Establishment of Dominance Hierarchies and Cooperation: A Game-Theoretic Perspective

Karan Jain
Brasenose College
University of Oxford

A thesis submitted for the degree of
MSc in Mathematical Modelling and Scientific Computing
September 2016

Acknowledgements

I would like to thank my supervisors - Mason Porter and Cameron Hall for the invaluable help and guidance they provided during this thesis. I would also like to thank Cameron Hall for providing me with his code, and helping me make my own modifications for the purposes of this thesis. I would like to thank Kathryn Gillow for her help over the year.

Abstract

Social interactions among animals can take many forms. While some animals are known to cooperate with each other, others are characterised by dominance hierarchies. This thesis is concerned with the study of these two forms of social organisation amongst animal groups. We model the problem using a game-theoretic framework.

We investigate the evolution of animal behaviour under the influence of natural selection and mutation. We study pairwise conflicts between animals from two distinct populations. We model natural selection using a discrete map which increases the proportion of animals that perform better relative to other animals in the population, and model mutations by adding a stochastic element to the discrete map.

We begin our analysis with the investigation of dominance hierarchies that can develop in the Hawk–Dove game. We study the different factors that can affect the formation of dominance hierarchies.

In order to gain insight into the evolution of cooperation, we generalise our model to incorporate animals that interact with each other repeatedly in one generation. We use this model to investigate the conditions under which cooperation can survive in the populations as a stable long run solution.

Contents

Chapter 1

Introduction

In nature, animals regularly come into conflict with other members of their own species, or with members of some other species. Conflict in the animal world can be over resources, or over outcomes. The former kind arises when two or more individuals compete for something that they both need but which is limited supply. For example, ants from two neighbouring nests fight for foraging space. Conflict over outcome occurs, for example, in predator–prey systems. There are several ways in which these conflicts may be resolved, which include violent fights, peaceful exchanges and division of labour. Different social interactions lead to different forms of social organisation. In this thesis, we study two contrasting, but commonly occurring forms of social organisations: cooperation and dominance hierarchies.

Dominance hierarchies are normally established between animals/groups/species which have a conflict of interest due to limited availability of a resource. They occur in many group of animals – including birds, fish, mammals, and insects [1, 2, 3] – and they can be of various kinds [4]. Hierarchies are often established because it is too costly for animals to fight with each other for prolonged periods [4, 28]. Once such hierarchies are established, subordinates give way to dominant members and allow them to take resources without a fight. For example, in baboons, older, dominant males have access to all the females in the group [4]. While subordinate, younger males have no access to females, they still benefit from being a part of the hierarchy as it provides an efficient means of defence against predation. This raises many interesting questions: How are dominance hierarchies established? What are the factors that govern their establishment? What are the conditions under which subordinates can destabilise the hierarchy? In this thesis, we conduct a mathematical investigation of these problems.

Asymmetries (in age, sex, size, physiology and levels of aggression) between animals has been found to be a key factor governing the establishment of dominance hierarchies [5]. Most studies examining dominance hierarchies, however, have focused on intra-species interactions, in which the level of asymmetries between animals is small. The issue of inter-species interactions (or other interactions involving participants with high levels of asymmetry of some form) is rarely addressed, even though such occurrences are known to be widespread in the animal world [6, 7, 8]. In this thesis, we will investigate the social organisation resulting from interactions between two distinct populations of animals, such that significant asymmetries exist between animals in the two populations.

Asymmetries are often reflected in the fighting ability of animals. For example, size or mass is often correlated to the fighting ability of the animal [5]. Such a scenario can then be modelled using the 'Hawk–Dove' framework developed by Maynard Smith [11], which we discuss further in Chapter 3. In particular, we are interested in the case in which animals in both populations are worse off fighting as compared to being subordinate in a dominance hierarchy, but one is more so than the other (due to the asymmetry).

Conflict among animals, be it over a resource of limited availability, or over outcomes (in male-female conflicts, for example), do not always lead to dominance hierarchies. Conflicts are often resolved by means of 'cooperation'. An animal's behaviour is defined as cooperative if it provides a benefit to another animal (recipient), but is costly for the donor (at least in the short-term) [9]. For example, vampire bats are known to regurgitate blood that they have obtained to feed a starving member of their colony, even when it is not related to the donor [62]. In meerkats, vigilance is undertaken by particular individuals which take turns to go to a high look out point such as a tree and keep watch for predators while the others feed [30].

Darwin recognized the problem that cooperation poses for his theory of evolution by natural selection, which favours individuals who have the greatest personal reproductive success. Thus, the reasons why an individual animal in a population would be motivated to cooperate are not clear. Trivers [34] showed that conditional behaviour can promote cooperation in a world in which individuals can recognize and remember others. He coined the term 'reciprocal altruism' for such behaviour. Reciprocal altruism, in other words, implies that giving up or sharing a resource on one occasion is beneficial in the long run because of the many future occasions in which a cooperator receives resources in return, and thus leads to cooperation amongst animals. Since then, the evolution of reciprocal altruism has received a great deal of theoretical attention [13, 14, 15]. This idea was further developed in the ingenious computational tournaments of Axelrod and Hamilton [35], who presented programmers with a challenge: in a game of repeated interactions in which cooperation leads to mutual gain, but exploitation of other cooperators leads to an even greater gain (known as the 'Prisoner's dilemma' game, which we shall discuss in further detail in Chapter 2), design a winning behavioural strategy. Despite the submission of some complex programs, the winning strategy was very simplistic: "Tit-for-Tat" (TFT). TFT strategy cooperates in the first round, and it subsequently mirrors the play of its opponent from the previous round. TFT strategy shows reciprocal altruism, as when TFT encounters a cooperator they enjoy a cooperative interaction, but TFT will not allow its cooperation to be exploited by a defecting partner.

Since Axelrod's tournament, several authors have argued that cooperation is likely to evolve whenever pairs of individuals interact repeatedly over a long time period [65, 16, 17]. Axelrod showed that in a population of animals which interact with each other repeatedly and using the TFT strategy, no other behavioural strategy (or 'mutant' strategy) could obtain a better payoff, provided the proportion of animals that adopt the mutant strategy is small [65]. One possible criticism of this analysis is that it considers the robustness of TFT only against isolated mutations. It does not consider the robustness of TFT against several different types of mutations that might invade the population at the same time. In this thesis, we address this discrepancy by investigating the robustness of TFT (and other strategies that lead to cooperation) against continuous random mutations in the population (discussed in further detail in Chapter 3).

In this thesis, we investigate the establishment and robustness of dominance hierarchies and cooperation between two distinct population of animals in conflict with each other. We do so in an evolutionary environment in which the strategies/behaviours that do better than others reproduce faster, and hence have a greater representation in the next generation of the populations, which is a consequence of Darwin's maxim of 'survival of the fittest'. Because the evolution of animal behaviour depends critically on how others in a population are behaving, game theory provides a convenient tool for modelling it. Game theory (or classical game theory) is a formal methodology to study the interaction of self interested decision-makers [44]. It was first formalised by Von Neumann and Morgenstern [42] to model human economic behaviour. In particular, we use the framework of evolutionary game theory, which is an adaptation of game theory for applications in evolution. It was developed mainly by Maynard Smith and Price [11, 18]. Since then, evolutionary game theory has found applications in biology, including the study of sexual allocation and parental investment [22], inter-species competition for resources [19], animal disperal [20], plant growth and reproduction [21] and microbial communities [23]. Its applications beyond biology, especially in the study of human social and economic behaviour [24, 52], are becoming increasingly popular.

The rest of the thesis is organised as follows: In Chapter 2, we give a brief introduction of the classical and evolutionary game theoretic concepts that we will be using in our work. In Chapter 3, we present a model to formulate the problem of asymmetric contests between animal populations using the language of game theory. We use Chapter 4 to present our results, as well as discuss some of the deficiencies of our model. After introducing some additional theory, we attempt to address these deficiencies by generalizing our model in Chapter 6. We then present our results of this extended model in Chapter 7, before concluding the thesis with a discussion of the results and a conclusion.

Chapter 2

Game Theory

2.1 Classical Game Theory (CGT)

In this section, we review some of the basic definitions and concepts of classical game theory (CGT). The definitions in this section closely follow the ones given by Shoham and Brown [44].

One of the most famous examples in game theory is the *Prisoner's Dilemma* (PD). This game can be motivated by considering two suspects ('players') A and B who have been taken into custody. Suppose that a district attorney is sure that they have committed a crime together but does not have enough evidence. They are interrogated in separate rooms and cannot communicate with each other. Both of the suspects are simultaneously given the opportunity to either betray the other by testifying against him/her or to cooperate with him/her by remaining silent. The offer is: (a) if A and B betray each other (or 'defect'), each of them serve two years in prison; (b) if A betrays B but B remains silent (or 'cooperates'), A will be set free and B will serve three years in prison; (c) if B betrays A but A remains silent, B will be set free and A will serve three years in prison; and (d) if both A and B remain silent, both of them serve one year in prison each (on a lesser charge). The PD game can be represented succinctly using the bi-matrix representation in Fig. 2.1. We call such a representation the *normal form* of a game.

Note that the bi-matrix in Fig. 2.1 specifies the following essential features of the game: (a) the set of players, (b) the possible actions available to each player, and (c) the rule determining the outcome of every possible game ending. Formally, we can now define a normal-form representation of a game (similar to the definition given by Shoham and Brown [44]).

		Player B	
		Defect	Cooperate
Player A	Defect	$-2, -2$	$0, -3$
	Cooperate	$-3, 0$	$-1, -1$

Figure 2.1: Normal form representation of the PD game.

Definition 2.1.1. The *normal-form* representation of an N-player game is defined by a tuple $G = (I, \mathcal{A}, \mathbf{u})$, where

1. $I = \{1, 2, \ldots, N\}$ is a finite set of players;
2. $\boldsymbol{\mathcal{A}} = \mathcal{A}_1 \times \cdots \times \mathcal{A}_N$, where $\mathcal{A}_i$ is a finite set of *actions* available to player i. We will refer to $\boldsymbol{\mathcal{A}}$ as the action space. Each vector $\mathbf{a} = (a_1, \cdots, a_N) \in \boldsymbol{\mathcal{A}}$ is called an *action profile*;
3. $\mathbf{u} = (u_1, \cdots, u_N)$, where $u_i : \boldsymbol{\mathcal{A}} \to \mathbb{R}$ is a real-valued payoff function for player i.

The normal-form representation for a particular game is not unique. In particular, the representation is equivalent up to a linear transformation of the vector of payoff functions.

When employing a *pure action*, a player deterministically picks a single action from his/her individual action space. However, there is no fundamental reason why players need to play only pure actions. It is reasonable to assume that the players can select randomly from their set of pure actions, using some probabilistic rule. This leads to the idea of a *mixed action*.

Definition 2.1.2. A *mixed action* for player i is a function $\sigma_i : \mathcal{A}_i \to [0, 1]$. It assigns a probability $\sigma_i(a_i) \geq 0$ to each pure action $a_i \in \mathcal{A}_i$. In addition, it is required to satisfy $\sum_{a_i \in \mathcal{A}_i} \sigma_i(a_i) = 1$.

Note that the set of all pure actions is a subset of the set of all mixed actions. We denote the set of all mixed action profiles $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_N)$ by $\boldsymbol{\Sigma} = \Sigma_1 \times \cdots \times \Sigma_N$. For an action profile $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_N)$, we let $\boldsymbol{\sigma}_{-i}$ denote the $(N-1)$-component vector of actions of all players excluding i, and we therefore write $\boldsymbol{\sigma} = (\sigma_i, \boldsymbol{\sigma}_{-i})$. We use a similar notation for pure action profiles. We generalise the definition of payoff functions to define payoffs over a profile of mixed actions as follows:

$$u_i(\boldsymbol{\sigma}) = u_i(\sigma_1, \ldots, \sigma_N) = \sum_{\mathbf{a} \in \boldsymbol{\mathcal{A}}} [\sigma_1(a_1)\sigma_2(a_2) \cdots \sigma_N(a_N)] u_i(\mathbf{s}). \tag{2.1}$$

Note that in equation (2.1), we assume that the players are randomizing independently.

An outcome of the PD game can be obtained by appealing to the *rationality* of the players. We say a player is *rational* if he/she seeks to play in a manner that maximises his/her own payoff [43]. Consider the PD game from the perspective of player A's payoff maximisation. If player B defects, A is better off defecting. More formally, we say that defecting is A's *best response* to player B's action of defecting.

Definition 2.1.3. An action $\sigma_i \in \Sigma_i$ is a *best response* to the action profile $\boldsymbol{\sigma}_{-i} \in \boldsymbol{\Sigma}_{-i}$ if $u_i(\sigma_i, \boldsymbol{\sigma}_{-i}) \geq u_i(\tilde{\sigma}_i, \boldsymbol{\sigma}_{-i})$ for all $\tilde{\sigma}_i \in \Sigma_i$.

If player B cooperates, player A is again better off defecting. In other words, action D gives a strictly greater payoff than action C for player A (regardless of the opponent's action choice).

Definition 2.1.4. An action $\sigma_i \in \Sigma_i$ is *strictly dominated* for player i if there exists a mixed action $\tilde{\sigma}_i \in \Sigma_i \backslash \{\sigma_i\}$ such that for all $\mathbf{a}_{-i} \in \boldsymbol{\mathcal{A}}_{-i}$, we have $u_i(\tilde{\sigma}_i, \mathbf{a}_{-i}) > u_i(\sigma_i, \mathbf{a}_{-i})$. In this case, we say that $\tilde{\sigma}_i$ *strictly dominates* σ_i.

If a strictly dominated action exists, it is reasonable to assume that a player will not play it as a consequence of the player's rationality.

We now consider the PD game from player B's perspective. Note that the PD game is symmetric; that is, we can permute the players while keeping the payoff functions intact. Therefore, by the logic used for player A previously, player B will always defect. We thus expect both players to defect. We have obtained an outcome of the PD game, namely (D, D). Note that neither player can unilaterally deviate from this action profile and achieve a greater payoff. Generalisation of this notion leads to the idea of a *Nash equilibrium.*

Definition 2.1.5. An action profile $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_N) \in \boldsymbol{\Sigma}$ is a *Nash equilibrium* (NE) if for all i and $\tilde{\sigma}_i \in \Sigma_i$, we have $u_i(\sigma_i, \boldsymbol{\sigma}_{-i}) \geq u_i(\tilde{\sigma}_i, \boldsymbol{\sigma}_{-i})$.

The action profile (D, D) is a Nash equilibrium of the PD game. In a Nash equilibrium, each player's action is a best response to those actions of his/her opponents that are components of the equilibrium. The following existence theorem for NE's was proved by Nash [45]: *Every game with a finite number of players and action space has at least one Nash equilibrium.*

The following proposition (by Shoham and Brown [44]) is useful for finding mixed action Nash equilibria.

Proposition 2.1.1. *For an action profile $\boldsymbol{\sigma}^*$, define $\mathcal{A}_i^* := \{a_i \in \mathcal{A}_i \mid \sigma_i^*(s_i) > 0\}$ as the set of pure actions that player i plays with positive probability according to $\boldsymbol{\sigma}^*$. Then, $\boldsymbol{\sigma}^*$ is a Nash equilibrium if and if only for all $i \in I$*

1. $u_i(a_i, \boldsymbol{\sigma}^*_{-1}) = u_i(a'_i, \boldsymbol{\sigma}^*_{-i})$ *for all* $a_i, a'_i \in \mathcal{A}_i^*$*;*
2. $u_i(a_i, \boldsymbol{\sigma}^*_{-i}) \geq u_i(a'_i, \boldsymbol{\sigma}^*_{-i})$ *for all* $a_i \in \mathcal{A}_i^*$ *and* $a'_i \in \mathcal{A}_i$*.*

For the proof of Proposition 2.1.1, the reader should consult Appendix A. The necessary condition in Proposition 2.1.1 implies that in any Nash equilibria, a player must be indifferent over the pure actions he/she is randomizing over. This places a restriction on the mixed actions of his/her opponents, which can be used to calculate the the mixed action profile of the Nash equilibrium (if it exists).

2.2 Evolutionary Game Theory (EGT)

Historically, economics was the original area of application for game theory. However, Maynard Smith [11] showed that it provides a very natural framework to model evolution and animal behaviour. Classical game theory (CGT) requires the modelling of an agent's self-interested behaviour. This measure is provided by *utility theory* [44], which values a wide variety of different outcomes (such as financial rewards and the risk of death) on a single scale. Maynard Smith replaced the concept of utility with the more natural concept of *Darwinian fitness*.

Definition 2.2.1. [11] *Darwinian fitness* of an animal in a population is defined as the payoff (measured in number of offspring) following a contest with another randomly chosen member of the population.

Note that the fitness of a action being played by an animal depends on the frequency of other types in the population. Hereafter, we will refer to the agents in the EGT framework as 'animals', to differentiate them from the agents in CGT (which are referred to as players). In Maynard Smith's original treatment of evolutionary game theory [11], there were two critical shifts from CGT. We discuss these two shits below, which are retained in our model:

1. **Action.** In CGT, players have action sets (both pure and mixed) from which they choose. In EGT, the action sets of a population of animals consist of genotypic variants from which animals inherit exactly one variant. Each animal is genetically programmed to play its unique pure action throughout its lifetime.

We can therefore define the *type* of an animal in a population using its unique pure action. Animals cannot play mixed actions, so they cannot randomize over different pure actions.

2. **Rationality of agents.** In CGT, it is assumed that the agents playing a game always act in a way that maximizes their utility and that they are capable of arbitrarily complex deductions towards that end [52]. In contrast, the concept of rationality is not suited for the framework of EGT. Animals are assumed to play their genetically inherited action throughout their life, regardless of the payoffs that they receive as a result of that. Consequently, the equilibrium concepts in CGT (like Nash equilibrium) requires rational agents making strategic decisions. In EGT, equilibrium is reached through the forces of *Darwinian natural selection* (which selects better performing actions/types) and *mutation* (which selects actions/types at random), and the ensuing population dynamics.

In this thesis, we are interested in studying contests between two distinct populations of animals, denoted by A and B. In particular, animals in population A only contest against animals in population B, and not against each other. Similarly, animals in population B contest only against animals in A, and not against each other. Let N_A and N_B be the size of the populations A and B, respectively. Denote the set of n pure actions (or types) available to animals in population A by $S_A = \{a_1, \ldots, a_n\}$, and let the set of m pure actions (or types) available to animals in population B by $S_B = \{b_1, \ldots, b_m\}$. In a conflict between an animal from population A of type a_i, and an animal in population B of type b_j, the payoff to animals of type a_i and b_j are $u_A(a_i, b_j) = u_{ij}$ and $u_B(b_j, a_i) = v_{ji}$, respectively. Denote the payoff matrices for A and B by $U = (u_{ij})$ (an $n \times m$ matrix) and $V = (v_{ji})$ (an $n \times m$ matrix), respectively.

