

Discussion Paper No. 858

THE STRENGTH OF
A LITTLE PERFECTION

by

Ehud Kalai*

and

Alejandro Neme**

February 1989

(Revised September 1989)

* Department of Managerial Economics and Decision Sciences, J. L. Kellogg Graduate School of Management, Northwestern University, 2001 Sheridan Road, Evanston, Illinois 60208.

**IMASL, Universidad Nacional de San Luis, Chacabuco y Pedernera, 5700 San Luis, Argentina. This work was done while this author was visiting at the Department of Managerial Economics and Decision Sciences, J. L. Kellogg Graduate School of Management, Northwestern University.

Kalai's research is supported by National Science Foundation Economics Grant No. SES-8720342.

The authors wish to thank Ehud Lehrer, Roger Myerson, Ariel Rubinstein, and Nancy Stokey for helpful comments and discussions.

Abstract

"The Strength of a Little Perfection"

by Ehud Kalai and Alejandro Neme

The paper deals with three related issues.

1. It introduces a measure of partial subgame perfection for equilibria of repeated games.
2. It illustrates that the folk-theorem discontinuity generated by small complexity costs, as exhibited by Abreu and Rubinstein disappears in the presence of any level of perfection.
3. It shows that reactive strategy equilibria, such as tit-for-tat, cannot be subgame perfect, even partially so. As a corollary, this shows a need to use full automata rather than exact automata when studying complexity and perfection in repeated games.

1. Introduction

The contributions in this paper can be divided into three related parts. The first consists of the introduction of a measure of subgame perfection for equilibrium strategies of extensive form games. The second part deals with the effects of complexity costs on the set of partially-perfect equilibrium outcomes of repeated games. The third part illustrates that, generically, for repeated games, partially-perfect equilibria using reactive strategies do not exist. As a corollary, the last statement illustrates that when studying repeated games and perfection, the use of exact automata imposes a serious restriction.

Our measure of perfection ranks every Nash equilibrium in an extensive form game as being p -subgame perfect with $p = 0, 1, 2, \dots, \infty$. 0-subgame perfect equilibria are precisely Nash equilibria, and ∞ -subgame perfect strategies are the subgame perfect equilibria in the original sense of Selten (1975). Equilibria that are p -subgame perfect can be considered as partially perfect. It seems that, in some games, partial perfection is a more reasonable requirement than full perfection or no perfection. In this sense we are taking an approach opposite to Harsanyi and Selten (1988) who push the rationality assumption to its limits. Rather, we follow an approach similar to Aumann (1988) of relaxing the rationality assumption in an effort to allow theoretical predictions that better fit common sense predictions.

The idea behind partial perfection is to break away from the convention that all histories should be treated the same. Under existing game theory, when optimizing, one only distinguishes between two classes of histories:

those that are on the equilibrium path, when considering Nash equilibria, or the class of all histories, when considering full subgame perfection. Yet simple examples and recent new ideas in game theory regarding bounded rationality and forward induction (see, for example, Selten, 1978; Rosenthal, 1981; Kohlberg and Mertens, 1986; and Aumann, 1988; and see Binmore, 1987, for a general discussion and additional references), seem to indicate that past behavior is important for considerations of future behavior. In particular, in this paper we want to allow a weakening the requirement of full rationality (i.e., subgame perfection) after histories containing a large number of deviations from the equilibrium play path. Our notion is related to, yet different from, Marschak and Selten's (1978) notion of paraperfection.

In the following section we discuss and give examples that illustrate and motivate our notion of partial perfection. In Section 3 we give a formal definition of the measure of perfection for infinitely repeated games and show how it relates to the usual folk theorems for such games. (See Aumann and Shapley (1976), Rubinstein (1979), Fudenberg and Maskin (1986), and Ben-Porath and Peleg (1987).) An extension of partial subgame perfection to partial sequential equilibria as defined by Kreps and Wilson (1982) would be of interest. However, one would have to deal with a comparison of the likelihoods of several histories that lead to an information set before deciding on relaxing the assumption that agents who follow the information set must be maximizers.

The second part of this paper deals with issues of complexity in infinitely repeated games. Aumann's (1981) suggestion of using automata to deal with the complexity of repeated game strategies has been receiving

considerable attention over the last several years (see, for example, Neyman, 1985; Ben-Porath, 1986; Kalai and Stanford, 1988; see also Kalai 1989, and Sorin, 1988, for surveys of these results). In Rubinstein (1986) and Abreu and Rubinstein (1989), the definition of the game was modified to incorporate complexity costs into the payoffs of the players. They obtained striking results, showing that even if the complexity costs are minimal, i.e., considered as secondary to utility in a lexicographic ordering, the equilibria of the game are drastically changed. For example, in the prisoners' dilemma game, the set of payoffs of the equilibrium outcomes described by the folk theorems shrinks to a one-dimensional set consisting of two straight line segments.

We find, however, that it is the lack of perfection of the Abreu-Rubinstein players that brings about this discontinuity. Players who are first concerned with their payoffs after deviations from the equilibrium path and only then with minimizing their complexity costs, do generate the full folk theorems. We show that even a minimal degree of perfection (defined using our new measure of perfection) on the part of the players fully recovers the folk theorems in the prisoners' dilemma game. Thus, the discontinuity discussed by Abreu and Rubinstein exists, but it occurs as one moves from no perfection to a minimal level of perfection. These results are discussed in Section 5.

Another situation where a minimal level of perfection drastically affects the outcomes of the game is when the players are restricted to use only reactive strategies (see Aumann, 1981; Stanford, 1986; and Kalai-Samet-Stanford, 1988). A strategy of a player is reactive if, after any two histories that are identical in the actions of his opponents (his own may

differ), he "reacts" with the same action. For example, the famous tit-for-tat strategy in repeated prisoner's dilemma games is of this type. These strategies are of special interest now because of the increasing attention paid to playing games through exact automata, i.e., automata whose input consists only of the actions of the opponents. The results discussed in Section 6 show that, generically, in the presence of a minimal level of perfection, the restriction to reactive strategies is very limiting. In the infinitely repeated prisoners' dilemma game, for example, they must yield the noncooperative action repeatedly.

2. Examples of Partial Perfection

We consider first the following one shot two-person game.

	Nash	Middle	Safe
Nash	6,6	6,0	0,5
Middle	0,6	0,0	6,5
Safe	5,0	5,6	5,5

Notice that the only pure strategy Nash equilibrium of this game is the pair (Nash,Nash), yielding the payoffs (6,6). The safe strategy guarantees each player a payoff of 5, and the middle strategy is risky. Consider now the above game being repeated 200 times with the standard information structure where, prior to every stage, each player is told all the previous actions of both players. Since the only (pure) strategy Nash equilibrium of

the one shot game is the pair of actions (Nash,Nash), it is easily deduced by backwards induction that the only full subgame perfect equilibrium of the repeated game prescribes playing the pair of actions (Nash,Nash) after every history of action combinations.

