6

Multiagent Interactions

So far in this book, we have been focusing on the problem of how to build an individual agent. Except in passing, we have not examined the issues associated in putting these agents together. But there is a popular slogan in the multiagent systems community;

> There's no such thing as a single agent system.

The point of the slogan is that interacting systems, which used to be regarded as rare and unusual beasts, are in fact the norm in the everyday computing world. All but the most trivial of systems contains a number of sub-systems that must interact with one another in order to successfully carry out their tasks. In this chapter, I will start to change the emphasis of the book, from the problem of 'how to build an agent', to 'how to build an agent society'. I begin by defining what we mean by a multiagent system.

Figure 6.1 {from Jennings (2000)) illustrates the typical structure of a multiagenl system. The system contains a number of agents, which interact with one another through communication. The agents are able to act in an environment; different agents have different 'spheres of influence', in the sense that they will have control over - or at least be able to influence - different parts of the environment. These spheres of influence may coincide in some cases. The fact that these spheres of influence may coincide may give rise to dependency relationships between the agents. For example, two robotic agents may both be able to move through a door - but they may not be able to do so simultaneously. Finally, agents will also typically be linked by other relationships. Examples might be 'power' relationships, where one agent is the 'boss' of another.

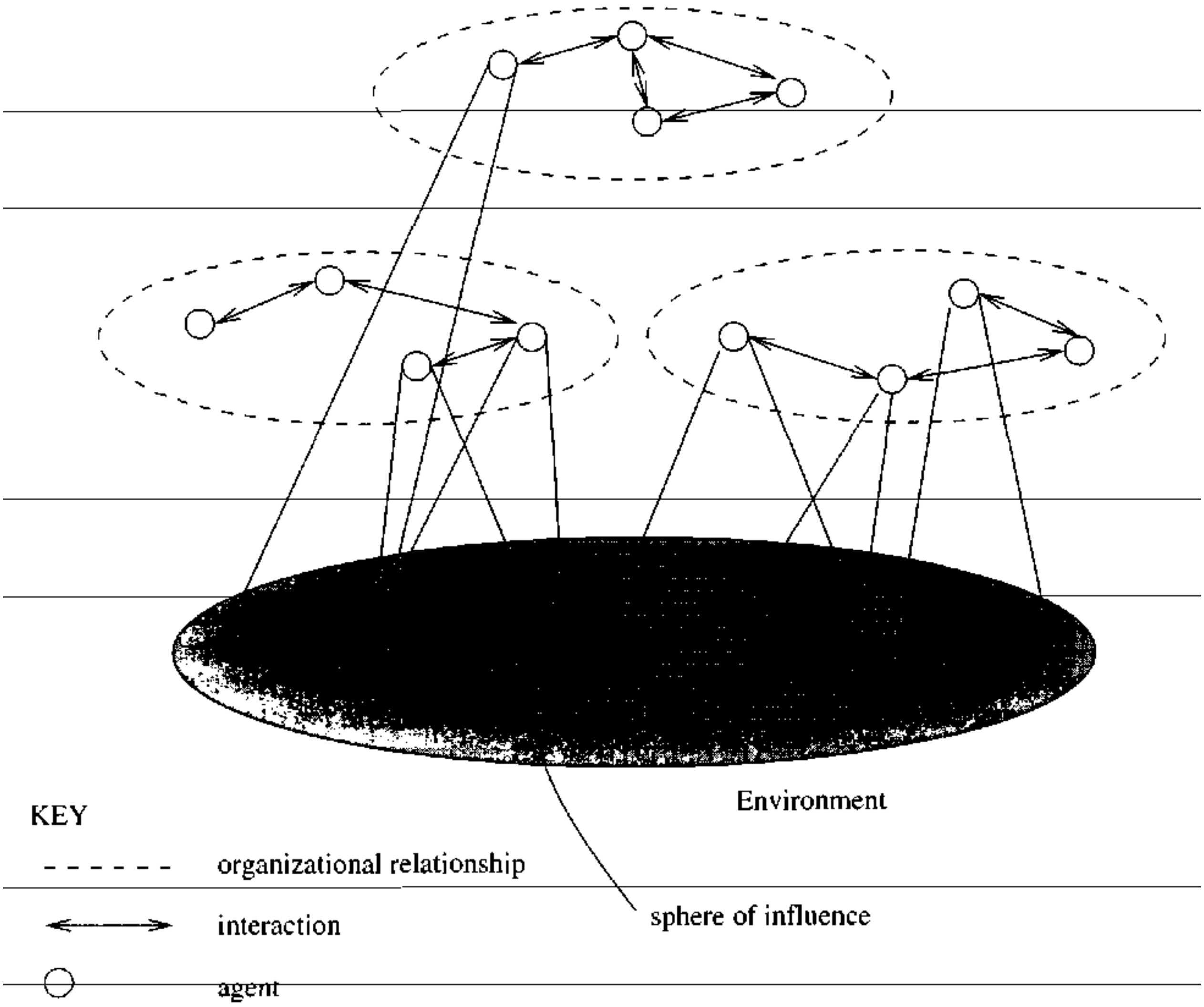

Figure 6.1 Typical structure of a multiagent system.

The most important lesson of this chapter - and perhaps one of the most important lessons of multiagent systems generally - is that when faced with what appears to be a multiagent domain, it is critically important to understand the *type* of interaction that takes place between the agents. To see what I mean by this, let us start with some notation.

6.1 Utilities and Preferences

