3. Scope and Method

This thesis provides some general results for *n*-player games; however, most of the research has been focused on 2-player 2-strategy (2x2) games. In several cases, it has been convenient to illustrate the obtained findings using specific types of 2x2 games, and for that purpose I have often selected 2x2 social dilemma games[7]. The first section of this chapter explains what social dilemmas are and how they can be formalised as 2x2 games; it also gives a brief account of some of the most relevant results obtained within each of the main branches of deductive game theory on the most famous 2x2 social dilemma, i.e. the Prisoner's Dilemma, and of how these results relate to empirical findings. The second section of this chapter outlines the range of formal methods that have been used to analyse the models developed in this thesis.

3.1. Social dilemmas

Social dilemmas are social interactions where everyone enjoys the benefits of collective action, but any individual would gain even more without contributing to the common good (provided that the others do not follow her defection). Social dilemmas are by no means exclusive to human interactions: in many social contexts, regardless of the nature of their component units, we find that individual interests lead to collectively undesirable outcomes for which there is a feasible alternative where every individual would be better off. The problem of how to promote cooperation in these situations without having to resort to central authority has been fascinating scientists from a broad range of disciplines for decades. The value of understanding such a question is clear: in the social and biological sciences, the emergence of cooperation is at the heart of subjects as diverse as the first appearance of life, the ecological functioning of countless environmental interactions, the efficient use of natural resources, the development of modern societies, and the sustainable stewardship of our planet. From an engineering point of view, the problem of understanding how cooperation can emerge and be promoted is crucial for the design of efficient decentralized systems where collective action can lead to a common benefit despite the fact that

[7] In chapter 5 I also investigate an *n*-player social dilemma.

individual units may (purposely or not) undermine the collective good for their own advantage.

At the most elementary level, social dilemmas can be formalised as two-person games where each player can either cooperate or defect. For each player i, the payoff when they both cooperate (R_i, for *Reward*) is greater than the payoff obtained when they both defect (P_i, for *Punishment*); when one cooperates and the other defects, the cooperator obtains S_i (*Sucker*), whereas the defector receives T_i (*Temptation*). Assuming no two payoffs are equal, the essence of a social dilemma is captured by the fact that both players prefer any outcome in which the opponent cooperates to any outcome in which the opponent defects ($\min(T_i, R_i) > \max(P_i, S_i)$), but they both can find reasons to defect. In particular, the temptation to cheat (if $T_i > R_i$) or the fear of being cheated (if $S_i < P_i$) can put cooperation at risk. There are three well-known social dilemma games: Chicken, Stag Hunt, and the Prisoner's Dilemma. In Chicken the problem is greed but not fear ($T_i > R_i > S_i > P_i$; $i = 1, 2$); in Stag Hunt, the problem is fear but not greed ($R_i > T_i > P_i > S_i$; $i = 1, 2$); and finally, both problems coincide in the paradigmatic Prisoner's Dilemma ($T_i > R_i > P_i > S_i$; $i = 1, 2$).

Social dilemmas have been studied from different perspectives, including empirical approaches (both experimental and field studies), discursive theoretical work, game theory, and computer simulation. Within the domains of these four approaches much of the work has been devoted to the study of the Prisoner's Dilemma (PD) or variations of it, often leading to conflicting conclusions (particularly relevant is the conflict between empirical work and classical game theory).

The most widespread results about the PD come from classical game theory. When the PD is played once by instrumentally rational agents, the expected outcome is bilateral defection: rational players do not cooperate since there is no belief that a player could hold about the other player's strategy such that it would be optimal to cooperate (the cooperative strategy is strictly dominated by the strategy of defecting). The situation is very different when the game is played repeatedly. In the (finite or infinitely) repeated game, the range of possible strategies and outcomes is much wider and defecting in every round is no longer a dominant strategy. In fact, in the repeated PD, there is not necessarily one best strategy irrespective of the opponent's strategy. As an example, Kreps et al. (1982) showed that a cooperative outcome can be sustained in the finitely repeated PD if a rational player believes that there is at least a small probability that the other player is playing "Tit for Tat" (TFT)[8].

