GAME THEORY - ITS APPLICATIONS TO ETHICAL DECISION MAKING

STEFANO CAVAGNETTO
BRUCE GAHIR

The application of game theory according to Hargreaves-Heap and Varonfakis (1995) to understand human behaviour, and in particular ethical behaviour, is a valuable development, as game theory has gradually become one of the key frameworks to assist us in the understanding of social sciences. Esther (1982) and Aumann and Hart (1992) show that there are several studies that indicate the importance of a game theoretic framework in advancing our understanding of social behaviour and evolutionary sciences. Although the application of game theory in the above areas has largely been not formalised, its application in the fields of ethical conduct and human behaviour is at present developed in several respects with the gradual assistance of advances in related areas such as evolutionary biology and our understanding of group social behaviour.

Game theory tends to be dominated by its assumptions concerning the non-rational nature of the dynamics existing within social structures and these processes are generally ignored or are sometimes difficult to model. Assumptions that are made concerning the cognitive abilities and beliefs about individuals are generally unrealistic, and this at times is the main criticism of game theoretic applications in our understanding of ethical behaviour. These are the lines followed by Solomon (1999) in his discussion on game theory when it is used to model scenarios in business ethics and in general in the business area. In particular, game theory tends to excessively emphasise the destructive obsession of the quantifiable outcomes and an artificial concept of competition. This modelling of competition is not exhaustive of human behaviour, but rather an oversimplification of the social dynamics in place in many different human contexts.

The first objective of this paper is to offer a brief sketch of the historical developments that have taken place in game theory and some of its fundamental concepts through an analysis of the most important games that have found an application to the field of ethics. A second aim is that of providing the reader with a review of developments relating to evolutionary game theory that has deeply assisted our comprehension of ethical norms and their emergence in group behaviour. The final part of the paper will be dealing with some conclusive remarks from a methodological point of view on the main issues concerning applications of game theory to ethics.

A SHORT HISTORY OF GAME THEORY

Game theory's history can be roughly divided into three different periods. The first period is characterised by the development of the study of games linked with the development of research in probability theory and the development of probability calculus due to Pierre de Fermat and Blaise Pascal. In the literature this moment begins with a famous correspondence between the two French mathematicians dated 1654 (Hyksova, 2004). About sixty years later, the mathematician James Waldegrave had an important part in the creation of the concept that is nowadays known as an example of a mixed strategy solution of a matrix game. Waldegrave in a written correspondence with Pierre Remond de Montmort on the 13th of November, 1713, offered his solution for the card game *Le Her*. This game was quite popular in the 18th century, and it is a card game played by two players. In the literature they are usually referred to as Peter and Paul, the dealer and the receiver, respectively (Epstein, 2009). It begins when Peter deals Paul one card that is randomly chosen from an ordinary deck of 52 cards. After that, he then gives a card to himself. Both players cannot see the cards dealt to each other. The purpose is to keep the card of a value that it is higher than the value of the opponent (the king is the highest and the ace the lowest). If Paul is not happy with his card value, he can ask his adversary to swap with him, with the only exception when Peter has a king. On the other side, if Peter, the dealer, is not happy with the card value that he has after the exchange, he is then entitled to swap it for a card randomly picked from the remaining pack of cards. At this point, Peter is permitted to take a new card if and only if the new card is not a king, in which case he must keep his own previous card. Following the exchanges, Peter and Paul compare the two cards that they are holding. The player with the higher card value is the winner. In the case that both Peter and Paul hold the same card value, then the deck holder is the winner, i.e. Peter.

This game was previously studied by de Montmort and Nicholas Bernoulli in terms of winning strategies. Both scholars arrived at the conclusion that the deck holder should change each card carrying a value inferior to 8, while the opponent should change each card carrying a value less than 7. In questionable situations de Montmort was of the idea that no guidelines could be given, while Bernoulli thought that both players should change. On his side Waldegrave was investigating the possibility of proving the existence of a strategy that could maximise the probability of a player's win, independently from the strategy selected by the other opponent. Nowadays this is known as "minimax principle", and it will be discussed later on in this paper. As remarked in Hyksova (2004), the mixed strategy proposed by Waldegrave for the solution can be formulated in terms of Black and White chips and it can be shortened as follows: the deck holder (Peter) has to adopt the strategy <change 8 and lower> with the probability $P = 0.375$ and the strategy <keep 8 and higher> with the probability $P = 0.625$. On the other side the opponent, Paul, has to select the strategy <change 7 and lower> with the probability $P = 0.625$ and the strategy <keep 7 and higher> with the probability $P = 0.375$. As pointed out in Hyksova (2004), despite the fact that de Montmort included in his publication all the correspondence relating to *Le Her*, together with Waldegrave's contribution, this solution proposed by Waldegrave stayed ignored for a very long time.

The second period goes approximately from 1913 to 1928 and is characterised by the work of the famous set theorist Ernst Zermelo, the French mathematician Emile Borel, and by the first researcher in this area, the naturalised American mathematician John Von Neumann. In some sense the first theorem of game theory was proven by Zermelo. The theorem claims that in a game of chess three different outcomes are possible: white can force a win, black can force a win, or both sides can force at least a draw. The first modern formulation of a mixed strategy and the minimax solution was given by Borel between 1921 and 1927. He gave the minimax solution games with two players and with three or five possible strategies. Finally, by 1928 von Neumann proved the minimax theorem for zero-sum games, two players, and finitely many pure strategies for each player. In particular, when mixed strategies are taken into account, this form of games have exactly one (rational) payoff vector. From the notational point of view, these works introduced into the field of the well-known extensive normal form for a game (Hyksova, 2004).

The third era starts with the investigations carried out by Von Neumann and the economist Oskar Morgenstern at the beginning of the 1940s during the Second World War. In traditional accounts of the history of game theory, it is recognised in the literature that the comprehensive monograph *Theory of Games and Economic Behavior* published in 1944 by J. Von Neumann and O. Morgenstern developed the modern mathematical framework for the discipline. Undeniably, this publication represents the first important modern milestone in the development of this branch of applied mathematics and is normally regarded as being the starting point of game theory as an autonomous mathematical discipline. The monograph starts with a thorough formulation of economical problems and explains what might be the potential of game theory and its possible applications in modelling economic situations.

