique state-dependent utility. The result does not assume that there are prizes with constant value by introducing a new kind of hypothetical kind of act in which both the prize and the state of nature are determined by an auxiliary experiment.

Our final paradox is the one proposed by the physicist William Newcomb. Consider an opaque box and a transparent box. An agent may choose one or the other taking into account the following: The transparent box contains one thousand dollars that the agent plainly sees. The opaque box contains either nothing or one million dollars, depending on a prediction already made. The prediction was about the agent's choice. If the prediction was that the agent will take both boxes, then the opaque box is empty. On the other hand, if the prediction was that the agent will take just the opaque box, then the opaque box contains a million dollars. The prediction is reliable. The agent knows all these features of his decision problem. So, we can depict the agent's options as follows:

	Prediction of one-boxing	Prediction of two-boxing
Take only one box	$\$M$	$0
Take two boxes	$\$M + \$T$	$\$T$

It is clear that two-boxing dominates one-boxing (the prizes of two-boxing are better than the prizes of one-boxing in each state of nature). So, two-boxing is the adequate choice according to dominance. Given the hypothesis of reliability of prediction, a prediction of one-boxing has a high probability given one-boxing. Similarly, a prediction of two-boxing has a high probability given two-boxing. Therefore, one-boxing's expected utility exceeds two-boxing's expected utility. One-boxing is the rational choice according to the principle of expected-utility maximization. Should one be a one-boxer or a two-boxer?

The formula used to calculate expected utility in the second case is: $U(A) = \sum_{I=1,n} \rho(S_i \mid A)\, u(A, S_i)$, where A is an act, and S_i are relevant states of nature. Joyce and Gibbard argue that two-boxing can be rationalized if one appeals to a different way of calculating expected utility:

$$UC(A) = \sum_{I=1,n} \rho(A > S_i)\, u(A, S_i) = \sum_{I=1,n} \rho(S_i \setminus A)\, u(A, S_i),$$

where the connective ">" is a counterfactual conditional and $\rho(S_i \setminus A)$ indicates a deviant type of conditional probability (called *imaging*) proposed by David Lewis. This type of conditional probability can be articulated in a paradox-free manner such that the probability of a counterfactual conditional coincides with the corresponding conditional probability (that is that $\rho(A > S_i) = \rho(S_i \setminus A)$). Classical conditional probability cannot satisfy this equation on pain on triviality (this was shown also by Lewis in a seminal paper that appeared in 1976: Probabilities of Conditionals and Conditional Probabilities, *Philosophical Review* 85, 297–315). The bibliographical notes below contain various useful pointers to recent papers debating the tenability of the corresponding notion of causal decision theory.

Suggested Further Reading

A classical and still rather useful book presenting the received view in decision theory is Savage's influential and seminal book: *The Foundations of Statistics*, Dover Publications; 2 Revised edition (June 1, 1972). A slightly more accessible but pretty thorough textbook presentation of Savage's account and beyond is the monograph by David Kreps: *Notes on the Theory of Choice*, Westview Press (May 12, 1988).

The classical essay by Daniel Ellsberg introducing his now famous paradox continues to be a very important source in this area: "Risk, Ambiguity and the Savage Axioms," *Quarterly Journal of Economics*, 75: 643-669, 1961. Isaac Levi presented a unified normative view of both Allais and Ellsberg in: "The Paradoxes of Allais and Ellsberg," *Economics and Philosophy*, 2: 23-53, 1986.

This solution abandons ordering rather than independence unlike the solution proposed by Ellsberg himself. Solutions abandoning independence have been in general more popular. Some of the classical papers in this tradition appear in the bibliography of the paper by Gilboa and Marinacci reprinted here. Many of the responses to Allais have been descriptive rather than normative. Prospect theory is a classical type of response along these lines. We reprint here an article that intends to present a unified descriptive response to both Allais and Ellsberg. The reader can find an excellent, thorough and mathematically mature presentation of the contemporary state of the art in Prospect theory in a recent book published by Peter Wakker: *Prospect Theory for Risk and Ambiguity*, Cambridge University Press, Cambridge, 2011.

The debate about the foundations of causal decision theory is in a way still open. An excellent presentation of causal decision theory can be found in an important book by Jim Joyce: *The Foundations of Causal Decision Theory*, Cambridge Studies in Probability, Induction and Decision Theory, Cambridge, 2008. Joyce has also written an interesting piece answering challenges to causal decision theory: "Regret and Instability in Causal Decision Theory," forthcoming in the second volume of a special issue of *Synthese* devoted to the foundations of the decision sciences (eds.) Horacio Arlo-Costa and Jeffrey Helzner. This special issue contains as well an essay by Wolfgang Spohn that intends to articulate the main ideas of causal decision theory by appealing to techniques used in Bayesian networks: "Reversing 30 Years of Discussion: Why Causal Decision Theorists should be One-Box." This is a promising line of investigation that has also been considered preliminary in an insightful article by Christopher Meek and Clark Glymour: "Conditioning and Intervening", *British Journal for the Philosophy of Science* 45, 1001-1021, 1994. With regard to issues related to actual causation a useful collection is the book: *Causation and Counterfactuals*, edited by J. Collins, N. Hall and L.A. Paul, MIT Press, 2004.

Finally a paper by Joseph Halpern and Judea Pearl offers a definition of actual causes using structural equations to model counterfactuals, Halpern, J. Y. and Pearl, J. (2005) "Causes and explanations: a structural-model approach. Part I: Causes", *British Journal for Philosophy of Science* 56:4, 843-887. This paper articulates ideas about causation based on recent work on Bayesian networks and related formalisms. Current work in this area seems to point to an unification of causal decision theory and an account of causation based on Bayesian networks.