We define the state of the population A by the row vector $\mathbf{x} = (x_1, \ldots, x_n)$, where x_i gives the number of animals of type a_i in population A. In addition, we require $\sum_{i \in \{1,\ldots,n\}} x_i = N_A$. Similarly, we define the state of population B by the row vector $\mathbf{y} = (y_1, \ldots, y_m)$, and require $\sum_{j \in \{1,\ldots,m\}} y_i = N_B$. The proportional distribution of the different types in population A and B is then given by

$$\boldsymbol{\sigma}_A = \frac{\mathbf{x}}{N_A} = \left(\frac{x_1}{N_A}, \ldots, \frac{x_n}{N_A}\right) \quad \text{and} \quad \boldsymbol{\sigma}_B = \frac{\mathbf{y}}{N_B} = \left(\frac{y_1}{N_B}, \ldots, \frac{y_m}{N_B}\right),$$

respectively. Note that we can think of $\boldsymbol{\sigma}_A$ and $\boldsymbol{\sigma}_B$ as mixed actions that assign a probability to each of the pure actions. In other words, a population distribution can be represented using a mixed action. We will call a mixed action associated with a population distribution as the *mixed action representation* of the population.

Similarly, we can think of the pure action a_i as a row vector of length n, such that all its elements are zero other than the i^{th} element, which is equal to one. Similarly, we can think of the pure action b_j as a row vector of length m, such that all its elements are zero other than the i^{th} element, which is equal to one. Let $\pi_A^{a_i}(\mathbf{y})$ represent the Darwinian fitness of type a_i in population A, relative to a population B in state $\mathbf{y}$. Similarly, let $\pi_B^{b_j}(\mathbf{x})$ represent the Darwinian fitness of type b_j in population B, relative to a population A in state $\mathbf{x}$. Note that the fitness of animals in population A does not depend on the frequency distribution of the different types in A, which is a consequence of our assumption that animals in A do not fight against each other. Similarly, the fitness of the different types in population B is independent of the frequency distribution of different types in B.

From Definition 2.2.1, we deduce that $\pi_A^{a_i}(\mathbf{y})$ is given by the expected payoff obtained by the animal of type a_i following a contest with a randomly chosen member of population B. Therefore, using (2.1), we get

$$\begin{aligned}\pi_A^{a_i}(\mathbf{y}) &= u_A(a_i, \boldsymbol{\sigma}_B)\\ &= u_A\left((0,\ldots,1,\ldots,0),(y_1/N_B,\ldots,y_m/N_B)\right)\\ &= \sum_{j=1}^{m}\frac{y_j u_A(a_i,b_j)}{N_B}\\ &= a_i^T U\boldsymbol{\sigma}_B, \qquad (2.2)\end{aligned}$$

where we used that fact that u_A is a bilinear function, as a result of (2.1). Similarly, the fitness of type $b_j \in S_B$ against a population A in state $\mathbf{x}$ is given by

$$\pi_B^{b_j}(\mathbf{x}) = u_B(b_j, \boldsymbol{\sigma}_A) = b_j^T V\boldsymbol{\sigma}_A. \qquad (2.3)$$

Chapter 3

The Basic Model

We introduce our basic model in this section. The term basic is used to distinguish it from the *extended model* that we consider in Section 6. We model pairwise contests, over several generations, between animals drawn from two distinct and finite populations. We label the two different populations by A and B. In each generation, every animal in population A is uniformly randomly matched (for an infinite number of times) with animals in population B (and vice versa). During each pairwise matching, animals play a two-player game, which we call the *stage or base* game. In our model, we only consider the case in which the stage game is a two-player simultaneous-move game (like the PD game in Section 2.1).

In this thesis, we investigate the *generic Hawk–Dove* (GHD) game. The name GHD is inspired by the 'Hawk–Dove' game, which was initially proposed by Maynard Smith [11]. In Fig. 3.1, we show the bi-matrix associated with GHD games. Using the notation for Maynard Smith's Hawk–Dove game, we denote the pure action space for players in both population A and B by $\{\text{Hawk}(H), \text{Dove}(D)\}$ (for the GHD game). The fact that animals in both roles A and B have the same actions available to them is purely for notational convenience. In particular, being Hawk in population could mean something different from being a Hawk in population B.

		Animal B	
		Hawk (H)	Dove (D)
Animal A	Hawk (H)	P_A, P_B	T, S
	Dove (D)	S, T	R, R

Figure 3.1: Bi-matrix of payoffs for the generic Hawk–Dove game $T > R > S$.

The GHD game can be interpreted as follows: in a contest between two animals, an animal behaving like a Hawk corresponds to it escalating and continuing to do so until injured or until opponent retreats. Dove however corresponds to not displaying any aggression and retreating at once if the opponent escalates. The payoff labels (T, R, S, P_A and P_B) are borrowed from the commonly used labels in the literature for PD games. In particular, if both animals play Dove, then the resource is shared equally between the two contestants (both animals receive a 'reward' R). If one animal plays Hawk while the other plays dove, the Hawk obtains the resource (with value equal to the 'temptation' T) and the Dove retreats before being injured (and thus gets the 'sucker's payoff' S). If both animals play Hawk, then each contestant can injure the opponent and obtain the resource (resulting in punishment payoffs P_A and P_B). The values of P_A and P_B reflect the expected gain/loss of fitness for animals in position A and B, respectively, from such an escalated contest. Note that in general, the values of P_A and P_B are not equal. This asymmetry reflects the difference in fighting abilities between animals in different roles. For example, if the role corresponds to size, the larger animal would sustain less injuries in a contest as compared to the smaller animal.

We are interested in studying two distinct populations of animals that interact over many generations in the GHD game. In particular, we are interested in the long-run behaviour of the population distributions (including the conditions under which equilibrium is reached and its stability). As we discussed in Assumption 2, Chapter 2.2, equilibrium selection in EGT is determined by two forces: *natural selection* and *mutation*.

Natural selection. Animals are viewed as genetically coded with an action. Those actions (or animal types) that perform better in a generation (relative to the other members in the population) produce more offspring and their proportion in the next generation increases. This is a consequence of the Darwinian evolutionary principle of 'the survival of the fittest'. One can think of selection as a biological mechanism in which the (Darwinian) fitness determines the number of descendants, so the share of better strategies increases. We model the inheritance of animal types over generations using a probabilistic rule that respects the Darwin's maxim of natural selection, which is discussed in further detail in Chapter 3.1.

Mutation. Mutation gives the other main ingredient of our evolutionary model. In contrast to (natural) selection, mutation is relatively rare, and it generates strategies at random, be they better or worse. Most of the current models for studying animal conflicts using EGT model mutations as isolated events [11, 52, 51]. Such models are based on the concept of *evolutionary stable action*, which was first proposed by Maynard Smith [11]. An evolutionarily stable action is one which if adopted by a population in a given environment, cannot be invaded by any alternative action that is initially rare. In other words, suppose all animals are genetically programmed to play a certain action in a game. Let's assume that a small population share of animals, who are likewise programmed to play some other action, are injected into the original population. The original action is said to be evolutionarily stable action if, for each such mutant action, there exists a positive invasion barrier such that if the population share of animals playing the mutant action falls below this barrier, then the incumbent action earns a higher payoff than the mutant action. Note however, that the evolutionary stability concept assumes that mutations occur as isolated phenomena. However, it is possible that a population can be invaded by several different types of mutants at the same time. We therefore model mutations by introducing a stochastic element to the Darwinian dynamic process, which is inspired by the work done by Kandori *et al.* [54] and Young *et al.* [53]. The formal details are discussed in Chapter 3.1.

Before we move on to give the formal description of our basic model, we discuss its main assumptions:

1. **Contests.** All contests are between a pair of animals; one of which is from population A, and the other is from population B. In addition, each animal from a particular population contests a resource against a uniformly randomly matched individual from the opposing population. Therefore, an animal in population A has an equal probability of meeting each different member of population B, and vice-versa.

2. **Asymmetry.** There is no mixing between the different populations, and animals know for certain which population they are in. In other words, animals from population A cannot move to population B, and vice versa.

3. **Resource.** The resource being contested for is limited and divisible.

4. **Reproduction.** In both the present model and the model presented in Chapter 6, we assume that the animals reproduce asexually.

5. **Memory.** In each generation, each animal plays the stage game an infinite number of times against members of the opposing population. Let's call each round in this infinite series a *period* of the infinitely repeated stage game. In

our basic model, we assume that animals have no memory; that is, they do not use the results of previous periods to change their strategies for future periods. We will relax this assumption in the extended model, presented in Chapter 6.

3.1 Formal Description of the Basic Model

Consider populations A and B of size N_A and N_B, respectively. Let both populations A and B consist of two types of animals, H and D. We consider the population distribution at discrete *generations* $n = 1, 2, \dots$. At the beginning of generation n, each animal in population A and B *inherits* (explained below) its type (Hawk or Dove) for the generation. Let x^n be the number of Hawk in population A at generation n, and y^n the number of Hawks in population B at generation n. This defines the state of both the populations at generation n can be expressed by $z^n = (x^n, y^n)$. Note that $z^n \in Z = \{0, 1, \dots, N_A\} \times \{0, 1, \dots, N_B\}$, where Z is the state space of the system.

In each generation, every animal in population A plays the stage game an infinite number of times, each time with a uniformly randomly chosen animal from population B. Similarly, each animal from population B plays the stage game an infinite number of times against a uniformly randomly chosen animals from population A. We denote the payoff matrices of A and B by U and V, respectively, where

$$U = \begin{bmatrix} P_A & T \\ S & R \end{bmatrix}, \qquad V = \begin{bmatrix} P_B & T \\ S & R \end{bmatrix}. \tag{3.1}$$

Let π_0 be the fitness of each animal (in either role) at the start of the generation. Additionally, let $\pi_A^H(y^n)$ be the fitness of an animal from population A and of type H, relative to a population B that is in state y^n. Note that the fitness of animals in role A is a function of the state of the population B only. We similarly define $\pi_A^D(y^n)$, $\pi_B^H(x^n)$, and $\pi_B^D(x^n)$. At generation n, let $\boldsymbol{\sigma}_A^n$ and $\boldsymbol{\sigma}_B^n$ denote the proportional distribution of different types in population A and B, respectively. We then have

$$\boldsymbol{\sigma}_A^n = \left(\frac{x^n}{N_A}, 1 - \frac{x^n}{N_A}\right), \; \boldsymbol{\sigma}_B^n = \left(\frac{y^n}{N_B}, 1 - \frac{y^n}{N_B}\right).$$

Using (2.2) and (2.3), one can calculate the expected fitness obtained by an animal, following an infinite number of contests against the opposing population in a given generation. If we let $\mathbf{i} = (1, 0)$ and $\mathbf{j} = (0, 1)$ to be the unit row vectors, the expected fitnesses are

$$\pi_A^H(y^n) = \pi_0 + \mathbf{i}^T U \boldsymbol{\sigma}_B^n, \qquad \pi_A^D(y^n) = \pi_0 + \mathbf{j}^T U \boldsymbol{\sigma}_B^n, \tag{3.2}$$

$$\pi_B^H(x^n) = \pi_0 + \mathbf{i}^T V \boldsymbol{\sigma}_A^n, \qquad \pi_B^D(x^n) = \pi_0 + \mathbf{j}^T V \boldsymbol{\sigma}_A^n. \tag{3.3}$$

The value of π_0 is selected such that the fitness (for all types and positions) is strictly positive in all generations.

Recall that we assumed that animals reproduce asexually, in numbers proportional to their fitness. This can be modelled by a probabilistic map $\mathbf{f} = (f_A, f_B)$, where f_A and f_B give the number of H types in the next generation in populations A and B, respectively. For a population distribution (x^n, y^n) at generation n, the number of H types in the next generation, $f_A(x^n, y^n)$, is drawn from the binomial distribution with parameters N_A and p_A, where p_A is the probability of success in each trial of the binomial distribution. This probability, p_A, is a function of the state (x^n, y^n), and is equal to the ratio of the total fitness of H types, which is $x^n\pi_A^H(y^n)$, to the total fitness obtained by all animals in that generation, which is $x^n\pi_A^H(y^n) + (N_A - x^n)\pi_A^D(y^n)$. By similarly defining the selection map for population B, $f_B(x^n, y^n)$, we arrive at the following selection dynamics:

$$f_A(x^n, y^n) \sim \mathrm{Bin}(N_A, p_A), \qquad \text{where } p_A = \frac{x^n\pi_A^H(y^n)}{x^n\pi_A^H(y^n) + (N_A - x^n)\pi_A^D(y^n)}, \tag{3.4}$$

$$f_B(x^n, y^n) \sim \mathrm{Bin}(N_B, p_B), \qquad \text{where } p_B = \frac{y^n\pi_B^H(x^n)}{y^n\pi_B^H(x^n) + (N_B - y^n)\pi_B^D(x^n)}. \tag{3.5}$$

We refer to this map as the *selection map.* Note that the definition of selection map ensures that extinct types stay extinct (apart from mutation, which we discuss below).

Earlier in this section we discussed that natural selection and mutation are the two major forces of change in evolutionary dynamics. Natural selection is governed by the mapping $\mathbf{f}$. We now discuss the procedure through which mutations enter the populations. Our model of mutations is inspired by the work of Kandori *et al.* [54] and Young *et al.* [53]. We assume that at the start of each generation, after all of the animals (in both populations A and B) have inherited their respective types (according to the selection map), each animal changes its type (independently of others) with probability ϵ. In particular, every animal in population A that inherited the H type at the start of a particular generation changes its type to D with probability ϵ (independently of the other animals in the population). By a similar mechanism, each D type in population A can change or 'flip' its inherited type at the start of each generation, independent of the other animals in the population. Similarly, each type in population B can 'flip' its type at the start of every generation. After the mutations have occurred, the distribution of the different types in both the populations stays the same for the rest of the generation. We can incorporate the mutations in the evolutionary system that describes the state of the population from one generation to the next. This yields a non linear stochastic difference equation.

$$(x^{n+1}, y^{n+1}) = \underbrace{(f_A(x^n, y^n), f_B(x^n, y^n))}_{\text{selection}} + \underbrace{(q^n, r^n) - (s^n, t^n)}_{\text{mutation}}, \tag{3.6}$$

where w_n, x_n, y_n and z_n have the binomial distributions:

$$q^n \sim \mathrm{Bin}(N_A - f_A(\mathbf{z}^n), \epsilon), \quad r^n \sim \mathrm{Bin}(N_B - f_B(\mathbf{z}^n), \epsilon), \tag{3.7}$$

$$s^n \sim \mathrm{Bin}(f_A(\mathbf{z}^n), \epsilon), \qquad t^n \sim \mathrm{Bin}(f_B(\mathbf{z}^n), \epsilon). \tag{3.8}$$

Chapter 4

Numerical Experiments

We investigate the GHD game, using the basic model. We have divided the chapter into sections according to the number of Nash equilibria that the stage game supports. In particular, if the GHD game has a unique pure action Nash equilibria, the state of the population converges towards it. For example, if $P_A \geq P_B > S$ in the GHD game, the unique Nash equilibria of the game is (H, H). Therefore, Dove types in both population A and B will go extinct, under natural selection, and the state of the populations will eventually settle at (N_A, N_B). In addition, simulations reveal that such a state is robust against mutations (see Appendix B). Therefore, when the stage game has a unique pure action Nash equilibria, the dynamics are trivial. For a detailed analysis of all cases of the GHD game with unique Nash equilibria, and the accompanying simulations, the reader should refer to Appendix B.

We would like to remark at this stage that the PD game discussed in Chapter 2 also has a unique pure action Nash equilibria, namely (Defect, Defect). In other words, our basic model suggests that under the influence of natural selection and mutations, animals in a population will always defect, and no cooperation will be seen eventually. This, however, is not in consistence with real world observations of animals and humans. The fields of biology and social sciences are ripe with examples in which animals cooperate with each other, even though they possess the ability to defect and obtain a higher payoff, as has already been discussed in the introductory chapter. We shall address this deficiency of the basic model in Chapter 6.

In the remainder of this chapter, we focus on the investigation of the cases in which the stage game has either 2 pure action NE. In other words, we focus our attention on the GHD game with parameters $T > R > S > P_A \geq P_B$. This case corresponds to the scenario in which both animals incur a net loss from fighting. Note that when $S > P_A \geq P_B$, both animals earn a lower payoff from fighting (that is, when both play Hawk), as compared to retreating in the face of an opponent that escalates (that is, play Dove in response to opponent's Hawk action). In such a case, fighting is not beneficial for either player, and both would benefit from being part of a dominance hierarchy (be it as subordinates or dominants). Dominance hierarchy here refers to the case when all animals in one population are of type Hawk and all animals in the other population are Doves. For example, if all animals in population A are of type Hawk and all animals in population B are of type Dove, we say that a dominance hierarchy exists with population A as the dominants and population B as the subordinates. The $P_A \geq P_B$ corresponds to the fact that fighting, although costly for both populations, is more so for population B.

Two Nash equilibria of the stage game consist of pure actions: namely, (H, D) and (D, H). The third is a mixed action Nash equilibrium. As a consequence of Proposition 2.1.1, we know that in any mixed action Nash equilibrium, either agent (A or B) is indifferent over the actions over which it is randomising over. This can be used to calculate the mixed action Nash equilibria. Consider the state of the population $z = (x, y) \in Z$. The differences between the fitnesses of Hawks and Doves, in populations A and B, are given by

$$\begin{aligned}\pi_A^H(y) - \pi_A^D(y) &= (T - R) - \frac{y}{N_B}(T + S - P_A - R),\\ \pi_B^H(x) - \pi_B^D(x) &= (T - R) - \frac{x}{N_A}(T + S - P_B - R),\end{aligned}$$

respectively. Let y^* denote the critical value of y, for which the fitness of Hawks and Doves is equal in population A. Similarly, let x^* be the critical value of x at which the fitness of Doves and Hawks in population B is the same. By Theorem 2.1.1, (x^*, y^*) is the mixed action Nash equilibrium of the stage game, and is equal to:

$$x^* = \frac{N_A(T - R)}{T + S - P_B - R}, \; y^* = \frac{N_B(T - R)}{T + S - P_A - R}. \tag{4.1}$$

Note that x^* and y^* are the critical levels of populations for which the following conditions are true

$$\operatorname{sign}(\pi_A^H(y) - \pi_A^D(y)) = \operatorname{sign}(y^* - y), \qquad \forall\, y \in \{1, \dots, N_B\}, \tag{4.2}$$

$$\operatorname{sign}(\pi_B^H(x) - \pi_B^D(x)) = \operatorname{sign}(x^* - x), \qquad \forall\, x \in \{0, 1, \dots, N_A\}. \tag{4.3}$$

In other words, the fitness of Hawks is strictly greater than that of Doves in population A if the number of Hawks in population B is strictly less than y^*. Similarly, x^* is the critical level of Hawks in population A above which Hawks in population B are favoured over Doves.

