Abstract

Standard non-cooperative game theory is not selective enough to discriminate among many of the possible outcomes in infinitely repeated reciprocity games. However, experimental evidence, especially in symmetric games, suggests strongly that people arrive at only a few outcomes. Theoretical justification for these particular outcomes is usually derived from focal and axiomatic solution concepts, yet these procedures are inherently ad-hoc. Here the outcome of a population game with evolutionary dynamics is examined in order to theatrically justify experimentally observed behavior in a less ad-hoc manner. It is shown that under an assumption of limited rationality, specifically limited memory, there is a unique global equilibrium in the Replicator Dynamics. The unique equilibrium contains a trio of outcomes: non-cooperative Nash play, payoff irrational play, and cooperative turn-taking, which roughly match the outcomes observed experimentally.

JEL classification: C62, C72, C73.
Keywords: Bounded Rationality, Evolution, Replicator, Reciprocity.

A Bounded Rationality, Evolutionary Model for Behavior in Two Person Reciprocity Games

Jeffrey Prisbrey
Universitat Pompeu Fabra[1]

July 1993.

[1]This paper has benefited from the comments of Richard McKelvey and Thomas Palfrey. Please direct correspondence to: Universitat Pompeu Fabra, Balmes 132, 08008 Barcelona, Spain.

1 Population Games and the Replicator Dynamic

Due to the Folk Theorem, a generic infinitely repeated game has many equilibria. The multitude of equilibria is a problem for theorists because a justifiable and non-arbitrary method of eliminating the majority of the equilibria has not been found. Experimental evidence suggests that players in infinitely repeated, symmetric reciprocity games usually succeed in establishing a pattern of alternation, *e.g.* Prisbrey (1992). The player's actions have given a clue as to which of the equilibria should remain after elimination. Presented here is a method which eliminates most of the experimentally unobserved equilibria.

The method is based on a mathematical model of evolution, developed in biology, called the Replicator Dynamic. The Replicator Dynamic supposes a large population of players, each endowed with a particular strategy. Each player in the population lives (plays a game), creates offspring identical to itself, and then dies. The mixture of player types within the population changes from generation to generation as the population grows and depends upon the success that each player has in creating offspring. In the Replicator Dynamic, each player begets a number of offspring that is proportional to that player's lifetime fitness, or payoff. The result is that later generations have a higher proportion of players endowed with high payoff strategies.

Suppose that the initial population has every possible strategy represented in it. Then, if the Replicator Dynamic is ever in equilibrium, meaning that the mixture of player types remains the same from generation to generation, the strategies that remain in the equilibrium have been justified in a Darwinistic sense.

The problem with the Replicator Dynamic is creating an initial population with every possible strategy in it. In the case of an infinitely repeated game, there are infinitely many possible strategies making it necessary to have a population of infinite size. The analysis of the dynamics on such a population are beyond the state of the art. For any analysis to succeed, there must be only a finite number of possible strategies. It is impossible, then, without further assumptions, to use the Replicator Dynamic as a method of justification, at least in the infinitely repeated reciprocity game. Here, it will be assumed that players have a finite memory. This *bounded rationality constraint* uniformly limits the number of possible strategies.

Application of the model to the infinitely repeated, symmetric reciprocity game succeeds in the sense that there is only one possible equilibrium. The equilibrium encompasses the exchange of favors as well as a behavior associated with short run payoff maximization and a behavior which could be coined as irrational (although none of the players are rational in any sense). All of these behaviors are seen in the experiments reported in Prisbrey (1992). It is not a complete success, however, because the behaviors are not seen in the same ratios and furthermore, the irrational behavior is not an equilibrium in the standard sense.

These types of population games have been studied before, perhaps the best known examples are the papers by Axelrod (1979) and Axelrod and Hamilton (1981) which reported on Repeated Prisoner's Dilemma tournaments. In these tournaments, various people, most of them professional scientists, submitted computer programs which were, in essence, strategies in the repeated Prisoner's Dilemma. Together, the programs made an artificial population which competed by playing a repeated Prisoner's Dilemma in round robin fashion. After competing each strategy was reproduced based on their relative scores, the higher a strategy's score, the higher that strategy's representation in latter generations. They found that the strategy tit-for-tat displaced the other submitted strategies.

A variety of papers focusing on the dynamics of the tournaments followed. Blad (1986), Hirshleifer and Martinez Coll (1988), Mueller (1987) and Young and Foster (1991) use Replicator Dynamics to justify or determine equilibria in three-strategy Prisoner's Dilemma and perturbed Prisoner's Dilemma games. The strategies considered were: All Defect, All Cooperate, and some variant of tit-for-tat (grim for example). The cooperative outcome was an equilibrium in almost all settings in these works. Smale (1980) applied the Replicator Dynamics to a setting where players only remembered a summary of the past (an average of their past payoffs). He found that if the players play *good* strategies, then cooperation is a globally stable equilibrium.

A different approach was taken by Miller (1989) who used an optimization technique called the Genetic Algorithm. A Genetic Algorithm takes a subset of the possible strategies as a population. The possible strategies, in turn, are determined by the computing power available. The population then evolves much like populations under the Replicator Dynamic. The difference is that each member of the population faces a probability of random mutation (be it gene specific or crossover) before or after the next generation is formed. Miller found that "…cooperative strategies …tend to proliferate throughout the population under [certain conditions]."[1]

One criticism of these lines of research is that all of them apply their dynamic models to subsets of the possible available strategies. Furthermore, the subsets are determined in relatively arbitrary ways. In the three strategy dynamic models, for example, no reason is given for considering tit-for-tat while not considering, at the same time, the grim strategy.[2] This criticism becomes more powerful when the works of Boyd and Lorberbaum (1987) and Nachbar (1989) are taken into account. Boyd and Lorberbaum showed that, contrary to previous optimistic research, no pure strategy is evolutionarily stable in the infinitely repeated Prisoner's Dilemma. This finding depends upon the fact that all possible strategies have a chance of being played. Nachbar showed that the limit of the Replicator Dynamic in a two-stage Prisoner's Dilemma has everybody defecting (although All Defect is not the only strategy in the limit).

