Chapter 1

AN INTRODUCTION TO GAME THEORY AND DECISION THEORY

Simon Parsons
Department of Computer Science, University of Liverpool, Liverpool L69 7ZF, United Kingdom.
s.d.parsons@csc.liv.ac.uk

Michael Wooldridge
Department of Computer Science, University of Liverpool, Liverpool L69 7ZF, United Kingdom.
m.j.wooldridge@csc.liv.ac.uk

Abstract In the last few years, there has been increasing interest from the agent community in the use of techniques from decision theory and game theory. Our aim in this article is to briefly summarise the key concepts of decision theory and game theory and explain why they are useful when building agent-based systems.

Keywords: Game theory, decision theory, multi-agent systems

1. INTRODUCTION

Decision theory (Raiffa, 1968) is a means of analysing which of a series of options should be taken when it is uncertain exactly what the result of taking the option will be. Decision theory concentrates on identifying the "best" decision option, where the notion of "best" is allowed to have a number of different meanings, of which the most common is that which maximises the expected benefit to the decision maker. Since self-interested entities are assumed to be acting best when maximising expected benefits, decision theory is often claimed to be able to make the most *rational* choice. Overall, decision theory provides a powerful tool with which to analyse scenarios in which a decision must be made.

Now, it is widely believed that the crucial issue in designing autonomous agents is how to provide those agents with the ability to select the best action from a range of possible actions. Frequently the agents in question are operating in an unpredicatble, and hence uncertain, environment, and therefore decision theory seems a natural tool to use to analyse their behaviour.

Game theory (Binmore, 1992) is a close relative of decision theory, which studies interactions between self-interested entities. In particular, it studies the problems of how interaction *strategies* can be designed that will maximise the welfare of an entity in an encounter, and how *protocols* or *mechanisms* can be designed that have certain desirable properties. In the same way that decision theory can be claimed to provide a means of making rational decisions under uncertainty, so game theory can be claimed to provide a rational means of analysing interactions. Notice that decision theory can be considered to be the study of *games against nature*, where nature is an opponent that does not seek to gain the best payout, but rather acts randomly.

In multi-agent systems, the issue of designing interaction strategies and mechanisms is very important, and so it comes as no surprise to learn that game theory has been widely applied. Many of these applications have been to analyse negotiation and co-ordination mechanisms.

2. DECISION THEORY

Classical decision theory, so called to distinguish it from a number of non-classical theories which have grown up in the last few years, is a set of mathematical techniques for making decisions about what action to take when the outcomes of the various actions are not known. Although the area grew up long before the concept of an intelligent agent was conceived, such agents are canonical examples of the decision makers which can usefully employ classical decision theory.

2.1. PROBABILITY THEORY

An agent operating in a complex environment is inherently uncertain about that environment; it simply does not have enough information about the enviroment to know either the precise current state of its environments, nor how that environment will evolve. Thus, for every variable X_i which captures some aspect of the current state of the environment, all the agent typically knows is that each possible value x_{i_j} of each X_i has some probability $\Pr(x_{i_j})$ of being the current value of X_i.

Writing $\mathbf{x}$ for the set of all x_{i_j}, we have:

$$\Pr : x \in \mathbf{x} \mapsto [0, 1]$$

and

$$\sum_j \Pr(x_{i_j}) = 1$$

In other words, the probability $\Pr(x_{i_j})$ is a number between 0 and 1 and the sum of the probabilities of all the possible values of X_i is 1. If X_i is known to have value x_{i_j} then $\Pr(x_{i_j}) = 1$ and if it is known not to have value x_{i_j} then $\Pr(x_{i_j}) = 0$.

While this mathematical definition of probability is rather straightforward, the same cannot be said of the semantics of probability. Indeed, there is no universal agreement on what probabilities mean. Of the various conflicting schools of thought (Shafer, 1992), there are two main positions. The first, historically, interprets a probability as a frequency of occurrence. This *frequentistic* approach takes that fact that an event a has a probability of 0.356 to mean that 0.356 of the time it will occur. This sounds like a very straightforward interpretation until you consider in more detail how such a probability can be measured[1] The second, *Bayesian*, position suggests that a probability is related to the odds that a rational person will bet on the event in question. Thus the probability of a being 0.356 means that a rational person will pay $0.356 to bet on a happening if they win $1 if a occurs. This approach solves some of the problems of the frequentistic interpretation (especially when estimating the probability of unique events) but has its own problems[2]. We will say no more about semantics here—all we say applies whatever semantics one chooses to use.

Given two of these variables, X_1 and X_2, then the probabilities of the various values of X_1 and X_2 may be related to one another. If they are not related, a case we distinguish be referring to X_1 and X_2 as being *independent*, then for any two values x_{1_i} and x_{2_j}, we have:

$$\Pr(x_{1_i} \wedge x_{2_j}) = \Pr(x_{1_i}) \Pr(x_{2_j})$$

If the variables are not independent, then:

$$\Pr(x_{1_i} \wedge x_{2_j}) = \Pr(x_{1_i} | x_{2_j}) \Pr(x_{2_j})$$

where $\Pr(x_{1_i} | x_{2_j})$ is the probability of X_1 having value x_{1_i} given that X_2 is known to take value x_{2_j}. Such *conditional probabilities* capture the relationship between X_1 and X_2, representing, for instance, the fact that x_{1_i} (the value "wet", say, of the variable "state of clothes") becomes

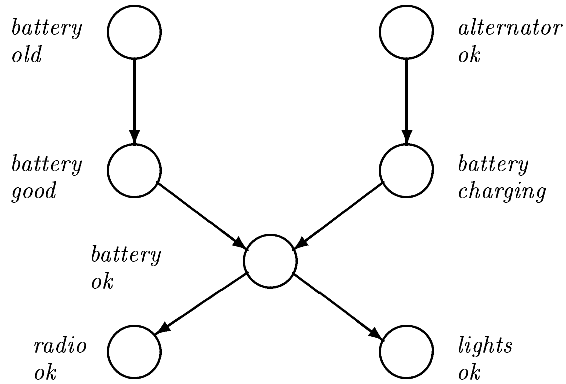

Figure 1.1. An example Bayesian network.

much more likely when x_{2_j} (the value "raining" of the variable "weather condition") is known to be true.

If we take the set of these X_i of which the agent is aware, the set $\mathbf{X}$, then for each pair of variables in $\mathbf{X}$ we can establish whether the pair are independent or not. We can then build up a graph in which each node corresponds to a variable in $\mathbf{X}$ and an arc joins two nodes if the variables represented by those nodes are not independent of each other. The resulting graph is known as a Bayesian network[3] (Pearl, 1988), and the graphical structure provides a convenient computational framework in which to calculate the probabilities of interest to the agent. In general, the agent will have some set of variables whose values it can observe, and once these observations have been taken, will want to calculate the probabilities of the various values of some other set of variables.

Figure 1.1 is an example of a fragment of a Bayesian network for diagnosing faults in cars. It represents the fact that the age of the battery (represented by the node *battery old*) has a probabilistic influence on how good the battery is, and that this in turn has an influence on whether the battery is operational (*battery ok*), the latter being affected also by whether the alternator is working and, as a result, whether the battery is recharged when the car moves. The operational state of the battery affects whether the radio and lights will work. In this network it is expected that the observations that can be carried out are those relating to the lights and the radio (and possibly the age of the battery), and that the result of these observations can be propagated through the network to establish the probability of the alternator being okay and the battery being good. In this case these latter variables are the ones which we are interested in since they relate to fixing the car.

