6 Simple Decisions

This chapter introduces the notion of *simple decisions*, where we make a single decision under uncertainty.[1] We will study the problem of decision making from the perspective of *utility theory*, which involves modeling the preferences of an agent as a real-valued function over uncertain outcomes.[2] This chapter begins by discussing how a small set of constraints on rational preferences can lead to the existence of a utility function. This utility function can be inferred from a sequence of preference queries. We then introduce the maximum expected utility principle as a definition of rationality, a central concept in *decision theory* that will be used as a driving principle for decision making in this book.[3] We show how decision problems can be represented as decision networks and show an algorithm for solving for an optimal decision. The concept of value of information is introduced, which measures the utility gained through observing additional variables. The chapter concludes with a brief discussion of how human decision making is not always consistent with the maximum expected utility principle.

[1] Simple decisions are simple compared to sequential problems, which are the focus of the rest of the book. Simple decisions are not necessarily simple to solve, though.

[2] Schoemaker provides an overview of the development of utility theory. See P. J. H. Schoemaker, "The Expected Utility Model: Its Variants, Purposes, Evidence and Limitations," *Journal of Economic Literature*, vol. 20, no. 2, pp. 529–563, 1982. Fishburn surveys the field. See P. C. Fishburn, "Utility Theory," *Management Science*, vol. 14, no. 5, pp. 335–378, 1968.

[3] A survey of the field of decision theory is provided by M. Peterson, *An Introduction to Decision Theory*. Cambridge University Press, 2009.

6.1 Constraints on Rational Preferences

We began our discussion on uncertainty in chapter 2 by identifying the need to compare our degree of belief in different statements. This chapter requires the ability to compare the degree of desirability of two different outcomes. We state our preferences using the following operators:

- $A \succ B$ if we prefer A over B.
- $A \sim B$ if we are indifferent between A and B.
- $A \succeq B$ if we prefer A over B or are indifferent.

Just as beliefs can be subjective, so can preferences.

In addition to comparing events, our preference operators can be used to compare preferences over uncertain outcomes. A *lottery* is a set of probabilities associated with a set of outcomes. For example, if $S_{1:n}$ is a set of outcomes and $p_{1:n}$ are their associated probabilities, then the lottery involving these outcomes and probabilities is written as

$$[S_1 : p_1; \ \ldots; \ S_n : p_n] \tag{6.1}$$

The existence of a real-valued measure of utility emerges from a set of assumptions about preferences.[4] From this utility function, it is possible to define what it means to make rational decisions under uncertainty. Just as we imposed a set of constraints on beliefs, we will impose some constraints on preferences:[5]

- *Completeness.* Exactly one of the following holds: $A \succ B$, $B \succ A$, or $A \sim B$.
- *Transitivity.* If $A \succeq B$ and $B \succeq C$, then $A \succeq C$.
- *Continuity.* If $A \succeq C \succeq B$, then there exists a probability p such that $[A : p; B : 1 - p] \sim C$.
- *Independence.* If $A \succ B$, then for any C and probability p, $[A : p; C : 1 - p] \succeq [B : p; C : 1 - p]$.

These are constraints on *rational preferences*. They say nothing about the preferences of actual human beings; in fact, there is strong evidence that humans are not always rational (a point discussed further in section 6.7). Our objective in this book is to understand rational decision making from a computational perspective so that we can build useful systems. The possible extension of this theory to understanding human decision making is only of secondary interest.

[4] The theory of expected utility was introduced by the Swiss mathematician and physicist Daniel Bernoulli (1700–1782) in 1738. See D. Bernoulli, "Exposition of a New Theory on the Measurement of Risk," *Econometrica*, vol. 22, no. 1, pp. 23–36, 1954.

[5] These constraints are sometimes called the *von Neumann–Morgenstern axioms*, named after the Hungarian-American mathematician and physicist John von Neumann (1903–1957) and the Austrian-American economist Oskar Morgenstern (1902–1977). They formulated a variation of these axioms. See J. von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior*. Princeton University Press, 1944. Critiques of these axioms are discussed by P. Anand, "Are the Preference Axioms Really Rational?" *Theory and Decision*, vol. 23, no. 2, pp. 189–214, 1987.

6.2 *Utility Functions*

Just as constraints on the comparison of the plausibility of different statements lead to the existence of a real-valued probability measure, constraints on rational preferences lead to the existence of a real-valued *utility* measure. It follows from our constraints on rational preferences that there exists a real-valued utility function U such that

- $U(A) > U(B)$ if and only if $A \succ B$, and
- $U(A) = U(B)$ if and only if $A \sim B$.