Since assuming players are instrumentally rational is not enough to narrow the set of solutions of the repeated PD sufficiently, common knowledge of rationality is brought into play. Assuming common knowledge of rationality it can be proved using backwards induction that a series of bilateral defections is the only possible outcome of the finitely repeated PD (Luce and Raiffa, 1957)[9]. Put differently, any two strategies which are an optimal response to each other necessarily lead to a series of bilateral defections in the finitely repeated game. However, when the number of rounds is not limited in advance, a very wide range of possible outcomes where the two players are responding optimally to each other's strategy still exists, even when assuming that the two players have detailed pre-planned strategies and these are common knowledge. Specifically, the "Folk Theorem" states that any individually-rational outcome[10] can be a Nash equilibrium in the infinitely-repeated PD if the discount rate of future payoffs is sufficiently close to one. In this case, orthodox game theory has little to say about the dynamics leading a set of players to one among many possible equilibria.

When classical game theoretical solutions of the PD and related games have been empirically tested, disparate anomalies have been found (see, for example, work reviewed by Colman (1995) in chapters 7 and 9, Roth (1995), Ledyard (1995), and Camerer (2003)). Generally, empirical studies have found that there is a wide variety of factors in addition to economic payoffs that affect our behaviour, and also that, while it is not easy to establish cooperation, levels of cooperation tend to

[8] This is the strategy consisting of starting by cooperating, and thereafter doing what the other player did on the previous move.

[9] For a detailed analysis of the finitely repeated Prisoner's Dilemma, see Raub (1988).

[10] An outcome giving each player at least the largest payoff that they can guarantee receiving regardless of the opponents' moves.

be higher than those predicted by classical game theory (see e.g. Dawes and Thaler, 1988). The explanation of the clash between classical game theory and empirical evidence is, of course, that the assumptions required to undertake a game theoretical analysis do not hold: economic payoffs do not readily correspond to preferences (e.g. considerations of fairness frequently influence behaviour); actual preferences are sometimes neither consistent nor static nor context-independent; players' cognitive capabilities are indeed limited, and players' assumptions of others' preferences and rationality assumed by game theory are therefore often wrong.

Research on the PD within evolutionary game theory was boosted by the computer simulations and empirical studies undertaken by Axelrod (1984). Axelrod's work represents a key event in the history of research on the PD. By inviting entries to two repeated PD computer tournaments, Axelrod studied the success of different strategies when pitted against themselves, all the others, and the random strategy. The strategy TFT won both tournaments and an extension of the second one. The extension, called ecological analysis, consisted of calculating the results of successive hypothetical tournaments, in each of which the initial proportion of the population using a strategy was determined by its success in the preceding tournament. Axelrod explains that TFT's success is due to four properties: TFT is nice (it starts by cooperating), provocable (it retaliates if its opponent defects), forgiving (it returns to play cooperatively if the opponent does so), and clear (it is easy for potentially exploitative strategies to understand that TFT is not exploitable). TFT's success is even more striking when one realises that it can never get a higher payoff than its opponent. Though severely criticised by some game theorists for drawing excessively on computer simulation and being partially flawed, Axelrod's work is widely accepted to have greatly stimulated analytical work within the domain of evolutionary game theory and further research on the PD using computer simulation. Findings on the repeated PD from evolutionary game theory are summarised by Bendor and Swistak (1995; 1998) and Gotts et al. (2003b); in particular, Gotts et al. (2003b) conclude that the assumptions about the dynamics of competition between strategies in mainstream EGT make the analytical results much less plausible as good approximations in social than in biological contexts. Gotts et al. (2003b) have also extensively reviewed work on social dilemmas using computer simulation.