One way to uniformly limit the number of strategies under consideration in an infinitely repeated setting is to apply a bounded rationality constraint. The constraint is a logical one to consider given the comments of Aumann and Sorin (1989) who write:

> The first hint that bounded recall might have something to do with cooperation came in the summer of 1978. Aumann and Kurz, with the help of Jonathan Cave …worked out a version of the infinitely repeated Prisoner's Dilemma with memory one; this means that each player can base his action only on what his opponent did at the previous stage – he has "forgotten" everything else. This results in an 8×8 bi-matrix game; iterated removal of weakly dominated strategies yields a unique

[1] Miller (1989), p. 12.

[2] Mueller (1987) attempts an argument by showing that he considers as a third strategy the strategy that in some sense punishes optimally.

> strategy pair, in which both players start by playing "friendly" and continue with "tit-for-tat" thereafter. The outcome is cooperative, both players always playing "friendly."[3]

However, Aumann and Sorin is not a paper concerned with population dynamics, and so it proceeds down a different path.

The effects of a one period recall will be considered here, only the payoff structure will not be that of a Prisoner's Dilemma. Instead, the analysis will focus on the Reciprocity Game. This paper is, in a sense, an answer to Rapoport (1988) who laments about the "…persistent hegemony of Prisoner's Dilemma …" and claims that "…it is evident that there is enough to do in this area [of 2×2 games] for an army of investigators."[4]

2 The Environment

Let G be a symmetric, two-person, strategic game with finite action spaces $A_i = A_j = \{a, b\}$ and payoff matrix

$$M = \begin{bmatrix} \alpha & \beta \\ \eta & \gamma \end{bmatrix}$$

where the top row and first column correspond with the choice of action a and the bottom row and second column with action b. Let G^∞ be the supergame made up of an infinite sequence of plays of game G.

A history or memory of length k for player i is defined as $h_i^k \in \Pi_{l=1}^k A_j$. Notice that under this definition, player i only has a memory of the last k actions of player j; player i does not remember his own actions. Let $s^k \in \mathcal{S}^k : h^k \mapsto A$ (with subscripts suppressed) be a function that maps a player's memory into an action. Call s^k a strategy with a bounded memory of length k and let $S^k \subset S^\infty$ be the set of all k length strategies. Let $S^B \equiv S^1$.

Another way to think of the set S^B is as the set of strategies which can be implemented by a two-state automaton, such automatons are commonly called Moore machines. A Moore machine, here from player i's point of view, consists of a quadruple, $\{H, q_0, f, \lambda\}_i$, where,

1. H is a finite set of histories or states,
2. q_0 is an initial state,
3. $f : H \times A_j \mapsto H$ is a transition function, and
4. $\lambda : H \mapsto A_i$ is a behavior function.

In this particular case, it is convenient to suppress H and f and explicitly enumerate λ. This should cause no confusion because $H = \{a, b\}$ and f maps A_j directly into H, *i.e.* $f(a) = a$ and $f(b) = b$. This convention allows a machine to be written as a triple, for example $\{a, a, a\}$,

[3] Aumann and Sorin (1989), p. 9.
[4] Rapoport (1988), pp. 400 – 401.

where the first represents q_0 the initial move of the machine, the second represents $\lambda(a)$, the move that the machine chooses if its opponent chooses action a, and the third represents $\lambda(b)$, the action chosen if its opponent chooses action b. The machine $\{a, a, a\}$ plays action 'a' on the first move, and then plays action 'a' regardless of the action its opponent chooses. There are eight possible two stage machines with these characteristics and they correspond directly with the strategies in the set S^B. Number the eight machines as in Table 1.

Suppose that two players, who are limited to choosing strategies in S^B or equivalently to choosing one of the eight machines, meet and play G^∞. Because of the finite strategies, the sequence of play eventually cycles, with the longest cycle being four stages. For example, if player i chooses machine $s_5 = \{a, a, b\}$ and player j chooses machine $s_8 = \{b, b, a\}$, then the sequence of plays will be $\{(a, b), (b, b), (b, a), (a, a), (a, b), (b, b), \ldots\}$, with the first of each pair in the sequence being player i's move. Player i's sequence of payoffs will be $\{\beta, \gamma, \eta, \alpha, \beta, \gamma, \ldots\}$; the payoffs will also cycle. Define the function $\pi : S^B \times S^B \mapsto \Re$ as player i's average cycle payoff. For this example,

$$\pi(s_5, s_8) = \frac{1}{4}(\beta + \gamma + \eta + \alpha).$$

As an alternative example, consider the payoff if player i had chosen s_1 and player j had chosen s_2. In this case, the sequence of play will be $\{(a, b), (a, a), (a, a), \ldots\}$. After the first stage, the machines play (a, a) forever. The average cycle payoff to player i is,

$$\pi(s_1, s_2) = \alpha.$$

The application of the bounded rationality constraint and the particular definition of the payoff functions has transformed the infinitely repeated game G^∞ into a single period game with an 8×8 payoff matrix, Π. Π is shown in Table 2

3 The Replicator

The following notation is inspired by Taylor and Jonker (1978). Consider a population of N risk-neutral, payoff maximizing players who interact in randomly matched pairs. Let n_i be the number of players who choose strategy i. The population can then be represented as a point $\boldsymbol{p}$ in the eight-dimensional simplex Δ, with $p_i = n_i/N$ and $\sum_{i=1}^{8} p_i = 1$.

Assume that there is exponential growth or decay. Specifically, $\frac{dn_i}{dt} = r_i n_i$, where r_i is the current growth rate for n_i. Growth in the population follows $\frac{dN}{dt} = \bar{r}N$, where $\bar{r}$ is the average growth rate.

By differentiating $p_i = \frac{n_i}{N}$,

$$\begin{aligned} \frac{dp_i}{dt} &= \frac{\frac{dn_i}{dt}}{N} - \frac{n_i \frac{dN}{dt}}{N^2} \\ &= \frac{r_i n_i}{N} - \frac{\bar{r} n_i}{N} \\ &= p_i(r_i - \bar{r}). \end{aligned}$$

Now, assume that the growth rate of players with strategy i is equivalent to the expected payoff, or fitness, of player i. In other words, $F(i|\boldsymbol{p}) = \sum_{l=1}^{8} p_l \pi(s_i, s_l)$, which is the expected payoff of player i, is equivalent to r_i. Similarly, $F(\boldsymbol{p}|\boldsymbol{p}) = \sum_{i=1}^{8} p_i F(i|\boldsymbol{p})$, which is the expected payoff of a random member of the population, is equivalent to $\bar{r}$. Then by substitution,

$$\frac{dp_i}{dt} = p_i[F(i|\boldsymbol{p}) - F(\boldsymbol{p}|\boldsymbol{p})]. \tag{1}$$

Now, $\frac{dp_i}{dt}$ is the instantaneous change in the proportion of players using strategy i. Note that $\frac{dp_i}{dt} > 0$ if and only if $F(i|\boldsymbol{p}) > F(\boldsymbol{p}|\boldsymbol{p})$, and $\frac{dp_i}{dt} < 0$ if and only if $F(i|\boldsymbol{p}) < F(\boldsymbol{p}|\boldsymbol{p})$. Hence, the proportion of players using strategy i rises (or falls) with time only if the expected payoff of strategy i is greater than (or less than) the expected payoff of a random member of the population. If the expected payoff to strategy i is the same as the expected payoff to a random member of the population, then $\frac{dp_i}{dt} = 0$.