The utility function is unique up to a *positive affine transformation*. In other words, for any constants $m > 0$ and b, $U'(S) = mU(S) + b$ if and only if the preferences induced by U' are the same as U. Utilities are like temperatures: you can compare temperatures using Kelvin, Celsius, or Fahrenheit, all of which are affine transformations of each other.

It follows from the constraints on rational preferences that the utility of a lottery is given by

$$U([S_1 : p_1; \; \ldots; \; S_n : p_n]) = \sum_{i=1}^{n} p_i U(S_i) \tag{6.2}$$

Example 6.1 applies this equation to compute the utility of outcomes involving a collision avoidance system.

Suppose that we are building a collision avoidance system. The outcome of an encounter of an aircraft is defined by whether the system alerts (A) and whether a collision occurs (C). Because A and C are binary, there are four possible outcomes. So long as our preferences are rational, we can write our utility function over the space of possible lotteries in terms of four parameters: $U(a^0, c^0)$, $U(a^1, c^0)$, $U(a^0, c^1)$, and $U(a^1, c^1)$. For example,

$$U([a^0, c^0 : 0.5; \quad a^1, c^0 : 0.3; \quad a^0, c^1 : 0.1; \quad a^1, c^1 : 0.1])$$

is equal to

$$0.5U(a^0, c^0) + 0.3U(a^1, c^0) + 0.1U(a^0, c^1) + 0.1U(a^1, c^1)$$

Example 6.1. A lottery involving the outcomes of a collision avoidance system.

If the utility function is bounded, then we can define a *normalized utility function*, where the best possible outcome is assigned utility 1 and the worst possible outcome is assigned utility 0. The utility of each of the other outcomes is scaled and translated as necessary.

6.3 *Utility Elicitation*

In building a decision-making or decision support system, it is often helpful to infer the utility function from a human or a group of humans. This approach is called *utility elicitation* or *preference elicitation*.[6] One way to go about doing this is to fix the utility of the worst outcome $\underline{S}$ to 0 and the best outcome $\overline{S}$ to 1. So long as the utilities of the outcomes are bounded, we can translate and scale the utilities without altering our preferences. If we want to determine the utility of outcome S, then we determine probability p such that $S \sim [\overline{S} : p; \underline{S} : 1 - p]$. It then follows that $U(S) = p$. Example 6.2 applies this process to determine the utility function associated with a collision avoidance problem.

[6] A variety of methods for utility elicitation are surveyed by P. H. Farquhar, "Utility Assessment Methods," *Management Science*, vol. 30, no. 11, pp. 1283–1300, 1984.

In our collision avoidance example, the best possible event is to not alert and not have a collision, and so we set $U(a^0, c^0) = 1$. The worst possible event is to alert and have a collision, and so we set $U(a^1, c^1) = 0$. We define the lottery $L(p)$ to be $[a^0, c^0 : p;\ \ a^1, c^1 : 1 - p]$. To determine $U(a^1, c^0)$, we must find p such that $(a^1, c^0) \sim L(p)$. Similarly, to determine $U(a^0, c^1)$, we find p such that $(a^0, c^1) \sim L(p)$.

Example 6.2. Utility elicitation applied to collision avoidance.

It may be tempting to use monetary values to infer utility functions. For example, if we are building a decision support system for managing wildfires, it may be tempting to define a utility function in terms of the monetary cost incurred by property damage and the monetary cost for deploying fire suppression resources. However, it is well known in economics that the utility of wealth, in general, is not linear.[7] If there were a linear relationship between utility and wealth, then decisions should be made in terms of maximizing expected monetary value. Someone who tries to maximize expected monetary value would have no use for insurance because the expected monetary values of insurance policies are generally negative.

[7] H. Markowitz, "The Utility of Wealth," *Journal of Political Economy*, vol. 60, no. 2, pp. 151–158, 1952.

Instead of trying to maximize expected wealth, we generally want to maximize the expected utility of wealth. Of course, different people have different utility functions. Figure 6.1 shows an example of a utility function. For small amounts of wealth, the curve is roughly linear, where \$100 is about twice as good at \$50. For larger amounts of wealth, however, the curve tends to flatten out; after all,

\$1000 is worth less to a billionaire than it is to the average person. The flattening of the curve is sometimes referred to as *diminishing marginal utility*.

When discussing monetary utility functions, the three terms listed here are often used. To illustrate this, assume that A represents being given \$50 and B represents a 50 % chance of winning \$100.