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 should be. For the moment, we will not be concerned with where these preferences come from; just assume that they are the preferences of the agent's user or owner. Next, we will assume that there is a set $\Omega = \{\omega_1, \omega_2, \dots\}$ of 'outcomes' or 'states' that the agents have preferences over. To make this concrete, just think of these as outcomes of a game that the two agents are playing.

We formally capture the preferences that the two agents have by means of *utility functions,* one for each agent, which assign to every outcome a real number, indicating how 'good' the outcome is for that agent. The larger the number the better from the point of view of the agent with the utility function. Thus agent i's preferences will be captured by a function

$$u_i : \Omega \to \mathbb{R}$$

and agent *j*'s preferences will be captured by a function

$$u_j : \Omega \to \mathbb{R}.$$

(Compare with the discussion in Chapter 2 on tasks for agents.) It is not difficult to see that these utility function lead to a *preference ordering* over outcomes. For example, if to and to' are both possible outcomes *in*Ω, and $u_i(\omega) \geqslant u_i(\omega')$, then outcome to is *preferred* by agent *i* at least as much as to'. We can introduce a bit more notation to capture this preference ordering. We write

$$\omega \succeq_i \omega'$$

as an abbreviation for

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

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

$$\omega \succ_i \omega'$$

as an abbreviation for

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

In other words,

to >*i* to' if and only if $u_i(\omega) \geqslant u_i(\omega')$and not $u_i(\omega) = u_i(\omega')$.

We can see that the relation $\succeq_i$ really is an ordering, over Ω, in that it has the following properties.

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

Transitivity: if $\omega \succeq_i$ to', and $\omega' \succeq_i$ to", then to $\succeq_i$ to".

Comparability: for all to $\in$ Ω and to' $\in \Omega$ wehave that either to $\succeq_i$ to' or to' $\succeq_i$ ω.

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

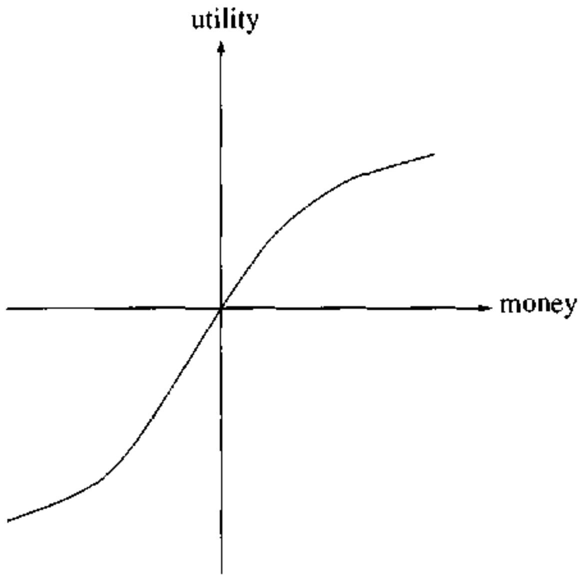

Figure 6.2 The relationship between money and utility.