Equation 1 implies a dynamic in continuous time on the simplex Δ. Given an initial state or initial population in Δ, the dynamic describes a particular trajectory.

Assumption 1 *Every initial population is a point* $\boldsymbol{p}$ *located in the interior of the simplex* Δ*.*

The assumption means that every possible strategy has at least some representation in the population.

Definition 1 *An* ***equilibrium*** *is any population* $\boldsymbol{p}$ *such that* $\frac{dp_i}{dt} = 0$ *for all i.*

Definition 2 *Given an equilibrium* $\boldsymbol{p}$*,* $\boldsymbol{p}$ *is* ***asymptotically stable*** *if a trajectory that passes through* $\boldsymbol{p}'$ *converges to* $\boldsymbol{p}$ *with time, for all* $\boldsymbol{p}'$ *in an open neighborhood around* $\boldsymbol{p}$*.*

Definition 3 *Given an equilibrium* $\boldsymbol{p}$*,* $\boldsymbol{p}$ *is* ***globally asymptotically stable*** *if a trajectory that passes through* $\boldsymbol{p}'$ *converges to* $\boldsymbol{p}$ *with time, for all* $\boldsymbol{p}'$ *in the interior of* Δ*.*

Playing a random member of a population $\boldsymbol{p}$ is like playing against a mixed strategy $\boldsymbol{q}$ where $\boldsymbol{p} = \boldsymbol{q}$. A particular mixed strategy will be denoted $\boldsymbol{q}_i$ and will be treated in the obvious way by the function F.

Assumption 2 *The payoff matrix, M, is such that* $\eta > \frac{\eta+\beta}{2} > \gamma > max\{\beta, \alpha\}$*.*

The assumption means action a is strictly payoff dominated by action b and assures that in what follows $x \in (1/2, 1)$. Furthermore, the assumption defines the properties necessary for the game to be a symmetric Reciprocity Game.

Now, which strategies are of interest in a Reciprocity game? The strategies s_1 and s_2 always play a, which is a dominated strategy in the game G. The strategies s_3 and s_4 always play action b, which is the dominant strategy in the game G. There are three ways alternation can occur: if a player with strategy s_5 meets a player with strategy s_6, or if two players with strategies s_7 or s_8 meet. Only the first of these ways is consistent with the idea of reciprocation.

Define the point $\boldsymbol{q}^e = \boldsymbol{p}^e = [0, 0, 0, 0, x, (1-x), 0, 0]$, with x such that the following equality holds:

$$x\alpha + (1-x)\frac{\beta+\eta}{2} = x\frac{\eta+\beta}{2} + (1-x)\gamma.$$

If the population is at point $\boldsymbol{p}^e$, then the only strategies present are strategies s_5 and s_6.

Lemma 1 *The point* $\boldsymbol{p}^e$ *is an equilibrium.*

Proof: This is true since,

$$\begin{aligned} F(5|\boldsymbol{p}^e) &= x\alpha + (1-x)\frac{\beta+\eta}{2} \\ &= x\frac{\eta+\beta}{2} + (1-x)\gamma \\ &= F(6|\boldsymbol{p}^e) \\ &= F(\boldsymbol{p}^e|\boldsymbol{p}^e) \end{aligned}$$

implies $\frac{dp_i^e}{dt} = 0$ for all i.

□

So, an equilibrium with both s_5 and s_6 players present exists. In this equilibrium, every time an s_5 player meets and s_6 player, there will be Alternation. Of course, meetings between s_5 and s_6 players are not the only types of meetings that occur. When an s_6 player meets another s_6 player, the sequence of play is $\{\ldots, (b, b), \ldots\}$. At each stage, the players myopicly choose the Dominant Strategy Nash equilibrium. It is also an subgame-perfect Nash equilibrium in the game G^∞, although every other equilibrium has Pareto Superior payoffs. The last type of meeting which could occur is between two s_5 players. In this case, the sequence of play is $\{\ldots, (a, a), \ldots\}$ which is not subgame perfect Nash equilibrium play; it will be called Irrational.

Lemma 2 *The equilibrium* $\boldsymbol{p}^e$ *is locally asymptotically stable.*

Proof: By Lemma 1 $\boldsymbol{p}^e$ is an equilibrium. In equilibrium, the expected payoff to a random member is $x\frac{\eta+\beta}{2} + (1-x)\gamma$. Suppose point $\boldsymbol{p}'$ is an element of an open neighborhood around $\boldsymbol{p}^e$. If the trajectory through $\boldsymbol{p}'$ converges with time to $\boldsymbol{p}^e$ then $\boldsymbol{p}^e$ is locally asymptotically stable.

Two conditions must be meet for the trajectory to converge to $\boldsymbol{p}^e$. First, no strategy with positive weight at point $\boldsymbol{p}'$ can have a higher fitness when playing $\boldsymbol{q}^e$, the equilibrium mixed strategy, than $\boldsymbol{q}^e$ itself.

Second, if any strategy happens to do equally as well as $\boldsymbol{q}^e$, then it must be the case that that strategy is in the support of $\boldsymbol{q}^e$ and $\boldsymbol{q}^e$ must do better when playing $\boldsymbol{q}'$ than $\boldsymbol{q}'$ itself, where $\boldsymbol{q}'$ is the mixed strategy associated with the point $\boldsymbol{p}'$.

Formally,

- For all $p_i' > 0$, $F(s_i|\boldsymbol{p}^e) \leq F(\boldsymbol{q}^e|\boldsymbol{p}^e)$ and,

- if $F(s_i|\boldsymbol{p}^e) = F(\boldsymbol{q}^e|\boldsymbol{p}^e)$, then $q_i > 0$ and $F(s_i|\boldsymbol{p}') < F(\boldsymbol{q}^e|\boldsymbol{p}')$.

