7. Discussion

In broad terms, most of the results presented in the previous 3 chapters can be seen as logical deductive inferences of the form:

"Set of assumptions A" IMPLIES "Set of (deduced) statements B" [7-1]

As a matter of fact, any computer simulation and any mathematical derivation can be seen as a logical inference that establishes the truth of a set of statements B (e.g. the output of a model, or a derived mathematical result) given the assumption that a set of statements A (expressed in e.g. computer code, or as a set of equations) are true.

Deductive logical inferences are more useful the greater the generality of the set of assumptions A, and the greater the scope and level of detail of the set of deduced statements B. As an example, consider the results presented in chapter 4 on the dynamics of the Bush-Mosteller reinforcement model. These results advance previous work by Cross (1973) and by Börgers and Sarin (1997) because the results derived in this thesis are valid not only for positive stimuli, but also for negative ones; thus, the generality of the set of assumptions investigated in this thesis is greater. Similarly, the results presented in that same chapter are an advancement of (parts of) the work conducted by Macy and Flache (2002) and Flache and Macy (2002) on the Bush-Mosteller model because the level of detail of the characterisation of this model's dynamics is significantly greater in this thesis.

The logical inferences derived in this thesis can be applied in a number of useful ways. This chapter outlines 5 ways in which the research conducted in the previous chapters can be usefully applied to contribute to the advancement of human knowledge.

7.1. Direct application of the derived inferences

The simplest application of the logical statement "A implies B" relates to the case where A is thought, postulated, or demonstrated to be true. If a set of individuals are playing a certain game using one of the decision-making algorithms investigated in this thesis (e.g. the Bush-Mosteller reinforcement learning algorithm), then the results obtained in the previous chapters can be used to *predict* the (dynamic) outcome of the game, and also how this outcome may change when certain conditions (e.g. the magnitude of the payoffs or the speed at which players learn) are modified. Similarly, since "A implies B" is logically equivalent to "Not B implies Not A", if the observed results are deemed significantly different from B, then logical statement [7-1] can be used to infer that A cannot be true.

7.2. Assessment of the importance of assumptions in similar models

Another way in which logical statement [7-1] can be meaningfully used concerns the identification of crucial assumptions in inferences of the type "Set of assumptions A2 implies set of statements B2". Consider the case where sets A and A2 contain a large number of identical assumptions. An example of this would be two models of the same game: one of the models (A2) assumes common knowledge of rationality among the players, whereas the other model (A) assumes that players make decisions following the Bush-Mosteller reinforcement learning approach. Comparing the set of deduced results B and B2 will be illuminating: any difference between B and B2 can be attributed to the differences between A and A2. Thus, inference [7-1] can be used to assess the impact of various assumptions in models that are similar to the one defined by the set of assumptions A, but not the same. A clear illustration of this type of inference in the literature is given by Flache and Hegselmann (1999), who compare two models that differ only in the decision-making algorithm used by a set of players confronting the same spatial social dilemma setting: in one of the models, players use (partially) rational strategies that cooperate whenever reciprocal cooperation can be sustained as a rational equilibrium in the 2-player game they play (i.e. whenever the "shadow of the future" (Axelrod, 1984) is powerful enough); in the other model, players use a reinforcement learning rule based on Bush and

Mosteller's (1955) principles. In particular, Flache and Hegselmann (1999) show that under a wide range of conditions, the reinforcement learners need more time than the (partially) rational players to form stable cooperative relationships. This line of work was further developed by Hegselmann and Flache (2000), who compared rational behaviour and the Bush-Mosteller reinforcement learning rule over all possible symmetric 2x2 prisoner's dilemma games.