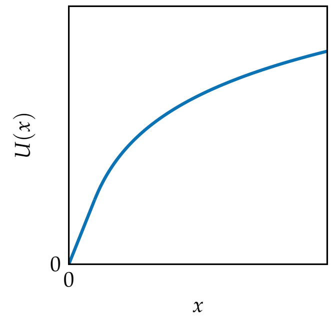


Figure 6.1. The utility of wealth x is often modeled as linear for small values and then concave for larger values, exhibiting risk aversion.

- *Risk neutral*. The utility function is linear. There is no preference between \$50 and the 50 % chance of winning \$100 ($A \sim B$).
- *Risk seeking*. The utility function is convex. There is a preference for the 50 % chance of winning \$100 ($A \prec B$).
- *Risk averse*. The utility function is concave. There is a preference for the \$50 ($A \succ B$).

There are several common functional forms for modeling risk aversion of scalar quantities,[8] such as wealth or the availability of hospital beds. One is *quadratic utility*:

$$U(x) = \lambda x - x^2 \tag{6.3}$$

where the parameter $\lambda > 0$ controls the risk aversion. Since we generally want this utility function to be monotonically increasing when modeling the utility of quantities like wealth, we would cap this function at $x = \lambda/2$. After that point, the utility starts decreasing. Another simple form is *exponential utility*:

$$U(x) = 1 - e^{-\lambda x} \tag{6.4}$$

with $\lambda > 0$. Although it has a convenient mathematical form, it is generally not viewed as a plausible model of the utility of wealth. An alternative is the *power utility*:

$$U(x) = \frac{x^{1-\lambda} - 1}{1 - \lambda} \tag{6.5}$$

with $\lambda \geq 0$ and $\lambda \neq 1$. The *logarithmic utility*

$$U(x) = \log x \tag{6.6}$$

with $x > 0$ can be viewed as a special case of the power utility where $\lambda \to 1$. Figure 6.2 shows a plot of the power utility function with the logarithmic utility as a special case.

[8] These functional forms have been well studied within economics and finance. J. E. Ingersoll, *Theory of Financial Decision Making*. Rowman and Littlefield Publishers, 1987.

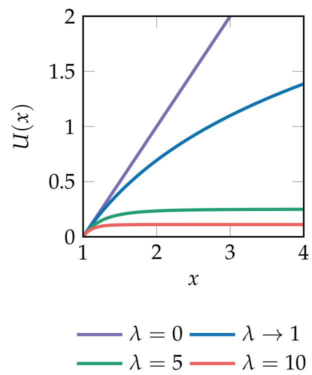


Figure 6.2. Power utility functions.

6.4 Maximum Expected Utility Principle

We are interested in the problem of making rational decisions with imperfect knowledge of the state of the world. Suppose that we have a probabilistic model $P(s' \mid o, a)$, which represents the probability that the state of the world becomes s', given that we observe o and take action a. We have a utility function $U(s')$ that encodes our preferences over the space of outcomes. Our *expected utility* of taking action a, given observation o, is given by

$$EU(a \mid o) = \sum_{s'} P(s' \mid a, o) U(s') \tag{6.7}$$

The *principle of maximum expected utility* says that a rational agent should choose the action that maximizes expected utility:

$$a^* = \arg\max_a EU(a \mid o) \tag{6.8}$$

Because we are interested in building rational agents, equation (6.8) plays a central role in this book.[9] Example 6.3 applies this principle to a simple decision problem.

[9] The importance of the maximum expected utility principle to the field of artificial intelligence is discussed by S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 4th ed. Pearson, 2021.

6.5 Decision Networks

A *decision network*, sometimes called an *influence diagram*, is a generalization of a Bayesian network to include action and utility nodes so that we may compactly represent the probability and utility models defining a decision problem.[10] The state, action, and observation spaces in the previous section may be factored, and the structure of a decision network captures the relationships between the various components.

[10] An extensive discussion of decision networks can be found in F. V. Jensen and T. D. Nielsen, *Bayesian Networks and Decision Graphs*, 2nd ed. Springer, 2007.

Decision networks are composed of three types of nodes:

- A *chance node* corresponds to a random variable (indicated by a circle).
- An *action node* corresponds to a decision variable (indicated by a square).
- A *utility node* corresponds to a utility variable (indicated by a diamond) and cannot have children.

Example 6.3. Applying the principle of maximum expected utility to the simple decision of whether to bring an umbrella.

