Projektbereich B
Discussion Paper No. B–243

Dynamic Stability in the Repeated Prisoners' Dilemma Played by Finite Automata*

Karl H. Schlag

April 1993
Revised December 1993

Wirtschaftstheoretische Abteilung III
Friederich-Wilhelms-Universität Bonn
Adenauerallee 24–26
D–53113 Bonn
Germany
e-mail: schlag@glider.econ3.uni-bonn.de

* I would like to thank Avner Shaked for his useful comments. Financial support from the Deutsche Forschungsgemeinschaft, Sonderforschungsbereich 303 at the University of Bonn is gratefully acknowledged.

Abstract

We investigate the replicator dynamics of the repeated Prisoners' Dilemma played by finite automata. The players discount repeated game payoffs and incur a cost which is proportional to the number of states in the automaton they use.

An initial result is that the singleton set that contains "Defect for Ever" is the only asymptotically stable set containing a pure strategy.

We then search for asymptotically stable sets when the dynamics are restricted to initial distributions that contain some given types in their support. It is shown that "Tat for Tit" is the only pure strategy (up to look-a-likes) besides "Defect for Ever" that is contained in such a set when the discount factor is sufficiently close to one and the cost per state is arbitrarily small. "Tat for Tit" when playing against itself will defect first and then cooperate forever.

1. Introduction

The Prisoners' Dilemma is a symmetric two person game with two strategies C and D (C="cooperate" and D="defect") where the payoffs to the row player $\pi:\{C,D\}\times\{C,D\}\rightarrow\Re$ are such that $\pi(D,C)>\pi(C,C)>\pi(D,D)>\pi(C,D)$. We will sometimes assume $2\pi(C,C)>\pi(D,C)+\pi(D,D)$ and/or $\pi(D,C)-\pi(C,C)>\pi(D,D)-\pi(C,D)$. Table I displays a possible constellation of the payoffs in which all conditions hold, this example will be used for illustrative purposes throughout this paper.

Table I: An example for payoffs $\pi()$ of the Prisoners' Dilemma, C="cooperate", D="defect".

	C	D
C	3, 3	−1, 5
D	5, −1	0, 0

We study the repeated Prisoners' Dilemma played by finite automata (or Moore machines, for an introduction see Hopcraft and Ullman [11]). The underlying game we consider fits into the framework developed by Rubinstein [15] and continued by Abreu and Rubinstein [1]. Two players choose each an automaton and thereafter the two automata play the repeated game against each other on behalf of the players. The individuals have preferences over the repeated game payoffs and the complexity of the automaton they choose.
This complexity is measured by the number of states of the automaton. The resulting game will be referred to as the meta game.

Abreu and Rubinstein [1] characterize the pure strategy Nash equilibrium payoffs for a general class of preferences in the meta game. In particular they consider lexicographic preferences where players prefer higher repeated game payoffs (measured by the limit of the means payoff criterion) and when indifferent prefer automata with less states. We will call these preferences the patient lexicographic preferences. Abreu and

Rubinstein [1] show that under these preferences the individual payoffs to pure strategy symmetric Nash equilibria lie dense in the interval $[\pi(D,D),\pi(C,C)]$.

We aim to analyze the meta game in an evolutionary framework. Maynard Smith and Price [13] first developed the concept of an Evolutionarily Stable Strategy (ESS) as an intuitive condition for a strategy to survive in an evolutionary process. Later it was shown that an ESS is an asymptotically stable strategy in the so called replicator dynamics (see Taylor and Jonker [19], Zeeman [22]). The fact that an ESS might not always exist has since led to various weaker concepts. One area of research is to find a weaker concept that is still sufficient for asymptotic stability of a strategy or a set of strategies with respect to the replicator dynamics (see Thomas [20], Schlag [16]). Other undertakings weaken the concept on intuitive grounds, without considering the replicator dynamics (e.g., Binmore and Samuelson [5], Swinkels [18]).

Central to the derivation of the replicator dynamics is a biologically motivated scenario of reproduction. Recently several models of interacting and learning agents have been constructed which show that the replicator dynamics also appear in economic contexts. While Binmore and Samuelson [6] and Börgers and Sarin [7] explicitly specify individual behavior, Schlag [17] implicitly determines which adjustment rules agents will use. Although very different in their setup, each of these models leads to dynamic processes that approximate the replicator dynamics.

The paper by Binmore and Samuelson [5] presented the initial spark for this research. They apply static evolutionary stability theory to the meta game assuming lexicographic preferences. Binmore and Samuelson [5] generalize the ESS condition to incorporate for these lexicographic preferences and define the concepts modified ESS (MESS) and polymorphous MESS. Their concepts are meant as an intuitive condition for a strategy or collection of strategies to survive "some" evolutionary process. However a direct connection to a dynamic adjustment process is not presented. On closer examination, various difficulties arise when trying to incorporate lexicographic preferences into such continuous dynamic systems. In section 6 this point will be elaborated on and the results of Binmore and Samuelson [5] will be compared to ours.

We will consider a slightly different payoff function in the meta game. The payoff a player receives (denoted by $E_{\delta,c}(\ )$) is the repeated game payoff minus a cost due to the degree of complexity of the automaton he uses. The repeated game payoff will be measured as the normalized expected value of the discounted future payoffs (discount factor δ, $\delta<1$), the cost of complexity of an automaton will be proportional to the number of the states it uses (c will be the cost per state, $c>0$). The constants δ and c will be referred to as the parameters.

We extend the results of Abreu and Rubinstein [1] to the meta game with preferences defined by this payoff function. We show that the individual payoffs to symmetric Nash equilibria of the meta game converge to and become dense in the interval $[\pi(D,D),\pi(C,C)]$ as the discount factor and the cost of complexity tend to 1 and 0 respectively.

The focus of this paper will be to analyze how the above static result without evolution relates to the strategies or sets of strategies that may have nice properties in an evolutionary process. Instead of defining or adapting an intuitive static concept we will perform a direct analysis of the stability of a population evolving according to the replicator dynamics with respect to rare mutations. We will combine sufficient conditions for stability and asymptotic stability like weak ESS and ESS with the direct dynamic analysis of the trajectories of the replicator dynamics. Along the way various useful dynamic stability concepts will be introduced. Features that will make the analysis interesting from the dynamic viewpoint will be the countable number of pure strategies of the meta game (automata) and the multiplicity of automata that have the same equilibrium path.

The replicator dynamics were originally derived from the approximation of the following discrete time process. Consider an infinite population of individuals, each endowed with a pure strategy (in our case an automaton). Individuals are pair-wise randomly matched to play the repeated Prisoners' Dilemma. The discount factor δ introduced earlier has a slightly different interpretation in this setup. In each round of the repeated game, $(1-\delta)$ is the conditional probability that the repeated game is over for all players simultaneously given that the game has lasted up to that round. Because δ is assumed to be strictly smaller than one, the expected number of rounds the repeated game lasts is finite. So in this context, δ can be considered the continuation probability. However, to stay in the framework of static game theory we will continue to call δ the discount factor. After the repeated game is over, the individuals reproduce proportionally to the relative payoff they received in the repeated game. After that they die and their offspring is randomly matched to play the repeated game, and so on. For a more detailed introduction to and derivation of the replicator dynamics, see (van Damme [21]). For some derivations of this process in economic models, see [6,7,17].

We will focus on the stability properties of sets instead of that of individual strategies. In this context a set of population distributions (i.e., mixed strategies) L will be said to "resist" small mutations if after an arbitrarily but sufficiently small frequency of mutants enters the population, the population distribution stays close to the initial population distribution and eventually evolves back to a distribution in the set L. In the framework of dynamical systems this is equivalent to searching for asymptotically stable sets of the replicator dynamics. An asymptotically stable set is a closed set of strategies such that trajectories starting close stay close and eventually converge to an element of the set. An asymptotically stable strategy is a singleton asymptotically stable set.

In order to simplify the discussion we will call a strategy that is associated with a symmetric Nash equilibrium a symmetric Nash equilibrium strategy. Our initial interest is to find out which of the unbounded number of symmetric Nash equilibrium strategies characterized by Abreu and Rubinstein [1] are in such a set (elements of asymptotically stable sets are Nash equilibrium strategies). Therefore we will focus our analysis on such sets that contain a monomorphic population (i.e., a pure strategy).

Consider the automaton that always plays "D" called "Defect for Ever" (see figure 1).

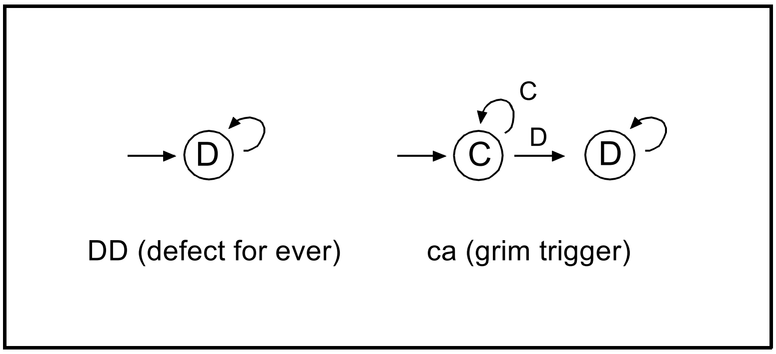

Figure 1: The automata "Defect for Ever" and "Grim Trigger".

For any

δ<1 and c>0, the characteristics of the Prisoners' Dilemma imply that ("Defect for Ever", "Defect for Ever") is a strict Nash equilibrium. Therefore "Defect for Ever" is an ESS and hence an asymptotically stable strategy (theorem A1). We show that {"Defect for Ever"} is the only asymptotically stable set containing a pure strategy when δ<1 and c>0. The reason for its uniqueness is illustrated in the following example that contains the basic intuition and structure of the rest of the paper.

Consider a population consisting of types "Tat for Tit" (CC), CA and DCC - we will be using the notation from [5] whenever applicable. These strategies are graphically represented in figure 2.

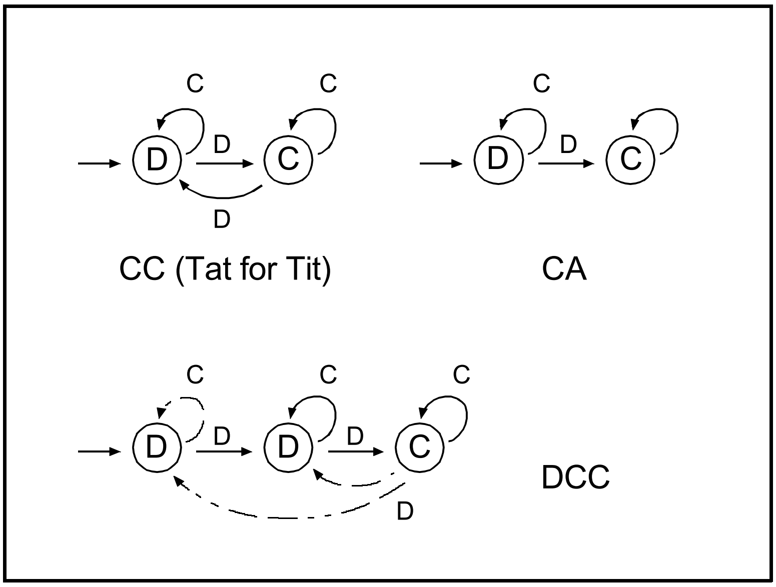

Figure 2: The automata "Tat for Tit" (CC), a toothless look-a-like (CA) and the "automaton" DCC that can take advantage of CA.

The automata CA and CC are quite similar, when playing among themselves will defect first and then cooperate forever. Additionally they both take advantage of the very simple automata that always cooperates. However, unlike "Tat for Tit", CA does not protect itself from deviations in the cooperation cycle, hence Binmore and Samuelson [5] call it toothless. Automata that only differ on transitions not used when matched against each other will play an important role in our analysis. They will be called look-a-likes, hence CA and CC are look-a-likes. DCC represents actually the class of all automata that have the basic structure graphed in figure 2, the transitions with dotted lines are optional.

These automata take advantage of the toothlessness of CA and cooperate for ever against CC. The payoffs to the row player among these three strategies in the example of table I are given in table II.

Table II: Payoffs $E_{\delta,c}(\ )$ to the row player in the repeated Prisoners' Dilemma of table I with the strategies CC, CA and DCC.

	CC	CA	DCC
CC	$3\delta-2c$	$3\delta-2c$	$-(1-\delta)\delta+3\delta^3-2c$
CA	$3\delta-2c$	$3\delta-2c$	$-\delta-2c$
DCC	$(1-\delta)5\delta+3\delta^3-3c$	$5\delta-3c$	$3\delta^2-3c$

Notice that CC cannot be in an asymptotically stable set. Every strategy with support in {CC,CA} is a fixed point of the dynamics and hence must be in an asymptotically stable set containing CC, especially CA must be in this set. On the other hand, elements of an asymptotically stable set must be a best response to themselves (see lemma A4 in appendix) and CA is not one ($E_{\delta,c}(DCC,CA)>E_{\delta,c}(CA,CA)$). So toothless look-a-likes deter asymptotic stability.

It turns out that "Defect for Ever" is the only symmetric Nash equilibrium strategy without toothless look-a-likes and hence as a singleton set the unique asymptotically stable set containing a pure strategy.

We are interested in preferences that approach the patient lexicographic preferences. However when the parameters (δ,c) are sufficiently close to (1,0), we show that "Grim Trigger" can take over in any neighborhood of "Defect for Ever". Therefore one might say that the ability of "Defect for Ever" to resist arbitrary mutations vanishes when this limit is taken. The automaton "Grim Trigger" (see figure 1) cooperates in each round until its opponent defects. From then on it defects for ever.

On closer examination, the requirement that a set must resist an arbitrary mutation is too stringent for the setup in this paper. Although the evolutionary model associated with the replicator dynamics assumes a one time entrance of mutants, the intuition follows more the line that mutation pressure will be sufficiently small relative to the dynamic selection pressure. Schlag [16] demonstrates that this more general model can be incorporated into the present setup as long as each mutation is sufficiently small. How small each mutation must be depends on the current population distribution.

In continuous dynamics like the replicator dynamics a type never vanishes although its frequency might tend to zero over time. So if the set of strategies is not too large, after a sufficient amount of time we might expect that every type will be present in the population. Our strategy space (denoted by S_p), the set of all automata for playing the repeated Prisoners' Dilemma, has a countable infinite number of elements. Here it is not so obvious that <u>all</u> strategies will be present although a certain number will be. Following these remarks, an analysis of trajectories that contain a pre-specified set becomes interesting. Certain strategies with very small frequencies will receive a very important role in the analysis. As we will see later, it can happen that a set of population distributions can only resist an arbitrary mutation if a certain type (or set of types) is present in the current mutation or was present in an earlier mutation.

We will adapt the definition of asymptotic stability to this weaker condition, fixing some set of types and only considering populations whose support contains that set. We will call these sets asymptotically stable given previous intruders. We show that if such a set is connected then it is a maximal connected set of stable equilibria.

In the following we must be more specific about the values of the parameters δ and c we want to consider. We want our analysis to go in the line of the patient lexicographic preferences approach of Abreu and Rubinstein [1]. What is the intuition behind these preferences? First of all, the repeated game payoffs should be the main impact on the payoff of the meta game and only close to a "tie" should complexity considerations play a role. Following this intuition for any fixed discount value we will analyze the game for any (strictly) positive cost below some cap. Secondly, the emphasis is on extremely patient players. We too will focus on the repeated game effects and consider the case where the discount factor δ is very close to one. To simplify notation, a subset of the parameters that is compatible with these two conditions will be called lex-patient.