7.3. Selection, parameterisation, and validation of models

A third way in which the research conducted in this thesis contributes to the advancement of human knowledge concerns the interdependent processes of selecting, parameterising, and validating a model. A model is an abstraction of a real-world system that allows us to establish inferences about how the real-world system or certain aspects of it operate. Any model represents a compromise between realism and manageability (Intriligator et al., 1996, p. 13). Ideally, one would like to have a model that captures the essence of the target system (i.e. the model is realistic) and, at the same time, enables us to draw insights and conclusions that could not be derived from direct observation of the target system (i.e. the model is manageable). A perfectly manageable model that is not realistic is not useful; similarly, a realistic model that is not manageable (i.e. it does not yield new insights) is useless. This thesis has increased the manageability of several models that have received empirical support, thus improving their applicability. In this way, the work reported in this thesis enhances game theorists' toolkit of models that can be usefully employed to study real-world systems.

The task of selecting one particular model often includes considering various different alternatives. Naturally, the choice of criteria for the comparison of models depends on the purpose of the modelling exercise. Models in game theory are often compared with the aim of understanding what decision-making processes may be generating an observed pattern of play (see e.g. Feltovich (2000) and Camerer (2003)). For that purpose, one is often interested in studying the models' ability to reproduce observed statistical signatures and to predict patterns of play to a satisfactory extent. To conduct this assessment, the models to be compared need to be parameterised first. The following section outlines how to do this.

7.3.1. Parameterisation of models

As explained in section 3.2.2, the models investigated in this thesis can all be meaningfully formalised as Markov processes. The implicit assumption when parameterising a model with a set of observed data is that such data have been generated by the (appropriately parameterised) model. The challenge when parameterising the models studied in this thesis is that they represent systems where the state is not a variable that can be observed, i.e. the Markov chain is *hidden*. What is available to an observer is the pattern of play (i.e. the decisions made by the players), which is a stochastic process governed by the underlying Markov chain, but different from it. As an example, consider the Bush-Mosteller model of reinforcement learning. As explained in chapter 4, the model can be formalised as a Markov chain $\{X_k\}_{k\geq 0}$ whose state is fully specified by a two-dimensional vector [$p_{1,C}$, $p_{2,C}$], where $p_{i,C}$ is player i's probability to cooperate. The sequence of actual decisions made by the players is another stochastic process $\{Y_k\}_{k\geq 0}$ which is linked to the hidden Markov chain $\{X_k\}_{k\geq 0}$ in the sense that X_k governs the distribution of the corresponding Y_k. Since only $\{Y_k\}$ is observed, any statistical inference about the unknown parameters of the Markov chain $\{X_k\}$ must be done in terms of $\{Y_k\}$. Fortunately, methods to parameterise hidden Markov chains have been developed remarkably in the last few years. An excellent introduction to conduct this type of parameterisation is given by Cappé et al. (2005). In addition to the analysis of the pattern of play, it could well be the case that the value of certain parameters can be inferred using various other methods, like purpose-designed experiments, questionnaires or interviews with the players. These methods may be more reliable, simpler and, in any case, constitute a source of potentially very useful information that does not decrease the validity of the quantitative methods described above; thus, it seems most advisable to conduct them, if at all feasible.

7.3.2. Selection, validation, and applicability of models

Once the models to be compared have been parameterised, the process of selecting one can proceed. This is an activity that is strongly linked with the process of model validation. In broad terms, models are compared with the aim of selecting the best one of them according to some set of criteria, whereas validating the selected model is studying whether this (best) model is “good enough” for the intended purpose. Thus, it seems natural that the same techniques used to pick out the best model are appropriate to assess its validity too.

A model is valid to the extent that it provides a satisfactory range of accuracy consistent with the intended application of the model (Kleijnen, 1995)[39]. As mentioned above, models in game theory are often constructed with the aim of understanding what decision-making processes may be generating an observed pattern of play. In that context, validation often refers to the process of assessing how well a model is capturing the essence of its empirical referent. As mentioned above, one should not forget that a simple approach to validate a model about how certain individuals played a game is actually asking that same question to the individuals themselves[40]. Unfortunately, this does not seem to be a common approach in the literature of experimental game theory, even though it seems clear that it has the potential to contribute significantly to the design of more realistic models. The long tradition of introspective theoretical work in classical game theory may be at the root of this apparent lack of interaction with experimental subjects.