As explained in the previous chapter, there are many different models in the branch of learning game theory, and their predictions for social dilemma games are far from uniform. In very general terms, models that have been designed to converge to Nash equilibria predict uncooperative solutions (see e.g. Fudenberg and Levine, 1998), while models including players who satisfice predict cooperative outcomes for certain ranges of aspiration thresholds (e.g. Karandikar et al., 1998; Bendor et al., 2001a, 2001b). There are also learning models where players do not satisfice and which lead to cooperative solutions; an interesting example is given by Erev and Roth (2001). Erev and Roth (2001) point out that the performance of reinforcement learning models in explaining human behaviour in games that facilitate reciprocation (i.e. games where players can coordinate and benefit from mutual cooperation, like the Prisoner's Dilemma) had traditionally been remarkably less successful than in explaining other types of games (e.g. zero-sum games and games with unique mixed strategy equilibria, see McAllister, 1991; Mookherjee and Sopher, 1994; Roth and Erev, 1995; Mookherjee and Sopher, 1997; Chen and Tang, 1998; Erev and Roth, 1998; Erev et al., 1999). As mentioned above, many people do learn to cooperate in the repeated Prisoner's Dilemma, whilst most simple models of reinforcement learning used in experimental game theory predicted uncooperative outcomes. Interestingly, Erev and Roth (2001) show that such a result does not reflect a limitation of the reinforcement learning approach but derives from the fact that previous models used to fit experimental data assumed that players can only learn over immediate actions (i.e. stage-game strategies) but not over a strategy set including repeated-game strategies (like e.g. tit-for-tat).

3.2. Method

In the following chapters we characterise the dynamics of various stochastic systems using a range of different techniques. The typical system investigated in this thesis contains a (potentially variable) finite number of players who interact to get certain payoffs, and are subject to stochasticity (either in their individual behaviour or in the dynamics of the population they belong to). In these systems, each of the players can adapt its behaviour (i.e. learn), or the population of players as a whole adapts through an evolutionary process. The payoff obtained by each of these players depends on the actions undertaken by other players; this feature is what makes game theory a useful framework to study the system.

This thesis makes extensive use of two distinct approaches to analyse the dynamics of these systems: computer simulation and mathematical analysis. As in Gotts et al. (2003a), it will be shown by example that mathematical analysis and simulation studies should not be regarded as alternative and even opposed approaches to the formal study of social systems, but as complementary. They are both extremely useful tools to analyse formal models, and they are complementary in the sense that they can provide fundamentally different insights on the same model (and also on one same question using different models, as argued by Gotts (2003b)). Chapter 4 will clearly illustrate the fact that the level of understanding gained by using these two techniques together could not be obtained using either of them on their own. Furthermore, each technique can produce both problems and hints for solutions for the other. The following explains how these two techniques have been used in this thesis.

3.2.1. Computer simulation

Simulations can usually provide an explicit and fully accurate representation of the original system and its stochastic dynamics. In this way, simulations allow us to explore the properties of formal models that are intractable using mathematical analysis, and they can also provide fundamentally new insights even when such analyses are possible.

The specific modelling technique used in this thesis is called agent-based modelling (ABM). ABM is a modelling paradigm with the defining characteristic that entities within the target system to be modelled –and the interactions between them– are explicitly and individually represented in the model (Edmonds, 2000). Because of this, ABM is especially appropriate to simulate game theoretical models, where the description of the system in terms of the behavioural and adaptive rules of the individual players is usually very simple. Clearly, running a stochastic agent-based model in a computer provides a formal proof that a particular micro-specification is *sufficient* to generate the pattern of behaviour that is observed during the simulation. However, one is usually interested not only in how the system *can* behave, but also in determining how the system behaves *in general*, which involves finding the probability distribution of different patterns. For this, it becomes necessary to run a large number of simulations with different random seeds and appropriately chosen initial conditions (see e.g. section 6.5.1). Most often, simulations cannot provide general closed-form results about how the system behaves, or about how it responds to changes in the parameter space. Thus, there is great value in complementing simulation with mathematical analysis.

In the work reported in this thesis simulation is often used as a starting point. There are two reasons for this. First, the very nature of the systems analysed here (see beginning of section 3.2) means that they can be easily described (and implemented) within an ABM framework. Secondly, the models developed here have not been designed to be mathematically tractable, but to study phenomena that we considered particularly interesting; thus, at least at first, they often seem to be mathematically intractable. Mathematical work is then used to analyse the patterns observed in the initial simulations, and this analysis sometimes leads to the production of simpler models that exhibit similar behavioural patterns and which are amenable to more detailed mathematical analysis. An example of this interaction between simulation and mathematical analysis is the development of deterministic approximations (i.e. simpler models) of the stochastic dynamics of a more complex system (e.g. see chapter 4). Simulation and mathematical analyses are therefore used complementarily: with simulation allowing us to explore intractable models, to extract the key features of such models, and to build new simpler models that still keep such key features; and mathematical work illuminating the behaviour of the initial models, and providing in-depth analyses of the simpler models. In many cases simulations have also suggested promising ways of pursuing new theoretical results.

