

# Learning to Learn, Pattern Recognition, and Nash Equilibrium\*

Doron Sonsino<sup>†</sup>

*Technion, Israel Institute of Technology, Haifa, Israel*

Received May 23, 1995

The paper studies a large class of bounded-rationality, probabilistic learning models on strategic-form games. The main assumption is that players “recognize” cyclic patterns in the observed history of play. The main result is convergence with probability one to a fixed pattern of pure strategy Nash equilibria, in a large class of “simple games” in which the pure equilibria are nicely spread along the lattice of the game. We also prove that a necessary condition for convergence of behavior to a mixed strategy Nash equilibrium is that the players consider arbitrarily long histories when forming their predictions. *Journal of Economic Literature* Classification Numbers: C72, D83. © 1997 Academic Press

## 1. INTRODUCTION

Modern economic theory makes extensive use of equilibrium concepts, Nash (1950) equilibrium in particular, when characterizing the solutions to models of strategic interaction. In a Nash equilibrium each player is assumed to correctly anticipate his opponents’ strategic behavior when playing the game. This strong assumption motivates the recent literature on “learning in games.”

Typically, the literature studies the case where a fixed game is played repeatedly by payoff maximizing players. In every stage of the repeated game, each player observes some information about his opponents’ realized play at that stage; he then tries to utilize this information in choosing an optimal action for the continuation game. The basic postulate of the

\* The paper is based on parts of my dissertation at Stanford University. I thank David Kreps for his advice, comments, and support. I thank Dimitri Vayanos for comments and suggestions. I have also benefited from conversations with Jeremy Bulow, Donald Brown, Peter Reiss, and John Roberts. I thank an anonymous referee for many important suggestions. I also thank the department of economics at the University of Arizona for its hospitality while I was working on this paper.

<sup>†</sup> E-mail: sonsino@ie.technion.ac.il

literature is that the players use the observations from the past to extrapolate the future. In this sense, the players try to “learn” their rivals’ future strategies from the realized path of play in the past. The different models studied in the literature vary in the specific assumptions made on the learning rules used by the agents,<sup>1</sup> the objective maximized by the agents, and the information revealed after each stage of the game.

The basic motivation underlying this paper is to suggest an intuitively appealing set of assumptions concerning the learning behavior of agents with bounded rationality and to study the dynamics of such a learning process. In particular, we assume that agents can choose their learning behavior independently throughout the repeated game, and that agents “learn to learn” in the sense of trying to choose the “best” learning model in every stage of the repeated interaction.

The specific framework we adopt is as follows: Agents with bounded rationality play repeatedly a fixed, finite strategic-form game. Each agent knows his payoff function, but doesn’t necessarily know the payoff functions of his rivals. After each stage-game, the agents observe the pure strategies played at that stage; each player examines the history of play up to the current stage, and tries to predict how the other agents will play at the next round. Formally, the realized prediction takes the standard form of a probability distribution over the pure strategy profiles of the player’s opponents. The agents are strictly myopic so that in every stage of the game they play a best response to their realized predictions.

Our main result is convergence with probability one to a pure strategy Nash equilibrium pattern, in a large class of games in which the pure equilibria are “nicely spread.” That is, when the topology of the stage game is such that potential pure strategy solutions (Kreps, 1990) exist, and are conveniently spread along the lattice of the game, then—with probability one—the players will eventually adopt one of these points or, more generally, a finite pattern of such equilibria as a convention on “how to play the game.”

In a repeated play of the “Battle of the Sexes” (Fig. 2.2), for example, the players may converge to a symmetric behavior pattern where one pure strategy Nash equilibrium is played on the even dates, and the other pure equilibrium is played on the odd dates. Our model is the first bounded rationality learning model showing that convergence to this intuitively appealing “taking turns” solution is theoretically possible. The experimen-

<sup>1</sup> All through the paper we use “players” and “agents” interchangeably to denote the set of individuals playing the repeated game. We use the female pronoun “she” when talking about the rows’ player (Rows) and the matrices’ player (Mats) in a given game; we use “he” when referring to a generic player or the columns’ player (Cols) in a given game.