One common technique to quantify the extent to which a model is capturing the essence of a pattern of play consists in studying the models’ ability to reproduce observed statistical signatures and to predict patterns of play to a satisfactory extent. This is an issue extensively studied in the systems identification literature (Söderström and Stoica, 1989; Ljung, 1999). The general approach to validate a model is based on an in-depth analysis of its prediction error, which is a measure of the disparity between the observed data and the model’s predicted output. If possible, the preferred option is to evaluate the model performance using a set of

[39] See a complete epistemic review of the validation problem in Kleindorfer et al. (1998).

[40] Work outside the literature in experimental game theory suggests that players’ responses may vary depending on *when* they are asked to describe their reasoning processes (Ericsson and Simon, 1980). People tend to verbalise what they are doing more accurately when asked *while they solve a problem* rather than when asked *some time after having tackled the problem*.

data different from the data employed to parameterise the model (i.e. the estimation data). If, on the other hand, the prediction error has to be calculated using the estimation data, there are a number of model selection criteria (e.g. Akaike's information criterion (Akaike, 1969) and minimum description length (Rissanen, 1978)) designed to avoid biases and pitfalls (e.g. overparameterisation and overfitting) by adding certain correcting terms to the computed prediction error (Ljung, 1999, p. 507). These correcting approaches are especially relevant when comparing models that have different number of parameters. An important part of the validation exercise is then the analysis of residuals (i.e. the part of the validation data that the model could not reproduce). This analysis minimally consists in plotting the residuals, computing basic statistics on them, analysing their structure, and conducting tests of independence. The precise purpose of the modelling exercise will dictate what other tests will be useful.

At this point it is worth addressing a criticism that the Bush-Mosteller model investigated in chapter 4 of this thesis has recently received, and which relates to its applicability. Bendor et al. (2007) argue that the BM model (and many others) have "little empirical content" because "such models imply that virtually anything can happen" (see reply by Macy and Flache (2007)). They prove their point showing that any outcome of the game can be sustained as a stable outcome by some pure SRE. Their proof of this result consists in setting an aspiration threshold below the lowest payoff of the game. As shown in chapter 4, once a certain value for the aspiration threshold is chosen, it is not generally true that any outcome can be sustained by an SRE. In fact, it is straightforward to see that any value for the aspiration threshold above the minimum payoff will preclude at least one outcome from being sustained by an SRE. Thus, their criticism refers to a Bush-Mosteller model where players have aspiration thresholds below the minimum payoff they can receive. In our view, the aspiration threshold is a parameter whose value can be estimated using empirical methods by e.g. using the theory of inference in hidden Markov chains mentioned in the previous section. The fact that it is possible to find a specific value for the aspiration threshold such that any outcome can be supported by an SRE is not a drawback of the model, since the value of the aspiration threshold can be inferred from empirical observation, and most of the values this parameter can take induce a process where not every outcome can be sustained by an SRE. An analogy that comes to mind is Newton's theory of gravitation: this theory provides (in particular) a mapping between the height at which an object is released and the time that the object takes to hit the ground (time = f(height)). Similarly, this thesis has characterised the (non-trivial) mapping between the parameters of the Bush-Mosteller model (in particular, the aspiration threshold) and the dynamics of the resulting process (in particular, the characterisation of the set of SREs):

Set_of_SREs = function(Aspiration_Threshold).

It is indeed true that for any given outcome, one can always find an aspiration threshold so the outcome is supported by an SRE. Similarly, in Newton's theory of gravitation, for any time t_0 one can always find a height h_0 such that $f(h_0) = t_0$, but this does not seem to be a drawback of the theory.

Bendor et al.'s (2007) criticism seems to be unjustified even in the case where aspiration thresholds are so low that any outcome can be sustained by an SRE. As explained in chapter 4, even in the case where there is a positive probability that any outcome will be played indefinitely, this probability is generally different for different outcomes and depends on a number of factors (e.g. initial conditions, aspiration thresholds, and learning rates). The exact probability of approaching each possible SRE can be estimated to any degree of accuracy using the methods explained in chapter 4. Thus, the Bush-Mosteller model yields predictions that can be falsified, even when aspiration thresholds are below the minimum payoff.

