alternatives according to the weighted sum calculated by multiplying the weight of each attribute with its value. The weights are real numbers between zero and one, which together sum up to one. Obviously, this type of criterion makes sense only if the degree to which each alternative satisfies any given attribute can be represented at least on an interval scale, i.e., if it makes sense to measure value in quantitative terms. Let us, for the sake of the argument, suppose that this is the case for the numbers in table above, and suppose that all attributes are assigned equal weights, i.e., 1/4. This implies that the value of alternative a_1 is $1/4 \cdot 1 + 1/4 \cdot 3 + 1/4 \cdot 1 + 1/4 \cdot 2 = 7/4$. Analogous calculations show that the value of a_2 is 2, while that of a_3 is also 2. Since we defined the ranking by stipulating that a higher number is better than a lower, it follows that a_2 and a_3 are better than a_1.

Another major sub-field of contemporary decision theory is *social choice theory*. Social choice theory seeks to analyze collective decision problems: How should a group aggregate the preferences of its individual members into a joint preference ordering? In this context, a group could be any constellation of individuals, such as a married couple, a number of friends, the members of a club, the citizens of a state, or even all conscious beings in universe. A *social choice problem* is any decision problem faced by a group, in which each individual is willing to state at least ordinal preferences over outcomes. Once all individuals have stated such ordinal preferences we have a set of *individual preference orderings*. The challenge faced by the social decision theorist is to somehow combine the individual preference ordering into a *social preference ordering*, that is, a preference ordering that reflects the preferences of the group. A *social state* is the state of the world that includes everything that individuals care about, and the term *social welfare function* (SWF) refers to any decision rule that aggregates a set of individual preference orderings over social states into a social preference ordering over those states. The majority rule used in democratic elections is an example of a SWF.

The most famous technical result in social choice theory is Arrow's impossibility theorem, according to which there is no SWF that meets a set of relatively weak normative conditions. A natural interpretation is that social decisions can never be rationally justified, simply because every possible mechanism for generating a social preference ordering – including the majority rule – is certain to violate at least one of Arrow's conditions. This result received massive attention in academic circles, and in the 1960s and 70s, many people took the theorem to prove that "democracy is impossible". However, the present view is that the situation is not that bad. By giving up or modifying some of Arrow's conditions one can formulate coherent SWFs that are not vulnerable to his impossibility result. Today, the theorem is interesting mainly because it opened up an entirely new field of inquiry.

Cross References

►Bayesian Statistics
►Decision Theory: An Overview
►Imprecise Probability
►Loss Function
►Multicriteria Decision Analysis
►Multiple Statistical Decision Theory
►Philosophical Foundations of Statistics
►Statistics and Gambling

References and Further Reading

Allais M (1953) Le comportement de l'homme rationnel devant le risque: critique des postulates et axiomes de l'ecole Américaine. Econometrica 21:503–546

Arrow KJ (1951) Social choice and individual values, 2nd edn (1963). Wiley, New York

Kahneman D, Tversky A (1979) Prospect theory: an analysis of decisions under risk. Econometrica 47:263–291

Peterson M (2009) An introduction to decision theory. Cambridge University Press, Cambridge

Ramsey FP (1926) Truth and Probability, Ramsey, 1931. In: Braithwaite RB (ed) The foundations of mathematics and other logical essays, Ch. VII. Kegan Paul, London, pp 156–198

Savage LJ (1954) The foundations of statistics, 2nd edn (1972, Dover). Wiley, New York

Trench, Trubner & Co., Harcourt, Brace and Company, New York

von Neumann J, Morgenstern O (1947) Theory of games and economic behavior, 2nd edn. Princeton University Press (1st edn without utility theory)

Decision Theory: An Overview

Sven Ove Hansson
Professor and Head
Royal Institute of Technology, Stockholm, Sweden

Decision theory (see also ►Decision Theory: An Introduction) is the systematic study of *goal-directed behavior under conditions when different courses of action (options) can be chosen*. The focus in decision theory is usually on the outcome of decisions as judged by pre-determined criteria or, in other words, on means-ends rationality. Decision theory has developed since the middle of the twentieth century through contributions from several academic disciplines. In this overview over fundamental decision theory, the focus will be on how decisions are represented and on decision rules intended to provide guidance for decision-making. Finally two paradoxes will be presented in order to exemplify the types of issues that are discussed in modern decision theory.