Consider, however, a hypothetical history of length 100 along which player one took many different actions yet player two kept repeating her safe action. Consider now the forthcoming 101st play. The subgame perfectionist's assumption that both players will follow with the pair of actions (Nash,Nash) seems too strong. Player one will be concerned with having to assume that player two will play Nash and player two, even if she planned to play Nash, would be concerned with player one backing up from the Nash play because of his concern about her. It seems that the large number of earlier deviations from the equilibrium actions would give justification to the possibility of players breaking away from the (Nash,Nash) combination imposed on them by subgame perfection. Partially perfect equilibria for this game can be constructed where the players play Nash along the equilibrium path but play safe after a specified number of deviations were observed.

A multiperson example in which the number of deviations from the equilibrium path is important is described by the following game of mutual effort. Assume n players standing in a line. A hat is passed to the first player and he is offered the options of putting \$1 into it, or not. The hat is then passed to the next player, who has the same two options. The hat continues to be passed down the line in this way. Each player, knowing the choices of all his predecessors, has to choose between the same two options. After the last player has made his choice, the following payoffs materialize. If all n players participated, they each receive \$2, making a

net profit of \$1. Otherwise, nobody receives any money--i.e., the participating players lose a dollar and the other players break even.

It is easy to see by backwards induction that the only subgame perfect equilibrium is for all players to participate. But if $n \geq 2$, then a complete constant nonparticipation strategy (every player does not participate after every history) is also a Nash (not subgame perfect) equilibrium. Moreover, when n is large this nonparticipation equilibrium seems intuitively reasonable. The subgame perfectionists' challenge of this nonparticipation equilibrium rests on the irrationality of the last player after considering a hypothetical chain of $n - 1$ earlier deviations from nonparticipation. Thus, as n gets larger, more hypothetical deviations must be considered before the last player should consider changing his constant nonparticipation strategy. Also this large number of necessary deviations is of concern to earlier players. Their rationale in changing to participation depends crucially on later players' participation. Thus, not only should the last player consider earlier deviations, but earlier players considering participation should be concerned that the later players do not ignore hypothetical long chains of deviations and trust their followers to do so, and so on.

It seems that for large n , a small amount of irrationality in a sense similar to that of Aumann (1988) would make nonparticipation very rational and break the appealing rationale of subgame perfection. This irrationality will be placed on the common knowledge assumption that players consider arbitrarily long chains of deviations when testing the validity of a strategy.

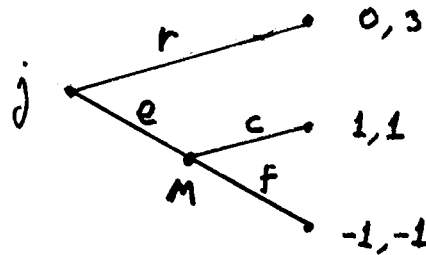
The measure of subgame perfection suggested in this paper supplies some

useful terminology for a discussion of the equilibria in the above examples. Given a strategy combination f we classify the histories of the games into classes $H_0 \subset H_1 \subset H_2 \dots$ described as follows. H_p contains all the histories during which there was a total of p or less deviations from f when adding up across all players and all periods. A strategy combination is defined to be p -subgame perfect if the players maximize their payoffs, i.e., play a Nash equilibrium, in all subgames following histories h in H_p but are not necessarily assumed to maximize after histories outside H_p . It is easy to see in our first example that if we allow strategies that are p -subgame perfect, with $p < 100$, then action combinations other than (Nash,Nash) are allowed after the history with many deviations described there. In the game of mutual effort the nonparticipation equilibrium can be seen to be $(n - 2)$ subgame perfect. And, indeed, as n gets larger this strategy becomes more intuitively appealing as well as more subgame perfect.

Our notion of partial perfection is similar to Marschak and Selten's (1978) notion of paraperfection. Their notion relaxes the assumption that an individual player is a maximizer after he has defected once from the equilibrium play path. With p -subgame perfection we allow p -deviations before the maximization assumption is disposed of. However, after observing the specified number of deviations, we dispose of the entire maximization assumption without targeting only the deviating players. In this sense, being 1-SPE is a weaker requirement than being paraperfect and results using an assumption of minimal perfection (1-SPE), as is done in Sections 5 and 6 of this paper, are stronger than their paraperfect counterparts. Other ways of discriminating among histories should be studied. One could discriminate according to the percentage of deviations in histories and could also take

into consideration the seriousness (in utility loss) of deviations. Again, our measure of 1-SPE as a way to measure a minimal level of perfection is weaker than the minimal perfection that one would get with these other measures.

An example of a different nature where partial perfection predicts "reasonable" equilibria is the well-known chain store paradox of Selten (1978). Here there is one major player, the chain store denoted by M, and 20 other players, $j = 1, 2, \dots, 20$, who are potential entrants to markets currently controlled by M. In each of 20 periods, $t = 1, 2, \dots, 20$, the chain store faces the player $j = t$ as a potential entrant. The two players' game between them is the following:



The story is that the j 'th player decides whether to enter (e) the market or refrain (r) from entering. If he enters, the chain store has to decide whether to fight him (f) or cooperate with him (c). After this stage game is completed with payoffs to j and M as indicated in the above figure, an identical game is played between $j + 1$ and M. In each stage there is perfect information about all the choices made by all previous players.

As was argued by Selten, the only subgame perfect equilibrium has all opponents $j = 1, 2, \dots, 20$ entering the market, and the chain store cooperating with all of them. However, as Selten pointed out, outcomes with the chain store fighting early entrants, no entry in the middle periods, and

entry and cooperation toward the end, seem reasonable. Partially perfect equilibria can indeed generate such outcomes. We can define p -perfect equilibria in which entry does not occur for the first $20 - p$ periods because the chain store will fight such early entrants, and only in the last p -periods opponents enter and the chain store cooperates. These strategies are defined as follows:

1. After every history containing a cooperating move by the chain store, opponents enter and the chain store cooperates.
2. After histories containing q entries that are all fought back by the chain store, the new opponents refrain from entry until the last $p - q$ periods. During these last $p - q$ periods, opponents enter and the chain store cooperates.

A straightforward induction argument shows that the above strategies constitute p -subgame perfect equilibria.

3. Notations and Conventions

Let $G = (A, u)$ be an n -person game, where A_i , the set of actions of player i , is a nonempty set and $A = \times_{i=1}^n A_i$; $u = (u_1, \dots, u_N)$ is a vector of utility functions with each $u_i: A \rightarrow \mathbb{R}$.

We first review the terminology and notations needed to describe the repeated game G with the average payoff criterion, \bar{G}^∞ , and the discounted repeated game, \bar{G}_α^∞ . Let $H = \bigcup_{m=0}^\infty H^m$ be the set of all histories where $H^m = A \times \dots \times A$, the m -fold Cartesian product of A , is the set of all histories of length m , and $H^0 = \{e\}$ is a singleton set consisting of the empty history. For both games, a strategy for player i is a function $f_i: H \rightarrow A_i$. Let F_i be the set of all individual strategies of player i , and

$F = \times_{i=1}^N F_i$ be the set of strategy vectors. A path is a sequence $p = (p(1), p(2), \dots, p(t), \dots)$, where $p(t) \in A$, $t = 1, 2, \dots$. Given a strategy vector $f \in F$, we define the path of f as follows:

$$p(f) = (p(f)(1), \dots, p(f)(t), \dots)$$

where $p(f)(1) = f(e)$ and $p(f)(t) = f(p(f)(1), \dots, p(f)(t-1))$. With the above construction we define the utility function for the game $\bar{G}^\infty(A, u)$:

$$\bar{u}_i(f) = \lim_{T \rightarrow \infty} 1/T \sum_{t=1}^T u_i(p(f)(t)).$$

When the above limit does not exist we can use the lim sup, lim inf, or Banach limits. In the cases of interest to us, the limit always exists and all these concepts coincide.