The work also lays down the foundations and a suitable framework of an axiomatic utility theory that is a classic reference for further research. The monograph also includes the general formal description of a game of strategy and an extensive study of the theory of finite two-player zero-sum games and certain types of zero-sum cooperative games with n players. The work of Von Neumann and Morgenstern widely contributed to the development of game theory as an independent field of research and thoroughly placed a great emphasis on its potential applicative aspects.

A second crucial milestone in the history of game theory is symbolised by John Forbes Nash's doctoral dissertation devoted to a full mathematical investigation of the theory of non-cooperative games. The idea of equilibrium point (today called Nash equilibrium) was discussed and its existence was demonstrated. The most important achievements of his dissertation were published in a short note, (Nash, 1950), and in a second more exhaustive paper published a year later (Nash, 1951); in the literature usually the name of Nash is also associated to a second idea that of Nash bargaining solution regarding two-player cooperative games for a certain type of games known as "cooperative games with no transferable payoffs". In this respect Nash formulated a system of axioms that a solution should satisfy as a set of requirements and demonstrated the existence of a unique solution with these requirements for the chosen set (Nash, 1951).

Definitely one of the first clear uses of game theory outside of the economic field was the application to political sciences. Remarkable in this sense was the work of L. Shapley and M. Shubik, *A Method for evaluating the distribution of power in a Committee System*, published in 1954. In this paper the authors investigate a possible solution for cooperative games in order to establish and determine the UN Security Council members' power.

In the late 1950s after Shapley's and Shubik's work, many other publications appeared that contributed to shape the literature of political sciences and made game theory central to the field in modelling several scenarios related to legislature, elections, politics of interest groups, lobbies, bargaining, etc. Nowadays game theory offers the suitable terminology and methods to resolve particular and general problems. Two fundamental works should be mentioned, the first is the work by R.D. Luce and A. A. Rogow, *A Game-Theoretic Analysis of Congressional Power Distributions for a Stable Two-Party System*, and the work by W. H. Riker, *A Text of the Adequacy of the Power Index*, published in 1956 and in 1959, respectively.

Current monographs on political sciences regard game theory as an indispensable tool and component of the discipline; the interested reader can see for example as reference works by J. D. Morrow and by P. Ordeshook (Morrow, 1994; Ordeshook, 1986). Even though game theory is not seen as the best provider solution tool for any given problem, and indeed it cannot provide an optimal solution to all problems, political scientists regard it as a powerful tool for analysis of a given political and social situation. As remarked in Hyksova (2002), in this framework game theory is regarded as a strong tool that "induces the decision-maker to think rationally and without emotions; this, in itself, often yields a general acceptable solution".

Currently game theory is one of the most important tools for the study of conflict and cooperation in the world of living beings ranging from animals to plants as well. Game theory has been used for the analysis, modelling, and understanding of several zoological applications, from fight to cooperation and communication of animals, and from analysis and modelling of coexistence of alternative traits to distributions of individuals in their habitats in different species. Many applications can be found also in botany where the theory has been applied successfully to seed dispersal, seed germination, root competition, nectar production, and flower size differentiation. The 1960s saw the appearance of different sparse and isolated research employing game theory to approach a few problems in biology, but most likely the cornerstone and ground-breaking work was *The Logic of Animal Conflict* by John Maynard Smith and George Robert Price. This work, published in 1973, inspired a large number of research endeavours and utilisation of game theory in evolutionary biology and contributed to create a solid framework for the further independent development of the field.

Maynard Smith's work *Evolution and the Theory of Games* appeared in 1982 summarises well the subsequent developments of these applications and their success in evolutionary biology. Indeed, *Evolution and the Theory of Games* demonstrated how game theory could help to offer the most insightful explanations of the theory of evolution and the principles of behaviour of animals and plants when they are analysed in a context where they can mutually interact. In the end, most likely evolutionary biology turned out to provide as a field of research and projects development the most promising applications of game theory (Hyksova, 2002).

INTRODUCTION TO SOME FUNDAMENTAL CONCEPTS AND THE PRISONER'S DILEMMA

Game theory can generally be explained as a framework that allows one to derive outcomes with assumed preferences for agents concerned. These outcomes may arise due to interactions amongst agents where certain preferences, called utilities, are key drivers.

The analysis of games depends on the outcomes and expectations of players, or agents. Situations where the expectations of other agents depend upon expectations of one-self will be generally governed by required outcomes and performances called utilities. Let us first examine the concept of utility prevalently used in game theory. An agent with required needs is generally defined as an economic agent. If such agents are considered as rational then their rational decision-making process can be described following the concept of utility. Utility here is described as a measure of an increase or decrease in the subjective welfare that is attained as a result of desired outcomes.

'Welfare' here is defined as some form of normative index indicating a measure that depicts the well being of an agent, for example, the welfare of different countries based on their per-capita income. In the areas of economics and game theory, where the focus is more on the relative welfare concerning people, the idea of 'utility' tends to denote a measure of subjective psychological fulfilment.

Game theory involves the domain of formed reasoning and as a result the idea of 'utility' requires a framework that is grounded in formal terms. The 'utility function' is therefore defined as a mapping that maps agents ordered preferences to real numbers. Let us suppose, for example, that:

*Agent A has a preference to outcome "X"
compared to outcome "Y" and outcome "Y" compared to outcome "Z".*

This scenario can be mapped using numbers and a utility mapping function as follows:

Outcome X = 3

Outcome Y = 2

Outcome Z = 1

In this example, we note that the property mapped by the utility function is order. In a game theoretic framework, a game is any situation where agents act in a way to maximise their utility by anticipating the response to their actions by other agents. Agents participating in games are called players. Players are assumed to have capacities that are collectively referred to as 'rationality'. A player is defined to be economically rational if and only if:

- 1.** A player evaluates outcomes in a game and ranks these outcomes with respect to their contributions to his/her welfare.
- 2.** A player calculates which sequence of actions leads to which outcomes.
- 3.** A player selects actions from a set of alternatives that produce the most preferred outcomes.

In such cases we may describe an agent as 'choosing' actions from a selection of possibilities. Such a kind of 'economic rationality' may be part of the behavioural dispositions of the agent. In this case, each player in the game is facing a choice and a strategy is defined as a pre-determined programme of play. In what are termed as 'simple games', from the point of view of logical structure, agents are supposed to have perfect information; for instance, chess is such a game.