Suppose that we are trying to decide whether to bring an umbrella on our vacation given the weather forecast for our destination. We observe the forecast o, which may be either rain or sun. Our action a is either to bring our umbrella or leave our umbrella. The resulting state s' is a combination of whether we brought our umbrella and whether there is sun or rain at our destination. Our probabilistic model is as follows:

o	a	s'	$P(s' \mid a, o)$
forecast rain	bring umbrella	rain with umbrella	0.9
forecast rain	leave umbrella	rain without umbrella	0.9
forecast rain	bring umbrella	sun with umbrella	0.1
forecast rain	leave umbrella	sun without umbrella	0.1
forecast sun	bring umbrella	rain with umbrella	0.2
forecast sun	leave umbrella	rain without umbrella	0.2
forecast sun	bring umbrella	sun with umbrella	0.8
forecast sun	leave umbrella	sun without umbrella	0.8

As shown in the table, we assume that our forecast is imperfect; rain forecasts are right 90 % of the time and sun forecasts are right 80 % of the time. In addition, we assume that bringing an umbrella does not affect the weather, though some may question this assumption. The utility function is as follows:

s'	$U(s')$
rain with umbrella	-0.1
rain without umbrella	-1
sun with umbrella	0.9
sun without umbrella	1

We can compute the expected utility of bringing our umbrella if we forecast rain using equation (6.7):

$$EU(\text{bring umbrella} \mid \text{forecast rain}) = 0.9 \times -0.1 + 0.1 \times 0.9 = 0$$

Likewise, we can compute the expected utility of leaving our umbrella if we forecast rain using equation (6.7):

$$EU(\text{leave umbrella} \mid \text{forecast rain}) = 0.9 \times -1 + 0.1 \times 1 = -0.8$$

Hence, we will want to bring our umbrella.

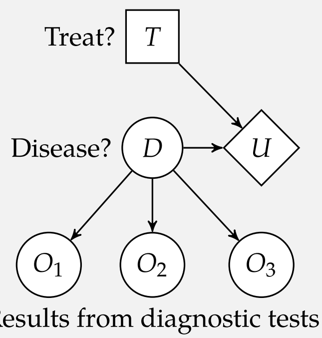
There are three kinds of directed edges:

- A *conditional edge* ends in a chance node and indicates that the uncertainty in that chance node is conditioned on the values of all its parents.
- An *informational edge* ends in an action node and indicates that the decision associated with that node is made with knowledge of the values of its parents. (These edges are often drawn with dashed lines and are sometimes omitted from diagrams for simplicity.)
- A *functional edge* ends in a utility node and indicates that the utility node is determined by the outcomes of its parents.

Like Bayesian networks, decision networks cannot have cycles. The utility associated with an action is equal to the sum of the values at all the utility nodes. Example 6.4 illustrates how a decision network can model the problem of whether to treat a disease, given the results of diagnostic tests.

We have a set of results from diagnostic tests that may indicate the presence of a particular disease. Given what is known about the tests, we need to decide whether to apply a treatment. The utility is a function of whether a treatment is applied and whether the disease is actually present. Conditional edges connect D to O_1, O_2, and O_3. Informational edges are not explicitly shown in the illustration, but they would connect the observations to T. Functional edges connect T and D to U.

T	D	$U(T, D)$
0	0	0
0	1	-10
1	0	-1
1	1	-1



Example 6.4. An example of a decision network used to model whether to treat a disease, given information from diagnostic tests.

Solving a simple problem (algorithm 6.1) requires iterating over all possible decision instantiations to find a decision that maximizes expected utility. For each instantiation, we evaluate the associated expected utility. We begin by instantiating the action nodes and observed chance nodes. We can then apply any inference algorithm to compute the posterior over the inputs to the utility function. The expected utility is the sum of the values at the utility nodes. Example 6.5 shows how this process can be applied to our running example.

```
struct SimpleProblem
    bn::BayesianNetwork
    chance_vars::Vector{Variable}
    decision_vars::Vector{Variable}
    utility_vars::Vector{Variable}
    utilities::Dict{Symbol, Vector{Float64}}
end

function solve(𝒫::SimpleProblem, evidence, M)
    query = [var.name for var in 𝒫.utility_vars]
    U(a) = sum(𝒫.utilities[uname][a[uname]] for uname in query)
    best = (a=nothing, u=-Inf)
    for assignment in assignments(𝒫.decision_vars)
        evidence = merge(evidence, assignment)
        φ = infer(M, 𝒫.bn, query, evidence)
        u = sum(p*U(a) for (a, p) in φ.table)
        if u > best.u
            best = (a=assignment, u=u)
        end
    end
    return best
end
```

Algorithm 6.1. A simple problem as a decision network. A decision network is a Bayesian network with chance, decision, and utility variables. Utility variables are treated as deterministic. Because variables in our Bayesian network take values from $1:r_i$, the utility variables are mapped to real values by the `utilities` field. For example, if we have a utility variable `:u1`, the `i`th utility associated with that variable is `utilities[:u1][i]`. The `solve` function takes as input the problem, evidence, and an inference method. It returns the best assignment to the decision variables and its associated expected utility.