The Representation of Decisions

The standard representation of a decision problem requires that we specify the alternatives available to the decision-maker, the possible outcomes of the decision, the values of these outcomes, and the factors that have an influence on the outcome.

Alternatives

To decide means to choose among different alternatives (options). In some decision problems, the set of alternatives is *open* in the sense that new alternatives can be invented or discovered by the decision-maker. A typical example is your decision how to spend tomorrow evening. In other decision problems, the set of alternatives is *closed* so that no new alternatives can be added. Your decision how to vote in the upcoming elections will probably be an example of this. There will be a limited number of alternatives (candidates or parties) between which you can choose.

In real life, many if not most decisions come with an open set of alternatives. In decision theory, however, alternative sets are commonly assumed to be closed. The reason for this is that closed decision problems are much more accessible to theoretical treatment. If the alternative set is open, a definitive solution to a decision problem is not in general available.

In informal deliberations about decisions, we often refer to alternatives that can be combined with each other. Hence, when deciding how to spend tomorrow evening you may begin by choosing between eating out and going to the cinema, but end up deciding to do both. In decision theory, the alternatives are assumed to be *mutually exclusive*, i.e., no two of them can both be realized. However, this difference is not very important since you can always convert a set of compatible alternatives to a set of mutually exclusive ones. This is done by listing all the possible combinations (in this example: eating out and not going to the cinema, eating out and going to the cinema, etc.).

States of Nature

The effects of a decision depend not only on what choice the decision-maker makes, but also on various factors beyond the decision-maker's control. Some of these extraneous factors constitute *background information* that the decision-maker has access to. Others are unknown. They may depend, for instance, on the actions of other persons and various natural events.

In decision theory, it is common to summarize the various unknown extraneous factors into a number of cases, called *states of nature*. The states of nature include decisions by other persons. This is a major difference between decision theory and game theory. In game theory, decisions by several persons that may compete or cooperate are treated on a par with each other in the formal representation. In decision theory, the focus is on one decision-maker, and the actions and choices of others are treated differently, namely in the same way as natural events.

As an example, consider a young boy, Peter, who makes up his mind whether or not to go to the local soccer ground to see if there is any soccer going on that he can join. The effect of that decision depends on whether there are any soccer players present. In decision theory, this situation is described in terms of two states of nature, "players present" and "no players present."

Outcomes

The possible *outcomes* of a decision are determined by the combined effects of a chosen alternative and the state of nature that materializes. Hence, if Peter goes to the soccer ground and there are no players present, then the outcome can be summarized as "walk and no soccer," if he goes and there are players present then the outcome is "walk and soccer," and if he does not go then the outcome is "no walk and no soccer."

Decision Matrices

The alternatives, the states of nature, and the resulting outcomes in a decision can be represented in a *decision matrix*. A decision matrix is a table in which the alternatives are represented by rows and the states of nature by columns. For each alternative and each state of nature, the decision matrix assigns an outcome (such as "walk, no soccer" in our example). The decision matrix for Peter's decision is as follows:

	No soccer players	Soccer players
Go to soccer ground	Walk, no soccer	Walk, soccer
Stay home	No walk, no soccer	No walk, no soccer

Such a matrix provides a clear presentation of the decision, but it does not contain all the information that the decision-maker needs to make the decision. The most important missing information concerns how the outcomes are valued.

Value Representation

When we make decisions, or choose among options, we try to obtain as good an outcome as possible, according to some standard of what is good or bad. The choice of a value-standard for decision-making is usually not considered to fall within the subject matter of decision theory. Instead, decision theory assumes that such a standard is available from other sources, perhaps from moral philosophy.

There are two major ways to express our evaluations of outcomes. One of these is *relational representation*. It is expressed in terms of the three comparative value notions, namely "better than" (*strong preference*, >), "equal in value to" (*indifference*, ≡), and "at least as good as" (*weak preference*, ≥). These three notions are interconnected according to the following two rules:

1. A is better than B if and only if A is at least as good as B but B is not at least as good as A. (A > B if and only if A ≥ B and not B ≥ A.)
2. A is equally good as B if and only if A is at least as good as B and B is at least as good as A. (A ≡ B if and only if A ≥ B and B ≥ A.)