tal literature indeed demonstrates that real subjects may rapidly learn to alternate when playing games with an equilibrium structure similar to that in the Battle of the Sexes (Rapoport *et al.*, 1976).<sup>2</sup>

We suggest two postulates concerning the agents' learning behavior:

*Postulate A.* Confused learning and random prediction rules. Our players can choose their learning behavior independently. In every stage of the game, each player has the freedom to use complicated formal methods, simple rules of thumb, or even plain intuition in deriving the next period's prediction. Each agent may use different prediction models at different dates. Different agents can use different prediction rules in every date. Typically, the agents feel confused when searching for the "right" prediction. The confusion is enhanced whenever a player realizes that his previous prediction was wrong and lead him to a suboptimal play. The player may then abort a prediction model that failed, and search for an alternative rule that seems to fit better the observed path of play.

Formally, we assume that for every history of the game, the prediction rule of each player can be described by a probability distribution over the space of possible predictions. Moreover, we stipulate that—because of the strategic confusion described above—the model is of "rich support" so that with some small probability  $\epsilon$  the players can behave in any reasonable way; e.g., follow the Cournot (1838) best response dynamics, in every stage of the repeated game.

As an interpretation, we suggest an underlying Bayesian model where each agent's learning-behavior-type is drawn independently in every stage of the game. Our probabilistic prediction rules might then be considered the direct product of this Bayesian model. The  $\epsilon$  compatibility with best response dynamics discussed above, for example, follows if we assume that there is always a positive probability  $\epsilon$  that all players will be of the Cournot-type. With this interpretation, the main convergence result shows that for almost every realization of types, the players will eventually learn to play an equilibrium pattern.

*Postulate B.* Learning to learn through strategic pattern recognition. Our players "learn to learn" by adjusting their learning behavior to the recent developments in the observed path of play. In particular, the players recognize cyclic strategic patterns. As a simple example, consider the case where three strategy profiles, say  $C, A, B$  have been played repeatedly, for a long time, in that specific order, so that the history of play at some stage is  $\dots C, A, B, C, A, B, C, A, B, C, A, B, C, A, B, C, A$ . We claim that the players

<sup>2</sup> In those experiments, however, the payoff matrix was common knowledge among the players. Since we don't need this strong assumption for our convergence result, the experimental data backs up our model in a restricted way.

must “recognize” the repeated pattern if it has been repeated successively with no interruptions a large enough number of times.<sup>3</sup> Upon recognizing the pattern, each player assumes that his opponents will keep following the pattern by playing their part in the profile  $B$  on the next stage of the game. Being strictly myopic, he plays a best response to  $B$  on the next round. When  $B$  is not a pure strategy Nash equilibrium, then—by recognizing the pattern—the agents contradict it, since realized play is the best response to  $B$  which is different from the pattern’s prediction. If, however,  $B$  is a pure strategy Nash equilibrium, realized play conforms to the pattern’s prediction, and expectations are being fulfilled. The players might thus expect that the same pattern will be followed again at the next stage of the game. This demonstrates the basic intuition behind our convergence to an equilibrium-pattern result.

We say “learning to learn through strategic pattern recognition” rather than “learning through strategic pattern recognition” since the players switch to (what they conceive to be) a better learning model whenever they recognize a pattern in the observed play. Moreover, our specific assumptions are flexible enough to accommodate dynamic, nonstationary strategic pattern recognition; e.g., if at some stage of the repeated game an “unfamiliar” pattern starts appearing successively in the observed play, then it may take the players a long time, say 60 successive repetitions, to recognize the pattern. If, after a short interruption, the same pattern starts repeating successively once again, then our agents may recognize this familiar pattern earlier, say after 8 successive repetitions.

Our pattern recognition model, however, deals only with the simplest case where the agents recognize cyclic strategic patterns that repeat successively at the observed path of play. Since, broadly interpreted, pattern recognition may encompass almost any form of inductive reasoning and learning, this is far from being a complete model. We hope that the paper will help initiate more research in this field.

