Chapter 5

Utility and rational decision-making

The ability of making appropriate decisions in any given situation is clearly a necessary prerequisite for the survival of any animal and, indeed, even simple animals are generally able to take the correct action even in difficult situations, provided that they are operating in their natural environment. While the concept of **rational decision-making** has been studied in ethology (and, more recently, in robotics), the theory of rational decision-making was formalized within the framework of economics, particularly in the important work by von Neumann and Morgenstern [28]. Their work also remains one of the cornerstones of **game theory**.

The choices facing a decision-maker within the framework of economics, can often be illustrated by means of **lotteries**, at least in cases where the number of **consequences** (or **outcomes**) is finite. Intuitively, one may think that the expected payoff, i.e. the amount of money that is likely to be gained by participating in the lottery, may determine a person's inclination to do so. However, things are a bit more complicated than that. As a first example, consider a lottery in which, with probability $p_1 = 0.5$ one would gain \$3 (i.e. with the **consequence** $c_1 = +\$3$), say, and, with probability $p_2 = 1 - p_1$, one would have to pay \$2. (outcome c_2 =-\$2). Thus, the expected payoff from this bet would be

$$P = p_1 c_1 + p_2 c_2 = 0.5 \times 3 - 0.5 \times 2 = 0.5. \tag{5.1}$$

Thus, it is likely that most people would accept this bet since the expected payoff is larger than zero. However, consider now a lottery with the same probabilities p_1 and p_2, but with the consequences c_1 =\$300,000 and c_2 =-\$200,000. In this case, the expected payoff would be \$50,000, a considerable amount of money, yet most people would be disinclined to accept the bet, given the risk of losing \$200,000.

As a second example, consider a situation where a person must carry out a potentially lethal task in order to gain \$10,000. If the person's total wealth is \$0 it is possible that he would accept the task, regardless of the risk. On the other hand, if the person's total wealth is \$1,000,000, it would hardly be worth taking any risk for a measly additional \$10,000. Thus, clearly, the *amount* that can be gained is not the sole determinant, or even the most important one, when contemplating what action to take in the situations just described.

5.1 Utility

In order to underline further the fact that the expected payoff alone is not what determines one's inclination to accept a bet, consider a repeated lottery in which a fair coin (equal probability for heads and tails) is tossed repeatedly, and where the player receives 2^k dollars if the first head, say, occurs after k tosses of the coin. The probability p_k of this event occurring equals $(1/2)^k$. Thus, the expected payoff from playing this lottery would be

$$P = \sum p_k c_k = \frac{1}{2}2^1 + \frac{1}{4}2^2 + \ldots + \left(\frac{1}{2}\right)^k 2^k + \ldots, \tag{5.2}$$

which is infinite! Thus, *if* the expected payoff P was all that mattered, a player should be willing to pay any sum of money, however large, in order to participate in this lottery, since the expected payoff would be larger. Clearly, something is wrong here: expected payoff is not, by itself, sufficient to determine a player's behavior. The situation just described is called the **St. Petersburg paradox**, and was formulated by Bernoulli. He a proposed a way of resolving the paradox, by postulating that it is not the expected payoff, in itself, that determines a player's behavior, but rather his *perception* of the amount of money gained. Bernoulli postulated that the subjective value of \$ N equals $\log_{10} N$. Thus, if this is the case, the subjective value of the outcome of the lottery would be

$$P_s = \frac{1}{2}\log_{10} 2 + \frac{1}{4}\log_{10} 4 + \ldots + \left(\frac{1}{2}\right)^k \log_{10} 2^k + \ldots, \tag{5.3}$$

which is finite (and equal to 0.60206). The subjective value of a certain amount of money, set arbitrarily to the logarithm of the amount by Bernoulli, is a special case of the concept of **utility**, which can be used for weighing different situations against each other and, thus, to decide which action to take.

In fact, it has been *proven* (rigorously) by von Neumann and Morgenstern [28] that, given certain assumptions that will be listed below, there exists a **utility function** which maps members c_i of the set of outcomes to a numerical value $u(c_i)$, called the utility of c_i, which has the following properties:

1. $u(c_1) > u(c_2)$ if and only if the person prefers[1] c_1 to c_2,

[1] If (and only if) the person is indifferent between c_1 and c_2, then $u(c_1) = u(c_2)$.