For the game G_α^∞ ($0 < \alpha < 1$):

$$u_i^\alpha(f) = (1 - \alpha) \sum_{t=1}^\infty \alpha^{t-1} u_i(p(f)(t))$$

for $i = 1, \dots, N$, and for all strategy vectors $f \in F$.

In everything that follows, we will use the notation G^∞ in a statement whenever the statement holds for both \bar{G}^∞ and G_α^∞ . Similarly, $u_i(f)$ will be used for both \bar{u}_i and u_i^α in their respective games.

For two histories h_1 and h_2 , of length m_1 and m_2 , respectively, we write $h_1 h_2$ to denote their concatenation, i.e., the history of length $m_1 + m_2$ starting with the elements of h_1 followed by those of h_2 . We make the convention that $eh = he = h$ for every history h .

Given an individual strategy f_i of G^∞ and a history h we let $f_i|h$ denote the strategy induced by f_i after the history h ,

$$f_i|h(\bar{h}) = f_i(h\bar{h}).$$

Similarly for a strategy vector f , $f|h = (f_1|h, \dots, f_N|h)$. A strategy vector f is a Nash equilibrium if for every i , f_i is a best response to f_{-i} , i.e., $u_i(f) \geq u_i(g_i; f_{-i})$ with $(g_i; f_{-i})$ denoting the strategy obtained from f by replacing f_i by g_i . f is a subgame perfect equilibrium if for every i , f_i is a perfect response to f_{-i} , i.e., for every history h , $f_i|h$ is a best response to $f_{-i}|h$. We define the complexity (see Kalai and Stanford, 1988) of a strategy f_i , $\text{comp } f_i$, to be the cardinality (number of states) of the set $\{f_i|h: h \in H\}$. We need to underline that the cardinality of the smallest automaton implementing f_i is $\text{comp}(f_i)$.

4. A Measure of Perfection

Given a strategy $f \in F$, we denote by $H_p(f)$ the set of all histories that have no more than p deviations from f , i.e.,

$$H_p(f) = \{h = (h^1, \dots, h^m) \in H: \sum_{t=1}^m D_t \leq p\} \text{ with}$$

$$D_t(h, f) = \#\{i: f_i(h^1, \dots, h^{t-1}) \neq h_i^t\} \text{ for } t \leq \text{length}(h).$$

Definition 1: A strategy $f = (f_1, \dots, f_N) \in F$ is a p-subgame perfect equilibrium, p-SPE (p a nonnegative integer), of the game $G^\infty(A, u)$ if for all i , f_i is a p-perfect response to f , i.e., $f_i|h$ is the best response to $f_{-i}|h$ for all $h \in H_p(f)$.

We let $\bar{u}(p\text{-SPE})$, $\bar{u}(\text{NE})$, $\bar{u}(\text{SPE})$ denote the sets of payoffs corresponding to the respective equilibrium concepts in the game \bar{G}^∞ , and the equivalent concepts with u^α , denote these sets in the game G_α^∞ .

Properties: Let G^∞ be an infinitely repeated game.

1. f is a Nash equilibrium if and only if it is 0-SPE;
2. If f is p -SPE, then it is $(p - 1)$ -SPE for $p = 1, 2, \dots$;
3. f is a SPE if and only if it is a p -SPE for $p = 0, 1, 2, \dots$;
4. for \bar{G}^∞ , $\bar{u}(p\text{-SPE}) = \bar{u}(\text{SPE})$ for $p = 0, 1, 2, \dots$;
5. for G_α^∞ , $\lim_{\alpha \rightarrow 1} u^\alpha(p\text{-SPE}) = \lim_{\alpha \rightarrow 1} u^\alpha(\text{SPE})$ for $p = 0, 1, 2, \dots$.

Properties 1, 2 and 3 follow immediately from the definitions.

Property 4 follows from the average payoff folk theorems (see Aumann and Shapley, 1976, and Rubinstein, 1979). Property 5 follows from the folk theorem with discounting (see Fudenberg and Maskin, 1986, and Ben-Porath and Peleg, 1987). Fudenberg and Maskin (1987) showed that the coincidence of $u^\alpha(\text{SPE})$ and $u^\alpha(\text{NE})$ occurs from some $\bar{\alpha}$ on with $\bar{\alpha} < 1$. Thus, the same coincidence occurs with all $u^\alpha(p\text{-SPE})$. Also, for a special case of the prisoners' dilemma game, $u^\alpha(\text{NE}) = u^\alpha(\text{SPE})$ for all α and therefore they must also equal $u^\alpha(p\text{-SPE})$ for all p . In general, however, we can have $u^\alpha(p\text{-SPE}) \not\supseteq u^\alpha((p + 1)\text{-SPE})$.

Abreu (1988) defined a concept of simple strategy and proved that, in the discounting case, every perfect equilibrium path is the outcome of some perfect simple strategies. It is easy to see that the structure of these strategies imply that together, Abreu's simplicity and 1-subgame perfection, are sufficient conditions for full subgame perfection.

We now define the width of a strategy vector, $w(f)$, a notion that is closely related to its degree of subgame perfection. This measure also captures a notion of simplicity closely related to Abreu's.

Recalling that $\bigcup_{p \geq 0} H_p(f) = H$, we let

$$w(f) = \min\{j: \{f|h: h \in H_j\} = \{f|h: h \in H\}\}$$

with the convention that $w(f) = \infty$ if the set of j 's above is empty. In other words, the width of f answers the question of how many deviations from the path of f must be considered in order to discover all the strategy vectors induced by f . It is easy to see that any constant strategy has width 0. But also any strategy vector that "uses only one path" has zero width. Strategy vectors consisting of Abreu's simple strategies all have width one. For example, any strategy vector that plays along a fixed path in the prisoners' dilemma game and trigger to the constant noncooperative strategy under any deviation has width one. But also strategies that trigger to different induced strategies, depending, for example, on when deviations occur, have width one.

Property 6: If f is of width w and it is a w -SPE, then it is SPE.

This property follows immediately from the definitions and shows, for example, that to check for full subgame perfection in an Abreu-type strategy vector, it suffices to check for 1-subgame perfection.

5. Perfection and Complexity Costs

In order to study the effects of complexity costs on the outcome of a

repeated game, Rubinstein (1986) and Abreu and Rubinstein (1989) incorporated those costs into the payoffs of the game and compared the new set of outcomes with the ones of the original unmodified game. Of a special interest to them was the case where the complexity costs are small relative to the overall payoff of the game. In order to do so they constructed lexicographic preferences for players who consider their payoffs first, but would prefer to attain any payoff using the least complex strategy. For the prisoners' dilemma game, they showed that even this minimal lexicographic modification of preferences has severe reduction in the set of equilibrium payoffs. It changes from the full set of feasible individually rational payoffs to a one-dimensional subset of it consisting of one or two straight line segments.