Our two behavior postulates are new to the learning in games’ literature.<sup>4</sup> The existing literature can be roughly divided as follows: Some papers take a general approach to learning, study a large class of learning

<sup>3</sup> The computer sciences literature on pattern recognition typically deals with Bayesian-learning type of problems that are quite different from what we call *strategic* pattern recognition in this paper. The literature on machine learning however deals with similar issues when analyzing problems in “sequence extrapolation” (Laird, 1994; Laird and Ronald, 1994). Yet, that literature typically searches for deterministic algorithms that may be used for successful sequence predictions in general, while our goal here is to suggest an intuitively appealing descriptive model of strategic pattern recognition.

<sup>4</sup> Sanchirico (1993) has been working concurrently on a rational probabilistic learning model with an entropy assumption that is similar to our confused learning. Sanchirico’s agents don’t recognize strategic patterns, however, and in a repeated play of the Battle of the Sexes, they converge to play a pure Nash equilibrium of the game.

models on a large class of games, and accordingly prove general, weak results (e.g., Milgrom-Roberts, 1991). Other papers take a stylized approach, impose structure on the learning models, restrict the class of games under consideration, and then prove strong convergence results (e.g., Krishna, 1992). In this paper we prove a strong convergence result for a large class of probabilistic learning models on a large class of games. Furthermore, our basic behavioral postulates are intuitively appealing; we believe they would survive empirical testing.

Section 2 of the paper presents two introductory examples that motivate the model and demonstrate the main results. Section 3 provides the formal setting. Section 4 studies the implications of our first behavioral postulate. We prove that a confused search for “the right way to play a game” must always reach a pure strategy Nash equilibrium, when the stage game belongs to some large class of “simple games.” In Section 5, we superimpose a pattern recognition scheme on the confused learning model of Section 4. The resulting model complies with our two behavioral postulates. We prove that it converges with probability one to a Nash equilibrium pattern when the stage game is simple.

Section 6 suggests that convergence of behavior to a mixed strategy Nash equilibrium of the stage game is “incompatible with strategic pattern recognition.” This implies that a necessary condition for convergence of behavior to a mixed equilibrium is that the agents consider arbitrarily long histories when forming their predictions. In Section 7, we add one assumption to the model to get a general (not restricted to simple games), weaker convergence result. Section 8 discusses different properties of the model and considers possible generalizations. Section 9 is a short discussion of the main results.

## 2. EXAMPLES

### 2.1. *Probabilistic Learning and Reachability of Pure Equilibria*

The existing literature on bounded rationality learning in games typically imposes, at the outset, some arbitrary set of assumptions on the players' learning behavior. In many cases, it is even assumed that the players are born to use some fixed learning heuristic like fictitious play (Brown, 1951) or Cournot dynamics (Cournot, 1838) and they follow that given model zealously; no deviations can ever occur. In many applications, however, stationary fictitious play learning and stationary Cournot learning might lead the players to repeated cyclic behavior and score a very low grade in predicting their actual behavior in every stage of the game.

Consider first the fictitious play learning heuristic: In every stage of the game each agent plays a best response to the empirical distribution derived from the observed history of play. The play at the first round is chosen arbitrarily.<sup>5</sup>

The game in Fig. 2.1.1 has a unique pure strategy Nash equilibrium:  $(B, D, L)$ . It is a Pareto efficient equilibrium that pays each player his maximal payoff. Yet, it is easy to show that for most initial conditions fictitious play “converges” to the cycle  $(A, E, L), (A, E, R), (B, E, R), (B, F, R), (C, F, R), (C, F, L), (C, E, L), (A, E, L)$  in the sense that the agents play only the strategies in this cycle and change strategies according to the fixed order in the cycle (from  $(A, E, L)$  to  $(A, E, R)$ , from  $(A, E, R)$  to  $(B, E, R)$ , and so on ...). Moreover, the number of successive repetitions of each strategy in the cycle, in the  $n$ th occurrence of the cycle, increases rapidly in  $n$ . At some stage of the repeated interaction, the observed history thus shows that the strategy profile  $(A, E, L)$  has been played for close to 100,000 periods successively, the strategy profile  $(A, E, R)$  has followed for more than 120,000 rounds successively, and so on ... . An outside observer facing these dynamics can do an asymptotically perfect job in predicting the players’ behavior in every stage of the game by using the alternative Cournot learning heuristic; i.e., by assuming that whatever has been played at the last stage will be played again at the current stage. The fact that the players keep following the fixed fictitious play learning rule although it does such a poor job in predicting their actual behavior seems unreasonable and provides a convincing argument against stationary fictitious play learning in this example.