Typically the variables an agent will be interested in are those that relate to its goals. For instance, the agent may be interested in choosing an action that will allow it to achieve a goal, and might therefore be interested in choosing that action which has the greatest chance of succeeding in achieving that goal. When the agent has many goals it could achieve, this strategy could be extended to make the agent choose to achieve the goal which has the greatest chance of being achieved, and to do this by applying the action which gives this greatest chance.

However, building an agent which follows this strategy is somewhat shortsighted since the agent will not consider the value of the goals, and will therefore choose a goal which is easy to achieve, but worthless, over a goal which is hard to achieve but very valuable. To take account of this problem, decision theory ke account of an agent's preferences. To do this, it makes use of the idea of *utility*.

2.2. UTILITY THEORY

We start from the assumption that each agent (or decision maker) has its own preferences and desires about how the world is. For the moment, we will not be concerned with where these preferences come from[4]; we will just assume that they are the preferences of the agent's user or owner. Next, we will assume that there is a set $\Omega = \{\omega_1, \omega_2, \ldots\}$ of "outcomes" or "states" that the agents have preferences over.

We will formally capture the preferences that an agent has by means of a *utility function*, which assigns to every outcome a real number, indicating how "good" the outcome is. The larger the number the better from the point of view of the agent with the utility function. Thus the preferences of an agent i will be captured by a function

$$u_i : \Omega \rightarrow \Re$$

It is not difficult to see that such a utility function leads to a *preference ordering* over outcomes. For example, if ω and ω' are both possible outcomes in Ω, and $u_i(\omega) \geq u_i(\omega')$, then outcome ω is *preferred* by agent i at least as much as ω'. We can introduce a bit more notation to capture this preference ordering. We write

$$\omega \succeq_i \omega'$$

as an abbreviation for

$$u_i(\omega) \geq u_i(\omega')$$

Similarly, if $u_i(\omega) > u_i(\omega')$, then outcome ω is *strictly preferred* by agent i over ω'. We write

$$\omega \succ_i \omega'$$

as an abbreviation for

$$u_i(\omega) > u_i(\omega').$$

In other words,

$$\omega \succ_i \omega' \text{ iff } u_i(\omega) \geq u_i(\omega') \text{ and not } u_i(\omega) = u_i(\omega').$$

We can see that the relation $\succeq_i$ really is a (partial) ordering, in that it has the following properties:

Reflexivity: For all $\omega \in \Omega$, we have that $\omega \succeq_i \omega$.

Transitivity: If $\omega \succeq_i \omega$, and $\omega' \succeq_i \omega''$, then $\omega' \succeq_i \omega''$.

Comparability: For all $\omega \in \Omega$, and $\omega' \in \Omega$ we have that either $\omega \succeq_i \omega'$ or $\omega' \succeq_i \omega$.

The strict preference relation will satisfy the second and third of these properties, but will clearly not be reflexive.

Undoubtedly the simplest way to think about utilities is as money; the more money, the better. However, it is deceptive to think that this is all that utilities are. Utility functions are a way of representing an agent's preferences. They *do not* simply equate to money.

To see why this is the case, suppose (and this really *is* a supposition) that the authors have US$500 million in the bank, while you, the reader, are absolutely penniless. A generous and rich benefactor appears, with one million dollars, which he generously wishes to donate to one or more of us. If the benefactor gives the dollar to the authors, what will the increase in the utility of our situation be? Well, we will have more money, so there will clearly be some increase in our utility. But there will not be much: after all, there is not much that you can do with US$501 million that you cannot do with US$500 million. In contrast, if the benefactor gave the money to you, the increase in your utility would be enormous; you would go from having no money at all to being a millionaire. That is a *big* difference[5].

This works the other way as well. Suppose the authors are in *debt* to the tune of US$500 million; well, there is frankly not that much difference in utility between owing US$500 million and owing US$499 million; they are both pretty bad. In contrast, there is a very big difference between being US$1 million in debt and not being in debt at all. A graph of the relationship between utility and money is shown in Figure 1.2.

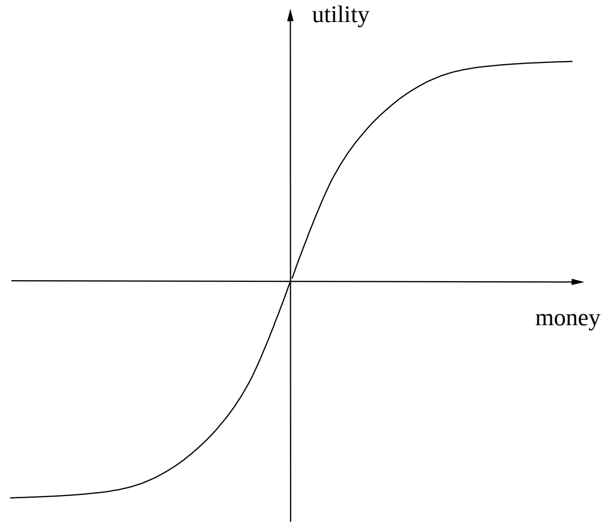

Figure 1.2. The relationship between money and utility.

So, to summarise, a utility is a value which is associated with a state of the world, and which represents the value that the agent places on that state of the world. Utilities provide a convenient means of encoding the preferences of an agent; as von Neumann and Morgenstern (von Neumann and Morgenstern, 1944) showed, it is possible to define utility functions that faithfully encode preferences such that a state S_i is preferred to S_j, if and only if it has a higher utility for the agent exactly as described above.

2.3. EXPECTED UTILITY

Now, we can consider that our agent has a set of possible actions $\mathbf{A}$, each member A_i of which has a range of possible outcomes since the actions are not deterministic. The value of taking a particular action will depend upon what the state of the world is—it is of little value carrying a surfboard when taking a trip across the Sahara, but it is extremely valuable carrying a surfboard when taking a trip across Bondi Beach—and so in choosing which action to undertake, our agent will need to look at the value of $U(S_j)$ where S_j is the state it is in after the action. Doing this for each possible action, the agent can then choose the action which leads to the state it values most. We can certainly build an agent which works in this way, and it would unerringly choose to achieve the goal with the highest value as encoded by its utility function. However it would be just as flawed as an agent which only tried to achieve the most likely goal, trying to achieve the most valuable goal irrespective of the difficulty of that goal.

To build more sensible agents we combine probability and utility calculations for each action and calculate the *expected utility* of each. This amounts to calculating a weighted average of the utility of each outcome, where the weight is the probability of that outcome given the action being performed. Since each outcome is itself a state, we have:

$$EU(A_i) = \sum_{S_j \in \mathbf{S}} \Pr(S_j|A_i)\, U(S_j)$$

where $\mathbf{S}$ is the set of all states. The agent then selects action A^* where:

$$A^* = arg \max_{A_i \in \mathbf{A}} \sum_{S_j \in \mathbf{S}} \Pr(S_j|A_i)\, U(S_j)$$

Now, these states which are being considered here are just particular instantiations of the set of state variables $\mathbf{X}$. Thus the probabilities in this calculation are just the probabilities of the X_i having particular values given the actions.

