Introduction to Game Theory

A. J. Ganesh

Feb. 2013

1 What is a game?

A game is a model of strategic interaction between *agents* or *players*. The agents might be animals competing with other animals for food or territory or mates, they might be firms competing for profits, and so on. The agents have a number of different *actions* among which they can choose. Typically, the benefit or *payoff* that an agent receives for a given action will depend on the actions chosen by other agents. This is the interaction element in the game. The assumption is that players seek to maximise their payoffs. However, it is not obvious how to do so as their payoff depends not only on their own action, but also on the actions of other players. Game theory is the study of the action choices agents might make in a given model of interaction. It is used across a range of disciplines including economics, biology, psychology, engineering and computer science. In the following, we shall confine ourselves to games in which there are a finite number of players, and each player has a finite number of actions (possible different for different players) from which to choose.

Another area of behavioural sciences which seeks to predict agent behaviour is *decision theory*. In decision theory, agents are faced with an uncertain environment (or one about which they have incomplete information) and need to select an action. The key difference between decision theory and game theory is that, in decision theory, the payoff to an agent depends only on its own action and the true state of the environment, and not the actions of other agents. In that sense, there is no interaction between the agents, and no need to think strategically about the possible behaviours of other agents. Decision theory problems are sometimes called "games against

nature", where nature chooses the state of the environment according to some pre-specified probability distribution.

The objective in modelling a social or biological interaction as a game is to try and make predictions about the likely behaviour of the players. It is worth emphasising that game theory is meant to be descriptive, not normative. Even though our behavioural predictions may be reached by reasoning about how players *ought* to act, they are under no moral obligation to do so. If predictions don't match up to observations, then it points to an error, either in the way the interaction has been modelled, or in the assumptions implicit in reasoning about player decision-making.

There are two common ways of describing a game, called the *normal form* (or strategic form) and the *extensive form*. The normal form is more naturally used to describe games in which the players move (choose actions) simultaneously, without knowing the choices of other players, while the extensive form is better suited for games in which players move sequentially and we need to explicitly specify the information available to them before they move. We will mostly focus on normal form games here.

A game in normal form can be described by specifying the set of players, the set of actions (also called pure strategies), and the payoff to each player arising from each possible *strategy profile*, namely, the set of strategies chosen by each of the players. More formally, let us denote the set of players by $I = \{1, 2, 3, \ldots, |I|\}$, the actions available to player i, or its pure strategy space, by S_i, and the payoff to player i corresponding to a pure strategy profile $\mathbf{s} = (s_1, s_2, \ldots, s_{|I|})$ by $u_i(\mathbf{s})$. If there are only two players, the payoffs can be conveniently represented by a matrix, as we now illustrate.

Examples of two-player games

1. *Hawk-Dove game:* There are two players, each of whom can play one of two strategies, Hawk or Dove. The payoff matrix is given below; the first number in each cell is the payoff to the row player, and the second to the column player.

$$
\begin{array}{c}
 \\
H \\
D
\end{array}
\begin{pmatrix}
\phantom{(\frac{u}{2}-c,} H {2}-c)} & D \\
(\frac{u}{2} - c, \frac{u}{2} - c) & (u, 0) \\
(0, u) & (\frac{u}{2}, \frac{u}{2})
\end{pmatrix}
$$

This game is intended to model the following interaction. Two agents are competing to gain access to a resource, which has total value u. If

2

both adopt a conciliatory strategy (dove), then they share the resource equitably. If one is conciliatory while the other is aggressive (plays Hawk), then the Hawk gets all of the resource while the Dove gets nothing. If both play Hawk, they get into a conflict, the outcome of which ends up with sharing the resource equally, but incurring a penalty c for the conflict. Note that c may be smaller or bigger than $u/2$; negative payoffs are allowed.

This could be a simplified model of two animals competing for food or territory. The payoff matrix shown involves assumptions of symmetry. In practice, there may be differences in fitness between the animals, which can be incorporated easily by making the payoffs asymmetric.