Is stationary Cournot learning more plausible? Clearly it is not. In the game of Figure 2.1.1, for most initial conditions, stationary Cournot learning converges to the cycle  $(A, E, L), (A, E, R), (B, E, R), (B, F, R), (C, F, R), (C, F, L), (C, E, L), (A, E, L)$ . The players will repeatedly cycle, playing  $(A, E, L)$  for one period, switching to  $(A, E, R)$  on the next period, moving to  $(B, E, R)$  at the following stage, etc. . . . Again we conclude that stationary learning is unreasonable in this example.

Since fictitious play and Cournot dynamics seem, a priori, reasonable learning heuristics, their potential instability demonstrates that the bounded rationality learning problem doesn’t obtain any ex ante clear solutions. We therefore suggest that (at the advanced stages of a repeated game) the players will feel basically confused when searching for the right way to play the game. This motivates our probabilistic approach to learning and the first behavioral postulate. We further suggest that a probabilistic, confused learning model should always reach a pure strategy Nash equilibrium in

<sup>5</sup> For slightly different definitions of fictitious play learning see Fudenberg and Kreps (1993) and Monderer and Shapley (1993).

L	D	E	F	R	D	E	F
A	0,0,0	2,2,1	0,0,0	A	0,0,0	1,2,2	0,0,0
B	2,2,2	0,0,0	0,0,0	B	0,0,0	2,1,2	1,2,2
C	0,0,0	1,2,2	2,1,2	C	0,0,0	0,0,0	2,2,1

FIGURE 2.1.1

games similar to that in Fig. 2.1.1. Consider, in particular, a perturbed Cournot learning model where the players follow the Cournot learning model with high probability in every stage of the game, but each player  $i$  may, with some small probability  $\epsilon$ , tremble by recognizing strategic patterns of length 2 in the observed past play; i.e., if the last strategy played at some stage of the game is  $A$  and somewhere in the observed past play  $A$  has been succeeded by  $B$ , then—with probability  $\epsilon$ — $i$  might play a best response to strategy  $B$  on the next stage of the game. Assume further that different players can recognize different patterns in a given history of the game; e.g., if the observed history ends with the string  $A, B, C, D, A, C, A$ , then one player might play a Cournot best response to  $A$  on the next stage of the game, another player might recognize the pattern  $A, B$  and play a best response to  $B$  at that stage, while a third player may recognize the pattern  $A, C$  and plays a best response to  $C$  on the next stage of the game.

The table in Fig. 2.1.2 shows that the perturbed model might lead the players from the Cournot cycle to the pure equilibrium  $(B, D, L)$ . To understand the example it is enough to focus on those dates where the players play the strategy  $(C, E, L)$ .

\* At the first such date  $t = 1$ ,  $(C, E, L)$  is followed by the best response  $(A, E, L)$ .

\* At the second such date  $t = 8$ , all the players recognize the pattern  $(C, E, L), (A, E, L)$  and play the best response to  $(A, E, L), (A, E, R)$  at  $t = 9$ .

\* At the third appearance of  $(C, E, L)$  (at  $t = 12$ ), Rows recognize the pattern  $(C, E, L), (A, E, R)$  and thus decides to play her best response to  $(A, E, R)$ , Row  $B$ , at the next stage. Assuming that Cols and Mats play a Cournot best response to  $(C, E, L)$  at that stage, realized play is  $(B, E, L)$ .

$t$	Strategy	Explanation
1	$(C, E, L)$	Starting Point
2	$(A, E, L)$	Cournot BR
3	$(A, E, R)$	Cournot BR
4	$(B, E, R)$	Cournot BR
5	$(B, F, R)$	Cournot BR
6	$(C, F, R)$	Cournot BR
7	$(C, F, L)$	Cournot BR
8	$(C, E, L)$	Cournot BR
9	$(A, E, R)$	All players recognize $(C, E, L), (A, E, L)$
10	$(B, F, R)$	All players recognize $(A, E, R), (B, E, R)$
11	$(C, F, L)$	All players recognize $(B, F, R), (C, F, R)$
12	$(C, E, L)$	Cournot BR
13	$(B, E, L)$	Rows Recognizes $(C, E, L), (A, E, R)$ Cols and Mats play Cournot BR
14	$(A, D, R)$	Cournot BR
15	$(A, E, R)$	Cournot BR
16	$(C, F, R)$	All players recognize $(A, E, R), (B, F, R)$
17	$(C, E, L)$	All players recognize $(C, F, R), (C, F, L)$
18	$(B, D, L)$	Rows recognizes $(C, E, L), (A, E, R)$ Cols recognizes $(C, E, L), (B, E, L)$ Mats plays Cournot BR