Harking back to the discussion of Bayesian networks above, we can think of the X_i as being structured as a graph, dropping the distinction between variables and the nodes in the graph which represent them. The A_i can be brought into the graph as well, as a different kind of node (square, perhaps, in contrast to the usual round ones relating to the X_i) linked to the X_i whose values they influence. We can also incorporate utilities. This time we only require a single node (a hexagon, to keep it distinct from the others), and this is linked to those X_i which affect its value. Such a graphical structure neatly captures all the dependencies in an expected utility calculation, and is known as an influence diagram (Howard and Matheson, 1984).

Figure 1.3 is an example of a small influence diagram capturing a decision problem which a company has to make about its research and development budget. Since the budget is the thing the decision is being made about, it is represented by a square *decision node*. This is linked to the factors it directly effects, namely the technical success of the company's products and their overall profitability, that latter being captured by the hexagonal *value node*. The remaining nodes are *chance nodes* and represent the other factors which relate to the decision. These are just like nodes in a Bayesian network. Given a particular instantiation of the decision node, the relevant values can be propagated through the network, using an algorithm such as Shacter's graph reduction algorithm (Shachter, 1986) to establish the expected utility of the decision.

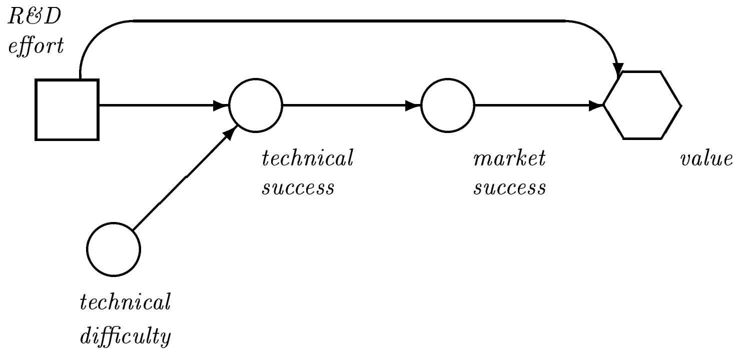

Figure 1.3. An example influence diagram.

2.4. DECISION THEORY AND AGENTS

Given that the basic mechanisms of decision theory fit so neatly into the context of intelligent agents, it is perhaps surprising that they have not been more widely employed in the field. However, agent systems which use decision theory seriously (that is adopting the notions of probability and utility) are rather scarce. One sub-area of decision theory is, however, becoming popular and that is the field of Markov decision processes (MDPs), discussed in detail in (Boutilier et al., 1999). In essence an MDP is an iterative set of classical decision problems. Consider a state of the world as a node in a graph. Carrying out an action in that state will result in a transition to one of a number of states, each connected to the first state by an arc, with some probability, and incur some cost. After a series of transitions a goal state may be reached, and the sequence of actions executed to do this is known as a *policy.* Solving an MDP amounts to finding a minimal cost policy for moving from some initial state to a goal state.

MDPs capture many of the facets of real world problems, but unrealistically assume that whatever system is solving the MDP knows at every point what state it is in. This amounts to assuming that it is possible to measure some aspect of the world and from this measurement tell precisely what state the world is in. This is rarely the case; it is far more likely is that from the measurement something can be uncertainly inferred about the world. In such a situation, the states of an MDP are replaced by beliefs about those states, and we have a partially observable Markov decision process (POMDP). Because they can capture so many real situations, POMDPs are currently a hot topic in agent research, despite the fact that they are intractable for all but the smallest problems.

3. GAME THEORY

Game theory is a branch of economics that studies interactions between self-interested agents. Like decision theory, with which it shares many concepts, game theory has its roots in the work of von Neumann and Morgenstern (von Neumann and Morgenstern, 1944). As its name suggests, the basic concepts of game theory arose from the study of games such as chess and checkers. However, it rapidly became clear that the techniques and results of game theory can equally be applied to *all* interactions that occur between self-interested agents.

The classic game theoretic question asked of any particular multi-agent encounter is: What is the best — most rational — thing an agent can do? In most multi-agent encounters, the overall outcome all depend critically on the choices made by all agents in the scenario. This implies that in order for an agent to make the choice that optimises its outcome, it must reason *strategically*. That is, it must take into account the decisions that other agent may make, and must assume that they will act so as to optimise their own outcome. Game theory gives us a way of formalising and analysing such concerns.

3.1. MULTIAGENT ENCOUNTERS

First, let us simplify things by assuming that we have just two agents; things tend to be much more complicated when we have more than two. Call these agents i and j respectively. Each of the agents is assumed to be *self-interested*. That is, each agent has its own preferences and desires about how the world is, and these are encoded as utilities. We also need to introduce a model of the environment in which these agents will act. The idea is that out two agents will simultaneously choose an action to perform in the environment, and as a result of the actions they select, an outcome in Ω will result. The *actual* outcome that will result will depend on the particular *combination* of actions performed. We will also assume that the agents have no choice about whether to perform an action — they have to simply go ahead and perform one. Further, it is assumed that they cannot see the action performed by the other agent.

To make the analysis a bit easier, we will assume that each agent has just two possible actions that it can perform. We will call these two actions C, for "cooperate", and "D", for "defect". (The rationale for this terminology will become clear below.) Let $Ac = \{C, D\}$ the set of these actions. The way the *environment* behaves is then determined by a function

$$Env : \underbrace{Ac}_{\text{agent } i\text{'s action}} \times \underbrace{Ac}_{\text{agent } j\text{'s action}} \to \Omega$$

In other words, on the basis of the action (either C or D) selected by agent i, and the action (also either C or D) chosen by agent j.

Here is an example of an environment function:

$$\begin{array}{ll} Env(D,D) = \omega_1 & Env(D,C) = \omega_2 \\ Env(C,D) = \omega_3 & Env(C,C) = \omega_4 \end{array} \tag{1}$$

This environment maps each combination of actions to a *different* outcome. This environment is thus sensitive to the actions that each agent performs. At the other extreme, we can consider an environment that maps each combination of actions to the *same* outcome.

$$\begin{array}{ll} Env(D,D) = \omega_1 & Env(D,C) = \omega_1 \\ Env(C,D) = \omega_1 & Env(C,C) = \omega_1 \end{array} \tag{2}$$

In this environment, it does not matter what the agents do: the outcome will be the same. Neither agent has any influence in such a scenario. We can also consider an environment that is only sensitive to the actions performed by one of the agents.

$$\begin{array}{ll} Env(D,D) = \omega_1 & Env(D,C) = \omega_2 \\ Env(C,D) = \omega_1 & Env(C,C) = \omega_2 \end{array} \tag{3}$$

In this environment, it does not matter what agent i does: the outcome depends solely on the action performed by j. If j chooses to defect, then outcome ω_1 will result; if j chooses to cooperate, then outcome ω_2 will result.