A variety of methods have been developed to make evaluating decision networks more efficient.[11] One method involves removing action and chance nodes from decision networks if they have no children, as defined by conditional, informational, or functional edges. In example 6.5, we can remove O_2 and O_3 because they have no children. We cannot remove O_1 because we treated it as observed, indicating that there is an informational edge from O_1 to T (although it is not drawn explicitly).

6.6 *Value of Information*

We make decisions based on what we observe. In many applications, it is natural to want to quantify the *value of information*, which is how much observing additional variables is expected to increase our utility.[12] For example, in the disease treatment

[11] R. D. Shachter, "Evaluating Influence Diagrams," *Operations Research*, vol. 34, no. 6, pp. 871–882, 1986. R. D. Shachter, "Probabilistic Inference and Influence Diagrams," *Operations Research*, vol. 36, no. 4, pp. 589–604, 1988.

[12] R. A. Howard, "Information Value Theory," *IEEE Transactions on Systems Science and Cybernetics*, vol. 2, no. 1, pp. 22–26, 1966. Applications to decision networks can be found in: S. L. Dittmer and F. V. Jensen, "Myopic Value of Information in Influence Diagrams," in *Conference on Uncertainty in Artificial Intelligence (UAI)*, 1997. R. D. Shachter, "Efficient Value of Information Computation," in *Conference on Uncertainty in Artificial Intelligence (UAI)*, 1999.

Example 6.5. Decision network evaluation of the diagnostic test problem.

We can use equation (6.7) to compute the expected utility of treating a disease for the decision network in example 6.4. For now, we will assume that we have the result from only the first diagnostic test and it came back positive. If we wanted to make the knowledge of the first diagnostic test explicit in the diagram, then we would draw an informational edge from O_1 to T, and we would have

$$EU(t^1 \mid o_1^1) = \sum_{o_3}\sum_{o_2}\sum_{d} P(d, o_2, o_3 \mid t^1, o_1^1) U(t^1, d, o_1^1, o_2, o_3)$$

We can use the chain rule for Bayesian networks and the definition of conditional probability to compute $P(d, o_2, o_3 \mid t^1, o_1^1)$. Because the utility node depends only on whether the disease is present and whether we treat it, we can simplify $U(t^1, d, o_1^1, o_2, o_3)$ to $U(t^1, d)$. Hence,

$$EU(t^1 \mid o_1^1) = \sum_{d} P(d \mid t^1, o_1^1) U(t^1, d)$$

Any of the exact or approximate inference methods introduced in the previous chapter can be used to evaluate $P(d \mid t^1, o_1^1)$. To decide whether to apply a treatment, we compute $EU(t^1 \mid o_1^1)$ and $EU(t^0 \mid o_1^1)$ and make the decision that provides the highest expected utility.

application in example 6.5, we assumed that we have only observed o_1^1. Given the positive result from that one diagnostic test alone, we may decide against treatment. However, it may be beneficial to administer additional diagnostic tests to reduce the risk of not treating a disease that is really present.

In computing the value of information, we will use $EU^*(o)$ to denote the expected utility of an optimal action, given observation o. The value of information about variable O', given o, is

$$VOI(O' \mid o) = \left(\sum_{o'} P(o' \mid o) EU^*(o, o')\right) - EU^*(o) \tag{6.9}$$

In other words, the value of information about a variable is the increase in expected utility if that variable is observed. Algorithm 6.2 provides an implementation of this.

```
function value_of_information(𝒫, query, evidence, M)
    ϕ = infer(M, 𝒫.bn, query, evidence)
    voi = -solve(𝒫, evidence, M).u
    query_vars = filter(v->v.name ∈ query, 𝒫.chance_vars)
    for o′ in assignments(query_vars)
        oo′ = merge(evidence, o′)
        p = ϕ.table[o′]
        voi += p*solve(𝒫, oo′, M).u
    end
    return voi
end
```

Algorithm 6.2. A method for computing the value of information of a query `query` given observed chance variables and their values `evidence`. The method additionally takes a simple problem 𝒫 and an inference strategy `M`.

The value of information is never negative. The expected utility can increase only if additional observations can lead to different optimal decisions. If observing a new variable O' makes no difference in the choice of action, then $EU^*(o, o') = EU^*(o)$ for all o', in which case equation (6.9) evaluates to 0. For example, if the optimal decision is to treat the disease regardless of the outcome of the *diagnostic test*, then the value of observing the outcome of the test is 0.