The other major method to express our evaluations of outcomes is *numerical representation*. It consists in assigning numbers to the possible outcomes; such that an outcome has a higher number than another if and only if it is preferred to it. In an economic context, willingness to pay is often used as a measure of value. If a person is prepared to pay, say $500 for a certain used car and $250 for another, then these sums can be used to express her (economic) valuation of the two vehicles.

However, not all values are monetary. According to some moral theorists, all values can instead be reduced to one unit of measurement, *utility*. This entity may or may not be identified with units of human happiness. According to utilitarian moral theory, decision-makers should, at least in principle, always (try to) maximize total utility.

Decision theorists often use numerical values as abstract tools in the analysis of decisions. These values may be taken to represent utilities, but only in a rather abstract sense since they are not based on any method to measure utilities.

Once we have a numerical representation of value, we can replace the verbal descriptions of outcomes in a decision matrix by these values. In our example, suppose that Peter likes to play soccer but does not like walking to the soccer ground and back home. Then his utilities may be representable as follows:

	No soccer players	Soccer players
Go to soccer ground	0	10
Stay home	3	3

Mainstream decision theory is almost exclusively devoted to problems that can be expressed in matrices of this type, *utility matrices* (payoff matrices).

Probability or Uncertainty

Decisions are often categorized according to how much the decision-maker knows beforehand about what state of nature will in fact take place. In an extreme case, the decision-maker knows for sure which state of nature will obtain. If, in the above example, Peter knows with certainty that there are players at the soccer ground, then this makes his decision very simple. The same applies if he knows that there are no players. Cases like these, when only one state of nature needs to be taken into account, are called *decision-making under certainty*.

Non-certainty is usually divided into two categories, called *risk* and *uncertainty*. A decision is made under risk if it is based on exact probabilities that have been assigned to the relevant states of nature; otherwise it is made under uncertainty. Decisions at the roulette table are good examples of decisions under risk since the probabilities are known (although some players do not pay much attention to them). A decision whether to marry is a good example of a decision under uncertainty. There is no way to determine the probability that a marriage will be successful, and presumably few prospective brides or grooms would wish to base their decision on precise probability estimates of marital success or failure.

In some cases, we do not even have a full list of the relevant states of affairs. Hence, decisions on the introduction of a new technology have to be made without full knowledge of the possible future social states in which the new technology will be used. Such cases are referred to as decision-making under *great uncertainty*, or *ignorance*. This adds up to the following scale of knowledge situations in decision problems:

Certainty	It is known what states of nature will occur
Risk	The states of nature and their probabilities are known
Uncertainty	The states of nature are known but not their probabilities
Great uncertainty, ignorance	Not even the states of nature are known

The probabilities referred to in decision theory may be either objective or subjective. In some applications, reliable estimates of probabilities can be based on empirically known frequencies. As one example, death rates at high exposures to asbestos are known from epidemiological studies. In most cases, however, the basis for probability estimates is much less secure. This applies for instance to failures of a new, as yet untried technological device. In such cases we have to resort to subjective estimates of the objective probabilities. Some decision theorists deny the existence of true objective probabilities and regard all probabilities as expressions of degrees of belief, which are of course strictly subjective.

In cases when exact probabilities are not known, uncertainty can be expressed with various, more complex measures.

Binary measures: The probability values are divided into two groups, possible and impossible values (or attention-worthy and negligible values). Usually, the former form a single interval. Then the uncertainty can be expressed in terms of an interval, for instance: "The probability of a nuclear war in the next thirty years is between 10 and 25 per cent."

Multivalued measures: A numerical measure is used to distribute plausibility over the possible probability values. This measure may (but need not) be a (second-order) probability measure. Then, instead of just saying that the probability is between 10% and 25%, we can say that there is a 5% probability that the probability is between 17% and 18%, a 4% probability that it is between 18% and 19%, etc.

Robustness measures: The more certain we are about a probability, the less willing we are to change our estimate of it. Therefore, willingness to change one's estimate of a probability when new information arrives can be used as a measure of uncertainty.

Decision Rules

Decision theorists have developed a series of decision rules, intended to ensure that decisions are made in a systematic and rational way.

The Maximin Rule

Among the decision rules that are applicable without numerical information, the *maximin* rule is probably the most important one. For each alternative, we define its *security level* as the worst possible outcome with that alternative. The maximin rule urges us to choose the alternative that has the highest security level. In other words, we *maxi*mize the *min*imal outcome.