The interesting story begins when we put an environment together with the preferences that agents have. To see what we mean by this, suppose we have the most general case, characterised by (1), where both agents are able to exert some influence over the environment. Now let us suppose that the agents have utility functions defined as follows:

$$\begin{array}{llll} u_i(\omega_1) = 1 & u_i(\omega_2) = 1 & u_i(\omega_3) = 4 & u_i(\omega_4) = 4 \\ u_j(\omega_1) = 1 & u_j(\omega_2) = 4 & u_j(\omega_3) = 1 & u_j(\omega_4) = 4 \end{array} \tag{4}$$

Since we know that every different combination of choices by the agents are mapped to a different outcome, we can abuse notation somewhat by writing the following:

$$\begin{array}{llll} u_i(D,D)=1 & u_i(D,C)=1 & u_i(C,D)=4 & u_i(C,C)=4 \\ u_j(D,D)=1 & u_j(D,C)=4 & u_j(C,D)=1 & u_j(C,C)=4 \end{array} \tag{5}$$

We can then characterise agent i's preferences over the possible outcomes in the following way:

$$(C,C) \succeq_i (C,D) \succ_i (D,C) \succeq_i (D,D) \tag{6}$$

Now, consider the following question:

> If you were agent i in this scenario, what would you choose to do — cooperate or defect?

In this case, the answer is pretty unambiguous. Agent i prefers *all* the outcomes in which it cooperates over *all* the outcomes in which it defects. Agent i's choice is thus clear: it should cooperate. It does not matter what agent j chooses to do.

For agent j, the story is the same: we can write j's preferences as follows.

$$(C,C) \succeq_i (D,C) \succ_i (C,D) \succeq_i (D,D)$$

In just the same way, agent j prefers all the outcomes in which *it* cooperates over all the outcomes in which it defects. Notice that in this scenario, neither agent has to expend any effort worrying about what the other agent will do: the action it should perform does not depend in any way on what the other does.

If both agents in this scenario act rationally, that is, they both choose to perform the action that will lead to their preferred outcomes, then the "joint" action selected will be (C, C): both agents will cooperate.

Now suppose that, for the same environment, the agents' utility functions were as follows:

$$\begin{array}{llll} u_i(D,D)=4 & u_i(D,C)=4 & u_i(C,D)=1 & u_i(C,C)=1 \\ u_j(D,D)=4 & u_j(D,C)=1 & u_j(C,D)=4 & u_j(C,C)=1 \end{array} \tag{7}$$

Agent i's preferences over the possible outcomes are thus as follows:

$$(D,D) \succeq_i (D,C) \succ_i (C,D) \succeq_i (C,C)$$

Agent j's preferences are:

$$(D,D) \succeq_i (C,D) \succ_i (D,C) \succeq_i (C,C)$$

In this scenario, agent i can do no better than to defect. The agent prefers *all* the outcomes in which it defects over *all* the outcomes in

	i defects	i cooperates
j defects	4 (i), 4 (j)	0 (i), 5 (j)
j cooperates	1 (i), 1 (j)	3 (i), 3 (j)

Table 1.1 . A payoff matrix with dominant strategies.

which it cooperates. Similarly, agent j can do no better than defect: it also prefers all the outcomes in which it defects over all the outcomes in which it cooperates. Once again, the agents do not need to engage in *strategic* thinking (worrying about what the other agent will do): the best action to perform is entirely independent of the other agent's choice. I emphasise that in most multiagent scenarios, the choice an agent should make is not so clear cut; indeed, most are much more difficult.

We can neatly summarise the previous interaction scenario by making use of a standard game theoretic notation known as a *payoff matrix*, as in Table 1.1. The way to read such a payoff matrix is as follows. Each of the four cells in the matrix corresponds to one of the four possible outcomes. For example, the top-right cell corresponds to the outcome in which i cooperates and j defects; the bottom-left cell corresponds to the outcome in which i defects and j cooperates. The payoffs received by the two agents are written in the cell. The value in the top-right of each cell is the payoff received by player i (the *column player*), while the value in the bottom left of each cell is the payoff received by agent j (the *row player*). Thus in Table 1.1, if i cooperates and j defects, j gets 5 and i gets 0. As payoff matrices are standard in the literature, and are a much more succinct notation than the alternatives, we will use them as standard in the remainder of this chapter.

Before proceeding to consider any specific examples of multiagent encounter, let us introduce some of the theory that underpins the kind of analysis we have informally discussed above.

3.2. DOMINANT STRATEGIES AND NASH EQUILIBRIA

Given a particular multiagent encounter involving two agents i and j, there is one critically important question that both agents want answered: *What should I do?* We have already seen some multiagent encounters, and informally argued what the best possible outcome should be. In this section, we will define some of the concepts that are used in answering this question.

The first concept we will introduce is that of *dominance.* To understand what is meant by dominance, suppose we have two subsets of Ω, which we refer to as Ω_1 and Ω_2 respectively. We will say that Ω_1 *dominates* Ω_2 for agent i if *every* outcome in Ω_1 is preferred over *every* outcome in Ω_2. For example, suppose that:

- $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$;
- $\omega_1 \succ_i \omega_2 \succ_i \omega_3 \succ_i \omega_4$;
- $\Omega_1 = \{\omega_1, \omega_2\}$; and
- $\Omega_2 = \{\omega_3, \omega_4\}$.

Then Ω_1 strongly dominates Ω_2 since $\omega_1 \succ_i \omega_3$, $\omega_1 \succ_i \omega_4$, $\omega_2 \succ_i \omega_3$, and $\omega_2 \succ_i \omega_4$. However, Ω_2 does not strongly dominate Ω_1, since (for example), it is not the case that $\omega_3 \succ_i \omega_1$. Formally, Ω_1 strongly dominates Ω_2 iff the following condition is true:

$$\forall \omega_1 \in \Omega_1, \forall \omega_2 \in \Omega_2, \text{ we have } \omega_1 \succ_i \omega_2.$$

Now, in order to bring ourselves in line with the game theory literature, we will start referring to actions (members of the set Ac) as *strategies.* Given any particular strategy s for an agent i in a multiagent interaction scenario, there will be a number of possible outcomes. Let us denote by s^* the outcomes that may arise by i playing strategy s. For example, referring to the example environment in equation (1), from agent i's point of view we have $C^* = \{\omega_3, \omega_4\}$, while $D^* = \{\omega_1, \omega_2\}$.

Now, we will say a *strategy* s_1 dominates a strategy s_2 if the set of outcomes possible by playing s_1 dominates the set possible by playing s_2, that is, if s_1^* dominates s_2^*. Again, referring back to the example of (6), it should be clear that, for agent i, "cooperate" strongly dominates "defect". Indeed, as there are only two strategies available, the cooperate strategy is *dominant*: it is not dominated by any other strategy. The presence of a dominant strategy makes the decision about what to do extremely easy: the agent guarantees its best outcome by performing the dominant strategy. In following a dominant strategy, an agent guarantees itself the best possible payoff.

Another way of looking at dominance is that if a strategy s is dominated by another strategy s', then a rational agent will not follow s (because it can guarantee to do better with s'). When considering what to do, this allows us to *delete* dominated strategies from our consideration, simplifying the analysis considerably. The idea is to iteratively consider each strategy s in turn, and if there is another remaining strategy that strongly dominates it, then to delete strategy s from consideration. If we end up with a single strategy remaining, then this will be the dominant strategy, and is clearly the rational choice. Unfortunately, for many interaction scenarios, there will not be a strongly dominant strategy; after deleting strongly dominated strategies, we may find more than one strategy remaining. What to do then? Well, we can start to delete *weakly* dominated strategies. A strategy s_1 is said to weakly dominate strategy s_2 if every outcome s_1^* is preferred at least as much as every outcome s_2^*. The problem is that if a strategy is only weakly dominated, then it is not necessarily irrational to use it; in deleting weakly dominated strategies, we may therefore "throw away" a strategy that would in fact have been useful to use.