What is utility?

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

To see what I mean by this, suppose (and this really *is* a supposition) that I have US$500 million in the bank, while you are absolutely penniless. A rich benefactor appears, with one million dollars, which he generously wishes to donate to one of us. If the benefactor gives the money to me, what will the increase in the utility of my situation be? Well, I have more money, so there will clearly be *some* increase in the utility of my situation. 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 go from having no money at all to being a millionaire. That is a *big* difference.

This works the other way as well. Suppose I am 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 6.2.

6.2 Multiagent Encounters

Now that we have our model of agent's preferences, we need to introduce a model of the environment in which these agents will act. The idea is that our 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. Thus *both* agents can influence the outcome. 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 $Ac = \{C, D\}$ be the set of these actions. The way the *environment* behaves is then determined by a function

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

(This is essentially a state transformer function, as discussed in Chapter 2.) 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 an outcome will result.

Here is an example of an environment function:

$$\tau(D,D) = \omega_1, \quad \tau(D,C) = \omega_2, \quad \tau(C,D) = \omega_3, \quad \tau(C,C) = \omega_4. \tag{6.1}$$

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.

$$\tau(D,D) = \omega_1, \quad \tau(D,C) = \omega_1, \quad \tau(C,D) = \omega_1, \quad \tau(C,C) = \omega_1. \tag{6.2}$$

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.

$$\tau(D,C) = \omega_2, \tag{6.3}$$

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 I mean by this, suppose we have the most general case, characterized by (6.1), 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:

$$\begin{aligned} &= 1, \quad u_i(\omega_2) = \quad u_i(\omega_3) = 4, \quad = 4. \\ u_j(\omega_1) &= 1, \quad u_j(\omega_2) = 4, \quad u_j(\omega_3) = 1, \quad u_j(\omega_4) = 4. \end{aligned} \tag{6.4}$$

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:

$$\left.\begin{array}{llll} u_i(D,D) = 1, & u_i(D,C) = 1, & u_i(C,D) = 4, & u_i(C,C) = 4, \\ u_j(D,D) = 1, & u_j(D,C) = 4, & u_j(C,D) = 1, & u_j(C,C) = 4. \end{array}\right\} \quad (6.5)$$

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

$$C,C \succeq_i C,D \succ_i D,C \succeq_i D,D.$$

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 (I hope), 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.

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:

$$\left.\begin{array}{llll} u_i(D,D) = 4, & u_i(D,C) = 4, & u_i(C,D) = 1, & u_i(C,C) = 1, \\ u_j(D,D) = 4, & u_j(D,C) = 1, & u_j(C,D) = 4, & u_j(C,C) = 1. \end{array}\right\} \quad (6.6)$$

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

$$D,D \succeq_i D,C \succ_i C,D \succeq_i C,C.$$

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 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 emphasize that in most multiagent scenarios, the choice an agent should make is not so clear cut; indeed, most arc much more difficult.

We can nearly summarize the previous interaction scenario by making use of a standard game-theoretic notation known as a *payoff matrix:*

	i defects	i cooperates
j defects	4 4	1 4
j cooperates	4 1	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).* 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.

6.3 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 by i 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, a set of outcomes Ω_1 strongly dominates set Ω_2 if the following condition is true:

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

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 (6.1), 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 (6.5), 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 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. We will not take this discussion further; see the *Notes and Further Reading* section at the end of this chapter for pointers to the literature.

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 rive 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 arc a pair of strategics in Nash equilibrium for agents i and j, respectively, and that agent i 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;* and

(2) some interaction scenarios have more than one Nash equilibrium.

Despite these negative results, Nash equilibrium is an extremely important concept, and plays an important role in the analysis of multiagent systems.

Competitive and Zero-Sum Interactions

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

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

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:

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

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). Perhaps games like chess - which are a highly stylized form of interaction - are the only real-world examples of zero-sum encounters.

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.

Enough abstract theory! Let us now apply this theory to some actual multiagent scenarios. First, let us consider what is perhaps the best-known scenario: the *prisoner's dilemma*.

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. Before reading any further, stop and think about this scenario: if you were one of the prisoners, what would you do? (Write down your answer somewhere, together with your reasoning; after you have read the discussion below, return and see how you fared.)