2. u is affine, i.e.

$$u(pc_1 + (1-p)c_2) = pu(c_1) + (1-p)u(c_2), \tag{5.4}$$

for any value of $p_1 \in [0, 1]$.

Furthermore, as shown by von Neumann and Morgenstern, u is unique up to a positive linear transformation, i.e. if a function v also describes a person's preferences, then $v = \alpha_1 u + \alpha_2$, where $\alpha_1 > 0$.

Clearly, there is no unique set of preferences, valid for all persons: one person may prefer a consequence c_1 to another consequence c_2, whereas another person's preferences may be exactly the opposite. Thus, utility tells us nothing about a person's preferences. However, it *does* tell us that, given that the preferences can be stated in a consistent way (see below), there exists a function u which can serve as a **common currency** in decision-making, i.e. when weighing different options against each other.

As previously mentioned, the existence of a utility function with the properties listed above depends on certain axioms, namely

Axiom 1 (Ordering) Given two outcomes c_1 and c_2 an individual can decide, and remain consistent, concerning his preferences, i.e. whether he prefers c_1 to c_2 (denoted $c_1 > c_2$), c_2 to c_1, or is indifferent (denoted $c_1 \sim c_2$).

Axiom 2 (Transitivity) If $c_1 \geq c_2$ and $c_2 \geq c_3$ then $c_1 \geq c_3$.

Axiom 3 (The Archimedean axiom) If $c_1 > c_2 > c_3$, there exists a $p \in [0, 1]$ such that $pc_1 + (1-p)c_3 > c_2$ and a $q \in [0, 1]$ such that $c_2 > qc_1 + (1-q)c_3$.

Axiom 4 (Independence) For all outcomes c_1, c_2, and c_3, $c_1 \geq c_2$ if and only if $pc_1 + (1-p)c_3 \geq pc_2 + (1-p)c_3$ for all $p \in [0, 1]$.

If these four axioms hold[2] it is possible to prove the existence of a utility function with the properties listed above (the proof will not be given here). However, most decision-making situations involve uncertainty, i.e. many different outcomes are possible. Such situations can also be handled, since it follows (not completely trivially, though, at least not for the most general case) from the utility function properties listed above that the **expected utility** $U(c)$ of a mixed consequence $c = p_1c_1 + p_2c_2 + \ldots p_nc_n$, where p_k is the probability for consequence c_k, is given by

$$U(c) = \sum p_k u(c_k), \tag{5.5}$$

[2]Note that von Neumann and Morgenstern used slightly different axioms, which, however, amounted to essentially the same assumptions as those underlying axioms 1-4 listed here.

so that a consequence $c^{\mathrm{I}} = \sum p_k c_k$ is preferred to another consequence $c^{\mathrm{II}} = \sum q_k c_k$, if and only if $U(c^{\mathrm{I}}) > U(c^{\mathrm{II}})$.

Returning to the axioms above, it should be noted that none of them is trivial, and they have all been challenged in the literature. For example, the ordering and transitivity axioms may be violated in cases where a person has very vague preferences among a certain set of outcomes, or (more importantly) where the outcomes differ considerably in their implications, making it hard to express clear preferences.

The Archimedean axiom also leads to problems when some of the outcomes are extremely negative. For example if c_1 consists of winning \$2, c_2 consists of winning \$1, and c_3 means that the person is killed, it is quite evident that $c_1 > c_2 > c_3$. Thus, by axiom 3, there must be a value of p such that $pc_1 + (1-p)c_3 > c_2$, or, in other words, that the individual faced with the consequences $c_1, c_2,$ and c_3 would prefer the possibility (however small) of death to the prospect of winning \$1 with certainty, if only the probability of winning \$2 is sufficiently close to 1. Given the small amounts of money involved it seems, perhaps, unlikely that one would accept even the slightest risk of being killed in order to gain a measly \$2. On the other hand, people *do* accept risking their lives on a daily basis, for example by driving a car, in order to gain time (or money).