The value of information only captures the increase in expected utility from making an observation. A cost may be associated with making a particular observation. Some diagnostic tests may be inexpensive, such as a temperature reading; other diagnostic tests are more costly and invasive, such as a lumbar puncture. The value of information obtained by a lumbar puncture may be much greater than that of a temperature reading, but the costs of the tests should be taken into consideration.

Value of information is an important and often-used metric for choosing what to observe. Sometimes the value of information metric is used to determine an appropriate sequence of observations. After each observation, the value of information is determined for the remaining unobserved variables. The unobserved variable with the greatest value of information is then selected for observation. If there are costs associated with making different observations, then these costs are subtracted from the value of information when determining which variable to observe. The process continues until it is no longer beneficial to observe any more variables. The optimal action is then chosen. This greedy selection of observations is only a heuristic; it may not represent the truly optimal sequence of observations. The optimal selection of observations can be determined by using the techniques for sequential decision making introduced in later chapters.

6.7 *Irrationality*

Decision theory is a *normative theory*, which is prescriptive, not a *descriptive theory*, which is predictive of human behavior. Human judgment and preference often do not follow the rules of rationality outlined in section 6.1.[13] Even human experts may have an inconsistent set of preferences, which can be problematic when designing a decision support system that attempts to maximize expected utility.

[13] Kahneman and Tversky provide a critique of expected utility theory and introduce an alternative model called *prospect theory*, which appears to be more consistent with human behavior. D. Kahneman and A. Tversky, "Prospect Theory: An Analysis of Decision Under Risk," *Econometrica*, vol. 47, no. 2, pp. 263–292, 1979.

Example 6.6 shows that certainty often exaggerates losses that are certain compared to losses that are merely probable. This *certainty effect* works with gains as well. A smaller gain that is certain is often preferred over a much greater gain that is only probable, in a way that the axioms of rationality are necessarily violated.

Example 6.7 demonstrates the *framing effect*, where people decide on options based on whether they are presented as a loss or as a gain. Many other cognitive biases can lead to deviations from what is prescribed by utility theory.[14] Special care must be given when trying to elicit utility functions from human experts to build decision support systems. Although the recommendations of the decision support system may be rational, they may not exactly reflect human preferences in certain situations.

[14] Several recent books discuss apparent human irrationality. D. Ariely, *Predictably Irrational: The Hidden Forces That Shape Our Decisions*. Harper, 2008. J. Lehrer, *How We Decide*. Houghton Mifflin, 2009.

Tversky and Kahneman studied the preferences of university students who answered questionnaires in a classroom setting. They presented students with questions dealing with the response to an epidemic. The students were to reveal their preference between the following two outcomes:

- A: 100 % chance of losing 75 lives
- B: 80 % chance of losing 100 lives

Most preferred B over A. From equation (6.2), we know

$$U(\text{lose } 75) < 0.8U(\text{lose } 100) \tag{6.10}$$

They were then asked to choose between the following two outcomes:

- C: 10 % chance of losing 75 lives
- D: 8 % chance of losing 100 lives

Most preferred C over D. Hence, $0.1U(\text{lose } 75) > 0.08U(\text{lose } 100)$. We multiply both sides by 10 and get

$$U(\text{lose } 75) > 0.8U(\text{lose } 100) \tag{6.11}$$

Of course, equations (6.10) and (6.11) result in a contradiction. We have made no assumption about the actual value of $U(\text{lose } 75)$ and $U(\text{lose } 100)$—we did not even assume that losing 100 lives was worse than losing 75 lives. Because equation (6.2) follows directly from the von Neumann–Morgenstern axioms given in section 6.1, there must be a violation of at least one of the axioms, even though many people who select B and C seem to find the axioms agreeable.

Example 6.6. An experiment demonstrating that certainty often exaggerates losses that are certain relative to losses that are merely probable. A. Tversky and D. Kahneman, "The Framing of Decisions and the Psychology of Choice," *Science*, vol. 211, no. 4481, pp. 453–458, 1981.

Tversky and Kahneman demonstrated the *framing effect* using a hypothetical scenario in which an epidemic is expected to kill 600 people. They presented students with the following two outcomes:

- E: 200 people will be saved.
- F: $1/3$ chance that 600 people will be saved and $2/3$ chance that no people will be saved.

The majority of students chose E over F. They then asked them to choose between the following:

- G: 400 people will die.
- H: $1/3$ chance that nobody will die and $2/3$ chance that 600 people will die.

The majority of students chose H over G, even though E is equivalent to G and F is equivalent to H. This inconsistency is due to how the question is framed.

Example 6.7. An experiment demonstrating the framing effect. A. Tversky and D. Kahneman, "The Framing of Decisions and the Psychology of Choice," *Science*, vol. 211, no. 4481, pp. 453–458, 1981.