7.4. Modelling frameworks

As explained in chapter 2, there is nowadays a whole universe of models that abandon the demanding assumptions of classical game theory on players' rationality and beliefs. These models make different assumptions regarding the meaning of payoffs, the amount of information that players can access, players' computational capabilities, and the level at which the dynamics are described (i.e. population adaptation vs. individual learning), to mention a few. The formal analysis of these models is often quite challenging, and consequently most of the research conducted until now has focused on characterising the dynamics of each of these non-trivial models in relative isolation. There is obviously a lot to be gained from comparing different models, but our lack of in-depth knowledge of their dynamics has meant that this comparison has had to be postponed. Fortunately, nowadays the number of models that have been thoroughly analysed seems to be sufficient to justify initiating the process of creating frameworks –i.e. meta-models– where alternative models would arise as particular cases.

An example of a useful framework that has been proposed within the field of learning game theory is Flache and Macy's (2002) general reinforcement learning (GRL) framework. Flache and Macy's (2002) framework integrates a smoothed version of the Erev-Roth model (see section 4.1) and the Bush-Mosteller model as particular cases. The GRL framework has a parameter that measures the level of fixation in the decision-making algorithm. When this fixation parameter equals 0, the framework reduces to the Bush-Mosteller model, whereas if the parameter equals 1, the obtained model is Erev and Roth's. The use of the GRL framework enabled Flache and Macy to conduct a transparent and fruitful comparison of the two models and also to uncover hidden assumptions in both models.

An example of a framework within the field of evolutionary game theory is EVO-2x2. As explained in chapter 6, EVO-2x2 is a computer simulation modelling framework designed to formally investigate the evolution of strategies in 2x2 symmetric games under various competing assumptions. EVO-2x2 enables the user to set up and run many computer simulations (effectively many different models) aimed at investigating the same question using alternative assumptions. Thus, EVO-2x2 provides a single coherent framework within which results obtained from different stochastic finite models can be contrasted and compared, as illustrated in section 6.5.2.

The development of frameworks is useful not only to assess the impact of various assumptions in theoretical terms, but also to inform experimental research. By making differences between models explicit, frameworks can facilitate the design of experiments targeted at identifying the type of models that may be most adequate in a certain situation. Frameworks can also help to identify the factors (i.e. types of assumption) that may have the greatest impact in the outcome of a social interaction. Thus, the use of frameworks may facilitate the interaction between game theorists and empirically-driven social scientists, from which game theory would benefit so much. The ideal result of this interaction would be a framework encompassing various models as particular cases, where the differences between the models were made explicit, and where each model were annotated with indications about the type of context for which the model may be most adequate.

A discussion about frameworks raises the question of whether evolutionary and learning game theory could be integrated into a single discipline. The derivation of a significant number of theoretical results relating various learning models with different versions of the replicator dynamics (e.g. Börgers and Sarin, 1997; Posch, 1997; Hopkins, 2002; Hopkins and Posch, 2005) would seem to suggest that the integration of these two fields may be within reach (Weibull, 1998). However, the integrative theoretical results tend to establish analogies at a very high level of abstraction. A representative example is given by Börgers and Sarin (1997), who demonstrate that the continuous time limit approximation of the dynamics of the Bush-Mosteller learning model (which cannot be used to characterise its asymptotic behaviour, as demonstrated in chapter 4) converges to the replicator dynamics of evolutionary game theory. These types of result are certainly useful, as they provide non-biological interpretations of evolutionary models, and evolutionary interpretations of learning models. However, the number of assumptions that are needed to align models from the two disciplines tend to decrease the applicability of the obtained inferences significantly. Thus, it seems that there are many frameworks that can be usefully developed at lower level of abstractions before the integration of learning and evolutionary game theory can take place.