2. *Prisoners' dilemma:* A crime was committed jointly by two people, and the perpetrators are caught and questioned separately by the police. If neither confesses to the crime, then the police don't have enough evidence to convict them. If one of them confesses, then both can be convicted, and the confessor will get a much lighter sentence. If both confess, both will get an intermediate sentence. We denote the action space of each player by $\{C, D\}$, where C denotes that the player co-operates with the other (doesn't confess), while D denotes that the player defects (confesses). The payoffs are denoted by

$$
\begin{array}{cc}
 & \begin{array}{cc} C & \hspace{1.5em} D \end{array} \\
\begin{array}{c} C \\ D \end{array} &
\left(\begin{array}{cc}
(0,0) & (-a,-1) \\
(-1,-a) & (-b,-b)
\end{array} \right)
\end{array}
$$

It would typically be assumed that $1 < b < a$.

3. *Stag hunt:* Two hunters must decide simultaneously whether to hunt for stag or for hare. It requires co-operation to catch a stage, so they will only succeed if both decided to hunt for stag. If that happens, they share it equally. On the other hand, whether both or either decide to hunt for hare, they will catch one hare each, and get to keep it. A stag is more valuable and provides $2x$ times as much food as a hare, for some $x > 1$. So the payoff matrix is given by

$$
\begin{array}{cc}
 & \begin{array}{cc} S & \hspace{1.5em} H \end{array} \\
\begin{array}{c} S \\ H \end{array} &
\left(\begin{array}{cc}
(x,x) & (0,1) \\
(1,0) & (1,1)
\end{array} \right)
\end{array}
$$

Though it was not modelled as a game, this particular example was discussed by Rousseau in his *Discourse on the Origin and Basis of Equality among Men.*

4. *Parental Investment game:* The parents of a brood have a choice between staying and caring for the brood, or leaving and trying to mate again. If both parents choose to stay, the brood is certain to survive, and provides a payoff of 1, say, to each parent. If only one parent stays, the brood survives with probability $p < 1$, and provides an expected payoff p to *each parent.* In addition, the parent who left can get an expected payoff q from possibly mating again. A parent who stays behind cannot mate again in that season. Thus, with actions denoted C (stay behind or co-operate) and D (defect or leave), the pay-off matrix is given by

$$
\begin{array}{cc}
 & \begin{array}{cc} C & \quad\quad D \end{array} \\
\begin{array}{c} C \\ D \end{array} &
\left(\begin{array}{cc}
(1,1) & (p, p+q) \\
(p+q, p) & (p, p)
\end{array} \right)
\end{array}
$$

5. *Paper-Scissors-Rock:* This is a popular children's game. There are two players, and they have to simultaneously choose an action from the set $\{P, S, R\}$. Scissors cut paper, paper covers rock, and rock blunts scissors. The payoff is 1 if you win, -1 if you lose, and 0 for a draw (both players choose the same action). Thus, the payoff matrix is

$$
\begin{array}{cc}
 & \begin{array}{ccc} P & \quad\; S & \quad\; R \end{array} \\
\begin{array}{c} P \\ S \\ R \end{array} &
\left(\begin{array}{ccc}
(0,0) & (-1,1) & (1,-1) \\
(1,-1) & (0,0) & (-1,1) \\
(-1,1) & (1,-1) & (0,0)
\end{array} \right)
\end{array}
$$

6. *Cournot competition:* This is an example where the action space is not finite. There are two (or more) firms, each of which has to decide what quantity $q_i \geq 0$ of a good to produce in a certain period. The cost for firm i to produce quantity q_i is given by some cost function $c_i(q_i)$, which is an increasing function. At the end of the period, once the goods are produced, the market price is a decreasing function of the total supply, $p(q_1 + q_2 + \ldots + q_{|I|})$. The payoff to player i is its profit, $q_i p(\sum_j q_j) - c_i(q_i)$.

7. *Congestion games* This is not a single game, but a wide class of games sharing certain common features. Games of this type are widely used

to model transport and communication networks. Let us illustrate with an example of a transport network. There are agents who want to travel between different origin-destination pairs in the network. The actions available to them are a choice of different routes between their origin and destination. The actions chosen by the different agents result in different amounts of traffic flowing along different links in the network. Each link has a delay or cost function associated with it, which specifies how much delay each unit of traffic flowing along it will encounter, as a function of the total traffic on the link. These delay functions, together with the knowledge of which links make up each route, can be used to work out the total delay incurred by each agent. The payoff to an agent is taken to be the negative of the delay or cost incurred by the agent.