We will find a lex-patient set such that "Tat for Tit" is contained in an asymptotically stable set given previous intruders for each parameter in the set. The condition of previous intruders will require that DCC (i.e., one out of its class) is contained in the support of the mutation.

To illustrate this result consider a population consisting of types CC, CA and (a type of the class of) DCC. Here the previous intruders condition is equivalent to assuming interior mutations, i.e., restricting the analysis to interior trajectories. We will show that there exists a lex-patient set such that if γ is the maximal frequency of CA in a symmetric Nash equilibrium strategy with support in {CA,CC} then $\{(1-\alpha)CC+\alpha CA,\ 0\leq\alpha\leq\gamma\}$ is stable and attracting w.r.t. interior trajectories in $\Delta\{CC,CA,DCC\}$. The phase diagram is sketched in figure 3.

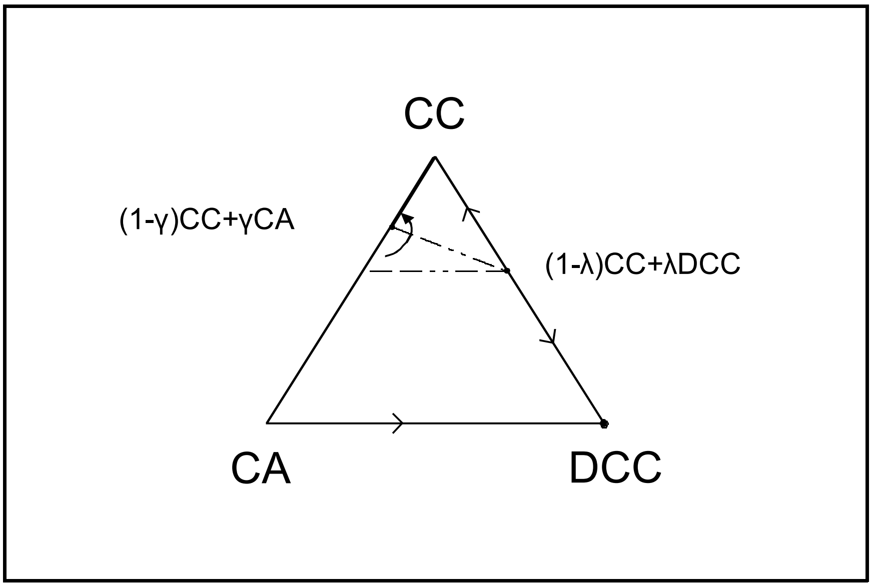

Figure 3: Qualitative phase diagram of the population consisting of the automata CC, CA and DCC for sufficiently large $\delta<1$ and $0<c<(1-\delta)[\pi(D,D)-\pi(C,D)]$.

Of course "Defect for Ever" is also contained in an asymptotically stable set given previous intruders. However we show that unlike "Defect for Ever" the ability of the asymptotically stable set given previous intruders containing "Tat for Tit" to resist arbitrary mutations does not vanish as (δ,c) goes to $(1,0)$.

Under an additional condition on the payoffs of the Prisoners' Dilemma we show that there is no other pure strategy that is contained in an asymptotically stable set given previous intruders for all parameters in some lex-patient set. Notice that in our above example DCC is an asymptotically stable strategy in a population consisting of types CC, CA and DCC. However we will find other strategies whose existence hinder (any type in) DCC from being in an asymptotically stable set given previous intruders when (δ,c) is close to (1,0).

How is this result proven? Connected asymptotically stable sets given previous intruders must be maximal connected sets of stable equilibria since stable means that trajectories starting close stay close. On the border of a maximal connected stable set containing a pure strategy other than "Defect for Ever" or a look-a-like of CC, toothless look-a-likes can gang up with mutants to get higher payoffs than the pure strategy by skipping its initial phase. This causes the look-a-likes to reproduce faster than the pure strategy and thus create an unstable population.

Hence under certain conditions on the payoffs of the Prisoners' Dilemma "Defect for Ever" and "Tat for Tit" with its look-a-likes are the only strategies contained in an asymptotically stable set given previous intruders for all parameters in a lex-patient set.

2. Preliminaries

In this section we will introduce the replicator dynamics, some general dynamic stability concepts alongside some basic notation. Consider a symmetric two person game $\Gamma(S,E)$ with the countable set of pure strategies $S:=\{e^i, i=1,2,...\}$ and the bilinear payoff function $E:\Delta S\times\Delta S\rightarrow\Re$ where ΔS is the set of probability distributions on S, i.e., $\Delta S:=\{x=(x_i)_{i\in\mathbb{N}} \text{ s.t. } x_i\geq 0 \text{ and } \sum_{i=1}^{\infty} x_i=1\}$, where $x_i=x(e^i)$ is the frequency of e^i. For $x\in\Delta S$ let R(x) be the set of pure strategy best responses to the strategy x, i.e., $R(x):=\{e^i\in S \text{ s.t. } E(e^i,x)\geq E(z,x) \text{ for all } z\in\Delta S\}$. For $x\in\Delta S$ let C(x) be the support of x, i.e., $C(x):=\{e^i\in S \text{ s.t. } x_i>0\}$. To simplify notation we will make no difference between the pure strategy $e\in S$ and the distribution in ΔS assigning probability one to e (i.e., $S\subseteq\Delta S$). We will sometimes call a strategy $x\in\Delta S$ such that $C(x)\subseteq R(x)$ a <u>symmetric Nash equilibrium strategy</u> since (x,x) is a Nash equilibrium under these assumptions.

The replicator dynamics (RD) on ΔS for continuous time and pure strategy types is as follows (see [19,22]):

$$x^0=\bar{x} \text{ and } \frac{d}{dt}x^t_i=[E(e^i,x^t)-E(x^t,x^t)]x^t_i,\ i=1,2,...,\ t\geq 0,$$

where $\bar{x}\in\Delta S$ is the initial state and x^t_i is the frequency of the type using strategy e^i ($e^i\in S$) at time t. It can be shown that for each $\bar{x}\in\Delta S$ the above differential equation defines a unique function $x:\Re^+\rightarrow\Delta S$. To simplify notation we will drop the parameter t from the expressions (e.g., $x=x^t$).

As a measure of distance in ΔS we will consider the l^1 norm, i.e., $dist(x,y):=\sum_{e\in S}|x(e)-y(e)|$ where $|z|$ denotes the absolute value of $z\in\Re$. We choose the l^1 norm because then the distance a monomorphic population is "moved" by a mutation of mass ϵ only depends on ϵ. For any $\epsilon>0$ and $x\in\Delta S$ let $U_\epsilon(x)$ be the open ball of radius ϵ around x, i.e., $U_\epsilon(x)=\{y\in\Delta S \text{ s.t. } dist(x,y)<\epsilon\}$.

An important characteristic of the replicator dynamics that we will use later is the continuity of its gradients. This property implies that given any trajectory leading from a to b ($a,b\in\Delta S$) and any $\epsilon>0$ there exists a strategy $a'\in U_\epsilon(a)$ such that the trajectory starting in a' intersects $U_\epsilon(b)$.

The subsequent dynamic stability concepts will be relevant for our analysis. A set $L\subseteq\Delta S$ is called <u>attracting</u> if there exists an open neighborhood W of L such that each trajectory starting in W converges to L ($W\subseteq\Delta S$). The maximal such set W will be called the <u>basin of attraction of L</u>. A strategy $p\in\Delta S$ is called <u>stable</u> if for every open neighborhood U of p there exists an open neighborhood V of p such that the trajectories starting in V do not leave U. A set $L\subseteq\Delta S$ is called <u>stable</u> if each $p\in L$ is stable. A strategy $p\in\Delta S$ is called <u>unstable</u> if it is not stable. A set $L\subseteq\Delta S$ is called an <u>asymptotically stable set</u> if it is closed, attracting and stable. A singleton asymptotically stable set is called an <u>asymptotically stable strategy</u>.

In the following we add some notes on the above definitions. A trajectory starting in X converges to L ($L,X\subseteq\Delta S$) if for any $\overline{x}\in X$ and any sequence $(t_k)_{k\in\mathbb{N}}$ converging to infinity ($t_k\in\Re$), $\inf\{dist(x^{t_k},z), z\in L\}\rightarrow 0$ as $k\rightarrow\infty$ where x^t solves the replicator dynamics starting at $x^0=\overline{x}$. Trajectories starting in the basin of attraction of an attracting set must not converge to a single point. However if the set is also stable then each element is and therefore trajectories will converge to an element of the set. The definition of stability is slightly stronger than the classical one (see e.g. Bhatia and Szegö [4]): in the standard definition the set as a whole must be stable, not necessarily each point. The standard definition of stability was modified following the intuition given in the introduction of what it means that a set of population distributions can "resist" small mutations. Since the definition of stability and that of attracting is preserved under closure w.l.o.g. we require additionally to the standard definition (Bhatia and Szegö [4]) for an asymptotically stable set to be closed. We find it intuitive to include rest points on the border of the set into the definition of an asymptotically stable set. Finally, notice that a stable strategy must be a rest point.

Since an asymptotically stable set does not have arbitrarily close rest points outside the set we immediately obtain the following lemma.

LEMMA 2.1:

Connected asymptotically stable sets are maximal connected sets of rest points.

3. The analysis

The goal of this section is to characterize the asymptotically stable sets containing a pure strategy of the repeated Prisoners' Dilemma played by finite automata.

Let S_p be the set of all finite automata for playing the repeated Prisoners' Dilemma (which is countable) and $E':\Delta S_p\times\Delta S_p\rightarrow\Re$ be such that $E'(x,y)$ is the normalized expected value of the discounted future payoffs to the player using x when matched against a player using y ($x,y\in\Delta S_p$, discount rate $\delta<1$). More specifically, let e^k be matched against e^j ($e^k,e^j\in S_p$) and let $z^t_k,z^t_j\in\{C,D\}$ be such that the player using e^k (e^j) plays z^t_k (z^t_j) in the t-th round, t=1,2,.. . Then $E'(e^k,e^j):=(1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}\pi(z^t_k,z^t_j)$ and $E'(\ )$ can be extended to $\Delta S_p\times\Delta S_p$ as a bilinear function. Alternatively let δ be the continuation probability, i.e., δ is the probability that the game is played in round r+1 given that it has lasted up to round r (r=1,..). Then $E'(\ )$ can be considered the normalized expected aggregated payoff.

The payoff function of the meta game of selecting an automaton to play the repeated Prisoners' Dilemma, denoted by $E(\ )$, will be the normalized expected value of the repeated game payoffs minus a fixed cost for each state the chosen automaton uses. Formally, for $(\delta,c)\in(0,1)\times(0,1)$ let $E_{\delta,c}:\Delta S_p\times\Delta S_p\rightarrow\Re$ such that for $e^j\in S_p$ and $y\in\Delta S_p$, $E_{\delta,c}(e^j,y):=E'(e^j,y)-|e^j|c$ where $|e^j|$ denotes the number of states in e^j and $c>0$ the cost of a state. Again this can be extended to $\Delta S_p\times\Delta S_p$. Notice that since $E_{\delta,c}(\ )$ is continuous in (δ,c) and the limit exists as $(\delta,c)\rightarrow(1,0)$, $E_{1,0}(\ )$ is well defined.

The number of states an automaton has is considered a one time investment and hence this cost is assumed to be separable from the repeated game payoffs $E'(\ )$. The assumption that cost is linear in the number of states is just to simplify notation and does not enter the results. The constants δ and c will be referred to as the parameters of the game. To make clear which parameters are considered we will write $R_{\delta,c}(x)$ for the set of pure strategy best responses.

We will start out by quoting the relevant results of Abreu and Rubinstein [1], characterizing the pure symmetric Nash equilibria in the repeated Prisoners' Dilemma played by finite automata.

Fix $\delta<1$ and $c>0$. The following statements are easy to check. An automaton e^j

matched against itself has an initial phase followed by a cycle. So if the initial phase has k states and the cycle has l states then $k+l \leq |e^j|$ ($k,l \in \mathbb{N}$). The corresponding individual payoff is a convex combination of $\pi(D,D) - |e^j|c$ and $\pi(C,C) - |e^j|c$. If e^j is a pure symmetric Nash equilibrium strategy then it uses each of its states in some round when matched against itself, i.e., $e^j \in R_{\delta,c}(e^j)$ implies $k+l = |e^j|$. Furthermore a Nash equilibrium strategy maximizes repeated game payoffs against itself, i.e., if $e^j \in R_{\delta,c}(e^j)$ then $e^j \in R_{\delta,0}(e^j)$.

The next theorem characterizes the set of Nash equilibrium payoffs and follows almost directly from one stated in (Abreu and Rubinstein [1]). It states for discounted payoffs and linear costs that the set of individual payoffs to pure strategy symmetric Nash equilibria becomes dense in the interval $[\pi(D,D),\pi(C,C)]$ as the discount factor and the cost of complexity tend to one and zero respectively.

THEOREM 3.1:

Let $(\delta_k,c_k)_{k\geq 0}$ be a sequence such that $(\delta_k,c_k) \in (0,1)\times(0,1)$ and $(\delta_k,c_k) \to (1,0)$ as $k \to \infty$ and let $D(\delta,c) = \{E_{\delta,c}(e^j,e^j) \text{ s.t. } e^j \in R_{\delta,c}(e^j)\}$. Then

$$\sup_y \inf_x \{\mathrm{dist}(x,y),\ x \in D(\delta_k,c_k),\ y \in [\pi(D,D),\pi(C,C)]\} \to 0 \text{ as } k \to \infty.$$

PROOF:

Abreu and Rubinstein [1] show for the "limit of means" payoff criterium and lexicographic preferences that every rational convex combination of $\pi(D,D)$ and $\pi(C,C)$ can be reached as individual payoff of a pure strategy Nash equilibrium. We will adopt their proof directly. Consider the strategies that Abreu and Rubinstein [1] use. They are of the form graphed in figure 4, each block represents a finite sequence of states prescribing the play written in the center of the block.

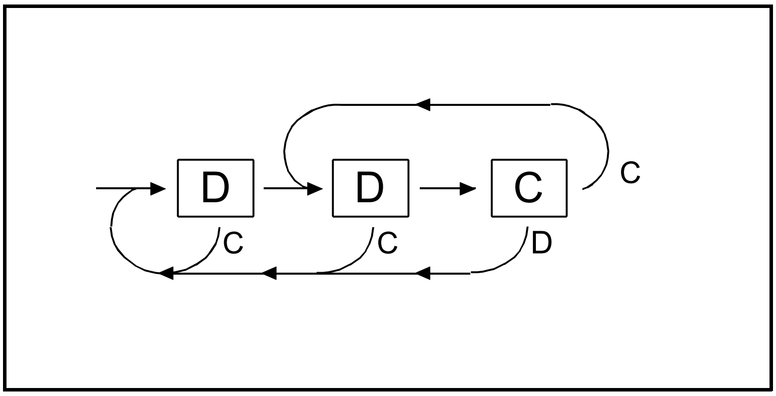

Figure 4: Basic features used to construct symmetric Nash equilibrium strategies.