Note that the expected value of a random variable $X \sim \text{Bin}(N, p)$ is $\mathbb{E}[X] = np$. From (3.4), we therefore deduce that

$$\begin{aligned} \frac{x^n}{\mathbb{E}[f_A(x^n, y^n)]} &= \frac{x^n \pi_A^H(y^n) + (N_A - x^n)\pi_A^D(y^n)}{N_A \pi_A^H(y^n)} \\ &= \frac{x^n}{N_A} + \frac{\pi_A^D(y^n)}{\pi_A^H(y^n)} \left(1 - \frac{x^n}{N_A}\right). \end{aligned} \tag{4.4}$$

This implies,

$$\text{sign}(\pi_A^H(y^n) - \pi_A^D(y^n)) = \text{sign}(\mathbb{E}[f_A(x^n, y^n)] - x_n), \tag{D1}$$

that is, if Hawks obtain a higher payoff than Doves at generation n, the expected number of Hawks in the generation $n+1$ is greater than at generation n. Using (3.5), a similar condition can be derived for population B:

$$\text{sign}(\pi_B^H(x^n) - \pi_B^D(x^n)) = \text{sign}(\mathbb{E}[f_B(x^n, y^n)] - y_n). \tag{D2}$$

Note that condition (D1) and (D2) are a manifestation of Darwin's principle of the 'survival of the fittest'. Using equations (4.2), (4.3), (D1), and (D2), we can deduce the following expressions:

$$\text{sign}(\mathbb{E}[f_A(x^n, y^n)] - x_n) = \text{sign}(y^* - y^n), \qquad \forall\, n = \{1, 2, \dots\}, \tag{4.5}$$

$$\text{sign}(\mathbb{E}[f_B(x^n, y^n)] - y_n) = \text{sign}(x^* - x^n), \qquad \forall\, n = \{1, 2, \dots\}. \tag{4.6}$$

Equations (4.5) and (4.6) show that if the number of Hawks in a population is less than the critical value, then the expected number of Hawks in the next generation will be greater or equal to the number of Hawks in this generation.

The selection map has five fixed points; namely $(0, 0)$, (N_A, N_B), $\alpha = (N_A, 0)$, $\beta = (0, N_B)$ and $z^* = (x^*, y^*)$. The first two out of these five are trivial, as they are not a characteristic feature of the payoff matrices. The other three, namely α, β and z^* correspond to the Nash equilibria of the GHD game, with z^* corresponding to the mixed action Nash equilibrium, and α and β corresponding to the pure action Nash equilibria. Given a fixed point $z \in Z$ of the selection map $\mathbf{f}$, we define its *basin of attraction* as that set $\{\tilde{z} \in Z \mid \exists\, n \in \mathbb{N} \text{ s.t. } \mathbf{f}^n(\tilde{z}) = z\}$; that is, it is the set of states $\tilde{s}$ that eventually approach z in the absence of mutations. Note that since the selection map we use is probabilistic, we work henceforth with its expected values. That is, whenever we write $\mathbf{f}(z)$, we mean the expected value of $\mathbf{f}(z)$. Conditions (4.6) and (4.6) divide the state space Z into four distinct regions, which we label in Fig. 4.1 as regions 1, 2, 3, and 4. Note that for the case under consideration ($S > P_A \geq P_B$),

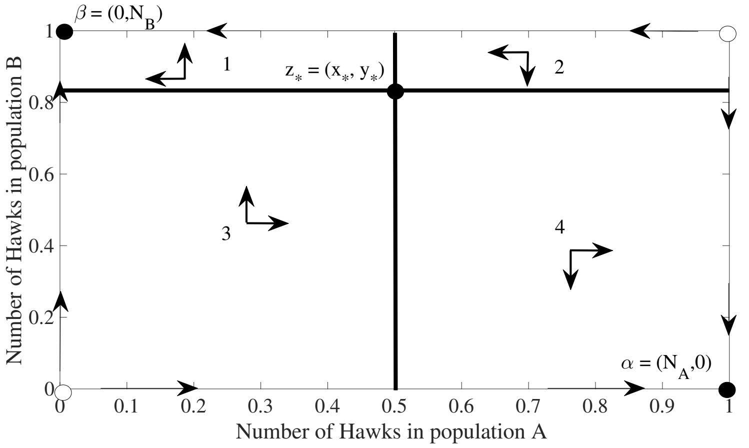


Figure 4.1: The state space Z, along with the five absorbing states. Solid dots are the non-trivial fixed points. The four regions 1, 2, 3, and 4. The arrows show the direction of the vector field $\mathbf{f}$.

both x^* and y^* are strictly bounded between 0 and 1, and therefore each region has a non-zero area.

A population distribution contained within region 1 of Fig. 4.1 has less Hawks in population A than the critical level x^*, and more Hawks in population B than the critical level y^*. Therefore, from conditions (4.5) and (4.6), the proportion of Doves in the next generation in population A increases; in contrast, in population B the proportion of Doves decreases. Therefore, a trajectory of the population distribution starting in subregion 1 will move in the positive direction along the vertical direction and in the negative direction in the horizontal direction (as shown in Fig. 4.1). This is true for trajectories starting on the edges of subregion 1 as well (except at the state s^*). A trajectory in subregion 1 will therefore move towards the fixed point $\beta = (0, N_B)$. Consequently, subregion 1 (except z^*) is a subset of the basin of attraction of the fixed state β. Using similar arguments, one can verify that subregion 4, including the edges but excluding z^*, is a subset of the basin of attraction of the fixed point $\alpha = (N_A, 0)$. We will therefore refer to the fixed points α and β as *attractors.*

Let's now consider a population state in subregion 3. Note that such a state contains fewer Hawks in population A than the critical level x^*. Similarly, the number of Hawks in population B in such a state is less than the critical level y^*. Therefore, according to conditions (4.5) and (4.6), the expected number of Hawks is greater in the next generation as compared to this generation, for both population A and B. A trajectory of the state of the population, with its initial state in subregion 3, will therefore move in the positive direction on both the vertical and horizontal axes (as

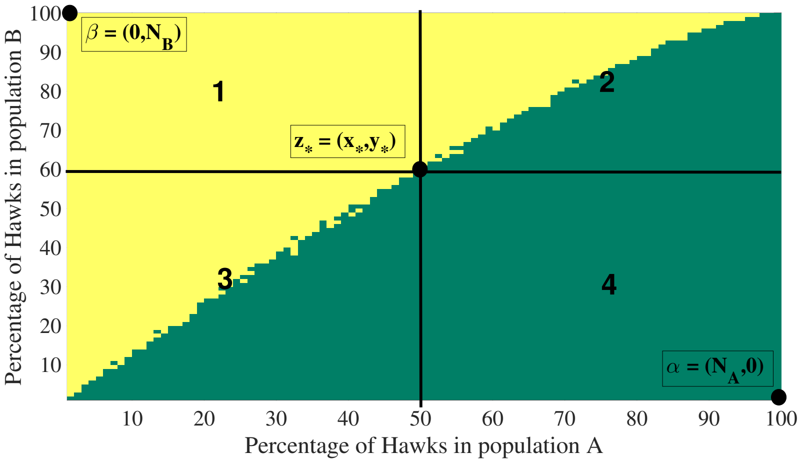


Figure 4.2: State space Z along with the four regions 1, 2, 3, and 4. Trajectories that started in the yellow region converged to α, and the trajectories that started in green region converged to β. Simulation parameters: $T = 1$, $R = 1/2$, $S = 0$, $P_A = -1/3$, $P_B = -1/2$, $\pi_0 = 1$, $\epsilon = 0$ and $N_A = N_B = 10^4$.

shown in Fig. 4.1). Using similar arguments, one can verify that a trajectory starting in subregion 2 will move in the negative direction both horizontally and vertically (as shown in Fig. 4.1).

The attractor to which a particular trajectory leads to depends on the initial state and the relative speed of adjustments of the dynamical process; that is, the expected rate of increase (over the generations) in better performing type in A vs. rate of increase in better performing type in B. In Fig. 4.2 shows the convergence of trajectories on a 100×100 lattice (for the selection map $\mathbf{f}$), where the colour of the $(i, j)^{\text{th}}$ term shows the state to which the trajectory converges in the absence of mutations, starting from the initial state of $(iN_A/100, jN_B/100)$[1]. The trajectories that started in the yellow region converged to β, and the trajectories that started in the green region converged to α. It can be seen from Fig. 4.2 that the absorbing state z^* behaves like a saddle-point.

Let's now investigate the behaviour of the system (3.6) in the presence of mutations. Note that when the mutation rate is 0, the system has two attractors to which trajectories in the state space converge (α and β). In the system with a small but non-zero mutation rate, simulations show that the position of β shifts vertically

---

[1] All simulations in this Chapter have been performed using code that was written by the author of this thesis, but is based on a much more extensive code written by Dr. Cameron Hall. In particular, the author of this thesis added the features that simulate the selection map $\mathbf{f}$ and the mutations.

downward and horizontally to the right, and the position of α shifts vertically upward and horizontally to the left. Denote the shifted attractors by $\bar{\alpha}(\epsilon)$ and $\bar{\beta}(\epsilon)$, in order to differentiate them from α and β. Note that $\bar{\beta}(\epsilon = 0) = \beta$ and $\bar{\alpha}(\epsilon = 0) = \alpha$. To see why this shift occurs, consider the variables (q^n, r^n), and (s^n, t^n) in (3.6), which control the levels of mutation in the system. Note that the expected value of a random variable drawn from a binomial distribution with parameters n and p is np. Therefore,

$$\mathbb{E}[q^n - s^n] = (N_A - f_A(x^n, y^n))\epsilon - f_A(x^n, y^n)\epsilon = (N_A - 2f_A(x^n, y^n))\epsilon, \tag{4.7}$$

$$\mathbb{E}[r^n - t^n] = (N_B - f_B(x^n, y^n))\epsilon - f_B(x^n, y^n)\epsilon = (N_B - 2f_B(x^n, y^n))\epsilon. \tag{4.8}$$

In other words, if $f_A(x^n, y^n)/N_A > 1/2$ (that is, if population A has more Hawks than Doves), we expect mutations to increase the number of Doves in A (because $\mathbb{E}[q^n - s^n] < 0$). Similarly, if the number of Doves in population A is greater than that of Hawks, then we expect mutation to increase the number of Hawks in A. Using similar arguments for population B, we thus conclude that in the state space Z (see Fig. 4.1), mutations drive the system towards the center of the state space $(N_A/2, N_B/2)$. In particular, in regions 1 and 4 of the state space, mutations act against selection pressures. Therefore, for small mutation rates, we expect the attractors α and β to move off the corner of the state space into its interior, to a point where the opposing forces of natural selection and mutation balance each other. In Fig. 4.3, we show six different trajectories that converge to $\bar{\alpha}(\epsilon)$ and $\bar{\beta}(\epsilon)$ for $\epsilon = 0.01$. For large mutation rates, we expect mutations to completely overcome the selection pressures. This in turn would cause both of the attractors to move very close to the midpoint $(N_A/2, N_B/2)$.

When the mutation rate is zero, the two attractors α and β are the fixed points of the selection map $\mathbf{f}$. This follows from (3.4) and (3.5), which ensure that extinct types stay extinct. For example, a trajectory that reaches $\beta = (0, N_A)$, will stay at β forever. However, when the mutation rate is non-zero, a trajectory starting near $\bar{\beta}(\epsilon)$ moves towards it, due to selection pressures. Once the trajectory has reached β, it does not have to stay there. There exists a non-zero probability that it can move away from $\bar{\beta}(\epsilon)$, because of mutations. If enough mutations occur, it can even jump to the basin of attraction of $\bar{\alpha}(\epsilon)$.

For small mutation rates, the fraction of time that the system spends in $\bar{\alpha}(\epsilon)$ and $\bar{\beta}(\epsilon)$ in the long-run depends upon three factors: (a) The rate of mutations, (b) the degree of asymmetry in the payoffs (P_A versus P_B), and (c) the selection map $\mathbf{f}$. We investigate the system by varying factors (a) and (b) to gauge their importance. We

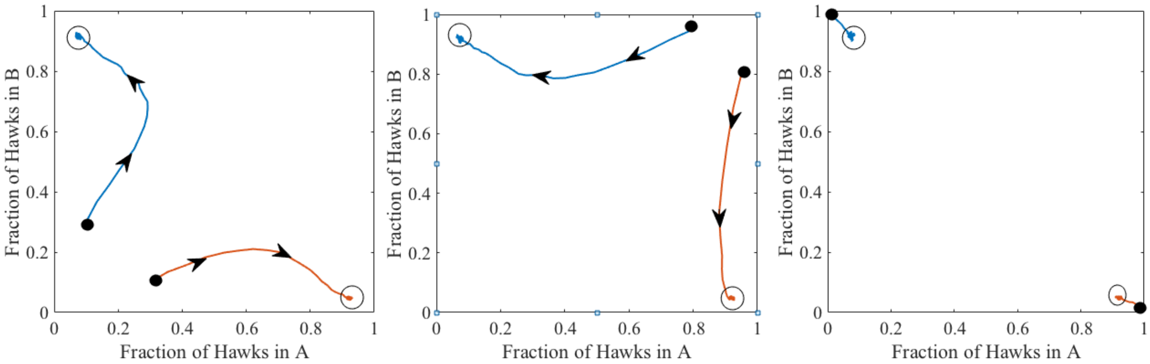


Figure 4.3: Simulation showing six different trajectories in the state space Z. The bold circles, hollow circles and the arrows have been added to the trajectories to highlight different aspects about them. The bold circles denote the initial state of the trajectory, hollow circles show the shifted attractors (due to non-zero mutation rate) to which the trajectories converge, and the arrows show the direction of the trajectories. Simulation parameters: $T = 1$, $R = -1/2$, $S = 0$, $P_A = -1/2$, $P_B = -1/3$, $\pi_0 = 1$, $\epsilon = 0.01$ and $N_A = N_B = 10^4$.

used the parameters $T = 1$, $R = 1/2$, $S = 0$, $\pi_0 = 1$ and $P_A \geq P_B$ for all our numerical simulations. Note that we can use these values for T, R and S because the normal form representation of a game is equivalent upto a linear transformation of the payoff functions (as was discussed in Chapter 2). For small asymmetries in the payoffs for populations A and B ($P_A = -0.48,\ P_B = -0.5$) and mutation rates $\epsilon < 0.025$, we conducted 10 simulations. It was found that $\bar{\alpha}$ and $\bar{\beta}$ were very close to α and β, respectively. For mutation rates $\epsilon > 0.03$, we conducted more than 10 simulation experiments, and on each occasion it was found that given enough time (about 1000 generations), trajectories always converged to $\bar{\alpha}$, regardless of the starting position. For mutation rates in the region between $0.025 < \epsilon < 0.03$, our simulation experiments reveal that the trajectories initially starting near $\bar{\beta}$ converged to it and stayed there for the length of the simulation, or they 'jumped' to $\bar{\alpha}$ after some time (see Fig. 4.4). On the other hand, trajectories that started near $\bar{\alpha}$ and converged to it initially, but were never found to 'jump' to $\bar{\beta}$ (in roughly 20 simulations), indicating that its stability is much more robust in the presence of mutations. Note, however, it would be wrong to conclude from these results that a jump from $\bar{\alpha}$ to $\bar{\beta}$ cannot occur, as there is always a non-zero probability of this occurring. What can be concluded from our simulations though, is that the probability of jumping from $\bar{\beta}$ to $\bar{\alpha}$ is much greater as compared to jumping the other way round.

We did a similar analysis the case of large asymmetries. We used the same value for P_B as before ($P_B = -1/2$), but increased the value of P_A to -1/10. Note that

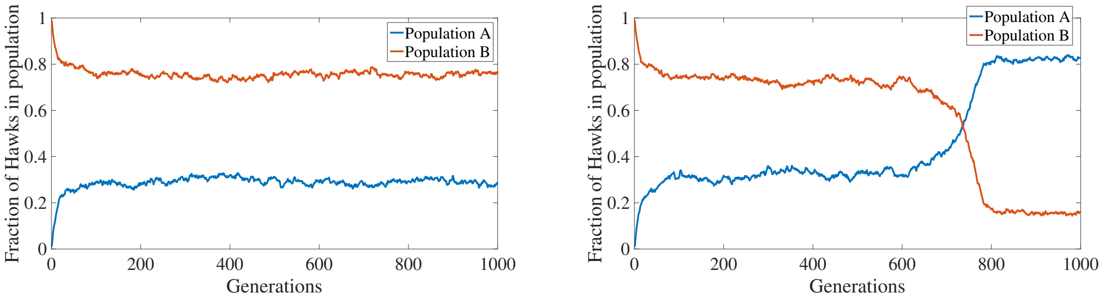


Figure 4.4: Two instance of a simulation experiment carried out with the parameters: $T = 1$, $R = -1/2$, $S = 0$, $P_A = -1/2$, $P_B = -1/3$, $\pi_0 = 1$, $\epsilon = 0.029$ and $N_A = N_B = 10^4$. In the initial state, population A was composed entirely of Doves and B entirely of Hawks. The first figure shows the state of the population settling at $\bar{\beta}(\epsilon)$ at roughly the 10th generation, and then staying there for the length of the experiment (1000 generations). The second figure shows another experiment in which the population starts from the same initial state, settles at $\bar{\beta}(\epsilon)$ until roughly the 600th period, but jumps to $\bar{\alpha}(\epsilon)$ after that.

increasing the value of P_A corresponds to increasing the value of y^* (see condition (4.1)). For example, when $P_B = -1/10$, using the parameter values $T = 1$, $R = 1/2$ and $S = 0$, the value of y^* is 0.98. The vertical height to region 1 (Fig. 4.2) would thus be 0.02. In other words, the basin of attraction of β would be much smaller than that of α. In the presence of a small mutation rate, therefore, we expect most trajectories to towards $\bar{\alpha}(\epsilon)$, and not to $\bar{\beta}(\epsilon)$. In the 10 numerical simulations that we conducted with $P_A = -1/10$ and $P_B = -1/2$, we observed that all trajectories converged to $\bar{\alpha}$ within a 100 generations for mutation rates as low as $\epsilon = 0.005$.

Discussion

In summary, in this section, we investigated the GHD game for parameter values in which two pure action Nash equilibira exist. The type of animal conflicts investigated were those in which animals from both populations incur a net loss from fighting. We found that in such conflicts, almost always a dominance hierarchy will be established, which would depend on three factors: Mutation rates, size of the asymmetry and the speed of adjustment of the selection map in the two populations. We investigate the first two of the three factors, and we found that large mutation rates and high levels of asymmetries in the payoff favour the establishment of the dominance hierarchies with the dominant animals being the ones that are favoured by the asymmetry. However, when the asymmetry is small, the weaker animals (which we took to be population B in the analysis above) can also dominate the hierarchy, especially when the mutation rates are low. This phenomena has also been observed to occur in nature [36, 37].