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 (6.1). Abstracting from the scenario above, we can write down the utility functions for each agent in the following payoff matrix:

	i defects	i cooperates
j defects	2 (j), 2 (i)	0 (j), 5 (i)
j cooperates	5 (j), 0 (i)	3 (j), 3 (i)

Note that the numbers in the payoff matrix do not refer to years in prison. They capture how good an outcome is for the agents - the shorter jail term, the better. In other words, the utilities are

$$u_i(D,D) = 2, \quad u_i(D,C) = 5, \quad u_i(C,D) = 0, \quad u_i(C,C) = 3,$$
$$u_j(D,D) = 2, \quad u_j(D,C) = 0, \quad u_j(C,D) = 5, \quad u_j(C,C) = 3,$$

and the preferences are

$$D,C \succ_i C,C \succ_i D,D \succ_i C,D,$$
$$C,D \succ_j C,C \succ_j D,D \succ_j D,C.$$

What should a prisoner do? The answer is not as clear cut as the previous 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 0. So the best payoff I can be *guaranteed* to get if I cooperate is 0.
- 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 0, whereas if I defect, the worst case is that I will get 2.

- I would prefer a guaranteed payoff of 2 to a guaranteed payoff of 0, so 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 strongly 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.*

Is this the best they can do? *Naive intuition says not.* Surely if they both *cooperated,* then they could do belief - 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 maximizes your utility - is to defect. The conclusion seems inescapable: the rational thing to do in the prisoner's 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 prisoner's dilemma may seem an abstract problem, but it turns out to be very common indeed. In the real world, the prisoner's 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 (i.e. get rid of their weapons), or defect (i.e. 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 'sucker's 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 of *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 prisoner's dilemma, in order to 'recover' cooperation (Binmore, 1992, pp. 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 maximize 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 maximize our welfare without any consideration for the welfare of our peers. Similarly, in many scenarios, we would be happy to trust our peers to recognize 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 (i.e. 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 they have paid is not verified. The fact that such a system works would appear to be evidence that reiving on spontaneous cooperation can work. But the fact that such a system works does not mean that it 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 (defect) will be better off than those that pay honestly (cooperate). 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 recognize that cooperation is the best outcome. For example, suppose the two prisoners are twins, unseparated since birth; then, it is argued, if their thought processes are sufficiently aligned, they will both recognize the benefits of cooperation, and behave accordingly. The answer to this is that it implies there are not actually two prisoners 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, p. 311). 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 - and game theorist Ken Binmore certainly did at the UKMAS workshop in December 1998 - that we might indeed be happy to risk cooperation as opposed to defection when faced with situations where the sucker's payoff really does not 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 maximizes our welfare, and defect.

The shadow of the future

Lest the discussion has so far proved too depressing, it should be emphasized 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, p. 358). 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, 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, predetermined, commonly known 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 predetermined, 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. To understand how this idea works, we will now turn to one of the best-known pieces of multiagent systems research: Axelrod's prisoner's dilemma tournament.

Axelrod's tournament

Robert Axelrod was (indeed, is) a political scientist interested in how cooperation can arise in societies of self-interested agents. In 1980, he organized a public tournament in which political scientists, psychologists, economists, and game theoreticians were invited to submit a computer program to play the iterated prisoner's dilemma. Each computer program had available to it the previous choices made by its opponent, and simply selected either C or *D* on the basis of these. Each computer program was played against each other for five games, each game consisting of two hundred rounds. The 'winner' of the tournament was the program that did best *overall,* i.e. best when considered against the whole range of programs. The computer programs ranged from 152 lines of program code to just five lines. Here are some examples of the kinds of strategy that were submitted.

ALL-D. This is the 'hawk' strategy, which encodes what a game-theoretic analysis tells us is the 'rational' strategy in the finitely iterated prisoner's dilemma: always defect, no matter what your opponent has done.

RANDOM. This strategy is a control: it ignores what its opponent has done on previous rounds, and selects either C or *D* at random, with equal probability of either outcome.

TIT-FOR-TAT. This strategy is as follows:

(1) on the first round, cooperate;

(2) on round $t > 1$, do what your opponent did on round $t - 1$.