Abreu and Rubinstein restricted their players to use finite complexity strategies. Neme and Quintas (1988) showed that without this restriction the discontinuity disappears.

Our purpose in this section is to shed some additional light on this discontinuity phenomenon. Our players, unlike those of Abreu-Rubinstein, would incorporate the possibilities of deviations, or mistakes, into their lexicographic preferences, and would place the complexity costs further down the line in the consideration. More specifically, a player would prefer strategy f_i to \bar{f}_i as a reply to f_{-i} if f_i yields him a higher payoff after every history of the game (with at least one strict inequality), or if f_i and \bar{f}_i yield exactly the same payoffs after all histories, but the complexity of f_i is strictly smaller. In other words, these players are first concerned with the payoffs they receive, in general or even after deviations, and only after that with saving on complexity costs. We will

show that for these types of players, the full folk theorems are recovered.

One may be concerned that we are going to an extreme opposite the one of Abreu-Rubinstein's who did not have their players guard their payoffs after deviations at all. It may be reasonable to be concerned with one's payoffs after a small number of deviations. However, to consider any unbounded number of deviations before bringing the complexity costs down may be unreasonable. The partial perfection concept defined in the previous section enables us to deal with this issue. We will show that players who put primary importance on their payoffs at equilibrium or after a small number of deviations (even one) and only lexicographic secondary importance on their complexity costs will generate the full set of equilibria described by the folk theorem.

Definition 2: A strategy $f = (f_1, \dots, f_N)$ is a simplified p-subgame perfect equilibrium (simplified p-SPE) of $G^\infty(A, u)$ if it is p-SPE and $\text{comp } f_i \leq \text{comp } g_i$ for all g_i such that $g_i|h$ is best response to $f_{-i}|h$ for every $h \in H_p(f)$.

Properties: Let G^∞ be an infinitely repeated game.

- i. The set of simplified 0-SPE coincide with the Nash equilibria of the lexicographic preferences defined by Abreu and Rubinstein.
- ii. If f is a simplified p-SPE and it is also $(p + 1)$ -SPE then it is a simplified $(p + 1)$ -SPE.

However, unlike the sets of p-SPE, there is no monotonicity in p of the sets of (simplified p-SPE).

Let P be the prisoners' dilemma game, represented by the following

bimatrix:

	C	N
C	1, 1	a, d
N	d, a	0, 0

Figure 1

such that $a < 0$ and $d > 1$. We let $A_j = \{C, N\}$, $A = A_1 \times A_2$ be the set of action combinations, and $u_i: A \rightarrow \mathbb{R}$ be the utility functions.

We define the following sets of payoff vectors:

$$U = \{u(a): a \in A\},$$

$$V^* = \{x: x = \sum_{j=1}^p r_j x_j, x_j \in U, r_j \in Q_+, j = 1, \dots, p, \sum_{j=1}^p r_j = 1\}$$

where Q_+ denotes the set of nonnegative rational numbers.

$V = V^* \cap \mathbb{R}_{++}^2$ is the set of individually rational feasible payoff combinations with rational coefficients. While the Abreu-Rubinstein results show that $\lim_{\alpha \rightarrow 1} u^\alpha$ (simplified 0-SPE) is a very small subset of V , the following theorem shows that for $p = 1, 2, \dots, \infty$, $\lim_{\alpha \rightarrow 1} u^\alpha$ (simplified p-SPE) = V .

Theorem 1: Consider the repeated prisoners' dilemma game with discounting, P_α^∞ . For all $x \in V$ there exists a strategy $f = (f_1, f_2)$ and $\bar{\alpha} < 1$ such that:

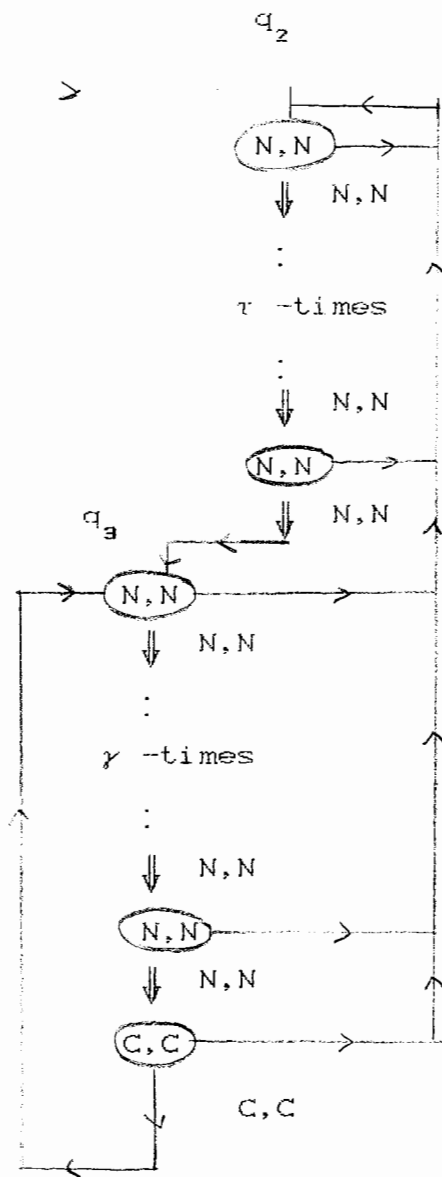
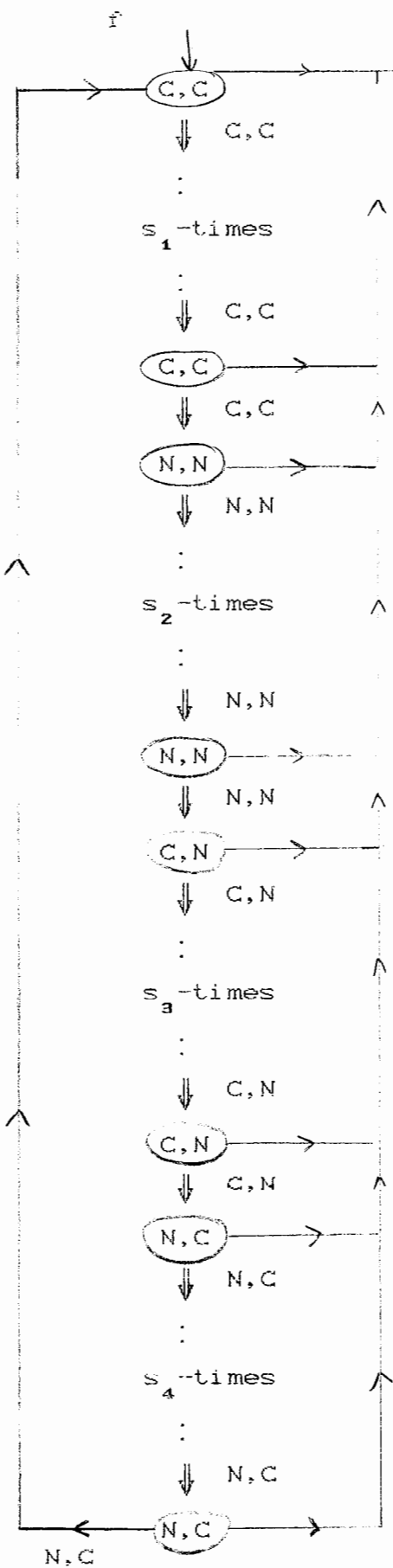
(i) $\lim_{\alpha \rightarrow 1} u_i^\alpha(f) = x_i$, $i = 1, 2$; (ii) for $p = 1, 2, \dots$ if $\bar{\alpha} < \alpha < 1$, f is a simplified p -SPE; (iii) $\text{comp } f_i < \infty$, and with $\text{width}(f_i) = 1$, $i = 1, 2, \dots$.