The fact that these two conditions are sufficient for asymptotic stability is due to Taylor and Jonker (1978).

Note that it is enough to consider only pure strategies with positive weight because any mixed strategy will have a payoff that is a linear combination of the payoffs to pure strategies. If all the pure strategies satisfy the previous two conditions, then any mixed strategy will as well.

The payoffs of all strategies that might have positive weight in a disturbed state against the equilibrium strategy are:

- $F(1,\boldsymbol{p}^e) = F(2,\boldsymbol{p}^e) = x\alpha + (1-x)\alpha = \alpha$
- $F(3,\boldsymbol{p}^e) = F(4,\boldsymbol{p}^e) = x\gamma + (1-x)\gamma = \gamma$
- $F(5,\boldsymbol{p}^e) = F(6,\boldsymbol{p}^e) = x\frac{\eta+\beta}{2} + (1-x)\gamma$
- $F(7,\boldsymbol{p}^e) = F(8,\boldsymbol{p}^e) = x\frac{\alpha+\beta+\eta+\gamma}{4} + (1-x)\frac{\alpha+\beta+\eta+\gamma}{4} = \frac{\alpha+\beta+\eta+\gamma}{4}$

None of these strategies does better against $\boldsymbol{q}^e$ than $\boldsymbol{q}^e$ itself. Both s_5 and s_6 do equally as well, however, so how these strategies do against themselves and how $\boldsymbol{q}^e$ does against them must also be considered.

- $F(\boldsymbol{q}^e,5) = F(\boldsymbol{q}^e,6) = x\frac{\eta+\beta}{2} + (1-x)\gamma$
- $F(5,5) = \alpha$
- $F(6,6) = \gamma$

Because the payoff to $\boldsymbol{q}^e$ is higher than the payoff to strategy s_5 when both play against s_5 and because its payoff is higher than the payoff to s_6 when both play s_6, and because no other strategy does as well against it as itself, the equilibrium is locally stable.

□

Hence the equilibrium $\boldsymbol{p}^e$ is resistant to small shocks or invasion by small numbers of players with a different strategy. But what if there is a large shock or if there are large numbers of players with other strategies in the initial population? To determine what happens in these cases, the following two lemmas will be used.[5]

Lemma 3 *Given s_i and s_j such that $F(i,\boldsymbol{p}) < F(j,\boldsymbol{p})$ for all $\boldsymbol{p} \in \Delta$, and any $\boldsymbol{p}^0$ in the interior of* Δ, $\lim_{t\to\infty} p_i = 0$.

[5]The two Lemmas show that it is possible to iteratively eliminate strictly dominated strategies in this case. A general theorem encompassing this result can be found in Samuelson and Zhang (1992).

Proof: First, note that because p_i and p_j are in Δ, $\frac{p_i}{p_j}$ is bounded below by 0. Because of this, it is enough to show that $\lim_{t\to\infty} \frac{p_i}{p_j} = 0$. Time is continuous and runs from 0 to ∞. Consider an infinite sequence of points in time $\{t_1, t_2, \ldots, t_n, \ldots\}$, such that $t_n > t_{n-1}$. Given any initial population, it is possible to determine the values of p_i and p_j at any point t_n. Define a second sequence by $T_n = \frac{p_i(t_n)}{p_j(t_n)}$.

Now, to show that the sequence of T_ns is monotonicly decreasing, it is sufficient to show that

$$\frac{\mathrm{d}p_i/p_j}{\mathrm{d}t} = (p_j\frac{dp_i}{dt} - p_i\frac{dp_j}{dt})/p_j^2 < 0$$

which implies that $\frac{\frac{dp_i}{dt}}{p_i} < \frac{\frac{dp_j}{dt}}{p_j}$.

Because $F(i|\boldsymbol{p}) < F(j|\boldsymbol{p})$,

$$\begin{array}{ccc} \frac{\frac{dp_i}{dt}}{p_i} & = & [F(i|\boldsymbol{p}) - F(\boldsymbol{p}|\boldsymbol{p})] \\ & \text{and} & \\ \frac{\frac{dp_j}{dt}}{p_j} & = & [F(j|\boldsymbol{p}) - F(\boldsymbol{p}|\boldsymbol{p})] \\ & \Downarrow & \\ \frac{\frac{dp_i}{dt}}{p_i} & < & \frac{\frac{dp_j}{dt}}{p_j}. \end{array}$$

Because the sequence T_n is monotonicly decreasing and bounded below by 0, it must converge, and because it converges, $\lim_{t\to\infty} \frac{p_i}{p_j}$ must also converge.

Suppose the limit converges to a point x greater than zero. Then, at point x, $\frac{\frac{dp_i}{dt}}{p_i} = \frac{\frac{dp_j}{dt}}{p_j}$. This implies that $F(i|\boldsymbol{p}) = F(j|\boldsymbol{p})$ which is a contradiction. So, the limit point must be 0.

□

Lemma 4 *Consider population* $\boldsymbol{p}$ *with* $p_k < \varepsilon$, *strategies* s_l *and* $\boldsymbol{q}$ *such that* $F(l|i) \leq F(\boldsymbol{q}|i)$ *for all* $i \neq k$ *and* $F(l|i) < F(\boldsymbol{q}|i)$ *for at least one* $i \neq k$. *If* ε *is small enough so that* $F(l|\boldsymbol{p}) < F(\boldsymbol{q}|\boldsymbol{p})$ *and if* p_k *is never bigger than* ε, *then for* $\boldsymbol{p}^0$ *in the interior of* Δ, $\lim_{t\to\infty} p_l = 0$.

Proof: First, note that because p_l and $\boldsymbol{q}$ are in Δ, $\frac{p_l}{\sum_{i|q_i>0} p_i}$ is bounded below by 0. Because of this, it is enough to show that $\lim_{t\to\infty} \frac{p_l}{\sum_{i|q_i>0} p_i} = 0$. Time is continuous and runs from 0 to ∞. Consider an infinite sequence of points in time $\{t_1, t_2, \ldots, t_n, \ldots\}$, such that $t_n > t_{n-1}$. Given any initial population, it is possible to determine the values of p_l and $\sum_{i|q_i>0} p_i$ at any point t_n. Define a second sequence by $T_n = \frac{p_l(t_n)}{\sum_{i|q_i>0} p_i(t_n)}$.