The maximin rule has often, and quite accurately, been described as a cautious rule. It has also been described as pessimistic, but that is an unfortunate terminology, since caution and pessimism are quite independent of each other.

As an example of the maximin rule, consider the following variant of the soccer example from above:

	No soccer players	Soccer players
Go to soccer ground	Walk, no soccer	Walk, soccer
Stay home	No walk, no soccer	No walk, no soccer

The preferences are:

Walk, soccer
is better than
No walk, no soccer
is better than
Walk, no soccer

The security level of Stay home is "no walk, no soccer" whereas that of Go to soccer ground is "walk, no soccer". Since the former is better than the latter, in order to maximize the security level, Peter would have to stay at home. Consequently, this is what the maximin rule recommends him to do.

Even though the maximin rule can be applied to relational value information as above, it is easier to apply if the value information is presented in numerical form. Again, consider the following utility matrix:

	No soccer players	Soccer players
Go to soccer ground	0	10
Stay home	3	3

Here, the security level of Stay home is 3 whereas that of Go to soccer ground is 0. Since 3 is larger than 0, the maximin rule recommends Peter to stay at home, just as in the relational presentation of the same example.

The Maximax Rule

The best level that we can at all obtain if we choose a certain alternative is called its *hope level*. According to the *maxi-max* rule, we should choose the alternative whose hope level (best possible outcome) is best. Just like the maxi-min rule, the maxi-max rule can be applied even if we only have relational (non-numerical) value information. Consider again the soccer example. The hope level of

Stay home is "no walk, no soccer," and that of Go to soccer ground is "walk, soccer" that Peter values higher. Hence, the maximax rule urges Peter to go to the soccer ground. Similarly, in the numerical representation, Stay home has the hope level 3 and Go to soccer ground has 10; hence again Peter is advised to go to the soccer ground.

The maximax rule has seldom been promoted. Contrary to the maximin rule, it is often conceived as irrational or as an expression of wishful thinking. It is indeed hardly commendable as a general rule for decision-making. However, in certain subareas of life, taking chances may be beneficial, and in such areas behavior corresponding to the maximax rule may not be irrational. Life would probably be much duller unless at least some of us were maximaxers on at least some occasions.

The Cautiousness Index

There is an obvious need for a decision criterion that does not force us into the extreme cautiousness of the maximin rule or the extreme incautiousness of the maximax rule. A middle road is available, but only if we have access to numerical information. We can then calculate a weighted average between the security level and the hope level, and use this weighted average to rank the alternatives. Let us again consider the numerical presentation of the soccer example:

	No soccer players	Soccer players
Go to soccer ground	0	10
Stay home	3	3

For each alternative A, let *min*(A) be its security level and *max*(A) its hope level.

In our example, *min*(Go to soccer ground) = 0 and *max*(Go to soccer ground) = 10. If we choose to assign equal weight to the security level and the hope level, then the weighted value of Go to soccer ground is $0.5 \times 0 + 0.5 \times 10 = 5$. Since *min*(Stay home) = *max*(Stay home) = 3, the weighted average value of Stay home is 3. Hence, with these relative weights, Peter is recommended to go to the soccer ground. More generally speaking, each alternative A is assigned a value according to the following formula:

$$\alpha \times min\,(\mathrm{A}) + (1 - \alpha) \times max\,(\mathrm{A})$$

If $\alpha = 1$, then this rule reduces to the maximin criterion and if $\alpha = 0$, then it reduces to the maximax criterion. The index α is often called the *Hurwicz α index*, after economist Leonid Hurwicz who proposed it in 1950. It is also often called the *optimism-pessimism index*, but the latter terminology should be avoided since the index represents the degree of (un)cautiousness rather than that of optimism. It can more appropriately be called the *cautiousness index*.

Minimax Regret

Utility information also allows for another decision criterion that puts focus on how an outcome differs from other outcomes that might have been obtained under the same state of affairs, if the decision-maker had chosen another alternative. In our example, if Peter stays home and there are players at the soccer ground, then he has made a loss that may give rise to considerable regret. If he goes to the soccer ground and there is no one there to play with him, then he has also made a loss, but a smaller one. The decision rule based on these considerations is usually called the *minimax regret* criterion. It also has other names, such as *minimax risk*, *minimax loss*, and *minimax*.