It follows that for any given number of strategies in the cycle if the initial phase is long enough then there exists an $\epsilon>0$ such that the corresponding automaton is a symmetric Nash equilibrium strategy for any $(\delta,c)\in(0,1)^2\cap\{(\delta',c') \text{ s.t. } 1-\delta'+c'<\epsilon\}$. The statement of theorem 3.1 now follows directly from the fact that the payoff function $E_{\delta,c}(\ )$ of the meta game approximates the limit of the means payoff as δ goes to 1 and c goes to 0.

□

Next we will present a necessary condition for an automaton to be a pure symmetric Nash equilibrium strategy. The theorem states that such an automaton must start out by playing "D" (defection).

THEOREM 3.2:

Let $\delta<1$, $c>0$ and $e^j\in S_p$. If (e^j,e^j) is a Nash equilibrium then e^j plays "D" in the first round.

PROOF:

For $a,b\in S_p$ let a_t (b_t) be the strategy that a (b) plays in round t when a and b are matched ($a_t,b_t\in\{C,D\}$, $t=1,2,..$). For $T\geq 1$ let $P_{\delta,c}^T(a,b):=(1-\delta)\sum_{t=T}^{\infty}\delta^{t-T}E_{\delta,c}(a_t,b_t)$ be the continuation payoff starting at round T (following Binmore and Samuelson [5]). Let (e^j,e^j)

be a pure strategy Nash equilibrium ($e^j \in S_p$) and let T* be the first round in which $P_{\delta,c}{}^T(e^j,e^j)$ achieves its minimum. Because (e^j,e^j) is a Nash equilibrium and c>0 it follows that e^j uses each state when matched against itself. By the minimality of $P_{\delta,c}{}^{T*}(e^j,e^j)$ the strategy played at round T* must be self enforcing, i.e. e^j plays D in round T*.

Assume that e^j starts out by cooperating. Then $T^* \geq 2$. Notice that e^j cooperates at time T^*-1 when matched against itself, otherwise $P_{\delta,c}{}^T(e^j,e^j)$ would achieve its minimum at time T^*-1.

On the other hand similar to the above argument that e^j cannot protect its action at time T*, e^j cannot protect its action at time T^*-1. Therefore the strategy played in round T^*-1 must be self enforcing too. This contradicts the fact that e^j cooperates at time T^*-1 and hence T*=1.

□

Looking back at the proof we showed that any pure symmetric Nash equilibrium strategy achieves its minimal continuation payoff at time 0. Therefore given any pure symmetric Nash equilibrium strategy, if each defection from the equilibrium path is punished back to the first state then this altered strategy is a symmetric Nash equilibrium strategy too. Defecting from the equilibrium path means that the opponent plays a different action than the pure symmetric Nash equilibrium strategy does.

When an automaton is matched against itself not every transition or even every state will be used. Automata that cannot be distinguished when matched among themselves will play a very important part in this paper.

For $e^j \in S_p$ let $Q(e^j)$ be the set of all automata that cannot be distinguished from e^j w.r.t. the play in the repeated game when matched against e^j. Formally $Q(e^j) := \{e^k \in S_p$ such that when e^k is matched against e^j, they both play the same strategy in each round$\}$. A strategy $e \in Q(e^j)$ such that $|e^j|=|e|$ will be called a <u>look-a-like</u> of e^j. Notice that if $e^k \in Q(e^j)$ then $Q(e^k)=Q(e^j)$. So "look-a-like" is a mutual property - look-a-likes essentially have the same states and differ only on transitions not used against each other. For $(\delta,c) \in (0,1) \times (0,1)$ we will call $x \in \Delta S_p$ a (symmetric) Nash equilibrium look-a-like of e^j if x is a symmetric Nash equilibrium strategy whose support only contains look-a-likes of e^j, i.e., $C(x) \subseteq Q(e^j) \cap R_{\delta,c}(x)$.

The next lemma demonstrates the important role of look-a-likes: an asymptotically stable set that contains a pure strategy must also contain all of its look-a-likes.

LEMMA 3.3:

Let $\delta<1$ and $c>0$. If $L\subseteq\Delta S_p$ is an asymptotically stable set that contains $e^j\in S_p$ then $\{x\in\Delta S_p \text{ s.t. } C(x)\subseteq Q(e^j)\}\subseteq L$.

PROOF:

Since $\{x\in\Delta S_p \text{ s.t. } C(x)\subseteq Q(e^j)\}$ is connected and any $x\in\Delta S_p$ such that $C(x)\subseteq Q(e^j)$ is a rest point, following lemma 2.1 the proof is complete.

□

Part of the next theorem follows directly from this lemma and lemma A4. Each element of an asymptotically stable set must be a symmetric Nash equilibrium strategy. But not every look-a-like of an automata with a cooperation state in its equilibrium path is one. Therefore the only candidate for a pure strategy contained in an asymptotically stable set is "Defect for Ever" (see figure 1).

THEOREM 3.4:

Let $\delta<1$ and $c>0$. "Defect for Ever" is an ESS and as a singleton set the only asymptotically stable set containing a pure strategy.

PROOF:

a) Since (DD,DD) is a strict Nash equilibrium ($c>0$), "Defect for Ever" is an ESS and by theorem A1 it follows that it is an asymptotically stable strategy.

b) Let $L\subseteq\Delta S_p$ be an asymptotically stable set containing $e^i\in S_p$. Assume that $e^i\neq$"Defect for Ever". Then $e^i\neq$"cooperate for ever" too since this is not a symmetric Nash equilibrium strategy (part ii) of lemma A4).

Every symmetric Nash equilibrium strategy with at least two states has a cooperation state since $c>0$, hence also e^i. Consider the look-a-like e of e^i that stays in its cooperation state whenever the opponent plays "D". Consequently e is not a symmetric Nash equilibrium strategy and therefore cannot be in L. On the other hand $e \in Q(e^i)$ and by lemma 3.3, e must be in L. This results in a contradiction and hence proves the theorem. □

We would like to examine the basin of attraction of "Defect for Ever". In the context of a population resisting arbitrary mutations we are interested in the maximal size of a neighborhood of "Defect for Ever" that is defined independently of the other types present in the population and which is contained in its basin of attraction. We will show that any open ball contained in the basin of attraction of "Defect for Ever" shrinks to the empty set as (δ,c) approaches (1,0).

Consider a population consisting of "Defect for Ever" and "Grim Trigger". These strategies are graphed in figure 1 and the payoffs they receive against each other in the example of table I are written in table III.

Table III: Payoffs to the row player in the repeated Prisoners' Dilemma of table I between the strategies DD ("Defect for Ever") and ca ("Grim Trigger").

	DD	ca
DD	$0-c$	$5(1-\delta)-c$
ca	$-(1-\delta)-2c$	$3-2c$

A frequency of more than

$$\frac{[\pi(D,D)-\pi(C,D)](1-\delta)+c}{[\pi(D,D)-\pi(C,D)+\pi(C,C)-\pi(D,C)](1-\delta)+[\pi(C,C)-\pi(D,D)]\delta}$$

"Grim Trigger" types (referring to the parameters in table I, more than $(1+c-\delta)/(4\delta-1)$) will disrupt the monomorphic population "Defect for Ever" and the population will converge to a population consisting of only "Grim Trigger" types. As (δ,c) goes to (1,0) less and less of "Grim Trigger" are necessary to disrupt "Defect for Ever".

We therefore obtain the following result.

THEOREM 3.5:

For $\delta<1$ and $c>0$ let $\epsilon(\delta,c)>0$ be such that trajectories starting in $U_{\epsilon(\delta,c)}(DD)$ converge to DD. It follows that $\epsilon(\delta,c)\to 0$ as $(\delta,c)\to(1,0)$.

The proof follows directly from the statement made above. Notice that by the continuity of the replicator dynamics the theorem is not changed if the trajectories are restricted to the interior of ΔS_p.

So the asymptotic stability of "Defect for Ever" depends crucially on (δ,c), the closer the parameters are to $(1,0)$, the smaller a mutation must be that "Defect for Ever" can resist. In the limit its asymptotic stability vanishes.

This unfavorable property causes us to pursue our analysis of the replicator dynamics to find strategies with properties that are more "stable" with respect to the parameters.

4. Interior mutations

In this section we will introduce some additional concepts in order to be able to analyze populations that can "resist" any mutation containing a pre-specified set of types, say $J\subseteq S$. We will adapt the dynamic stability concepts to this alternative setup in the following way. In the neighborhoods in the definitions of stability, attracting and asymptotic stability we will only consider strategies with support containing J. This is well defined concept since the support of a strategy remains the same if it is evolving according to the replicator dynamics. These alternative definitions will get the suffix "<u>given previous intruders (in J)</u>". If various different types in X ($X\subseteq S$) are each sufficient as a previous intruder we will add the suffix "<u>given previous intruders containing a type in X</u>. As an example, the set $L\subseteq\Delta S$ is called <u>attracting given previous intruders in J</u> if there exists an open neighborhood W of L such that each trajectory starting in $x^\circ\in W$ with $J\subseteq C(x^\circ)$ converges to L.

Adding the suffix "given previous intruders" results in a weakening of the concept.

The following lemma gives the relation between the concepts with and without the restriction of the support.

LEMMA 4.1:

Let $J \subseteq S$.

i) A strategy $p \in \Delta S$ is stable if and only if it is stable given previous intruders in J.

ii) A set $L \subseteq \Delta S$ that is attracting given previous intruders in J need not be attracting.

PROOF:

We may assume that J=S, i.e. the case of interior trajectories.

Part i): All we need to show is the "if" statement. Assume that a trajectory starting on the border of ΔS close to a strategy that is stable given previous intruders leads away from it. Then by the continuity property of the replicator dynamics (see section 2) a trajectory starting in the interior of ΔS also leads away. This contradicts the stability given previous intruders.

Part ii): If δ is sufficiently large and c is sufficiently small in the repeated Prisoners' Dilemma with types CC, CA and DCC then there exists $\gamma \in (0,1)$ such that $\{(1-\alpha)CC+\alpha CA$ s.t. $0 \leq \alpha \leq \gamma\}$ is an asymptotically stable set given previous intruders in DCC. However it is not attracting since each $x \in \Delta\{CA,CC\}$ is a rest point (see introduction or theorem 5.4).

□

An easy conclusion from the above lemma is that if a strategy is unstable it cannot be made stable given previous intruders if the set of types in the population is enlarged. On the other hand, a set that is not attracting given previous intruders might become attracting w.r.t. the interior if the set of strategies is enlarged.

The next result is useful for finding asymptotically stable sets given previous intruders.

LEMMA 4.2:

Let $J \subseteq S$ and $L \subseteq \Delta S$ be a connected asymptotically stable set given previous intruders in J. Then

i) L is a maximal connected set of stable strategies,

ii) L may not have arbitrarily close weak ESS's outside of the set but

iii) L may have arbitrarily close rest points outside of L but their support cannot contain J.

PROOF:

Part i) and iii) follow directly from the definitions. Part ii) then follows by theorem A2.

□

5. Asymptotically stable sets given previous intruders

Following the previous section we will now relax the condition of asymptotic stability. We continue the analysis of section 3, now analyzing the asymptotically stable sets given previous intruders that contain a pure strategy.

Up to now all we imposed on the parameters δ and c of the game was that $\delta<1$ and $c>0$. We are not interested in results that only hold for specific values but in results that are "qualitatively" true for a range of parameters (what we mean by "qualitatively" true will become clear later on). There are two conditions we will impose on such a range. We are interested in the approximation of the lexicographic preferences where players prefer higher repeated game payoffs and only when indifferent prefer automata with less states. Therefore we demand that the results should "qualitatively" remain unchanged if the cost of complexity is decreased. Additionally we are interested in values of the discount factor (continuation probability) close to 1, i.e., the results should not depend on the fact that the discount factor may not approach 1. Sets of parameters satisfying these two conditions will be called lex-patient.

This is formalized in the following definition. We will call a set $W\subseteq(0,1)\times(0,1)$ lex-patient if

i) $(\delta,c)\in W$ implies $(\delta,c')\in W$ for all $c'\in(0,c)$ and

ii) $(1,0)\in\partial W$ (i.e., in the closure of W).

The main goal of the following theorems will be to show that the set of Nash equilibrium look-a-likes of "Tat for Tit" is an asymptotically stable set given previous intruders.

First we will introduce two more look-a-likes of CC, namely AA and AC as defined in figure 5.

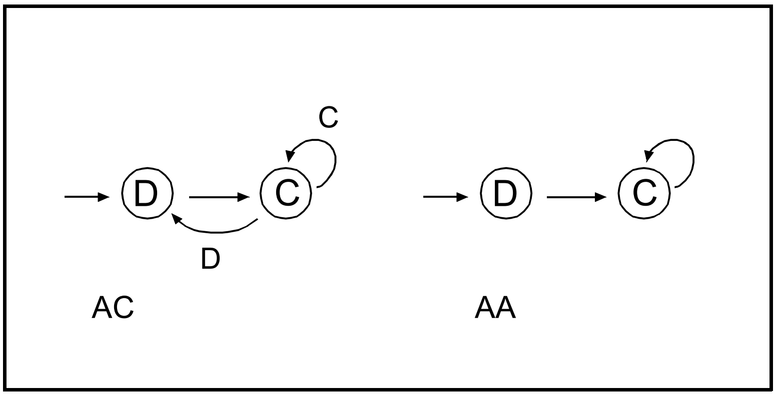

Figure 5: The automaton AC and a toothless look-a-like (AA).

It follows that the look-a-likes of "Tat for Tit" are AC, AA and CA (see also figure 2). Notice that AC can be a symmetric Nash equilibrium strategy for values of (δ,c) close to $(1,0)$ whereas this is not possible for AA since it is toothless. More specifically, a best response of AC for all parameters (δ,c) arbitrarily close to $(1,0)$ must be a look-a-like if and only if $2\pi(C,C)>\pi(D,C)+\pi(D,D)$ and $0<c<(1-\delta)[\pi(D,D)-\pi(C,D)]$. Notice that the first condition holds if and only if (CC,CC) is a Nash equilibrium for all (δ,c) sufficiently close to $(1,0)$. The second condition is necessary to ensure that "Cooperate for Ever" is not a best response of AC.

We will now characterize the set of Nash equilibrium look-a-likes of "Tat for Tit" and characterize a set that is contained in its basin of attraction. Especially, the basin of attraction does not converge to the empty set as (δ,c) goes to $(1,0)$.

THEOREM 5.1:

Assume $2\pi(C,C)>\pi(D,C)+\pi(D,D)$. Let $\rho^* = \dfrac{2\pi(C,C)-\pi(D,C)-\pi(D,D)}{4\pi(C,C)-\pi(D,C)-2\pi(D,D)-\pi(C,D)}$,

for $(\delta,c)\in(0,1)^2$ $\gamma(DCC) = \dfrac{[\pi(C,C)-\pi(D,C)]\delta(1-\delta)+[\pi(C,C)-\pi(D,D)]\delta^2(1-\delta)+c}{\pi(D,C)\delta^2-\pi(D,D)\delta^2(1-\delta)-\pi(C,C)\delta^3}$,

$T=\{(\delta,c)\in(0,1)^2 \text{ s.t. } 0<c<(1-\delta)[\pi(D,D)-\pi(C,D)]\}$,

$L(\delta,c)=\{x\in\Delta\{AA,CA,AC,CC\} \text{ s.t. } x(AA)+x(CA)\leq\gamma(DCC)\}$ and

$B(\epsilon)=\{x\in\Delta S_p \text{ s.t. } x(AC)+x(CC)>1-\epsilon,\ x(DCC)>0\}$.