In this game a player tends to choose actions by reflecting upon the series of responses and counter responses of the opponent and then considers which of these responses will result in the highest utility. Let us illustrate this further by a famous example called the Prisoner's Dilemma (PD). Usually in the literature the game is attributed to Merrill Flood and Melvin Dresher as the type of games that were discussed extensively at the RAND Corporation in the 1950s during the Cold War in the context of the exploration of different strategies to be pursued in increased development relating to global nuclear weapons.

The name "Prisoner's Dilemma" was given by the Canadian mathematician Albert Tucker (Poundstone, 1992). In 1950, Tucker presented the ideas of Flood and Dresher to an audience of psychologists at Stanford University, and in order to make these concepts more accessible to an audience of non-mathematicians, proposed a situation that describes a scenario where the police have arrested two possible suspects (Axelrod, 1984). They believe that these suspects have committed an armed robbery. However, the police are in an unfortunate situation as they lack evidence to support their claim about their robbery being committed together by these two possible suspects. However, the police have enough evidence to convict each of the suspects for one year for theft of the getaway car that was used. So, during the questioning process the following proposals are made to both the suspects by the police:

- 1.** If Suspect A (SA) confessed to the robbery claiming that Suspect B (SB) was involved and SB does not confess to his collaboration with SA, then SA will go free and SB will be convicted to twenty years, a similar situation also results if SB confesses and SA remains silent.
- 2.** If both SA and SB confess then, each of them will get five years.
- 3.** If neither SA nor SB confesses, then they each will be convicted for one year because of the car theft.

These options are represented in the matrix in Figure 1.

		PRISONER B	
		Confess	Remain Silent
PRISONER A	Confess	5,5	0,20
	Remain Silent	20,0	1,1

Figure 1: Pictorial representation of the Prisoner's Dilemma.

If we represent these options in term of utility functions representing Suspect A as player one and Suspect B as player two, then we have the following situation:

Go free = 4
Sentenced for 1 year = 3
Sentenced for 5 years = 2
Sentenced for 20 years = 0

If we now utilise these numbers to express each player's payoffs in the different options represented above, we can represent them on a single matrix as follows:

		PLAYER 2 (B)	
		Confess	Refuse
PLAYER 1 (A)	Confess	2,2	4,0
	Refuse	0,4	3,3

Figure 2: A Prisoner's Dilemma matrix of payoffs.

We note that each cell in the matrix provides a payoff to both players, and player 1's payoff appears as a first numbers in each of the pairs. So for example, (2, 2) represents five years in prison for each suspect, this matrix can therefore be compared with the one in Figure 1.

The situation is therefore such that each player is able to evaluate their two possible actions by comparing and contrasting the personal payoffs that each will receive, and these are represented by the numbers in each pair shown in the cells of the matrix in Figure 2. Now, let us consider the situation from the point of view of player 2. If player 2 confesses, then based on the utilities and options in the matrix, player 1 gets a payoff of 2 due to confessing and a payoff of 0 if he refuses to confess; therefore player 1 is better off confessing independently from what player 2 does in this situation. So, in a similar way player 2 also comes to an evaluation that points to confessing. So, we note that if both players confess then both will be convicted for 5 years, and in this case, we say that confessing strictly dominates refusing to confess for both players, so the dominant strategy is to confess.

First of all, let us notice that both players can arrive at this outcome by using a mechanical process called iterated elimination. Using the matrix in Figure 2 we can see that for player 1 the payoffs for each of the cells in the top row are higher in value than those in the bottom row; therefore for player 1 the utility maximising strategy is not to choose the options in the bottom row, this means, denying to confess, independently from what player 2 does. So, as player 1 will not play the bottom row strategy, we can simply erase the bottom row or shift along the orange arrows to the higher pay off options along the matrix as indicated in Figure 2.

It is now obvious from Figure 2 that player 2 will also not refuse to confess because the payoff from confessing is higher than refusing to do so. Again, in this case as before we can erase the one cell column on the right of the matrix or again move along the orange shaded arrows which will now converge to the point of the light green shaded equilibrium cell. Through this iterative procedure we can determine almost mechanically the dominant strategy, in this case that is represented by the strategy pair <confess, confess> for both suspects.

The order in which the rows are deleted in the above process does not really matter; we would arrive at the same solution even if we had started by deleting the right column first. The strategies chosen by agents reflect the ones that lead to higher payoffs and these point to the situation of joint confession as the solution for this PD game where both players are being motivated by being economically rational. When the PD is generally discussed, one often hears that the police inspector must separate the suspects into different rooms, and this is done with the aim of preventing any mutual communication between the suspects. Because in the case of possible communication the suspects would surely realise that each of them is better off if both refuse to confess, and this means that they could make an agreement to do so; collusion is therefore barred in this PD example. We will discuss later examples where cooperation can be introduced in such PD type examples.

In general, the PD models a scenario where there are advantages from cooperation (each player would prefer the strategy <refuse, refuse> than they both select the strategy <confess, confess> but every player has an incentive to a 'free ride' (namely to select the second strategic option) independently from the behaviour of the other player. Despite the attractiveness of the name of the game, its value goes well beyond comprehending the incentives for prisoners to select the strategy <confess, confess>, but because there are numerous situations that have analogous logical structures. Every time each of two players has two possible strategies, let's say R (denoting the strategy 'Remain silent') and C (denoting the strategy to 'Confess'), the first player prefers (C, R) to (R, R) to (C, C) to (R, C), and the second player prefers (R, C) to (R, R) to (C, C) to (C, R), the dilemma represents the situation that the player is facing. A classic example in economy where the situation can be easily modelled using a variant of the PD is the model of a duopoly where two companies manufacture an identical good, for which every company charges either a low price or a high price. In this situation every company aims at the highest possible profit. In this case, the two strategies 'Charge High Price' and 'Charge Low Price' correspond to the strategies 'Remain Silent' and 'Confess'. A reader interested on several different examples derived from the PD and a discussion in depth about the assumptions around the incentives for this type of games can consult Osborne (2002).

Let us now suppose that the two suspects do value each others' future lives; if this is the case, then we could reasonably assume that such a condition will be reflected in their respective utility functions and also in their payoffs, and it is therefore the logic of the way the PD situation is set up that produces the inefficient outcome and not the psychological makeup of the two suspects. Following the terminology of game theory, the optimal (2, 2) position is referred to as the equilibrium position, technically known as the Nash Equilibrium (NE) from the name of the mathematician John Forbes Nash. Given the strategies of all other players in the game, we say that a set of strategies is a Nash Equilibrium (NE) when no player could improve their payoff. Therefore, if iterative elimination takes us to a unique outcome, this is the game's unique NE.