FIGURE 2.1.2



	Col A	Col B
Row A	1,2	0,0
Row B	0,0	2,1

FIG. 2.2. Battle of the sexes.

\* At the fourth appearance of  $(C, E, L)$  (at  $t = 17$ ), Rows recognizes the pattern  $(C, E, L), (A, E, R)$  and chooses Row  $B$  for the next stage of the game, Cols recognizes the pattern  $(C, E, L), (B, E, L)$  and accordingly chooses Col  $D$ . Assuming further that Mats plays a Cournot best response to  $(C, E, L)$ , we get the Nash equilibrium  $(B, L, D)$  at  $t = 18$ .

In Section 4 we generalize this example and prove that a probabilistic learning model that always assigns a positive probability to the perturbed Cournot behavior described above must reach a pure equilibrium in some large class of games.

## 2.2. Convergence to Equilibrium Patterns

The well known “Battle of the Sexes” of Fig. 2.2 has two symmetric pure strategy Nash equilibria:  $(A, A)$  is Cols’ favorite equilibrium while  $(B, B)$  is Rows’ favorite. The repeated play of the game has a natural, intuitively appealing solution where one equilibrium, say  $(A, A)$ , is played on the odd dates and the other equilibrium,  $(B, B)$ , is played on the even dates. Yet, it is easy to prove that if fictitious play or Cournot learning ever hit a strict pure strategy Nash equilibrium, then the players keep playing that equilibrium forever.<sup>6</sup> These learning heuristics can thus never converge to correlated solutions like that suggested above.

Consider, however, a probabilistic learning model where the players may recognize patterns of length 2 as illustrated in Example 2.1. Assume further that because of the players’ confusion about the strategic behavior of their opponents “anything” can happen (with positive probability) at the first three stages of the repeated game. In particular, there is a positive probability that realized play at these three rounds will be  $(A, A), (B, B), (A, A)$ . Consider the case where this, in fact, is the observed path of play at the first three rounds. Assuming that both players happen to recognize the pattern  $(A, A), (B, B)$  in that history, and thus play a best response to the pattern’s prediction  $(B, B)$  at the fourth stage of the game, we get the extended history:  $(A, A), (B, B), (A, A), (B, B)$ . Repeating that argument

<sup>6</sup> For a proof that fictitious play satisfies this property see Fudenberg and Kreps (1993).

iteratively we may get an arbitrarily long history where the players repeatedly play the *pure equilibrium pattern*  $(A, A), (B, B)$  time after time with no interruptions. Assuming further that the probability of recognizing the pattern  $(A, A), (B, B)$  converges to one as the number of previous successive repetitions increases, we get a scenario where the players converge to the equilibrium pattern  $(A, A), (B, B)$ . The main result of this paper is that a confused learning model with strategic pattern recognition must converge to an equilibrium-pattern when the stage game belongs to a large class of “simple” games.

### 3. THE MODEL: NOTATION AND DEFINITIONS

#### 3.1. *The Stage Game*

Let  $\Gamma = (I, \langle S^i \rangle_{i \in I}, \langle \pi^i \rangle_{i \in I})$  be a finite strategic-form game, where  $I$  is a finite set of players,  $S^i$  is a finite set of pure strategies for player  $i$ , and  $\pi^i: \times_{i \in I} S^i \mapsto \mathfrak{R}$  is the payoff function for agent  $i$ .  $S = \times_{i \in I} S^i$  is used to denote the space of pure strategy profiles in the game.  $S^{-i} = \times_{j \in I \setminus i} S^j$  denotes the pure strategy profiles of  $i$ 's opponents.  $s^i \in S^i$ ,  $s \in S$ , and  $s^{-i} \in S^{-i}$  are used to denote an arbitrary element of the corresponding space.