Then there exists $\delta_0<1$ and $\rho:(\delta_0,1)\to(0,1)$ such that for any $(\delta,c)\in T\cap\{\delta>\delta_0\}$,

i) $L(\delta,c)=\{x\in\Delta S$ s.t. $C(x)\subseteq R_{\delta,c}(x)\cap Q(CC)\}$,

ii) $L(\delta,c)\subseteq B(\rho(\delta))$,

iii) any trajectory starting in $B(\rho(\delta))$ converges to $L(\delta,c)$ and

iv) $\rho(\delta)\to\rho^*$ as $\delta\to 1$.

Notice that $T\cap\{\delta\geq\delta_0\}$ is a lex-patient set, that $\gamma(DCC)\to 0$ and $L(\delta,c)\to\Delta\{AC,CC\}$ as $(\delta,c)\to(1,0)$, that $\rho^*\in(0,½)$ and in the special example of a Prisoners' Dilemma given in table I, that $\rho^*=1/8$.

We will be needing the following notation. For $x\in\Delta S_p$ let $\alpha(x)$ and $\beta(x)$ be such that $\alpha(x)=x(AA)+x(CA)$ and $1-\alpha(x)-\beta(x)=x(AC)+x(CC)$. It follows that $\beta(x)$ is the measure of the strategies that are not look-a-likes of CC. Furthermore let $e^*(x)$, $e(x)$ and $q(x)$ be the mean strategies of x in $\{AC,CC\}$, $\{AA,CA\}$ and $S\backslash Q(CC)$, provided that they exist. Formally, if $\alpha(x)+\beta(x)<1$ then let $e^*(x)=\dfrac{x(AC)AC+x(CC)CC}{x(AC)+x(CC)}$, if $\alpha(x)>0$ then let

$e(x)=\dfrac{x(AA)AA+x(CA)CA}{x(AA)+x(CA)}$ and if $\beta(x)>0$ then let $q(x)=\dfrac{\sum_{a\in S\backslash Q(CC)} x(a)a}{\sum_{a\in S\backslash Q(CC)} x(a)}$. It follows that

$e^*(x)\in\Delta\{AC,CC\}$, $e(x)\in\Delta\{AA,CA\}$, and $q(x)\in\Delta\{S_p\backslash Q(CC)\}$ (when well defined), $\alpha(x),\beta(x)\in[0,1]$ and $x=[1-\alpha(x)-\beta(x)]e^*(x)+\alpha(x)e(x)+\beta(x)q(x)$. Notice that $e^*(x)$, $e(x)$ and $q(x)$ will generally change over time if x is the mean strategy of a population evolving according to the replicator dynamics.

PROOF of theorem 5.1:

Let $\delta_1\in(0,1)$ be such that $\pi(C,C)(1+\delta)>\pi(D,C)+\delta\pi(D,D)$ whenever $\delta_1\in(\delta,1)$. Fix $(\delta,c)\in T\cap\{\delta>\delta_1\}$. It follows that $R_{\delta,c}(AC)=R_{\delta,c}(CC)=Q(CC)$. We will first find the maximal frequencies of AA and CA that can be in a Nash equilibrium look-a-like of CC. Fix $e^*\in\Delta\{AC,CC\}$ and $e\in\Delta\{AA,CA\}$. Since AA and CA are not symmetric Nash equilibrium strategies there will be a strategy in $S_p\backslash Q(CC)$ that is the "reason" why a larger proportion of e cannot be in a symmetric Nash equilibrium strategy with support on $\{e^*,e\}$.

For $q \in S_p \backslash Q(CC)$ such that $E_{\delta,c}(q,e) > E_{\delta,c}(e^*,e^*)$ let $\gamma(q) = \dfrac{E_{\delta,c}(e^*,e^*) - E_{\delta,c}(q,e^*)}{E_{\delta,c}(q,e) - E_{\delta,c}(q,e^*)}$. It follows that $E_{\delta,c}(q,(1-\gamma(q))e^* + \gamma(q)e) = E_{\delta,c}(e^*,e^*)$. Notice that $1/\gamma(q)$ is defined for all $q \in S_p \backslash Q(CC)$ and that the definition of $\gamma(q)$ is consistent with the definition of $\gamma(DCC)$ in the statement of the theorem. Consider $\inf\{\gamma(q) \text{ s.t. } q \in S_p \backslash Q(CC), E_{\delta,c}(q,e) > E_{\delta,c}(e^*,e^*)\}$. By the properties of finite automata it follows that "infimum" can be replaced by "minimum". Moreover it follows that $\{DCC\} = \text{argmin}\{..\}$. Consequently for $x \in \Delta Q(CC)$, if $\alpha(x) < \gamma(DCC)$ then $R_{\delta,c}(x) = Q(CC)$, if $\alpha(x) = \gamma(DCC)$ then $R_{\delta,c}(P) = \{DCC\} \cup Q(CC)$ and if $\alpha(x) > \gamma(DCC)$ then $E_{\delta,c}(DCC,x) > E_{\delta,c}(x,x)$. Therefore $x \in \Delta\{AA,AC,CA,CC\}$ is a symmetric Nash equilibrium strategy if and only if $x(AA) + x(CA) \le \gamma(DCC)$ which proves part i).

The objective for the proof of part iii) will be to show for sufficiently large δ and $0 < c < (1-\delta)[\pi(D,D) - \pi(C,D)]$ that there exists $\rho(\delta) > 0$ such that $L(\delta,c) \subseteq B(\rho(\delta))$ and for any $x \in B(\rho(\delta))$ the mean frequency of AC and CC will increase, i.e., $[E_{\delta,c}(e^*(x),x) - E_{\delta,c}(x,x)](1 - \alpha(x) - \beta(x)) > 0$. Once this is shown the proof of iii) will be complete. A trajectory starting in $B(\rho(\delta))$ will converge to $\Delta Q(CC)$. Note that at this point we can not claim that the trajectory converges to a strategy. However since $E_{\delta,c}(DCC,x) > E_{\delta,c}(CC,CC)$ for $x \in \Delta Q(CC)$ such that $\alpha(x) > \gamma(DCC)$ and the fact that $x(DCC) > 0$ when $x \in B(\epsilon)$ it follows that the trajectory will not run arbitrarily close to $\{x \in \Delta Q(CC) \text{ s.t. } \alpha(x) > \gamma(DCC)\} = \Delta Q(CC) \backslash L(\delta,c)$ (for a formal argument construct a Lyapunov function). Therefore each trajectory starting in $B(\rho(\delta))$ will converge to $L(\delta,c)$.

Some calculations:

For $e^*, e, q \in \Delta S_p$ such that $C(e^*) \subseteq \{AC,CC\}$, $C(e) \subseteq \{AA,CA\}$ and $C(q) \cap Q(CC) = \varnothing$ we obtain

$$\begin{aligned}
E_{\delta,c}(x,x) = {} & E_{\delta,c}(e^*,e^*) + \beta[E_{\delta,c}(e^*,q) - 2E_{\delta,c}(e^*,e^*) + E_{\delta,c}(q,e^*)] \\
& + \alpha\beta[E_{\delta,c}(e,q) - E_{\delta,c}(e^*,q) + E_{\delta,c}(q,e) - E_{\delta,c}(q,e^*)] \\
& + \beta^2[E_{\delta,c}(e^*,e^*) - E_{\delta,c}(q,e^*) + E_{\delta,c}(q,q) - E_{\delta,c}(e^*,q)].
\end{aligned}$$

$$\begin{aligned}
E_{\delta,c}(e^*,x) - E_{\delta,c}(x,x) = {} & \beta[[E_{\delta,c}(e^*,e^*) - E_{\delta,c}(q,e^*)] \\
& - \alpha[E_{\delta,c}(e,q) - E_{\delta,c}(e^*,q) + E_{\delta,c}(q,e) - E_{\delta,c}(q,e^*)] \\
& - \beta[E_{\delta,c}(e^*,e^*) - E_{\delta,c}(q,e^*) + E_{\delta,c}(q,q) - E_{\delta,c}(e^*,q)]] \\
= {} & \beta[E_{\delta,c}(e^*,e^*) - E_{\delta,c}(q,e^*)][1 - \alpha/\mu - \beta/\lambda] \text{ where}
\end{aligned}$$

$$\frac{1}{\lambda(q)}=1+\frac{E_{\delta,c}(q,q)-E_{\delta,c}(e^*,q)}{E_{\delta,c}(e^*,e^*)-E_{\delta,c}(q,e^*)}$$ and

$$\frac{1}{\mu(q)}=\frac{E_{\delta,c}(q,e)-E_{\delta,c}(q,e^*)+E_{\delta,c}(e,q)-E_{\delta,c}(e^*,q)}{E_{\delta,c}(e^*,e^*)-E_{\delta,c}(q,e^*)}=\frac{1}{\gamma(q)}+\frac{E_{\delta,c}(e,q)-E_{\delta,c}(e^*,q)}{E_{\delta,c}(e^*,e^*)-E_{\delta,c}(q,e^*)}\ .$$

Notice that if $E_{\delta,c}(q,q)>E_{\delta,c}(e^*,q)$ then $\lambda(q)\in(0,1)$ and $E_{\delta,c}(q,(1-\lambda)e^*+\lambda q)=E_{\delta,c}(e^*,(1-\lambda)e^*+\lambda q)$. Especially if $e^*=CC$ then $(1-\lambda(q))CC+\lambda(q)q$ is a rest point of the replicator dynamics (compare to figure 2 in which $\gamma=\gamma(DCC)$ and $\lambda=\lambda(DCC)$ are included.

Fix $(\delta,c)\in T\cap\{\delta>\delta_0\}$. We will now show that there exists $\rho°(\delta)>0$ such that for all $q\in\Delta S_p$ such that $C(q)\cap Q(CC)=\varnothing$, $1/\lambda(q)$ is bounded above by $1/\rho°(\delta)$. This is equivalent to showing that $\lambda(q)$ is bounded below by $\rho°(\delta)$ for all $q\in S_p\backslash Q(CC)$ such that $E_{\delta,c}(q,q)>E_{\delta,c}(e^*,q)$.

Notice that for $t\in[0,1]$ and $q^i\in\Delta S$ such that $C(q)\cap Q(CC)=\varnothing$, $i=1,2$, $\frac{1}{\lambda((1-t)q^1+tq^2)}\leq\max\{\frac{1}{\lambda(q^1)},\frac{1}{\lambda(q^2)}\}$. Moreover $1/\lambda(q)$ is decreasing in c and for fixed q, $1/\lambda(q)$ achieves its maximum when varying e* in $\Delta\{AC,CC\}$ on $\{AC,CC\}$. Therefore it is enough to show $1/\lambda(q)\leq 1/\rho°(\delta)$ for $c=0$, $e^*\in\{AC,CC\}$ and any $q\in S_p\backslash Q(CC)$.

Fix $e^*\in\{AC,CC\}$ and let $c=0$. From the properties of finite automata it follows that there exists $q^*\in S_p\backslash Q(CC)$ that maximizes $1/\lambda(q)$. For demonstration purposes we will assume that e* and q* are the only strategies and derive some properties of q*. The automaton q* will deviate from e*'s behavior in some round and then "know" if it is matched against e* or q*. We may assume that q* will deviate from e*'s behavior in either round one or round two. If q* learns that it is matched against q* it will cooperate for ever from then on. If it learns that it is matched against e* who is going to play D, q* will play D too.

Let k be the first round in which e* will cooperate (play "C") after q* learns that it is matched against e*. Together with the remarks above either q* will forever cooperate or will forever defect from then on. It can be shown (which also seems intuitive) that q* will play D for ever on once it reveals the identity of e*.

Let $\rho^\circ(\delta)=\lambda(q^*)$. The above characterization of q* implies that $1/\lambda(q)\leq 1/\rho^\circ(\delta)$ and $\rho^\circ(\delta)\to\rho^*$ as $\delta\to 1$.

Similarly there exists $\rho'(\delta)>0$ such that for all $z\in\Delta S_p$ such that $C(z)\cap Q(CC)=\emptyset$, $1/\mu(z)$ is bounded above by $1/\rho'(\delta)$. The proof of this statement follows very closely the one above. Let c=0, e∈{AA,CA}, e*∈{AC,CC} and let z* be the maximizer of $1/\mu(z)$ for $z\in S_p\backslash Q(CC)$. It can be shown (which again seems intuitive) that z* will cooperate for ever once it recognizes that it is matched against e and defects for ever once it knows that it is matched against e*. With $1/\rho'(\delta)=1/\mu(z^*)$ we obtain that $1/\mu(z)\leq 1/\rho'(\delta)$ and $1/\rho'(\delta)\to 1/\rho^*$ as $\delta\to 1$.

With the above bounds we are now able to show that the mean frequency of AC and CC is increasing. Let $\rho(\delta):=\min\{\rho^\circ(\delta),\rho'(\delta)\}$. Since $\gamma(DCC)\to 0$ as $(\delta,c)\to(1,0)$ there exists $\delta_0\in[\delta_2,1)$ such that $L(\delta,c)\subseteq B(\rho(\delta))$ for all $(\delta,c)\in T\cap\{\delta>\delta_0\}$. For $(\delta,c)\in T\cap\{\delta>\delta_0\}$ and $x\in B(\rho(\delta))\backslash L(\delta,c)$ we obtain $0<\alpha(x)+\beta(x)<\rho(\delta)$, $\beta(x)>0$ and therefore $[E_{\delta,c}(e^*,x)-E_{\delta,c}(x,x)](1-\alpha-\beta)=\beta(1-\alpha-\beta)[E_{\delta,c}(e^*,e^*)-E_{\delta,c}(q,e^*)][1-\alpha/\mu-\beta/\lambda]$ $\geq\beta(1-\alpha-\beta)[E_{\delta,c}(e^*,e^*)-E_{\delta,c}(q,e^*)][1-(\alpha+\beta)/\rho(\delta)]>0$. This finishes the proof of part ii) and iii).

Finally part iv) follows directly from the fact that $\rho(\delta)$ converges to ρ^* as δ goes to 1.

□

The next result follows almost directly from the proof of the above theorem.

COROLLARY 5.2:

Assume $2\pi(C,C)>\pi(D,C)+\pi(D,D)$. There exists $\delta_1<1$ such that $\Delta\{AC,CC\}$ is stable for all $(\delta,c)\in T\cap\{\delta>\delta_1\}$.

PROOF:

It was shown in the proof of theorem 5.1 that there exists a $\delta_0<1$ such that for any

$(\delta,c)\in T\cap\{\delta>\delta_0\}$, any $x\in B(\rho(\delta))$, $E_{\delta,c}(e^*(x),x)\geq E_{\delta,c}(x,x)$. However since $x=(1-\alpha-\beta)e^*(x)+\alpha e+\beta q$, we did not show that $E_{\delta,c}(z^*,y)\geq E_{\delta,c}(y,y)$ for all $z^*\in\Delta\{AC,CC\}$ and all $y\in B(\rho(\delta))$. Especially setting $z^*=AC$ this would mean that the frequency of AC always increases in $B(\rho(\delta))$. This is not true. It is easy to find a strategy q and trajectory in $\Delta\{AA,CA,AC,CC,q\}$ such that the frequency of AC decreases over time although the frequency of the mean strategy of AC and CC increases (as shown in the proof in general).

With a bit more effort and using continuity it can be shown that the above statement is true in a neighborhood of z^*, i.e., for any $(\delta,c)\in T\cap\{\delta>\delta_0\}$ and any $z^*\in\Delta\{AC,CC\}$ there exists an $\epsilon>0$ such that for any $y\in U_\epsilon(z^*)$, $E_{\delta,c}(z^*,y)\geq E_{\delta,c}(y,y)$. Lemma A3 then implies that z^* is stable.