It turns out that congestion games are rather special. In general, agents in games pull in different directions. What is good for one agent may turn out to be bad for another (Paper Scissors Rock being a classic example), and so it is hard to determine the stable outcomes of the game. But in congestion games, it turns out that we can identify a global objective function, or potential function, and that action choices that minimise this function also maximise the payoffs (or minimise the delays) for the individual agents. In a certain sense, such games are naturally co-operative; agents who act in their own self-interest end up helping other agents too. The equilibrium outcomes of these games can be determined by solving the associated global optimisation problem. Morevoer, many adaptive algorithms have good convergence properties for this class of games.

Some of the above examples may appear simplistic, but they nevertheless capture many of the essential features of strategic interactions. As an example, Prisoner's Dilemma illustrates the benefits of co-operation and Hawk-Dove the cost of conflict, but also illustrate the benefits to individuals of deviating from this "social optimum". Paper-Scissors-Rock models situations like the co-evolution of pathogens and immune systems.

Next, we shall briefly discuss a few examples of games in extensive form. In fact, we can use the same examples as above. First, consider the Hawk-Dove game. Suppose $c > u/2$, so that conflict is penalised heavily. Now, if one of the players gets to move first, and decides to play Dove, the second mover has a choice between playing Hawk and getting a payoff of u, or playing Dove and getting $u/2$. Clearly, under the assumption of rationality, the

second player will choose Hawk and the first player will get a payoff of 0. On the other hand, if the first player chooses to play Hawk, then the second player will be forced to choose Dove (as the resulting pay-off of 0 is bigger than the payoff of $(u/2) - c$ to playing Hawk), and the first player will get u. Hence, the first player will always choose to play Hawk. On the other hand, if $c < u/2$, the first player will always choose to play Dove. In the next section, we will see that this is quite different to what happens if both players move simultaneously.

Next, consider the parental investment game again but suppose that the decisions are made sequentially rather than simultaneously. If one of the parents is given *first-mover advantage*, then that radically changes the likely behaviours. If the first player defects, then the actions available to the second player have payoffs p for staying back and raising the brood, and q for abandoning it and attempting to mate again. If $p > q$, it is rational for the second player to raise the brood. Knowing this, the payoffs to the first player are 1 for staying back and sharing the raising of the brood, and $p + q$ for abandoning it. If $p + q > 1$, then it is rational for the first player to always abandon the brood. This is very different from the setting where the decisions are made simultaneously. In fact, the parental investment game can be generalised to consider multiple periods, with choices being made simultaneously in each period. This is a more realistic model of the situation faced by many birds or animals, which have the choice to abandon at any time in the season. Abandoning later increases the chances of survival of the young, but decreases the chances of another successful mating in that season.

We finish this section with one final example, known as the *Ultimatum Game*. This is a two-player game. Player 1 is given a fixed sum of money, say £100, to be shared between the players. He can make any proposal to player 2, say a 50-50 split, or a 90-10 split in his favour, or anything else. If player 2 accepts the proposed split, say 90-10, then both players get the corresponding amounts, £90 and £10 in this case. If player 2 rejects the offer, both players get 0.

2 Solution concepts for games

The examples above demonstrate that there is typically no single action that unambiguously maximises a player's payoff irrespective of what other

players do. How then should players choose their actions? It is natural to expect that, in doing so, players should try and anticipate the possible choices of other players. A first standard assumption in game theory is that the action spaces and payoffs of all players are common knowledge. Here, *common knowledge* is a technical term which means not only that all players know them, but also that they know that other players know them, know that other players know that they know them, and so on ad infinitum. A second standard assumption is that players are perfectly rational and have unbounded computational power, and know that other players are the same. These assumptions may seem unrealistic to the point of absurdity. But we will start with them, because they allow us to make testable predictions. Later, we will think about how they might be weakened.

The first solution concept we will look at is that of a minimax or maximin strategy. Here, a player would like to maximise its payoff in the worst case situation, i.e., corresponding to the worst-case strategy choices of the other players. In certain kinds of games, it turns out that this leads to an equilibrium - the strategy chosen by each player based on this reasoning leads to the worse-case choice for the others.