Now, to show that the sequence of T_ns is monotonicly decreasing, it is sufficient to show that

$$\frac{\mathrm{d}p_l / \sum_{i|q_i>0} p_i}{\mathrm{d}t} = (\sum_{i|q_i>0} p_i \frac{dp_l}{dt} - p_i \sum_{i|q_i>0} \frac{dp_i}{dt})/(\sum_{i|q_i>0} p_i)^2 < 0$$

which implies that $\frac{\frac{dp_l}{dt}}{p_l} < \frac{\sum_{i|q_i>0} \frac{dp_i}{dt}}{\sum_{i|q_i>0} p_i}$.

Because $F(l|\boldsymbol{p}) < F(\boldsymbol{q}|\boldsymbol{p})$,

$$\frac{\frac{dp_l}{dt}}{p_l} = [F(l|\boldsymbol{p}) - F(\boldsymbol{p}|\boldsymbol{p})]$$

and

$$\frac{\sum_{i|q_i>0} \frac{dp_i}{dt}}{\sum_{i|q_i>0} p_i} = [F(i|\boldsymbol{p}) - F(\boldsymbol{p}|\boldsymbol{p})]$$

$$\Downarrow$$

$$\frac{\frac{dp_l}{dt}}{p_l} < \frac{\sum_{i|q_i>0} \frac{dp_i}{dt}}{\sum_{i|q_i>0} p_i}.$$

Because the sequence T_n is monotonicly decreasing and bounded below by 0, it must converge, and because it converges, $\lim_{t\to\infty} \frac{p_l}{\sum_{i|q_i>0} p_i}$ must also converge.

Suppose the limit converges to a point x greater than zero. Then, at point x, $\frac{\frac{dp_l}{dt}}{p_l} = \frac{\sum_{i|q_i>0} \frac{dp_i}{dt}}{\sum_{i|q_i>0} p_i}$. This implies that $F(l|\boldsymbol{p}) = F(\boldsymbol{q}|\boldsymbol{p})$ which is a contradiction. So, the limit point must be 0.

□

And now, the main result:

Theorem 1 *The equilibrium $\boldsymbol{p}^e$ is globally asymptotically stable.*

Proof: The strategy s_4 strictly dominates the strategies s_1 and s_2, $F(s_4|\boldsymbol{p}) > F(s_1|\boldsymbol{p})$ and $F(s_4|\boldsymbol{p}) > F(s_2|\boldsymbol{p})$ for all $\boldsymbol{p}$ in Δ. By Lemma 3, p_1 and p_2 go monotonicly to zero as t goes to infinity. In particular, for any small positive number ε, at some point in time, p_1 and p_2 will both be less then ε.

Now, either s_4 or $\boldsymbol{q}^e$ strictly dominates s_7 with regard to all strategies except s_1 and s_2. If

$$\frac{\alpha + \beta + \eta + \gamma}{4} - \beta > 0$$

then $\boldsymbol{q}^e$ is strictly dominant. If the inequality does not hold, then clearly,

$$\gamma - \frac{\alpha + \beta + \eta + \gamma}{4} > 0$$

implying that s_4 is strictly dominant.

Because the inequalities above are strict, it is possible to pick an ε small enough so that if the weights of the strategies s_1 and s_2 are less than ε, either $F(s_4|\boldsymbol{p}) > F(s_7|\boldsymbol{p})$ or $F(\boldsymbol{q}^e|\boldsymbol{p}) > F(s_4|\boldsymbol{p})$. Lemma 4 then implies that p_7 goes to zero as t goes to infinity.

Similarly, Lemma 4 can be used to show that p_8 and then p_3 and p_4 go to zero as t goes to infinity; each time $\boldsymbol{q}^e$ is the mixed strategy needed in the Lemma.

We are now left with only two pure strategies that can have weight greater than ε, s_5 and s_6. Suppose p_5 is very small, then $F(s_5|\boldsymbol{p}) = (\beta + \eta)/2 + \delta_1$ and $F(s_6|\boldsymbol{p}) = \gamma + \delta_2$, for some δ_1 and δ_2 small. If p_5 is small, then p_5 will grow with time. What if p_5 is large? Then $F(s_5|\boldsymbol{p}) = \alpha + \delta_1$ and $F(s_6|\boldsymbol{p}) = (\beta + \eta)/2 + \delta_2$, for some δ_1 and δ_2 small. If p_5 is large, then p_5 will decay with time.

In any case, the trajectory through any initial point must eventually come within any neighborhood of $\boldsymbol{p}^e$, and by Lemma 2 converge to $\boldsymbol{p}^e$.

□

No matter what the initial population is (as long as it is in the interior of Δ), the Replicator Dynamic will converge to an equilibrium with only s_5 and s_6 players. In this equilibrium there will be three types of sequences of play: Alternation, Dominant Strategy Nash, and Irrational.

3.1 Other Symmetric Games

What types of problems occur if Assumption 2 is not met? Well, suppose $\alpha > \gamma$, which covers the case of the Prisoner's Dilemma. Then there is no strategy that strictly dominates another. While $\boldsymbol{p}^e$ is still an equilibrium, it is not a globally asymptotically stable equilibrium. The same result occurs if $\beta > \gamma$, or if $\beta > \eta$; these cases cover the game of Chicken. A numerical example encompassing both of these alternatives will be given later.

Recall Aumann and Sorin's application of a one period recall to the infinitely repeated Prisoner's Dilemma. Aumann and Sorin justify their result by the elimination of strategies based on weak dominance alone. Unfortunately, under the Replicator Dynamic weak dominance alone is not enough to assure that a particular strategy's representation in the population goes to zero. While Nachbar (1988) does prove a theorem which gives positive convergence results in a subset of weakly dominant solvable games, his result cannot be applied in Aumann and Sorin's example. The difficulties encountered in weakly dominant solvable games are covered well in Nachbar (1988) and interested readers are referred there.

3.2 Asymmetric Games

There are inherent similarities between the Battle of the Sexes and Reciprocity Games which might lead you to believe that a similar result could be obtained in the Battle of the Sexes. There is a problem, however – the Battle of the Sexes is an asymmetric game. There are two approaches in the modeling of asymmetric games as population games: analyze two distinct populations, and make the game symmetric through random population assignment.