TIT-FOR-TAT was actually the simplest strategy entered, requiring only five lines of Fortran code.

TESTER. This strategy was intended to exploit computer programs that did not punish defection: as its name suggests, on the first round it tested its opponent by defecting. If the opponent ever retaliated with defection, then it subsequently played TIT-FOR-TAT. If the opponent did not defect, then it played a repeated sequence of cooperating for two rounds, then defecting.

JOSS. Like TESTER, the JOSS strategy was intended to exploit 'weak' opponents. Tt is essentially TIT-FOR-TAT, but 10% of the time, instead of cooperating, it will defect.

Before proceeding, consider the following two questions.

(1) On the basis of what you know so far, and, in particular, what you know of the game-theoretic results relating to the finitely iterated prisoner's dilemma, which strategy do you think would do best overall?

(2) If you were entering the competition, which strategy would you enter?

After the tournament was played, the result was that the overall winner was TIT-FOR-TAT: the simplest strategy entered. At first sight, this result seems extraordinary. It appears to be empirical proof that the game-theoretic analysis of the iterated prisoner's dilemma is wrong: cooperation *is* the rational thing to do, after all! But the result, while significant, is more subtle (and possibly less encouraging) than this. TIT-FOR-TAT won because the overall score was computed by taking into account *all* the strategies that it played against. The result when TIT-FOR-TAT was played against ALL-D was exactly as might be expected: ALL-D came out on top. Many people have misinterpreted these results as meaning that TIT-FOR-TAT is the optimal strategy in the iterated prisoner's dilemma. *You should be careful not to interpret Axelrod's results in this way.* TIT-FOR-TAT was able to succeed because it had the opportunity to play against other programs that were also inclined to cooperate. Provided the environment in which TIT-FOR-TAT plays contains sufficient opportunity to interact with other 'like-minded' strategies, TIT-FOR-TAT can prosper. The TIT-FOR-TAT strategy will not prosper if it is forced to interact with strategies that tend to defect.

Axelrod attempted to characterize the reasons for the success of TIT-FOR-TAT, and came up with the following four rules for success in the iterated prisoner's dilemma.

(1) **Do not be envious.** In the prisoner's dilemma, it is not necessary for you to 'beat' your opponent in order for you to do well.

(2) **Do not be the first to defect.** Axelrod refers to a program as 'nice' if it starts by cooperating. He found that whether or not a rule was nice was the single best predictor of success in his tournaments. There is clearly a risk in starting with cooperation. But the loss of utility associated with receiving the sucker's payoff on the first round will be comparatively small compared with possible benefits of mutual cooperation with another nice strategy.

(3) **Reciprocate cooperation and defection.** As Axelrod puts it, 'TIT-FOR-TAT represents a balance between punishing and being forgiving' (Axelrod, 1984, p. 119): the combination of punishing defection and rewarding cooperation seems to encourage cooperation. Although TIT-FOR-TAT can be exploited on the first round, it retaliates relentlessly for such non-cooperative behaviour. Moreover, TIT-FOR-TAT punishes with *exactly* the same degree of violence that it was the recipient of: in other words, it never 'overreacts' to defection. In addition, because TIT-FOR-TAT is *forgiving* (it rewards cooperation), it is possible for cooperation to become established even following a poor start.

(4) Do **not** be **too clever.** As noted above, TIT-FOR-TAT was the simplest program entered into Axelrod's competition. Either surprisingly or not, depending on your point of view, it fared significantly better than other programs that attempted to make use of comparatively advanced programming techniques in order to decide what to do. Axelrod suggests three reasons for this:

(a) the most complex entries attempted to develop a model of the behaviour of the other agent while ignoring the fact that this agent was in turn watching the original agent - they lacked a model of the reciprocal learning that actually takes place;

(b) most complex entries over generalized when seeing their opponent defect, and did not allow for the fact that cooperation was still possible in the future - they were not *forgiving;*

(c) many complex entries exhibited behaviour that was too complex to be understood - to their opponent, they may as well have been acting randomly.

