Chapter 2

Iterated Prisoner's Dilemma and Evolutionary Game Theory

Siang Yew Chong[1], Jan Humble[2], Graham Kendall[2], Jiawei Li[2,3], Xin Yao[1]

University of Birmingham[1], University of Nottingham[2], Harbin Institute of Technology[3]

2.1. Introduction

The prisoner's dilemma is a type of non-zero-sum game in which two players try to maximize their payoff by cooperating with, or betraying the other player. The term non-zero-sum indicates that whatever benefits accrue to one player do not necessarily imply similar penalties imposed on the other player. The Prisoner's dilemma was originally framed by Merrill Flood and Melvin Dresher working at RAND Corporation in 1950. Albert W. Tucker formalized the game with prison sentence payoffs and gave it the "Prisoner's Dilemma" name. The classical prisoner's dilemma (PD) is as follows:

> Two suspects, A and B, are arrested by the police. The police have insufficient evidence for a conviction, and, having separated both prisoners, visit each of them to offer the same deal: if one testifies for the prosecution against the other and the other remains silent, the betrayer goes free and the silent accomplice receives the full 10-year sentence. If both stay silent, the police can sentence both prisoners to only six months in jail for a minor charge. If each betrays the other, each will receive a two-year sentence. Each prisoner must make the choice of whether to betray the other or to remain silent. However, neither prisoner knows for sure what choice the other prisoner will make. So the question this dilemma poses is: What will happen? How will the prisoners act?

The general form of the PD is represented as the following matrix [Scodel et al. (1959)]:

		Prisoner 2	
		Cooperate	**Defect**
Prisoner 1	**Cooperate**	*(R,R)*	*(S,T)*
	Defect	*(T,S)*	*(P,P)*

where *R, S, T, and P* denote Reward for mutual cooperation, Sucker's payoff, Temptation to defect, and Punishment for mutual defection respectively, and $T > R > P > S$ and $R > 1/2(S+T)$. The two constraints motivate each player to play noncooperatively and prevent any incentive to alternate between cooperation and defection [Rapoport (1966, 1999)].

Neither prisoner knows the choice of his accomplice. Even if they were able to talk to each other, neither could be sure that he could trust the other. The "dilemma" faced by the prisoners here is that, whatever the other does, each is better off confessing than remaining silent. However, the payoff when both confess is worse for each player than the outcome they would have received if they had both remained silent. Traditional game theory predicts the outcome of PD be mutual defection based on the concept of Nash equilibrium. To defect is dominant because if both players choose to defect, no player has anything to gain by changing their own strategy [Hardin (1968); Nash (1950, 1951, 1996)].

In the Iterated Prisoner's Dilemma (IPD) game, two players have to choose their mutual strategy repeatedly, and have memory of their previous behaviors. Because players who defect in one round can be "punished" by defections in subsequent rounds and those who cooperate can be rewarded by cooperation, the appropriate strategy for self-interested players is no longer obvious in IPD games. If the precise length of an IPD is known to the players, then the optimal strategy is to defect on each round (often called All Defect of AllD) [Luce and Raiffa (1957)]. This single rational play strategy which is deduced from propagating the single stage Nash equilibrium of mutual defection backwards through every stage of the game prevents players from cooperating to achieve higher payoffs [Selten (1965, 1983, 1988); Noldeke and Samuelson (1993)]. If the game has infinite length or at least the players are not aware of the length of the game, backward induction is no longer effective and there exists the possibility that cooperation can take place. In fact, there is still controversy about whether or not backward induction can be applied to infinite (or finite) IPDs [Sobel (1975, 1976); Kavka (1986); Becker and Cudd (1990); Binmore

(1997); Binmore et al. (2002); Bovens (1997)]. However, in IPD experiments, it was not uncommon to see people cooperate to gain a greater payoff not only in repeated games but even in one-shot games [Cooper et al. (1996); Croson (2000); Davis and Holt (1999); Milinski and Wedekind (1998)]. Traditional game theory interprets the cooperation phenomena in IPDs by means of reputation [Fudenberg and Maskin (1986); Kreps and Wilson (1982); Milgrom and Roberts (1982)], incomplete information [Harsanyi (1967); Kreps et al. (1982; Sarin (1999)], or bounded rationality [Anthonisen (1999); Harborne (1997); Radner (1980, 1986); Simon (1955, 1990); Vegaredondo (1994)].

Evolutionary game theory differs from classical game theory in respect of focusing on the dynamics of strategy change in a population more than the properties of strategy equilibrium. In evolutionary game theory, IPD is an ideal experimental platform for the problem as to how cooperation occurs and persists, which is considered to be impossible in the static or deterministic environment. IPD attracted wide interest after Robert Axelrod's famous book 'The Evolution of Cooperation'. In 1979, Robert Axelrod organized a prisoner's dilemma tournaments and solicited strategies from game theorists [Axelrod (1980a, 1980b)]. Each of the 14 entries competed against all others (including itself) over a sequence of 200 moves. The specific payoff function used is as follows.

		Prisoner 2	
		Cooperate	**Defect**
Prisoner 1	**Cooperate**	*(3,3)*	*(0,5)*
	Defect	*(5,0)*	*(1,1)*

The winner of the tournament was 'tit-for-tat' (TFT) submitted by Anatol Rapoport. TFT always cooperate on the first move and then mimics whatever the other player did on the previous move. In a second tournament with 62 entries, again the winner was TFT. Axelrod discovered that 'greedy' strategies tended to do very poorly in the long run while 'altruistic' strategies did better when PD were repeated over a long period of time with many players. Then genetic algorithms were introduced to show how these altruistic strategies evolve in the populations that are initially dominated by selfishness. The prisoner's dilemma is therefore of interest to the social sciences such as economics, politics and sociology, and to the biological sciences such as ethology and evolutionary biology, as well to the applied mathematics such as evolutionary computing. Many social and natural processes, for example arm race between states and price setting for duopolistic firms, have been abstracted into models in which independent groups or individuals are engaged in PD games [Brelis (1992); Bunn and Payne (1988); Hauser (1992); Hemelrijk (1991); Surowiecki (2004)].

The optimal strategy for the one-shot PD game is simply defection. However, in the IPD game the optimal strategy depends upon the strategies of the possible opponents. For example, the strategy of Always Cooperate (AllC) is dominated by the strategy of Always Defect (AllD), and AllD is optimal in a population consisting of AllD and AllC. However, in a population consisting of AllD, AllC, and TFT, AllD is not necessarily the optimal strategy. It appears that all the strategies in the population determine which strategy is optimal. Although TFT was proved to be efficient in lots of IPD tournaments and was long considered to be the best basic strategy, it could be defeated in some specific circumstances [Beaufils, Delahaye and Mathieu (1996); Wu and Axelrod (1994)]. Therefore, there is lasting interest for game theorists to find optimal strategies or at least novel strategies which outperform TFT in IPD tournaments.

Since Axelrod, two types of approaches are developed to test the efficiency or robustness of a strategy and further to derive optimal strategies:

(1) Round-robin tournaments.

(2) Evolutionary dynamics.

Round-robin tournament shows the efficiency of a strategy in competing with others, while ecological simulation illustrates the evolutionary robustness of a strategy in terms of the number of descendants or survivability in a certain environment. Lots of novel strategies have been developed and analyzed by means of these approaches.

By using round-robin tournaments, the interactions between different strategies can be observed and analyzed. If the statistical distribution of opposing strategies can be determined an optimal counter-strategy can be derived mathematically. For example, if the population consists of 50% TFT and 50% AllC, the optimal strategy should cooperate with TFT and defect with AllC in order to maximize the payoff. It is easy to design such a strategy that defects in the first two moves, and then plays always *C* if the opponent defected on the second move, otherwise plays always *D*. A similar concept in analyzing optimal strategy is *Bayesian Nash equilibrium* which is widely used in experimental economics [Bedford and Meilijson (1997); Gilboa and Schmeidler (2001); Kagel and Roth (1995); Kalai and Lehrer (1993); Rubinstein (1998);]. In evolutionary dynamics, the processes like natural selection are simulated where individuals with low scores die off, and those with high scores flourish. The evolutionary rule that describes what future states follow from the current state is fixed and deterministic: for a given time interval only one future state follows from the current state [Katok and Hasselblatt (1996)]. The common methodology of the evolution rule is replicator equations that assume infinite populations, continuous time, complete mixing and that strategies breed true. Given a population of strategies and the dynamic equations, the evolutionary process can be simulated, and how strategies evolve in the population over a short or long time period can be shown. Optimal strategies can be developed in this way [Axelrod (1987); Darwen and Yao (1995, 1996, 2001); Lindgren (1992); Miller (1996)].

2.2. Strategies in IPD tournaments

Axelrod is the first who attempts to search for efficient strategies by means of IPD tournament [Axelrod (1980a, 1980b)]. TFT had long been studied as a strategy of IPD game [Komorita, Sheposh and Braver (1968); Rapoport and Chammah (1965)]. However, it is after Axelrod's tournaments that TFT become well-known.

According to Axelrod, several conditions are necessary for a strategy to be successful. These conditions include:

Nice

The most important condition is that the strategy must be "nice". That is, it will not defect before its opponent does. Almost all of the top-scoring strategies are nice. Therefore a selfish strategy will never defect first.

Retaliating

Axelrod contended that a successful strategy must not be a blind optimist. It must always retaliate. An example of a non-retaliating strategy is AllC. This is a very bad choice, as "nasty" strategies will ruthlessly exploit such strategies.

Forgiving

Another quality of successful strategies is that they must be forgiving. Though they will retaliate, they will fall back to cooperating if the opponent does not continue to defect. This stops long runs of revenge and counter-revenge, thus maximising payoffs.

Clear

The last quality is being clear, that is making it easier for other strategies to predict its behavior so as to facilitate mutually cooperation. Stochastic strategies, however, are not clear because of the uncertainty in their choice.

In a further study, Axelrod noted that just a few of the 62 entries in the second tournament have reasonably influence on the performance of a given strategy. He utilized eight strategies as opponents for a simulated evolving population based on a genetic algorithm approach [Axelrod (1987)]. The population consisted of deterministic strategies that use outcomes of the three previous moves to determine a current move. The simulation was conducted using a population of 20 strategies from a total of 2^{70} strategies executed repeatedly against the eight representatives. Mutation and crossover were used to generate new strategies. The typical results indicated that populations initially generated mutual defection, but subsequently evolved toward mutual cooperation. Moreover, most of the strategies that evolved in the simulation actually resemble TFT, having the properties of 'Nice', 'Forgiving', and 'Retaliating'.

Although TFT has been considered to be the most successful strategy in IPD for several decades, there still is some controversy about it. There seems to be a lack of theoretical explanation for the strategies like TFT in traditional game theory. TFT is not subgame perfect, and there are always subgame perfect equilibria that dominate TFT according to the Folk Theorem [Binmore (1992); Hargreaves and Varoufakis (1995); Myerson (1991); Rubinstein (1979); Selten (1965, 1975)]. On the other hand, whether or not TFT is the most efficient singleton strategy in IPD game is still unclear; therefore, many researchers are attempting to develop novel strategies that can outperform TFT.

2.2.1. Heterogeneous TFTs

Since TFT had such success in IPD tournaments and experiments, it is natural to draw the conclusion that TFT may be improved by slightly modifying its rule. Many heterogeneous TFTs have been developed in order to overcome TFT's shortcoming or to adapt to a certain environment, for example IPD with noise. Among these strategies, Tit-for-Two-Tats (TFTT), Generous TFT (GTFT), and Contrite TFT (CTFT) are examples.

A situation that TFT does not handle well is a long series of mutual retaliations evoked by an occasional defection. The deadlock can be broken if the co-player behaves more generously than TFT and forgives at least one defection. TFTT retaliates with defection only after two successive defections and thus attempts to avoid becoming involved in mutual retaliations. Usually, TFTT performs well in a population with more cooperative strategies but does poorly in a population with more permanently defective strategies. Similar to TFTT, Benevolent TFT (BTFT) always cooperates after cooperation and normally defects after defection, but occasionally BTFT responds to defection by cooperation in order to break up a series of mutual obstruction [Komorita, Sheposh and Braver (1968)]. In experiments of Manarini (1998) and Micko (1997), fixed interval BTFT strategies were shown to be superior to, or at least equivalent to, TFT in terms of cooperation as well as in terms of cumulative pay-off. However, BTFT tends to produce irregularly alternating exploitations and sometimes resort mutual retaliations.