A specific class of games we look at are what are called *two-person zero sum games*. An example is Paper Scissors Rock. If you look at the payoff matrix, you can see that for any actions chosen by the two players, the sum of their payoffs is always zero. In other words, what one player wins is exactly what the other loses. For this class of games, it turns out that there is an equilibrium involving maximin strategies. But before we show this, we need to expand the definition of the strategy space.

Our discussion so far has centred on the deterministic choice of individual actions by players. These are also called *pure strategies*. But players also have the option of choosing randomly from the available actions, possibly with different probabilities. If the actions available to player i are specified by a finite set S_i, then a probability distribution p_i on S_i specifies what we call a *mixed strategy*. Thus, for instance, in Paper Scissors Rock, a mixed strategy for player 1 is the probability distribution $(0.1, 0.6, 0.3)$ on (PSR), while $(0.8, 0.2, 0)$ is a different mixed strategy for player 2. The strategy space for player i is thus not just the actions in S_i, but all probability distributions on S_i.

In a two person game, let $\mathbf{p}_1$ and $\mathbf{p}_2$ denote the strategies of players 1 and 2. Let $u_1(x, y)$ and $u_2(x, y)$ denote the payoffs to the two players correspond-

ing to the pure strategy profile (x, y), i.e., if they choose actions x and y respectively. The, the expected payoff to player $i \in \{1, 2\}$ under the mixed strategy profile $(\mathbf{p}_1, \mathbf{p}_2)$ is given by

$$u_i(\mathbf{p}_1, \mathbf{p}_2) = \sum_{x \in S_1} \sum_{y \in S_2} p_1(x) p_2(y) u_i(x, y). \tag{1}$$

Note that, in writing the probability that player 1 chooses action x and player 2 chooses action y as $p_1(x) p_2(y)$, we are implicitly assuming that these choices are independent. This is natural if we assume that these choices are made simultaneously, and hence cannot be co-ordinated.

In the maximin setting, each player tries to maximise its own payoff. Thus, player 1 seeks to choose $\mathbf{p}_1$ to solve the problem

$$\max_{\mathbf{p}_1} \min_{\mathbf{p}_2} u_1(\mathbf{p}_1, \mathbf{p}_2), \tag{2}$$

while player 2 seeks to solve

$$\max_{\mathbf{p}_2} \min_{\mathbf{p}_1} u_2(\mathbf{p}_1, \mathbf{p}_2). \tag{3}$$

If it is a zero-sum game, then $u_2(\mathbf{p}_1, \mathbf{p}_2) = -u_1(\mathbf{p}_1, \mathbf{p}_2)$, so the goal of the second player can be rewritten as

$$\max_{\mathbf{p}_2} \min_{\mathbf{p}_1} -u_1(\mathbf{p}_1, \mathbf{p}_2) = - \min_{\mathbf{p}_2} \max_{\mathbf{p}_1} u_1(\mathbf{p}_1, \mathbf{p}_2). \tag{4}$$

Expanding out $u_1(\mathbf{p}_1, \mathbf{p}_2)$ according to (1), it can be shown that (2) and (3) are linear programming problems; in fact, they are primal and dual versions of the same problem and so, by linear programming duality, they have the same solution for $\mathbf{p}_1$ and $\mathbf{p}_2$. Such a combination of mixed strategies, which simultaneously solve the maximin problem for both players, is called a saddle point of the game.

Let us return to the example of Paper Scissors Rock. If $\mathbf{p}_1$ and $\mathbf{p}_2$ denote the strategies of the two players, then the payoff to player 1 is given by

$$u_1(\mathbf{p}_1, \mathbf{p}_2) = p_1(P)[p_2(R) - p_2(S)] + p_1(S)[p_2(P) - p_2(R)] + p_1(R)[p_2(S) - p_2(P)].$$

Now consider the strategy for player 1 which assigns equal probability $1/3$ to each of the possible actions. It is easy to see from above that the expected payoff to player 1 is zero, irrespective of the strategy $\mathbf{p}_2$ chosen by player 2. In particular, its payoff is zero in the worst case.