Although we have not investigated the effects of the relative speed of adjustments of the selection map on the establishment of dominance hierarchies, it is clear that it can have a dramatic effect. For instance, consider a selection map such that in region 2 of Fig. 4.1, animals in population B adjust much slower than animals in population A. In addition, assume that the selection map is such that in region 3, animals in population B adjust much faster than population A. In such a case, the basin of attraction of β would be the union of region 1, and most of regions 2 and 3, while the basin of attraction of α is the region 4. For small mutation rates, it would be easier for trajectories starting near $\bar{\alpha}(\epsilon)$ to escape to $\bar{\beta}(\epsilon)$, but not the other way round. Therefore, even though the asymmetry in the payoffs favours animals in population A, the hierarchy might be controlled by animals in B in the long run.

Chapter 5

Repeated Games

In the basic model (see Chapter 3), we studied two populations of animals such that, in each generation, every animal in either population plays the stage game an infinite number of times, each time against a random individual from the other population. In other words, each animal plays an *infinitely-repeated game* against the other population in every generation. One of the assumptions of the basic model was that the animals were memory less; that is, in each of round of the repeated game, their action does not depend on the prior action responses.

In Chapter 6, we will relax this assumption to allow animals to possess a memory of size 1, in that they have the ability to change their response based on the opponent's action in the previous round of the stage game. To do so, we need to formalise the notion of *repeated games.* We present definitions that closely follow the ones given by Fujiwara [43].

Definition 5.0.1. A *infinitely repeated* game G^T consists of a set I of players playing the simultaneous-move game $G = \{I, \boldsymbol{\mathcal{A}}, \mathbf{u}\}$ for a given number of times T. We refer to G as the *stage game.* If T is a positive integer, the game is said to be *finitely repeated.* It is *infinitely repeated* if $T = \infty$.

In this and the subsequent chapters, we only consider repeated games in which the stage game is a two-player simultaneous-move game (for example the PD game).

We call the play of each round of the stage game a *period* of the repeated game. In a repeated game, a player plays a particular action from his action space in each period of the repeated game. We refer to the sequence of actions taken by a particular player in a repeated game as its *strategy.* It is important to keep in mind the difference between an action and a strategy. An action is a decision taken by a particular player in a particular period of the repeated game. A strategy for a player in a repeated game is a complete contingency plan for every period of the game.

To formally define the concept of a pure strategy for a repeated game, we first need to specify what is observed by the players after each period. We assume that each player observes the pure actions of its opponents from the preceding periods. We can then define a pure strategy for a player as a function which prescribes, for all periods of the game and for every possible sequence of pure-action profiles (that is, observations) up to that period, an action of the stage game. To formalize this discussion, consider an infinitely repeated game G^{∞}, where $G = \{I, \boldsymbol{\mathcal{A}}, \mathbf{u}\}$ is the stage game with the set of players $I = \{1, 2, \dots, N\}$, the set of actions available to each player is $\boldsymbol{\mathcal{A}} = \{\mathcal{A}_i\}_{i \in I}$ and the payoff functions are $\mathbf{u} = \{u_i\}_{i \in I}$. Define a sequence of pure action profiles until period $t \in \{1, 2, \dots\}$ a *history* up to period t. Therefore, the set of histories until period t, denoted by H_t, is given by

$$H_t = (\mathcal{A}_1 \times \mathcal{A}_2 \times \cdots \times \mathcal{A}_N)^{t-1} = \left(\prod_{i=I} \mathcal{A}_i \right)^{t-1}$$

By convention, the set of histories until the first period is the empty set $\emptyset$.

Definition 5.0.2. A pure strategy s_i of player i in G^{∞} is a sequence $s_i = (s_{i1}, s_{i2}, \dots)$ such that for each period $t \in \{1, 2, \dots\}$,

$$s_{it} : \left(\prod_{i=I} \mathcal{A}_i \right)^{t-1} \to \mathcal{A}_i.$$

Definition 5.0.3. A strategy $s_i = (s_{i,1}, s_{i,2}, \dots)$ for player i is *memory-less* if for all histories $h, h' \in H_t$, we have $s_{i,t}(h) = s_{i,t}(h')$. If a strategy is not, we call it a *memory strategy.*

We denote the set of all possible pure strategies available to player i by S_i, and define $\mathbf{s} = (s_1, s_2, \dots, s_N) \in \mathbf{S}$ (where $\mathbf{S} = S_1 \times S_2 \times \cdots \times S_N$) as a pure strategy profile. Note that although we have defined a pure strategy for a player i to be a vector of functions $s_i = (s_{i1}, s_{i2}, \dots)$, it is sometimes convenient to think s_{it} as the action of player i at the t^{th} period, assuming that the last period of the game has already been completed. Having defined pure strategies for a repeated game, we can define mixed strategies for a player in much the same way as for a stage games (see Definition 2.1.2, Chapter 2). It is a probability distribution over the set of all pure strategies. We use σ_i to denote mixed strategies of the player i in a repeated game. Similarly, other definition of Chapter 2, including that of best response action, strictly dominated action, and NE action profile can be extended to best response strategy, strictly dominated strategy, and NE strategy profile, respectively.

We now need to define a measure of the payoffs for repeated games. In an infinitely repeated game, a player receives an infinite sequence of payoffs. Therefore, we have to define the payoffs for an infinite game in so that for any sequence of payoffs of the stage game, the total payoff is always finite. We therefore introduce a *discount factor* $\delta \in (0,1)$, which we assume to be common to all players [43]. We can then define the payoff function $\mathbf{v} = (v_1, \ldots, v_N)$ as

$$\mathbf{v} : \mathbf{S} \to \mathbb{R}^N \text{ such that } v_i(\mathbf{s}) = \sum_{t=1}^{\infty} \delta^{t-1} u_i(s_{1t}, \ldots, s_{Nt}) \ . \tag{5.1}$$

Note that this definition ensures that the payoff for an infinitely repeated game is finite.

One can interpret the discount factor as follows: At any given period of the game, it is the probability that the game continues to the next period. If the game stops at some period (with probability $1-\delta$), the future payoffs are 0. Such an interpretation is convenient because we no longer need to assume that players play an infinite number of games. We assume that at the start of the game, the players involved do not know for certain the total number of periods of the repeated game. The expected number of periods/rounds of the stage game are given by $1+\delta+\delta^2+\cdots = 1/(1-\delta)$. Henceforth, we denote infinite games by $G^{\infty}(\delta)$.

Chapter 6

The Extended Model

Recall in Chapter 4, we discussed that when $P_A \geq P_B \geq S$, the GHD game has a unique pure action NE profile (Hawk, Hawk). In such a case, all Doves in both population A and B will become extinct eventually, regardless of the initial state of the population. Note that the PD game discussed in Chapter 2 (see Fig. 6.2) is a special case of the GHD game with $P_A \geq P_B \geq S$. In particular, the PD game has a unique Nash equilibria (Defect, Defect). If the stage game used in the basic model (see Chapter 3) was the PD game, all cooperators would thus become extinct eventually. Therefore, our basic model predicts that if animal conflicts are modelled using the PD game, cooperation cannot be sustained.

Experimental results from biology and social sciences, however, contradict the above predictions. It has been observed that animals cooperate with each (equivalent to playing Dove in the GHD framework), despite having the option to change their behaviour and obtain the temptation payoff, which is strictly greater than the payoff recieved when both cooperate. For example, cooperative behaviour has been observed in honeybees [60] and vervet monkeys [61]. It has been observed that vampire bats regurgitate blood that they have obtained to give it to a hungry member of their colony [62]. Experiments conducted on human subjects playing a repeated Prisoner's Dilemma game have exhibited similar patterns of cooperation [64].

In this section, we show that the contradictions between the results of the basic model and existing experimental observations can be resolved by assuming that

1. **Infinitely repeated games**. The players play an infinite number of rounds (or 'periods') of the PD game at each generation.
2. **Memory**. Animals condition their response in each round based on the preceding responses of their opponents.

For a formal game theoretic definition of infinitely repeated games, the reader should refer to Appendix 5. In addition to these two assumptions, the first four assumptions of the basic model (the assumptions about contests, asymmetries, resource and reproduction), discussed in Chapter 3, are kept intact in the extended model. Note that the memory assumption of the extended model is different from that of the basic model. Although we assumed an infinite play of the stage game in the basic model, such an assumption was not necessary for the results of Chapter 3 and 4. However, the assumption of infinitely repeated games is crucial for the studying the evolution of cooperation, and thus is an important assumption in the extended model. The framework of the interactions in the extended model is the same as that of the basic model. In particular, we study two distinct population's of animals over successive generations $n = 1, 2, \ldots$. During each generation, every animal from either population plays a stage game for an infinite number of times, each time with a uniformly randomly chosen opponent of the opposing population.

We investigate the emergence and maintenance of cooperation in a Prisoner's Dilemma (PD) stage game, which provides a simple and convenient framework to study cooperation. Note that the PD game corresponds to the GHD game with parameters $T > R > P_A = P_B > S$ and $2R > T + S$. Because we henceforth only deal with the special case of the PD game, we refer to the strategies available to both the players as 'cooperate' (C) and 'defect' (D). Note that the condition $P_A = P_B$ ensures that the game is symmetric. As discussed in Chapter 2, the representation of a PD game is equivalent upto a linear transformation of its payoff functions. Therefore, for the rest of the work, we concentrate on the numerical values of the PD game shown in Fig 6.1.

		Animal B	
		Defect (D)	Cooperate (C)
Animal A	Defect (D)	$1, 1$	$3, 0$
	Cooperate (C)	$0, 3$	$2, 2$

Figure 6.1: A particular example of the PD game.

We have already noted that the assumption (1) of infinitely repeated game is required because if the stage game has only one Nash equilibrium, then cooperation cannot be sustained. Similarly, without sssumption (2), one cannot sustain cooperation in the PD game, regardless of whether a finite (or infinite) number of rounds are played at each generation. This is because without memory, one can only have two types of players, those who always cooperate in each round of the PD game, and those who always defect. Because cooperators are strictly dominated by defectors, cooperation cannot be sustained in such a system, as was discussed in the beginning of Chapter 4 and Appendix B. However, if individuals/animals are allowed to have memory, there are various setting in which cooperation can be sustained. For example, in the late 1970s, political scientist Robert Axelrod held a computer tournament [65] designed to investigate the PD game. Contestants in the tournament submitted computer programs that would compete in an repeated PD game of 200 rounds, using the payoff matrix in Figure 2.1. The competing strategies were not allowed to use the fact that the game ended in a finite number of periods, so it was essentially an infinite period game. The 14 entries were paired with each other in a round-robin tournament. Some of the strategies were rather intricate. An example is one that in each move models the behaviour of the other player as a Markov process, and then uses Bayesian inference to select what appear to be the best choice for the long run. However, the result of the tournament was that the highest mean score was attained by the simplest of all strategies submitted: *Tit-for-tat* (TFT). An animal that plays the Tit-for-tat strategy cooperates in the first round of the game and then copies the strategy of its opponent from the previous round. In the second tournament, in which 62 entries were submitted, Tit-for-tat emerged as the winner. Note that Tit-for-tat can sustain cooperation between animals. In nature, behaviour similar to the Tit-for-tat strategy has been observed in a variety of different species [66, 67].

The robustness of TFT in individual contests against other strategies, including the ones which are much more sophisticated than TFT, is illustrated by Axelrod's tournaments. However, the robustness in an ecological sense – that is, robustness of a population of animals of type TFT against random mutations, is still a matter of debate. In addition to TFT, there are several other one-memory strategies that can lead to cooperation between animals, and that perform well in individual contests against a wide variety of other memory strategies. We develop a model to investigate, through numerical experiments, the evolution and robustness of such memory strategies under the influence of natural selection and mutation.

6.1 Formal description of the extended model

We only study one-memory strategies. That is, strategies which depend only on the history of the action profiles used in the previous round (which we will refer to as the *play* of the previous period). Note that in a pairwise contest with two types of animals, with two actions (Defect and Cooperate) available to either of them, there are four possible outcomes. If we let a and b denote the action of an animal and its opponent, respectively, the four possible outcomes are $ab \in$ {DD, DC,CD,CC}. Note that whenever we write $ab \in$ {DD,DC,CD,CC} as the outcome of the PD game, the first element of ab (that is, a) will always represent the action of player A, and the second element will always represent the action of player B. In addition to the initial action, a one-memory strategy must specify the action to be played for each of these four outcomes. Because there are two distinct actions available to an animal at any period, there are a total of $2^5 = 32$ one-memory strategies. One can represent one-memory strategies using the five-bit notation $\mathbf{s} = (s_0|s_{DD}, s_{DC}, s_{CD}, s_{CC})$, where s_0 is the probability of defecting in the first/initial period.

Note that not all of the 32 strategies are distinct. For example, the strategy that prescribes cooperation initially as well as in response to all possible outcomes (which we call the 'All C' strategy), has a bitwise representation of $(0|0, 0, 0, 0)$. However, the strategy $(0|1, 0, 0, 0)$, is equivalent to the All C strategy. This is because it only defects if the outcome in the previous period is (D,D), which can never occur. After eliminating multiple occurrences of the same strategy, 26 distinct strategies remain. We show them in Table. C.1 in Appendix C.

Note that individual animals are only allowed to have pure one-memory strategies. In other words, if $(s_0|, s_, s_2, s_3, s_4)$ is the one-memory type of an animal, then $s_0, s_1, s_2, s_3, s_4 \in \{0, 1\}$. Therefore, both population A and B can have 26 different types of animals, and the state of a dynamical system at any given generation is specified by $\mathbf{z} = (\mathbf{x}, \mathbf{b} = \mathbf{y})^T$, where $\mathbf{x} = (x_1, \ldots, x_{26})$ and $\mathbf{y} = (y_1, \ldots, y_{26})$ are row vectors such that the i^{th} element gives the number of animals of type $i \in \{1, \ldots, 26\}$ in population A and B, respectively. Note that $\mathbf{x}$ and $\mathbf{y}$ must satisfy the conditions $\sum_{i=1}^{26} x_i = N_A$ and $\sum_{i=1}^{26} y_i = N_B$. We denote the state space of the dynamical system by Z, where

$$Z = \{(\mathbf{x}, \mathbf{y}) \mid \mathbf{x}, \mathbf{y} \in \mathbb{N}^{26} \text{ and} \sum_{i=1}^{26} x_i = N_A, \ \sum_{i=1}^{26} y_i = N_B\}.$$

We are interested in the long-term behaviour of the state of the populations. That is, we want to examine the population dynamics over the generations $n = 0, 1, 2, \ldots$. Let the state of the population in the n^{th} generation be $\mathbf{z}^n = (\mathbf{x}^n, \mathbf{y}^n) \in Z$, where $\mathbf{z}^0$ denotes the initial state of the population. During each generation, each animal in population A plays infinitely many rounds of the PD game, each time with a uniformly randomly chosen member of population B. Similarly, each animal in population B plays an infinitely repeated PD game against uniformly randomly chosen opponent from population A. Note that playing against uniformly randomly chosen members of the opposing population is mathematically equivalent to playing against the mixed strategy which represents the opposing population. Recall that the mixed strategy representation of a population gives the frequency weighted average of the occurrence of the different pure one-memory strategies in the population. Let σ_A^n and σ_B^n denote the one-memory mixed strategy representation of population A and B, respectively, in the n^{th} generation. Denoting the i^{th} one-memory strategy in Table C.1 by $s_i = (s_{i,0}|s_{i,1}, s_{i,2}, s_{i,3}, s_{i,4})$, we have

$$\sigma_A^n = \frac{1}{N_A}\left(\sum_{i\in\{1,\ldots,26\}} s_{i,0}x_i^n \middle| \sum_{i\in\{1,\ldots,26\}} s_{i,1}x_i^n, \ldots, \sum_{i\in\{1,\ldots,26\}} s_{i,4}x_i^n\right),$$
$$\sigma_B^n = \frac{1}{N_B}\left(\sum_{i\in\{1,\ldots,26\}} s_{i,0}y_i^n \middle| \sum_{i\in\{1,\ldots,26\}} s_{i,1}y_i^n, \ldots, \sum_{i\in\{1,\ldots,26\}} s_{i,4}y_i^n\right).$$

We now describe the procedure through which we calculate payoffs in our model. We only describe the case for animals in population A, and the description for population B is analogous. In generation n, one can calculate the payoff for the i^{th} type $s_i = (s_{i,0}|s_{i,1}, s_{i,2}, s_{i,3}, s_{i,4})$ in population A as the discounted infinite sum of payoffs against the mixed strategy σ^n. The initial action sequence for the infinitely repeated game is selected according to the first bitwise element of s_i and σ^n. For example, the outcome DC is selected with probability $s_{i,0}(1-\sigma_{B,0}^n)$, and the outcome CD is selected with probability $(1-s_{i,0})\sigma_{B,0}^n$. Given an outcome for a particular period t, one calculates the outcome of period $t+1$ using the probabilities prescribed by the strategies s_i and σ_B^n. For example, if the previous outcome is DC, the probability of occurrence of the outcomes (DD, DC, CD, CC) is

$$\left(s_{i,2}\sigma_{B,3}^n,\ s_{i,2}(1-\sigma_{B,3}^n),\ (1-s_{i,2})\sigma_{B,3}^n,\ (1-s_{i,2})(1-\sigma_{B,3}^n)\right).$$

Given an outcome in a particular period, the payoff is specified by the payoff bi-matrix of the PD game (plus some base fitness π_0 to ensure payoffs are always positive [1]). Denote the total payoff from the infinitely repeated PD game for animal of the i^{th}

---

[1] We need to ensure that the fitness of animals is always positive, even if the payoff received from a game is not. As was discussed in Chapter 3, the fitness of a species is never zero. For example, in an animal conflict over a territory, if an animal doesn't win the territory, the payoff received is zero. But it can always raise its offspring in a less favourable territory, and thus has a nonzero fitness

type in population A by $\pi_A^i(\mathbf{b}_n)$. It then follows that $\pi_A^i(\mathbf{b}_n)$ is the discounted sum of the infinite number of payoffs, according to the formula (5.1). One can similarly calculate the payoff/fitness gained by the i^{th} type in population B, which we denote by $\pi_B^i(\mathbf{a}_n)$.