From the amount of space we have devoted to discussing it, you might assume that the prisoner's dilemma was the *only* type of multiagent interaction there is. This is not the case-

Other Symmetric 2×2 Interactions

Recall the ordering of agent i's preferences in the prisoner's dilemma:

$$D,C \succ_i C,C \succ_i D,D \succ_i C,D.$$

This is just one of the possible orderings of outcomes that agents may have. If we restrict our attention to interactions in which there are two agents, each agent has two possible actions (C or D), and the scenario is *symmetric,* then there are 4! = 24 possible orderings of preferences, which for completeness I have summarized in Table 6.1. (In the game-theory literature, these are referred to as symmetric 2×2 games.)

In many of these scenarios, what an agent should do is clear-cut. For example, agent *i* should clearly cooperate in scenarios (1) and (2), as both of the outcomes in which i cooperates are preferred over both of the outcomes in which *i* defects. Similarly, in scenarios (23) and (24), agent *i* should clearly defect, as both outcomes in which it defects are preferred over both outcomes in which it cooperates. Scenario (14) is the prisoner's dilemma, which we have already discussed at length, which leaves us with two other interesting cases to examine: the *stag hunt* and the *game of chicken.*

The stag hunt

The stag hunt is another example of a social dilemma. The name stag hunt arises from a scenario put forward by the Swiss philosopher Jean-Jacques Rousseau in his 1775 *Discourse on Inequality.* However, to explain the dilemma, I will use a scenario that will perhaps be more relevant to readers at the beginning of the 21st century (Poundstone, 1992, pp. 218, 219).

> You and a friend decide it would be a great joke to show up on the last day of school with some ridiculous haircut. Egged on by your clique, you both *swear* you'll get the haircut.

Table 6.1 The possible preferences that agent i can have in symmetric interaction scenarios where there are two agents, each of which has two available actions, C (cooperate) and D (defect); recall that X, Y means the outcome in which agent i plays X and agent j plays Y.

Scenario	Preferences over outcomes	Comment
1.	$C,C \succ_i C,D \succ_i D,C \succ_i D,D$	cooperation dominates
2.	$C,C \succ_i C,D \succ_i D,D \succ_i D,C$	cooperation dominates
3.	$C,C \succ_i D,C \succ_i C,D \succ_i D,D$	
4.	$C,C \succ_i D,C \succ_i D,D \succ_i C,D$	stag hunt
5.	$C,C \succ_i D,D \succ_i C,D \succ_i D,C$	
6.	$C,C \succ_i D,D \succ_i D,C \succ_i C,D$	
7.	$C,D \succ_i C,C \succ_i D,C \succ_i D,D$	
8.	$C,D \succ_i C,C \succ_i D,D \succ_i D,C$	
9.	$C,D \succ_i D,C \succ_i C,C \succ_i D,D$	
10.	$C,D \succ_i D,C \succ_i D,D \succ_i C,C$	
11.	$C,D \succ_i D,D \succ_i C,C \succ_i D,C$	
12.	$C,D \succ_i D,D \succ_i D,C \succ_i C,C$	
13.	$D,C \succ_i C,C \succ_i C,D \succ_i D,D$	game of chicken
14.	$D,C \succ_i C,C \succ_i D,D \succ_i C,D$	prisoner's dilemma
15.	$D,C \succ_i C,D \succ_i C,C \succ_i D,D$	
16.	$D,C \succ_i C,D \succ_i D,D \succ_i C,C$	
17.	$D,C \succ_i D,D \succ_i C,C \succ_i C,D$	
18.	$D,C \succ_i D,D \succ_i C,D \succ_i C,C$	
19.	$D,D \succ_i C,C \succ_i C,D \succ_i D,C$	
20.	$D,D \succ_i C,C \succ_i D,C \succ_i C,D$	
21.	$D,D \succ_i C,D \succ_i C,C \succ_i D,C$	
22.	$D,D \succ_i C,D \succ_i D,C \succ_i C,C$	
23.	$D,D \succ_i D,C \succ_i C,C \succ_i C,D$	defection dominates
24.	$D,D \succ_i D,C \succ_i C,D \succ_i C,C$	defection dominates

A night of indecision follows. As you anticipate your parents' and teachers' reactions. . .you start wondering if your friend is really going to go through with the plan.

Not that you do not want the plan to succeed: the best possible outcome would be for both of you to get the haircut.

The trouble is, it would be awful to be the *only* one to show up with the haircut. That would be the worst possible outcome.