Before going on, let us see some of the assumptions that play a crucial role in the Nash equilibrium theory. As in the case of the rational decision maker, it is assumed that every player select the best possible action offered. Of course, in a given game scenario the best possible available action depends for each player on the actions selected by other players. This implies that when a player will choose an action, that player will have in mind the actions that the other players will select. This means that each player will create a belief based on the actions of the other players.

Of course the obvious question here is how this can be done, on which basis can such a belief be created? The main assumption here is that each player has a certain set of beliefs that is based upon their experience in the past in playing the game and that this experience is sufficient and extensively rich enough to know how their opponents in the game will behave. Nobody tells the player how the other players will behave and which course of actions will be chosen; however, the past involvement of the player leads to some level of certainty about these future actions (Osborne, 2002).

This point is quite subtle but very important. The assumption is that every participant in the game has experience in playing it; one can assume that the player views every round of the game in isolation. The result of this is twofold. On one side, the specific player does not familiarise with the behaviour of specific adversaries, and as a consequence, does not make the actions of the strategy chosen depending on the actions of the specific opponent they are facing. On the other side, the player does not expect that the current actions deriving from the selected strategy will affect the other participant's future behaviour (Osborne, 2002).

The best way to imagine this situation is to consider the following idealised situation. "For every player in the game there is a population of many decision-makers who may, on any occasion, take that player's role. In each play of the game, players are selected randomly, one from each population. Thus each player engages in the game repeatedly, against ever-varying opponents" (Osborne, 2002). Thus it is the experience of the player that guides the player to certain beliefs concerning the actions of characteristic opponents and not any detailed group of opponents.

To sum up, the concept outlined above has two components. First, each player selects an action consistently with the model of rational choice adopted, given the set of beliefs about the actions of the other players. Second, the beliefs of each player concerning the actions of the other players are correct. As remarked in Osborne (2002), these two elements are the founding conceptual elements embodied in the notion of Nash equilibrium described above.

Having elaborated upon some basic concepts and their role in game theory we will turn now to a specific application in the field of ethics.

GAME THEORY AND ETHICAL DECISION MAKING

We can date back the application of game theory to ethics around 1954 when the philosopher Richard Braithway gave a lecture titled "Theory of games as a tool for the moral philosopher". He argued that several of the problems encountered in the areas that discuss distributive justice in ethics have similar structures as the bargaining problem. The application of game theory to ethics was not really an original development during this period, elements of game theory analysis of certain ethical problems can be found in the works of several philosophers such as Thomas Hobbes and David Hume, while further details of this work can be found in Gauthier (1969), Kavka (1986), Hampton (1986), Van der Schraaf (1998). Several years after the work of Braithway, Brain Barry published his work "Political Argument", followed by David Lewis' work "Convention" (Barry, 1965; Lewis, 1969) that provided further details to Braithway's ideas.

However, it is not until the late 1960s when the first publication in a series by David Gauthier appeared providing formalised detail to the bargaining situation and its application to ethics. Surveying the literature, concerning the applications of game theory to ethics, one can generally classify the areas being discussed into three main parts:

1. The functionalist approach - where the focus of game theory is to establish the functions of morality.
2. The contractarianism approach - where game theory is used to formalise social contract theory.
3. The evolutionary approach - where evolutionary game theory is utilised to recover and establish the origin of moral norms.

Below we shall discuss these separate approaches further in some detail, elaborate on their outcomes, and present some of the problems that authors have commented upon.

THE FUNCTIONALIST APPROACH

A good example to consider when discussing the functionalist approach is the "Gunner" example presented by Edna Ullmann-Margalit, in her work "The Emergence of Norms", the interested reader can see Ullman-Margalit (1977).

In this work Ullman-Margalit argues for the emergence of moral norms and argues that such norms provide agents with a framework to cooperate and coordinate their actions in certain scenarios and situations where the following of self-interest would otherwise prevent such a situation from occurring. She takes the example of two artillery men facing a choice to either stay and use their weapons and fight against the enemy or flee for their safety. If both decide to stay and face the enemy, they may have a high chance of being injured in the attack, but there is a high probability that they will stop the enemy from advancing further.

However, if one artillery man stays and the second one flees, the brave artillery man is certain to die in battle while the other one will escape safely. So, each artillery man has a choice of either fleeing or staying and fighting, this is represented in Figure 3 as a matrix.

		GUNNER NO. 2	
		Stay	Flee
GUNNER NO. 1	Stay	2,2	0,3
	Flee	3,0	1,1

Figure 3: Illustrating the Ullmann-Margalit Example.

The choices are represented as rows for artillery man 1 (A1) and columns for artillery man 2 (A2). We note that, as before in Figure 2, each cell in the matrix is shown to represent the particular outcome of each possible pair of choices facing the artillery men.

The first number represents how A1 could rank his outcome taking into account the relative possible outcomes of others, ranks are here represented as numbers denoting utility, and the second number in the pair represents the ranking of A2. Let us now consider the situation of A1, suppose that A2 decides to stay and face the enemy, to defend his position, in this case it is best for A1 to flee, he will survive without getting hurt. The matrix shows that he will obtain a higher ranking which is 3 rather than 2, ($3 > 2$). A similar situation results if we consider A2 fleeing, then it is best for A1 to flee also, he will survive the battle, although he may face imprisonment for fleeing in the face of the enemy attack.

So, for each individual artillery man it would be at an advantage to flee, despite the actions of the other. Following the reasoning of iterative elimination carried out with the PD example, we note that from the point of view of utility maximisation, both artillery men would be better off if they stayed in their respective positions and faced the enemy, even though this may seem paradoxical showing that the outcome of individual rational action is being suboptimal. Ullmann-Margalit argues that in this example rationality appears to point to outcomes that are not optimal, morality binds the individual gunners to stay and fight, to face the enemy and defend their positions, so it seems as if the function of morality is to assist us in preventing the failures of rationality.

Ullmann-Margalit claims that it is the function of the artillery men that binds them to stay and not to follow the rational choice of fleeing and saving their lives. We note that there are several criticisms that can be made towards such a functional approach towards our understanding of moral behaviour. First, there are recognised hurdles with the functionalist approach when it is utilised to explain areas of moral behaviour in the social sciences. One needs to be aware of the fact that if a practice or an institution follows certain functions this does not necessarily explain the origin or the maintenance of such a function. For example, the function of the public education system may be to provide an opportunity for the citizens of the state to obtain an education. However, until there is a causal connection between the apparent function and its bringing about into existence and maintaining the education system, no real explanation has been provided by following such reasoning.