Allowing some percentage of the other player's defections to go unpunished has been widely accepted as a good way to cope with noise [Molander (1985); May (1987); Axelrod and Dion (1988); Bendor et al. (1991); Godfray (1992); Wu and Axelrod (1994)]. A reciprocating strategy such as TFT can be modified to forgive the other player's defection with a certain ratio in order to decrease the influence of noise. GTFT behaves like TFT but cooperates with the probability of $q = \min[1-(T-R)/(R-S),(R-P)/(T-P)]$ when it would otherwise defect. This prevents a single error from echoing indefinitely. For example, in the case of T=5, R=3, P=1, and S=0, q=1/3. GTFT is said to take over the dominant position of the population of homogeneous TFT strategies in an evolutionary environment with noise [Nowak and Sigmund (1992)].

In a noisy environment, retaliating unintended defection often leads to permanent bilateral retaliation. Therefore, forgiving defection evoked by unintended defection allows a quick way to recover from error. It is based upon the idea that one shouldn't be provoked by the other player's response to one's own unintended defection [Sugden (1986); Boyd (1989)]. The strategy of CTFT has three states: "contrite", "content" and "provoked". It begins in a content state, with cooperation and stays there unless there is a unilateral defection. If it was the victim while content, it becomes provoked and defects until a cooperation from other player causes it to become content. If it was the defector while content, it becomes contrite and cooperates. When contrite, it becomes content only after it has successfully cooperated. CTFT can correct its unintended defection in a noisy environment. If one of two CTFT players defects, the defecting player will contritely cooperate on the next move and the other player will defect, and then both will be content to cooperate on the following move. However, CTFT is not effective at correcting the other player's error. For example, if CTFT is playing TFT and the TFT player defected by accident, the retaliation will continue until another error occurs. In an ecological simulation with noise, GTFT and CTFT competed with the 63 rules of the

Second Round of the Computer Tournament for the Prisoner's Dilemma [Axelrod (1984)]. CTFT is the dominant strategy, becoming 97% of the population at generation 2000 [Wu and Axelrod (1994)].

2.2.2. Pavlov (Win-Stay Lose-Shift)

A possible drawback of TFT is that it performs poorly in a noisy environment. Assume that a population of TFT strategies plays IPD with one another in a noisy environment, where every choice may be occasionally implemented in error. Although a TFT strategy cooperates with its twin at the beginning, it would get out of cooperation as soon as the other player's action is misinterpreted, and then this induces the other player's defection in the next round. Therefore, after an error, the result of the game turns out to be a *CD*, *DC*, *CD* … cycle. If a second error happens, the outcome is as likely to fall into defection as it is to resume cooperation. Cooperation between TFT strategies is easy to break even in the case of low noise frequency [Donninger (1986); Kraines and Kraines (1995)].

The Pavlov strategy, also known as Win-Stay Lose-Shift or Simpleton [Rapoport and Chammah (1965)], has been shown to outperform TFT in the environment with noise [Fudenberg and Maskin (1990); Kraines and Kraines (1995, 2000)]. Pavlov cooperates when both sides have cooperated or defected on the previous move, and defects otherwise. Pavlov, as well as TFT, are a type of memory-one strategies where players only remember and make use of their own move and their opponent's move on the last round. The major difference between Pavlov and TFT is that Pavlov will choose COOPERATE after a defection as against TFT's DEFECT, and this helps Pavlov resume cooperation with those cooperative strategies, such as TFT, in a noisy environment. When restricted to an environment of memory-one agents interacting in iterated Prisoners Dilemma games with a 1% noise level, Pavlov is the only cooperative strategy and one of the very few that cannot be invaded by a similar strategy [Nowak and Sigmund (1993, 1995)].

Simulation of evolutionary dynamics of win-stay lose-shift strategies shows that these strategies are able to adapt to the uncertain environment even when the noise level is high [Posch (1997)]. In simulated stochastic memory-one strategies for the IPD games, Nowak and Sigmund (1993, 1995) report that cooperative agents using a Pavlov type strategy eventually dominate a random population. Memory-one strategies can be expressed in the form of $S(p_1, p_2, p_3, p_4)$, where p_1 denotes the probability of playing C (*Cooperate*) after a *CC* outcome, p_2 denotes the probability of playing C after a *CD* outcome, p_3 denotes the probability of playing C after a *DC* outcome, and p_4 denotes the probability of playing *C* after a *DD* outcome. Most of the well-known strategies can be expressed in this form. For example, AllC = *S*(1, 1, 1, 1), AllD = *S*(0, 0, 0, 0), TFT = *S*(1, 0, 1, 0), Pavlov = *S*(1, 0, 0, 1). Noise is conveniently introduced by restricting the conditional probabilities p_i to range between 0 and 1. For example, *S*(0.999, 0.001, 0.999, 0.001) is a TFT strategy with 0.001 probability of being misinterpreted. In a computer simulation with a population using the totally random strategy *S*(0.5, 0.5, 0.5, 0.5), win-stay lose-shift strategy shows its evolutionary robustness in noisy environment. After each 100 generations from a total of 10^7 generations, 10^5 mutant strategies that are generated at random are introduced. Simulation results show that the populations are dominated by win-stay lose-shift strategy in 33 of a total of 40 simulations. TFT strategies perform poorly in large part because they do not exploit overly cooperate strategies.

Simulations reveal that Pavlov loses against AllD but can invade TFT, and that Pavlov cannot be invaded by AllD [Milinski (1993)].

2.2.3. Gradual

The *Gradual* strategies are like TFT but respond to the opponent with a gradual pattern. This strategy acts as TFT, except when it is time to forgive and remember the past. It uses cooperation on the first move and then continues to do so as long as the other player cooperates. Then after the first defection of the other player, it defects one time and cooperates two times; after the second defection of the opponent, it defects two times and cooperates two times, ... after the n^{th} defection it reacts with n consecutive defections and then calms down its opponent with two cooperations [Beaufils, Delahaye and Mathieu (1996)].

Both round-robin competitions and ecological evolution experiments are conducted in order to compare the performance of Gradual with TFT. Gradual wins in experiments where round-robin competitions are conducted with several well-known strategies, such as TFT and GRIM. In ecological evolutionary experiments, gradual and TFT have the same type of evolution, with the difference of quantity in favor of gradual, which is far away in front of all other survivors when the population is stabilised. However, it is efficient to demonstrate that TFT is not always the best, but not efficient to prove that Gradual always outperforms TFT. Gradual receives fewer points than TFT while interacting with AllD because Gradual forgives too many defections. Therefore, if there are lots of defecting strategies like AllD in the competition, it would be possible that TFT outperforms Gradual in this case.

Beaufils, Delahaye and Mathieu (1996) try to improve the performance of Gradual by using a genetic algorithm. 19 different genes are used and a fitness function evaluates the quality of the strategies. Several new strategies are found after 150 generations of evolution. One of them beats Gradual and TFT in round-robin tournament, as well as in an ecological simulation. In the two cases it has finished first just in front of Gradual, TFT being two or three places behind, with a wide gap in the score, or in the size of the stabilised population.

The evolution dynamics of populations including Gradual has also been studied in Delahaye and Mathieu (1996), Doebeli and Knowlton (1998), Glomba, Filak, and Kwasnicka (2005), Beaufils, Delahaye, and Mathieu (1996).

2.2.4. Adaptive Strategies

From the viewpoint of automation, the strategies in IPD games can be regarded as automatic agents with or without feedback mechanisms. Most well-known IPD strategies are not adaptive because their responses to any certain opponent are fixed. It is impossible to improve their performance since the parameters of their responding mechanism cannot be adjusted. However, there are still some strategies in IPDs which are adaptive. Although there is still no experimental evidence of adaptive strategies outperforming non-adaptive ones in IPD games, adaptive strategies are worth studying since creatures with higher intelligence are all adaptive.

There have been two approaches to developing adaptive strategies. Firstly, adaptive mechanisms can be implemented by making the parameters of a non-adaptive strategy adjustable. Secondly, new adaptive strategies can be developed by using evolutionary computation, reinforcement learning, and other computational techniques [Darwen and Yao (1995, 1996)].

Tzafestas (2000a, 2000b) introduced adaptive tit-for-tat (ATFT) strategy that embedded an adaptive factor into the conventional TFT strategy. ATFT keeps the advantages of tit-for-tat in the sense of retaliating and forgiving, and implements some behavioural gradualness that would show as fewer oscillations between Cooperate and Defect. It uses an estimate of the opponent's behavior, whether cooperative or defecting, and reacts to it in a tit-for-tat manner. To represent degrees of cooperation and defection, a continuous variable named '*world*' which ranges from 0 (total defection) to 1 (total cooperation) is applied. The ATFT strategy can then be formulated as a simple model:

If (opponent played C in the last cycle) then
        world = world + r(1-world)*
else
        world = world + r(0-world)*
If (world >= 0.5) play C, else play D

r is the adaptation rate here. The TFT strategy corresponds to the case of r=1 (immediate convergence to the opponent's current move). Clearly, ATFT is an extension of the conventional TFT strategy. By simulating the spatial IPD games between ATFT, AllD, AllC, and TFT on 2D grid, it shows that ATFT is fairly stable and resistant to perturbations. Since the use of a fairly small adaptation rate r will allow more gradual behavior, ATFT tends to be more robust than TFT in a noisy environment.

Since evolutionary computation has been widely used in simulating the dynamics of IPD games, it is natural to consider obtaining IPD strategies directly by using evolutionary approaches [Lindgren (1991); Fogel (1993); Darwen and Yao (1995, 1996)]. Axelrod (1987) studied how to find effective strategies by using genetic algorithms as simulation method. He established an initial population of strategies that is deterministic and uses the outcome of the three previous moves to make a choice in the current move. By means of playing IPD games between one another, successful strategies are selected to have more offspring. Then the new population will display patterns of behavior that are more like those of the successful strategies of the previous population, and less like those of the unsuccessful ones. As the evolution process continues, the strategies with relatively high scores will flourish while the unsuccessful strategies die out. Simulation results show that most of the strategies that were evolved in the simulation actually resemble TFT and does substantially better than TFT. However, it would not be accurate to say that these strategies are better than TFT because they are probably not very robust in other environments [Axelrod (1987)].

Many researchers have found that evolved strategies may lack robustness, i.e., the strategies did well against the local population, but when something new and innovative appeared they fail [Lindgren (1991); Fogel (1993)]. Darwen and Yao (1996) applied a technique to prevent the genetic algorithm from converging to a single optimum and attempted to develop new IPD strategies without human intervention. It concludes that adding static opponents to the round robin tournament improves the results of final population.