Similar to the basic model, inheritance of strategies by a new generation is governed by a map $\mathbf{f} : Z \to Z$ (which we call the *selection map*), where

$$\mathbf{f}(\mathbf{z}^n) = \begin{pmatrix} f_{A,1}(\mathbf{z}^n), & \cdots, & f_{A,26}(\mathbf{z}^n) \\ f_{B,1}(\mathbf{z}^n), & \cdots, & f_{B,26}(\mathbf{z}^n) \end{pmatrix}. \tag{6.1}$$

Similar to the basic model, we define the map so that given a generation n, the proportion of a particular type in generation $n+1$ is equal to the total fitness achieved in the infinitely repeated PD game by that type in generation n, divided by the sum of the total fitness received by all types in generation n. Let $\boldsymbol{\pi}_A = (\pi_A^i)_{i\in\{1,\ldots,26\}}$ and $\boldsymbol{\pi}_B = (\pi_B^i)_{i\in\{1,\ldots,26\}}$ denote the vector of payoff functions of different types in population A and B, respectively. Then, for all $i \in \{1, \ldots, 26\}$

$$f_{A,i}(\mathbf{z}^n) \sim \mathrm{Mn}(N_A, (p_1, \ldots, p_{26})), \text{ where } p_i = \frac{x_i^n \pi_A^i(\mathbf{y}^n)}{\mathbf{x}^n \cdot \boldsymbol{\pi}_A(\mathbf{y}^n)}, \tag{6.2}$$

$$f_{B,i}(\mathbf{z}^n) \sim \mathrm{Mn}(N_B, (q_1, \ldots, q_{26})), \text{ where } q_i = \frac{y_i^n \pi_B^i(\mathbf{x}^n)}{\mathbf{y}^n \cdot \boldsymbol{\pi}_B(\mathbf{x}^n)}, \tag{6.3}$$

where $X \sim \mathrm{Mn}(N, (p_1, \ldots, p_k))$ denotes a random variable X drawn from the multinomial distribution with N trials and $(p_1, \ldots, p_k)$ event probabilities ($\sum p_i = 1$). Note that such a map respects the Darwinian principle of natural selection, because the expected proportion of a animal type in generation $n+1$ is given by the proportion of the payoff received by that type to the total payoff received by all types in generation n.

Recall that in the basic model discussed in Chapter 3, we modelled mutations based on the work done by Kandori *et al.* [54] and Young *et al.* [53]. We now generalise the notion of mutations used in the basic model. We assume that at the start of each generation, after all of the animals have inherited their respective types, each animal changes its type (independently of others) with probability ϵ. Therefore, the state of the next generation $\mathbf{z}_{n+1} = (\mathbf{a}_{n+1}, \mathbf{b}_{n+1})$ is given by

$$\mathbf{x}^{n+1} = \sum_{i=1}^{26} X_i, \qquad \text{where} \qquad X_i \sim \mathrm{Mn}\left(f_{A,i}(\mathbf{z}_n), \left(\frac{\epsilon}{25}, \ldots, \underbrace{1-\epsilon}_{i^{\text{th}} \text{ element}}, \ldots, \frac{\epsilon}{25}\right)\right),$$

$$\mathbf{y}^{n+1} = \sum_{i=1}^{26} Y_i, \qquad \text{where} \qquad Y_i \sim \mathrm{Mn}\left(f_{B,i}(\mathbf{z}_n), \left(\frac{\epsilon}{25}, \ldots, \underbrace{1-\epsilon}_{i^{\text{th}} \text{ element}}, \ldots, \frac{\epsilon}{25}\right)\right).$$

6.2 Some important strategies for the infinitely repeated PD game.

Recall from Chapter 4 that the fixed points of the selection map were given by the NE profiles of the stage game. In the extended model, the animals play an infinitely repeated PD game. By definition, a NE profile is one in which no player can unilaterally change from the strategy prescribed to it by that profile and expect to gain a higher payoff. Therefore, the set of attractors of the dynamical system $\mathbf{z}^{n+1} = \mathbf{f}(\mathbf{z}^n)$ (where $\mathbf{f}$ is given by (6.1) and (6.2)) will be a subset of the NE strategy profiles of the infinitely repeated PD game. To this end, we discuss a few relevant NE profiles of the infinitely repeated PD game.

Because we can linearly transform the payoff matrix associated with the PD game, we can without loss of generality use the numerical values indicated in Fig. 6.1 for the PD game. We now discuss the characteristics of a few of the important one-memory strategies from Table C.1 that we will use in Chapter 7.

All D. (number 26 in Table C.1) The All D strategy, with bitwise representation $(1|1,1,1,1)$ prescribes D initially and in response to all possible outcomes. Because (D,D) is a Nash equilibrium of the PD game, by Definition NE profiles, the strategy profile (All D, All D) is an NE profile of the infinitely repeated PD game.

All D FC. (number 25 in Table C.1) This strategy is called All D FC because it first cooperates (hence FC), and then behaves like All D. This strategy has the bitwise representation $(0|1,1,1,1)$. It is not part of any NE profile because it is strictly dominated by the All D strategy. To see this, consider the strategy profile (All D FC, All D FC), that is, both players A and B play All D FC. In such a case, the infinite sequence of outcomes of the infinitely repeated PD game are CC, CC, CC, …, where the first element in a outcome ab represents the action used by A, and the second element represents the action used by B. In this case, the payoff for both animals are $2/(1-\delta)$ (using (5.1)). Let's now suppose that player A deviates from this strategy profile and plays according to All D instead. If that happens, the infinite sequence of outcomes of the infinitely repeated PD game are DC, DD, DD, … . The payoff for player A therefore is $3 + (2\delta)/(1-\delta^2)$, which is strictly greater than $2/(1-\delta)$ for all $\delta \in (0,1)$. Player A can deviate unilaterally from the strategy profile (All D FC, All D FC) and gain a higher payoff. Therefore, this strategy profile is not an NE.

Swindler. (number 21 in Table C.1) This strategy has the bitwise representation of $(0|1,0,1,1)$. It prescribes C initially. Thereafter, it prescribes D, unless the outcome in the preceding period was DC. Similar to All D FC, it can be shown that Swindler can never be part of a NE profile, since it is strictly dominated by All D strategy for all values of $\delta \in (0, 1)$.

All C. (number 1 in Table C.1) This strategy has the bitwise representation of $(0|0, 0, 0, 0)$. It prescribes C initially and in response to all possible outcomes. Similar to All D FC, it can be shown that All C can never be part of a NE profile, since it is strictly dominated by All D strategy for all values of $\delta \in (0, 1)$.

Punisher. (number 5 in Table C.1) This strategy has the bitwise representation of $(0|0, 0, 1, 0)$. It prescribes C initially. Thereafter, it prescribes C, unless the outcome in the preceding period was CD. Similar to All D FC, it can be shown that Punisher can never be part of a NE profile, since it is strictly dominated by All D strategy for all values of $\delta \in (0, 1)$.

Tit-for-Tat (TFT). (number 9 in Table C.1) This strategy has the bitwise representation of $(0|1, 0, 1, 0)$. We have already noted that the strategy TFT is known to perform well against a range of other strategies. It prescribes cooperation in the first period, and replication of the opponent's behaviour (from the previous period) thereafter. Axelrod [65] showed that for discount factors $\delta \geq 1/2$, the strategy profile (TFT,TFT) is a NE of the infinitely repeated PD game. Axelrod's full proof has been included in Appendix D.

Pavlov. (number 13 in Table C.1) This strategy has the bitwise representation of $(0|0, 1, 1, 0)$. It prescribes cooperation initially. Thereafter, if the payoff from the preceding round was 1 (punishment payoff) or 0 (sucker's payoff), it 'flips' its action. If the payoff from the preceding round was 2 (reward payoff) or 3 (temptation payoff), it prescribes the same action as the preceding round. It can be shown that Pavlov is strictly dominated by the All D strategy (using arguments similar to that of All D FC).

Grim Trigger (GT). (number 9 in Table C.1) This strategy has the bitwise representation of $(0|1, 1, 1, 0)$. It prescribes cooperation initially and continues to cooperate until itself or its opponent defects, after which defects. The Grim Trigger strategy imposes the most severe punishment available for the smallest departure from cooperation–namely a response of eternal defection. Using exactly the same arguments as for the strategy profile (TFT,TFT) (in Appendix D), it can be shown that the strategy profile (GT,GT) is a NE profile of the infinitely repeated PD game for $\delta \geq 1/2$.

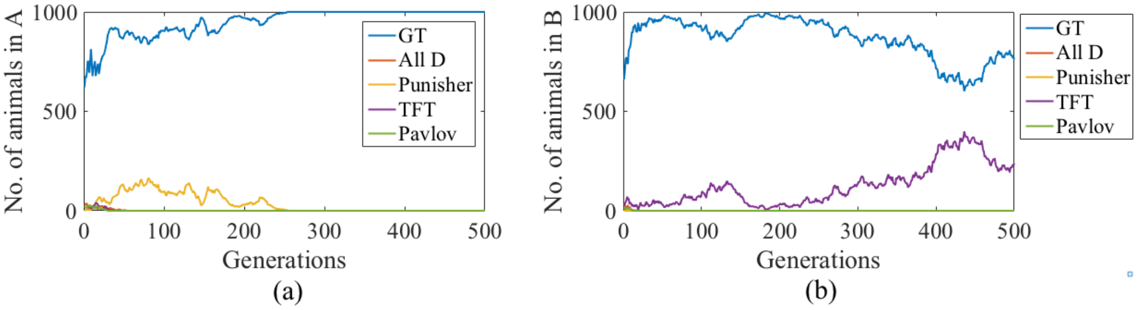
In summary, out of all the one-memory strategies we investigated in this section, we found that only the strategy profiles (All D, All D), (TFT, TFT) and (GT,GT) can be NE profile of the infinitely repeated PD game. Out of these (All D, All D) is a NE profile for all $\delta \in (0,1)$, while (TFT, TFT) and (GT, GT) are NE profiles if and only if $\delta \in [1/2, 1)$. If the populations A and B are composed entirely of All Ds, then all contests between animals have the outcome DD and the payoff for both animals in each contest is 1. In contrast, if the populations A and B are consist of TFT types only, then all contest outcomes are CC, and the payoff for each animal from a contest is 2. This is because all TFT type animals start by cooperating, and cooperate in response to the outcome CC of the preceding period.

Chapter 7

Numerical Experiments

In this section, we investigate one-memory strategies that support cooperative behaviour, using the framework of the PD game. We do so using numerical simulations. Note that one-memory strategies are not dependent on the total number of rounds of the stage game that have occurred. Since it is not possible to simulate a infinite PD game, we need to ensure that our simulations run for a large enough period such that the payoffs beyond that period are insignificant and can be ignored safely. If the discount factor is approximately 0.99, the payoffs at the 1000^{th} period are discounted by a factor of $\delta^{1000} \sim 10^{-5}$. Payoffs beyond this period are therefore insiginificant, and can be ignored. In our simulation, we will therefore use δ values less than 0.99 and the total number of periods of the stage game T equal to 1000. Note that a δ greater than 0.99 requires the use of a larger T, which significantly increases the computation time.

Note that for $\delta < 1/2$, the expected number of periods in the stage game is $1/(1-\delta) < 2$. In such a case, the value of the payoffs received beyond the first few payoffs is insignificant. Therefore, strategies that defect will flourish over the ones that cooperate in the first few periods. Even strategies like TFT and GT, which promote cooperation by 'threats' of punishment to opponents, that is, by defecting against opponents that defect, will not be able to sustain cooperation. This is reflected in the fact that (TFT,TFT) and (GT,GT) are not NE profiles if $\delta < 1/2$. The case of $\delta < 1/2$ is similar to playing the PD game once (or playing the infinitely repeated game with memory-less strategies), a scenario which was modelled in Chapter 4. It was observed in that chapter that cooperation in the scenario of a one-shot PD game cannot be sustained. Therefore, we only study the case when $\delta \in (1/2, 1)$ for the rest of this chapter.



(a) (b)

Figure 7.1: The first and the second sub-figure show the number of the peaceful strategies, namely GT, All D, Punisher, TFT and Pavlov in population A and B, respectively. In their initial states, both populations contained mostly GT type animals ($\approx 65\%$), and the rest of the population was randomly distributed over the other types. Simulation parameters: $\delta = 0.75$, $\epsilon = 0$ and $N_A = N_B = 100$.

7.0.1 Mutation rate $\epsilon = 0$

We conducted 10 simulations with both populations A and B composed mostly of GT types initially, while the rest of the population was uniformly randomly distributed over all of the 25 other types of one-memory strategies shown in Table C.1. The parameter values used were $\delta = 0.75$,and $\epsilon = 0$. Fig. 7.1 shows 1 instance of the numerical experiments[1]. In all of the 10 experiments that were done, we observed two kinds of behaviour:

1. All types other than GT became extinct in one or both of the populations within about 100 periods (see Fig. 7.1 (a)).

2. One or both of the population distributions fluctuated between the set of five strategies $\{$GT, All D, Punisher, TFT, Pavlov$\}$. We show the latter behaviour in population B in the Fig. 7.1 (b), in which the population is fluctuates between GT and TFT, while all other types are extinct.

We refer to the type which is present in a majority at the initial generation in a population as the *indigenous type*. For example, in Fig. 7.1, the indigenous type in both populations A and B was GT. To see why the population distributions evolve according to the two kinds of behaviours described above for large discount rate $\delta = 0.75$ and indigenous type GT, recall from our discussion in Chapter 6.2 that

---

[1] All simulations in this Chapter have been performed using code that was written by the author of this thesis, but is based on a much more extensive code written by Dr. Cameron Hall. In particular, the author of this thesis added the features that simulate the selection map $\mathbf{f}$ and the mutations.

(GT, GT) is an NE profile of the infinitely repeated PD game. This means that no other one-memory strategy can obtain a higher payoff against GT than GT does against itself. In a majority GT population, the probability of an animal being paired against a GT is high. Therefore, if GT is the indigenous type in both population A and B, then no other one-memory strategy can *invade* (that is, replaced, by virtue of obtaining a higher payoff) the GT types under the selection map $\mathbf{f}$.

However, there can exist other strategies that obtain exactly the same payoff against GT as GT does against itself. Note that in contest between two GTs, the sequence of outcomes is $CC, CC, CC, \ldots$. Therefore, for a strategy to obtain the same payoff against GT, it will need to start a contest by cooperating. In addition, it will need to cooperate if the previous outcome was CC. All strategies of the form $(0|\cdot,\cdot,\cdot,0)$ fit this criteria. From Table C.1, Appendix C, it can be confirmed that the set of one-memory strategies which fulfill this criteria, which we call the set of *peaceful strategies*, is $\{$GT, All C, Punisher, TFT and Pavlov$\}$. We will call a strategy *non peaceful* if it not in the set of peaceful strategies. Once the majority GT population has driven all strategies not in the peaceful set to extinction, the population of GTs can then be *infiltrated*, by one or more of the peaceful types. Infiltrated here means that a indigenous type is replaced by another strategy type that obtains exactly the same payoff as the indigenous type does against itself. Note that an invading strategy replaces an indigenous type due to natural selection. However, an infiltrating strategy is not favoured over the indigenous type by natural selection. The distribution of the population between the indigenous and infiltrating type(s) will thus change with chance. This phenomena, known as *genetic/neutral drift*, is a well known phenomena in the study of evolution [39, 40, 41]. In other words, neutral drift is the phenomena of infiltration of the indigenous population, by a type that earns the same payoff against the indigenous type as the indigenous type does against itself. Note that unlike natural selection, neutral drift does not have a definitive direction. However, unlike mutations, it is not completely random either, since it works only on a small subset of strategies.

Note that GT is not the only peaceful strategy that is part of an NE profile. (TFT,TFT), is a NE profile for $\delta \geq 1/2$ (see Chapter 6.2). We conducted more than 10 numerical simulations with $\delta = 0.75$, $\epsilon = 0$, and TFT as the indigenous type in both populations A and B. Similar to the behaviour of (GT, GT), we observed that (TFT, TFT) drove all non-peaceful strategies to extinction, and then the populations drifted between the different peaceful strategies. However, for simulations carried out with indigenous populations of All C, Pavlov or Punisher, it was observed on all occasions that the indigenous population was invaded by strategy types like All D, All D FC and Swindler. This is because none of the strategies All C, Pavlov, or Punisher is a Nash equilibria, as was shown in Chapter 6.2.

If the indigenous population is GT or TFT, the population eventually settles down to cooperation; that is, all animals cooperate with each other. It is therefore instructive to define the *mean payoff* at a generation for a population.

$$\text{Mean payoff in generation } n = \frac{\text{Sum of payoffs recieved by all animals over generation } n}{\text{Size of the population}} \times (1-\delta).$$

Note that if most of the animals in both the populations in a generation cooperate, that is, if most of the individual contest outcomes are CC, then the average payoff for both populations would be close to 2, and we say that the populations are in the *cooperative state*. Similarly, if most of the animals defect, and the individual outcomes are DD, then the average payoff for both populations will be close to 1, and we say that the populations are in the *defecting state*.

To summarise, for high discount rates (δ roughly equal to 0.75 and greater) and $\epsilon = 0$, our simulations show that if the indigenous type is GT or TFT, then all non-peaceful strategies are driven to extinction. The population distributions then fluctuate between the peaceful strategies. All animals in the population therefore cooperate with each other (all outcomes of individual contests are CC). If the discount factor is close to 1/2, or the indigenous type is a peaceful strategy other than TFT and GT, the indigenous population is immediately invaded by strategies like All D, All D FC (which we discuss in further detail in Chapter). In this case, all outcomes of individual contests were DD.

7.1 Mutation rate $\epsilon > 0$

We now investigate the behaviour of predominately GT (the results for TFT are similar) population in the presence of mutations. We investigated the system for extreme values of the parameters δ and ϵ to gauge their effect on the system. In more than 20 simulations that we conducted with $\delta = 0.95$ and $\epsilon = 0.001$, and GT as the indigenous population, we observed that eventually the entire population was composed of peaceful types, and the proportion of non-peaceful types was close to zero. Non-peaceful types were not extinct, though, because of mutations. But their proportion was small in the entire length of the simulation. These experiments

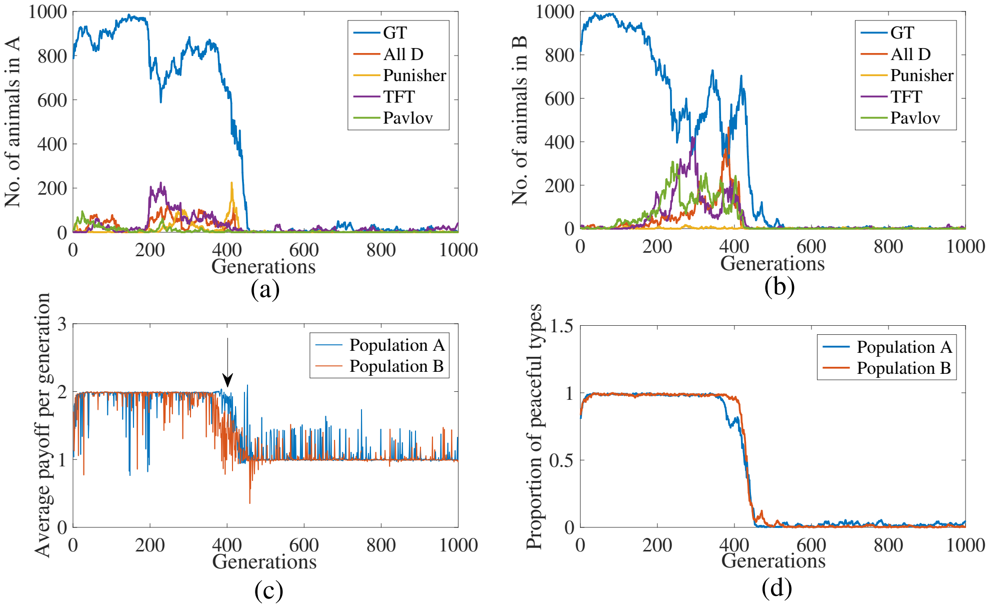

(a) (b) (c) (d)

Figure 7.2: Simulation with GT as the initial indigenous strategy ($\approx 75\%$), discounting-factor ($\delta = 0.75$), and non-zero mutation rate of ($\epsilon = 0.005$).

suggest that if δ is greater than 0.95 and mutation rate is smaller than 0.001, GT will always drive the populations A and B towards the peaceful strategies. The mean payoff for both populations would be equal close to 2 in this case, indicating that most outcomes of most individual contests were CC. We also carried out more than 20 simulations with small discount rates $\delta = 0.51$ and large mutation rates $\epsilon = 0.05$. In all simulations we observed that regardless of whether the indigenous type was GT or TFT, it was invaded by non-peaceful strategies. The average payoff in this case for both populations was therefore 1, indicating that most outcomes of most individual contests were DD.