Proof: Because $x \in V$ we can find nonnegative integers s_1, s_2, s_3, s_4 and b such that

$$x = \sum_{j=1}^4 s_j y_j / b \text{ with } b = \sum_{j=1}^4 s_j > 0$$

and $(y_1, y_2, y_3, y_4) = ((1, 1), (0, 0), (a, d), (d, a))$. We describe the strategy vector $f = (f_1, f_2)$ required for the proof of the theorem by a pair of (full) automata (see Kalai and Stanford, 1988). Since the automata corresponding to f_1 and f_2 have the same number of states and identical transition functions, we describe them in one graph.

The graph should be interpreted as follows. Circles correspond to states of automata. The pairs of actions written in the circles represent, respectively, the actions that players one and two take at their corresponding states. Thus, if we cover all the second coordinates inside the circles we obtain the automaton description of the strategy f_1 . Similarly, covering the first coordinates will describe f_2 . Arrows in the graph represent the transition rules of the two automata. A labeled arrow coming out of the state represents the next state the automata move to if the labeled pair of actions were taken. The unlabeled arrows coming out of a state represents the transition rules following all action combinations that were not covered by labeled ones. Notice that all of our labeled arrows correspond to paths produced by states where the unlabeled arrows represent deviations. So f_1 , for example, prescribes the following actions



to player one. Starting at the top left state it prescribes cooperation. Then if it observes a pair of cooperative actions it moves down to a similar cooperating state. However, if it observed any deviation from (C,C), by either player, it moves to the right to play the induced strategy g . In g it starts by a state recommending noncooperation to player one and moving down when (N,N) is observed but back to itself otherwise.

Notice that the play path of f yields in the limit (as $\alpha \rightarrow 1$) the payoff x , and that the complexity of $f_i < \infty$ and width $(f_i) = 1$ for $i = 1, 2, \dots$. Thus we only have to prove that f is a simplified p-SPE for $p = 1, 2, \dots$. (The strategy q_2 is of the type used by Abreu and Rubinstein (1989) in their constructions.)

We choose γ to satisfy $\gamma > d$, $\gamma > 2b$, $x_i > u_i(C,C)/\gamma + 1$, and $\tau \geq 1$. Thus, for sufficiently large α , $u_i^\alpha(q_2) < u_i^\alpha(q_3) < u_i^\alpha(q_1)$, for $i = 1, 2$.

Observation 1: For a sufficiently large α , starting at any combination of states of the path of f (the circles directly under q_1) yields both players a higher payoff than starting at any other combination of states (not directly under q_1). Also for sufficiently large α , starting on any circle on the path of q_3 yields every player a higher payoff than starting on the circles of q_2 before entering q_3 . The circles of q_2 before q_3 yield decreasing payoffs as we move down.

From Observation 1 we obtain:

Observation 2: For sufficiently large α , f is a SPE and therefore a p-SPE for $p = 0, 1, 2, \dots$

It follows from Observation 2 and the second property of a simplified p-SPE that it suffices to show that f is a simplified 1-SPE. Because of the symmetry of the players it now suffices to show that if \bar{f}_1 is a best response to f_2 after all histories $h \in H_1(f)$ then $\text{comp}(\bar{f}_1) \geq \text{comp}(f_1)$. Notice that any of the circles in Figure 2 can be reached after histories in $H_1(f)$. Thus, the following claim will complete the proof.

Claim: If a strategy \bar{f}_1 satisfies $u(\bar{f}_1|h, f_2|h) = u(f|h)$ for all $h \in H_1(f)$, then for any two histories $h, h' \in H_1(f)$ leading to different circle in Figure 2, $\bar{f}_1|h \neq \bar{f}_1|h'$.

Notice that the path of $(\bar{f}_1|h, f_2|h)$ and $f|h$ agree for all $h \in H_1(f)$.

Suppose that $\bar{f}_1|h = \bar{f}_1|h'$. Because the paths of $f|h, f|h'$ are different, we can find a nonnegative integer number t_0 and a history $\tilde{h} = (\tilde{h}_1, \dots, \tilde{h}_{t_0})$ such that:

$$\begin{aligned} (\bar{f}_1, f_2)(h) &= \tilde{h}_1 = (\bar{f}_1, f_2)(h') \\ (\bar{f}_1, f_2)(h(\tilde{h}_1, \dots, \tilde{h}_{t_0-1})) &= \tilde{h}_{t_0} = (\bar{f}_1, f_2)(h'(\tilde{h}_1, \dots, \tilde{h}_{t_0-1})) \\ f_2(h\tilde{h}) &\neq f_2(h'\tilde{h}). \end{aligned}$$

Clearly the claim holds for pairs h, h' both leading to different circles in the path of q_2 .

Consider now the case when one history, say h , lead to a circle in q_1 , and the other one, h' , leads to a circle in q_2 .

Since $\hat{h} = \tilde{h}(f(h'\tilde{h}))$ has one deviation for $f|h$, then we have that $f|h\hat{h}$

starts in the first circle in q_2 . The history \hat{h} does not have deviations for $f|h'$, hence $h|\hat{h}$ leads to a circle in q_2 which is different from the first.

Because $h\hat{h}, h'h \in H_1(f)$ and lead to different circles in q_2 we have that $\bar{f}_1|h\hat{h} \neq \bar{f}_1|h'h$. So $\bar{f}_1|h \neq \bar{f}_1|h'$.

The remaining case is when h, h' lead to a different circles in q_1 . In this case, $f|h\hat{h}$ starts in a circle in q_2 and $f|h'h$ does it in q_1 . But this case was already analyzed. []

We now move to the case of the repeated prisoners' dilemma game but with the average payoff criterion, \bar{P}^∞ and show again the folk theorem with simplified equilibria.