You're not above enjoying your friend's embarrassment. If you *didn't* get the haircut, but the friend did, and looked like a real jerk, that would be almost as good as if you both got the haircut.

This scenario is obviously very close to the prisoner's dilemma: the difference is that in this scenario, mutual cooperation is the most preferred outcome, rather than you defecting while your opponent cooperates. Expressing the game in a payoff matrix (picking rather arbitrary payoffs to give the preferences):

	i defects	i cooperates
j defects	1 1	0 2
j cooperates	2 0	3 3

It should be clear that there are *two* Nash equilibria in this game: mutual defection, or mutual cooperation. If you trust your opponent, and believe that he will cooperate, then you can do no better than cooperate, and vice versa, your opponent can also do no better than cooperate. Conversely, if you believe your opponent will defect, then you can do no better than defect yourself, and vice versa.

Poundstone suggests that 'mutiny' scenarios are examples of the stag hunt: 'We'd all be better off if we got rid of Captain ~~Bligh~~, but we'll be hung as mutineers if not enough of us go along' (Poundstone, 1992, p. 220).

The game of chicken

The game of chicken (row 13 in Table ~~6.1~~) is ~~characterized~~ by agent i having the following preferences:

$$D,C \succ_i C,C \succ_i C,D \succ_i D,D.$$

As with the stag hunt, this game is also closely related to the prisoner's dilemma. The difference here is that mutual defection is agent i's most feared outcome, rather than i cooperating while j defects. The game of chicken gets its name from a rather silly, macho 'game' that was supposedly popular amongst juvenile delinquents in 1950s America; the game was immortalized by James Dean in the film *Rebel Without a Cause*. The purpose of the game is to establish who is ~~bravest~~ out of two young thugs. The game is played by both players driving their cars at high speed towards a cliff. The idea is that the least brave of the two (the 'chicken') will be the first to drop out of the game by steering away from the cliff. The winner is the one who lasts longest in the car. Of course, if *neither* player steers away, then both cars fly off the cliff, taking their foolish passengers to a fiery death on the rocks that undoubtedly lie below.

So, how should agent i play this game? It depends on how brave (or foolish) i believes j is. If i believes that j is braver than i, then i would do best to steer away from the cliff (i.e. cooperate), since it is unlikely that j will steer away from the cliff. However, if i believes that j is *less* brave than i, then i should stay in the car; because j is less brave, he will steer away first, allowing i to win. The difficulty arises when both agents mistakenly believe that the other is less brave; in this case, both agents will stay in their car (i.e. defect), and the worst outcome arises.

Expressed as a payoff matrix, the game of chicken is as follows:

	i defects	*i* cooperates
j defects	0 0	1 3
j cooperates	3 1	2 2

It should be clear that the game of chicken has two Nash equilibria, corresponding to the above-right and below-left cells. Thus if you believe that your opponent is going to drive straight (i.e. defect), then you can do no better than to steer away from the cliff, and vice versa. Similarly, if you believe your opponent is going to steer away, then you can do no better than to drive straight.

[6.7 Dependence Relations in Multiagent Systems

Before leaving the issue of interactions, I will briefly discuss another approach to understanding how the properties of a multiagent system can be understood. This approach, due to Sichman and colleagues, attempts to understand the *dependencies* between agents (Sichman *et al.*, 1994; Sichman and Demazeau, 1995). The basic idea is that a dependence relation exists between two agents if one of the agents requires the other in order to achieve one of its goals. There are a number of possible dependency relations.

Independence. There is no dependency between the agents.

Unilateral. One agent depends on the other, but not vice versa.

Mutual. Both agents depend on each other with respect to the same goal.

Reciprocal dependence. The first agent depends on the other for some goal, while the second also depends on the first for some goal (the two goals are not necessarily the same). Note that mutual dependence implies reciprocal dependence.

These relationships may be qualified by whether or not they are *locally believed* or *mutually believed.* There is a locally believed dependence if one agent believes the dependence exists, but does not believe that the other agent believes it exists. A mutually believed dependence exists when the agent believes the dependence exists, and also believes that the other agent is aware of it. Sichman and colleagues implemented a *social reasoning system* called DepNet (Sichman *et al*, 1994). Given a description of a multiagent system, DepNet was capable of computing the relationships that existed between agents in the system.