We now investigate the system for values of parameters δ and ϵ in between these extreme values. Fig. 7.2 shows one instance of more than 20 numerical simulations that were done with the parameter values $\delta = 0.75$ and $\epsilon = 0.005$. Although not visible in Fig. 7.2, within the first 20 generations, the indigenous type GT drove all non-peaceful types to almost extinction (they can't be made completely extinct because of mutations). This is because the mean payoff, shown in Fig. 7.2 (c), rose to 2 within the first 20 generations, indicating that most of the population is composed of peaceful strategies. The number of GTs in the population then fluctuated between other peaceful strategies due to neutral drift (see Fig. 7.2 (a), (b)). One can see in Fig. 7.2 (d) that the average payoffs for the population up to generation 400 is 2, indicating cooperation (with some sparsely populated downward spikes due to mutations). This is due to the fact that both the populations are composed of almost entirely of peaceful strategies, even though the population distributions are constantly changing because of neutral drift.

Just before the 400th generation, there is a sharp increase in the number of All C types in population B (see Fig. 7.2, (b)). There is also a significant number of Pavlov types (about 20%) in population B at this point. We already noted in Chapter 6.2 that all peaceful strategies other than GT and TFT are strictly dominated by the All D strategy. Therefore, strategies like All D can take advantage of the strategies All C, Pavlov and Punisher. In Fig. 7.2 (c), we have highlighted (with an arrow) the generation at which the mean payoff for generation B begins to drop from its constant value of 2. Note that at this point, the mean payoff for population A is greater than that of B. This corresponds to the All D mutants in A taking advantage of Pavlov and All C types in population B. Therefore, natural selection allows non-peaceful mutants in A to increase their proportion in the population, which causes a sharp decline the proportion of peaceful types in A (see Fig. 7.2 (d)). In particular, the population is invaded by strategies that favour defection, which is evident from the fact that the average payoff for both populations is 1 after about the 450th generation. Once population A is composed primarily of types that favour defection, peaceful types in B are worse off as compared to non-peaceful types, and their numbers will decrease as well. We will discuss in further detail about which non-peaceful strategies are responsible for this change of state in Section 7.3.

Neutral drift clearly plays a vital role in describing the long-term behaviour of the system in Fig. 7.2. We conducted several simulations with parameters in the range $0.65 \leq \delta \leq 0.85$ and $0.001 \leq \epsilon \leq 0.01$, and using GT (similar simulations were performed for TFT) as the indigenous type. In all simulations, behaviour similar to the simulation shown in Fig. 7.2 was observed. We summarise this behaviour as follows: First, the indigenous type (GT or TFT) removes most of non-peaceful strategies from both the populations. After this, the population distribution of the different peaceful strategies fluctuates because of neutral drift, although the proportion of peaceful strategies in both populations A and B stays close to 1. The mean payoff of both populations stays constant at 2 (other than the occasional spikes caused by mutations). Once there is a large enough proportion of All C, Pavlov and Punisher in either population A or B, non-peaceful mutants (like All D) from the other population take advantage of these three peaceful strategies, and the proportion of peaceful strategies quickly reduces to 0, and the mean payoff reduces to 1.

7.2 Strategies responsible for establishing and destabilising cooperative populations

Prior to this section, we investigated the mechanism through which cooperation in populations is established by GT or TFT strategies. We also investigated the mechanism through which cooperation in populations was eventually destabilised. However, we have not yet investigated the role of individual strategies in the establishment and destabilisation of cooperation. For instance, it is not clear whether strategies other than GT or TFT can establish cooperation. We have also not investigated the individual strategies that invade the peaceful strategies All C, Pavlov, and Punisher.

We have already established that the cooperative populations, which can be sustained by Grim Trigger and TFT, and can be destabilised because of the entry of the other peaceful strategies (All C, Pavlov, and Punisher). To see if one-memory strategies other than TFT or GT could sustain a population with mean payoff 2, we performed simulations with parameters $\delta = 0.75$, $\epsilon = 0.005$, and using all one-memory strategies from Appendix C, except for TFT and GT. 24 such experiments were performed, with each one using a different indigenous strategy. In all such simulations we observed that no indigenous type was able to sustain the population in the cooperative state, and in each simulation the population settled into the defecting state within 50 generations. Five instances of such results are shown in Appendix E, Fig. E.1. These results suggest that only TFT and GT strategies are capable of sustaining populations in which the mean payoff is 2.

In order to investigate the importance of the strategies All C, Pavlov, and Punisher, in destabilising populations in the cooperative state, we performed more than 20 simulations in the absence of these three strategies, and with parameter values $\delta = 0.75$, $\epsilon = 0.005$ and indigenous type GT. Fig. E.2, Appendix E shows five instances of these simulations. In all simulations, we observed that the population never left the cooperative state during the length of the simulation (1000 generations). These simulations should be compared with Fig. 7.2, which we performed with exactly the same parameter values but with all of the one-memory strategies, including All C, Pavlov, and Punisher. It was because of these three strategies that the populations A and B left the cooperative state in Fig. 7.2. Simulations in E.2 provide further support for the suggestions of Chapter 7.1 that the primary cause of destabilisation of cooperative populations established by GT or TFT is infiltration by All C, Pavlov, and Punisher.

We noted in Chapter 7.1 that for simulations performed for parameter values

$0.065 \leq \epsilon \leq 0.085$, $0.001 \leq \epsilon \leq 0.01$, and indigenous type GT or TFT, the stability of a population with mean payoff 2 was upset within a 1000 generations. In all simulations, we observed that the strategies that were responsible for causing the populations to leave the cooperative state were one or more of the following three strategies: All D, All D FC and Swindler. These strategies invaded the cooperative populations when the proportion of All C, Pavlov, or Punisher was high enough. Note that only All D strategy out of three strategies All D, All D FC and Swindler, is a NE of the infinitely repeated PD game, as was discussed in Chapter 6.2. We performed more than 20 simulations with all one-memory one strategies of Table C.1 but without All D, All D FC and Swindler, to establish if there were any other non-peaceful strategies that could upset the cooperative populations established by TFT or GT. We used the parameter values $\delta = 0.75$, $\epsilon = 0.005$, and GT as the indigenous type (the same parameter values as 7.2, except now the simulations are performed without All D, All D FC and Swindler). We show three such instances of simulations in Fig. 7.3. In all such simulations, we observed that the cooperative state of populations was stable for most of the length of the simulation, with intermittent periods in which the cooperative state was temporarily destabilised. For example, these temporary deviations occur between the 100^{th} and the 300^{th} generation in the simulation shown in the top panel of Fig. 7.3, between the 450^{th} and 800^{th} generation in the middle panel, and between the 600^{th} and 850^{th} generation in the bottom panel. However, in contrast to the simulation in Fig. 7.2, the deviation from the cooperative state was not for long periods of time. In fact, in Fig. 7.2, in which All D, All D FC and Swindler were present, the mean payoff of the populations settled at 1 after the population left the cooperative state. In contrast, in Fig. 7.3, the mean payoff fluctuates between 1 and 2 when the populations are not in the cooperative state, but returns to its almost constant value of 2 once the cooperative state is re-established.

In the simulations of Fig. 7.3, the temporary deviations from the cooperative state occurred when the peaceful population neutrally drifted towards All C, Pavlov, or Punisher. These strategies were invaded by other non-peaceful strategies (other than All D, All D FC and Swindler, which have been included in these particular simulations), but the simulations suggest that none of these mutant non-peaceful strategies are capable of establishing the defecting state. The mean payoff therefore fluctuates between 1 and 2 for some time, but eventually returns to 2. This simulation suggest that out of the current set of strategies, only All D, All D FC and Swindler are capable of sustaining convergence to the defecting state.

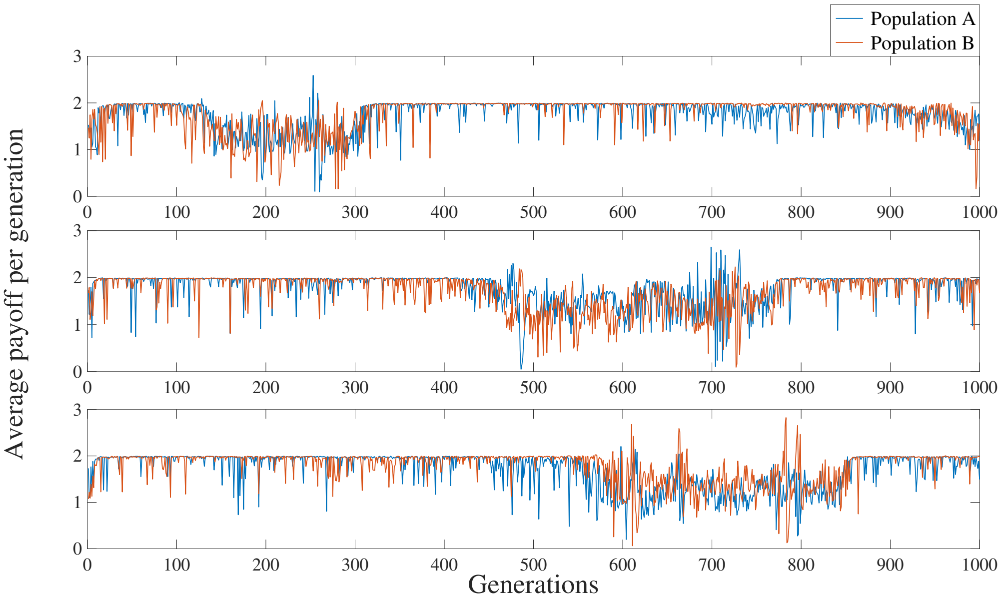


Figure 7.3: Simulations carried out with all one-memory strategies, excluding All D, All D FC and Swindler. Indigenous population was TFT ($\approx 65\%$), $\epsilon = 0.005$, $\delta = 0.75$.

7.3 Investigating why a population with mean payoff 1 is more stable than one with mean payoff 2

Note that in Chapter 7.1, we conducted more than 20 simulations with parameter values $0.65 \leq \delta \leq 0.85$, $0.001 \leq \epsilon \leq 0.01$, and indigenous type GT and TFT. In all such simulations, the stability of the cooperative state was upset within a 1000 generations. When this happened, the population was invaded by All D, All D FC and Swindler types, and the mean payoff for the population settled at the defecting state. However, the stability of the defecting state was not observed to be upset in any of the simulations. In addition, we carried out 5 simulations with the same parameter values $\delta = 0.75$ and $\epsilon = 0.005$ for 10^4 generations. But even over this longer simulation period, a population moving from the defecting state was never found to move to the cooperative state. This raises the question the following question: Why is the defecting state more stable than the cooperative state?

In particular, we know from Chapter 7.1 that neutral drift plays an important part in the destabilisation of the cooperative state. But neutral drift does not seem to affect the defecting state. For example, if the indigenous type in population A and B is All D, then strategy with the bitwise representation $(1|1, 0, 0, 0)$ can infiltrate the All D population. This is because this infiltrating strategy starts by defecting and plays defect in response to the outcome DD of the preceding period. Therefore, it receives the same payoff against All D as All D does against itself. Once this strategy infiltrates All D indigenous population, it can be invaded by strategies like GT. This is because in a contest of GT vs. $(1|1,0,0,0,0)$, the payoff for GT is $(3\delta)/(1-\delta^2)$, while that of $(1|1,0,0,0)$ is 1. Since $(3\delta)/(1-\delta^2) > 1$ for all $\delta \in (1/2, 1)$, the type $(1|1,0,0,0)$ will be invaded by GT, which might lead to the establishment of the cooperative state. However, such a scenario was never observed in the simulations performed in Chapter 7.1.

To investigate this, we conducted more than 20 simulations with parameter values $\delta = 0.75$ and $\epsilon = 0.005$, and with all one-memory strategies expect All D FC and Swindler. In Appendix E, Fig. E.3, we show 5 instances of these simulations. In Fig. 7.4, we show the same simulation presented in the top panel of Fig. E.3. Note that the population in Fig. 7.4 started in the cooperative state, the state settled in the defecting state at roughly the 250^{th} period, and the cooperative state was restored at roughly the 300^{th} period. This phenomena, of the state of the population switching between the cooperative and defective state was observed on several of the 20 simulations. These simulations seem to suggest that stability of the cooperative state is enhanced if All D FC and Swindler are not included in the simulation. In fact, the cooperative state was found to be re-established after the population was in the defecting state between roughly the 250^{th} and the 300^{th} period. This phenomena was not observed for any of the simulations with parameter values $0.065 \leq \epsilon \leq 0.085$, $0.001 \leq \epsilon \leq 0.01$ and including the All D FC and Swindler strategies.

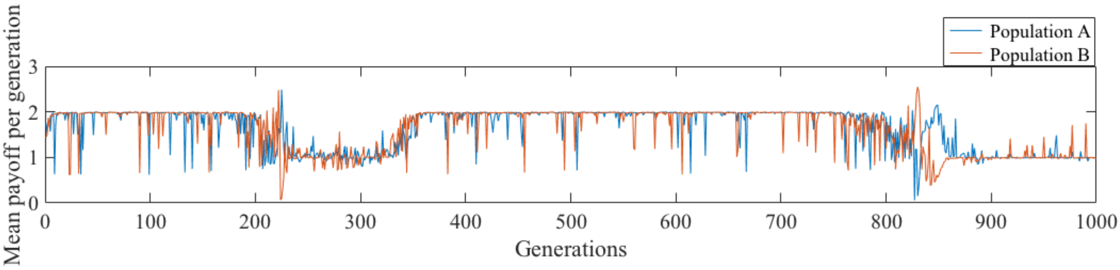


Figure 7.4: In this figure, we display an instance of more than 20 simulations that were performed with all the one-memory strategies of Appendix C, Table C.1 except All D FC and Swindler. Indigenous population was GT ($\approx 65\%$), $\epsilon = 0.005$, $\delta = 0.75$.

We performed simulations similar to Fig. 7.4 by excluding All D strategy instead of All D and Swindler (parameters were $\epsilon = 0.005$, $\delta = 0.75$ and indigenous type GT). In all of the 20 simulations that were done (five instances are shown in E.4 in Appendix E), behaviour similar to the one seen in Fig. 7.4 was observed. That is, the state of populations switched between the cooperative and defect state.

Our simulations (performed with parameters $\epsilon = 0.005$, $\delta = 0.75$) suggest that the only strategies that can establish a population in which the mean payoff is 1 are All D, All D FC and Swindler. In addition, we observed that if all of these states are present, then a population in the defecting state never moves to a cooperative state. However, if All D is not present, or alternatively if All D FC and Swindler are not present, we observed that the population switched between the cooperative and defecting states.

One possible explanation for this behaviour is as follows: Assume initially that the indigenous type in population A and B is All D. We discussed in Chapter 6.2 that (All D, All D) is a NE profile, and hence it cannot be invaded by any other strategy. Note however, that any strategies with the bitwise representation $(1|1, \cdot, \cdot, \cdot)$ can infiltrate a primarily All D population. This is because in a contest against All D, these strategies defect initially and in response to the outcomes DD from the preceding periods. They receive the same payoff as All D against itself, and therefore can infiltrate All D. All possible one-memory strategies that fit this criteria are $(1|1, 0, 0, 0)$, $(1|1, 0, 1, 0)$, $(1|1, 0, 0, 1)$ and $(1|1, 0, 1, 1)$. Each of these strategies is strictly dominated by All D FC or Swindler. For example, consider a contest of All D FC against $(1|1, 0, 0, 0)$. The payoff for All D FC against $(1|1, 0, 0, 0)$ is $(3\delta)/(1-\delta^2)$, while the payoff of $(1|1, 0, 0, 0)$ against itself is $1/(1 - \delta)$. Since $(3\delta)/(1 - \delta^2) \geq 1/(1 - \delta)$ for all $\delta \in (1/2, 1)$, we conclude that All D FC strictly dominates $(1|1, 0, 0, 0)$. Using similar arguments, it can be shown that all infiltrators of All D are strictly dominated by All D FC or Swindler. Therefore, if a population comprising of primarily All D types is infiltrated, the proportion of strategies All D FC and Swindler in the population will increase.

The only strategy that strictly dominates All D FC is All D. This is because All D FC defects in all periods after the initial one. Therefore, for any strategy that strictly dominates it will have to defect in all periods, including the initial period. Therefore, only All D can invade All D FC. Similarly, it can be shown that the only strategies that strictly dominate Swindler are All D FC and All D. Now consider the strategies that can potentially infiltrate All D FC. For them to receive the same payoff as All D FC does against itself, their bitwise representation will be of the form $(0|1, \cdot, \cdot, 1)$. All such strategies are strictly dominated by All D. Similarly, it can be shown that all strategies that have the potential to infiltrate Swindler are strictly dominated by All D. In other words, if an indigenous All D FC or Swindler population is infiltrated, All D types will flourish, and their proportion in the population will increase.

In summary, our discussion above shows that neutral drift cannot upset a population in the defecting state. This is because such a population comprises of the three types: All D, All D FC, and Swindler. If initially the indigenous type is All D, then it cannot be invaded by any other strategy. If it is infiltrated by any other strategy, All D FC and Swindler will ensure the infiltrators do not thrive, as they strictly dominate all potential infiltrators of the All D indigenous population. The All D FC (or Swindler) indigenous population can only be invaded by All D. Also, if All D FC (or Swindler) populations is infiltrated, the infiltrators will be removed by All D.