If the four axioms are accepted, however, the resulting utility function can be used as a powerful tool in decision-making. In order to do so, one must first be able to construct the utility function, a procedure that will now be described by means of an example. Consider a case in which a young man (A) in a bar is contemplating whether or not to approach a beautiful young woman sitting at the counter. He envisions two possible consequences: Either c_1, in which the girl is single and falls in love with him[3], or c_2 in which she is not single and where A will be severely beaten by her jealous boyfriend. Clearly, c_1 is preferable to c_2, and let us suppose that A has assigned utilities[4] such that $u(c_1) = 5$ and $u(c_2) = -10$. By the affinity property of the utility function, the expected utility of the **mixed consequence** c_m where, say, the probability p of c_1 is 0.2 and the probability $q = 1-p$ of c_2 is 0.8 (it is assumed no other potential consequences exist, once the decision to approach the young woman has been made) equals

$$U(c_m) = U(pc_1 + qc_2) = pu(c_1) + qu(c_2) = 0.2 \times 5 + 0.8 \times (-10) = -7. \quad (5.6)$$

Using the same procedure, utility values can be derived for any mixed consequence (i.e. for any value of $p \in [0, 1]$). Next, consider the consequence c of not approaching the young woman at all, and living with the knowledge of not

[3]The latter part of the consequence does not logically follow from the former, but for (considerable) simplicity, it will be assumed that this is the case nevertheless.

[4]The exact numerical values do not matter, as long as $u(c_1) > u(c_2)$.

having tried. Clearly c_1 is preferred to c and, unless A is hopelessly in love, c is preferred to c_2. Thus,

$$u(c_1) > U(c) > u(c_2), \tag{5.7}$$

but how should $U(c)$ be determined? Because of the affinity property of the utility function, it is clear that there exists some value of p such that

$$U(c) = U(pu(c_1) + (1-p)u(c_2)) = pu(c_1) + (1-p)u(c_2). \tag{5.8}$$

Thus, in other words (and by the first property of the utility function), given some time to think, A will come up with a value of p at which he is indifferent between the mixed consequence $pc_1 + (1-p)c_2$ and c. Let's say the point of indifference occurs at $p = 0.1$. Then

$$U(c) = 0.1 \times 5 + 0.9 \times (-10) = -8.5. \tag{5.9}$$

Thus, the expected utility for the consequence c has now been determined, and the expected utility of any other consequence preferred to c_2 but not to c_1 can be computed in a similar way. Note that there is no right answer - the exact shape of the utility function depends on A's preferences. For example, if he were more cautious, he would perhaps conclude that the point of indifference occurs at $p = 0.5$ rather than $p = 0.1$, in which case $U(c) = -2.5$.

As a more formal example, consider a lottery in which a person (B) is indifferent between the consequence c_{30} of a certain gain of \$30 and a the mixed consequence of gaining \$100 with probability $p = 0.5$ and gaining nothing with probability $p = 0.5$. Assume further that B has assigned utility $u(c_0) = 0$ and $u(c_{100}) = 1$ to the consequences of gaining \$0 and \$100, respectively. Simplifying the notation somewhat by writing $u(x)$ for $u(c_x)$ one then finds

$$U(30) = 0.5 \times u(0) + 0.5 \times u(100) = 0.5. \tag{5.10}$$

In other words, the **certainty equivalent** for the mixed consequence in Eq. (5.10) is equal to \$30. Proceeding in a similar way, using certainty equivalents, the expected utility for any amount \$x in the range $[0, 100]$ can be determined, and the utility function can be plotted. A common shape of the resulting curve is shown in Fig 5.1. The curve shows the utility function for a person who is **risk-averse**, i.e. who would prefer a sure payoff of x \$ to a lottery resulting in the same *expected* payoff. If a person is **risk-neutral** the utility function will be be a straight line through the origin and, similarly, for a **risk-prone** person, the utility function would bend upwards. Put differently, risk-aversion implies that the second derivative $U''(x)$ of the utility function is negative.