Is there some other strategy for player 1 that can do better in the worst case, i.e., guarantee a strictly positive payoff no matter what strategy player 2 chooses? If so, then, by the symmetry of the game, there must also be a strategy for player 2 that guarantees a strictly positive payoff against any strategy of player 1. Now suppose the players play these two strategies against each other. Then they both gain strictly positive expected payoffs, and so the sum of the expected payoffs is strictly positive. But this is a contradiction because Paper Scissors Rock is a zero sum game, i.e., the sum of payoffs is zero for any combination of actions. Hence, the expected sum of payoffs is also zero.

We have thus shown that the strategy which assigns equal probability $1/3$ to each player is a maximin strategy for player 1. By symmetry, it is also a maximin strategy for player 2. If someone chose a different strategy, say playing Paper much more often, then an opponent who knew this (even if the opponent didn't know what they would play in a specific instance of the game) could exploit it to their advantage by giving greater weighting to Scissors. Thus, if the game were played repeatedly, such strategies would be penalised, and only the balanced strategy would be observed in equilibrium.

The reasoning in the last paragraph is rather loose. We have not defined equilibrium, or precisely defined the setting of a repeated game, but all we are trying to do at this point is to give some intuitive justification for the typical behaviours we might expect to observe in a population of agents playing this game.

Next, let us consider a non-zero sum game, say the Hawk-Dove game. Take $u = 10$ and $c = 8$ for concreteness. Then the strategy of playing Hawk gets a worst case payoff of -3, while that of playing Dove gets a worst case payoff of 0 (the worst case being when the opponent always plays Hawk). Thus, in this case, the maximin strategy is to play Dove. If both players adopted this strategy, then they would both get a payoff of 5. However, intuitively, this situation is unstable. If one of the players knows that the other is highly likely to play Dove, and is willing to risk a small chance of an unfavourable outcome, then that player can gain a payoff of 10 by playing Hawk.

By focusing on the maximin setting, we are implicitly assuming that the players are extremely risk-averse, which is not a realistic assumption. In the case of zero-sum games, this assumption is not needed. It turns out that the maximin choice is natural, and neither player can do better by *deviating unilaterally* from it (unlike in the Hawk-Dove game, for instance). This leads

us naturally to our next solution concept.

Nash equilibrium Consider a finite set of players $I = \{1, 2, \ldots, |I|\}$, with corresponding finite pure strategy sets $S_1, S_2, \ldots, S_{|I|}$, and a payoff function $\mathbf{u} = (u_1, u_2, \ldots, u_{|I|})$. Here $u_j(s_1, s_2, \ldots, s_{|I|})$ denotes the payoff to user j if the different players choose actions $s_1, s_2, \ldots, s_{|I|}$. A mixed strategy profile $(\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_{|I|})$ is called a *Nash equilibrium* of this game if no player has an incentive to deviate unilaterally from this profile, i.e., if

$$u_i((\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_{|I|}) \geq u_i(\mathbf{q}_i, \mathbf{p}_{-i}) \tag{5}$$

for all other probability distributions $\mathbf{q}_i$ on S_i. The notation $\mathbf{p}_{-i}$ denotes that all players j other than i continue to play the strategy $\mathbf{p}_j$.

This equilibrium concept was defined by John Nash, and is appropriate for all games, not just zero-sum games. However, does every game have such an equilibrium strategy profile? The answer turns out to be yes.

Theorem 1 *Every game with a finite number of players, and a finite pure strategy space for each player, has a Nash equilibrium in mixed strategies.*

We now take the Hawk-Dove game as an example, and show how to work out the Nash equilibrium. Suppose first that $c > u/2$. Let $\mathbf{p} = (p_H, p_D)$ denote the strategy adopted by player 1. What strategy should player 2 adopt in response? Let us work out the expected payoff to player 2 for the two pure strategies, Hawk and Dove. We have

$$\mathbb{E}[u_2(\mathbf{p}, H)] = p_H\left(\frac{u}{2} - c\right) + p_D u = \frac{u}{2} - c + p_D\left(\frac{u}{2} + c\right),$$

since $p_H + p_D = 1$, whereas

$$\mathbb{E}[u_2(\mathbf{p}, D)] = p_H \cdot 0 + p_D\frac{u}{2} = p_D\frac{u}{2}.$$

Comparing the payoffs, we see that player 2 should always play Hawk if

$$p_D c > c - \frac{u}{2}, \text{ i.e., } p_D > 1 - \frac{u}{2c}, \text{ i.e., } p_H < \frac{u}{2c},$$

and always play Dove if $p_H > u/(2c)$. Note that $u/(2c)$ lies between 0 and 1 by the assumption that $c > 2u$. If $p_H = u/(2c)$, then player 2 is indifferent between the two alternatives, and hence also between all mixed strategies. In particular, the mixed strategy $p_H = \frac{u}{2c}$, $p_D = 1 - \frac{u}{2c}$ is as good as any.