7.4 Discussion

In this chapter, we investigated the establishment and stability of cooperative and defecting states in populations A and B. Our numerical simulations suggest that the only one-memory strategies that can establish and sustain cooperative populations are GT and TFT, provided the mutation rate is low enough ($\epsilon < 0.01$) and the discount factor is high enough ($\delta > 0.7$). The simulations also suggest that the defecting state can only be established and sustained by strategies All D, All D FC, and Swindler. In particular, our numerical simulations show that for parameter values in the range $0.001 < \epsilon < 0.01$, $0.65 < \delta < 0.85$, neutral drift caused the populations to move from the cooperative state to the defecting state. In contrast, we never observed populations in the defecting state moving to the cooperative state (in simulations that used the full set of 26 one-memory strategies).

One possible reason behind this could be that the strategies All D, All D FC and Swindler (which are responsible for the defecting state) ‘coexist’ with each other. That is, if one of these three strategies is in the majority at a particular generation, the other two types will not allow infiltrators to creep into the population. In contrast, for the cooperative state, which is established by TFT or GT, no strategy exists that can enforce the cooperative state (that is, play C initially and in response to CC, so that the conflict outcomes are CC) and at the same time punish the infiltrators in TFT and GT majority populations.

Chapter 8

Conclusion

In this thesis, we develop a model to investigate the evolution of dominance hierarchies and cooperation under the evolutionary forces of natural selection and mutation. We study two distinct populations, and our model takes into account the asymmetries that might exist between them.

In conflicts in which animals in both populations incur a net loss from fighting, our analysis shows that the dominance hierarchy that is established in the long run depends on the mutation rate, the speed of adjustment of the population to natural selection, and the level of asymmetry in the payoffs. In particular, animals favoured by the asymmetry can under some conditions be the subordinates in the dominance hierarchy, especially when the mutation rates are low and the level of asymmetry small.

In order to gain insight into the evolution of cooperation, we generalise our model to incorporate one-memory strategies. We study the conditions under which cooperative behaviour in the populations can be established and sustained. Our analysis, based on numerical simulations, shows that neutral drift is a key factor in determining the long term behaviour of the populations, along with natural selection and mutation. Our analysis also shows that in the case of one-memory strategies, the defecting state is more likely to be observed in the long run (provide the discounting factor is small enough), as it is capable of resisting invasion by other strategies due to neutral drift.

There are several possible extensions to the work done in this thesis. For example, in the study of dominance hierarchies we assumed that the population to which an animal belongs is fixed, which is suitable for conflicts between predator–prey, male–female. However, there exist conflicts in the animal kingdom in which the roles of the animals depend on the outcome of the conflicts. Conflicts over territories is one such example. Another extension of the model is the incorporation of longer memory strategies. Including longer memory strategies could change the stability of the cooperative state dramatically. This is because there might exist longer memory strategies that co-exist with GT/TFT and prevent infiltrators to creep into predominantly GT/TFT populations. The effect of including memory strategies on the evolution of dominance hierarchies would also be an interesting extension of our work.

Appendices

Appendix A

Mixed Strategy Nash Equilibria

The proof of the following proposition was given by Shoham and Brown [44]. We have included the proof here for the convenience of the reader.

Proposition A.0.1. *For an action profile* $\boldsymbol{\sigma}^*$*, define* $S_i^* := \{s_i \in S_i \mid \sigma_i^*(s_i) > 0\}$ *as the set of pure strategies that player* i *plays with positive probability according to* $\boldsymbol{\sigma}^*$*. Then,* $\boldsymbol{\sigma}^*$ *is a Nash equilibrium if and if only for all* $i \in I$

1. $u_i(s_i, \boldsymbol{\sigma}^*_{-1}) = u_i(s_i', \boldsymbol{\sigma}^*_{-i})$ *for all* $s_i, s_i' \in S_i^*$*;*
2. $u_i(s_i, \boldsymbol{\sigma}^*_{-i}) \geq u_i(s_i', \boldsymbol{\sigma}^*_{-i})$ *for all* $s_i \in S_i^*$ *and* $s_i' \in S_i$*.*

Proof. **Necessary condition.** Assume that at least one of the above conditions fails to hold. In such a case, there exist strategies $s_i' \in S_i$ and $s_i \in S_i^*$ such that $u_i(s_i', \boldsymbol{\sigma}^*_{-i}) > u_i(s_i, \boldsymbol{\sigma}^*_{-i})$. We now construct a mixed action $\boldsymbol{\sigma}$ by setting $\sigma_i(\tilde{s}_i) = \sigma_i^*(\tilde{s}_i)$ for all $\tilde{s}_i \notin \{s_i', s_i\}$ and $\sigma_i(s_i') = \sigma_i^*(s_i) + \sigma_i^*(s_i')$ and $\sigma_i(s_i) = 0$. From the formula (2.1), it is evident that $u_i(\sigma_i, \boldsymbol{\sigma}^*_{-i}) > u_i(\sigma_i^*, \boldsymbol{\sigma}^*_{-i})$. Hence, σ_i^* is not a best response to $\boldsymbol{\sigma}^*_{-1}$ and is consequently, $\boldsymbol{\sigma}^*$ is not a Nash equilibrium.

Sufficient condition. Assume that both the conditions hold true for all players. For player i, the second condition ensures that any mixed action other than $\boldsymbol{\sigma}^*$ that places a positive probability on pure strategies not in S_i^* cannot be a better response to $\boldsymbol{\sigma}^*_{-i}$ than σ_i^*. In addition, the first condition ensures that any mixed action that defines a different distribution on the elements of S_i^* cannot do strictly better than $\boldsymbol{\sigma}^*$. Therefore if both conditions are satisfied, there is no action for player i that does strictly better against $\boldsymbol{\sigma}^*_{-i}$ than σ_i^*. This reasoning is true for all players $i \in I$. Therefore, $\boldsymbol{\sigma}^*$ is a Nash equilibrium. ☐

Appendix B

Unique pure strategy Nash equilibria.

Let's first study GHD game with $P_A \geq P_B > S$, in this case, the the strategy Hawk strictly dominates the Dove strategy (see Definition 2.1.4) for both populations A and B. Hawks in population A (or B) will get a better reward as compared to the Doves, regardless of the population composition of animals in population B (or A). Note that the only Nash equilibrium of the stage game is (Hawk, Hawk). Assume that in the initial state, both populations consist of a non-zero number of the type Hawk. In the absence of mutations ($\epsilon = 0$), the proportion of Hawks in both the populations will strictly increase over the generations, by definition of the selection map $\mathbf{f}$, until the state $s^* = (N_A, N_B)$ is reached. A population in state s^* will remain in state s^*, in the absence of mutations. The state s^* is therefore a fixed point of the selection map ($\mathbf{f}(s^*) = s^*$). There are three other fixed points of $\mathbf{f}$; namely $(0, 0)$, $(0, N_B)$ and $(N_A, 0)$. We shall refer to these absorbing states as 'trivial', because they are absorbing states by virtue of our choice of the selection map $\mathbf{f}$, and not a defining characteristic of the payoff matrices under consideration.

Since Hawks do strictly better than Doves, we expect s^* to be robust against mutations in the long-run. This is confirmed by numerical simulations conducted using the map $\mathbf{f}$. It was observed that for a 100×100 lattice, in which the $(i, j)^{\text{th}}$ term is the initial state $(iN_A/100, jN_B/100)$, every initial state converged to the fixed point (N_A, N_B) in the presence of mutations ($\epsilon = 0.01$), in a maximum of $\mathcal{O}(10)$ generations. Fig. B.1 shows two particular instances with different initial states. In summary, for the GHD game with $P_A \geq P_B > S$, the populations will spend all their times in the state s^* in the long run; that is, Doves in both populations A and B will

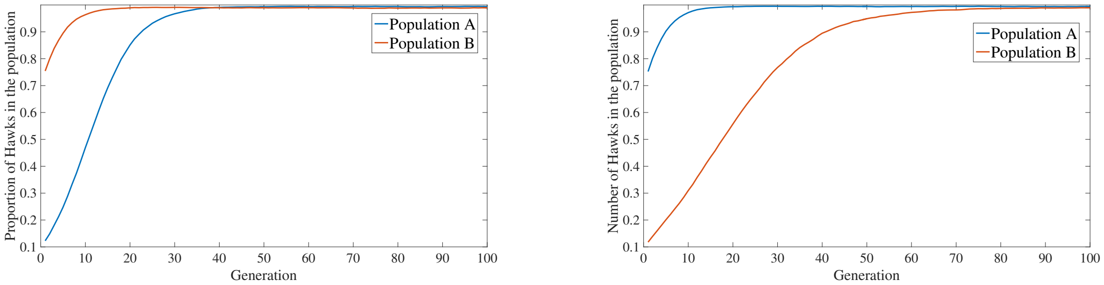


Figure B.1: Evolution of the populations A and B under the replicator style dynamics (Case (a)). Simulation parameters: $T = 1$, $R = 0.5$, $S = 0$, $P_A = 0.2$, $P_B = 0.1$, $\pi_0 = 1$, $\epsilon = 10^{-3}$, $N_A = N_B = 10^5$.

become extinct (eventually).

In the case (b) $P_A \geq S > P_B$, the strategy Hawk is strictly dominant for animals in role A, whereas no such strategy exists for animals in role B. The only Nash equilibrium of the stage game is therefore (Hawk, Dove). Similar to Case (a), it can be argued that the proportion of Hawks in A will increase, regardless of the composition of population B, and regardless of the mutation rate. Note that the strategy Dove is the best response by animals in role B to Hawks in role A. Therefore, once the population A is predominately composed of Hawks (which it eventually will be), Doves in B will do better as compared to Hawks. Thus, the only non-trivial fixed point of the dynamics in the absence of mutations is $(N_A, 0)$, and it is the only state that is robust against mutations. In simulation experiments similar to that of Case (a), it was found that for all initial states on the 100×100 lattice, in which the $(i, j)^{\text{th}}$ term is the initial state $(iN_A/100, jN_B/100)$, and it was observed that the system converged to the state $(N_A, 0)$ in the presence of mutations. Fig. B.2 shows the extreme case in which initially, the population of A consists of only Doves, while the population B is consists of only Hawks.

It is interesting to note that the solution obtained in our evolutionary model, for Case (a) and Case (b), is exactly the same as that predicted by CGT. Recall that in CGT, all players are assumed to be hyper-rational; that is, they are rational (always play strategies that maximize their payoffs), and know that the others are rational, and that the others know that everybody knows that the others are rational, *ad infinitum*. If player A and B were hyper-rational, they would not play the Dove strategy in Case (a), since it is strictly dominated by Hawk. In Case (b), A will not play strategy Dove (because it is strictly dominated), and will know that B will play Dove (since A knows that B knows that A is rational, and hence will know that A

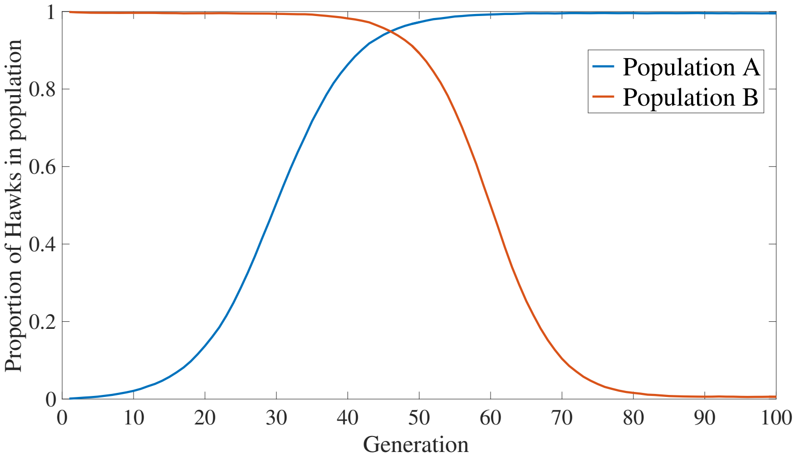


Figure B.2: Evolution of the populations A and B under the replicator style dynamics (Case (b)). Simulation parameters: $T = 1$, $R = 0.5$, $S = 0$, $P_A = 0.2$, $P_B = -0.2$, $\pi_0 = 1$, $\epsilon = 10^{-3}$, $N_A = N_B = 10^5$.

will play Hawk).

Appendix C

Table of all possible one-memory strategies

Strategy	Representation	Name
1	$(0\|0,0,0,0)$	All C
2	$(1\|0,0,0,0)$	All C FD
3	$(1\|1,0,0,0)$	Blinder
4	$(1\|0,1,0,0)$	Prodigal Son
5	$(0\|0,0,1,0)$	Punisher
6	$(1\|0,0,1,0)$	Punisher FD
7	$(0\|0,0,0,1)$	Rebel
8	$(1\|0,0,0,1)$	Rebel FD
9	$(0\|1,0,1,0)$	Tit-for-tat
10	$(1\|1,0,1,0)$	Tit-for-tat FD
11	$(0\|1,0,0,1)$	
12	$(1\|1,0,0,1)$	
13	$(0\|0,1,1,0)$	Pavlov
14	$(1\|0,1,1,0)$	Pavlov FD
15	$(0\|0,1,0,1)$	Tree frog
16	$(1\|0,1,0,1)$	Tree frog FD
17	$(0\|0,0,1,1)$	CD repeater
18	$(1\|0,0,1,1)$	DC repeater
19	$(0\|1,1,1,0)$	Grim Trigger
20	$(0\|1,1,0,1)$	
21	$(0\|1,0,1,1)$	Swindler
22	$(1\|1,0,1,1)$	Swindler FD
23	$(0\|0,1,1,1)$	
24	$(1\|0,1,1,1)$	
25	$(0\|1,1,1,1)$	All D FC
26	$(1\|1,1,1,1)$	All D

Figure C.1: Table of all possible distinct one-memory strategies. The representation $\mathbf{s} = (s_0|s_1, s_2, s_3, s_4)$ gives the probability of playing Hawk in response to the following outcomes, according to order: (Initial action | DD, DC, CD, CC). Note that DC denotes the outcome in which the player of type $\mathbf{s}$ defected in the last period and its opponent cooperated.

Appendix D

Proving that (TFT,TFT) and (GT,GT) are NE profiles

Consider the PD game with the following bi-matrix:

		Animal B	
		Defect (D)	Cooperate (C)
Animal A	Defect (D)	$1, 1$	$3, 0$
	Cooperate (C)	$0, 3$	$2, 2$

Let us investigate properties of some interesting strategies of the infinitely repeated PD game.

TFT

We have already noted that the strategy TFT (number 9 in Table C.1) is known to perform well against a range of other strategies. TFT is a strategy that prescribes cooperation in the first period, and replication of the opponent's behaviour (from the previous period) thereafter. Axelrod [65] proved that (TFT,TFT) is a Nash equilibrium profile for $\delta \geq 1/2$. For the convenience of the reader, we include his proof here. First note that the only alternative strategies we need to consider are $CCCC\ldots$; $DDDD\ldots$; and $DCDC\ldots$, because no other strategy can do better than all of these. To see this, note that TFT has a memory of only one game, and that the expected number of further games at any time is constant. If s is a strategy played against TFT, a play of C by s will reset the match to exactly the state it was in initially. Also, if the first play by s is D, a play of D at any later point will reset the match to its initial state. Note next that if s is the best reply to TFT, then it must make the same play as it did on the first move if the initial state ever recurs; if there were a better play in that state, it should have been played initially. It then follows that a best reply must have one of these forms:

- First play C; the initial state is then repeated on the second game, and C must be played again, and so on. That is, play $CCCC\ldots$, which is essentially the All C strategy. The payoff of a TFT against another TFT is $2+2\delta+2\delta^2+\cdots = 2/(1-\delta)$. The payoff of $CCCC\ldots$ against TFT would be the same, and hence $CCCC\ldots$ cannot invade (TFT,TFT).
- First play D, and then C; the initial state is then repeated on the third game, and D must be played again, and so on. That is, play $DCDCDC\ldots$. The payoff to $DCDCDC\ldots$ against TFT is $3+3\delta^2+3\delta^4+\cdots = (3)/(1-\delta^2)$. For the payoff for a TFT against another TFT to be greater than the payoff for $DCDCDC\cdots$, we therefore require
$$\delta \geq \frac{1}{2}. \tag{D.1}$$
- First play D, and then D; the initial state is repeated on the third game, and D must be played. That is, play $DDDD\cdots$. The payoff for $DDDD\cdots$ against TFT is $3+\delta+\delta^2+\cdots = 3+\delta 1/(1-\delta^2)$. For the payoff for a TFT against another TFT to be greater than the payoff for $DDDD\cdots$, we therefore require
$$\delta \geq \frac{1}{2}. \tag{D.2}$$

The strategy profile (TFT,TFT) is therefore a Nash equilibrium of the infinitely repeated PD game if and only if

$$\delta \geq \frac{1}{2}.$$

Grim Trigger

Another strategy that can sustain cooperation is the Grim Trigger (GT) strategy (strategy 19 in Appendix C), in which the animal starts out by cooperating, and continues to cooperate until itself or its opponent defects; after which it never cooperates. The Grim Trigger strategy imposes the most severe punishment available for the smallest departure from cooperation, namely a response of eternal defection. Similar to the strategy profile (TFT,TFT), it can be shown, by following the recipe discussed in Section 5, that the strategy profile (GT,GT) is a NE profile if and only if

$$\delta \geq \frac{1}{2}. \tag{D.3}$$

Pavlov Pavlov (strategy 13 in Appendix C) cooperates initially and then if the payoff from the last round was R (reward) or T (temptation) payoff was received in the last round, it repeats the last choice. Otherwise, it chooses the opposite choice. Consider the strategy All D against Pavlov. The payoff for All D in such a contest would be $3 + \delta + 3\delta^2 + \delta^3 + \cdots = (3 + \delta)/(1 - \delta^2)$. Since the payoff for a Pavlov against another Pavlov is $2 + 2\delta + 2\delta^2 + \cdots = 2/(1 - \delta)$, the profile (Pavlov, Pavlov) is a Nash equilibrium of the PD game if and only if

$$\frac{3 + \delta}{1 - \delta^2} \leq \frac{2}{1 - \delta} \quad \Longrightarrow \quad \delta \geq 1.$$

But since the discounting factor is always less than 1, we concluded that the profile (Pavlov, Pavlov) is not a Nash equilibrium profile of the PD game.