Optimal strategies can be determined only if the strategy of the opponent is known. By means of reinforcement learning, model-based strategies with the ability of on-line identification of an opponent can be built [Sandholm and Crites (1996); Freund et al. (1995); Schmidhuber (1996)]. How can a player acquire a model of its opponent's strategy? One possible source of information available for the player is the history of the game. Another possible source of information is observed games between the opponent and other agents. In the case of IPD games, a player can infer an opponent's model based on the outcome of the past moves and then adapts its strategy during the game. Reinforcement learning (RL) is based on the idea that the tendency to produce an action should be strengthened if it produces favorable results, and weakened if it produces unfavorable results [Watkins (1989); Watkins and Dayan (1992); Kaelbling and Moore (1996)]. A model-based RL approach generates expectation about the opponent's behavior by making use of a model of its strategy [Carmel and Markovitch (1997, 1998)]. It is well suited for use in IPD tournament against an unknown opponent because of its small computational complexity. The major problem in designing a model-based strategy (MBS) is the risk involved in the exploration, and thus the issue of *exploitation versus exploration*. An exploring action taken by the MBS tests unfamiliar aspects of the opponent which can yield a more accurate model of the opponent. However, this action also carries the risk of putting the MBS into a much worse position. For example, in order to distinguish the strategy ALLC from GRIM and TFT in IPD tournament, a MBS has to defect at least once and therefore loses the chance to cooperate with GRIM. The exploratory action affects not only the current payoff but also the future rewards [Berry and Fristedt (1985)]. There have been several approaches developed to solve this problem [Berry and Fristedt (1985); Gittins (1989); Sutton (1990); Narendra and Thathachar (1989); Kaelbling (1993); Moore and Atkeson (1993); Carmel and Markovitch (1998)]. Since possible strategies for a repeated game is usually infinite, computational complexity is another problem that needs to be addressed [Ben-porath (1990); Carmel and Markovitch (1998)]. There is seldom a record of an effective MBS in round-robin IPD tournaments. However, the strategy that won competition 4 in 2005 IPD tournament, Adaptive Pavlov, is such a strategy [Prisoner's dilemma tournament result (2005)]. Furthermore, it seems that each of the strategies that ranked above TFT incorporated a mechanism to explore the opponent.

2.2.5. Group Strategies.

In the 2004 IPD competition [20th-anniversary Iterated Prisoner's Dilemma competition], a team from Southampton University led by Professor N. Jennings introduced a group of strategies, which proved to be more successful than Tit-for-Tat (see chapter 9).

The group of strategies were designed to recognise each other through a known series of five to ten moves at the start. Once two Southampton players recognized each other, they would act as their 'master' or 'slave' roles – a master will always defect while a slave will always cooperate in order for the master to win the maximum points. If the program recognized that another player was not a Southampton entry, it would immediately defect to minimise the score of the oppositions. The Southampton group strategies succeeded in defeating any non-grouped strategies and won the top three positions in the competition [Prisoner's dilemma tournament result (2004)].

According to Grossman (2004), it was difficult to tell whether a group strategy would really beat TFT because most of the 'slave' group members received far lower scores than the average level and were ranked at the bottom of the table. The average score of the group strategies is not necessarily higher than that of TFT.

The significance of group strategies maybe lies in their evolutionarily characters. None of known strategies in IPD games is an evolutionarily stable strategy. [Boyd and Loberbaum (1987)] The strategies that are most likely to be evolutionarily stable, such as AllD or GRIM, can resist the invasion of some types of strategies but cannot resist the invasion of others. For example, a small group of TFT strategies can not invade a large population of AllD; however, STFT can do. There exists the possibility that TFT can successfully invade a population of AllD indirectly. Suppose that a large population of AllD is continuously attacked by small groups of STFT. Because every invasion makes a small positive proportion of STFT remain in the population of AllD, the number of STFT increases gradually. When the number of STFT is large enough, a small group of TFT can successfully invade and AllD will die out.

However, group strategies may be evolutionarily stable. By means of cooperating with group members and defecting against non-group members, a population of group strategies can prevent any foreigner from successfully invading. This is, perhaps, the real value of group strategies.

2.3. Evolutionary Dynamics in Games

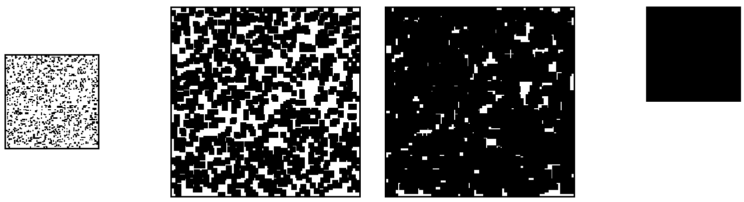
Traditional game theorists have developed several effective approaches to study static games based on the assumption of rationality. By using Neumann-Morgenstern utility, refinement of Nash equilibrium, and reasoning, both cooperative and non-cooperative games are analyzed within a theoretical framework. However, in the area of repeated games, especially in games where dynamics are concerned, few approaches from traditional game theory are available.

Evolutionary game theory provides novel approaches to solve dynamic games. If the precise length of an IPD is known to the players, then the optimal strategy is to defect on each round. If the game has infinite length or at least the players are not aware of the length of the game, there exists the possibility that cooperation happens [Dugatkin (1989); Darwen and Yao (2002); Akiyama and

Kaneko (1995); Doebeli, Blarer, and Ackermann (1997); Axelrod (1999); Glance and Huberman (1993, 1994); Ikegami and Kaneko (1990); Schweitzer (2002)].

Nowak and May (1992, 1993) showed that cooperators and defectors coexist in certain circumstances by introducing spatial evolutionary games, in which two types of players – cooperators who always cooperate and defectors who always defect are placed in a two-dimensional spatial array. In each round, every individual plays the PD game with its immediate neighbors. The selection scheme is that each lattice is occupied either by its original owner or by one of the neighbors, depending on who scores the highest total in that round, and so on to the next round of the game. Simulation results show that cooperators remain a considerable percentage of the population in some cases, and defector can invade any a lattice but can not occupy the whole area.

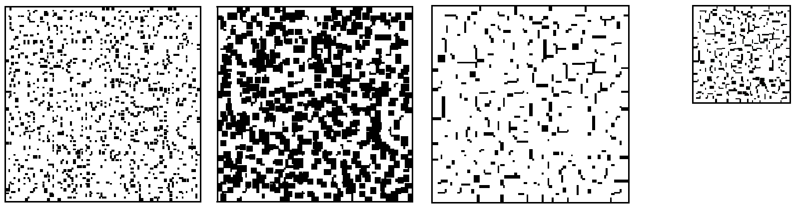
When the parameters of the payoff matrix are set to be $T = 2.8$, $R = 1.1$, $P = 0.1$, and $S = 0$ and the initial state is set to be a random mixture of the two types of strategies, the evolutionary dynamics of the local interaction model lead to a state where each player chooses the strategy Defect, the only ESS in the prisoner's dilemma. Fig.2.1 shows that the population converges to a state where everyone defects and no Cooperate strategy survives after 5 generations.



Generation 1 Generation 2 Generation 3 Generation 6

Fig.2.1 Spatial Prisoner’s Dilemma with the values $T = 2.8$, $R = 1.1$, $P = 0.1$, and $S = 0$ [Nowak and May (1993)].

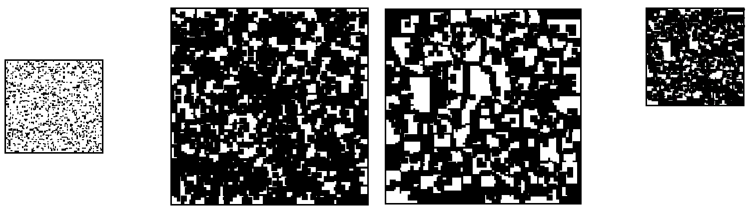
However, when the parameters of the payoff matrix are set to $T = 1.2$, $R = 1.1$, $P = 0.1$, and $S = 0$, the evolutionary dynamics do not converge to the stable state of defection. Instead, a stable oscillating state where cooperators and defectors coexist and some regions are occupied in turn by different strategies.



Generation 1 Generation 2 Generation 19 Generation 20

Fig.2.2 Spatial Prisoner's Dilemma with the values T = 1.2, R = 1.1, P = 0.1, and S = 0 [Nowak and May (1993)].

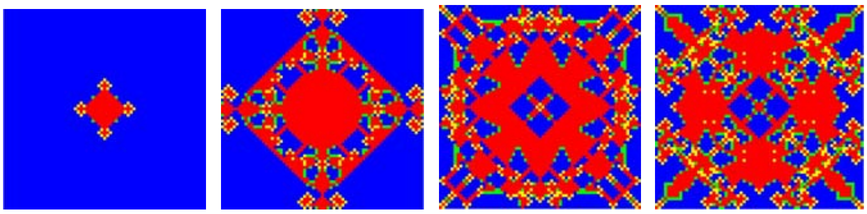
Moreover, when the parameters of payoff matrix are set to be T = 1.61, R = 1.01, P = 0.01, and S = 0, the evolutionary dynamics lead to a chaotic state: regions occupied predominantly by Cooperators may be successfully invaded by Defectors, and regions occupied predominantly by Defectors may be successfully invaded by Cooperators.



Generation 1 Generation 3 Generation 13 Generation 15

Fig.2.3 Spatial Prisoner's Dilemma with the values T = 1.61, R = 1.01, P = 0.01, and S = 0 [Nowak and May (1993)].

If the starting configurations are sufficiently symmetrical, this spatial version of the PD game can generate chaotically changing spatial patterns, in which cooperators and defectors both persist indefinitely. For example, if we set R = 1, P = 0.01, S = 0.0 and T = 1.4, and initial state is that every individual in a square 69×69 lattice is a cooperator except a defector in the middle of the lattice. The structure of the evolving lattice varies like a kaleidoscope, and the ever-changing sequences of spatial patterns can be very beautiful, as shown in Fig. 2.4. The role of the spatial interaction in the evolution of cooperation is further studied by Durrett and Levin (1998), Schweitzer, Behera, and Mühlenbein (2002), Ifti, Killingback, and Doebeli (2004).



Generation 10 Generation 40 Generation 4000 Generation 6000

Fig.2.4 Spatial Prisoner's Dilemma with the values T = 1.4, R = 1, P = 0.01, and S = 0, where Blue, Red, Green, and Yellow denote cooperators, defectors, new cooperators, and new defectors respectively [Nowak and May (1993)].

2.3.1. Evolutionary Stable Strategy

Just like the Nash equilibrium in traditional game theory, Evolutionarily Stable Strategy (ESS) is an important concept used in theoretical analysis of evolutionary games. According to Maynard Smith (1982), an ESS is a strategy such that, if all the members of a population adopt it, then no mutant strategy could invade the population under the influence of natural selection. ESS can be seen as an equilibrium refinement to the Nash equilibrium. Suppose that a player in a game can choose between two strategies: *I* and *J*. Let E(*J, I*) denote the payoff he receives if he chooses the strategy *J* while all other players choose *I*.

Then, the strategy I is evolutionarily stable if either

1. $E(I, I) > E(J, I)$, or
2. $E(I, I) = E(J, I)$ and $E(I, J) > E(J, J)$

is true for all $I \neq J$ [Maynard Smith and Price (1973); Maynard Smith (1982)].

Thomas (1985) rewrites the definition of ESS in a different form. Following the terminology given in the first definition above, we have

1. $E(I, I) \geq E(J, I)$, and
2. $E(I, J) > E(J, J)$

From this alternative form of definition, we find that ESS is just a subset of Nash equilibrium. The benefit of this refinement of Nash equilibrium is not just to eliminate those weak Nash equilibrium, but to provide an efficient mathematical tool for dynamic games. Following the concept of ESS, two approaches to evolutionary game theory have been developed. The first approach directly applies the concept of ESS to analyze static games. The second approach simulates the evolutionary process of dynamic games by constructing a dynamic model, which may take into consideration the factors of the population, replication dynamics, and strategy fitness.

As an example of using ESS in static games, consider the problem of the Hawk-Dove game. Two types of animals employ different means to obtain resources (a favorable habitat, for example)—Hawk always fights for some resources while Dove never fights. Let V denote the value of the resources, which can be considered the Darwinian fitness of an individual obtaining the resource, described by Maynard Smith (1982). Let $E(H,D)$ denote the payoff to a Hawk against a Dove opponent. If we assume that (1) whenever two Hawks meet, conflict eventually results and the two individuals are equally likely to be injured, (2) the cost of the conflict reduces individual fitness by some constant value C, (3) when a Hawk meets a Dove, the Dove immediately retreats and the Hawk obtains the resource, and (4) when two Doves meet the resource is shared equally between them, the payoff matrix for Hawk-Dove game will look like this,

	Hawk	*Dove*
Hawk	*((V-C/2,(V-C)/2)*	*(V,0)*
Dove	*(0,V)*	*(v/2,V/2)*

Then, it is easy to verify that the strategy Dove is not an ESS because there is $E(D,D) < E(H,D)$, which means that a pure population of Doves can be invaded by a Hawk mutant. In the case that the value V of the resource is greater than the cost C of injury, the strategy Hawk is an ESS because there is $E(H,H) > E(D,H)$, which means that a Dove mutant can not invade a group of Hawks. If $V < C$ is true, the Hawk-Dove game becomes the game of Chicken originated from the 1955 movie *Rebel without a cause*. Neither pure Hawk nor pure Dove is ESS in this game. However, there is an ESS if mixed strategies are permitted [Bishop and Cannings (1978)].