We have thus shown that these choices of p_H and p_D by each of the two players constitute a Nash equilibrium, and in fact constitute the only Nash equilibrium.

On the other hand, if $c < u/2$, then similar reasoning shows that the only Nash equilibrium is for both players to play Hawk. Note that, as a consequence, both players only get a payoff as $(u/2) - c$, whereas they could have got a bigger payoff of $u/2$ if they had both played Dove. This example demonstrates that a Nash equilibrium can be strictly worse for every player than some other strategy combinations. Unfortunately, those other combinations are not equilibria, and selfish agents cannot sustain them. A similar outcome obtains in the Prisoners Dilemma. It is left as an exercise to you to check that the only Nash equilibrium is for both players to Defect.

In Paper Scissors Rock, the maximin solution we obtained earlier, which assigned equal probability $1/3$ to each action, is a Nash equilibrium, and the only one for this game. This is also left to you as an exercise. In fact, it is more generally true in two-person zero sum games that the maximin solution (saddle point) is a Nash equilibrium.

A Nash equilibrium is appealing as a description of the likely behaviour of players in a game. Indeed, if all players are perfectly rational, then no player should expect any of their opponents to play a non-equilibrium strategy because a non-equilibrium strategy admits improvement. By this reasoning, the Nash equilibria are the only stable outcomes. Indeed, it is often the case that they are useful in predicting likely player behaviour. However, som qualifications are needed.

The first is that the theorem does not tell us that there is a unique Nash equilibrium. A game may have many Nash equilibria, even infinitely many. Sometimes, one of these may be much more "natural" than others, or more stable under small perturbations. But sometimes, there may be no good way of choosing one Nash equilibrium over another. In this case, the model does not yield precise predictions, or yields a number of different predictions. If this number is small, it could still be useful, as experiment can determine which of the equilibria are operative in the setting under study.

A second qualification is that the assumptions implicit in arguing that players can find a Nash equilibrium may be unrealistic. In the simple examples of games we have considered, it is plausible that agents with moderate cognitive resources can find the Nash equilibria, at least approximately. But there are games where this is not the case. There is a branch of computer science

dealing the computational complexity of finding Nash equilibria, and there are games in which this is provably hard, in the sense of NP-completeness. In such cases, a Nash equilibrium may not be a suitable predictor of the behaviour of human agents, much less birds or ants.

We will now address each of these objections in turn. First we present a refinement of the Nash equilibrium concept, known as an *evolutionarily stable strategy* (ESS). The motivation for this concept is the following. Consider, say, a two-player game, and suppose that there is a large population of agents which interact pairwise according to it. The Hawk-Dove game is a good example to keep in mind. The agents themselves may be quite simple, and incapable of sophisticated game theoretic reasoning. In fact, the strategy of each individual agent may be genetically pre-programmed; for example, a single agent may always play Hawk or always play Dove. We can then ask how frequently a given strategy is represented in the population, in equilibrium. The payoffs in this game are to be interpreted as a measure of biological fitness, say the mean number of offspring left by that individual in the next generation. The other ingredient needed by the model is how to pair up individuals within the population to play the game. The simplest is to assume that each agent plays against another agent chosen uniformly at random from the population. If the proportions of different strategies in the population are such that one strategy is better than another, then that strategy will become more prevalent in succeeding generations. Thus, at equilibrium, each surviving strategy (i.e., each strategy played by a non-zero fraction of the population) must be equally good, and at least as good as those strategies not played by anybody. Let us now formalise this intuition.