In this decision rule the degree of regret is measured as the difference between the utility obtained and the highest utility level that could have been obtained (in the same state of the world) if another alternative had been chosen. A *regret matrix* can quite easily be derived from a utility matrix: Replace each entry by the number obtained by subtracting it from the highest utility in its column. In our example, the regret matrix will be as follows:

	No soccer players	Soccer players
Go to soccer ground	3	0
Stay home	0	7

The minimax regret criterion advices the decision-maker to choose the option with the lowest maximal regret (to *min*imize *max*imal *regret*). In this case it recommends Peter to go to the soccer ground.

Just like the maximin rule, the minimax regret rule can be described as cautious, but they apply cautiousness to different aspects of the decision (the value of the actual outcome respectively its regrettableness). As this example shows, they do not always yield the same recommendation.

Expected Utility

None of the above decision rules requires or makes use of probabilistic information. When probabilities are available, the dominating approach is to maximize expected utility (EU). Then to each alternative is assigned a weighted average of its utility values under the different states of nature, with the probabilities of these states used as weights.

In the above example, suppose that based on previous experience Peter believes the probability to be 0.4 that there are players at the soccer ground. We can enter the probabilistic information into the column headings of the utility matrix as follows:

	No soccer players Probability 0.6	Soccer players Probability 0.4
Go to soccer ground	0	10
Stay home	3	3

The expected (probability-weighted) utility of going to the soccer ground is $0.6 \times 0 + 0.4 \times 10 = 4$, and that of staying at home is $0.6 \times 3 + 0.4 \times 3 = 3$. If Peter wants to maximize expected utility then he should, in this case, go to the soccer ground. Obviously, the recommendation would be different with other probabilities.

For a general formula representing expected utility, let there be n outcomes, to each of which is associated a utility and a probability. The outcomes are numbered, so that the first outcome has utility u_1 and probability p_1, the second has utility u_2 and probability p_2, etc. Then the expected utility is defined as follows:

$$p_1 \times u_1 + p_2 \times u_2 + \ldots + p_n \times u_n$$

Expected utility maximization based on subjective probabilities is commonly called *Bayesian decision theory*, or Bayesianism. (The name derives from Thomas Bayes, 1702–1761, who provided much of the mathematical foundations for probability theory). According to Bayesianism, a rational decision-maker should have a complete set of probabilistic beliefs (or at least behave as if she had one) and all her decisions should take the form of choosing the option with the highest expected utility.

Two Paradoxes of Decision Theory

Much of the modern discussion on decision theory has been driven by the presentation of paradoxes, i.e., situations in which decision-making criteria that seem to epitomize rationality nevertheless give rise to decisions that are contrary to most people's intuitions. The following two decision paradoxes serve to exemplify the kinds of philosophical problems that are discussed in the decision-theoretical research literature.

Ellsberg's Paradox

Daniel Ellsberg has presented the following decision problem: We have an urn that contains 30 red ball and 60 balls that are either black or yellow. The distribution between the latter two colors is unknown. A ball is going to be drawn at random from the urn. Before that is done you are offered bets by two persons.

Anne offers you to bet either on red or on black. If you bet on red, then you will receive € 100 if the drawn ball is red and nothing if it is either black or yellow. Similarly, if you bet on black, then you will get € 100 if the ball is black, and nothing if it is red or yellow.

Betty offers you to bet either on red-or-yellow or on black-or-yellow. If you bet on red-or-yellow, then you will get € 100 if the drawn ball is either red or yellow, but nothing if it is black. If you bet on black-or-yellow, then you will get € 100 if the drawn ball is either black or yellow, but nothing if it is red.

Most people, it turns out, prefer betting red to betting black, but they prefer betting black-or-yellow to betting red-or-yellow. It is fairly easy to show that this pattern is at variance with expected utility maximization, i.e., there is no way to assign utilities that would make this pattern compatible with the maximization of expected utility. Ellsberg's own conclusion was that decision-making must take into account factors not covered by probabilities and utilities, in particular the degree of uncertainty of the various probability estimates.

Another problem with this pattern is that it violates the *sure-thing principle* that is a much acclaimed rationality criterion for decisions. To introduce the principle, let A and B be two alternatives, and let S be a state of nature such that the outcome of A in S is the same as that of B. In other words, the outcome in case of S is a "sure thing," not influenced by the choice between A and B. The sure-thing principle says that if the "sure thing" (i.e., the common outcome in case of S) is changed, but nothing else is changed, then the choice between A and B is not affected.