We use  $\Delta(X)$  to denote the space of finite support probability distributions over a set  $X$ . Thus,  $\Sigma^i = \Delta(S^i)$  denotes the space of mixed strategies for agent  $i$ ,  $\Sigma = \times_{i \in I} \Sigma^i$  denotes the space of mixed strategy profiles in the game, and  $\Delta(S^{-i})$  denotes the space of probability distributions over  $i$ 's opponents' pure strategies. We sometimes call  $\Delta(S^{-i})$  player  $i$ 's predictions' space.  $\sigma^i \in \Sigma^i$ ,  $\sigma \in \Sigma$ , and  $y \in \Delta(S^{-i})$  are defined accordingly. Expressions of the form  $\sigma^i(s^i)$  are used to denote the probability assigned to the pure strategy  $s^i$  by the mixed strategy  $\sigma^i$ .

We extend the payoff functions in the standard way to the space  $\Delta(S)$ ;  $\pi^i(\sigma^i, y)$  then denotes the expected payoff to player  $i$  from playing a mixed strategy  $\sigma^i$  when his opponents play the (possibly correlated) strategy  $y$ . We use  $BR^i$  to denote the best response correspondence of player  $i$  in pure strategies; i.e., for every  $y \in \Delta(S^{-i})$ ,  $x \in BR^i(y)$  iff  $x \in \arg \max_{s \in S^i} \pi^i(s, y)$ ;  $BR^i(s^{-i})$  is defined accordingly. To simplify notation we sometimes use  $BR^i(s)$  to denote  $BR^i(s^{-i})$ ;  $BR(s)$  is used to denote the joint best response correspondence of the players; i.e.,  $\tilde{s} \in BR(s)$  iff  $\tilde{s}^i \in BR^i(s) \forall i \in I$ . We say that a sequence  $s_0, s_1, \dots, s_n$ , is a *BR sequence* iff  $s_j \in BR(s_{j-1})$  for every  $j = 1, \dots, n$ .

A strategy profile  $s$  is a pure strategy Nash equilibrium iff  $s \in BR(s)$ ; it is a strict equilibrium iff  $s = BR(s)$ . We use  $PN(\Gamma)$  to denote the set of

pure strategy Nash equilibria of  $\Gamma$ ;  $S \setminus \text{PN}(\Gamma)$  is used to denote the complementary set. We say that  $\Gamma$  is *generic* iff for every  $s \in S(\Gamma)$ ,  $\text{BR}(s)$  is a singleton.

### 3.2. The Learning Process

Throughout this paper we study the case where a generic, finite strategic-form game is played repeatedly at dates  $t = 1, 2, \dots$ . After each stage of the game, every player observes the pure strategy profile played at that stage. We use  $z_t$  to denote the history of the play up to the beginning of date  $t$ ; i.e.,  $z_t = s_1, s_2, \dots, s_{t-1}$ , where  $s_{t'} \in S$  for  $t' = 1, 2, \dots, t - 1$ .  $\mathcal{Z}_t$  is used to denote the space of histories of length  $t - 1$ ;  $\mathcal{Z}_\infty$  denotes the space of histories in the infinitely repeated game.

Let  $\mu_t^i: \mathcal{Z}_t \mapsto \Delta(\Delta(S^{-i}))$  be a *probabilistic prediction rule* for agent  $i$  at date  $t$ . That is, facing a given history  $z_t$ , agent  $i$  may choose different predictions for his opponents' behavior at the next round. We assume w.l.g. that the probabilistic rule of his choice is given by a finite support probability distribution over his predictions' space  $\Delta(S^{-i})$ :  $\mu_t^i(z_t) \in \Delta(\Delta(S^{-i}))$  (The finite support assumption is with no loss of generality since we are analyzing games with finite pure strategy spaces so that for every "well-defined" probability distribution  $\tilde{\mu}_t^i$  over  $\Delta(S^{-i})$  there exists a finite support probability distribution  $\mu_t^i$  that generates the same probabilistic behavior rule for agent  $i$  as  $\tilde{\mu}_t^i$ ). For every  $y \in \Delta(S^{-i})$  we use  $\mu_t^i(z_t)(y)$  to denote the probability assigned to the prediction  $y$  by the probabilistic prediction rule  $\mu_t^i(z_t)$ ;  $\text{supp}(\mu_t^i(z_t))$  is used to denote the (finite) support of  $\mu_t^i(z_t)$ ; i.e.,  $y \in \text{supp}(\mu_t^i(z_t))$  iff  $\mu_t^i(z_t)(y) > 0$ .