□

As mentioned earlier we aim to show that the set of Nash equilibrium look-a-likes of CC is an asymptotically stable set given previous intruders. Following theorem 5.1 and corollary 5.2 all we must show is that each Nash equilibrium look-a-like of CC that contains toothless look-a-likes is stable. Because we do not have a general theorem to prove this statement, we must consider the behavior of the trajectories directly.

Let $e^*\in\Delta\{AC,CC\}$, $e\in\Delta\{AA,CA\}$ and $q\in\Delta\{S_p\backslash Q(CC)\}$. Consider the replicator dynamics in $\Delta\{e^*,e,q\}$ as if e^*, e and q were pure strategies, denoted by RD(e^*,e,q).

Formally, RD(e^*,e,q) is defined by

$x^0=\overline{x}$ and $\frac{d}{dt}x^t(a)=[E(a,x^t)-E(x^t,x^t)]\,x^t(a)$, $a\in\{e^*,e,q\}$, $t\geq 0$, where $\overline{x}\in\Delta\{e^*,e,q\}$ is the initial state and $x^t(a)$ is the frequency of the type using strategy a ($a\in\{e^*,e,q\}$) at time t.

Notice that the trajectories of RD(e^*,e,q) can also be projected into ΔS_p, however the resulting trajectories in ΔS_p should not be confused with the trajectories of the replicator dynamics (RD). In RD(e^*,e,q) the strategies e^*, e and q stay fixed, in the replicator dynamics on ΔS_p the mean strategies on {AC,CC}, {AA,CA} and $S_p\backslash Q(CC)$ vary over time. In order to distinguish strategies in $\Delta\{e^*,e,q\}$ from mean strategies in ΔS_p we will write $x=(1-a-\beta)[e^*]+\alpha[e]+\beta[q]$ when $x\in\Delta\{e^*,e,q\}$.

The following lemma will simplify the calculation of the trajectories of the replicator dynamics.

LEMMA 5.3:

Let $X \subseteq \{(\alpha,\beta)$ s.t. $\alpha+\beta \leq 1,\ \alpha,\beta \geq 0\}$ be a closed polygon. Assume for each $e^* \in \Delta\{AC,CC\}$, $e \in \Delta\{AA,CA\}$ and $q \in S_p \backslash Q(CC)$ that the trajectory of RD(e^*,e,q) starting in $U' := \{(1-\alpha-\beta)[e^*]+\alpha[e]+\beta[q],\ (\alpha,\beta) \in X\}$ stays in U'. Then the trajectory of RD starting in $U := \{x \in \Delta S_p$ s.t. $(\alpha,\beta) \in X,\ x(\{AC,CC\})=1-\alpha-\beta,\ x(\{AA,CA\})=\alpha\}$ stays in U.

Before presenting the proof notice that U' might be a small neighborhood of some strategy in $\Delta\{[e^*],[e],[q]\}$ whereas U will always be a "cylinder" set in ΔS_p.

PROOF:

Since X is a polygon, i.e., an intersection of half spaces, so is U. Trajectories starting in U will stay in U if and only if the trajectories starting on the border of U do not lead out of U. Because U is a polygon, this holds if and only if the gradients on the border of U "point" towards U. The same can be stated of U'. Notice that the above is not necessarily true if X is not a polygon.

We will now relate the gradients of RD(e^*,e,q) to those of RD. It can be shown for any $x \in \Delta S_p$ that the gradient of RD in x is identical to the gradient of RD($e^*(x)$,e(x),q(x)) in $(1-\alpha(x)-\beta(x))[e^*]+\alpha(x)[e]+\beta(x)[q]$ projected into ΔS_p.

Combining the above two statements completes the proof of the lemma.

□

Now we are able to present the first of two major results of this paper, stating (together with theorem 5.1) that there exists a lex-patient set such that for all parameters in this set, the set of Nash equilibrium look-a-likes of "Tat for Tit" is an asymptotically stable set given previous intruders. The condition of previous intruders is satisfied if each initial state contains a type in the class of DCC.

THEOREM 5.4:

Assume $2\pi(C,C)>\pi(D,C)+\pi(D,D)$. Let T, $L(\delta,c)$ and $\gamma(DCC)$ be defined as in theorem 5.1. There exists $\delta^*<1$ such that for all $(\delta,c)\in T\cap\{\delta>\delta^*\}$, $L(\delta,c)$ is an asymptotically stable set given previous intruders containing a type in the class of DCC.

PROOF:

According to theorem 5.1 there exists $\delta_0<1$ and $\rho:(\delta_0,1)\rightarrow(0,1)$ such that for all $(\delta,c)\in T\cap\{\delta>\delta_0\}$, $L(\delta,c)$ with respect to initial states in $B(\rho(\delta))$. Following corollary 5.2, each $x\in\Delta(AC,CC\}$ is stable for these parameters. All that is left to show is that there exists $\delta^*\in[\delta_0,1)$ such that strategies in $L(\delta,c)\backslash\Delta\{AC,CC\}$ are stable for any $(\delta,c)\in T\cap\{\delta>\delta^*\}$.

First we will show that strategies on the border of $L(\delta,c)$ are stable, i.e., that there exists $\delta°<1$ such that for any $(\delta,c)\in T\cap\{\delta>\delta°\}$ and any $x\in\Delta S_p$ with $\alpha(x)=\gamma(DCC)$ and $\beta(x)=0$ is stable.

The proof of this statement is split up into three steps, enumerated one to three. We will first consider the replicator dynamics as if e*, e and q were pure strategies (i.e., RD(e*,e,q)). This reduces the number of pure strategies from countable infinite to three which substantially facilitates an analysis of the trajectories. We will show that trajectories in $\Delta\{e^*,e,q\}$ starting close to $(1-\gamma(DCC))[e^*]+\gamma(DCC)[e]$ will stay close. Then we will apply lemma 5.3 to show that the frequencies of the mean strategies on {AC,CC}, {AA,CA} and $S_p\backslash Q(CC)$ of a strategy starting close to the border of $L(\delta,c)$ and evolving according to RD will neither change very much. Finally we will show that the trajectories of RD must stay close too (in ΔS_p).

STEP 1:

Let $\delta_1\in[\delta_0,1)$ be such that for all $(\delta,c)\in T\cap\{\delta>\delta_1\}$, $\max\limits_{q\in\Delta\{S_p\backslash Q(CC)\}}\left\{\frac{1}{\lambda(q)},\frac{1}{\mu(q)}\right\}<\frac{2}{\rho*}$.

Such a δ_1 exists by theorem 5.1.

Let $(\delta,c)\in T\cap\{\delta>\delta_1\}$, $e^*\in\Delta\{AC,CC\}$, $e\in\Delta\{AA,CA\}$ and $q\in\Delta\{S_p\backslash Q(CC)\}$ be fixed. In the following we will consider the replicator dynamics as if e*, e and q were pure strategies (i.e., RD(e*,e,q)).

For any $P\in\Delta\{e^*,e,q\}$, let $\alpha=P(e)$ and $\beta=P(q)$ and vice versa, let $P(\alpha,\beta)=(1-\alpha-\beta)[e^*]+\alpha[e]+\beta[q]$ for $\alpha,\beta>0$ such that $\alpha+\beta\leq 1$. Following the proof of theorem 5.1 and the definition of δ_1, $E_{\delta,c}(e^*,P)-E_{\delta,c}(P,P)\geq 0$ (>0 if $\beta>0$ and) if $\alpha+\beta<\rho^*/2$ in which case the frequency of e* will increase in state P.

Using the notation in the proof of theorem 5.1, we obtain

$$E_{\delta,c}(q,P)-E_{\delta,c}(P,P)=E_{\delta,c}(q,e^*)-E_{\delta,c}(e^*,e^*)+\alpha[E_{\delta,c}(q,e)-E_{\delta,c}(q,e^*)]$$
$$+\beta[-E_{\delta,c}(e^*,q)+E_{\delta,c}(q,q)+2E_{\delta,c}(e^*,e^*)-2E_{\delta,c}(q,e^*)]$$
$$-\alpha\beta[E_{\delta,c}(e,q)-E_{\delta,c}(e^*,q)+E_{\delta,c}(q,e)-E_{\delta,c}(q,e^*)]$$
$$-\beta^2[E_{\delta,c}(e^*,e^*)-E_{\delta,c}(q,e^*)+E_{\delta,c}(q,q)-E_{\delta,c}(e^*,q)]$$
$$=[E_{\delta,c}(e^*,e^*)-E_{\delta,c}(q,e^*)](-1+\alpha/\gamma+\beta(1+1/\lambda-\alpha/\mu-\beta/\lambda))$$
$$=[E_{\delta,c}(e^*,e^*)-E_{\delta,c}(q,e^*)]((1-\beta)(\beta/\lambda-1)+\alpha(1/\gamma-\beta/\mu)).$$

For $P\in\Delta\{e^*,e,q\}$ and $a\in\{e^*,e,q\}$ let f(a) be the marginal increase of the frequency of "a" according to RD(e*,e,q), i.e., $f(a)=[E_{\delta,c}(a,P)-E_{\delta,c}(P,P)]P(a)$. With this notation, a measure of the direction of the gradient of RD(e*,e,q) evaluated at a strategy P with $\beta(P)>0$ and $\alpha(P)+\beta(P)<\rho^*/2$ is $\dfrac{f(q)}{f(e^*)}(\alpha,\beta)=\dfrac{(1-\beta)(\frac{\beta}{\lambda}-1)+\alpha(\frac{1}{\gamma}-\frac{\beta}{\mu})}{(1-\frac{\alpha}{\mu}-\frac{\beta}{\lambda})(1-\alpha-\beta)}$.

The rest of step 1 will be subdivided into three parts.

STEP 1.1

We now proceed to show that the rate at which q increases relative to the rate at which e* increases is arbitrarily small if trajectories start in a neighborhood of $P(\gamma(DCC),0)$.

For $P\in\Delta\{e^*,e,q\}$ and $\alpha_0\in(0,1)$ let $SH(s,\alpha_0)$ be the shaded area in figure 6 where $s:=\dfrac{P(q)}{P(e^*)}$ is the "direction" of the vector P originating in $(1-\alpha_0)[e^*]+\alpha_0[e]$.

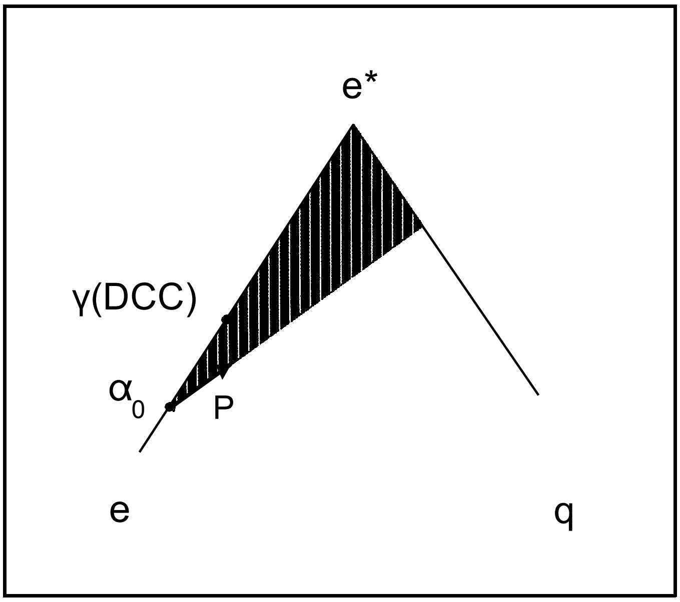

Figure 6: The set $SH(s,\alpha_0)$.

What characteristic does the shaded area have?

$$\begin{pmatrix} 1-\alpha-\beta \\ \alpha \\ \beta \end{pmatrix} - \begin{pmatrix} 1-\alpha_0 \\ \alpha_0 \\ 0 \end{pmatrix} = \begin{pmatrix} \alpha_0-\alpha-\beta \\ \alpha-\alpha_0 \\ \beta \end{pmatrix}.$$

So $P\in SH(s,\alpha_0)$ if and only if $\dfrac{\beta}{\alpha_0-\alpha-\beta} \le s$ and $SH(s,\alpha_0):=\{P(\alpha,\beta)$ s.t.

$\alpha+\beta(1+1/s)\le\alpha_0\}$.

We will find $\alpha^*\in(\gamma(DCC),1)$ and construct a continuous function $s^*()$ such that for $\alpha_0\in(\gamma(DCC),\alpha^*)$ and $(\delta,c)\in T\cap\{\delta>\delta_1\}$, trajectories starting in $SH(s^*(\alpha_0),\alpha_0)$ stay in it and the shaded area $SH(s^*(\alpha_0),\alpha_0)$ approaches the segment $\{(1-\theta)[e^*]+\theta[e],\ 0\le\theta\le\gamma(DCC)\}$ as α_0 approaches $\gamma(DCC)$.

Let $s^*(\alpha_0):=\dfrac{-1+\dfrac{\alpha_0}{\gamma(DCC)}}{(1-\dfrac{2\alpha_0}{\rho^*})(1-\alpha_0)}$ be defined for $\alpha_0\in[\gamma(DCC),\rho^*/2)$.

Notice that s*() is continuous and that $s^*(\gamma(DCC))=0$. Therefore $SH(s^*(\alpha_0),\alpha_0)\to\{(1-\theta)[e^*]+\theta[e],\ 0\leq\theta\leq\gamma(DCC)\}$ as $\alpha_0\to\gamma(DCC)$.

Consider the direction of the gradient along the border of $SH(s^*(\alpha_0),\alpha_0)$ that is in the interior of $\Delta\{e^*,e,q\}$, i.e., in all points $P(\alpha,\beta)$ such that $\alpha+\beta(1+\frac{1}{s^*(\alpha_0)})=\alpha_0$ and

$0\leq\beta\leq\beta(\alpha_0)$ where $\beta(\alpha_0)$ is defined such that $\beta(\alpha_0)(1+\frac{1}{s^*(\alpha_0)})=\alpha_0$. Note that $\beta(\alpha_0)\to 0$ as

$\alpha_0\to 0$. We will show that for all $(\delta,c)\in T\cap\{\delta>\delta_1\}$ there exists $\alpha°>\gamma(DCC)$ such that for all $\gamma(DCC)\leq\alpha_0<\alpha°$ and all $0\leq\beta\leq\beta(\alpha_0)$ $\frac{f(q)}{f(e^*)}(\alpha_0-\beta\left(1+\frac{1}{s^*(\alpha_0)}\right),\beta)\leq s^*(\alpha_0)$. (1)

Since the frequency of e* increases over time (1) will ensure that trajectories starting in $SH(s^*(\alpha_0),\alpha_0))$ stay in it.