As an example, suppose that a whimsical host wants to choose a dessert by tossing a coin. You are invited to choose between alternatives A and B. In alternative A, you will have fruit in case of heads and nothing in case of tails. In alternative B you will have pie in case of heads and nothing in case of tails. The decision matrix is as follows:

	Heads	Tails
A	Fruit	Nothing
B	Pie	Nothing

When you have made up your mind and announced which of the two alternatives you prefer, the whimsical host suddenly remembers that he has some ice cream, and changes the options so that the decision matrix is now as follows:

	Heads	Tails
A	Fruit	Ice cream
B	Pie	Ice cream

Since only a "sure thing" (an outcome that is common to the two alternatives) has changed between the two decision problems, the sure-thing principle demands that you do not change your choice between *A* and *B* when the decision problem is revised in this fashion. If, for instance, you chose alternative *A* in the first decision problem, then you are bound to do so in the second problem as well.

In this example, the sure-thing principle appears rational enough, and it would seem natural to endorse it as a general principle for decision-making. Ellsberg's paradox shows that is not quite as self-evident as it may seem to be at first sight.

Newcomb's Paradox

The following paradox was proposed by the physicist William Newcomb: In front of you are two boxes. One of them is transparent, and you can see that it contains $1,000. The other is covered, so that you cannot see its contents. It contains either $1,000,000 or nothing. You have two options to choose between. One is to take both boxes, and the other is to take only the covered box. A predictor who has infallible (or almost infallible) knowledge about your psyche has put the million in the covered box if he predicted that you will only take that box. Otherwise, it is empty.

Let us apply maximized expected utility to the problem. If you decide to take both boxes, then the predictor has almost certainly foreseen this and put nothing in the covered box. Your gain is $1,000. If, on the other hand, you decide to take only one box, then the predictor has foreseen this and put the million in the box, so that your gain is $1,000,000. In other words, maximization of expected utility urges you to take only the covered box.

There is, however, another plausible approach to the problem that leads to a different conclusion. If the predictor has put nothing in the covered box, then it is better to take both boxes than to take only one, since you will gain $1,000 instead of nothing. If he has put the million in the box, then too it is better to take both boxes, since you will gain $1,001,000 instead of $1,000,000. Thus, taking both boxes is better in all states of nature. (It is a *dominating* option.) It seems to follow that you should take both boxes, contrary to the rule of maximizing expected utilities.

The two-box strategy in Newcomb's problem maximizes the "real gain" of having chosen an option, whereas the one-box strategy maximizes the "news value" of having chosen an option. The very fact that a certain decision has been made in a certain way changes the probabilities that have to be taken into account in that decision.

In *causal decision theory*, expected utility calculations are modified so that they refer to real value rather than news value. This is done by replacing standard probabilities by some formal means for evaluating the causal implications of the different options. Since there are several competing philosophical views of causality, there are also several formulations of causal decision theory.

About the Author

Sven Ove Hansson is professor in philosophy and Head of the division of Philosophy, Royal Institute of Technology, Stockholm. He is Editor-in-Chief of *Theoria*. His research areas include value theory, decision theory, epistemology, belief dynamics and the philosophy of probability. He is the author of well over 200 articles in refereed journals. His books include *A Textbook of Belief Dynamics. Theory Change and Database Updating* (Kluwer 1999) and *The Structures of Values and Norms* (CUP 2001). He is a member of the Royal Swedish Academy of Engineering Sciences (IVA).

Cross References

►Bayesian Statistics
►Decision Theory: An Introduction
►Imprecise Probability
►Loss Function
►Multicriteria Decision Analysis
►Multiple Statistical Decision Theory
►Philosophical Foundations of Statistics
►Statistical Inference for Quantum Systems
►Statistics and Gambling

References and Further Reading

Gärdenfors P, Sahlin NE (eds) (1988) Decision, probability, and utility. Cambridge University Press, Cambridge

Hansson SO (2005) Decision theory. A brief introduction. http://www.infra.kth.se/~soh/decisiontheory.pdf

Luce RD, Raiffa H (1957) Games and decisions: introduction and critical survey. Wiley, New York

Resnik MD (1987) Choices – an introduction to decision theory. University of Minnesota Press, Minneapolis