Analyzing two distinct populations changes the dynamics dramatically. Consider what would happen in the case of the Reciprocity Game. Call the two populations the row population, ***rp***, and the column population, ***cp***. Let them evolve in the obvious way. Then the populations such that $\boldsymbol{rp} = \boldsymbol{cp} = \boldsymbol{p}^e$ would still be an equilibrium, but instead of being an global attractor, it would be a repeller. Any small deviation leading to a higher number of s_5 row players, for example, will cause the dynamics to flow towards populations consisting entirely of s_5 row players and s_6 column players. There are many other equilibria possible, each depending upon the initial populations. A global result is impossible.

Suppose that the game is made symmetric. The obvious way of accomplishing the task is to randomly choose one of each pair of players to be the row player and to let the other be the column player. A player's payoff would be their average payoff gotten as a row player plus their average payoff gotten as a column player divided by two. Alternatively, each player would face a payoff matrix consisting of cells which were the average payoff across both types given those actions. Specifically, suppose

$$M_c = \begin{bmatrix} \alpha_c & \beta_c \\ \eta_c & \gamma_c \end{bmatrix}$$

and

$$M_r = \begin{bmatrix} \alpha_r & \beta_r \\ \eta_r & \gamma_r \end{bmatrix},$$

where M_c was the payoff matrix faced by column players and M_r was the payoff faced by row players. Then the payoff matrix faced by a player in the version of this game played with random population assignment would be:

$$M = \begin{bmatrix} \frac{\alpha_c+\alpha_r}{2} & \frac{\beta_c+\beta_r}{2} \\ \frac{\eta_c+\eta_r}{2} & \frac{\gamma_c+\gamma_r}{2} \end{bmatrix}.$$

This method is an improvement over the two population method because it does not change the outcome predicted in the Reciprocity Game. In fact, any asymmetric game that meets Assumption 2 after having been made symmetric will meet all the assumptions required by Theorem 1.

Unfortunately, even with random population assignment, the Battle of the Sexes does not meet Assumption 2.

4 Examples

Consider the payoff matrix M_1,

$$M_1 = \begin{bmatrix} 3 & 3 \\ 7 & 4 \end{bmatrix}.$$

Then Theorem 1 holds with $x = 1/3$. Figure 1 shows a phase portrait for the initial generation that has all strategies with equal representation in the population.

In equilibrium, there are only three possible outcomes to a meeting between two players, call them: Alternation, Dominant Strategy Nash Play, and Irrational Play. Alternation occurs whenever a player with strategy s_5 meets a player with strategy s_6. The sequence of play in this case would be $\{(a,b),(b,a),(a,b),\ldots\}$. Alternation occurs with probability 4/9. Dominant Strategy Nash Play occurs whenever a player with strategy s_6 meets another player with strategy s_6. The sequence of play in this case would be $\{(b,b),(b,b),(b,b),\ldots\}$. Dominant Strategy Nash play occurs with probability 4/9. Irrational Play occurs whenever a player with strategy s_5 meets another player with strategy s_5. The sequence of play in this case would be $\{(a,a),(a,a),(a,a),\ldots\}$. The probability of this outcome is 1/9.

As an example of what happens if the payoff matrix is not constructed with the correct inequalities, consider the payoff matrix M_2,

$$M_2 = \begin{bmatrix} 4 & 3 \\ 7 & 2 \end{bmatrix}.$$

In this case, Theorem 1 does not hold. Figure 2 shows a phase portrait for the initial generation that has all strategies with equal representation in the population. Figure 3 shows a phase portrait for a different initial generation. Notice that the equilibria are different for these two initial generations.

5 Conclusion

It has been shown that a large class of two-player, bi-matrix games, both symmetric and asymmetric, have a unique equilibrium when they are modeled as population games containing players with bounded recall. The class is the set of all games which meet Assumption 2. In the unique equilibrium, both trading favors and short term maximization occur. A third irrational outcome also occurs. Normative justification for all three of these behaviors can be obtained from the Darwinistic maxim claiming that only the fittest should survive.

6 References

AKIN, E. AND J. HOFBAUER. 1982. "Recurrence of the Unfit." *Mathematical Biosciences* Vol. 61: pp. 51 – 62.

AUMANN, R. AND S. SORIN. 1989. "Cooperation and Bounded Recall." *Games and Economic Behavior* Vol. 1: pp. 5 – 39.

AXELROD, R. November 1979. "The Evolution of Cooperation in the Prisoner's Dilemma." Institute of Public Policy Studies Discussion Paper No. 143. Ann Arbor, Michigan: University of Michigan.

AXELROD, R. AND W. D. HAMILTON. March 27, 1981. "The Evolution of Cooperation." em Science Vol. 211: pp. 1390 – 1396.

BANKS, J. AND R. K. SUNDARAM. 1990. "Repeated Games, Finite Automata, and Complexity." *Games and Economic Behavior* Vol 2: pp. 97 – 117.

BLAD, M. C. 1986. "A Dynamic Analysis of the Repeated Prisoner's Dilemma Game." *International Journal of Game Theory* Vol. 15, No. 2: pp. 83 – 99.

BOYD, R. AND J. P. LORBERBAUM. May 7, 1987. "No Pure Strategy is Evolutionarily Stable in the Repeated Prisoner's Dilemma Game." *Nature* Vol. 327: pp. 58 – 59.

CRAWFORD, V. P. 1992. "Adaptive Dynamics in Coordination Games." University of California, San Diego Discussion Paper 92-02.

GILBOA, I. AND D. SAMET. 1989. "Bounded versus Unbounded Rationality: The Tyranny of the Weak." *Games and Economic Behavior* Vol. 1: pp. 213 – 221.

HIRSHLEIFER, J. AND J. C. MARTINEZ COLL. June 1988. "What Strategies Can Support the Evolutionary Emergence of Cooperation." *Journal of Conflict Resolution* Vol. 32, No. 2: pp. 367 – 398.

KALAI, E. AND W. STANFORD. March, 1988. "Finite Rationality and Interpersonal Complexity in Repeated Games." *Econometrica* Vol 56, No. 2: pp. 397 – 410.

MILLER, J. July 10, 1989. "The Coevolution of Automata in the Repeated Prisoner's Dilemma." Santa Fe Institute Working Paper 89-003.

MUELLER, U. December 1987. "Optimal Retaliation for Optimal Cooperation." *Journal of Conflict Resolution* Vol. 31, No. 4: pp. 692 – 724.

NACHBAR, J. H. December 1988. "An Ecological Approach to Economic Games." mimeograph.

NACHBAR, J. H. December 1989. "The Evolution of Cooperation Revisited." mimeograph.