An *evolutionarily stable state* is a dynamical property of a population to return to using a strategy, or mix of strategies, if it is perturbed from that strategy, or mix of strategies [Maynard Smith (1982)]. A population of ESS must be evolutionarily stable because it is impossible for any mutant to invade it. Many biologists and sociologists attempt to explain animal and human behavior and social structures in terms of ESS [Cohen and Machalek (1988); Mealey (1995)]. However, a dynamic game is not necessarily converging to a stable state in which ESS is prevalent. For example, using a spatial model in which each individual plays the Prisoner's Dilemma with his or her neighbors, Nowak and May (1992, 1993) show that the result of the game depends on the specific form of the payoff matrix.

Now imagine a population of players in a society where each one has to play Prisoner's Dilemma with another and whether or not one can survive and breed is determined by his payoff in the game. How will the population evolve? In order to show the evolutionary process of the population, a model of dynamics that takes time t into consideration is needed.

2.3.2. Genetic Algorithm

A genetic algorithm maintains a population of sample points from the search space. Each point is represented by a string of characters, known as genotype [Holland (1975, 1992, 1995)]. By defining a fitness function to evaluate them, genetic algorithm proceeds to initialize a population of solutions randomly, and then improve it through repetitive application of mutation, crossover, and selection operators.

The common methodology to study the evolutionary dynamics in games is through replicator equations. Replicator equations usually assume infinite populations, continuous time, complete mixing and that strategies breed true [Taylor (1979); Maynard Smith (1982); Weibull (1995); Hofbauer and Sigmund (1998)]. Originated from biology and then introduced into evolutionary game theory by Taylor and Jonker (1978), replicator equations provide a continuous dynamic model for evolutionary games.

Consider a population of *n* types of strategies, and let x_i be the frequency of type *i*. Let *A* be the $n \times n$ payoff matrix. With the assumptions that the population is infinitely large and strategies are completely mixed and x_i are differentiable functions of time *t*, a strategy's fitness, or expected payoff can be written as $(Ax)_i$ if strategies meet one another randomly. The average fitness of the population as a whole can be written as $x^T Ax$. Then, the replicator equation is

$$\dot{x}_i = x_i\left((Ax)_i - x^T Ax\right) \tag{2.1}$$

Evolutionary games with a replicator dynamic as described in (2.1) will converge to a result that strategies with strong fitness bloom in the population.

For the Prisoner's Dilemma, the expected fitness of the strategies Cooperate and Defect, E_C and E_D respectively, are

$$E_C = x_C R + x_D S \text{ , and } E_D = x_C T + x_D P \tag{2.2}$$

where x_C and x_D denote the proportions of the strategies of Cooperate and Defect in the population respectively. Let $\overline{E}$ denote the average fitness of the entire population, there is

$$\overline{E} = x_C E_C + x_D E_D \tag{2.3}$$

Then, the replicator equations for this game are

$$\frac{dx_C}{dt} = x_C(E_C - \overline{E}) \text{, } \frac{dx_D}{dt} = x_D(E_D - \overline{E}) \tag{2.4}$$

Since there is $T > R$ and $P > S$, $E_D - E_C = x_C(T - R) + x_D(P - S) > 0$ holds, and there must be $E_D > \overline{E} > E_C$. Therefore, there are $\frac{dx_C}{dt} < 0$ and $\frac{dx_D}{dt} > 0$. This means that the number of the strategies of Cooperate will always decline while the number of the strategies of Defect increases as the game goes on. Sooner or later, the proportion of the population choosing the strategy Cooperate will, in theory, become extinct.

Besides replicator dynamics, there exist other types of dynamics equations that can be used in modeling evolutionary systems [Akin (1993); Thomas (1985); Bomze (1998, 2002); Balkenborg and Schlag (2000); Cressman, Garay and Hofbauer (2001); Weibull (1995); Hofbauer (1996); Gilboa and Matsui (1991); Matsui (1992); Fudenberg and Levine (1998); Skyrms (1990); Swinkels (1993); Smith and Gray (1994)]. Lindgren (1995) and Hofbauer and Sigmund (2003) have given a comprehensive review of them.

In general, dynamic games are of great complexity. How an evolutionary system evolves depends not only on the population and dynamic structures but also on where the evolution starts. Because of dynamic interactions between multiple players, especially those players with intelligence, genetic algorithms may converge towards local optima rather than the global optimum. Also, operating on dynamic data sets is difficult as genomes begin to converge early on towards solutions which may no longer be valid for later data [Michalewicz (1999); Schmitt (2001)]. Analysis of the evolutionary dynamic systems is not just a problem of evolutionary game theory, but a new direction in applied mathematics [Garay and Hofbauer (2003); Gaunersdorfer (1992); Gaunersdorfer, Hofbauer, and Sigmund (1991); Hofbauer (1981, 1984, 1996); Krishna and Sjöström (1998); Plank (1997); Smith (1995); Zeeman (1993), Zeeman and Zeeman (2002, 2003)].

2.3.3. Strategies

What strategies should be involved in evolutionary dynamics is a difficult question. One approach is to take into consideration lots of representative strategies, for example Axelord (1984), Dacey and Pendegraft (1988), and Akimov and Soutchanski (1994), since it is impossible to enumerate all possible strategies. However, it is difficult to say what strategy should be included and which ones not, and there is little comparability between evolutionary processes with different strategies because the selection of strategies may have great influence on the outcome of the dynamics. Another approach is to study the interactions between specific strategies, for example Nowak and Sigmund (1990, 1992) and Goldstein and Freeman (1990). In this way, it is convenient to make clear the relationship between strategies in the evolutionary process; however, generality of complex evolutionary systems loses to some extent.

Strategies in PD games (or in non-PD games) can be characterized as either deterministic or stochastic. Deterministic strategies leave nothing to change and respond to the opponent with predetermined actions; stochastic strategies, however, leave some uncertainty in their choices.

Oskamp (1971) presents a thorough review of the early studies on the strategies involved in PD games and non-PD games, for example AllD, TFT, and lots of stochastic strategies that play *C* or *D* with some certain probabilities [Lave (1965); Bixenstine, Potash, and Wilson (1963); Solomon (1960); Crumbaugh and Evans (1967); Wilson (1969); Oskamp and Perlman (1965); Sermat (1967); Heller (1967); Knapp and Podell (1968); Lynch (1968); Swingle and Coady (1967); Whitworth and Lucker (1969)].

After Axelord's IPD tournament, memory-one strategies that interact with the opponent according to both sides' behavior in the previous move become prevalent. TFT, Pavlov, Grim Trigger, and many other memory-one strategies are analyzed in varies of environment: round-robin tournaments, evolutionary dynamics with or without noise [Nowak and Sigmund (1990, 1992, 1993); Pollock (1989); Wedekind and Milinski (1996); Milinski and Wedekind (1998); Sigmund (1995); Stephens (2000); Stephens, Mclinn and Stevens (2002); Sandholm and Crites (1996); Doebeli and Knowlton (1998); Brauchli, Killingback and Doebeli (1999); Sasaki, Taylor and Fudenberg (2000)].

No strategy has been shown to be superior in a dynamic environment, and even deterministic cooperators can invade defectors in specific circumstances. It is not sensible to discuss which strategy is best unless the context is defined. Comparing TFT with GTFT, Grim (1995) suggests that, in the non-stochastic Axelrod models, it is TFT that is the general winner; within a purely stochastic model, the greater generosity of GTFT pays off; in a model with both stochastic and spatial elements, a level of generosity twice that of GTFT proves optimal. Pavlov has an obvious advantage over TFT in noisy environments [Nowak and Sigmund (1993); Kraines and Kraines (1995)]. In an evolutionary process where AllC, AllD, TFT, and GTFT strategies are involved, evolution starts off toward defection but then veers toward cooperation. TFT strategies play a key role in invading the population of defectors. However, GTFT strategies and then more generous AllCs gradually become dominant once cooperation is widely established, and this provides an opportunity to AllD to invade again [Nowak and Sigmund (1992)]. Additionally, Selten and Stoecker (1986) have studied the end game behavior in finite IPD supergames, and find that cooperative behaviors last until shortly before the end of the supergame.

Machine Learning approaches have been introduced into evolutionary game theory to develop adaptive strategies, especially those for IPD games [Carmel and Markovitch (1996, 1997, 1998); Littman (1994); Tekol and Acan (2003); Hingston and Kendall (2004)]. Adaptive strategies, at least in theory, have obvious advantages over fixed strategies. Among the set of adaptive strategies, there may be an evolutionarily stable strategy for IPD games and potential winner of future IPD tournaments.

2.3.4. Population

Population size and structure are of great importance in evolutionary dynamics. In general, evolutionary processes in a large population are quite different from that in small populations [Maynard Smith (1982); Fogel and Fogel (1995); Fogel, Fogel and Andrew (1997, 1998); Ficici and Pollack (2000)].

Young and Foster (1991) have studied stochastic effects in a population consisting of three strategies: AllD, AllC, and TFT. They show that the outcome of the evolutionary process depends crucially on the amount of noise, which is inversely proportional to the population size. The more people there are, the more that random variations in their behavior are smoothed out in the population proportions. For large populations, the system tends to drift from TFT to AllC, which is then invaded by AllD. As a result, most of the players behave as AllD, even though initially most players may have started as TFT. They conclude that cooperation is viable in the short run, but not stable in the long run in a large population.

Boyd and Richerson (1988, 1989) suggest that reciprocity is unlikely to evolve in large groups as a result of natural selection because reciprocators punish defection by withholding future cooperation which will penalize other cooperators in the group. Boyd and Richerson (1990, 1992) analyze a model in which the punishment response to defection is directed solely at defectors. In this model, cooperation reinforced by retribution can lead to the evolution of cooperation in different ways. There is the possibility that strategies which cooperate and punish defectors, strategies which cooperate only if punished, and strategies which cooperate but do not punish coexist in the long run, as well as the possibility that only one type exists. As the group size grows larger, however, the conditions for co-operators' surviving becomes more difficult.

Glance and Huberman (1994) discuss how to achieve cooperation in groups of various sizes in *n*-person PD games and find that there are two stable points in large groups: either there is a great deal or very little cooperation. Cooperation is more likely in smaller groups than in larger ones and there is greater cooperation when players are allowed more communication with each other. Large random fluctuations are related to group size. Groups beyond a certain size may experience increased difficulty of informational exchange and coordination; further, reneging on contracts is possible to be prevalent as each member may expect that the effect of his/her action on other members will be diluted. However, Dugatkin (1990) finds that cooperation may invade large populations more easily than smaller ones, but it is likely to represent a smaller proportion of the population in larger groups. In order to consider the potential importance of the relationship between population size and cooperative behaviour, two N-person game theoretical models are presented. The results show that cooperation is frequently not a pure evolutionarily stable strategy, and that many metapopulations should be polymorphic for both cooperators and defectors.

It is well accepted that communication among members of a society leads to more cooperative behaviors [Insko et al. (1987); Orbell, Kragt, and Dawes (1988)]. Insko et al. (1987, 1988, 1990, 1993) explore the role of communication on interindividual-intergroup discontinuity in the context of the extended PD game that adds a third withdrawal choice to the usual cooperative and uncooperative choices, and interindividual-intergroup discontinuity is the tendency of intergroup relations to be more competitive and less cooperative than interindividual relations. The lesser tendency of individuals to cooperate when there is no communication with the opponent partially explains the group discontinuity.