Similarly, one may argue that even though moral norms may contribute to bringing about certain outcomes that are not realisable through uncoordinated individual rational action, no real explanation can be said to be existing unless one demonstrates that somehow the function of providing education plays a motivating force in contributing towards human action and is therefore causally effective in the consequences that may arise with possible beneficial outcomes, that are also optimal in a game theoretic framework.

Secondly, it is questionable if morality corresponds to mutually advantageous outcomes. We are often morally motivated to act in ways that cause disadvantages to people around us. A common example taken is the prohibition against slavery that is willingly entered into by an individual. Such a situation may have certain advantages to the slave as he will be able to acquire enough funds to settle due debts and for the master also as he would have some assistance for housework on a daily basis; however this is not only prohibited morally but is also illegal.

Thirdly, the functionalist viewpoint tends to assume that the moral demands appear to be in conflict with our individual rational decision making abilities, according to the above artillery example the morality of both artillery men is supposed to correct the pending sub-optimal option that results due to so-called 'rational deliberation'. However on the functionalist account of morality, moral agents appear to act irrationally as the artillery men exclude any considerations of guilt-avoidance or regret or even their patriotic feelings towards their country and fellow soldiers, this tends to beg the question, "Why be moral?"; functionalism does not appear to provide us with any answer to this question.

This leads us to the fourth and final criticism. The objectives of functionalism seem to be to provide us with explanations for the emergence and persistence of moral norms by observing the functions that are performed by agents. However, morality as a guide to action, the normative approach is not really focused upon. Clearly, one can argue and question if there is any difference between (i) ascertaining what the function(s) of morality are and (b) questioning whether a certain group of norms are to be followed in a situation facing the agent, but we are at a loss on how this question can be answered if a functionalist approach is taken as our viewpoint.

We note that in the artillery example discussed above, the question of moral motivation arises as a kind of deliberation that exists outside of individual rational choice, that is not really entertained as part of the rational choice process. Is there any way that we can understand morality as part of individual rational choice? The approach that has been used and the one that we will discuss below regards morality as the intended interaction between ideal circumstances and rational agents. This follows similar ideas to those developed in the 'social contract' where morality arises as a consequence of a kind of bargaining process. As a result of this, the introduction of game theory, primarily those areas concerned with bargaining and cooperative game theory have generated interest in the development of social contract theory. The works of John Harsanyi, Richard Braithwaite, John Rawls, Brian Barry and David Gauthier have formulated versions of such a theory (Harsanyi, 1955; Braithwaite, 1955; Barry, 1965; Rawls, 1971; Gauthier, 1986).

By utilising areas of bargaining theory into a game theoretic framework, there has been an attempt to show that:

1. When rational agents are in an idealised bargaining situation, they will tend to find an agreement on a distribution of benefits of cooperations amongst themselves.
2. There is a particular form to this distribution.
3. This distribution determines what is just.
4. Rational agents will agree with the conditions of the bargaining process.

Let us discuss how such a bargaining situation is characterised in the artillery example above. Without any co-operation, the artillery men are bound to flee and face captivity for the remainder of their lives. If we now suppose that it may be possible for the two artillery men to come to some kind of binding agreements (however, the problem is that it is not obvious how the benefits of such a co-operative agreement could be distributed) then we could consider an example where the artillery men could come to a situation where they decide to flee with a probability of $1/3$ and stay and face the enemy with a probability of $2/3$. The numbers $(0, 1, 2, 3)$ in the matrix in Figure 3 only represent the rankings of the outcomes, by assigning probabilities we are now providing greater information concerning the relative rankings of the different outcomes.

The utility of "2", for example, that stands for the co-operative outcome and means that the artillery man is in a position where he is indifferent between this particular outcome and the alternative "0" which stands for the worst outcome with a probability $1/3$ and "3" which is the best outcome with a probability $2/3$. In this case, we can define what is called a 'Bargaining Area' which represents a range of different outcomes. Suppose now, that the artillery men are provided with a pair of dice and they agree to throw both dice and be governed by the following conditions:

1. If a total is equal to 6 or less, then A1 will flee therefore achieving his utility ranking of 3.
2. If both dice total more than 6 then A1 will stay and face the enemy thereby achieving the worst utility ranking of 0.

The expected utility of this deal for A1 can now be calculated as follows:

$$5/12 (3) + 7/12 (0) = 1.25$$

While A2 can expect $3 - 1.25 = 1.75$ from this situation. Thus the artillery men can realise an entire range of outcomes forming the shaded green area called the bargaining area as shown in figure 4.

Figure 4: Illustrating the Bargaining Area.

From this bargaining area we note that the optimal outcome of the agreement between both gunners will be $(2, 2)$. Each outcome that provided the artillery men an expected utility of greater than 1 appears to be rationally acceptable. The application of bargaining theory to ethical reasoning, therefore, provides further detail to PD type cases, we now consider how this has been developed further by the work of David Gauthier in explaining the contractarian approach.

THE CONTRACTARIANISM APPROACH

The most influential areas of contractarian theory are those developed by David Gauthier presented in his seminal work, "Morals by Agreement" (Gauthier, 1986). This work utilises extensively game and bargaining theory ideas, some of which have been touched above and attempts to provide us with an answer to the question posed earlier, namely "Why be moral?". In this respect, we could say that Gauthier's contractarian theory does contain several differences compared with the work developed by Rawls, Harsanyi, and others, we discuss some of these ideas presented by Gauthier briefly here.

The first part of this work provides us with a general account of "practical reason" and the "condition of human-kind" in relation to moral motivation; Gauthier (1986) then continues to discuss the "principles of conduct" that rational agents should agree upon – a kind of contract similar to the social contract. The third part of his work is a much discussed account concerning "practical rationality"; indeed this part is essential to Gauthier's argument as it aims to demonstrate that everyone under "normal circumstances" has some reason to accept the constraints imposed by those principles, this part essentially presents the answer to the question "Why be moral?". Finally, Gauthier argues that the principles of conduct, that he details, are indeed principles of morality.