As mentioned in the introduction, a great effort has been made in this thesis to make sure that every computational experiment conducted here can be easily inspected, rerun, scrutinised, reimplemented, and modified by independent researchers. Given the amount of care put on this task, I place as much confidence on the results obtained using computer simulation as I do on the mathematical derivations.

3.2.2. Mathematical analysis

The original systems investigated in this thesis can all be meaningfully formalised as Markov processes. However, the (sometimes infinite) number of states and the nature of the transitions between different states often mean that traditional Markov analysis cannot be readily applied. In the presence of these difficulties, there are two approaches that have been followed to characterise these systems using mathematical analysis: (a) partial analysis of the original Markov process, and (b) in-depth analysis of an approximation to the original Markov process.

The partial analysis often starts by finding out whether the Markov process is ergodic. If the process is ergodic, this means that the probability of finding the system in each of its states in the long run is unique (i.e. initial conditions are immaterial). This probability is also the long-run fraction of the time that the system spends in each of its states. Although calculating such probabilities may be unfeasible, one can always estimate them using computer simulation (see e.g. section 6.5.1). If the process is not ergodic, one can try to identify its various transient and recurring classes (see e.g. sections 4.7 and 5.4). This task may involve using very specific techniques which may be adequate only for certain types of Markov processes. A particular feature of Markov processes that often determines which techniques may be most appropriate for their analysis is how (if at all) the speed of change (e.g. the rate of learning) itself varies with time. As an example, it will be shown in chapter 4 that when the magnitude of change remains constant in time (e.g. in models where learning does not fade away in time), results from the theory of distance diminishing models (Norman, 1968, 1972) can be particularly useful. Another useful analysis that can be conducted on non-ergodic Markov chains with various absorbing states consists in identifying which of these absorbing states are robust to small perturbations (Foster and Young, 1990; Young, 1993; Ellison, 2000). This sort of analysis has been conducted in sections 4.8 and 5.7.3.

A complementary approach to the partial analysis of the original Markov process consists in studying a simpler approximation to it. In this thesis I have made extensive use of mean-field approximations. The use of mean-field (or expected-motion) approximations to understand the dynamics of complex stochastic models is common in the game theoretical literature (see e.g. Fudenberg and Levine, 1998; Vega-Redondo, 2003). Note, however, that these are approximations whose validity may be constrained to specific conditions. As a matter of fact, there is a whole field in mathematics, namely stochastic approximation theory (Benveniste et al., 1990; Kushner and Yin, 1997), devoted to analysing under what conditions the *expected* and the *actual* motion of a system should become arbitrarily close in the long run. This is generally true for processes whose motion slows down at an appropriate rate (as explained by e.g. Hopkins and Posch (2005) when studying the Erev-Roth reinforcement model), but not necessarily so in other cases.

In any case, mean-field approximations can be very useful even when it is known that they cannot be used to characterise the dynamics of the system in the long-run. As an example, in chapter 4 we use the expected motion of the system to get insights about what areas of the state space may be particularly stable (or unstable), to identify their basins of attraction, to clarify the crucial assumptions of the model, to assess its sensitivity to various parameters, and to characterise and graphically illustrate the *transient* dynamics of the model. We also show that the expected-motion approximation, while valid over bounded time intervals, deteriorates as the time horizon increases. In fact, the approximation becomes very misleading when studying the *asymptotic* behaviour of the model.

It is also worth mentioning that mean-field approximations are often used in the literature not only to average stochasticity out, but also to average out heterogeneity among players (e.g. see the studies conducted by Galán and Izquierdo (2005), Edwards et al. (2003), Castellano, Marsili, and Vespignani (2000), and Huet et al (2007)). Such approximations provide simpler, more abstract models which are often amenable to mathematical analysis and graphical representation. However, as pointed out above, even though they are usually useful, one should not forget that the insights provided by these mathematical abstractions could be misleading.

To conclude, let us mention that a range of other mathematical techniques (e.g. Brouwer's fixed-point theorem in section 4.9, and graph theory in section 5.7.3) have been used to analyse various properties of the models developed in this thesis.