5.2 Rational decision-making

Once the utility functions for a given individual have been determined, the expected utility values can be used to guide behavior as follows: Given two

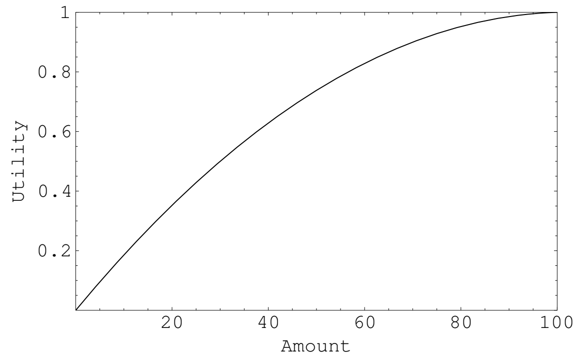

Figure 5.1: *A typical utility function, showing diminishing (sub-linear) utility values for larger amounts.*

possible actions a_1 and a_2, leading to the consequences c_1 and c_2, respecively, a **rational agent** will choose the action for which the corresponding expected utility is larger. Thus, if $U(c_1) > U(c_2)$, the agent would choose a_1, otherwise a_2 would be chosen. This behavior, in fact, *defines* a rational agent.

At this point, it should be noted that, even if all the axioms necessary for the existence of a utility function holds true, it is not always so that utilities can be computed in a simple way. In the examples considered above, the **state of nature** was known, i.e. the various consequences and their probabilities could be listed, and all that remained was the uncertainty due to randomness. However, in many situations, the state of nature is not known, and it is therefore much more difficult to assign appropriate utility values to guide behavior. The general theory of decision-making under uncertainty is, however, beyond the scope of this text.

5.2.1 Decision-making in animals

Animals (including humans) generally face the problem of scarce resources, making the ability to choose between different activities completely central to survival. For example, a common problem is to allocate time to various essential activities in such a way as to keep physiological variables (e.g. hunger, thirst, temperature) within acceptable limits.

Clearly, even simple animals are capable of rational behavior. It should be noted however, that rational behavior does *not* require rational thought. There are many examples (one of which will be described below) of rational behavior in animals that simply lack the brain power to contemplate their sensory input in any detail (let alone maintain complex internal states). In such cases, the rational selection of behavior occurs as a result of (evolutionary) design of the animal. Note that rational behavior does not automatically imply **intelligent behavior**. An agent is rational if it strives to maximize utility, but intelligent

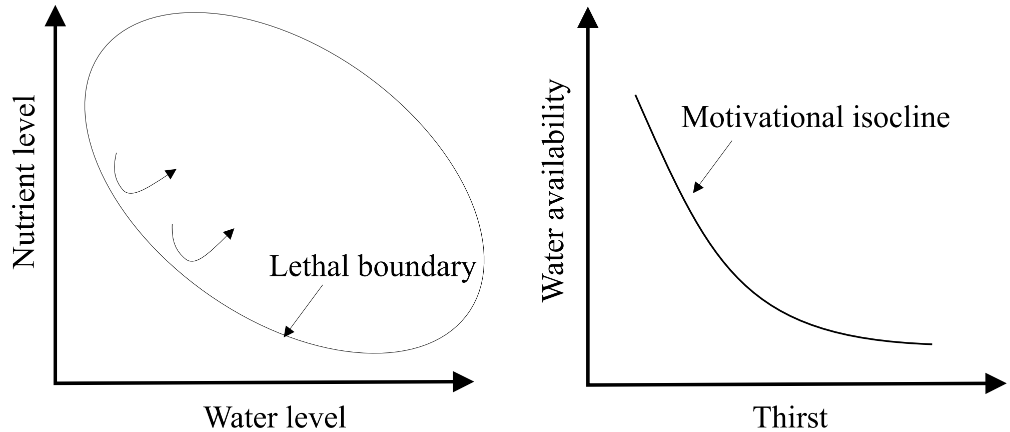

Figure 5.2: *Left panel: A physiological space. The curved arrows show typical trajectories followed as the animal comes sufficiently close to the lethal boundary. Right panel: A motivational space. The isocline joins points at which the tendency to perform the behavior in question (drinking, in this case) takes a given constant value.*

behavior will follow only if the animal's utility functions have been shaped (by evolution or as a result of learning) in an appropriate way.