Gauthier points out that we tend to misconceive practical rationality. He argues that the main aim of rationality is not to determine our principles of decision making – for example, to select the best option at every time when we are faced with choices. Gauthier asserts that our aim should be to reason in ways that are utility maximising. He develops his arguments in "Morals by Agreement" in terms of "dispositions to choose" and precisely by utilising the idea of "constrained maximisation" - the disposition to co-operate with others even when circumstances dictate that defecting is the more advantageous option as in the arbitrary example.

If we assume, as Gauthier does, that agents are at an advantageous position in certain circumstances when acting in certain ways that are not maximising utilities, the question we then face is to explain how to act as such a 'constrained maximiser' is economically rational. Gauthier argues that in certain circumstances if our tendencies to choose are rational, then it follows that our choices determined by these tendencies are also rational. For Gauthier, if a course of action is better, from a constrained point of view, then any other action that may follow may also be rational to adopt given certain conditions. However, to carry out such an action may not be in the best interests of all concerned.

Gauthier, therefore, seeks to establish that if a mode of deliberation is rational, then acting accordingly can also be rational, even if such an action requires performances that are not considered to be optimal from the game theoretical point of view. If Gauthier's arguments carry substance, we note that it can, therefore, be rational to adhere to certain norms even though they may require following some course of actions that are not the best from the point of view of utility maximisation. Gauthier's work has been developed further by different revisionist accounts of practical rationality (see Gauthier (1994), (1996), (1998a and b), and for revisionist accounts the interested reader can see for example McClennen (1990)). Gauthier's defence of 'constrained maximisation' represents a significant revision of standard game theoretic approaches to such decision making problems. Most game theorists argue that Gauthier's idea concerning "modes of deliberation" under "constrained maximisation" conditions can be modelled in a more complex kind of decision game theory (Binmore, 1994).

However, there are doubts whether a game theoretic approach to our comprehension of the bargaining process can assist us to predict the outcome of the deliberation of rational agents involved in required decisions given certain conditions. Such approaches to our understanding of the decision making process tends to assume that there is a unique outcome with rational motivations to such deliberations, while this may be plausible under certain circumstances, generally this is far from obvious in most situations.

Gauthier's development may be acceptable in explaining the emergence of fundamental norms arising from hypothetical situations in which large groups of agents are bargaining over principles. Nonetheless, one could say that it is questionable whether this is an appropriate methodology to follow in modelling the process of rational choice leading to the emergence of particular moral norms being followed. At this point, it is appropriate to consider an alternative viewpoint being offered by Evolutionary Game Theory (EGT). EGT is a third mode in which game theory has been developed to its application to ethics.

THE EVOLUTIONARY GAME THEORY APPROACH

Some important insights into how certain kinds of game equilibria can lead to the evolution of choices were offered by the development of Evolutionary Game Theory (EGT) and its application to ethics. Gintis (2000, 2009) states that "evolutionary game theory is a universal language for the unification of behavioural sciences". There are several good examples of such unifying works; Binmore (1998, 2005a) models social history as a series of convergent situations that arise due to efficient equilibria encountered in different transaction games.

EGT has also been successfully applied to the field of evolutionary biology where species and/or genes are treated as agents/players from a game theoretical viewpoint. In EGT, moves amongst a choice of possible options are not arrived at by economically rational agents. Instead, agents are disposed to particular strategies by their very nature, and the success and failure of a particular strategy is determined by the number of copies of itself that it will have to play in the games of succeeding generations, given a population in which other kinds of strategies that an agent acts are also distributed at particular frequencies. In this approach, the players represent strategies and the individuals are the executors of these strategies who get the benefits/costs associated with the different outcomes. To elaborate on this further, our discussion will now follow the work of Skyrms (2004), who introduces the idea of "replicator dynamics". As an example, let us first consider how natural selection works for any animal that may have heritable features.

From evidence provided in biological studies, it can be seen that any animal that may have heritable features that are likely to increase its expected number of offspring in a given environment, has the tendency of leaving a greater number of off-spring than others. So long as the environment remains stable, this off-spring are more likely to inherit the features in question, and therefore the proportion of these features in the population will gradually increase with successive generations. A natural question arising now is the following: how does game theory assist us to explain such evolutionary dominance?

One important characteristic of an organism's environment is to maximise reproductive fitness by adopting certain strategies that provide optimal solutions given the strategies of other populations. In EGT we no longer think of agents as choosing strategies when they leave one game to enter another; the focus is not to find the equilibrium of games, instead the aim is to seek equilibria that are further explored as they gradually alter over time. We therefore model strategies as competing against each other and we say that one strategy is more efficient than the other if it has more copies of itself in the next generation; this allows us to investigate the modification and change in a distribution of strategies in the population at the moment the sequence of games unfolds (Skyrms, 2004). So, we note that instead of understanding morality as the consequence of complex processes involving bargaining between informed agents following rationality, the focus now in EGT is rather to move away from such an approach and see morality as an unintended consequence of interactions between agents that arises due to repeated interactions between small groups. As a result of this, rather than assuming full information and rationality as a condition amongst agents, EGT requires less taxing assumptions concerning cognitive and deliberative skills of agents compared with earlier game theoretic models.

Let us now consider the following example to illustrate these ideas further. The example comes from Rousseau who describes the state of nature as being similar to a stag hunt (Rousseau 1964, pp. 166-167). Skyrms (2004) also considers a more contemporary treatment of a similar game situation and considers the stag hunt as modelling the social contract.

Imagine two hunters, who are faced with a hunting decision, they either choose to hunt for rabbits (during their hunt their chances of hunting for rabbits are not influenced by the actions of others around them) or they can choose to hunt for stag. However, since both hunters prefer to have venison for dinner, and if they were to hunt for stag they will only be successful if the other hunter also agrees to hunt for stag. So, the options open to them are to hunt for stags or to hunt for rabbits. Figure 5 illustrates the utility matrix for the different options. We noted earlier that in the PD case, there is a conflict between individual rationality and mutual benefit; in a similar way in the stag hunt example what is rational for one hunter to choose will depend upon his beliefs about what the other one will choose.

		NUMBER 1	
		Stag	Hare
NUMBER 2	Stag	3,3	0,2
	Hare	2,0	2,2

Figure 5: Illustrating the Example of the Stag Hunt.