PRISBREY, J. 1992. "An Experimental Analysis of Two Person Reciprocity Games." Social Science Working Paper No. 787: California Institute of Technology.

RAPOPORT, A. 1988. "Editorial Comments on the Article by Hirshleifer and Martinez Coll." *Journal of Conflict Resolution* Vol. 32, No. 2: pp. 399 - 401.

RUBENSTEIN, A. 1986. "Finite Automata Play the Repeated Prisoner's Dilemma." *Journal of Economic Theory* Vol 39: pp. 83 – 96.

SAMUELSON, L. AND J. ZHANG. 1992. "Evolutionary Stability in Asymmetric Games." *Journal of Economic Theory* Vol 57: pp. 363 - 391.

SMALE, S. November 1980. "The Prisoner's Dilemma and Dynamical Systems Associated to Non-Cooperative Games." *Econometrica* Vol. 48 No. 7: pp. 1617 – 1634.

TAYLOR, P. D. AND L. B. JONKER. 1978. "Evolutionarily Stable Strategies and Game Dynamics." *Mathematical Biosciences* Vol. 40: pp. 145 – 156.

YOUNG, H. P. AND D. FOSTER. 1991. "Cooperation in the Short and Long Run." *Games and Economic Behavior* Vol. 3: pp. 145 – 156.

s_1 :	$\{a, a, a\}$	s_2 :	$\{b, a, a\}$
s_3 :	$\{a, b, b\}$	s_4 :	$\{b, b, b\}$
s_5 :	$\{a, a, b\}$	s_6 :	$\{b, a, b\}$
s_7 :	$\{a, b, a\}$	s_8 :	$\{b, b, a\}$

Table 1: The eight machines or strategies contained in S^B.

$$\Pi = \begin{bmatrix} \alpha & \alpha & \beta & \beta & \alpha & \alpha & \beta & \beta \\ \alpha & \alpha & \beta & \beta & \alpha & \alpha & \beta & \beta \\ \eta & \eta & \gamma & \gamma & \gamma & \gamma & \eta & \eta \\ \eta & \eta & \gamma & \gamma & \gamma & \gamma & \eta & \eta \\ \alpha & \alpha & \gamma & \gamma & \alpha & \frac{\beta+\eta}{2} & \frac{\alpha+\beta+\eta+\gamma}{4} & \frac{\alpha+\beta+\eta+\gamma}{4} \\ \alpha & \alpha & \gamma & \gamma & \frac{\eta+\beta}{2} & \gamma & \frac{\alpha+\beta+\eta+\gamma}{4} & \frac{\alpha+\beta+\eta+\gamma}{4} \\ \eta & \eta & \beta & \beta & \frac{\alpha+\beta+\eta+\gamma}{4} & \frac{\alpha+\beta+\eta+\gamma}{4} & \frac{\alpha+\gamma}{2} & \beta \\ \eta & \eta & \beta & \beta & \frac{\alpha+\beta+\eta+\gamma}{4} & \frac{\alpha+\beta+\eta+\gamma}{4} & \eta & \frac{\gamma+\alpha}{2} \end{bmatrix}$$

Table 2: The environment Π.

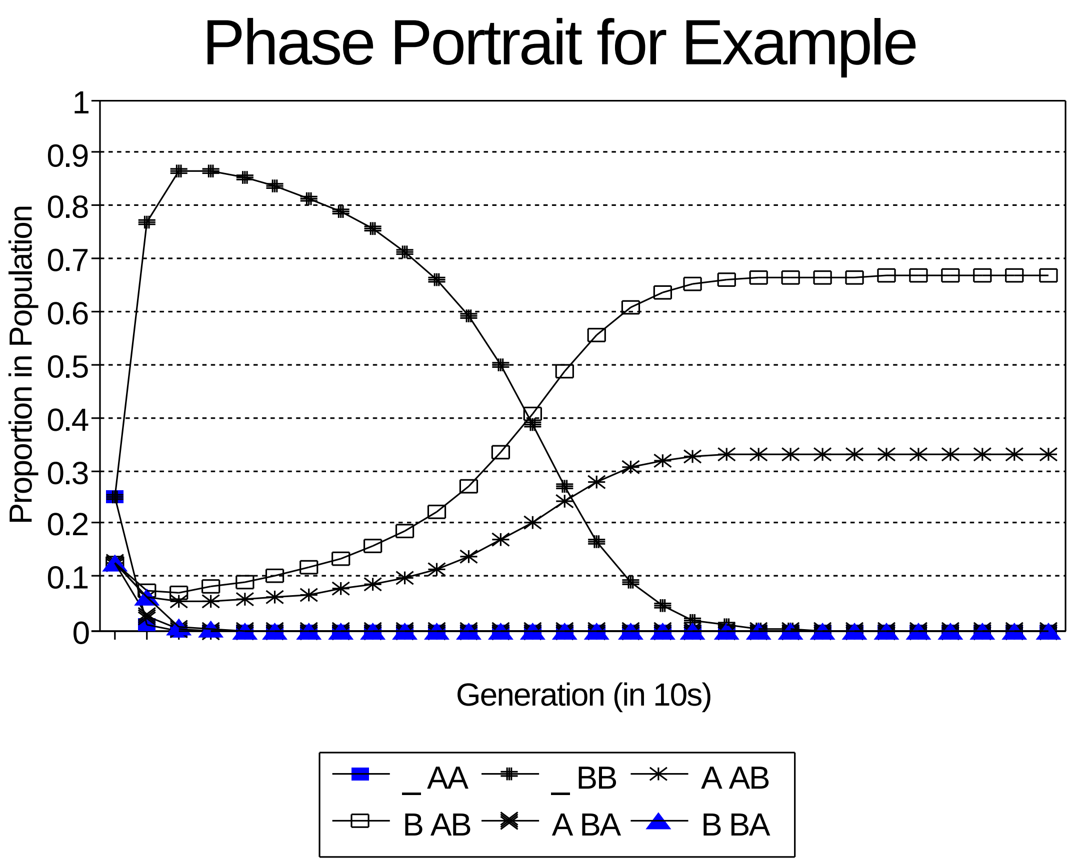

Figure 1: The phase portrait for payoff matrix M_1. The term _ AA stands for the sum of the representation of s_1 and s_2, _ BB is similar. The initial generation has all strategies equally represented in the population.