As was discussed above, rational behavior amounts to the maximization of a quantity called utility that functions as a common currency for the comparison of different behaviors in any given situation. Put in a different way, a rational animal should switch between activities when (and only when) the switch leads to an increase in (marginal) utility. In ethology, it is customary to work with **cost** rather than utility. Thus, in ethology, the goals of the animal are commonly expressed in terms of the minimization of cost (e.g. energy expenditure). However, the term **benefit** (negative cost) which is also used in ethology, corresponds to utility in economics.

As mentioned in Chapter 2, the physiological state of an animal can be represented by a point in a multi-dimensional physiological space, exemplified in the left panel of Fig. 5.2, in which limits for each variable (e.g. the level of some nutrient in the body) can be introduced. Similarly, the motivational state of the animal, which is generated by the combination of the physiological and perceptual states, can be represented as a point in a multi-dimensional space. In this space, **isoclines** determining a given strength of the tendency to display a given behavior, can be introduced. A simplified, two-dimensional case, involving a single behavior, is shown in the right panel of Fig. 5.2. As the animal displays a given behavior, its location in the motivational space will change, as a result of, for example, a change in the perceived stimuli or its physiological state (e.g. satiation as a result of eating).

At the points in the (multi-dimensional) motivational space where two such isoclines cross each other, a switch in behavior (e.g. from eating to drink-

Figure 5.3: *A Stentor. Reproduced with kind permission of Dr. Ralf Wagner.*

ing) can be observed. By design, an animal's motivations will generally be such as to keep it away from the lethal boundaries of the physiological space.

The problem of behavior selection is made more complicated by the fact that the consequence of a given choice may be hard to assess (cognitively, or by design) accurately in an unstructured environment. Furthermore, the switch between two activities may involve a **cost of changing**, further complicating the decision-making process.

Behavior selection in Stentor

Stentor is a simple, single-celled animal (see Fig. 5.3), that attaches itself to e.g. a rock, and uses its hair-like **cilia** to sweep nutrients into its trumpet-shaped mouth. Obviously, this animal has no nervous system, but is nevertheless capable of quite complicated self-preserving behaviors: Besides the feeding behavior (B_1), Stentor is equipped with four different avoidance behaviors (in response e.g. to the presence of a noxious substance). In increasing order of energy expenditure, the avoidance behaviors are (a) *turning away* (B_2), (b) *reversing the cilia*, which interrupts the feeding activity (B_3), (c) *contraction, followed by waiting* (B_4), and (d) *detachment*, i.e. breaking away from the object to which the animal is attached (B_5). Despite its simple architecture, Stentor is able to execute the various avoidance behaviors in a rational sequence, i.e. starting with B_2 and, if this behavor is insufficient to escape the noxious substance, proceeding with B_3 etc. However, the sequence of activation of the different avoidance behaviors is not fixed: Sometimes, B_2 is followed by B_4 instead of B_3 etc. How can such a simple animal be capable of such complex behavior, given its utter inability to reason about which activity to perform? It turns out, as described in detail by Staddon [41], that Stentor's behavior can be accounted for by a very simple model, involving several leaky integrator elements of the type described in Chapter 2. A two-parameter leaky integrator is described by the equation

$$\frac{\mathrm{d}U}{\mathrm{d}t} + aU(t) = bX(t), \tag{5.11}$$

where a and b are constants and X is the external stimulus. Now, let U_i denote the utility associated with executing B_i, and set $U_1 = C =$ constant. For the avoidance behaviors, let the utility functions be given by

$$\frac{\mathrm{d}U_i}{\mathrm{d}t} + a_i U_i(t) = b_i X(t), \quad i = 2, 3, 4, 5 \tag{5.12}$$

Now, given initial values of the utilities $U_2, U_3, U_4,$ and U_5 (here all set, arbitrarily, to zero), and the variation of X with time, the utility for each behavior can be computed at all times. Using a utility-maximizing (rational!) behavior selection procedure, where the behavior $B_{i_{\mathrm{sel}}}$ corresponding to maximum current utility is selected, i.e.