Let the two hunters be represented by the numbers 1 and 2 as in the diagram. So, the pair (Stag, Stag) will represent the utility ranking (3, 3) where both hunters hunt for stag. We note that there are two rational outcomes to the stag hunt:

1. Either both hunters hunt the stag as a team.
2. Each hunts rabbits by himself.

Each would prefer to cooperate in hunting the stag, but if the other player's motives or actions are uncertain, the rabbit hunt is a risk-free alternative. If they both go for option 1, they have to both cooperate to hunt and catch the stag. When we solved the PD, we saw that one strategy was better regardless of what the other player was doing. Let us see if this dominant solution concept works here. If hunter 2 goes for a stag then hunter 1 should go for a stag as well since $3 > 2$. But if hunter 2 is going after rabbits then hunter 1 should hunt rabbits as well since $2 > 0$. So, we note that sometimes hunter 1 prefers hunting stag and other times he prefers hunting rabbits; thus our dominant solution concept is not sufficient for this game and consequently we must introduce another method for attaining a solution to the game.

Remembering that conditions for attaining the Nash equilibrium are as follows:

1. A game is said to be in NE when no player has any incentive to deviate from his strategy given the strategies of all other players.
2. At least one NE exists in all finite games.

In our example let us look at the (stag, stag) outcome, and we note that this is a NE point following the iterative reasoning developed earlier. Now we note that if we consider (rabbit, rabbit) this too is a NE because hunter 1 would prefer to hunt for rabbits if the second hunter is hunting for rabbits as well. These examples illustrate that cooperation in game theoretic modelling is an important aspect to consider as the evolution of cooperation in such models is an important indicator for determining more than one NE.

We note from our discussion above that EGT can be utilised to explain how cooperation may emerge amongst rational agents to assist in reaching their aims. Skyrms (1996) and other authors have also shown that agents with self-interested motives may also develop reasoning frameworks such as the 'Golden Mean' and versions of Gauthier's 'constrained maximisation' idea under relevant circumstances. These developments tend to support the view that evolution tends to favour the emergence of certain patterns of behaviour that tend to conform to moral norms and also the development of cognitive frameworks that tend to support characteristics related to moral reasoning.

Most of the authors who have developed the evolutionary approach using game theory point to several advantages of such an approach compared to the ones we discussed earlier, and these advantages can be briefly listed as follows:

1. In EGT norms arise as unintended consequences due to repeated actions of the rational agents.
2. There is no consequential connection among the efficiency of outcomes and morality in the evolutionary game theory paradigm. We could therefore say that one advantage with respect to Gauthier's claims for 'constrained maximisation' is that since in EGT the focus is on equilibrium and not determining efficient situations concerning utility maximisation, the 'function' of moral norms are to select stable equilibria. So, following the norm is individually rational.

However, there are also several reasons that point us to be more careful concerning the claims and successes relating to the evolutionary approach. As the functionalist approach, the focus of EGT is one explanation to explain the emergence of norms and their stability, but it does not supply us with any kind of framework to criticise and evaluate these moral norms. It is, therefore, not clear and evident to what extent such a paradigm offers to us an alternative to the moral theories already available. Another important criticism that has been pointed at the application of EGT to ethical decision making is the idea of the rational agent; this criticism can be thought of as a general criticism relating to the application of the whole of game theory to our understanding of ethical norms. A rational agent is a kind of one-dimensional entity who is supposed to determine preference rankings over possible outcomes. However, differences of character and other psychological variables concerning the agent are not really considered in the application process. Recent developments in game theory have made efforts to incorporate the notion of the player's "reputation" to explain the emergence of cooperation in iterated plays of games such as that seen in the PD case, while further examples of this can be found in Kreps and Wilson (1982).

CONCLUSION

The aim of this paper has generally been two-fold. First, to provide a concise historical introduction to the field of game theory leading to the discussion of the Prisoner's Dilemma and the use of pay off matrices. Secondly, three approaches where the application of game theory to ethical decision making are considered and their criticisms lead us to the development of Evolutionary Game Theory. The functionalist approach presents the argument developed by Ullmann-Margalit that the function of morality is to assist us in preventing failures of rationality. However, the five examples presented following this discussion show us that the emergence of norms could be explained by a more evolutionary approach. Careful reasoning indicates that an evolutionary game theory approach to ethical decision making may provide us with a suitable framework that explains the emergence of norms used to guide agent behaviour during the ethical decision making process; however, such a framework does not assist us in formulating an evaluation of such norms for individual agents.

Despite the successful application of evolutionary game theory to decision making, a critical area that will require further development is to consider collection of agents forming a group, for example, a board of directors; in such a case the emergence of norms guiding ethical decision making will depend upon several different psychological factors concerning not only a single agent but the group of agents in question. Such differences will need to be incorporated into any kind of evolutionary game theoretic approach when considering collective ethical decision making. ■