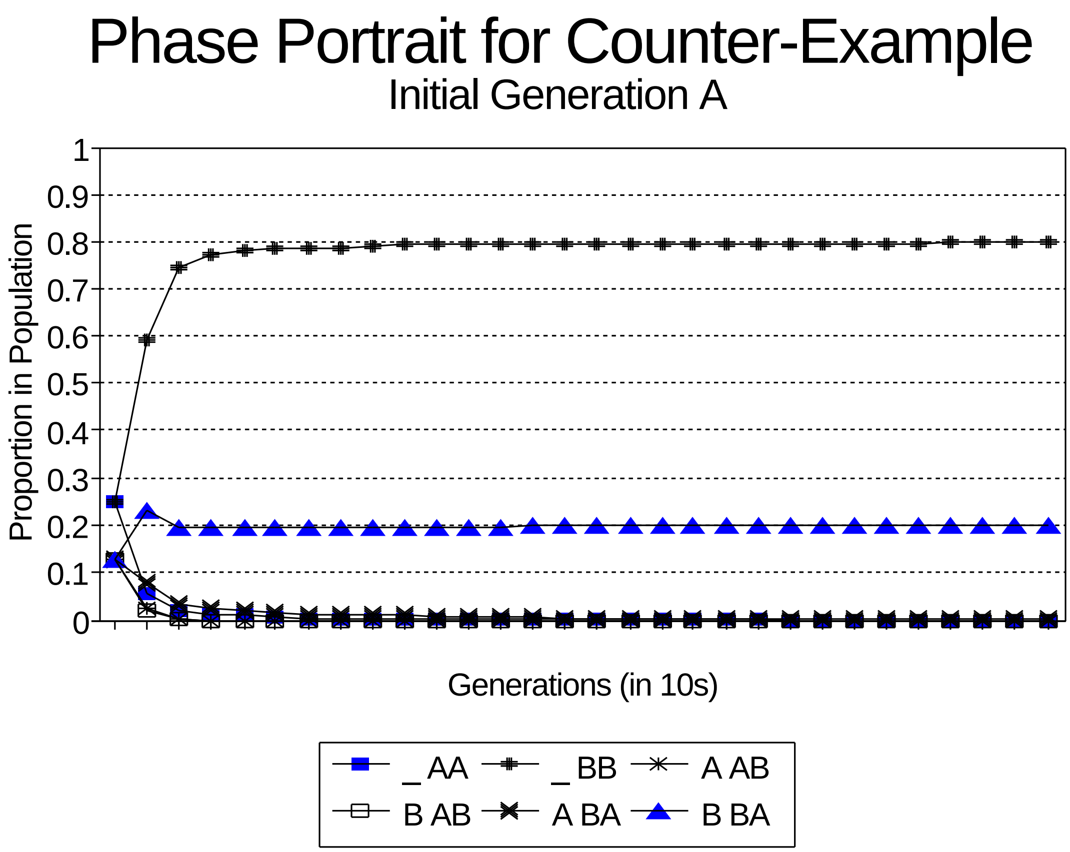

Figure 2: The phase portrait for payoff matrix M_2. The term _ AA stands for the sum of the representation of s_1 and s_2, _ BB is similar. The initial generation has all strategies equally represented in the population.

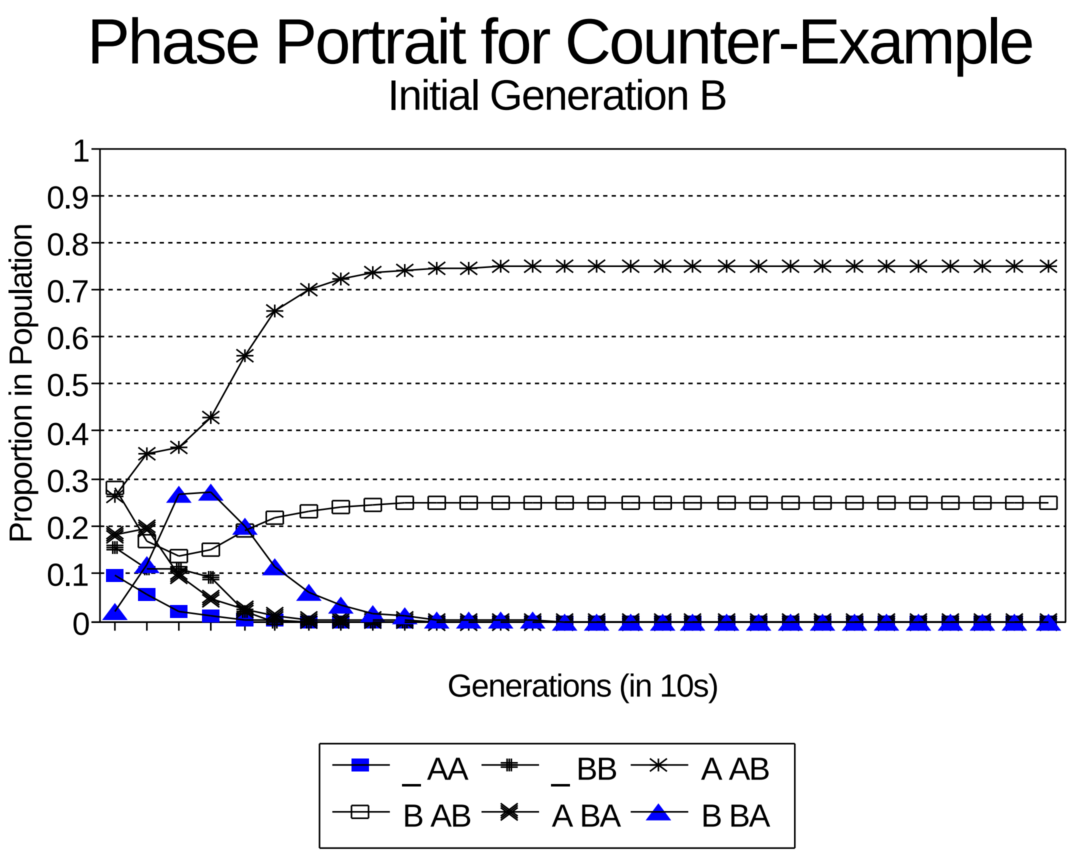

Figure 3: The phase portrait for payoff matrix M_2. The term _AA stands for the sum of the representation of s_1 and s_2, _BB is similar. The initial generation is $\boldsymbol{p}^0 = [0.048, 0.048, 0.078, 0.078, 0.264, 0.282, 0.182, 0.020]$.