Choice and refusal of partners may accelerate the emergence of cooperation. Experiments have shown that people who are given the option of playing or not are more likely to choose to play if they are themselves planning to cooperate. More cooperative players are more likely to anticipate that others will be cooperative [Orbell and Dawes (1993)]. Defecting players are possible to be alienated by cooperators [Schuessler (1989); Kitcher (1992); Batali and Kitcher (1994)]. In the N-person PD game, it may be that players can change groups if they don't satisfy the size of their groups [Hirshleifer and Rasmusen (1989)]. The option of choice and refusal of partners in IPD means that players will attempt to select partners rationally. Analytical studies reveal that the subtle interplay between choice and refusal in N-player IPD games can result in various long-run player interaction patterns: mutual cooperation; mixed mutual cooperation and mutual defection; parasitism; and wallflower seclusion. Simulation studies indicate that choice and refusal can accelerate the emergence of cooperation in evolutionary IPD games [Stanley, Ashlock, and Tesfatsion (1994); Stanley, Ashlock and Smucker (1995)].

The effects of freedom to play, reciprocity and interchange, coalitions and alliances, and various sizes of groups on evolution are also studied [Orbell and Robyn (1993); Alexander and Frans (1992); Glance and Bernardo (1994); Hemelrijk (1991)]. In a specific scenario, the prestructuration of the population may determine the evolution of the patterns of interaction that constitute the final social structure [Eckert, Koch, and Mitlöhner (2005)].

2.3.5. Selection Scheme

Evolutionary selection schemes can be characterized as either generational or steady-state schemes [Thierens (1997)]. Generational schemes that are widely used in evolutionary game theory mean that each generation of a population is replaced in one step by a new generation. In a system with a steady-state scheme only a small percentage of the population is replaced in each generation. Evolutionary selection schemes can be further subdivided as pure or elitist selection schemes in terms of whether or not there is an overlap between successive generations. Pure selection schemes allow no overlap between successive generations: all parents from previous generation are discarded and the next generation is filled entirely with offspring from these parents. In elitist schemes, subsequent generations may be the same: parents with higher fitness are transferred to the next generation and only poorly performing parents are replaced [Mitchell (1996)].

Pure selection schemes are commonly used in IPD research [Axelord (1987); Axelrod and Dion (1988); Huberman and Glance (1993); Akimov and Soutchanski (1994); Mill (1996)]. These schemes use fitness-proportional selection of the parents in combination with single-point crossover or use a random uniform simple set to select the fittest agent to produce offspring. A robust society of cooperators emerges only if the level of competition between the players is neither too small nor too large. In elitist selection schemes, the population is firstly shuffled randomly and partitioned into pairs of parents. Then, each pair of parents creates two offspring, and a local competition between parents and their offspring is held. Finally, the best two players of each pair of parents are transferred to the next generation [Thierens and Goldberg (1994)]. In this case, stable societies of highly cooperative players evolve. It shows that a suitable model of the selection process is of crucial importance in terms of simulating real-world economic situations [Ficici, Melnik, and Pollack (2000); Bragt, Kemenade and Poutré (2001)].

Selection is clearly an important genetic operator, but opinion is divided over the importance of crossover versus mutation. Some argue that crossover is the most important, while mutation is only necessary to ensure that potential solutions are not lost [Grefenstette, Ramsey and Schultz (1990); Wilson (1987)]. Others argue that crossover in a largely uniform population only serves to propagate innovations originally found by mutation, and in a non-uniform population crossover is nearly always equivalent to a very large mutation [Spears (1992)].

2.4. Evolution of Cooperation.

A fundamental problem in evolutionary game theory is to explain how cooperation can emerge in a population of self-interested individuals. Axelrod (1984, 1987) attributes the reason of emergence of cooperation to the 'shadow of the future': the likelihood and importance of future interaction. This implies that rewards from cooperation should be mutually expected payoff and to cooperate is a rational choice for self-interested individuals [Martinez-Coll and Hirshleifer (1991)]. Axelrod's work has been subjected to a number of criticisms because his conclusions obviously conflict with traditional game theory [Binmore (1994, 1998)], as Nachbar's criticism that "Axelrod mistakenly ran an evolutionary simulation of the finitely repeated Prisoners' Dilemma. Since the use of a Nash equilibrium in the finitely repeated Prisoners' Dilemma necessarily results in both players always defecting, we then wouldn't need a computer simulation to know what would survive if every strategy were present in the initial population of entries. The winning strategies would never co-operate." [Nachbar (1992)]. There are also arguments that the conflict stems from the assumption of Von Neumann-Morgenstern utility. According to Spiro (1988), the problem with Axelrod's argument is the oft-discussed problem of interpersonal utility comparison. Axelrod's argument, and all game theoretic modeling, welfare economics, and utilitarian moral philosophy, in fact, would require that it be possible for one to measure and compare the utilities of different people. The problem with this assumption is that it is quite impossible to construct a scale of measurement for human preferences [Rothbard (1997)].

Although evolutionary game theory is aimed primarily towards dynamic games, while traditional game theory deals with non-dynamic games, there are still area of intersection, for instance in the field of repeated games. Furthermore, although evolutionary game theory mainly depends on experiments and computer simulations, its theoretical foundations, i.e. individual utility (or preference) and payoff-maximizing, stem from traditional game theory. Controversies about Axelrod's work reflect the bifurcation between evolutionary approaches and the basic assumptions of game theory. Based on the assumption of 'rational players', traditional game theory regards a finite repeated game as a combination of many singleton games. 'Backward induction' is applied in order to dissect the link between these singleton games, and then each of them can be analyzed statically [Harsanyi and Selten (1988)]. The concept of backward induction was first employed by Von Neumann and Morgenstern (1944) and then developed by Selten (1965, 1975) based on Nash equilibrium. First, one determines the optimal strategy of the player who makes the last move of the game. Then, the optimal action of the next-to-last moving player is determined taking the last player's action as given. The process continues in this way backwards through time until all players' actions have been determined. Subgame perfect Nash equilibrium deduced directly from backward induction is an equilibrium such that players' strategies constitute a Nash equilibrium in every subgame of the original game [Aumann (1995)]. Selten proved that any game which can be broken into "sub-games" containing a subset of all the available choices in the main game will have a subgame perfect Nash equilibrium. In the case of a finite number of iterations in IPD games, the unique subgame perfect Nash equilibrium is AllD. However, many psychological and economic experiments have shown that subjects would not necessarily apply a strategy like AllD [Kahn and Murnighan (1993); McKelvey and Palfrey (1992); Cooper et al. (1996)]. Game theorists explain these experimental results in terms of incomplete information, reputation, and bounded rationality, which are all based on theoretical analysis [Harsanyi (1967); Kreps et al. (1982); Simon (1990); Bolton (1991); Bolton and Ockenfels (2000); Binmore et al. (2002); Samuelson (2001)]. In some sense, Axelrod's work is a parallel of these explanations, but it seems that his approach is absolutely different. Before a soundly theoretical explanation can be established, the problem of how cooperation emerges is left unsolved.

As to the problem of how cooperation can persist during evolution, sufficient evidence has been provided to support the point that cooperation can survive and flourish in a wide range of circumstances if only some conditions are satisfied. Nowak and Sigmund (1990) have shown that cooperation can emerge among a population of randomly chosen reactive strategies, as long as a stochastic version of TFT is added to the population. If cooperators can recognize each other with the help of some label they can increase their payoff by interacting selectively with one another [Frank (1988)]. Social norms aid in cooperation in many ways [Bendor and Mookherjee (1990); Kandori (1992); Sethi and Somanathan (1996)]. As to the influence of payoff variations, Mueller (1988) finds that payoff settings with increasing values of T relative to P promote cooperative behaviour; while Fogel (1993) regards that smaller values for T

promote the evolution of cooperative behaviour. Nachbar (1992) selects a payoff setting strongly favouring the relative reward of cooperating and finds that this setting elicits an increased degree of cooperation. Kirchkamp (1995) finds that the value of S becomes less important with longer memory. Also, the effects of population structure, repetition, and noise have been studied [Hirshleifer and Coll (1988); Mueller (1988); Boyd (1989); Marinoff (1992); Hoffmann (2001)].

To end, we note that Binmore (1998) stated:

> "*…One simply cannot get by without learning the underlying theory. Without any knowledge of the theory, one has no way of assessing the reliability of a simulation and hence no idea of how much confidence to repose in the conclusions that it suggests*".

There is still a need for an underlying theory for IPD tournaments. Evolutionary game theory has provided us with many experimental approaches; however, better theoretical explanations are still needed. Even though IPD tournaments have been run for over 40 years, we suspect there will be more as we search for new strategies and new theories which explain the complex interactions that take place.

Finally, this review has been restricted to the IPD literature. Even so, we have not been able to include every article and there are, no doubt, omissions. However we hope that this chapter has provided enough information for the interested reader to follow up on.

References

Akimov V. and Soutchanski M. (1994) Automata simulation of N-person social dilemma games, *Journal of Conflict Resolution*, **38**, pp. 138-148.

Akin E. (1993) The general topology of dynamical systems, American Mathematics Society, Providence.

Akiyama E. and Kaneko K. (1995) Evolution of cooperation, differentiation, complexity and diversity in an iterated three-person game, *Artificial Life*, **2**, pp. 293-304.

Alexander H. and Frans B. (1992) *Coalitions and Alliances in Humans and Other Animals*. Oxford: Oxford University Press.

Anthonisen N. (1999) Strong rationalizability for two-player noncooperative games, *Economic Theory*, **13**, pp. 143-169.

Aumann R. (1995) Backward Induction and Common Knowledge of Rationality, Games and Economic Behavior, **18**, pp. 6-19.

Axelrod R. (1980a) Effective choice in the prisoner's dilemma, *Journal of Conflict Resolution*, **24**, pp. 3-25.
Axelrod R. (1980b) More effective choice in the prisoner's dilemma, *Journal of Conflict Resolution*, **24**, pp. 379-403.
Axelrod R. M. (1984). *The Evolution of Cooperation* (BASIC Books, New York).
Axelrod R. (1987) The evolution of strategies in the iterated prisoner's dilemma, In Davis L., Genetic Algorithms and Simulated Annealing, pp. 32-41.
Axelrod R. (1999) *The Complexity of Cooperation: Agent-based Models of Competition and Collaboration*. University Press, Princeton, NJ.
Axelrod R. and Dion D. (1988) The further evolution of cooperation, *Science*, **242**, pp. 1385-1390.
Axelrod R. and Hamilton W. (1981) The evolution of cooperation, *Science*, **211**, 4489, pp. 1390-1396.
Balkenborg D. and Schlag K. (2000) Evolutionarily stable sets, International *Journal of Game Theory*, **29**, pp. 571-595.
Batali J. and Kitcher P. (1994) Evolutionary dynamics of altruistic behaviour in optional and compulsory versions of the iterated prisoner's dilemma, In Rodney A. and Maes P. Artificial Life IV. MIT Press, pp. 343-348.
Beaufils B., Delahaye J., and Mathieu P. (1996) Our meeting with gradual: A good strategy for the iterated prisoner's dilemma, *Proceedings of the Artificial Life V*, pp. 202-209.
Becker N. and Cudd A. (1990) Indefinitely repeated games: a response to Carroll, *Theory and Decision*, **28**, pp. 189-195.
Bendor J. and Mookherjee D. (1990) Norms, third-party sanctions, and cooperation, Journal of Law, *Economics, and Organization*, **6**, pp. 33-63.
Bendor R., Kramer M., and Stout S. (1991) When in doubt: cooperation in a noisy prisoner's dilemma, *Journal of Conflict Resolution*, **35**, pp. 691-719.
Ben-porath E. (1990) The complexity of computing a best response automaton in repeated games with mixed strategies, *Games and Economic Behavior*, **2**, pp. 1-12.
Berry D. and Fristedt B. (1985) *Bandit problems: sequential allocation of experiments*. Chapman and Hall, London.
Binmore K. (1992) *Fun and games*. Lexington, MA: D.C. Heath and Company.
Binmore K. (1994) *Playing fair game theory and the social contract I*. MIT Press.
Binmore K. (1997) Rationality and backward induction, *Journal of Economic Methodology*, **4**, pp. 23-41.
Binmore K. (1998) Review of R. Axelrod's 'The complexity of cooperation: agent based models of competition and collaboration', *Journal of Artificial Societies and Social Simulation*, **1**, 1

Binmore K., McCarthy J., Ponti G., Samuelson L. and Shaked A. (2002) A backward induction experiment, *Journal of Economic Theory*, **104**, pp. 48-88.