The next notion we shall discuss is one of the most important concepts in the game theory literature, and in turn is one of the most important concepts in analysing multiagent systems. The notion is that of *equilibrium*, and more specifically, *Nash equilibrium.* The intuition behind equilibrium is perhaps best explained by example. Every time you drive a car, you need to decide which side of the road to drive on. The choice is not a very hard one: if you are in the UK, for example, you will probably choose to drive on the left; if you are in the US or continental Europe, you will drive on the right. The reason the choice is not hard is that it is a Nash equilibrium strategy. Assuming everyone else is driving on the left, you can do no better than drive on the left also. From everyone else's point of view, assuming you are driving on the left then everyone else can do no better than drive on the left also.

In general, we will say that two strategies s_1 and s_2 are in Nash equilibrium if:

1 under the assumption that agent i plays s_1, agent j can do no better than play s_2; and

2 under the assumption that agent j plays s_2, agent i can do no better than play s_1.

The *mutual* form of an Equilibrium is important because it "locks the agents in" to a pair of strategies. *Neither agent has any incentive to deviate from a Nash equilibrium.* To see why, suppose s_1, s_2 are a pair of strategies in Nash equilibrium for agents i and j respectively, and that agent j chooses to play some other strategy: s_3 say. Then by definition, i will do no better, and may possibly do worse than it would have done by playing s_1.

The presence of a Nash equilibrium pair of strategies in a game might appear to be the definitive answer to the question of what to do in any given scenario. Unfortunately, there are two important results in the game theory literature which serve to make life difficult:

1 *Not every interaction scenario has a Nash equilibrium.*

If there is no Nash equilibrium, then it may be possible to look for a *mixed* equilibrium as described in the next section.

2 *Some interaction scenarios have more than one Nash equilibrium.*

Despite these negative results, the notion of a Nash equilibrium is an extremely important concept, and will be useful in the analysis that follows.

3.3. COMPETITIVE AND ZERO-SUM INTERACTIONS

Suppose we have some scenario in which an outcome $\omega \in \Omega$ is preferred by agent i over an outcome ω' if, and only if, ω' is preferred over ω by agent j. Formally,

$$\omega \succ_i \omega' \quad \text{if and only if } \omega' \succ_j \omega.$$

The preferences of the players are thus diametrically opposed to one-another: one agent can only improve its lot (i.e., get a more preferred outcome) at the expense of the other. An interaction scenario that satisfies this property is said to be *strictly competitive*, for hopefully obvious reasons.

Zero-sum encounters are those in which, for any particular outcome, the utilities of the two agents sum to zero. Formally a scenario is said to be zero sum if the following condition is satisfied:

$$u_i(\omega) + u_j(\omega) = 0 \quad \text{for all } \omega \in \Omega.$$

It should be easy to see that any zero sum scenario is strictly competitive. Zero sum encounters are important because they are the most "vicious" types of encounter conceivable, allowing for no possibility of cooperative behaviour. If you allow your opponent positive utility, then this means that you get *negative* utility — intuitively, you are worse off than you were before the interaction.

Games such as chess and chequers are the most obvious examples of strictly competitive interactions. Indeed, any game in which the possible outcomes are win or lose will be strictly competitive. Outside these rather abstract settings, however, it is hard to think of real-world examples of zero-sum encounters. War might be cited as a zero sum interaction between nations, but even in the most extreme wars, there will usually be at least *some* common interest between the participants (e.g., in ensuring that the planet survives).

	i defects	i cooperates
j defects	-3 3	1 -1
j cooperates	0 0	-1 1

Table 1.2 . A payoff matrix for a game with no pure equilibrium.

For these reasons, some social scientists are sceptical about whether zero-sum games exist in real-world scenarios (Zagare, 1984, p.22). Interestingly, however, people interacting in many scenarios have a tendency to treat them *as if they were zero-sum*. Below, we will see that in some scenarios — where there is the possibility of mutually beneficial cooperation — this type of behaviour can be damaging.

3.4. MIXED EQUILIBRIA

All the scenarios we have considered up to now have had solutions in which the agents pick a single outcome, a single row or column in the payoff matrix[6]. As discussed above, there are scenarios in which agents that choose such *pure* strategies can arrive at Nash equilibrium or other stable solutions. However, there are some situations in which pure strategies will *not* give a stable solution. As an example, consider the game decribed by the payoff matrix in Table 1.2.

This payoff matrix describes a zero sum game. If both i and j defect, then i loses 3 and j wins 3. If i defects and j cooperates, then both get a zero payoff. If i cooperates and j defects, i wins 1 and j loses 1, while if both cooperate then the payoff is reversed. Thus neither agent has one move it can make which is definitely better than the other—it all depends on what the other agent does. The result of this arrangement of payoffs is that neither agent can make a choice which holds up if the other agent somehow finds out what the first agent is intending to do (a situation that is sometimes referred to as having a *spyproof* strategy).

To see this, consider what each agent will chose to do. If i knew that j would defect, then i would choose to cooperate. However, if j knew that i would co-operate, then j would choose to cooperate. Similarly, if i knew that j would cooperate, then it would choose to defect, while if j knew that i was planning to defect, then j would defect as well. So any piece of information about what the other is intending (which, of course, could be established by either agent by thinking about what it would do in the other agent's shoes) will cause an agent to change its strategy.

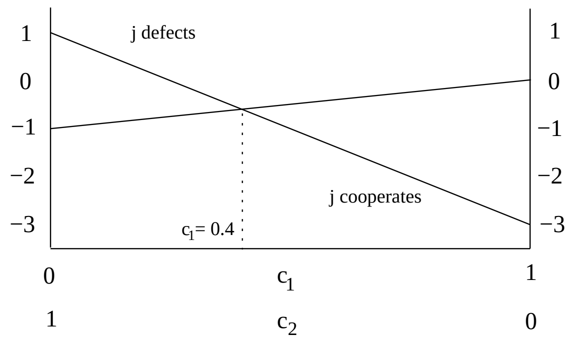

Figure 1.4. Determining a mixed strategy for Agent i.

So, rather than adopt a pure strategy, agents adopt a *mixed* strategy in which they take a random choice across a set of pure strategies. More formally, agent i picks a vector of probabilities over the columns in the payoff matrix that relate to the pure strategies it might choose. In general, where agent i has n possible strategies, it needs a vector:

$$c = (c_1, c_2, \dots, c_n)$$

where

$$\sum_k c_k = 1$$

and

$$c_k \geq 0$$

Agent i then picks strategy k with probability c_k. Any vector of probabilities gives a mixed strategy. The question is how to obtain a good mixed strategy. Clearly, for the 2×2 game we have here, i needs to compute the best values of c_1 and c_2—the probabilities of defecting and cooperating respectively. These will be the values which give i the highest expected payoff for its mixed strategy. The diagram in Figure 1.4 is one way of thinking about solving the problem of picking a mixed strategy for agent i, which is fundamentally a problem of determining the values of c_1 and c_2 which will maximise the agent's expected utility.