6.8 Summary

- Rational decision making combines probability and utility theory.
- The existence of a utility function follows from constraints on rational preferences.
- A rational decision is one that maximizes expected utility.
- Decision problems can be modeled using decision networks, which are extensions of Bayesian networks that include actions and utilities.
- Solving a simple decision involves inference in Bayesian networks and is thus NP-hard.
- The value of information measures the gain in expected utility should a new variable be observed.
- Humans are not always rational.

6.9 Exercises

Exercise 6.1. Suppose that we have a utility function $U(s)$ with a finite maximum value $\overline{U}$ and a finite minimum value $\underline{U}$. What is the corresponding normalized utility function $\hat{U}(s)$ that preserves the same preferences?

Solution: A normalized utility function has a maximum value of 1 and a minimum value of 0. Preferences are preserved under affine transforms, so we determine the affine transform of $U(s)$ that matches the unit bounds. This transform is

$$\hat{U}(s) = \frac{U(s) - \underline{U}}{\overline{U} - \underline{U}} = \frac{1}{\overline{U} - \underline{U}}U(s) - \frac{\underline{U}}{\overline{U} - \underline{U}}$$

Exercise 6.2. If $A \succeq C \succeq B$ and the utilities of each outcome are $U(A) = 450$, $U(B) = -150$, and $U(C) = 60$, what is the lottery over A and B that will make us indifferent between the lottery and C?

Solution: A lottery over A and B is defined as $[A:p; B:1-p]$. To satisfy indifference between the lottery and C ($[A:p; B:1-p] \sim C$), we must have $U([A:p; B:1-p]) = U(C)$. Thus, we must compute p that satisfies the equality

$$\begin{aligned} U([A:p; B:1-p]) &= U(C) \\ pU(A) + (1-p)U(B) &= U(C) \\ p &= \frac{U(C) - U(B)}{U(A) - U(B)} \\ p &= \frac{60 - (-150)}{450 - (-150)} = 0.35 \end{aligned}$$

This implies that the lottery $[A:0.35; B:0.65]$ is equally as desired as C.

Exercise 6.3. Suppose that for a utility function U over three outcomes A, B, and C, that $U(A) = 5$, $U(B) = 20$, and $U(C) = 0$. We are given a choice between a lottery that gives us a 50 % probability of B and a 50 % probability of C and a lottery that guarantees A. Compute the preferred lottery and show that, under the positive affine transformation with $m = 2$ and $b = 30$, that we maintain a preference for the same lottery.

Solution: The first lottery is given by $[A:0.0; B:0.5; C:0.5]$, and the second lottery is given by $[A:1.0; B:0.0; C:0.0]$. The original utilities for each lottery are given by

$$\begin{aligned} U([A:0.0; B:0.5; C:0.5]) &= 0.0U(A) + 0.5U(B) + 0.5U(C) = 10 \\ U([A:1.0; B:0.0; C:0.0]) &= 1.0U(A) + 0.0U(B) + 0.0U(C) = 5 \end{aligned}$$

Thus, since $U([A:0.0; B:0.5; C:0.5]) > U([A:1.0; B:0.0; C:0.0])$, we prefer the first lottery. Under the positive affine transformation $m = 2$ and $b = 30$, our new utilities can be computed as $U' = 2U + 30$. The new utilities are then $U'(A) = 40$, $U'(B) = 70$, and $U'(C) = 30$. The new utilities for each lottery are

$$\begin{aligned} U'([A:0.0; B:0.5; C:0.5]) &= 0.0U'(A) + 0.5U'(B) + 0.5U'(C) = 50 \\ U'([A:1.0; B:0.0; C:0.0]) &= 1.0U'(A) + 0.0U'(B) + 0.0U'(C) = 40 \end{aligned}$$

Since $U'([A:0.0; B:0.5; C:0.5]) > U'([A:1.0; B:0.0; C:0.0])$, we maintain a preference for the first lottery.

Exercise 6.4. Prove that the power utility function in equation (6.5) is risk averse for all $x > 0$ and $\lambda > 0$ with $\lambda \neq 1$.

Solution: Risk aversion implies that the utility function is concave, which requires that the second derivative of the utility function is negative. The utility function and its derivatives are computed as follows:

$$U(x) = \frac{x^{1-\lambda} - 1}{1 - \lambda}$$
$$\frac{dU}{dx} = \frac{1}{x^{\lambda}}$$
$$\frac{d^2U}{dx^2} = \frac{-\lambda}{x^{\lambda+1}}$$

For $x > 0$ and $\lambda > 0$, $\lambda \neq 1$, $x^{\lambda+1}$ is a positive number raised to a positive exponent, which is guaranteed to be positive. Multiplying this by $-\lambda$ guarantees that the second derivative is negative. Thus, for all $x > 0$ and $\lambda > 0$, $\lambda \neq 1$, the power utility function is risk averse.