Bishop, D. and Cannings, C. (1978) A generalized war of attrition, *Journal of Theoretical Biology*, **70**, pp. 85-124.

Bixenstine V., Potash H., and Wilson K. (1963) Effects of level of cooperative choice by the other player on choices in a Prisoner's Dilemma game, *Journal of Abnormal and Social Psychology*, **66**, pp. 308-313.

Bolton G. (1991) A comparative model of bargaining: theory and evidence, *The American Economic Review*, **81**, 5, pp. 1096-1136.

Bolton G. and Ockenfels A. (2000) ERC: a theory of equity, reciprocity, and competition, *The American Economic Review*, **90**, pp. 166-193.

Bomze I. (1998) Uniform barriers and evolutionarily stable sets, *Game Theory, Experience, Rationality*, pp. 225-244.

Bomze I. (2002) Regularity vs. degeneracy in dynamics, games, and optimization: a unified approach to different aspects, *SIAM Review*, **44**, pp. 394-414.

Boyd R. (1989) Mistakes allow evolutionary stability in the repeated prisoner's dilemma game, *Journal of Theoretical Biology*, **136**, 11, pp. 47-56.

Boyd R. (1992) The evolution of reciprocity when conditions vary, Harcourt A. and Frans B. (eds.) Alliance formation among male baboons: shopping for profitable partners. Oxford: Oxford University Press, pp. 473-489.

Boyd R. and Loberbaum J. (1987) No pure strategy is evolutionarily stable in the repeated Prisoner's Dilemma game, *Nature*, **327**, pp. 58-59.

Boyd R. and Richerson P. (1988) The evolution reciprocity in sizable groups, *Journal of Theoretical Biology*, **132**, pp. 337-356.

Boyd R. and Richerson P. (1989) The evolution of indirect reciprocity, *Social Networks*, **11**, pp. 213-236.

Boyd R. and Richerson P. (1990) Group selection among alternative evolutionarily stable strategies. *Journal of Theoretical Biology*, **145**, pp. 331-342.

Boyd R. and Richerson P. (1992) Punishment allows the evolution of cooperation (or anything else) in sizable groups, *Ethology and Sociobiology*, **13**, pp. 171-195.

Bovens L. (1997) The backward induction argument for the finite iterated prisoner's dilemma and the surprise exam paradox, *Analysis*, **57**, 3, pp. 179-186.

Bragt D., Kemenade C. and Poutré H. (2001) The influence of evolutionary selection schemes on the iterated prisoner's dilemma, *Computational Economics*, **17**, pp. 253-263.

Brauchli K., Killingback T. and Doebeli M. (1999) Evolution of cooperation in spatially structured populations, *Journal of Theoretical Biology*, **200**, pp. 405-417.

Brelis M. (1992) Reputed mobster defends his honor. *Boston Globe*, **1**, pp. 23.

Bunn G. and Payne R. (1988) Tit-for-tat and the negotiation of nuclear arms control, *Arms Control*, **9**, pp. 207-233.

Carmel D. and Markovitch S. (1996) Learning models of intelligent agents, *Proceedings of the 13th National Conference on Artificial Intelligence and the 8th Innovative Applications of Artificial Intelligence Conference*, **2**, pp. 62-67.

Carmel D. and Markovitch S. (1997) Model-based learning of interaction strategies in multi-agent systems, *Journal of Experimental and Theoretical Artificial Intelligence*, **10**, 3, pp. 309-332.

Carmel D. and Markovitch S. (1998) How to explore your opponent's strategy (almost) optimally, *Proceedings of the International Conference on Multi Agent Systems*, pp. 64-71.

Cohen L. and Machalek R. (1988) A general theory of expropriative crime: an evolutionary ecological approach, *American Journal of Sociology*, **94**, 3, pp. 465-501.

Cooper R., Jong D., Forsythe R., and Ross T. (1996) Cooperation without reputation: experimental evidence from prisoner's dilemma games, *Games and Economic Behavior*, **12**, 2, pp. 187–218.

Cressman R., Garay J. and Hofbauer J. (2001) Evolutionary stability concepts for N-species frequency-dependent interactions, *Journal of Theoretical Biology*, **211**, pp. 1-10.

Croson R. (2000) Thinking like a game theorist: Factors affecting the frequency of equilibrium play, *Journal of Economic Behavior and Organization*, **41**, 3, pp. 299–314.

Crumbaugh C. and Evans G. (1967) Presentation format, other-person strategies, and cooperative behaviour in the prisoner's dilemma, *Psychological Reports*, **20**, pp. 895-902.

Dacey R. and Pendegraft N. (1988) The optimality of Tit-For-Tat, *International Interactions*, **15**, pp. 45-64.

Darwen P. and Yao X. (1995) On evolving robust strategies for iterated prisoner's dilemma, Progress in Evolutionary Computation, volume 956 in Lecture Notes in Artificial Intelligence, Springer, pp. 276-292.

Darwen P. and Yao X. (1996) Automatic modularization by speciation, *IEEE International Conference on Evolutionary Computation*, pp. 88-93.

Darwen P. and Yao X. (2001) Why more choices cause less cooperation in Iterated Prisoner's Dilemma, *Proceedings of the 2001 IEEE Congress on Evolutionary Computation*.

Darwen P. and Yao X. (2002) Coevolution in iterated prisoner's dilemma with intermediate levels of cooperation: Application to missile defense, *International Journal of Computational Intelligence and Applications*, **2**, 1, pp. 83-107.

Davis D. and Holt C. (1999) Equilibrium cooperation in two-stage games: Experimental evidence, *International Journal of Game Theory*, **28**, 1, pp. 89-109.

Delahaye J. and Mathieu P. (1996) Etude sur les dynamiques du Dilemme Itéré des Prisonniers avec un petit nombre de stratégies : Y a-t-il du chaos dans le Dilemme pur?, Publication Interne IT-294, Laboratoire d'Informatique Fondamentale de Lille.

Doebeli M., Blarer A., and Ackermann M. (1997) Population dynamics, demographic stochasticity, and the evolution of cooperation, Proceedings of National Academy Society of USA, 94: 5167–5171.

Doebeli M. and Knowlton N. (1998) The evolution of interspecific mutualisms, Proceedings of the National Academy of Sciences, 95(15): 8676-8680.

Donninger C. (1986) Is it always efficient to be nice?, In Paradoxical effects of social behavior, edited by Dickmann A. and Mitter P., Heidelberg, Germany: Physica Verlag, pp. 123-134.

Dugatkin L. (1989) N-person games and the evolution of cooperation: a model based on predator inspection in fish, *Journal of Theoretical Biology*, **142**, pp. 123–135.

Dugatkin L. (1990) N-person Games and the Evolution of Co-operation: A Model Based on Predator Inspection in Fish, *Journal of Theoretical Biology*, **142**, pp. 123-135.

Durrett R. and Levin S. (1998) Spatial aspects of interspecific competition, *Theoretical Population Biology*, **53**, 1, pp. 30-43.

Eckert D., Koch S., and Mitlöhner J. (2005) Using the iterated prisoner's dilemma for explaining the evolution of cooperation in open source communities, *Proceedings of the First Conference on Open Source System*, pp. 186-191.

Ficici S., Melnik O., and Pollack J. (2000) A game-theoretic investigation of selection methods used in evolutionary algorithms, *Proceedings of the 2000 Congress on Evolutionary Computation*, **2**, pp. 880-887.

Ficici S. and Pollack J. (2000) Effects of finite populations on evolutionary stable strategies, *Proceedings of the 2000 Genetic and Evolutionary Computation*, pp. 927-934.

Fogel D. (1993) Evolving behaviors in the iterated prisoner's dilemma, *Evolutionary Computation*, **1**, 1, pp. 77-97.

Fogel D. and Fogel G. (1995) Evolutionary stable strategies are not always stable under evolutionary dynamics, *Evolutionary Programming IV*, pp. 565-577.

Fogel D., Fogel G., and Andrew P. (1997) On the instability of evolutionary stable strategies, *BioSystems*, **44**, pp. 135-152.

Fogel G., Andrew P., and Fogel D. (1998) On the instability of evolutionary stable strategies in small populations, *Ecological Modelling*, **109**, pp. 283-294.

Frank R. (1988) *Passions within reason. The strategic role of the emotions*, New York: W.W. Norton & Co.

Freund Y., Kearns M., Mansour Y., Ron D., Rubinfeled R., and Schapire R. (1995) Efficient algorithms for learning to play repeated games against computationally bounded adversaries, *Proceedings of the Annual Symposium on the Foundations of Computer Science*, pp. 332–341.

Fudenberg D. and Maskin E. (1986) The Folk Theorem in repeated games with discounting and incomplete information, *Econometrica*, **54**, pp. 533–554.

Fudenberg D. and Maskin E. (1990) Evolution and cooperation in noisy repeated games, *New Developments in Economic Theory*, **80**, pp. 274-279.

Fudenberg D. and Levine D. (1998) *The theory of learning in games*. MIT Press.

Garay B. and Hofbauer J. (2003) Robust permanence for ecological differential equations: minimax and discretizations, *SIAM Journal on Mathematical Analysis*, **34**, pp. 1007-1093.

Gaunersdorfer A. (1992) Time averages for heteroclinic attractors, *SIAM Journal on Applied Mathematics*, **52**, pp. 1476-1489.

Gaunersdorfer A., Hofbauer J., and Sigmund K. (1991) On the dynamics of asymmetric games, *Theoretical Population Biology*, **39**, pp. 345-357.

Gilboa I. and Matsui A. (1991) Social stability and equilibrium, *Econometrica*, **59**, pp. 859-867.

Gilboa I. and Schmeidler D. (2001) *A theory of case-based decisions*. Cambridge University Press.

Gittins J. (1989) *Multi-armed bandit allocation indices*. Wiley, Chichester, NY.

Glance N. and Huberman B. (1993) The outbreak of cooperation, *Journal of Mathematical sociology*, **17**, 4, pp. 281–302.

Glance N. and Huberman B. (1994) The dynamics of social dilemmas, *Scientific American*, **270**, pp. 76-81.

Glomba M., Filak T., and Kwasnicka H. (2005) Discovering effective strategies for the iterated prisoner's dilemma using genetic algorithms, *5th International Conference on Intelligent Systems Design and Applications*, pp. 356-363.

Godfray H. (1992) The evolution of forgiveness, *Nature*, **355**, pp. 206-207.

Goldstein J. and Freeman J. (1990) *Three-Way Street: Strategic Reciprocity in World Politics*. Chicago: University of Chicago Press.

Grefenstette J., Ramsey C., and Schultz A. (1990) Learning sequential decision rules using simulation models and competition, *Machine Learning*, **5**, pp. 355-381.

Grim P. (1995) The greater generosity of the spatialized prisoner's dilemma, *Journal of Theoretical Biology*, **173**, pp. 242-248.

Grossman W. (2004) New tack wins Prisoner's Dilemma, *Wired News*, Lycos.

Harborne S. (1997) Common belief of rationality in the finitely repeated prisoners' dilemma, *Games and Economic Behavior*, **19**, 1, pp. 133-143.

Hardin G. (1968) The tragedy of the commons, *Science*, **162**, pp. 1243-1248.

Hargreaves H. and Varoufakis Y. (1995) *Game theory: a critical introduction*. Routledge, London.

Harsanyi J. (1967) Games with incomplete information played by Bayesian players, *Management Science*, **14**, 3, pp. 159-182.

Harsanyi, J., and Selten, R. (1988) *A General Theory of Equilibrium Selection in Games*. Cambridge: MIT Press.