The diagram plots expected utility on the vertical axis(axes), and the probabilities c_1 and c_2 across the horizontal axis. If j chooses to play the first row, in other words to defect, then i will get either -3 or 1 depending on its choice of column. When this choice is made randomly, the expected payoff to i is the weighted sum:

$$y = -3c_1 + c_2$$

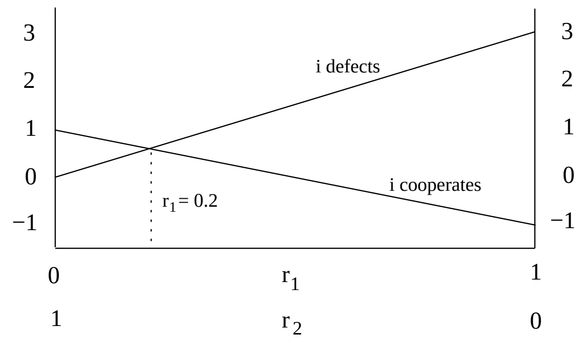

Figure 1.5. Determining a mixed strategy for Agent j.

which is the equation of the line joining 1 and -3 in Figure 1.4 (since $c_1 = 1 - c_2$). Similarly, if j chooses to play the second row, then the expected payoff to i is between 0 and -1, depending on the probability with which the two options available to i are selected, and is plotted by the other line in Figure 1.4. The intersection of the two lines gives the value of c_1 and c_2 (which is just $1 - c_1$) at which i is indifferent as to what j plays—whichever j chooses, the expected payoff to i is the same. This, then, yields a form of stability and the resulting values of c_1 and c_2 give the mixed strategy for i.

Agent j can analayse the problem in terms of a probability vector

$$r = (r_1, r_2)$$

and come up with a similar picture (Figure 1.5).

Now, let's consider the payoff's the players will expect. With i having mixed strategy (c_1, c_2) and j having (r_1, r_2), then the loss that i will expect to make will be:

$$\begin{aligned} L &= 3c_1r_1 + 0(1 - r_1)c_1 \\ &\quad - r_1(1 - c_1) + (1 - c_1)(1 - r_1) \\ &= 5c_1r_1 - 2r_1 - c_1 + 1 \end{aligned}$$

Now, assuming that i uses $c_1^* = 0.4$ as calculated above. Then:

$$\begin{aligned} L &= 5(0.4r_1) - 2r_1 - c_1 + 1 \\ &= 0.6 \end{aligned}$$

and i will have an expected loss of 0.6. Similarly, we can calculate the gain that j will expect to make as:

$$\begin{aligned} G &= 3c_1 r_1 + 0(1-c_1)r_1 \\ &\quad - c_1(1-r_1) + (1-c_1)(1-r_1) \\ &= 5c_1 r_1 - r_1 - 2c_1 + 1 \end{aligned}$$

If j picks $r_1^* = 0.2$ then:

$$G = 0.6$$

The neat thing is that the expected gain or loss for one agent does not depend upon the strategy of the other agent—once the correct mixed strategy has been determined, it no longer matters what the other agent is going to do, the first agent still has the same expected outcome. This result generalises to more general games, and Von Neumann's Minimax Theorem shows that you can always find a pair of mixed strategies x^* and y^* which result in i and j having the same expected value for the game.

This theory is sufficient background for the rest of the book, but before we leave the topic, it is worth considering the best-known multi-agent scenario: the *prisoner's dilemma*.

4. THE PRISONER'S DILEMMA

Consider the following scenario:

> Two men are collectively charged with a crime and held in separate cells. They have no way of communicating with each other or making any kind of agreement. The two men are told that:
>
> 1 if one of them confesses to the crime and the other does not, the confessor will be freed, and the other will be jailed for three years; and
>
> 2 if both confess to the crime, then each will be jailed for two years.
>
> Both prisoners know that if neither confesses, then they will each be jailed for one year.

We refer to confessing as defection, and not confessing as cooperating, because we are viewing the problem from the point of view of the prisoners—from the point of view of one of them, the best, most cooop-erative, thing the other can do is to not confess.

There are four possible outcomes to the prisoner's dilemma, depending on whether the agents cooperate or defect, and so the environment is of type (1). Abstracting from the scenario above, we can write down the utility functions for each agent in the payoff matrix of Table 4. Note that the numbers in the payoff matrix are *not* the length of the jail term.

	i defects	i cooperates
j defects	2 2	1 5
j cooperates	5 1	3 3

Table 1.3 . A payoff matrix for the Prisoner's Dilemma.

In other words, the utilities are as follows:

$$\begin{array}{llll} u_i(D,D)=2 & u_i(D,C)=5 & u_i(C,D)=1 & u_i(C,C)=3 \\ u_j(D,D)=2 & u_j(D,C)=1 & u_j(C,D)=5 & u_j(C,C)=3 \end{array}$$

And the preferences are:

$$\begin{array}{c} (D,C) \succ_i (C,C) \succ_i (D,D) \succ_i (C,D) \\ (C,D) \succ_j (C,C) \succ_j (D,D) \succ_j (D,C) \end{array}$$

So, what should a prisoner do? The answer is not as clear cut as the previous pure strategy examples we looked at. It is not the case a prisoner prefers all the outcomes in which it cooperates over all the outcomes in which it defects. Similarly, it is not the case that a prisoner prefers all the outcomes in which it defects over all the outcomes in which it cooperates.

The "standard" approach to this problem is to put yourself in the place of a prisoner, i say, and reason as follows:

> Suppose I cooperate. Then if j cooperates, we will both get a payoff of 3. But if j defects, then I will get a payoff of one. So the best payoff I can be *guaranteed* to get if I cooperate is 1.
>
> Suppose I defect. Then if j cooperates, then I get a payoff of 5, whereas if j defects, then I will get a payoff of 2. So the best payoff I can be guaranteed to get if I defect is 2.
>
> So, if I cooperate, the worst case is I will get a payoff of 1, whereas if I defect, the worst case is that I will get 2.
>
> Since I would prefer a payoff of 2 to a payoff of 1, I should defect.

Since the scenario is symmetric (i.e., both agents reason the same way), then the outcome that will emerge — if both agents reason "rationally" — is that both agents will defect, giving them each a payoff off 2.

Notice that neither strategy dominates in this scenario, so our first route to finding a choice of strategy is not going to work. Turning to Nash equilibria, there is a single Nash equilibrium of D, D. Thus under the assumption that i will play D, j can do no better than play D, and under the assumption that j will play D, i can also do no better than play D.

But is this the best they can do? *Naive intuition says not.* Surely if they both *cooperated*, then they could do better — they would receive a payoff of 3. But if you assume the other agent will cooperate, then the rational thing to do — the thing that maximises your utility — is to defect. The conclusion seems inescapable: the rational thing to do in the prisoners dilemma is defect, even though this appears to "waste" some utility. (The fact that our naive intuition tells us that utility appears to be wasted here, and that the agents could do better by cooperating, even though the rational thing to do is to defect, is why this is referred to as a dilemma.)