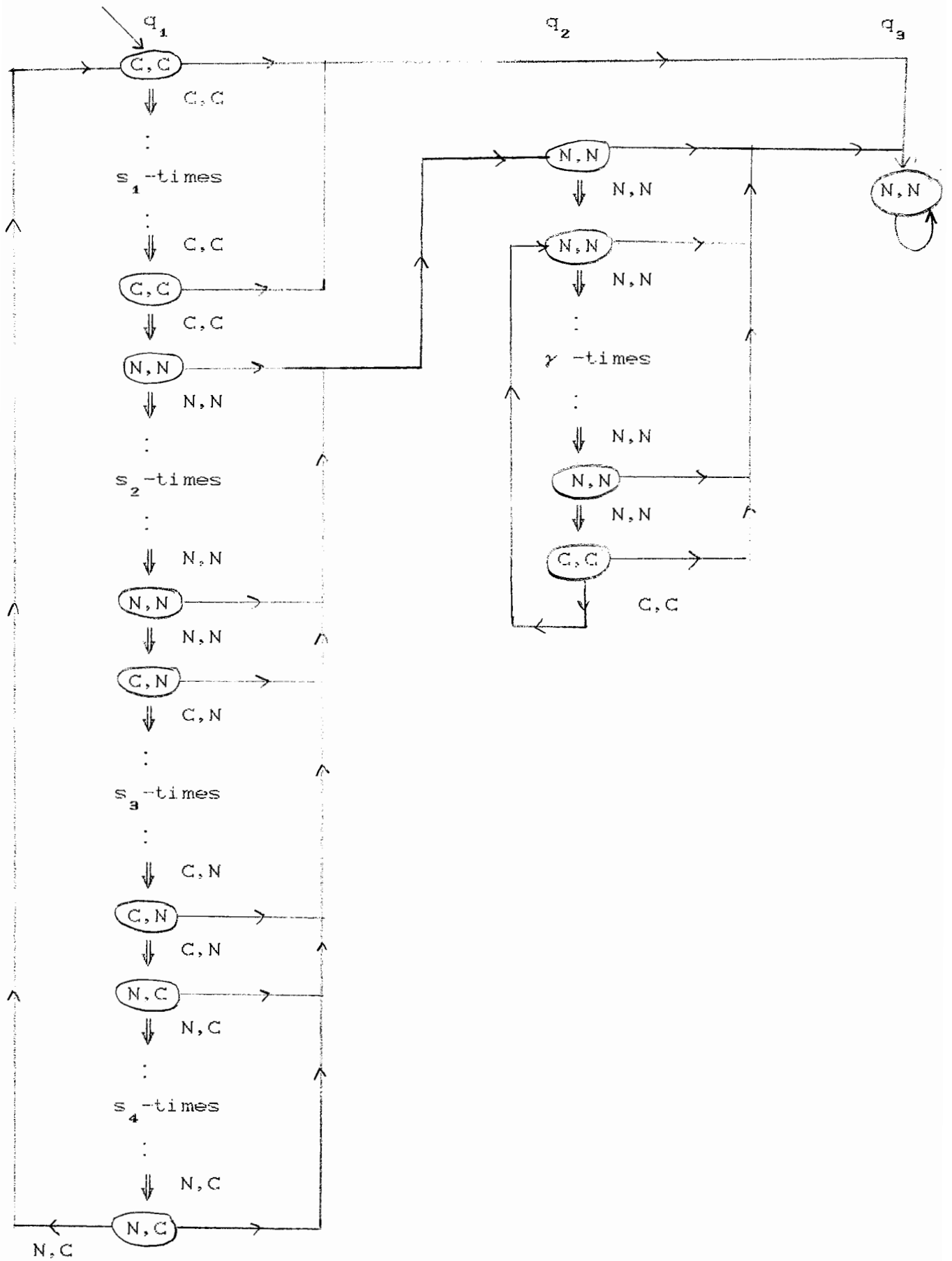
Theorem 2: Consider the repeated prisoners' dilemma game with the average payoff criterion, \bar{P}^∞ . For all $x \in V$ there exists a strategy vector $f = (f_1, f_2)$ satisfying: (i) $u(f) = x$; (ii) for $p = 1, 2, \dots$ f is a simplified p -SPE; and (iii) $\text{comp } f_i < \infty$, and $\text{width}(f_i) = 1$, $i = 1, 2$.

Proof: We construct the pair of equilibrium strategies f described by the following diagram. The proof is very similar to that of Theorem 1.

The graph is interpreted as in Theorem 1. Notice that the primary path of q_1 yields the payoff x , the complexity of $f_i < \infty$, and the width $(f_i) = 1$ for $i = 1, 2$.

Choosing γ sufficiently large, we have $u_i(q_3) < u_i(q_2) < u_i(q_1)$. Clearly, f is a SPE and therefore p -SPE for $p = 0, 1, 2, \dots$

By using an argument similar to that of Theorem 1, it suffices to show



that if \bar{f}_1 is a best response to f_2 after all histories $h \in H_1(f)$ then $\text{comp } \bar{f}_1 \geq \text{comp } f_1$. For such a strategy the following observations hold.

Observation 1: Let $h \in H_1(f)$ be a history leading f to a circle in q_1 or q_2 . Then the path of $f|h$ and $(\bar{f}_1|h, f_2|h)$ are the same.

Observation 2: Because the payoff $u(f|(C,N) = u(\bar{f}_1|(C,N), f_2|(C,N)) = 0$, there exists a history h^* such that the path of $(\bar{f}_1|h^*, f_2|h^*)$ is (N,N) for ever. The following claim will complete the proof.

Claim: Let \bar{f}_1 be a strategy satisfying $u(\bar{f}_1|h, f_2|h) = u(f|h)$ for all $h \in H_1(f)$: (i) for any history $h \in H_1(f)$ such that $f|h$ starts in the same circle on q_1 or q_2 ; then $\bar{f}_1|h \neq \bar{f}_1|h^*$; and (ii) for any two histories $h, h' \in H_1(f)$ leading to a different circle in q_1, q_2 , $\bar{f}_1|h \neq \bar{f}_1|h'$.

(i) Suppose that $\bar{f}_1|h = \bar{f}_1|h^*$. Since $\bar{f}_1(h) = N$ the strategy $f|h$ does not start in q_1' . If $f|h$ starts in q_2 , there exist $\tilde{h} = ((N,N), \dots, (N,N))$ such that $(\bar{f}_1, f_2)(h\tilde{h}) = (C,C)$. By Observation 2, $\bar{f}_1|h \neq \bar{f}_1|h^*$. Suppose that h leads to a circle in q_1'' , then there exists $\tilde{h} = ((N,N), \dots, (N,N))$ such that $h\tilde{h} \in H_1(f)$ and $f|h\tilde{h}$ starts in a circle in q_2 . So, by Observation 2, $\bar{f}_1|h \neq \bar{f}_1|h^*$.

(ii) Suppose that $\bar{f}_1|h = \bar{f}_1|h'$. By the same argument used in the proof of Theorem 1, we can assume, without loss of generality, that $f_2(h) \neq f_2(h')$. Clearly, the claim holds for pairs h, h' , both leading to a different circle in the path of q_1' , or both in the path of q_2 , or when one

history, h , leads to a circle in q_1'' , and the other one, h' , leads to a circle in q_2 .

Consider now when the history h leads to a circle in q_2'' and h' leads to a circle in q_2 . The action $(\bar{f}_1, f_2)(h')$ is a deviation for $f|h$, then $f|h(\bar{f}_1, f_2)(h')$ starts in the first circle in q_2 and $f|h'(\bar{f}_1, f_2)(h')$ starts in q_2 , but in a different circle. The histories $h(\bar{f}_1, f_2)(h'), h'(\bar{f}_1, f_2)(h') \in H_1(f)$, then $\bar{f}_1|h \neq \bar{f}_1|h'$.

The remaining cases are when the history $h \in H_1(f)$ leads to a circle in q_1'' and the other one $h' \in H_1(f)$ leads to a different circle in q_1'' , or a circle in q_1' .