Hauser M. (1992) Costs of deception: cheaters are punished in rhesus monkeys (Macaca mulatta). *Proceedings of the National Academy of Sciences*, **89**, pp. 12137-12139.

Heller J. (1967) The effects of racial prejudice, feedback strategy, and race on cooperative-competitive behaviour, *Dissertation Abstracts*, **27**, pp. 2507-2508.

Hemelrijk C. (1991) Interchange of 'Altruistic' Acts as an Epiphenomenon. *Journal of Theoretical Biology*, **153**, pp. 131-139.

Hingston P. and Kendall G. (2004) Learning versus evolution in iterated prisoner's dilemma, *Proceedings of Congress on Evolutionary Computation*, pp. 364-372.

Hirshleifer J. and Coll J. (1988) What strategies can support the evolutionary emergence of cooperation?, *Journal of Conflict Resolution*, **32**, 2, pp. 367-398.

Hirshleifer D. and Rasmusen E. (1989) Cooperation in a repeated prisoner's dilemma with ostracism, *Journal of Economic Behavior and Organization*, **12**, pp. 87-106.

Hofbauer J. (1981) On the occurrence of limit cycles in the Volterra-Lotka equation, *Nonlinear Analysis*, **5**, pp. 1003-1007.

Hofbauer J. (1984) A difference equation model for the hypercycle, SIAM *Journal on Applied Mathematics*, **44**, pp. 762-772.

Hofbauer J. (1996) Evolutionary dynamics for bimatrix games: a Hamiltonian system, *Journal of Mathematical Biology*, **34**, pp. 675-688.

Hofbauer J. and Sigmund K. (1998) *Evolutionary games and population dynamics*. Cambridge University Press.

Hofbauer J. and Sigmund K. (2003) Evolutionary game dynamics, *Bulletin of the American Mathematical Society*, **40**, pp. 479-519.

Hoffmann R. (2001) The ecology of cooperation, *Theory and Decision*, **50**, pp. 101-118.

Holland J. (1975) *Adaptation in Natural and Artificial Systems*. University of Michigan Press, Ann Arbor.

Holland J. (1992) Genetic algorithm, *Scientific American*, **267**, 4, pp. 44-50.

Holland J. (1995) *Hidden Order - How adaptation builds complexity*, Reading, Mass.: Addison-Wesley.

Huberman B. and Glance N. (1993) Evolutionary games and computer simulations, *Proceedings of the National Academy of Sciences*, **90**, pp. 7716-7718.

Ifti M., Killingback T., and Doebeli M. (2004) Effects of neighborhood size and connectivity on the spatial continuous prisoner's dilemma, Journal of *Theoretical Biology*, **231**, pp. 97-106.

Ikegami T. and Kaneko K. (1990) Computer symbiosis - emergence of symbiotic behavior through evolution, *Physica D*, **42**, pp. 235-243.

Insko C., Pinkley R., Hoyle R., Dalton B., Hong G., Slim R., Landry P., Holton B., Ruffin P., and Thibaut J. (1987) Individual-group discontinuity: the role of intergroup contact, *Journal of Experimental Social Psychology*, **23**, pp. 250-267.

Insko C., Hoyle R., Pinkley R., and Hong G. (1988) Individual-group discontinuity: the role of a consensus rule, *Journal of Experimental Social Psychology*, **24**, pp. 505-519.

Insko C., Schopler J., Hoyle R., Dardis G., and Graetz K. (1990) Individual-group discontinuity as a function of fear and greed, *Journal of Personality and Social Psychology*, **58**, pp. 68-79.

Insko C., Schopler J., Drigotas S., Graetz K., Kennedy J., Cox C., and Bornstein G. (1993) The role of communication in interindividual-intergroup discontinuity, *Journal of Conflict Resolution*, **37**, pp. 108-138.

Kaelbling L. (1993) *Learning in embedded systems*. The MIT Press, Cambridge, MA.

Kaelbling L. and Moore A. (1996) Reinforcement learning: a survey, *Journal of Artificial Intelligence Research*, **4**, pp. 237-285.

Kagel J. and Roth A. (1995) *The Handbook of Experimental Economics*. Princeton University Press.

Kahn L. and Murnighan J. (1993) Conjecture, uncertainty, and cooperation in Prisoners' Dilemma games: Some Experimental Evidence, *Journal of Economic Behavior and Organisms*, **22**, pp. 91–117.

Kalai E. and Lehrer E. (1993) Rational learning leads to Nash equilibrium *Econometrica*, **61**, 5, pp. 1019-1045.

Kandori M. (1992) Social norms and community enforcement, *The Review of Economic Studies*, **59**, 1, pp. 63-80.

Katok A. and Hasselblatt B. (1996) *Introduction to the modern theory of dynamical systems*. Cambridge ISBN 0521575575.

Kavka G. (1986) *Hobbesean Moral and Political Theory*. Princeton: Princeton University Press.

Kirchkamp O. (1995) Spatial Evolution of Automata in the Prisoners' Dilemma. University of Bonn SFB 303, Discussion Paper B-330.

Kitcher P. (1992) Evolution of altruism in repeated optional games, Working Paper of University of California at San Diego.

Knapp W. and Podell J. (1968) Mental patients, prisoners, and students with simulated partners in a mixed-motive game, *Journal of Conflict Resolution*, **12**, pp. 235-241.

Komorita S., Sheposh J., and Braver S. (1968) Power, the use of power, and cooperative choice in a two-person game, *Journal of Personality and Social Psychology*, **8**, pp. 134-142.

Kraines D. and Kraines V. (1995) Evolution of learning among Pavlov strategies in a competitive environment with noise, *The Journal of Conflict Resolution*, **39**, 3, pp. 439-466.

Kraines D. and Kraines V. (2000) Natural selection of memory-one strategies for the iterated Prisoner's Dilemma, *Journal of Theoretical Biology*, **203**, pp. 335-355.

Kreps D., Milgrom P., Roberts J., and Wilson R. (1982) Rational cooperation in the finitely repeated prisoner's dilemma, *Journal of Economic Theory*, **27**, pp. 245–252.

Kreps, D., and Wilson R. (1982) Reputation and imperfect information, *Journal of Economic Theory*, **27**, pp. 253–279.

Krishna V. and Sjöström T. (1998) On the convergence of fictitious play, *Mathematics Operations Research*, **23**, pp. 479-511.

Lave L. (1965) Factors affecting cooperation in the prisoner's dilemma, *Behavioral Science*, **10**, pp. 26-38.

Lindgren K. (1991) Evolutionary phenomena in simple dynamics, In Christopher G., et al. Santa Fe Institute Studies in the Sciences of Complexity. **10**, pp. 295-312.

Lindgren K. (1992) Evolutionary phenomena in simple dynamics, In Langton C. (ed.) *Artificial Life II*. Addison-Wesley.

Lindgren K. (1995) Evolutionary dynamics in game-theoretic models, The economy as an evolving complex system II, Santa Fe Institute.

Littman M. (1994) Markov games as a framework for multiagent reinforcement learning, *Proceedings of the 11th International Conference on Machine Learning*, pp. 157-163.

Luce R. and Raiffa H. (1957) *Games and decisions*. New York: Wiley.

Lynch G. (1968) Defense preference and cooperation and competition in a game, *Dissertation Abstracts*, **29**, pp. 1174.

Manarini S. (1998) The prisoner's dilemma, experiments for the study of cooperation. Strategies, theories and mathematical models, Ph.D. thesis, University of Padova.

Marinoff L. (1992) Maximizing expected utilities in the Prisoner's Dilemma, *Journal of Conflict Resolution*, **36**, 1, pp. 183-216.

Martinez-Coll J. and Hirshleifer J. (1991) The limits of reciprocity, *Rationality and Society*, **3**, pp. 35-64.

Matsui A. (1992) Best response dynamics and socially stable strategies, *Journal of Economic Theory*, **57**, pp. 343-362.

May R. (1987) More evolution of cooperation, *Nature*, **327**, pp. 15-17.

Maynard Smith J. and Price G. (1973) The logic of animal conflict, *Nature*, **246**, pp. 15-18.

Maynard Smith J. (1982) *Evolution and the Theory of Games*, Cambridge University Press.

McKelvey R. and Palfrey T. (1992) An experimental study of the centipede game, *Econometrica*, **60**, pp. 803-836.

Mealey L. (1995) The sociobiology of sociopathy: an integrated evolutionary model, *Behavioral and Brain Sciences*, **18**, 3, pp. 523-599.

Michalewicz Z. (1999) *Genetic Algorithms + Data Structures = Evolution Programs*, Springer-Verlag.

Micko H. (1997) Benevolent tit for tat strategies with fixed intervals between offers of cooperation, *Meeting of Experimental Psychologists*, pp. 250-256.

Micko H. (2000) Experimental Matrix games, In Open and Distance Learning-Mathematical Psychology, Institut für Sozial- und Persönlichkeitspsychologie, Universität Bonn.

Milgrom, P. and Roberts J. (1982): Predation, reputation and entry deterrence, *Journal of Economic Theory*, **27**, pp. 280-312.

Milinski M. (1993) Cooperation wins and stays, *Nature*, **364**, pp. 12-13.

Milinski M. and Wedekind C. (1998) Working memory constrains human cooperation in the prisoner’s dilemma, *Proceedings of the National Academy of Sciences of the United States of America*, **95**, 23, pp. 13755-13758.

Miller J. (1996) The coevolution of automata in the repeated prisoner’s dilemma, *Journal of Economic Behavior and Organization*, **29**, pp. 87-112.

Mitchell M. (1996) *An introduction to Genetic Algorithms*. The MIT Press, Cambridge MA.

Molander P. (1985) The optimal level of generosity in a selfish, uncertain environment, *Journal of Conflict Resolution*, **29**, pp. 611-618.

Moore A. and Atkeson C. (1993) Prioritized sweeping: reinforcement learning with less data and less real time, *Machine Learning*, **13**, pp. 103-130.

Mueller U. (1988) Optimal retaliation for optimal cooperation, *Journal of Conflict Resolution*, **31**, 4, pp. 692-724.

Myerson R. (1991) *Game Theory, Analysis of Conflict*. Cambridge, Harvard University Press.

Nachbar J. (1992) Evolution in the finitely repeated Prisoners' Dilemma, *Journal of Economic Behavior and Organization*, **19**, pp. 307-326.

Narendra K. and Thathachar M. (1989) *Learning automata: an introduction*. Prentice-Hall, Englewood Cliffs, NJ.

Nash J. (1950) Equilibrium points in n-person games, *Proceedings of the National Academy of the USA*, **36**, 1, pp. 48-49.

Nash J. (1951) Non-cooperative games, *The Annals of Mathematics*, **54**, 2, pp. 286-295.

Nash J. (1996) *Essays on Game Theory*. Elgar. Cheltenham.

Noldeke G. and Samuelson L. (1993) An evolutionary analysis of backward and forward induction, *Games and Economic Behaviour*, **5**, pp. 425-454.

Nowak M., Bonhoeffer S., and May R. (1994) More spatial games, *International Journal of Bifurcation and Chaos*, **4**, 1, pp. 33-56.

Nowak M. and May R. (1992) Evolutionary games and spatial chaos, *Nature*, **359**, pp. 826-829.

Nowak M. and May R. (1993) The spatial dilemmas of evolution, International *Journal of Bifurcation and Chaos*, **3**, pp. 35-78.

Nowak M. and Sigmund K. (1990) The evolution of stochastic strategies in the prisoner's dilemma, *Acta Applicandae Mathematicae*, **20**, pp. 247-265.

Nowak M. and Sigmund K. (1992) Tit for tat in heterogeneous populations, *Nature*, **359**, pp. 250-253.

Nowak M. and Sigmund K. (1993) A strategy of win-stay lose-shift that outperforms Tit-for-Tat in the Prisoner's Dilemma game, *Nature*, **364**, pp. 56-58.

Nowak M., Sigmund K. and El-Sedy E. (1995) Automata, repeated games, and noise, *Journal of Mathematical Biology*, **33**, pp. 703-722.