In order to show (1) we will use the following inequality and consider two cases, one in which q is such that $\gamma(q)$ is close to $\gamma(DCC)$, one in which it is not.

$$\frac{f(q)}{f(e^*)}(\alpha_0-\beta\left(1+\frac{1}{s^*}\right),\beta)\leq\frac{-1+\frac{\alpha_0}{\gamma(q)}+\beta\left[1+\frac{2}{\rho^*}-\frac{\alpha_0}{\mu}-\left(1+\frac{1}{s^*}\right)\frac{1}{\gamma(q)}\right]+\beta^2\left[\left(1+\frac{1}{s^*}\right)\frac{1}{\mu}-\frac{1}{\lambda}\right]}{(1-\frac{2}{\rho^*}\alpha_0)(1-\alpha_0)} \quad (2)$$

Assume that $\frac{1}{\gamma(q)}\geq\frac{1}{4\gamma(DCC)}$. Notice that the right hand side of (2) is bounded by

$s^*(\alpha_0)$ when $\beta=0$. Let $\alpha'\in(\gamma(DCC),\rho^*/2)$ be such that $1+\frac{2}{\rho^*}-\frac{\alpha_0}{\mu}-\left(1+\frac{1}{s^*}\right)\frac{1}{4\gamma(DCC)}<0$ for

all $\alpha_0\in(\gamma(DCC),\alpha')$. Such an α' exists because $s^*(\alpha_0)\to 0$ as $\alpha_0\to\gamma(DCC)$ and $1/\mu(q)$ is bounded below. Using (2), (1) follows.

Assume now that $\frac{1}{\gamma(q)}<\frac{1}{4\gamma(DCC)}$. Then for $\alpha_0\leq 2\gamma(DCC)$, $-1+\frac{\alpha_0}{\gamma(q)}\leq-\frac{1}{2}$. If additionally β is small, then (1) follows. So there exists $\alpha''>\gamma(DCC)$ such that (1) holds for all $\alpha_0\in(\gamma(DCC),\alpha'')$.

Therefore setting $\alpha^\circ=\min\{\alpha',\alpha''\}$, (1) is proven and step 1.1 is completed.

STEP 1.2

We will now show that in a neighborhood of $P(\alpha,0)$, $\alpha<\gamma(DCC)$ the gradients in the interior point toward $\Delta\{e^*,e\}$ for any q. Together with the fact that the frequency of e* increases, this will imply that trajectories starting in a neighborhood of $P(\alpha,0)$ stay in it.

The gradients on $\Delta\{e^*,e\}$ are null vectors since these are rest points of the dynamics. Consider the direction of gradients arbitrarily close to $\Delta\{e^*,e\}$, i.e., as β goes to zero and denote this by $s(q,\alpha)$, i.e., $s(q,\alpha)=\lim_{\beta\to 0}\frac{f(q)}{f(e^*)}(\alpha,\beta)=\frac{-1+\frac{\alpha}{\gamma}}{(1-\frac{\alpha}{\mu})(1-\alpha)}$.

Let $\epsilon\in(0,\gamma(DCC)/2)$ be fixed. By the maximality of DCC for $1/\gamma(\cdot)$ it follows that there exists $\sigma>0$ such that $s(q,\gamma(DCC)-\epsilon)<-2\sigma$ for all q. From the properties of finite automata for given (δ,c) that $1/\gamma$, $1/\lambda$ and $1/\mu$ are bounded below for all $q\in\Delta\{S_p\backslash Q(CC)\}$. Therefore using continuity there exists a neighborhood M' of $P(\gamma(DCC)-\epsilon,0)$ such that for any $P(\alpha,\beta)\in M'$, $\frac{f(q)}{f(e^*)}(\alpha,\beta)<-\sigma$. For $\omega\in(0,\epsilon)$ let

$M:=\{P(\alpha,\beta)\text{ s.t. }\alpha+\beta\leq\gamma(DCC)-\epsilon+\omega,\ \alpha+\beta(1+1/\sigma)\geq\gamma(DCC)-\epsilon-\omega\}$. The set M is graphically represented in figure 7.

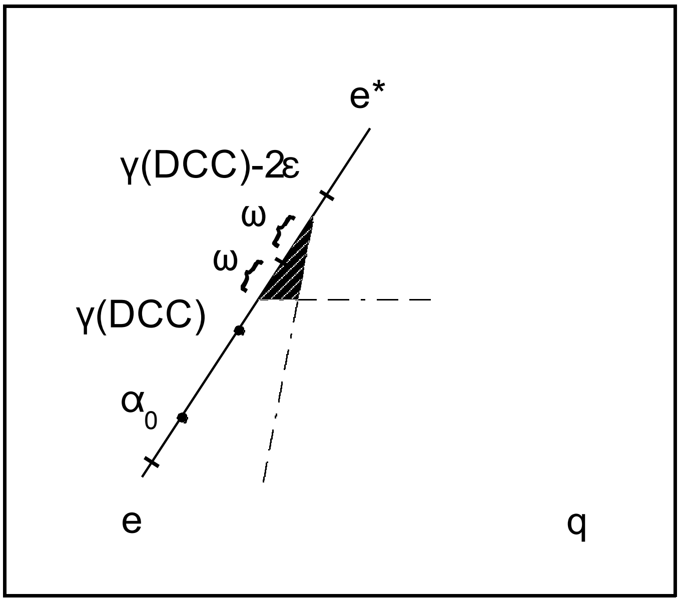

Figure 7: The set M.

It follows that ω can be chosen such that $M \subseteq M' \cap U_\epsilon(P(\gamma(DCC),0))$. Therefore M is a neighborhood of $P(\gamma(DCC)-\epsilon,0)$ and trajectories starting in M will stay in it and converge to a limit $P(\alpha,0)$ such that $\alpha<\gamma(DCC)$.

STEP 1.3:

Starting in a neighborhood of $P(\gamma(DCC),0)$ the frequency of q will at most barely increase while the frequency of e* increases steadily (step 1.1). We will now choose α_0 closes enough to $\gamma(DCC)$ such that either the trajectory will converge to a point $P(\alpha,0)$ with α close to $\gamma(DCC)$ or it will enter the neighborhood M of $P(\gamma(DCC)-\epsilon,0)$ from step 1.2 and from there on converge to a point $P(\alpha,0)$ with α close to $\gamma(DCC)-\epsilon$.

Let $\alpha_1 := \max\{\alpha_0 \leq \alpha^\circ \text{ s.t. } \alpha_0 \leq \gamma(DCC)+2\epsilon, \partial SH(s^*(\alpha_0),\alpha_0) \cap M \cap \{\beta>0\} \neq \emptyset\}$. Since $s^*(\alpha_0) \to 0$ as $\alpha_0 \to \gamma(DCC)$ such an α_1 exists. Let
$V := SH(s^*(\alpha_1),\alpha_1) \cap \{(\alpha,\beta) \text{ s.t. } \alpha+\beta(1+1/\sigma) \geq \gamma(DCC)-\epsilon-\omega\}$ (see figure 8).

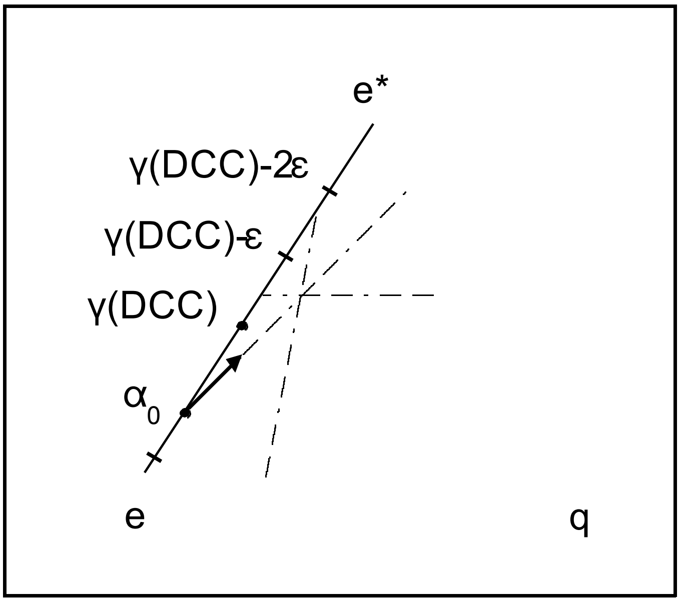

Figure 8: The construction of the neighborhood V of P(γ(DCC),0).

By the above construction, $V \subseteq U_{2\epsilon}(P(\gamma(DCC),0))$ and trajectories starting in V stay in it. Furthermore the definition of V is not dependent on the choice of $e^* \in \Delta\{AC,CC\}$, $e \in \Delta\{AA,CA\}$ or $q \in \Delta\{S_p \backslash Q(CC)\}$. Since $\epsilon \in (0,\gamma(DCC)/2)$ was arbitrary, we have shown that P(γ(DCC),0) is stable in RD(e*,e,q).

STEP 2:

In this step we apply lemma 5.3 to the results of step 1. Let $X=\{(\alpha,\beta)$ s.t. $P(\alpha,\beta) \in V\}$. Notice that X is a polygon. The conclusion is that the trajectories of RD starting in $V°=\{x \in \Delta S_p$ s.t. $(x(\{AA,CA\}),x(\{S_p \backslash Q(CC)\})) \in X\}$ will stay in V°. So the frequencies of the mean strategies on {AC,CC}, {AA,CA} and $S_p \backslash Q(CC)$ stay close.

STEP 3:

Finally we must rule out the case that the frequencies of the pure strategies do not stay close although the mean strategies do according to step 2. For this we will show that the mean frequency on {AC,CC} increases up to a positive constant at least as strong as the frequency of any pure strategy increases or decreases. Therefore if the mean strategy on {AC,CC} stays close, so will the frequencies of the pure strategies.

Let $(\delta,c) \in T \cap \{\delta > \delta_1\}$ be fixed and consider $e° \in Q(CC)$, $x \in \Delta S_p$ such that $x(e°)>0$

then $\frac{f(e^\circ)}{f(e^*)} = \frac{x(e^\circ)}{1-\alpha-\beta} + \frac{\beta x(e^\circ)[E_{\delta,c}(e^\circ,q)-E_{\delta,c}(e^*,q)]}{\beta(1-\alpha-\beta)(1-\frac{\alpha}{\mu}-\frac{\beta}{\lambda})[E_{\delta,c}(e^*,e^*)-E_{\delta,c}(q,e^*)]}$.

Since $\frac{|E_{\delta,c}(e^\circ,q)-E_{\delta,c}(e^*,q)|}{E_{\delta,c}(e^*,e^*)-E_{\delta,c}(q,e^*)}$ is bounded for all e*, e and q it follows that $|\frac{f(e^\circ)}{f(e^*)}|$ is bounded for all $x \in V^\circ$.

Similarly for $q^\circ \in S_p \backslash Q(CC)$,

$\frac{f(q^\circ)}{f(e^*)} = \frac{[E_{\delta,c}(q,x)-E_{\delta,c}(x,x)]x(q^\circ)}{f(e^*)} + \frac{[E_{\delta,c}(q^\circ,x)-E_{\delta,c}(q,x)]x(q^\circ)}{\beta(1-\alpha-\beta)(1-\frac{\alpha}{\mu}-\frac{\beta}{\lambda})[E_{\delta,c}(e^*,e^*)-E_{\delta,c}(q,e^*)]}$. Since

$x(q^\circ) \le x(q)$, $|\frac{f(q)}{f(e^*)}|$ is bounded close to $L(\delta,c)$ and $\frac{|E_{\delta,c}(q^\circ,x)-E_{\delta,c}(q,x)|}{E_{\delta,c}(e^*,e^*)-E_{\delta,c}(q,e^*)}$ is bounded it follows again that $|\frac{f(q^\circ)}{f(e^*)}|$ is bounded.

We have shown that for any $(\delta,c) \in T \cap \{\delta > \delta_1\}$, any $x \in \Delta\{AA,CA,AC,CC\}$ with $x(\{AA,CA\}) = \gamma(DCC)$ is stable.

Looking back at step 1.2 together with step 3 it follows that any $x \in \Delta\{AA,CA,AC,CC\}$ with $0 < x(\{AA,CA\}) < \gamma(DCC)$ is stable. This completes the proof. □

In the following we will analyze whether there exist asymptotically stable sets given previous intruders that contain a pure strategy other than DD, AC or CC.

Similar to lemma 3.3, look-a-likes also play an important role for asymptotically stable sets given previous intruders that contain a pure strategy. The following lemma presents some necessary conditions for a pure strategy, say $e^\circ \in S_p$, to be contained in an asymptotically stable set given previous intruders for all parameters in some lex-patient set. It states that the lex-patient set of parameters can be chosen such that the best responses of any pure symmetric Nash equilibrium look-a-like of e° are precisely e°'s look-a-likes and that each of the Nash equilibrium look-a-likes of e° is contained in such a "nice" set.

LEMMA 5.5:

Let $e^\circ \in S_p$ and let $W \subseteq (0,1)\times(0,1)$ be a lex-patient set. Assume that $L(\delta,c)$ is an asymptotically stable set given previous intruders that contains e° for each $(\delta,c)\in W$. Then there exists a lex-patient set $W' \subseteq W$ such that for all $(\delta,c)\in W'$

i) $R_{\delta,c}(e^j)=Q(e^\circ)$ when $e^j \in R_{\delta,c}(e^j)\cap Q(e^\circ)$ and

ii) $\{x\in\Delta S_p \text{ s.t. } C(x)\subseteq R_{\delta,c}(x)\cap Q(e^\circ)\}\subseteq L(\delta,c)$.

PROOF:

We will first show that there exists a $v^*<1$ such that $R_{\delta,0}(e^\circ)=Q(e^\circ)$ for all $\delta\in(v^*,1)$. Let $(\delta,c)\in W$. Since e° is contained in an asymptotically stable set given previous intruders, (e°,e°) is a Nash equilibrium. Following the result of Abreu and Rubinstein [1] (see note before theorem 3.1), e° is also a symmetric Nash equilibrium strategy for $(\delta,0)$.

Let $Q'(e^\circ)$ be the set of all strategies $e^k\in\Delta S_p$ such that e^k plays just like e° when matched against e°. Let $e^k\in Q'e^\circ)\backslash Q(e^\circ)$. Since (e°,e°) is a Nash equilibrium and $c>0$ it follows that $|e^k|>|e^\circ|$ and hence $E_{\delta,c}(e^\circ,e^\circ)>E_{\delta,c}(e^k,e^\circ)$ for all $(\delta,c)\in(0,1)^2$. Therefore we can ignore strategies in $Q'(e^\circ)$ when proving the existence of $v^*<1$.

Let $e^k\in S_p\backslash Q'(e^\circ)$ and let n be the round in which for the first time e^k's play differs from e°'s when matched against e°. Then

$$\Delta(\delta):=E_{\delta,0}(e^\circ,e^\circ)-E_{\delta,0}(e^k,e^\circ)=\alpha(1-\delta)\delta^{n-1}+p(\delta)\delta^n\geq 0 \text{ where } \alpha\neq 0 \text{ and } p(\delta)=\sum_{m=0}^{\infty}\beta_m\delta^m.$$ The reason for $\alpha\neq 0$ is that in the Prisoners' Dilemma there is a unique best reply in the stage game. Note that $\Delta(\ )$ is continuous.

From the identity theorem for polynomials it follows that Δ cannot be identical to 0. Furthermore because two finite automata playing against each other must enter a cycle, $p(\ )$ is a rational polynomial of finite degree in δ and hence Δ cannot oscillate at $\delta=1$. This shows that $v(e^k)=\inf\{\delta\geq 0 \text{ s.t. } \Delta(\delta')>0 \text{ for all } \delta'\in(\delta,1)\}$ exists and that $v(e^k)<1$.