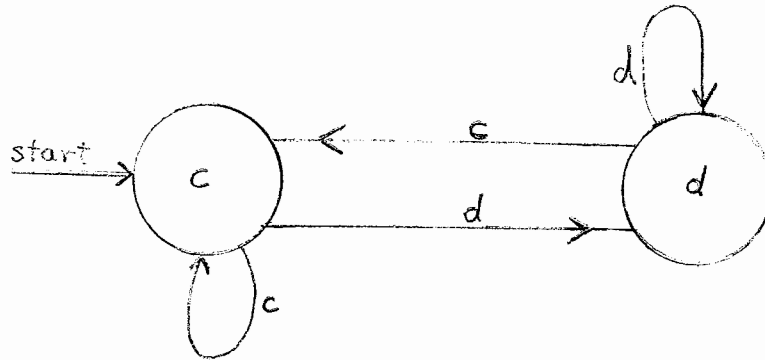
The action $(\bar{f}_1, f_2)(h')$ is a deviation for $f|h$. Then the history $h(\bar{f}_1, f_2)(h')$ leads to a circle in q_2 . On the other hand, the history $h'(\bar{f}_1, f_2)(h')$ leads to a circle in q_1' or q_1'' , because the action $(\bar{f}_1, f_2)(h')$ is not a deviation for $f|h'$. So, by the above cases, $\bar{f}_1|h \neq \bar{f}_1|h'$. []

6. Exact Automata, Full Automata, and Reactive Strategies

Two types of automata have been used in studies of strategic complexity in repeated games (see Kalai, 1987). The first type, exact automata, were suggested by Aumann (1981). Full automata, on the other hand, were introduced by Kalai-Stanford (1988). The difference between the two is that the input for an exact automaton of player i consists of action combinations of i 's opponents, excluding his own actions. On the other hand, the input to a full automaton, consists of the action combinations of all players, including i himself. This difference, which may seem minor at first, turns out to be important when we study issues of subgame perfection. This is the

topic of this section.

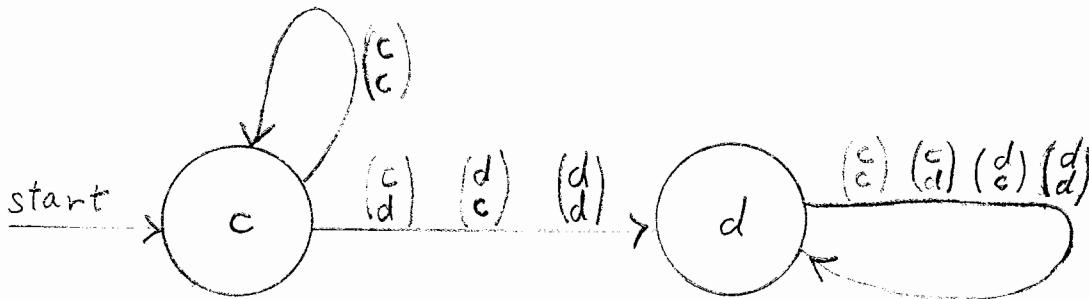
To illustrate the two types of automata, we use the following examples from the infinitely repeated prisoners' dilemma game. The "cooperate and then tit-for-tat your opponent" strategy for player one, c-tft, is described by the following exact automaton:



Here there are two states (of mind) of player 1, represented by the two circles. In the first and starting state he cooperates, while in the second he defects, as illustrated by the letters inside the circles. The letters along the arcs represent the actions of player 2; the arcs themselves represent the transitions of player 1 among his states. It is easy to see that this automaton, when it starts at the cooperative state, would play the c-tft strategy. Notice that if we start this automaton at the other state it would play a d-tft strategy.

For games with the average payoff criterion and for almost all discount parameters (with the exception of at most one, see Kalai-Samet-Stanford, 1988), a pair of c-tft strategies is not subgame perfect. The strategies that are used to exhibit subgame perfect equilibria are often the grim-trigger ones. Under such a strategy a player, say player 1, starts by cooperating and continues to do so after all histories which are fully

cooperative. If any defection from cooperation by any player (including himself) is observed anywhere along the history, then player 1 defects. Since player one's actions may depend on his own previous action, we must have both players' actions as input to his (full) automaton. A full automaton describing his grim-trigger strategies is given by the following figure.



Notice that only when both players cooperate does player 1 stay in the cooperative state: all other action combinations move him to the defecting state, which is absorbing.

One may think of exact automata as ones that assume that their own prescriptions are always followed and that deviations from them are impossible. Full automata, on the other hand, acknowledge the possibility of deviations and their input includes the information about such deviations. Full automata also allow for studying subgame perfection.

It turns out, however, that for the usual notion of Nash equilibria, the distinction between exact and full automata is irrelevant. Starting with an exact automaton of a player, we can define a complete strategy on the entire set of histories that extends the prescriptions of the automaton in a natural way. Given a history of action combinations, we let the

automaton follow the history, transiting through its states as if its recommendations were followed (even if they were not) and in effect ignoring the part of the input describing the player's own actions. After it reads the last action combination and it makes its transition, the player adopts the action prescribed by the last state of the automaton.

It is easy to see that his procedure yields strategies which we call reactive; see Aumann (1981), Stanford (1986), and Kalai-Samet-Stanford (1988). More formally, they are defined as follows.

Definition: A strategy of player i , f_i , is reactive if for any two histories, $h = h^1 h^2 \dots h^m$ and $\bar{h} = \bar{h}^1 \bar{h}^2 \dots \bar{h}^m$, $f_i(h) = f_i(\bar{h})$ whenever $h_{-i}^t = \bar{h}_{-i}^t$ for $t = 1, 2, \dots, m$, i.e., whenever the two histories coincide on the actions of the players other than i .

Any non-artificial extension of the prescriptions of an exact automaton to a full strategy yields a reactive strategy. So in order to study the scope of what can be attained by exact automata in a repeated game, we switch now to a discussion of what can be attained when we restrict ourselves to reactive strategies. We first show that for Nash equilibria (0-subgame perfect equilibria), the restriction does not matter. We consider any two-person infinitely repeated game.

Proposition 1: The payoff set of reactive (simplified) Nash equilibria coincides with the payoff set of (resp. simplified) Nash equilibria.

Proof: Let \mathcal{F} be the set of all strategies and R be the subset of reactive

strategies of the infinitely repeated game. We only have to show that the payoffs obtained by equilibria using any strategies can be obtained by equilibria using reactive strategies. Let $\psi_i: \mathcal{F} \rightarrow \mathbb{R}$ with $\psi_i(f_i)(h^1, \dots, h^n) = f_i(\tilde{h}^1, \dots, \tilde{h}^n)$ where $\tilde{h}^1 = (f_i(e), h_{-i}^1)$, $\tilde{h}^2 = (f_i(\tilde{h}^1), h_{-i}^2), \dots, \tilde{h}^n = (f_i(\tilde{h}_1, \dots, \tilde{h}^{n-1}), h_{-i}^n)$.

For all $f_i \in \mathcal{F}$, $i = 1, 2$:

- (i) $\text{comp } f_i \geq \text{comp } \psi_i(f_i)$.
- (ii) $P(f_1, f_2)(t) = P(\psi_1(f_1), \psi_2(f_2))(t)$ for all t (recall that $p(f)$ is the path resulting from f).