Appendix E

Some figures showing simulation experiments discussed in Chapter 7

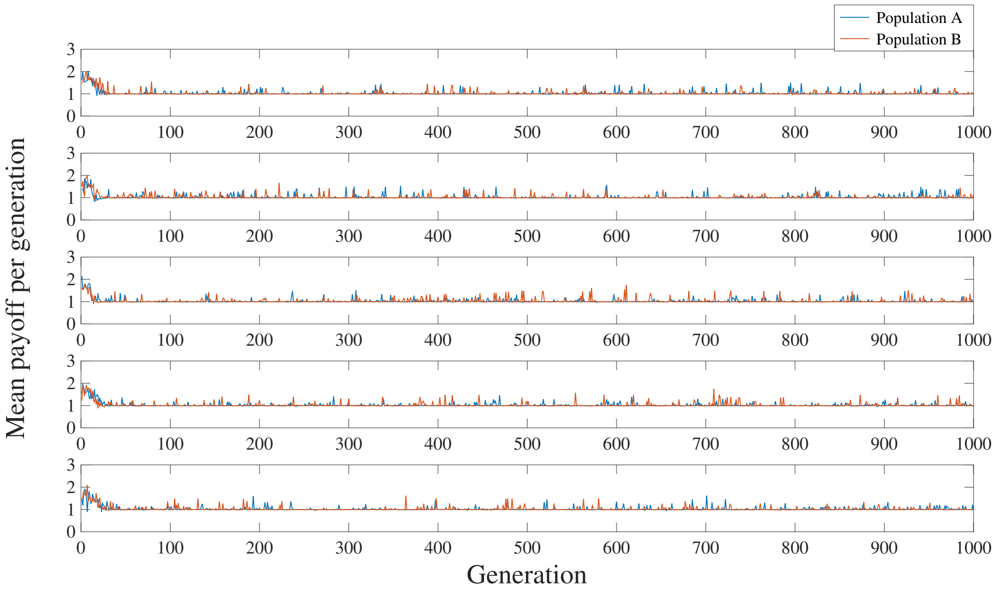


Figure E.1: In this figure, we display five instances 24 simulations that were performed with all the one-memory strategies of Appendix C, Table C.1 except TFT and GT. Simulation parameters used were $\epsilon = 0.005$ and $\delta = 0.75$. Indigenous type used for the five instances are as follows (from the top to the bottom panel): All C, Punisher, Pavlov, Pavlov FD (strategy number 6 in Appendix C, Table. C.1) and TFT FD (strategy number 10 in Appendix C, Table. C.1).

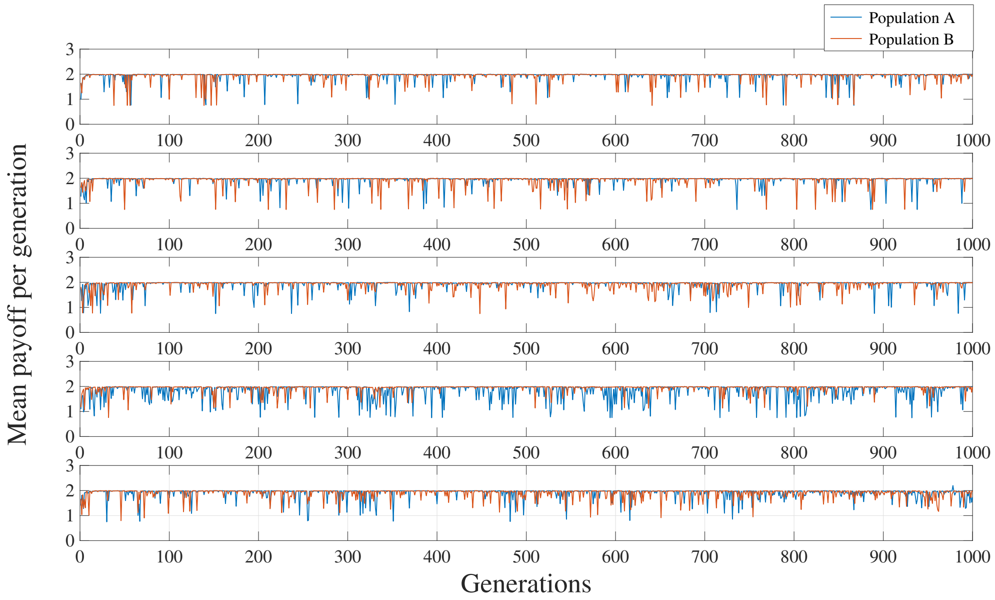


Figure E.2: In this figure, we display five instances of more than 20 simulations that were performed with all the one-memory strategies of Appendix C, Table C.1 except All C, Pavlov, and Punisher. Note that all three of the eliminated strategies belong to the peaceful set. Indigenous population was GT ($\approx 65\%$), $\epsilon = 0.005$, $\delta = 0.75$.

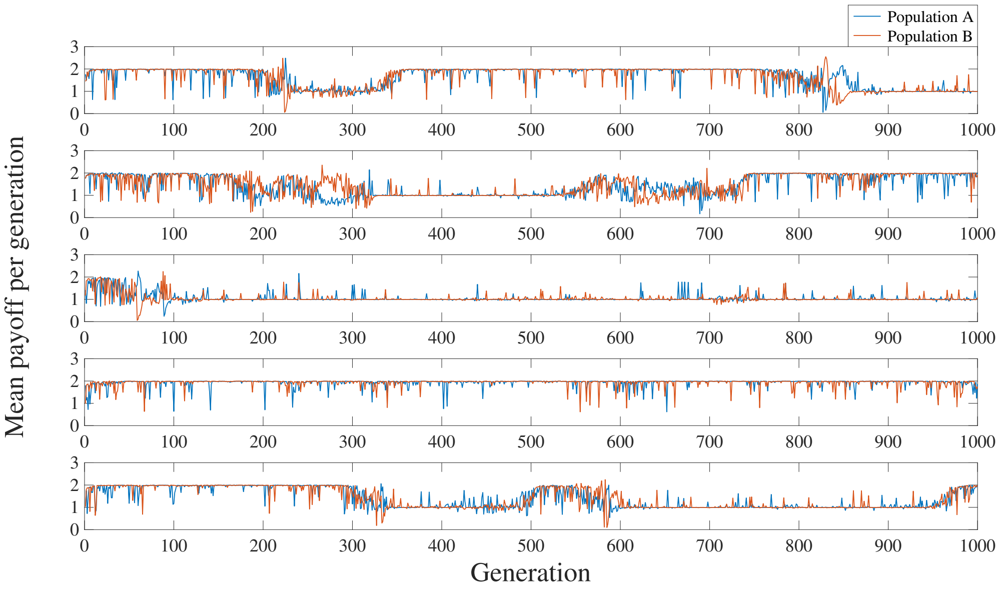


Figure E.3: In this figure, we display five instances of more than 20 simulations that were performed with all the one-memory strategies of Appendix C, Table C.1 except All D FC and Swindler. Indigenous population was GT ($\approx 65\%$), $\epsilon = 0.005$, $\delta = 0.75$.

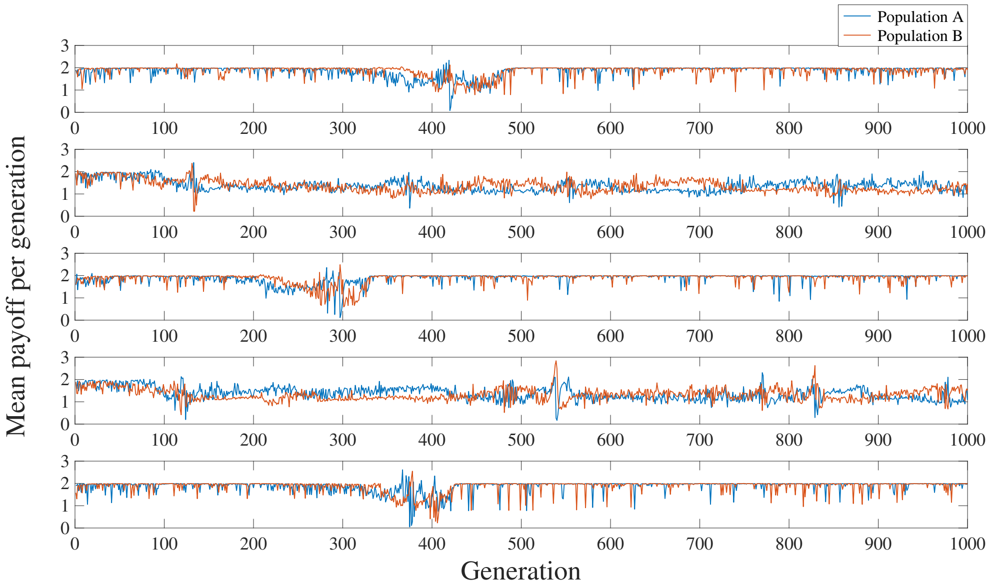


Figure E.4: In this figure, we display five instances of more than 20 simulations that were performed with all the one-memory strategies of Appendix C, Table C.1 except All D. Indigenous population was GT ($\approx 65\%$), $\epsilon = 0.005$, $\delta = 0.75$.

Bibliography

[1] Barnard, C. and Burk, T. *Dominance hierarchies and the evolution of individual recognition.* Journal of Theoretical Biology, Vol. 81, pp. 65–73 (1979).

[2] Broom, M., Koenig, A., and Bornes, C. *Variation in dominance hierarchies among group-living animals: modelling stability and the likelihood of coalitions.* Behavioural Ecology, Vol. 20, pp. 844–855 (2009).

[3] Rushen, G. *The peck orders of chickens: How do they develop and why are they linear?* Animal Behaviour, Vol, pp. 1129–1137 (1982).

[4] Arora, M.H. *Animal Behaviour.* Himalaya Publishing House (2008).

[5] French, A.R. and Smith, T.B. *Importance of Body Size in Determining Dominance Hierarchies among Diverse Tropical Frugivores.* Biotropica, Vol. 37, No. 1 , pp. 96-101 (2005).

[6] Clutton-Brock, T.H., Albon S.D. and Guiness, E.E. *Maternal dominance, breeding success and birth sex ratios in red deer.* Nature 308: 358 (1984).

[7] Savolainen, R. and Vepsalainen, K. *Niche differentiation of ant species within territories of the wood ant Formica polyctena.* Oikos 56: 3-16 (1989).

[8] Daily, G.C. and Ehrlich, P.R. *Influence of social status on individual foraging and community structure in a bird guild.* Oecologia (Berlin) 100: 153-165 (1994).

[9] Davies, N.B., Krebs, J.R. and West, S.A. *An introduction to Behavioural Ecology* Fourth Edition. Blackwell Science Ltd. (2012).

[10] Huntingford, F. and Turner, A. *Animal Conflict.* University Press, Cambridge (1987).

[11] Smith, J.M. *Evolution and the Theory of Games.* Cambridge University Press, Cambridge, UK, 1982.

[12] Clutton-Brock, T.H. and Parker, G.A. *Potential reproductive rates and the operation of sexual selection.* Quarterly Review of Biology , 67 , 437–455 (1992).

[13] Dawkins, M.S. *Do asymmetries destabilize the Prisoner's Dilemma and make reciprocal altruism likely?* Animal Behaviour, Vol. 80, pp. 339–341 (2010).

[14] Lehman, L. and Keller, L. *The evolution of cooperation and altruism: a general framework and classification of models.* Journal of Evolutionary Biology, Vol. 71, pp. 556–584 (2009).

[15] Nowak, M.A. *Five rules for the evolution of cooperation.* Science, Vol. 314, pp. 1560–1565 (2006).

[16] Axelrod, R. *The evolution of Cooperation.* Basic Books, Inc., Publishers (1984).

[17] Brown, J.S., Sanderson, M.J. and Michod, R.E. *Evolution of social behavior by reciprocation.* J. theor. Biol., 99 (1982), p. 319

[18] Maynard Smith, J. and Price, G.R. *The logic of animal conflict.* Nature 246 15–18 (1973).

[19] Lawlor, L.R. and Maynard Smith, J. *The coevolution and stability of competing species.* Am. Nat. 110, 79–99 (1976).

[20] Hamilton, W.D. and May, R.M. *Dispersal in stable habitats.* Nature, Lond. 269, 578–581 (1977).

[21] Mirmirani, M. and Oster, G. *Competition, kin selection and evolutionary stable strategies.* Theor. Pop. Biol. 13, 304–309.

[22] Charnov, E. *The Theory of Sex Allocation.* Princeton University Press (2009).

[23] Frey, E. *Evolutionary game theory: Theoretical concepts and applications to microbial communities.* Physica A 389 4265–4298 (2010).

[24] Friedman, D. *On economic applications of evolutionary game theory.* J Evol Econ 8: 15–43 (1998).

[25] Issa, F.A., Adamson, D.J. and Edwards, D.H. *Dominance hierarchy formation in juvenile crayfish procambarus clarkii.* Journal of Experimental Biology, Vol. 202(24), pp. 3497–3506 (1999).

[26] Guhl, A.M. *The social order of chickens.* Scientific American, Vol. 194(2), pp. 42–46 (1956).

[27] Brittany M.T., Jessie L.K. and Wendy A.K. *Interspecific Competition and Social Hierarchies in Frugivorous Neotropical Birds of Costa Rica.* Journal of Young Investigators, Vol. 29, Issue 2 (2015).

[28] Le Boeuf, B.J. *Male–male competition and reproductive success in elephant seals.* American Zoologist, 14: 163–176 (1974).

[29] Broom, M. *A unified model of dominance hierarchy formation and maintenance.* Journal of Theoretical Biology, Vol. 219, pp. 63-72 (2002).

[30] Pagel, M. and Dawkins, M.S. *Peck orders and group size in laying hens: futures contracts for non-aggression.* Behavioural Processes, Vol. 40, pp. 13–25 (1997).

[31] Kalmus, H. *Defence of source of food by bees.* Nature, London 148, 228 (1941).

[32] Bell, G. *Selection: The Mechanism of Evolution.* Oxford University Press (2008).

[33] Wingreen, N.S. and Levin, S.A. *Cooperation among Micro-organisms.* PLoS Biol. 2006 Sep; 4(9): e299 (2006).

[34] Trivers, R.L. *The evolution of reciprocal altruism.* Q Rev Biol 46:35–57 (1971).

[35] Axelrod R. and Hamilton W.D. *The evolution of cooperation.* Science 211:1390–1396 (1981).

[36] Hughes, M. *Size assessment via a signal in snapping shrimps.* Behavioural Ecology and Sociobiology, Vol 38, pp. 51–57 (1996).

[37] Wells, M.S. *Effects of body size and resource value on fighting behaviour in a jumping spider.* Animal Behaviour, Vol. 36, pp. 321–326 (1988).

[38] Dawkins, R. *The Selfish Gene* (2^{th} edition). Oxford University Press (1989).

[39] Lande, R. *Natural selection and random genetic drift in phenotypic evolution.* Evolution : 314-334 (1976).

[40] Hartl, D.L., Clark, A.G., and Clark, A.G. *Principles of population genetics.* Vol. 116. Sunderland: Sinauer associates (1997).

[41] Crow, J.F. *Wright and Fisher on inbreeding and random drift.* Genetics 184.3 : 609-611 (2010).

[42] Von Neumann, J. and Morgenstern, O. *Theory of Games and Economic Behaviour.* Princeton University Press (1953).

[43] Fujiwara-Greve, T. *Non-Cooperative Game Theory.* Springer, 2015.

[44] Shoham, Y. Leyton-Brown, K. *Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations.* Cambridge University Press, 2009.

[45] Nash, J. *Non-cooperative games.* Ann Math 54(2):286-295

[46] Gintis, H. *Game Theory Evolving.* Princeton University Press, 2000.

[47] Selten, R. *Reexamination of the perfectness concept for equilibrium points in extensive games.* Int J Game Theory 4(1):25–55

[48] Kuhn, H. (1953) *Extensive games and the problem of information.* In: Kuhn, Tucker (eds) Contributions to the theory of games, vol 2. Princeton University Press, Princeton, pp 123-216.

[49] Aumann, R. (1964) *Mixed and behaviour strategies in infinite extensive games.* In: Dresher, Shapley, Tucker (eds). Advances in game theory, annals of mathematic studies, vol 52. Princeton University Press, Princeton, pp 627-650.

[50] Maschler, M. Solan, E. Zamir, S. *Game Theory.* Cambridge University Press 2013.

[51] Hofbauer, J. and Sigmund, K. *Evolutionary Games and Population Dynamics.* Cambridge University Press, Cambridge, UK, 1998.

[52] Weibull, J. W. *Evolutionary Game Theory.* The MIT Press, Cambridge, Massachusetts, 1995.

[53] Young, P. *The evolution of cooperation.* Econometrica 61: 57–58 (1993)

[54] Kandori, M. Mailath, G. J. and Rob, R. *Learning, mutation, and long run equilibria in games.* Econometrica 61: 1019–1045 (1993)

[55] Smith, J. M. *Will a Sexual Population Wvolve to an ESS?* The American Naturalist, Vol. 117, No. 6 (Jun., 1981), 1015-1018.

[56] Serfozo, R. *Basics of Applied Stochastic Processes.* Springer-Verlag Berlin Heidelberg, 2009.

[57] Freidlin, M. I. Wentzell, A.D. *Random Perturbations of Dynamical Systems.* New York: Springer Verlag (1984).

[58] Selten, R. *Reexamination of the perfectness concept for equilibrium points in extensive games.* Int J Game Theory 4(1):25–55 (1975).

[59] Tirole, J. and Fudenberg, D. *Game Theory.* MIT Press; 1st Edition edition (1991).

[60] Taryp, D.R., Gilley, D.C. and Seeley, T.D. *Levels of selection in a social insect: a review of conflict and cooperation during honey bee (Apis mellifera) queen replacement.* Behavioural Ecology and Sociobiology, Vol. 55, pp. 513–523 (2004).

[61] Seyfarth, R.M. and Cheney, D.L. *Grooming, alliances and reciprocal altruism in vervet monkeys.* Nature, Vol. 308, pp. 541–543 (1984).

[62] Wilkinson, G.S. *Reciprocal food sharing in the vampire bat.* Nature, Vol. 308, pp. 181–184 (1984).

[63] Brosig, J. *Identifying cooperative behavior: some experimental results in a prisoner's dilemma game.* Journal of Economic Behavior and Organization Vol. 47, pp. 275–290 (2002).

[64] Ledyard, J.O. *Public Goods: A survey of Experimental Research.* In *Handbook of Experimental Economics,* ed. J.H. Kagel and A.E. Roth, pp. 111–194. Princeton, NJ: Princeton University Press (1995).

[65] Axelrod, C. and Hamilton, W.D. *The Evolution of Cooperation.* Science, New Series, Vol. 211(2289), pp. 1390–1396 (1981).

[66] Godard, R. *Tit for tat among neighbouring hooded warblers.* Behavioural Ecology and Sociobiology, Vol. 33, pp. 45–50 (1993).

[67] Milinski, M. *Tit for tat in sticklebacks and the evolution of cooperation.* Nature, Vol. 325, pp. 433–435 (1987).

[68] Freidman, J. W. *A Non-Cooperative Equilibrium in Supergames.* Review of Economic Studies, Vol. 38, pp. 1-12 (1971).

[69] Hegselmann, R. and Flache, A. *Rational and Adaptive Playing: A Comparative Analysis for All Possible Prisoner's Dilemmas.* Analyse and Kritik (2000).