Orbell J., Kragt A., and Dawes R. (1988) Explaining discussion-induced cooperation, *Journal of Personality and Social Psychology*, **54**, pp. 811-819.

Orbell J. and Dawes R. (1993) Social welfare, cooperator's advantage, and the option of not playing the game, *American Sociological Review*, pp. 787-800.

Orbell J. and Robyn M. (1993) Social welfare, cooperators' advantage, and the option of not playing the game. *American Sociological Review*, **58**, pp. :787-800.

Oskamp S. (1971) Effects of programmed strategies on cooperation in the prisoner's dilemma and other mixed-motive games, *The Journal of Conflict Resolution*, **15**, 2, pp. 225-259.

Oskamp S. and Perlman D. (1965) Factors affecting cooperation in a prisoner's dilemma game, *Journal of Conflict Resolution*, **9**, pp. 359-374.

Plank M. (1997) Some qualitative differences between the replicator dynamics of two player and n player games, *Nonlinear Analysis*, **30**, pp. 1411-1417.

Pollock G. (1989) Evolutionary Stability of Reciprocity in a Viscous Lattice. *Social Networks*, **11**, pp. 175-212.

Posch M. (1997) Win Stay–Lose Shift: An Elementary Learning Rule for Normal Form Games, Working Paper of Santa Fe Institute, http://ideas.repec.org/p/wop/ safire/97-06-056e.html.

Prisoner's dilemma tournament result (2004) http://www.prisoners-dilemma.com/results/cec04/ipd_cec04_full_run.html .

Prisoner's dilemma tournament result (2005) http://www.prisoners-dilemma.com/results/cig05/cig05.html .

Radner R. (1980) Collusive behaviour in non-cooperative epsilon-equilibria in oligopolies with long but finite lives, *Journal of Economic Theory*, **22**, pp. 136-154.

Radner R. (1986) Can bounded rationality resolve the prisoner's dilemma, In Mas- Colell A. and Hildenbrand W. Essays in Honor of Gerard Debreu, pp. 387-399.

Rapoport A. (1966) Optimal policies for the prisoner's dilemma, Technical Report No.50 Psychometric Laboratory, University of North California, MH-10006.

Rapoport A. (1999) *Two-person Game Theory*. Dover Publications, New York.

Rapoport and Chammah (1965) *Prisoner's dilemma: a study in conflict and cooperation*. Ann Arbor: University of Michigan Press.

Rothbard M. (1997) Toward a Reconstruction of Utility and Welfare Economics, *In The Logic of Action One: Method, Money, and the Austrian School*, pp. 211-55.

Rubinstein A. (1979) Equilibrium in super games with the overtaking criterion, *Journal of Economic Theory*, **21**, pp. 1-9.

Rubinstein A. (1998) *Modeling bounded rationality*. The MIT Press, 1998.

Samuelson L. (2001) Introduction to the evolution of preferences, *Journal of Economic Theory*, **97**, pp. 225-230.

Sandholm T. and Crites R. (1996) Multiagent reinforcement learning in the iterated Prisoner's Dilemma, *Biosystems*, **37**, 1-2, pp. 147-66.

Sarin R. (1999) Simple play in the prisoner's dilemma, *Journal of Economic Behavior and Organization*, **40**, 1, pp. 105–113.

Sasaki A., Taylor C. and Fudenberg D. (2000) Emergence of cooperation and evolutionary stability in finite populations, *Nature*, **428**, pp. 646-650.

Schmidhuber J. (1996) A general method for multi-agent learning and incremental self-improvement in unrestricted environments, In Yao X. (ed.) Evolutionary Computation: Theory and Applications. Scientific Publications Co.

Schmitt L. (2001) Theory of genetic algorithms, *Theoretical Computer Science*, **259**, pp. 1-61.

Schuessler R. (1989) Exit threats and cooperation under anonymity, *Journal of Conflict Resolution*, **33**, pp. 728-749.

Schweitzer F. (2002) *Modeling Complexity in Economic and Social Systems*. World Scientific, Singapore.

Schweitzer F., Behera L., and Mühlenbein H. (2002) Evolution of cooperation in a spatical prisoner's dilemma, *Advances in Complex Systems*, **5**, 2-3, pp. 269-299.

Scodel A., Minas J., Ratoosh P., and Lipetz M. (1959) Some descriptive aspects of two-person non-zero sum games, *Journal of Conflict Resolution*, **3**, pp. 114-119.

Selten, R. (1965) Spieltheoretische behandlung eines oligopolmodells mit nachfragetragheit, Zeitschrift fur die Gesamte Staatswissenschaft, **12**, pp. 301-324.

Selten, R. (1975) Reexamination of the perfectness concept for equilibrium points in extensive games, *International Journal of Game Theory*, **4**, pp. 25-55.

Selten R. (1983) Evolutionary stability in extensive two-person games, *Mathematical Social Science*, **5**, pp. 269-363.

Selten R. (1988) Evolutionary stability in extensive two-person games: correction and further development, *Mathematical Social Science*, **16**, pp. 223-266.

Selten R. and Stoecker R. (1986) End behaviour in sequences of finite Prisoner's Dilemma supergames: a learning theory approach, *Journal of Economic Behaviour and Organisation*, **7**, pp. 47-70.

Sethi R. and Somanathan E. (1996) The evolution of social norms in common property resource use, *The American Economic Review*, **86**, 4, pp. 766-788.

Simon H. (1955) A behavioral model of rational choice, *Quarterly Journal of Econometrics*, **69**, 1, pp. 99-118.

Simon H. (1990) A mechanism for social selection and successful altruism, *Science*, **250**, 4988, pp. 1665-1668.

Sermat V. (1967) Cooperative behaviour in a mixed-motive game, Journal of *Social Psychology*, **62**, pp. 217-239.

Sigmund K. (1995) *Games of Life: Explorations in Ecology, Evolution and Behaviour*. Penguin, Harmondsworth.

Skyrms B. (1990) The Dynamics of Rational Deliberation. Harvard UP.

Smith H. (1995) Monotone dynamical systems: an introduction to the theory of competitive and cooperative systems, *AMS Mathematical Surveys and Monographs*, **41**.

Smith R. and Gray B. (1994) Co-adaptive genetic algorithms: an example in Othello strategy, *Proceedings of the 1994 Florida Artificial Intelligence Research Symposium*, pp. 259-264.

Sobel J. (1975) Reexamination of the perfectness concept of equilibrium in extensive games, *International Journal of Game Theory*, **4**, pp. 25-55.

Sobel J. (1976) Utility maximization in iterated Prisoner's Dilemmas, *Dialogue*, **15**, pp. 38-53.

Solomon L. (1960) The influence of some types of power relationships and game strategies upon the development of interpersonal trust, *Journal of Abnormal and Social Psychology*, **61**, pp. 223-230.

Spears W. (1992) Crossover or mutation? *Foundations of Genetic Algorithms. 2, FOGA-92*, edited by Whitley D., California: Morgan Kaufmann.

Spiro D. (1988) The state of cooperation in theories of state cooperation: the evolution of a category mistake, *Journal of International Affairs*, **42**, pp. 205-225.

Stanley E., Ashlock D., and Smucker M. (1995) Iterated prisoner's dilemma with choice and refusal of partners: Evolutionary results, *Lecture Notes in Artificial Intelligence*, **929**, pp. 490-502.

Stanley E., Ashlock D., and Tesfatsion L. (1994) Iterated prisoner's dilemma with choice and refusal of partners, In Christopher G. *Artificial Life III*. Addison-Wesley, pp. 131-176.

Stephens D. (2000) Cumulative benefit games: achieving cooperation when players discount the future, *Journal of Theoretical Biology*, **205**, 1, pp. 1-16.

Stephens D., Mclinn C., and Stevens J. (2002) Discounting and Reciprocity in an Iterated Prisoner's Dilemma, *Science*, **298**, 5601, pp. 2216-2218.

Sugden, R. (1986) *The Economics of Cooperation*, Rights and Welfare. Basil Blackwell.

Surowiecki J. (2004) *The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business*, Economies, Societies and Nations. Little, Brown.

Sutton R. (1990) Integrated architectures for learning, planning, and reacting based on approximating dynamic programming, *Proceedings of the 7th International Conference on Machine Learning*, pp. 216-224.

Swingle P. and Coady H. (1967) Effects of the partner's abrupt strategy change upon subject's responding in the prisoner's dilemma, *Journal of Personality and Social Psychology*, **5**, pp. 357-363.

Swinkels J. (1993) Adjustment dynamics and rational play in games, *Games and Economic Behavior*, **5**, pp. 455-84.

Taylor, P. D. (1979). Evolutionarily stable strategies with two types of players, *Journal of Applied Probability*, **16**, pp. 76-83.

Taylor, P. and Jonker, L. (1978) Evolutionary stable strategies and game dynamics, *Mathematical Biosciences*, **40**, pp. 145-156.

Tekol Y. and Acan A. (2003) Ants can play Prisoner's Dilemma, *Proceedings of the 2003 Congress on Evolutionary Computation*, pp. 1151-1157.

Thierens D. (1997) Selection schemes, elitist recombination, and selection intensity, *Proceedings of the 7th International Conference on Genetic Algorithms*, pp. 152-159.

Thierens D. and Goldberg D. (1994) Elitist recombination: an integrated selection recombination GA, *Proceedings of the First IEEE Conference on Evolutionary Computation*, pp. 508-512.

Thomas B. (1985) On evolutionarily stable sets, *Journal of Mathematical Biology*, **22**, pp. 105-115.

Tzafestas E. (2000a) Toward adaptive cooperative behavior, *Proceedings of the Simulation of Adaptive Behavior Conference*, pp. 334-340.

Tzafestas E. (2000b) Spatial games with adaptive tit-for-tats, *Proceedings of the 6th Parallel Problem Solving from Nature (PPSN-VI)*, pp. 507-516.

Young H. and Foster D. (1991) Cooperation in the Short and in the Long Run, *Games and Economic Behavior*, **3**, pp. 145-156.

Vegaredondo F. (1994) Bayesian boundedly rational agents play the finitely repeated prisoner's dilemma, *Theory and Decision*, **36**, 2, pp. 187–206.

Von Neumann J. and Morgenstern O. (1944) *Theory of Games and Economic Behavior*. Princeton UP.

Watkins C. (1989) Learning from delayed rewards. Ph.D. thesis, King's College, Cambridge, UK.

Watkins C. and Dayan P. (1992) Q-learning, *Machine Learning*, **8**, 3, pp. 279-292.

Wedekind C**.** and Milinski M. (1996) Human cooperation in the simultaneous and the alternating Prisoner's Dilemma: Pavlov versus Generous Tit-for-Tat, *Proceedings of the National Academy of Sciences of the United States of America*, **93**, 7, pp. 2686-2689.

Weibull J. (1995) *Evolutionary Game Theory*. MIT Press, Cambridge, Mass.

Wilson W. (1969) Cooperation and the cooperativeness of the other player, *Journal of Conflict Resolution*, **13**, pp. 110-117.

Wilson W. (1987) Classifier systems and the animat problem, *Machine Learning*, **2**, pp. 199-228.

Whitworth R. and Lucker W. (1969) Effective manipulation of cooperation with college and culturally disadvantaged populations, *Proceedings of 77th Annual Convention of American Psychological Association*, **4**, pp. 305-306.

Wu J. and Axelrod R. (1995) How to cope with noise in the Iterated Prisoner's Dilemma, *Journal of Conflict Resolution*, **39**, pp. 183-189.

Zeeman M. (1993) Hopf bifurcations in competitive three dimensional Lotka-Volterra systems, *Dynamics and Stability of Systems*, **8**, pp. 189-217

Zeeman E., Zeeman M. (2002) An n-dimensional competitive Lotka-Volterra system is generically determined by its edges, *Nonlinearity*, **15**, pp. 2019-2032.

Zeeman E., Zeeman M. (2003) From local to global behavior in competitive Lotka-Volterra systems, *Transaction of American Mathematical Society*, **355**, pp. 713-734.