REFERENCES

- Aumann R. and Hart S., ed. (1992) *Handbook of Game Theory with Economic Applications*, vol. 1, Amsterdam: Elsevier.
- Axelrod, R. (1984) *The Evolution of Cooperation*. New York: Basic Books.
- Barry, B. (1965) *Political Argument, International library of philosophy and scientific method*. London: Routledge & Kegan Paul.
- Binmore, K. (1994) *Playing Fair (Game Theory and the Social Contract, vol 1)*. Cambridge: The MIT Press.
- Id. (1998) *Just Playing (Game Theory and the Social Contract, vol. 2)*. Cambridge: The MIT Press.
- Id. (2005) *Natural Justice*. New York: Oxford University Press.
- Braithwaite, R. B. (1955) *Theory of games as a tool for the moral philosopher*. Cambridge: Cambridge University Press.
- Brennan, G. and Philip P. (2004) *The Economy of Esteem: An Essay on Civil and Political Society*. Oxford: Oxford University Press.
- Cremaschi, S. (2005) *L'etica del Novecento*. Roma: Carocci.
- Id. (2007) *L'etica moderna*. Roma: Carocci.
- Dreier, J. (2004) *Decision Theory and Morality*. In *The Oxford Handbook of Rationality*. Oxford: Oxford University Press.
- Epstein, R. (2009) *The theory of Gambling and Statistical Logic*. Boston: Academic Press.
- Esther, M. S. (2004) 'How Psychology Made Its (Limited) Way Back Into Economics', *History of Political Economy*, 36, pp. 735-760.
- Frank, R. (1988) *Passions within Reason*. London: W. W. Norton & Company, Inc.
- Gauthier, D. (1967) 'Morality and Advantage', *Philosophical Review*, 76, pp. 460-475.
- Id. (1969) *The Logic of 'Leviathan': the Moral and Political Theory of Thomas Hobbes*. Oxford: Clarendon Press.
- Id. (1986) *Morals by Agreement*. Oxford: Clarendon Press.
- Id. (1994) 'Assure and Threaten', *Ethics*, 104, pp. 690-721.
- Id. (1996) 'Commitment and Choice: An Essay on the Rationality of Plans', in Farina, F., Hahn, F. and Vannucci, S. (eds.) *Ethics, Rationality, and Economic Behaviour*. Oxford: Oxford University Press, pp. 217-244.
- Id. (1998a) 'Intention and Deliberation', in Danielson, P. (ed.) *Modeling Rationality, Morality, and Evolution*. Oxford: Oxford University Press, pp. 41-54.
- Id. (1998b) 'Rethinking the Toxin Puzzle', in Coleman, J. and Morris, C. (eds.) *Rational Commitment and Social Justice: Essays for Gregory Kavka*. Cambridge: Cambridge University Press, pp. 47-58.
- Id. (2008) 'Friends, Reasons and Morals', in Verbeek, B. (ed.) *Reasons and Intentions*. Aldershot: Ashgate, pp. 17-36.
- Gibbard, A. (1990) *Wise Choices, Apt Feelings: A Theory of Normative Judgment*. Cambridge, MA: Harvard University Press.
- Hald, A. (2003) *A History of Probability and Statistics and Their Applications before 1750*. Hoboken, NJ: Wiley.
- Id. (1998) *A History of Mathematical Statistics from 1750 to 1930*. New York: Wiley.
- Hampton, J. (1986) *Hobbes and the Social Contract Tradition*. Cambridge: Cambridge University Press.
- Hardin, R. (1988) 'Bargaining for Justice', *Social Philosophy and Policy*, 5, pp. 65-74.
- Id. (2003) *Indeterminacy and Society*. Princeton: Princeton University Press.
- Hargreaves-Heap, S. and Varonfakis Y. (1995) *Game Theory: A Critical Introduction*. London: Routledge.
- Harsanyi, J. C. (1955) Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility. *Journal of Political Economy* 63: pp. 309-321.
- Kavka, G. (1986) *Hobbesian Moral and Political Theory*. Princeton: Princeton University Press.
- Kreps, D. M. and Robert W. (1982) 'Reputation and Imperfect Information', *Journal of Economic Theory*, 27(2), pp. 253-79.
- Kuhn, S. T. (2004) 'Reflections on Ethics and Game Theory', *Synthese*, 141(1), pp. 1-44.

- Lewis, D. (1969) *Convention: A Philosophical Study*. Cambridge: Harvard University Press.
- Mackie, J. (1977) *Ethics*. London: Penguin Books Ltd.
- McClennen, E. F. (1990) *Rationality and Dynamic Choice: Foundational Explorations*. Cambridge: Cambridge University Press.
- Mele, A. and Piers R., eds. (2004) *The Oxford Handbook of Rationality*. Oxford: Oxford University Press.
- Morris, C. W. (1999) 'What is this Thing Called 'Reputation'?', *Business Ethics Quarterly*, 9(1), pp. 87-102.
- Nash, J. (1950) 'The Bargaining Problem', *Econometrica*, 18(2), pp. 155-162.
- Osborne, M. J. (2002) *An Introduction to Game Theory*. Oxford: Oxford University Press.
- Poundstone, W. (1992) *Prisoner's Dilemma*. New York: Doubleday, 1992
- Rawls, J. (1971) *A Theory of Justice*. Cambridge: Harvard University Press.
- Ross, D. (2012) 'Game Theory', in Zalta, E.N. (ed.) *The Stanford Encyclopedia of Philosophy, Winter Edition*. [Online]. Available at: <http://plato.stanford.edu/archives/win2012/entries/game-theory/> (Accessed: 15 March 2014).
- Rousseau, Jean-Jacques. (1964) *Discours sur l'origine et les fondements dell'inégalité parmi les hommes*, vol. III, *Oeuvres complètes*. Paris: Éditions Gallimard.
- Rubinstein, A. (1982) 'Perfect Equilibrium in a Bargaining Model', *Econometrica*, 50(1): pp. 97-109.
- Skyrms, B. (1996) *Evolution of the Social Contract*. Cambridge: Cambridge University Press.
- Id. (2004) *The Stag Hunt and the Evolution of Social Structure*. Cambridge: Cambridge University Press.
- Smith, J. M. (1982) *Evolution and the Theory of Games*. Cambridge: Cambridge University Press.
- Solomon, R. (1999) 'Game Theory as a Model for Business and Business Ethics', *Business Ethics Quarterly*, 9(1), pp. 11-29.
- Sugden, R. (1986) *The Economics of Rights, Co-operation and Welfare*. Oxford: Basil Blackwell.
- Id. (1991) 'Rational Bargaining', in Bacharach, M. and Hurley, S. (eds.) *Foundations of Decision Theory*. Oxford: Basil Blackwell.
- Id. (2001) 'The Evolutionary Turn in Game Theory', *Journal of Economic Methodology*, 8(1), pp. 113-30.
- Ullmann-Margalit, E. (1977) *The Emergence of Norms*. Oxford: Oxford, University Press.
- Vanderschraaf, P. (1998) 'The Informal Game Theory in Hume's Account of Convention', *Economics and Philosophy*, 14(2), pp. 215-247.
- Velleman, J. D. (1997) 'Deciding How to Decide', in Culitty, G. and Gaut, B. (eds.) *Ethics and Practical Reason*. Oxford: Oxford University Press, pp. 29-52.
- Verbeek, B. (2002) *Instrumental Rationality and Moral Philosophy: An essay on the virtues of cooperation*. Dordrecht: Kluwer Academic Publishers.
- Id. (2007) 'The Authority of Norms', *American Philosophical Quarterly*, 44(3), pp. 245-258.
- Id. (2008) 'Conventions and Moral Norms: the Legacy of Lewis', *Topoi*, 27(1-2), pp. 73-86.
- Verbeek, B. and Morris, C. (2010) 'Game Theory and Ethics', in Zalta, E.N. (ed.) *The Stanford Encyclopedia of Philosophy, Summer Edition*. [Online]. Available at: <http://plato.stanford.edu/archives/sum2010/entries/game-ethics/> (Accessed: 15 March 2014).
- Von Neuman J. and Morgenstern O. (1944) *Theory of Games and Economic Behavior*. Princeton: Princeton University Press.