Let $v^*:=\sup\{v(e^k), e^k\in S_p\backslash Q'(e°)\}$. We will show that $v^*<1$ by showing that w.l.o.g. the supremum is only taken over finitely many e^k's. For $e^k\in S_p\backslash Q'(e°)$ the sign of Δ does not depend on the round in which e^k plays differently than e° for the first time. Therefore we may restrict our attention to automata that differ from e°'s behavior for the first time when e° reaches a state for the first time. Additionally the sign of Δ does not change if e^k plays like e° when e° is in the same state for the second time. With these two remarks and the fact that the automaton e° only has finitely many states it follows that $v^*<1$.

We now consider $c>0$. For $c>0$, $\delta\in(v^*,1)$ and $e^k\in S_p\backslash Q'(e°)$ let $k(\delta,e^k)=\sup\{c<1$ s.t. $E_{\delta,c}(e°,e°)>E_{\delta,c}(e^k,e°)\}$. By the continuity of the payoffs and the definition of v* it follows that $k(\delta,e^k)>0$. Let $k(\delta):=\sup\{c<1$ s.t. $R_{\delta,c}(e°)=Q(e°)\}=\inf\{k(\delta,e^i), e^i\in S_p\backslash Q'(e°)\}$. We will show that $k(\delta)>0$. Notice that if $|e^k|\geq|e°|$ then $E_{\delta,c}(e°,e°)>E_{\delta,c}(e^k,e°)$ holds for all $c>0$. Therefore $k(\delta)=\inf\{k(\delta,e^k)$ s.t. $e^k\in S_p\backslash Q'(e°)$ and $|e^k|<|e°|\}$. Since there are only finitely many $e^k\in S_p$ with less states than e° it follows that $k(\delta)>0$. Let $W(e°)=\{(\delta,c)$ s.t. $\delta\in(v^*,1)$, $c\in(0,k(\delta))\}$ then $W\cap W(e°)$ satisfies parts i) for e° (especially it is non empty).

We will now show that $\{x\in\Delta S_p$ s.t. $C(x)\subseteq R_{\delta,c}(x)\cap Q(e°)\}\subseteq L(\delta,c)$ for any $(\delta,c)\in W\cap W(e°)$. Let $J\subseteq S_p$ be the finite set such that $L(\delta,c)$ is an asymptotically stable set given previous intruders in J. For $y\in\Delta S_p$ such that $C(y)\subseteq R_{\delta,c}(y)\cap Q(e°)$ it follows that y is a weak ESS. Using lemma 4.2 and the fact that $\{x\in\Delta S_p$ s.t. $C(x)\subseteq R_{\delta,c}(x)\cap Q(e°)\}$ is connected, part ii) is shown.

Finally, just like e°, any $e^j\in S$ such that $e^j\in R_{\delta,c}(e^j)\cap Q(e°)$ is contained in $L(\delta,c)$ (part ii)). We now repeat the proof of part i) for each such e^j and obtain $W(e^j)$. Then with $W'=\{(\delta,c)$ s.t. $(\delta,c)\in W(e^j)$ whenever $e^j\in R_{\delta,c}(e^j)\cap Q(e°)\}\cap W$ (the intersection is over finitely many sets and therefore non empty), we obtain a W' that satisfies i) with $W'\subseteq W$.
□

We now come to the second major result in this paper. It states that under additional restrictions on the payoffs of the Prisoners' Dilemma, no pure strategy other than "Defect for Ever" (see theorem 3.4), "Tat for Tit" and its look-a-like AC (see theorem 5.4) is contained in an asymptotically stable set given previous intruders for all parameters in some lex-patient set.

THEOREM 5.6:

Assume $\pi(C,C)-\pi(D,D)>\pi(D,C)-\pi(C,C)>\pi(D,D)-\pi(C,D)$. Assume that $W\subseteq(0,1)\times(0,1)$ is a lex-patient set and $b^*\in S_p$ is contained in an asymptotically stable set given previous intruders for all $(\delta,c)\in W$. It follows that $b^*\in\{DD,AC,CC\}$.

PROOF:

Let the assumptions of the theorem hold. Let $W'\subseteq W$ be a lex-patient set satisfying the statement of lemma 5.5. By subtracting $\pi(D,D)$ from $\pi(r,t)$ for $r,t\in\{C,D\}$ w.l.o.g. we may assume that $\pi(D,D)=0$. Assume that $b^*\notin\{DD,AC,CC\}$.

Assume that b* has a cooperation cycle or that it has a cooperation node followed by a defection node in some part of its initial phase.

Then there exists $u,v\in\mathbb{N}$ such that b* plays the following strategies in the first u+v+1 rounds: u D's are followed by v C's that are followed by one D where $u\geq 1$ and if b* has a cooperation cycle then v=0, otherwise $v\geq 1$ (theorem 3.2 and $b^*\notin\{AC,CC\}$). Since each pure strategy Nash equilibrium look-a-like of b* must be in the asymptotically stable set given previous intruders (lemma 5.5), w.l.o.g. we may assume that b* punishes any play different from its own (except in the first round) to the first (initial) state (compare to the note after theorem 3.2). Additionally b* is such that its state in the second round is independent of its opponents play in the first. See figure 9 for a graphical representation of b*.

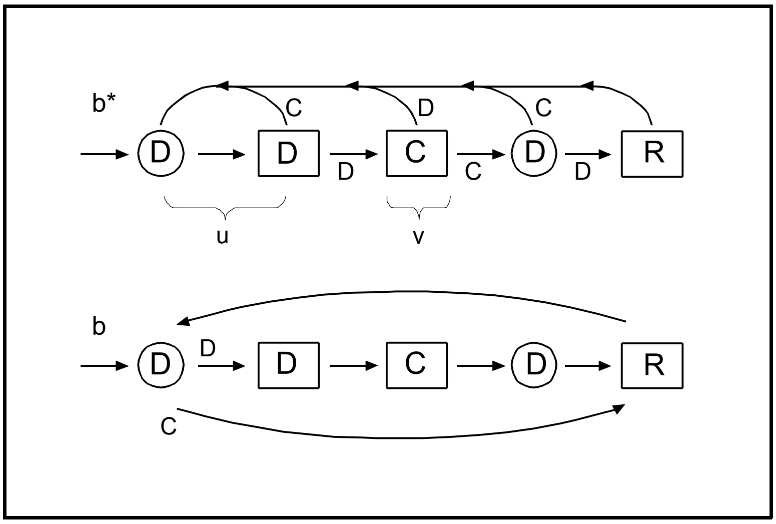

Figure 9: The symmetric Nash equilibrium strategy b*
and its toothless look-a-like b.

Let $b \in Q(b^*)$ be the look-a-like of b* that offers to jump to the state played in round $u+v+2$ if it receives C in the first round. The automaton b is graphed in figure 9.

Let $r=E_{1,0}(b^*,b^*)$. We will first show that
$$(u+v)r-v\pi(C,C)>-\pi(C,D). \tag{3}$$

Assume that $r=\pi(C,C)$. Then $v=0$ by assumption and equation (3) turns into $u\pi(C,C)>-\pi(C,D)$. From $2\pi(C,C)>\pi(D,C)$ and $\pi(D,C)+\pi(C,D)>\pi(C,C)$ it follows that $\pi(C,C)>-\pi(C,D)$ and since $u \geq 1$ (3) holds.

Now assume that $r<\pi(C,C)$ then $v \geq 1$ by assumption. In order for b* to be a symmetric Nash equilibrium strategy it must protect itself against an automaton that follows b*'s behavior in the first $u+v-1$ rounds and plays "D" in the $(u+v)$th round. Taking the limit as (δ,c) goes to $(1,0)$ we obtain that $r \geq [\pi(D,C)+(v-1)\pi(C,C)]/(u+v)$ which is equivalent to $(u+v)r-r\pi(C,C) \geq \pi(D,C)-\pi(C,C)$. Using the fact that $\pi(D,C)+\pi(C,D)>\pi(C,C)$ (3) follows.

We now aim to apply lemma A5. For this we will show that b is toothless and will find a toughest entrant of b* and b. In the following we will present two possible candidates for toughest entrants.

Candidate 1: Let $y \in S_p$ be an automaton that plays "C" in the first round and then imitates b's play (anticipating that b will jump to b*'s $u+v+2$'nd state) until it learns that it is matched against b*. After that it imitates b*'s play.

When y is matched against b and (δ,c) is close to $(1,0)$, $E_{\delta,c}(y,b)-E_{\delta,c}(b,b)$ is approximately equal to $\pi(C,D)+(u+v)r-v\pi(C,C)$. Following (3) there exists $\delta_1<1$ and $c_1>0$ such that $E_{\delta,c}(y,b)>E_{\delta,c}(b,b)$ for all $(\delta,c) \in (\delta_1,1)\times(0,c_1)$, especially b is toothless for these parameters.

Let $(\delta,c) \in W' \cap (\delta_1,1)\times(0,c_1)$. We will show that $E_{\delta,c}(b^*,y)<E_{\delta,c}(b,y)$.There exists a state s of y which is not the first in which b*'s behavior differs from b's when matched against y. Otherwise b* could shorten its initial phase. Assume that y plays "C" and b* plays "D" when y is in state s. Then b* punishes this deviation to its first state and y plays

the same strategies b* does from then on. Notice that this is as if an automaton played "D" when b* was in state s. By lemma 5.5 such a deviation does not payoff and therefore $E_{\delta,c}(b^*,y)<E_{\delta,c}(b,y)$.

Now consider the other case in which y plays "D" and b* plays "C" when y is in state s. As before this behavior will be translated to behavior against b*. Playing "C" in state s of b* is worse than playing "D" in the last cooperation state of b* prior to state s (on the equilibrium path of b*). This in turn does not pay off since b* is a Nash equilibrium strategy and therefore $E_{\delta,c}(b^*,y)<E_{\delta,c}(b,y)$.

Candidate 2: Let $z \in S_p$ be an automaton that plays C in the first round and then imitates b* until it learns that it is matched against b. Let k be the number of this round. Since z did not follow b's equilibrium path b punishes z back to b's first (initial) state. The automaton z is now assumed to play "C" and then imitate b's play. Since $E_{\delta,c}(y,b)>E_{\delta,c}(b,b)$ and $R(b^*)=Q(b^*)$, z maximizes its continuation payoff $P_{\delta,c}^{k+1}(z,b)$ in round k+1. Notice that $|z| \geq |b|$.

We will now show that $E_{\delta,c}(b,z)-E_{\delta,c}(b^*,z)>E_{\delta,c}(z,b)-E_{\delta,c}(b,b)$ if δ is sufficiently large. Since $E_{\delta,c}(b^*,z)=E_{\delta,c}(b,b)+(1-\delta)[\pi(D,C)-\pi(D,D)]$,
$\Delta(\delta)=[E_{\delta,c}(b,z)-E_{\delta,c}(b^*,z)]-[E_{\delta,c}(z,b)-E_{\delta,c}(b,b)]$
$=E_{\delta,c}(b,z)-E_{\delta,c}(z,b)-(1-\delta)[\pi(D,C)-\pi(D,D)]$
$=(1-\delta)[\pi(D,D)-\pi(C,D)+w\delta^{k-1}+(\pi(D,C)-\pi(C,D))\delta^k]+c(|z|-|b|)$ where
$w \in \{\pi(C,D)-\pi(D,C),\pi(D,C)-\pi(C,D)\}$. It follows that if δ is sufficiently large then $\Delta(\delta)>0$.

Consequently, δ is sufficiently large and $E_{\delta,c}(z,b)-E_{\delta,c}(b,b)>0$ then $E_{\delta,c}(b,z)-E_{\delta,c}(b^*,z)>0$.

Since $R(b^*)=Q(b^*)$, either y or z will be the toughest entrant of b* and b. In either case following the above calculations, the assumptions of lemma A5 are satisfied but its conclusion contradicted.

Therefore b* can neither have a cooperation cycle nor have a cooperation node followed by a defection node in its initial phase. If $b^* \notin \{DD,AC,CC\}$ and it does not have a cooperation cycle then b* will play "D" after it plays "C" in some round against itself. This must not be in the initial phase as assumed above. However, it can be shown that the inequalities above are not influenced by whether or not this occurs in the initial phase of b* and therefore the proof is complete.

☐

What happened on the border of L that caused it not to be a closed connected stable set when b* had more than one defection node? We constructed an entrant y (or z) that took advantage of one of its toothless look-a-likes b but at the same time did not punish b so that when matched against y, b got more than b* who is caught up in punishing y and thereby punishes himself.

In order to obtain this result we had to restrict the payoffs of the Prisoners' Dilemma ($\pi(D,C)-\pi(C,C)>\pi(D,D)-\pi(C,D)$ and $2\pi(C,C)>\pi(D,C)+\pi(D,D)$). The first of the two assumptions make a shortcut to the cooperation cycle possible when the look-a-like is matched against the mutant.

6. Discussion

Our aim was to analyze the repeated Prisoners' Dilemma played by finite automata in the continuous replicator dynamics. Especially we wanted to single out among the infinite set of pure symmetric Nash equilibrium strategies under patient lexicographic preferences (see [1]) the ones that are contained in a set that can "resist" small mutations. The payoffs of the Prisoners' Dilemma were specified such that "Tat for Tit" is a Nash equilibrium for the patient lexicographic preferences. The payoffs of the meta game were set up as the discounted repeated game payoffs minus a constant cost per state of the automaton used. It followed that "Defect for Ever" is the only pure strategy contained in a set that can resist an arbitrary mutation. We then searched for sets that can resist any mutation once some type(s) attempted to enter the population previously. Three automata ("Defect for Ever", "Tat for Tit" and its look-a-like AC) are contained in such a set when preferences are close to the patient lexicographic case. Furthermore, unlike "Defect for Ever", this property of the set containing "Tat for Tit" and AC does not vanish as the preferences approach the patient lexicographic preferences. Finally under an additional restriction on the payoffs of the Prisoners' Dilemma ($\pi(D,C)-\pi(C,C)>\pi(D,D)-\pi(C,D)$) there are no other sets that can resist any mutation given some previous mutations.

The first evolutionary analysis of the repeated Prisoners' Dilemma was undergone by Axelrod [2] in the form of an experiment. Although complexity considerations were not incorporated into the payoffs Axelrod [2] still argued that "Tit for Tat" won the tournament partially due to its simple structure.

Initial theoretical analysis of this repeated game followed the evolutionary stability theory developed by Maynard Smith and Price [13] and considered preferences based solely on repeated game payoffs. Consequently, strategies that behave differently off the equilibrium path cause Evolutionarily Stable Strategies (ESS) not to exist in this setup (see Boyd and Lorberbaum [9], Farell and Ware [10]). A further attempt along these lines was undergone by Kim [12] who discovered a Folk Theorem for the weaker definition of a limit ESS, a concept with no apparent connection to the replicator dynamics.

Binmore and Samuelson [5] were the first to include complexity considerations into an evolutionary analysis of the repeated Prisoners' Dilemma. They consider the meta game of choosing automata to play the repeated game. They select among the infinite pure symmetric Nash equilibria of the game with the help of their own static evolutionary concepts called modified Evolutionarily Stable Strategy (MESS) and polymorphous MESS. These concepts are adaptations of ESS and weak ESS to the lexicographic preferences on the repeated game payoffs and the number of states. Binmore and Samuelson [5] mainly use the limit of means payoff criterion for which they get substantial results.

When their concepts are applied to the discounted repeated game payoffs criterion it is easy to see that any pure symmetric Nash equilibrium strategy that punishes play that differs from its own to its first state is a MESS. Similarly polymorphous MESS can easily be found, especially {AC,CC} is a polymorphous MESS. So under discounted payoffs their concepts do not have enough bite, the set of payoffs achieved by MESS's is the same as that achieved by symmetric Nash equilibrium strategies.