Consider a symmetric two-player game. By symmetric, we mean that the action sets of both players are the same and $u_1(x, y) = u_2(y, x)$ for any two actions x and y. In words, if we switch the actions of the two players, then we also switch their payoffs. This is the most natural setting for considering an ESS as we assume that there is a large population, but individuals with the population are identical, i.e., interchangeable. The definition can be extended to allow for a fixed number of player types, with specified fractions of the population belonging to each of the types. As this doesn't change anything of substance, but makes the notation more complicated, we will confine ourselves to the single type setting. In this setting, we say that a mixed strategy $\mathbf{p}$ is an ESS if

$$u_1(\mathbf{p}, \mathbf{p}) > u_1(\mathbf{q}, \mathbf{p}) \text{ or} \tag{6}$$
$$u_1(\mathbf{p}, \mathbf{p}) = u_1(\mathbf{q}, \mathbf{p}) \text{ and } u_1(\mathbf{p}, \mathbf{p}) > u_1(\mathbf{q}, \mathbf{q}) \text{ for all } \mathbf{q} \neq \mathbf{p}.$$

Pure strategies are included in this definition because the strategy x is the same as the mixed strategy which assigns probability 1 to x and probability 0 to all other actions.

Let us deconstruct what the above equation says, using the Hawk-Dove game as an example. Let p_H and p_D denote the population-level frequencies of the Hawk and Dove strategies respectively. It doesn't matter exactly how these frequencies are realised, whether by each individual agent playing Hawk and Dove with those probabilities, or by a fraction p_H always playing Hawk and the residual fraction p_D always playing Dove, or some intermediate situation. We assume that, in each time period, each agent is paired up with another agent chosen uniformly at random from the population, that they play one instance of the Hawk-Dove game, and that the resulting payoff determines the number of offspring they have in the next time step. A higher payoff implies a higher fitness, hence more offspring. In particular, if there was a strategy $\mathbf{q}$ such that $u_1(\mathbf{q}, \mathbf{p}) > u_1(\mathbf{p}, \mathbf{p})$, then a mutant that played strategy $\mathbf{q}$ would have higher fitness than the population average, and would increase in prevalence in successive generations. Thus, such a mutant would be able to invade, implying that $\mathbf{p}$ cannot be a stable equilibrium. (Even a small perturbation, a tiny fraction of type $\mathbf{q}$ mutants, would eventually shift the population away from $\mathbf{p}$.) But this argument only tells us that the first inequality in (6) cannot be reversed. Why can we not have equality? Indeed, if we compare the definition of an ESS in (6) with that of a Nash equilibrum in (5), we observe that a key difference is that the inequality in (6) is strict. To see why this is so, consider the second line in the definition. This tells us that, if we have equality in the first line, we need an additional condition, namely that $u_1(\mathbf{p}, \mathbf{p}) > u_1(\mathbf{q}, \mathbf{q})$. Suppose that strategy $\mathbf{q}$ is fitness neutral, i.e., of the same fitness as $\mathbf{p}$ when playing against $\mathbf{p}$. Now, if there is a small fraction of $\mathbf{q}$ mutants, then, very rarely, a $\mathbf{q}$ will be paired up against another $\mathbf{q}$. If they can gain additional fitness from these rare interactions, then they will still be able to invade the population. The conditions in the definition of an ESS say that there is no such strategy, i.e., that an ESS cannot be invaded by any mutant strategy.

It can be seen by comparing the definitions of Nash equilibria and evolutionarily stable strategies that the requirements for an ESS are more stringent. In other words, every ESS is a Nash equilibrium, but a Nash equilibrium may not be an ESS. If a game has too many Nash equilibria, the ESS concept may be helpful in deciding which of these is a more likely outcome of a game. However, a disadvantage is that, while every game is guaranteed to

have a Nash equilibrium, a game may have no ESS.

3 Adaptive dynamics

A criticism of game theoretic models that we raised in the last section is that it is unrealistic to expect agents with limited cognitive capabilities to arrive at Nash equilibrium strategies purely by ratiocination. Given this, are Nash equilibria of much value at all as solution concepts?

Many of the real-world interactions that we are interested in modelling as games are not engaged in by the agents concerned on a one-off basis. Rather, they are repeated, even everyday, occurences in the lives of these agents. Therefore, it might be argued that agents learn Narch equilibrium strategies by trial and error rather than by pure reason. Can we figure out rules of thumb, or adaptation or learning mechanisms, that are simple enough to be plausible, but which have the property of converging to Nash equilibria? And if so, how quickly do the converge? These are the questions we shall address in this section. In the remainder of this section, we will consider the same game, played repeatedly by the same set of players, following some simple adaptation rule.