Consider $f = (f_1, f_2)$ a Nash equilibrium and assume that $(\psi_1(f_1), \psi_2(f_2))$ is not a Nash equilibrium. Then for some $g_i \in \mathcal{F}$, $u_i(g_i, \psi_{-i}(f_{-i})) > u_i(\psi_1(f_1), \psi_2(f_2))$. But then, by property (ii), $u_i(g_i, f_{-i}) > u_i(f_1, f_2)$, so $(\psi_1(f_1), \psi_2(f_2))$ is a reactive strategy Nash equilibrium yielding the same payoffs as f . Now consider a simplified Nash equilibrium f and suppose that $\psi(f) = (\psi_1(f_1), \psi_2(f_2))$ is not a simplified Nash equilibrium. Since $\psi(f)$ is a Nash equilibrium, there is a g_i with $u_i(g_i, \psi_{-i}(f_{-i})) = u_i(\psi(f))$ and $\text{comp}(g_i) < \text{comp}(\psi_i(f_i))$. But then, by (ii), $\text{comp}(g_i) < \text{comp}(f_i)$, $u_i(\psi_i(g_i), \psi_{-i}(f_{-i})) = u_i(g_i, f_{-i})$, yet, by (i), $u_i(g_i, f_{-i}) = u_i(f)$, so f is not simplified, a contradiction. []

The above proposition shows that all the payoffs (even the paths) that can be attained with Nash equilibria can be attained with Nash equilibria using only reactive strategies. And the same result holds for simplified Nash equilibria. Thus, the analysis of efficient automata playing the repeated game can be conducted with exact or full automata. Our next proposition illustrates that when a minimal level of perfection is

introduced to the discussion, the restriction to reactive strategies becomes severe. The proposition is a straightforward generalization of the main result in Kalai-Samet-Stanford (1988). We consider the infinitely repeated prisoners' dilemma game with discount parameters α , $0 < \alpha < 1$. A pair of strategies, $f = (f_1, f_2)$ is a discount robust equilibrium if it is an equilibrium of the game for all discount parameters β in some neighborhood of α . Thus, a non-discount robust equilibrium is a knife-edge situation that works only for isolated discount parameters.

Proposition 2: If f is a discount robust p -SPE using reactive strategies for some $p \geq 1$, then the path of f consists of repeated plays of the noncooperative action d .

Proof: The proof is essentially the same as the one for the case of full perfection. Assuming that f satisfies the conditions of the proposition for $p = 1$, then for any two histories, h and \bar{h} , that coincide on the actions of, say, player 1, $f_1|h = f_1|\bar{h}$. Therefore, by 1-SPE of f if $h, \bar{h} \in H_1$, $u_2(f|h) = u_2(f|\bar{h})$. Since $u_2(f|h)$ is a geometric series in the discount parameter, if f is discount robust then the equality of $u_2(f|h) = u_2(f|\bar{h})$ implies that the payoffs of $f|h$ and $f|\bar{h}$ must agree term by term along the play path. This, in turn, with the specific payoffs of the prisoners' dilemma game, implies that the utilities for player 1 must agree term by term along the play paths of $f|h$ and $f|\bar{h}$. It follows that $u_1(f|h) = u_1(f|\bar{h})$. So player 1 receives the same induced payoff after any two histories $h, \bar{h} \in H_1$. Thus, after histories $h \in H_0$ he must be maximizing the one-shot payoffs. Thus, he must be playing d along the equilibrium path. []

References

- Abreu, D. (1988), "On the Theory of Infinitely Repeated Games with Discounting," Econometrica, 56 (2), 383-396.
- Abreu, D. and A. Rubinstein (1989), "The Structure of Nash Equilibrium in Repeated Games with Finite Automata," Econometrica, 56 (6), 1259-1281.
- Aumann, R. J. and L. S. Shapley (1976), "Long Term Competition: A Game Theoretic Analysis," mimeo, Hebrew University.
- Aumann, R. J. (1981), "Survey of Repeated Games," in Essays in Game Theory and Mathematical Economics in Honor of Oskar Morgenstern, Wissenschaftsverlag, Bibliographisches Institut, Mannheim, Wien, Zurich 11-42.
- Aumann, R. J. (1988), "Irrationality in Game Theory," presented at the International Conference on Economic Theories of Politics, Haifa.
- Ben-Porath, E. (1986), "Repeated Games with Finite Automata," manuscript, Hebrew University.
- Ben-Porath, E. and A. Peleg (1987), "On the Folk Theorem and Finite Automata," manuscript, Hebrew University.
- Binmore, K. (1987), "Modeling Rational Players I," Economics and Philosophy, 3, 179-214.
- Fudenberg, D. and E. Maskin (1986), "The Folk Theorem in Repeated Games with Discounting or With Incomplete Information," Econometrica, 53 (3), 533-554.
- Fudenberg, D. and E. Maskin (1987), "Nash and Perfect Equilibria of Discounted Repeated Games," Discussion Paper No. 1301, Harvard University.

- Kalai, E., D. Samet and W. Stanford (1988), "A Note on Reactive Equilibria in the Discounted Prisoners' Dilemma and Associated Games," International Journal of Game Theory, 18 (3), 177-186.
- Kalai, E. and W. Stanford (1988), "Finite Rationality and Interpersonal Complexity in Repeated Games," Econometrica, 56 (2), 397-410.
- Kalai, E. (1987), "Bounded Rationality and Strategic Complexity in Repeated Games," forthcoming in the Proceedings of the International Workshop on Game Theory, Columbus, Ohio.
- Kohlberg, E. and J. F. Mertens (1986), "On the Strategic Stability of Equilibria," Econometrica, 54, 1003-1037.
- Kreps, D. and R. Wilson (1982), "Sequential Equilibria," Econometrica, 50, 863-894.
- Marschak, T and R. Selten (1978), "Restabilizing Resources, Inertia Supergames, and Oligopolistic Equilibria," Quarterly Journal of Economics, 42, 71-93.
- Neme, A. and L. Quintas (1988), "Equilibrium of Repeated Games with Cost of Implementation," Discussion Paper No. 763, Northwestern University.
- Neyman, A. (1986), "Bounded Complexity Justifies Cooperation in the Finitely Repeated Prisoners' Dilemma," Economic Letters, 19, 227-229.
- Rosenthal, R. (1981), "Games of Perfect Information, Predatory Pricing and the Chain-Store Paradox," Journal of Economic Theory, 25, 92-100.
- Rubinstein, A. (1979), "Equilibrium in Supergames with the Overtaking Criterion," Journal of Economic Theory, 21, 1-9.
- Rubinstein, A. (1986), "Finite Automata Play the Repeated Prisoner's Dilemma," Journal of Economic Theory, 39, 83-96.
- Selten, R. (1975), "Reexamination of the Perfectness Concept for Equilibrium

- Points in Extensive Games," International Journal of Game Theory, 4, 25-55.
- Selten, R. (1978), "The Chain Store Paradox," Theory and Decisions, 9, 127-159.
- Sorin, S. (1988), "Repeated Games with Bounded Rationality," forthcoming in Models of Incomplete Information and Bounded Rationality, Proceedings of NATO-ASI, Anacapri.
- Stanford, W. (1986), "Subgame Perfect Reaction Function Equilibria in Discounted Duopoly Supergames are Trivial," Journal of Economic Theory, 39, 226-232.