Exercise 6.5. Using the parameters given in example 6.3, compute the expected utility of bringing our umbrella if we forecast sun and the expected utility of leaving our umbrella behind if we forecast sun. What is the action that maximizes our expected utility, given that we forecast sun?

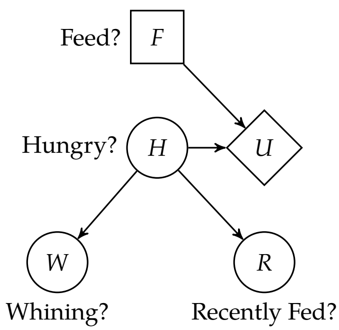
Solution:

$$EU(\text{bring umbrella} \mid \text{forecast sun}) = 0.2 \times -0.1 + 0.8 \times 0.9 = 0.7$$
$$EU(\text{leave umbrella} \mid \text{forecast sun}) = 0.2 \times -1.0 + 0.8 \times 1.0 = 0.6$$

The action that maximizes our expected utility if we forecast sun is to bring our umbrella!

Exercise 6.6. Suppose that we are trying to optimally decide whether or not to feed (F) our new puppy based on the likelihood that the puppy is hungry (H). We can observe whether the puppy is whining (W) and whether someone else has recently fed the puppy (R). The utilities of each combination of feeding and hunger and the decision network representation are provided here:

F	H	$U(F, H)$
0	0	0.0
0	1	−1.0
1	0	−0.5
1	1	−0.1



Given that $P(h^1 \mid w^1) = 0.78$, if we observe the puppy whining (w^1), what are the expected utilities of not feeding the puppy (f^0) and feeding the puppy (f^1)? What is the optimal action?

Solution: We start with the definition of expected utility and recognize that the utility depends only on H and F:

$$EU(f^0 \mid w^1) = \sum_h P(h \mid w^1)U(f^0, h)$$

Now, we can compute the expected utility of feeding the puppy given that it is whining and, in a similar fashion as before, the expected utility of not feeding the puppy given that it is whining:

$$EU(f^0 \mid w^1) = 0.22 \times 0.0 + 0.78 \times -1.0 = -0.78$$
$$EU(f^1 \mid w^1) = 0.22 \times -0.5 + 0.78 \times -0.1 = -0.188$$

Thus, the optimal action is to feed the puppy (f^1) since this maximizes our expected utility $EU^*(w^1) = -0.188$.

Exercise 6.7. Using the results from exercise 6.6, if $P(r^1 \mid w^1) = 0.2$, $P(h^1 \mid w^1, r^0) = 0.9$, and $P(h^1 \mid w^1, r^1) = 0.3$, what is the value of information of asking someone else if the puppy has recently been fed, given that we observe the puppy to be whining (w^1)?

Solution: We are interested in computing

$$\text{VOI}(R \mid w^1) = \left(\sum_r P(r \mid w^1)EU^*(w^1, r)\right) - EU^*(w^1)$$

We start by computing $EU(f \mid w^1, r)$ for all f and r. Following a similar derivation as in exercise 6.6, we have

$$EU(f^0 \mid w^1, r^0) = \sum_h P(h \mid w^1, r^0)U(f^0, h)$$

So, for each combination of F and R, we have the following expected utilities:

$$EU(f^0 \mid w^1, r^0) = \sum_h P(h \mid w^1, r^0)U(f^0, h) = 0.1 \times 0.0 + 0.9 \times -1.0 = -0.9$$
$$EU(f^1 \mid w^1, r^0) = \sum_h P(h \mid w^1, r^0)U(f^1, h) = 0.1 \times -0.5 + 0.9 \times -0.1 = -0.14$$
$$EU(f^0 \mid w^1, r^1) = \sum_h P(h \mid w^1, r^1)U(f^0, h) = 0.7 \times 0.0 + 0.3 \times -1.0 = -0.3$$
$$EU(f^1 \mid w^1, r^1) = \sum_h P(h \mid w^1, r^1)U(f^1, h) = 0.7 \times -0.5 + 0.3 \times -0.1 = -0.38$$

The optimal expected utilities are

$$EU^*(w^1, r^0) = -0.14$$
$$EU^*(w^1, r^1) = -0.3$$

Now, we can compute the value of information:

$$\text{VOI}(R \mid w^1) = 0.8(-0.14) + 0.2(-0.3) - (-0.188) = 0.016$$