Notes and Further Reading

Ken Binmore, in his lucid and entertaining introduction to game theory, *Fun and Games,* discusses the philosophical implications of the prisoner's dilemma at length (Binmore, 1992, p. 310-316). This text is recommended as a readable - albeit mathematically demanding - introduction to game theory, which provides extensive pointers into the literature.

There are many other interesting aspects of Axelrod's tournaments that I can only briefly mention due to space restrictions. The first is that of *noise.* I mentioned above that the iterated prisoner's dilemma is predicated on the assumption that the participating agents can see the move made by their opponent: they can see, in other words, whether their opponent defects or cooperates. But suppose the game allows for a certain probability that on any given round, an agent will *misinterpret* the actions of its opponent, and perceive cooperation to be defection and vice versa. Suppose two agents are playing the iterated prisoner's dilemma against one another, and both are playing TIT-FOR-TAT. Then both agents will start by cooperating, and in the absence of noise, will continue to enjoy the fruits of mutual cooperation. But if noise causes one of them to misinterpret defection as cooperation, then this agent will retaliate to the perceived defection with defection. The other agent will retaliate in turn, and both agents will defect, then retaliate, and so on, losing significant utility as a consequence. Interestingly, cooperation can be restored if further noise causes one of the agents to misinterpret defection as cooperation - this will then cause the agents to begin cooperating again! Axelrod (1984) is recommended as a point of departure for further reading; Mor and Rosenschein (1995) provides pointers into recent prisoner's dilemma literature; a collection of Axelrod's more recent essays was published as Axelrod (1997). A non-mathematical introduction to game theory, with an emphasis on the applications of game theory in the social sciences, is Zagare (1984).

Exercises

(1) **[Level 1.]**

Consider the following sets of outcomes and preferences:

- $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$;
- $\omega_6 \succ_i \omega_2 \succ_i \omega_3 \succ_i \omega_1 \succ_i \omega_5 \succ_i \omega_4$;
- $\Omega_1 = \{\omega_1, \omega_3\}$;
- $\Omega_2 = \{\omega_3, \omega_4\}$;
- $\Omega_3 = \{\omega_3\}$; and
- $\Omega_4 = \{\omega_2, \omega_6\}$.

Which of these sets (if any) dominates the others? Where neither set dominates the other, indicate this.

(2) **[Level 2.]**

Consider the following interaction scenarios:

	i defects	*i* cooperates
j defects	3 (j: 3)	4 (j: 2)
j cooperates	1 (j: 1)	2 (j: 4)

	i defects	*i* cooperates
j defects	−1 (j: −1)	2 (j: 1)
j cooperates	1 (j: 2)	−1 (j: −1)

	i defects	*i* cooperates
j defects	3 (j: 3)	4 (j: 2)
j cooperates	1 (j: 1)	2 (j: 4)

Now, for each of these scenarios,

- begin by informally analysing the scenario to determine what the two agents should do;
- classify the preferences that agents have with respect to outcomes;
- determine which strategies arc strongly or weakly dominated;
- use the idea of deleting strongly dominated strategies to simplify the scenario where appropriate;
- identify any Nash equilibria.

(3) [Class discussion.]

This is best done as a class exercise, in groups of three: play the prisoner's dilemma. Use one of the three as 'umpire', to keep track of progress and scores, and to stop any outbreaks of violence. First try playing the one-shot game a few times, and then try the iterated version, first for an agreed, predetermined number of times, and then allowing the umpire to choose how many times to iterate without telling the players.

- Which strategies do best in the one-shot and iterated prisoner's dilemma?
- Try playing people against strategies such as TIT-FOR-TAT, and ALL-D.
- Try getting people to define their strategy precisely in advance (by writing it down), and then see if you can determine their strategy while playing the game; distribute their strategy, and see if it can be exploited.

(4) [Level 2.]

For each of the scenarios in Table 6.1 that was not discussed in the text,

- draw up a payoff matrix that characterizes the scenario (remembering that these are *symmetric* interaction scenarios);
- attempt to determine what an agent should do;
- identify, if possible, a real-world interaction situation that corresponds to the abstract scenario.