The prisoners dilemma may seem an abstract problem, but it turns out to be very common indeed. In the real world, the prisoners dilemma appears in situations ranging from nuclear weapons treaty compliance to negotiating with one's children. Consider the problem of nuclear weapons treaty compliance. Two countries i and j have signed a treaty to dispose of their nuclear weapons. Each country can then either cooperate (= get rid of their weapons), or defect (= keep their weapons). But if you get rid of your weapons, you run the risk that the other side keeps theirs, making them very well off while you suffer what is called the "suckers payoff". In contrast, if you keep yours, then the possible outcomes are that you will have nuclear weapons while the other country does not (a very good outcome for you), or else at worst that you both retain your weapons. This may not be the best possible outcome, but is certainly better than you giving up your weapons while your opponent kept theirs, which is what you risk if your give up your weapons.

Many people find the conclusion of this analysis — that the rational thing to do in the prisoner's dilemma is defect — deeply upsetting. For the result *seems* to imply that cooperation can only arise as a result or *irrational* behaviour, and that cooperative behaviour can be exploited by those who behave rationally. The apparent conclusion is that nature really is "red in tooth and claw". Particularly for those who are inclined to a liberal view of the world, this is unsettling and perhaps even distasteful. As civilized beings, we tend to pride ourselves on somehow "rising above" the other animals in the world, and believe that we are capable of nobler behaviour: to argue in favour of such an analysis is therefore somehow immoral, and even demeaning to the entire human race.

Naturally enough, there have been several attempts to respond to this analysis of the prisoners dilemma, in order to "recover" cooperation (Binmore, 1992, p.355–382).

We Are Not All Machiavelli! The first approach is to argue that we are not all such "hard boiled" individuals as the prisoner's dilemma (and more generally, this kind of game theoretic analysis) implies. We are *not* seeking to constantly maximise our own welfare, possibly at the expense of others. Proponents of this kind of argument typically point to real-world examples of *altruism* and spontaneous, mutually-beneficial cooperative behaviour in order to justify their claim.

There is some strength to this argument: we do not, (or at least, most of us do not), constantly deliberate about how to maximise our welfare without any consideration for the welfare of our peers. Similarly, in many scenarios, we would be happy to trust our peers to recognise the value of a cooperative outcome without even mentioning it to them, being no more than mildly annoyed if we get the "sucker's payoff".

There are several counter responses to this. First, it is pointed out that many real-world examples of spontaneous cooperative behaviour are not really the prisoner's dilemma. Frequently, there is some built in mechanism that makes it in the interests of participants to cooperate. For example, consider the problem of giving up your seat on the bus. We will frequently give up our seat on the bus to an older person, mother with children, etc., apparently at some discomfort (= loss of utility) to ourselves. But it could be argued that in such scenarios, society has ways of punishing non-cooperative behaviour: suffering the hard and unforgiving stares of fellow passengers when we do not give up our seat, or worse, being accused in public of being uncouth!

Second, it is argued that many "counter examples" of cooperative behaviour arising do not stand up to inspection. For example, consider a public transport system, which relies on everyone cooperating and honestly paying their fare every time they travel, even though whether or not they have paid is not verified. The fact that such a system works (the buses turn up on time) would appear to be some evidence that relying on spontaneous cooperative can work. But the fact that such a system functions does not mean that the system is not exploited: it will be, and if there is no means of checking whether or not someone has paid their fare and punishing non-compliance, then all other things being equal, those individuals that do exploit the system will be better off. Unpalatable, perhaps, but true nevertheless.

The Other Prisoner is My Twin! A second line of attack is to argue that two prisoner's will "think alike", and recognise that cooperation is the best outcome. For example, suppose the two prisoners are twins, inseparable since birth; then, it is argued, if their though processes are sufficiently aligned, they will both recognise the benefits of cooperation, and behave accordingly. The answer to this is that it implies there are not actually two prisoner's playing the game. If I can make my twin select a course of action simply by "thinking it", then we are not playing the prisoner's dilemma at all.

This "fallacy of the twins" argument often takes the form "what if everyone were to behave like that" (Binmore, 1992, p311). The answer, as Yossarian pointed out in Joseph Heller's *Catch 22*, is that if everyone else behaved like that, you would be a damn fool to behave any other way.

People Are Not Rational! Some would argue[7], that we might indeed be happy to risk cooperation as opposed to defection when faced with situations where the sucker's payoff really does matter very much. For example, paying a bus fare that amounts to a few pennies does not really hurt us much, even if everybody else is defecting and hence exploiting the system. But, it is argued, when we are faced with situations where the sucker's payoff really *hurts* us — life or death situations and the like — we will choose the "rational" course of action that maximises our welfare, and defect.

The Shadow of the Future Lest the discussion so far prove too depressing, it should be emphasised that there are quite natural variants of the prisoner's dilemma in which cooperation *is* the rational thing to do. One idea is to *play the game more than once.* In the *iterated prisoner's dilemma*, the "game" of the prisoner's dilemma is played a number of times. Each play is referred to as a "round". Critically, it is assumed that each agent can see what the opponent did on the previous round: player i can see whether j defected or not, and j can see whether i defected or not.

Now, for the sake of argument, assume that the agents will continue to play the game *forever*: every round will be followed by another round. Now, under these assumptions, what is the rational thing to do? If you know that you will be meeting the same opponent in future rounds, the incentive to defect appears to be considerably diminished, for two reasons:

- If you defect now, your opponent can *punish* you by also defecting. Punishment is not possible in the one-shot prisoner's dilemma.

- If you "test the water" by cooperating initially, and receive the sucker's payoff on the first round, then because you are playing the game indefinitely, this loss of utility (one util) can be "amortized" over the future rounds. When taken into the context of an infinite

> (or at least very long) run, then the loss of a single unit of utility will represent a small percentage of the overall utility gained.

So, if you play the prisoner's dilemma game indefinitely, then cooperation is a rational outcome (Binmore, 1992, p358). The "shadow of the future" encourages us to cooperate in the infinitely repeated prisoner's dilemma game.

This seems to be very good news indeed, as truly one-shot games are comparatively scarce in real life. When we interact with someone, then there is often a good chance that we will interact with them in the future, and rational cooperation begins to look possible. However, there is a catch.

Suppose you agree to play the iterated prisoner's dilemma a *fixed* number of times, (say 100). You need to decide (presumably in advance) what your strategy for playing the game will be. Consider the last round (i.e., the 100th game). Now, on this round, you know — as does your opponent — that you will not be interacting again. In other words, the last round is in effect a one-shot prisoner's dilemma game. As we know from the analysis above, the rational thing to do in a one-shot prisoner's dilemma game is defect. Your opponent, as a rational agent, will presumably reason likewise, and will also defect. On the 100th round, therefore, you will both defect. But this means that the last "real" round, is 99. But similar reasoning leads us to the conclusion that this round will also be treated in effect like a one-shot prisoner's dilemma, and so on. Continuing this *backwards induction* leads inevitably to the conclusion that, in the iterated prisoner's dilemma with a fixed, pre-determined number of rounds, defection is the dominant strategy, as in the one-shot version (Binmore, 1992, p.354).