Best response dynamics The players start out playing an arbitrary strategy in tiem step 0. Each player can observe the past actions of all other players. In every subsequent time step n, each player i plays a best response (one with highest payoff), to the observed actions of the other players in time step $n - 1$; if there is more than one such action, ties are broken arbitrarily. Note that this update rule is "myopic", in that it only looks at the last time step, but nevertheless fairly natural.

Best response dynamics are not guaranteed to converge. Indeed, Paper Scissors Rock provides an obvious counterexample, where the strategies of the two players cycle indefinitely. However, if the dynamics converge, then the limit point is guaranteed to be a Nash equilibrium. This is because the limiting strategy profile is necessarily a best response to itself. But this is the defining condition of a Nash equilibrium.

Fictitious play This is, like best response, an attempt to do the best thing in hindsight. But where best response dynamics only looks at the last time step, fictitious play looks at the entire past. Each agent keeps track of the fraction of times in the past that every other agent has played each action

in its pure strategy space. It then pretends that these observed fractions represent exactly the mixed strategies adopted by the other players, and selects its best response to these mixed strategies. In practice, other players are also adapting their strategies, so their most recent actions might be more indicative of their current strategies than actions in the distant past. However, fictitious play does not attempt to account for this. Nevertheless, by not putting all its weight on the single most recent step, but taking into account more information from the past, it seems to do somewhat better than best response dynamics.

Fictitious play is not guaranteed to converge in arbitrary games. However, it has been shown that it does converge in congestion games, which are a practically important class of games. Moreover, if they converge at all, then the limit point is a Nash equilibrium, for much the same reason as with best response dynamics. If fictitious play converges, then eventually only strategies very close to the limit strategy are played. Hence, the observed average of other players' action choices are very nearly the fractions specified in the limiting strategy. Consequently, the limiting strategy is a best response to itself, which is the defining condition of a Nash equilibrium.

Replicator dynamics The last example we shall look at is motivated by the same reasoning as that behind the notion of an evolutionarily stable strategy (ESS). As in our discussion of the ESS, we will assume that the agents play a symmetric two-person game. As described then, this assumption can be relaxed by introducing a finite number of agent types, one for each player in a general non-symmetric game (albeit one with finitely many players and actions).

We assume that there is a large number of agents, and that these are paired randomly in each time step. We can analyse the dynamics in the limit when the number of agents tends to infinity, by considering the law of large numbers behaviour of the resulting system. Let $x_i(t)$ denote the fraction of agents playing strategy i at time step t. The expected payoff to each such player is $\sum_j x_j u_1(i, j)$ because x_j is the probability that it is paired up with an agent who plays j, and $u_1(i, j)$ is the payoff in that case. The sum is taken over all actions j in the pure strategy space S (which we don't index by the player number because it is the same for all players). Let us define

$$f_i(\mathbf{x}) = \sum_{j \in S} x_j u_1(i, j)$$

to be the expected payoff to, or fitness of, strategy i in environment $\mathbf{x}$. Here,

the vector $\mathbf{x}$ denotes the proportions of different strategies prevalent in the environment, and is what determines the fitness of any given strategy. Let us also denote the average fitness by

$$\phi(\mathbf{x}) = \sum_{i \in S} x_i f_i(\mathbf{x}).$$

Then, any strategies that are more fit than average will increase in prevalence in subsequent generations, while those that are less fit than average will decrease in prevalence. But this will change the environment, and hence also the fitness of different strategies.

How exactly should we model this change of prevalence from one generation to the next? We could try to construct a discrete-time model as in the previous two examples, but replicator dynamics have traditionally been described by the following continuous-time model:

$$\frac{d}{dt} x_i(t) = x_i(t)[f_i(\mathbf{x}(t)) - \phi(\mathbf{x}(t))]. \tag{7}$$

As with our previous adaptive schemes, these dynamics are not guaranteed to converge. Moreover, the replicator equations are nonlinear differential equations, and it is hard to say much about them in general. However, they are quite straightforward to write down, and to simulate for specific games.