$$i_{\mathrm{sel}} = \mathrm{argmax}(U_i), \tag{5.13}$$

the behavior of Stentor, including the variable activation sequence of the avoidance behaviors, can be modelled quite accurately, by selecting appropriate values for the constants a_i and b_i. An example is shown in Fig. 5.4. Here, it was assumed that the variation $X(t)$ of the noxious substance can be modelled as

$$\frac{\mathrm{d}X}{\mathrm{d}t} + k_1 X = X_1, \tag{5.14}$$

if B_1 is active, and

$$\frac{\mathrm{d}X}{\mathrm{d}t} + k_2 X = X_2, \tag{5.15}$$

if any other behavior is active. $k_1, k_2, X_1,$ and X_2 are non-negative constants, satisfying $X_1 > X_2$ (i.e. the amount of noxious substance tends towards higher values if no evasive action is taken). The left panel of the figure shows the variation of U_i for $i = 1, \ldots 5$. As can be seen in the figure, when X becomes sufficiently large, B_2 is activated for a short while. Next B_3, B_4, and B_5 are activated, in order. With the parameter values used in the example, the fall in X is quite slow. Thus, in subsequent activations of the avoidance behaviors, $B_2 - B_4$ are skipped, and the Stentor immediately activates B_5.

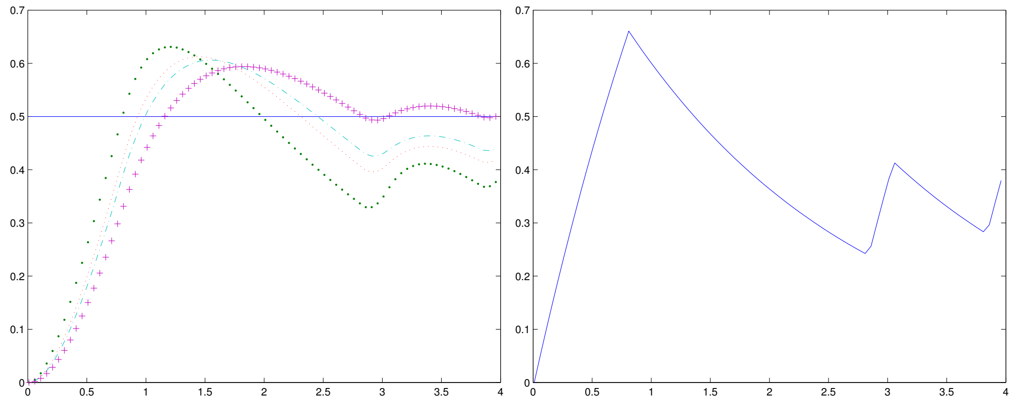

Figure 5.4: *Left panel: The variation of the utility functions for the 5 behaviors. Solid curve:* B_1*, large-dotted curve:* B_2*, small-dotted curve:* B_3*, dot-dashed curve:* B_4*, plus-sign curve:* B_5*. Right panel: The variation* $X(t)$ *of the concentration of the noxious substance.*

5.2.2 Decision-making in robots

The decision-making problems faced by autonomous robots are very similar to those faced by animals: an autonomous robot must complete certain tasks (e.g. delivery of objects in a factory), while avoiding collisions and while maintaining a sufficient level of energy in its batteries.

Guided by ethological results, using the principles of utility maximization (or cost minimization), McFarland [7] and McFarland and Bösser [8] have modelled elementary behavior selection in robots, using quadratic cost functions.

While the principle of utility maximization certainly is valid regardless of the number of choices (behaviors) available, the problem of assigning appropriate utility functions by hand becomes unmanageable in situations involving more than a few behaviors. In order to assign such functions, the designer of the robot must, in effect, tackle the daunting task of reasoning about the usefulness of various behaviors, in all relevant situations, and then assign appropriate utility values. An alternative procedure, more closely related to the principles guiding rational behavior in simple animals, has been developed by Wahde [43] and co-workers, who use evolutionary algorithms to evolve appropriate utility functions for behavioral selection. In the **utility function method** for behavioral selection, utility functions, usually in the form of polynomials in the state variables (comprising both external variables, such as e.g. the readings of IR sensors, and internal variables, corresponding e.g. to hormone levels), are evolved, and the principle of utility maximization is then used for the activation of behaviors. Methods for behavioral selection will be studied in the next chapter.