Whereas it seemed to be very good news that rational cooperation is possible in the iterated prisoner's dilemma with an infinite number of rounds, it seems to be very bad news that this possibility appears to evaporate if we restrict ourselves to repeating the game a pre-determined, fixed number of times. Returning to the real-world, we know that in reality, we will only interact with our opponents a finite number of times (after all, one day the world will end). We appear to be back where we started.

The story is actually better than it might at first appear, for several reasons. The first is that *actually* playing the game an infinite number of times is not necessary. As long as the "shadow of the future" looms sufficiently large, then it can encourage cooperation. So, rational cooperation can become possible if both players know, with sufficient probability, that they will meet and play the game again in the future.

The second reason is that, even though a cooperative agent can suffer when playing against a defecting opponent, it can do well overall provided it gets sufficient opportunity to interact with other cooperative agents.

5. SUMMARY

In this chapter we have discussed some of the basic concepts in decision theory and game theory. Our intention was simply to make the rest of the book comprehensible even for those who have not come across either decision theory, or game theory, or both, in the past, and this brief discussion stops short of being anything like a comprehensive introduction to the topics. To do the latter would take a book as long as this whole work (at the very least). However, for this who want to know more then there are a number of places to look.

For information on game theory, one of the best places to start is (Binmore, 1992). This is a very readable introduction from a leading expert, and, being intended as a textbook, starts at the very beginning and covers much of the game theory one is ever likely to need to know. From a more specifically agent-oriented perspective, (Rosenschein and Zlotkin, 1994) is a seminal work in the application of game theory to multi-agent encounters, and (Kraus, 2001) is a deep and rigorous investigation of similar matters. In addition, (Sandholm, 1999) gives a good general discussion of the use of game theory in designing multi-agent systems, as well as related topics such as auction theory. Initial work on the iterated prisoner's dilemma was carried out by Axelrod (Axelrod, 1984), and this is recommended as a point of departure for further reading on the topic. (Mor and Rosenschein, 1995) provides pointers into recent prisoner's dilemma literature. An non-mathematical introduction to game theory, with an emphasis on the applications of game theory in the social sciences, is (Zagare, 1984), and some of the early work in game theory is surveyed in (Schwalbe and Walker, 2001)[8].

The seminal introduction to decision theory is (Raiffa, 1968), and a good, if slightly dogmatic, alternative is (Lindley, 1975). (Smith, 1999) is not quite as detailed, but covers much of the relevant material and is very easy to read. Bayesian networks are covered in detail by (Pearl, 1988), the first book on the subject, and by (Castillo et al., 1997; Cowell et al., 1999; Jensen, 2001). The last two also cover influence diagrams and so deal with network models for the whole of decision theory (in contrast Pearl and Castillo *et al.* deal only with probability). Finally, (Puterman, 1994) is a comprehensive introduction to Markov decision processes, but for work in this area much of the most relevant material has only yet appeared in journal and conference papers.

Notes

1. You can calculate it as a ration of occurrence to the total number of events, but what counts as an occurrence? It is clear when calculating the probability of getting heads when tossing a coin, but much less obvious when, for example, estimating the probability of a particualr disease giving certain symptoms. Even trickier is the question of how many events you have to sum over. Typical definitions claim that you only truly know the probability if you sum over an infinite number of events which presents obvious practical diffculties.

2. The key step in the argument as to why betting rates are a suitable way of determing probabilities is that if the bettor misestimates the probability and so proposes a different bet—say paying $0.4 for the chance of winning $1 when *a* occurs—then the person they are betting with can exploit them. In particular, this second person can construct a *Dutch book*, a set of bets which can win an arbitrarily large amount of money from the bettor. Thus, the argument runs, the bettor will be motivated to get the probability right. The main problem with this argument is that it places a considerable cognitive burden on whoever is establishing the probability.

3. The notion of independence captured in the arcs of a Bayesian network is somewhat more complex than that described here, but the difference is not relevant for the purposes of this article. For full details, see (Pearl, 1988).

4. Indeed, in economics as a whole, very little thought is given to the question of where preferences come from.

5. To misquote Martin Amis, the difference is clear—a cool million dollars.

6. How might an agent choose more than one row in the kind of framework we have been considering? Well, as we will see in a very short while, it doesn't, but at the same time it doesn't choose a single row either...

7. Ken Binmore certainly did at the UKMAS workshop in December 1998.

8. They also provide the first English translation of what is generally regarded as the first paper on game theory.

References

Axelrod, R. (1984). *The Evolution of Cooperation*. Basic Books.

Binmore, K. (1992). *Fun and Games: A Text on Game Theory*. D. C. Heath and Company: Lexington, MA.

Boutilier, C., Dean, T., and Hanks, S. (1999). Decision-theoretic planning: structural assumptions and computational leverage. *Journal of Artificial Intelligence Research*, 11:1–94.

Castillo, E., Gutiérrez, J. M., and Hadi, A. S. (1997). *Expert Systems and Probabilistic Network Models*. Springer Verlag, Berlin, Germany.

Cowell, R. G., Dawid, A. P., Lauritzen, S. L., and Spiegelhalter, D. J. (1999). *Probabilistic Networks and Expert Systems*. Springer Verlag, Berlin, Germany.

Howard, R. A. and Matheson, J. E. (1984). Influence diagrams. In Howard, R. A. and Matheson, J. E., editors, *Readings on the Principles and Applications of Decision Analysis*, pages 719–762. Strategic Decisions Group, Menlo Park, CA.

Jensen, F. V. (2001). *Bayesian Networks and decision graphs.* Springer Verlag, New York, NY.

Kraus, S. (2001). *Strategic negotiation in multiagent environments.* MIT Press, Cambridge, MA.

Lindley, D. V. (1975). *Making Decisions.* John Wiley & Sons, Chichester, UK.

Mor, Y. and Rosenschein, J. S. (1995). Time and the prisoner's dilemma. In *Proceedings of the First International Conference on Multi-Agent Systems (ICMAS-95)*, pages 276–282, San Francisco, CA.

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

Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems; Networks of Plausible Inference.* Morgan Kaufmann, San Mateo, CA.

Puterman, M. L. (1994). *Markov Decision Processes.* John Wiley & Sons, New York, NY.

Raiffa, H. (1968). *Decision Analysis: Introductory Lectures on Choices under Uncertainty.* Addison Wesley, Reading, MA.

Rosenschein, J. S. and Zlotkin, G. (1994). *Rules of Encounter: Designing Conventions for Automated Negotiation among Computers.* The MIT Press: Cambridge, MA.

Sandholm, T. (1999). Distributed rational decision making. In Weiß, G., editor, *Multiagent Systems*, pages 201–258. The MIT Press: Cambridge, MA.

Schwalbe, U. and Walker, P. (2001). Zermelo and the early history of game theory. *Games and Economic Behaviour*, 34:123–137.

Shachter, R. (1986). Evaluating influence diagrams. *Operations Research*, 34:871–882.

Shafer, G. (1992). Can the various meanings of probability be reconciled? In Keren, G. and Lewis, C., editors, *A Handbook for Data Analysis in the Behavioural Sciences.* Lawrence Erblaum, Hillsdale, NJ.

Smith, J. Q. (1999). *Decision Analysis: A Bayesian Approach.* Springer Verlag, Berlin.

Zagare, F. C. (1984). *Game Theory: Concepts and Applications.* Sage Publications: Beverly Hills, CA.