In a different approach in the appendix of their paper Binmore and Samuelson [5] disregard complexity considerations and search for pure strategies that do not have an unstable strategy within $O(1-\delta)$ of them. They call such a pure strategy a GUESS. Although they do not present an example of such a strategy, they show that this concept has substantially more bite: each GUESS must have a cooperation cycle. Theorem 5.1 supplements their analysis, it implies that both "Tat for Tit" and AC are a GUESS when $2\pi(C,C)>\pi(D,C)+\pi(D,D)$ and δ is sufficiently large (set c=0 and use the continuity of the replicator dynamics).

Parallel to the development of the theory of our paper, Probst [14] modified the concept of a polymorphous MESS by assuming that a small frequency of "Cooperate for Ever" and "Defect for Ever" types are always apparent in the population. Probst [14] focuses on the limit of the means payoff criterion, just as most of [5] does.

The above two papers on the evolutionary properties of the repeated Prisoners' Dilemma with the number of states as the measure of complexity set up static conditions and calculate the set of strategies that satisfy them. As such they are models of equilibrium selection. Although the concepts are derived from the ESS condition, a relation to a dynamic adjustment process like the replicator dynamics is lacking. Various problems arise when trying to set up a continuous dynamic adjustment process like the replicator dynamics with their preference structure. These difficulties result from their use of the limit of means payoff criterion and of lexicographic preferences.

We argue that the limit of the means payoff criterion is not applicable in the context of evolutionary processes. Using the limit of the means payoff criterion means that the Prisoners' Dilemma is played an infinite number of times (with probability one). In the context of evolutionary processes each generation plays the repeated game. Since an infinite game cannot be played repeatedly this criterion seems inappropriate. We therefore chose to discount repeated game payoffs. The discount value δ can be considered a continuation probability and thereby making the expected number of rounds the repeated game is played finite (see also section 1). In this setup the repeated game can be played repeatedly.

Finally, the common argument of justifying the use of the limit of the means payoff criterion as the limit of the discounting criterion cannot be applied to dynamic stability analysis. Dynamic stability considerations are not interchangeable with taking the limit as the discount value tends to one which makes the respective results incomparable.

Why did we use a small cost of complexity instead of the patient lexicographic preferences used in [5]? Assuming these lexicographic preferences means that an individual prefers higher repeated game payoffs and only when indifferent prefers an automaton with less states. First of all there is no utility function representing these preferences, so the dynamics cannot be implemented directly. What would happen if our dynamic framework could be adjusted to select according to these preferences? Mutants never die out completely in the replicator dynamics. Consequently there will generally be types getting different payoffs which prevents the second condition to be implemented. Only in specific populations will this not be the case. So when applied to the dynamic analysis these lexicographic preferences ironically disregard complexity considerations in polymorphous populations because of arbitrarily small differences in the repeated game payoffs.

On the other hand, the assumption made in this paper of an arbitrarily small cost makes the complexity considerations bite when the individual is nearly indifferent between the repeated game payoffs. Taking the limit of the solutions of the analysis as the cost goes to zero is a method to incorporate the intuition behind the lexicographic preferences into a dynamic setting.

Under the assumptions of our paper $\Delta\{AC,CC\}$ is the unique "sensible" set emerging when taking the limit of the solutions as the cost goes to zero and the discount value goes to one. It can therefore be argued that $\Delta\{AC,CC\}$ is the unique set containing a pure strategy that can resist mutations in the evolutionary process driven by the replicator dynamics under preferences that approximate the patient lexicographic preferences.

What is the role of look-a-likes and "previous intruders"? Without complexity costs an ESS failed to exist because of look-a-likes (Boyd and Lorberbaum [9], Farell and Ware [10]). In the model with a cost for each state look-a-likes remain important for the analysis. Look-a-likes cannot be prevented from entering an asymptotically stable set and result in {"Defect for Ever"} being the only asymptotically stable set that contains a pure strategy. Look-a-likes can enter asymptotically stable sets given previous intruders too, however just as long as the population mean strategy is a Nash equilibrium. The object of the pure strategy Nash equilibrium is to protect the small frequencies of look-a-likes that can enter from being taken advantage of. In this assignment the "previous intruders" types help out. They limit the size of the connected set of rest points. How does this help for example in the case of "Tat for Tit" (CC)? In a neighborhood of $\Delta\{AC,CC\}$ the mean frequency of AC and CC will always increase when some non look-a-like of CC exists with small frequency. Thus the population is forced back to a limit with support on the look-a-likes of CC. If none of the types use an automaton of the class of DCC, the connected set of rest points containing CC might be so large that parts of it lie outside of the "gravitation field" of AC and CC, i.e., too close to CA and a mutation might lead away from the set.

This paper is the first explicit theoretical analysis of the dynamic adjustment process of a population playing a repeated game. As such it is a special example and might bring forward some critical remarks. In the following we wish to point out some details of the model and suggest possible lines for future research.

We focus our analysis mainly on "nice" sets that contain a pure strategy. This was done because it simplifies the analysis and because this was sufficient to make a comparison with related results. Abreu and Rubinstein [1] considered only pure strategy Nash equilibria because of "delicate issues" and problems of motivation of mixed strategies. Binmore and Samuelson [5] avoid sets without pure strategies by setting up a condition that forces pure strategies to exist in a polymorphous MESS.

Mixed strategy equilibria cannot be ruled out in a dynamic model because any population mean will generally be mixed and trajectories converging to a mixed strategy are generally the rule. Hence we aim to continue the present research in order also to analyze the nice sets that do not contain pure strategies.

Although mistakes in the form of mutations are allowed, the automata do not make mistakes when playing the repeated game. Introducing noise might make an interesting extension of the model.

Automata are frequently used to classify strategies. However the use of the number of states as a realistic measure of complexity has been repeatedly criticized. Alternative models using the number of transitions and the number of transitions and states have been analyzed (e.g. Banks and Sundaram [3]).

Much of the analysis relies on the smooth properties that arise through the fact that the population is infinite and time is continuous. As in the original evolutionary model we consider non overlapping generations. All individuals are matched simultaneously and the continuation probability δ smaller than one eventually causes them all to simultaneously stop playing the game and to reproduce. While in a one shot game this assumption is easier to justify, in repeated games played over a longer period of time intuition behind a mechanism stopping all of them from playing simultaneously is more difficult to imagine.

Due to the difficulty of the theoretical analysis of dynamic adjustment processes simplifications (some of which were mentioned above) that might be considered too simplistic are often necessary. Additionally the results are hard to convey in a few words. However the gain of such an analysis is that instead of presupposing the result by defining more or less arbitrary conditions the results of a dynamic analysis can often be surprising and open up new structures for equilibrium sets. An example is the emergence of the role of arbitrary small frequencies of DCC for the "stability" of our set.

REFERENCES:

1. D. Abreu and A. Rubinstein, The Structure of Nash Equilibrium in Repeated Games with Finite Automata, Econometrica 56 (1988), 1259-1282.

2. R. Axelrod, Effective Choice in the Prisoners' Dilemma, J. Conflict Resolution 24 (1980), 3-25.

3. J. Banks and R. Sundaram, Repeated Games, Finite Automata and Complexity, Games and Econ. Beh. 2 (1990), 97-117.

4. N. P. Bhatia and G. P. Szegö, "Stability Theory of Dynamical Systems," Springer-Verlag, Berlin, 1970.

5. K. Binmore and L. Samuelson, Evolutionary Stability in Repeated Games Played by Finite Automata, J. Econ. Theory 57(2) (1992), 278-305.

6. K. Binmore and L. Samuelson, "Muddling Through, Noisy Equilibrium Selection," Mimeo (1993).

7. T. Börgers and R. Sarin, "Learning Through Reinforcement and Replicator Dynamics," Mimeo (1993).

8. I. Bomze, Non-Cooperative 2-Person Games in Biology: a Classification, Int. J. Game Theory 15 (1986), 31-59.

9. R. Boyd and J. Lorberbaum, No Pure Strategy is Evolutionarily Stable in the Repeated Prisoners' Dilemma Game, Nature 327 (1987), 58-59.

10. J. Farell and R. Ware, Evolutionary Stability in the Repeated Prisoners' Dilemma Game, Theor. Pop. Biol. 36 (1988), 161-166.

11. J. Hopcraft and D. Ullman, "Introduction to Automata Theory, Language and Computation," Addison-Wesley, New York, 1979.

12. Y.-G. Kim, "Evolutionarily Stable Strategies in the Repeated Prisoners' Dilemma," Mimeo (1992).

13. J. Maynard Smith and G. Price, The Logic of Animal Conflicts, Nature 246 (1973), 15-18.

14. D. Probst, Evolution, "Automata and the Repeated Prisoners' Dilemma," Discussion Paper B-232, University Bonn, 1992.

15. A. Rubinstein, Finite Automata Play the Repeated Prisoners' Dilemma, J. Econ.

Theory 39 (1986), 83-96.

16. K. Schlag, "Evolutionary Stability in Games with Equivalent Strategies, Mixed Strategy Types and Asymmetries," Discussion Paper #912, Northwestern University, 1990.
17. K. Schlag, "Why Imitate, and if so, How?," Mimeo (1993).
18. J. Swinkels, Evolutionary Stability with Equilibrium Entrants, J. Econ. Theory 57 (1992), 306-332.
19. P. D. Taylor and L. B. Jonker, Evolutionary Stable Strategies and Game Dynamics, Math. Biosc. 40 (1978), 145-156.
20. B. Thomas, On Evolutionarily Stable Sets, J. Math. Biology 22 (1985), 105-115.
21. E. Van Damme, "Stability and perfection of Nash equilibria (2nd ed.)," Springer-Verlag, Berlin, 1991.
22. E. C. Zeeman, Population Dynamics from Game Theory, in "Global Theory of Dynamical Systems (Z. Nitecki, C. Robinson, Eds.)," Lecture Notes in Mathematics 819, 471-497, Springer-Verlag, Berlin, 1980.

A. Appendix

In this section we will state the popular "static" definitions ESS and weak ESS and quote the theorems relating them to the replicator dynamics. Let $\Gamma(S,E)$ be a symmetric game with the countable set of pure strategies S and the bilinear payoff function $E:\Delta S\times\Delta S\rightarrow\Re$.

DEFINITION: (Maynard Smith and Price [13])

A strategy $p\in\Delta S$ is called an Evolutionarily Stable Strategy (ESS) if for any $q\in\Delta S\backslash\{p\}$, $E(p,p)\geq E(q,p)$ where $E(p,p)=E(q,p)$ implies $E(p,q)>E(q,q)$.

It follows from the definition that if (p,p) is a strict Nash equilibrium (i.e., $\{p\}=R(p)$) then p is an ESS.

Consider the continuous replicator dynamics as defined in section 2.

THEOREM A1: (Taylor and Jonker [19], Zeeman [22])

If S is a finite set and $p\in\Delta S$ is an ESS then p is an asymptotically stable strategy.

DEFINITION: (Thomas [20])

A strategy $p\in\Delta S$ is called a weak Evolutionarily Stable Strategy (weak ESS) if for any $q\in\Delta S$, $E(p,p)\geq E(q,p)$ where $E(p,p)=E(q,p)$ implies $E(p,q)\geq E(q,q)$.

THEOREM A2: (Thomas [20])

If S is a finite set and $p\in\Delta S$ is a weak ESS then p is stable.

The following lemma is a stronger version of theorem A2 for countable infinite set of pure strategies.

LEMMA A3:

Let $p \in \Delta S$. If there exists $\epsilon > 0$ such that $E(p,x) \geq E(x,x)$ for all $x \in U_\epsilon(p)$ and equality implies that $C(x) \subseteq R(p)$ then p is stable.

The proof follows directly using the one by Thomas [20] that proves theorem A2.

LEMMA A4: (Bomze [8]) Let S be finite
i) If the trajectory starting in x converges to p then (p,p) is a Nash equilibrium in the game with the set of pure strategies restricted to C(x), i.e., $E(p,p) \geq E(e,p)$ for all $e \in C(x)$.
ii) If p is stable then (p,p) is a Nash equilibrium. The converse is false.

PROOF:

Part i) follows directly from the definition of the replicator dynamics.

Part ii): If (p,p) is not a Nash equilibrium then there exists an $a^i \in S$ such that $E(a^i,p) > E(p,p)$. Since p is a rest point it follows that $p(a^i)=0$. By continuity there exists an open neighborhood of p, U(p) such that $E(a^i,x) > E(x,x)$ for all $x \in U(p)$. Starting in $x \in U(p)$ with $x(a^i) > 0$ the frequency of a^i will increase as long as the population mean is in U(p). However this contradicts the fact that p is stable and $p(a^i)=0$.

A counterexample of a symmetric Nash equilibrium strategy that is not stable is easy to find (see e.g., [21, p.229]).

□

The following lemma is purely technical and relates to the payoff structure on the border of asymptotically stable sets given previous intruders. For any $e^j \in S$, let $Q(e^j)=\{v \in S \text{ s.t. } E(e^j,e^j)=E(e^j,v)=E(v,e^j)=E(v,v)\}$.

LEMMA A5:

Let S be finite, $L \subseteq \Delta S$ be an asymptotically stable set given previous intruders and let $e^i \in S$ be such that $e^i \in L$ and $R(e^i)=Q(e^i)$. Furthermore, let $e \in Q(e^i)$ be such that $e \notin R(e)$. Then $\beta^*=\min\{\beta \in (0,1) \text{ s.t. } R((1-\beta)e^i+\beta e) \backslash Q(e^i) \neq \varnothing\}$ exists and $E(e^i,q) \geq E(e,q)$ for all

$q \in R((1-\beta^*)e^i+\beta^* e) \setminus Q(e^i)$.

Such a $q \in R((1-\beta^*)e^i+\beta^* e) \setminus Q(e^i)$ we will call a toughest entrant of e^i and e.

PROOF:

Since $R(e^i)=Q(e^i)$ and $e \notin R(e)$, $\beta^* \in (0,1)$ exists. Let $L^\circ=\{(1-\alpha)e^i+\alpha e, 0 \leq \alpha \leq \beta^*\}$ and $v=(1-\beta^*)e^i+\beta^* e$. For any $x \in L^\circ$ it follows that $R(x)=Q(e^i)$ and therefore that x is a weak ESS. Since $v \in L$ together with Lemma A3 it follows that v is stable.

Let $q \in R((1-\beta^*)e^i+\beta^* e) \setminus Q(e^i)$. Assume that $E(e,q)>E(e^i,q)$ and consider the game and the dynamics restricted to $\Delta\{e^i,e,q\}$. By the definition of β^* and since $q \notin R(e^i)$ it follows that $E(q,(1-\alpha)e^i+\alpha e)>E(e^i,(1-\alpha)e^i+\alpha e)$ when $\alpha>\beta^*$. Therefore the symmetric Nash equilibrium strategies with support in $\{e^i,e\}$ are L°. Consider any trajectory starting arbitrarily close to v with $C(x)=\{e^i,e,q\}$ and $x(e)/x(e^i)>\beta^*/(1-\beta^*)$. Since e strictly dominates e^i in the interior of $\Delta\{e^i,e,q\}$, the frequency of e will always grow more than that of e^i. Together with part i) of lemma A4 it follows from the above that this trajectory will not converge to $\Delta\{e^i,e\}$. Therefore v cannot be stable in a population of types e^i, e and q. Together with the note after lemma 4.1 we get a contradiction.

□