7.5. Models as ‘tools to think with’

The formal models developed in this thesis have also been useful as ‘tools to think with’. The clearest example of this use of a model is illustrated in section 5.5, where the concept of iterative elimination of dominated outcomes was put forward. Iterative elimination of dominated outcomes is a logical process through which players can arrive at sensible (*i.e.* Pareto optimal) outcomes in games.

Dominated outcomes are outcomes which are not individually rational – i.e. there is at least one player who is obtaining a payoff below her *Maximin*. The idea behind the process of iterative elimination of dominated outcomes is that players cannot rationally accept outcomes where they are not obtaining at least their *Maximin* (rational players are not exploitable). When players who do not accept outcomes where they get a payoff lower than *Maximin* meet, they might learn by playing the game the fact that their opponent is not exploitable either. If this occurs, it will be mutual belief that dominated outcomes cannot be sustained because at least one of the players will not accept them. That inference (and the consequent disregard of dominated outcomes by every player) can make an outcome that was not previously dominated in effect be dominated. In other words, the concept of dominance can be applied to outcomes *iteratively* just as it is applied *iteratively* to strategies.

In this section we expand the philosophical basis of this process of reasoning by outcomes a bit further. As mentioned several times in this thesis, the history of classical game theory has been marked by the assumption that agents are instrumentally rational. However, except in strictly competitive games, defining rational behaviour in games is by no means straightforward (Colman, 1995). The challenge in game theory is that, in general, the definition of rational behaviour for any one player depends on the behaviour of potentially every other player in the game. As an example, in an iterated Prisoner's Dilemma game, the rational strategy against a player who always defects is to defect, but the rational strategy against a player who is known to play Tit for Tat may be to cooperate, if the number of rounds is sufficiently large.

Thus, in order to identify the rational course of action in a game, one is bound to partition the infinite set of possible behaviours that the other players may take according to some criterion, and then try to compute the best reply to each type of behaviour identified. Classical game theory partitions this universe of possible behaviours according to strategies. In this way, classical game theory defines rationality in terms of beliefs about the *strategy* that the other players may use: rational players do not choose dominated *strategies* because there is no belief about the other players' *strategies* such that selecting the dominated *strategy* is optimal. The partition of the "behaviour space" according to strategies is quite natural since, after all, it is strategies that players can choose.

On the other hand, players' measure of success –i.e. the obtained payoff– is not determined solely by their strategy, but by every player's strategy, i.e. by the resulting *outcome* of the interaction. Thus, it may also seem natural to assume that players do not think in terms of *strategies*, but in terms of *outcomes*. In other words, players may be willing to accept certain *outcomes* but not others. The models developed in chapter 5 triggered the idea of defining rationality partitioning the universe of possible behaviours according to *outcomes*, instead of strategies. This leads to the definition of the so-called outcome-based rationality. According to this definition, rational players do not accept dominated *outcomes*. Note that this definition is somewhat problematic, since the words "do not accept" already imply the existence of some dynamics. Remember, however, that the definition of rationality based on strategies also led to similarly worrying problems (e.g. the existence of many possible Nash equilibria).

Once outcome-based rationality is defined, one can develop the same concepts that were explained in section 2.2.2 using the new definition of rationality. Thus, one can define the process of iterative elimination of dominated outcomes, and also the concept of rationalisable outcomes.

The definition of outcome-based rationality has a certain intuitive appeal which becomes apparent when studying the Prisoner's Dilemma. The process of iterative elimination of dominated outcomes leaves mutual cooperation as the unique surviving outcome. The reasoning behind this logical process goes as follows: players are rational and therefore they will not accept the outcome where they receive the sucker's payoff. They also know that the other player is rational, so they acknowledge the fact that their counterpart is not going to be exploitable either. Once this is recognised by the two players, the rational course of action is to try to achieve mutual cooperation rather than mutual defection.

It seems clear that even though there is a clear causal link between strategies and outcomes, defining rationality in terms of outcomes rather than in terms of strategies leads to completely different results even in the simplest games. Section 5.2 explained how rational strategies may lead to outcomes that are not rational, whereas rational outcomes may be generated by strategies that are not rational. A more thorough account of the implications of outcome-based rationality is left for future work.