For any finite set  $X$  we use  $\text{card}(X)$  to denote the cardinality of  $X$ . Following the myopia assumption we assume that for every history  $z_t$ , for any realization of the random prediction rule  $\mu_t^i(z_t)$ , agent  $i$  plays a best response to his realized prediction.  $\phi_t^i(z_t)$  is used to denote the *probabilistic behavior rule* derived from the probabilistic prediction rule  $\mu_t^i(z_t)$  given the myopia assumption; i.e., assuming that  $\text{card}(\text{BR}^i(y)) = 1$  for every  $y \in \text{supp}(\mu_t^i(z_t))$ ,  $\phi_t^i: \mathcal{Z}_t \mapsto \Delta(S^i)$  is defined by  $\phi_t^i(z_t)(x) = \sum_{\{y \in \text{supp}(\mu_t^i(z_t)) \mid x \in \text{BR}^i(y)\}} \mu_t^i(z_t)(y)$ .<sup>7</sup>

We call the collection  $(\mu, \phi) = \langle \{\mu_t^i\}, \{\phi_t^i\}_{i \in I} \rangle$  a (*myopic*) *learning model* on  $\Gamma$ . Assuming further that the agents behave independently, we get that for each  $t$ , the collection  $\langle \{\phi_{t'}^i\}_{t'=1}^{t-1} \rangle_{i \in I}$  determines a distribution on  $\mathcal{Z}_t$ . By a standard application of Kolmogorov's extension theorem we can extend the process to  $\mathcal{Z}_\infty$ ;  $(\mathcal{Z}_\infty, \mathcal{F}, \mathcal{P}_{(\mu, \phi)})$  is used to denote the extended space.

<sup>7</sup> When  $\text{card}(\text{BR}^i(y)) > 1$  for some  $y$  in the support of  $\mu_t^i(z_t)$ , the corresponding myopic behavior rule also depends on the tie-breaking rule. Using  $q_{z_t}^i(\cdot \mid y): \text{BR}^i(y) \mapsto [0, 1]$  to denote a tie breaking rule for agent  $i$  (so that  $\sum_{\{x' \in \text{BR}^i(y)\}} q_{z_t}^i(x' \mid y) = 1$ ), we define  $\phi_t^i(z_t)(x) = \sum_{\{y \in \text{supp}(\mu_t^i(z_t)) \mid x \in \text{BR}^i(y)\}} \mu_t^i(z_t)(y) \cdot q_{z_t}^i(x \mid y)$ .

### 3.3. Strategic Patterns

We say that  $p$  is a *strategic pattern* in a game  $\Gamma$  iff  $p$  is a finite sequence of strategies in  $\Gamma$ . We use  $l(p)$  to denote the length of a strategic pattern  $p$ . That is, if  $p = s_1, s_2, \dots, s_n$ , then  $l(p) = n$ . When all the elements of  $p$  are in  $\text{PN}(\Gamma)$  we say that  $p$  is a *PN pattern*. When  $l(p) = 1$  we call  $p$  a *singleton pattern*.<sup>8</sup>

Since we will be dealing a lot with appearances of strategic patterns in given histories, we define special projection functions,  $\text{PRO}$ , that will be used to denote subsequences of histories and patterns:

Let  $\zeta = s_1, s_2, \dots, s_t$  be a sequence of length  $t$  of pure strategies in  $\Gamma$ .

- (1) For every  $0 \leq t' < t$ , let  $\text{PRO}_{t'}(\zeta) = s_{t'}$ .
- (2) For every  $0 \leq t_1 < t_2 \leq t$ , let  $\text{PRO}_{t_1, t_2}(\zeta) = s_{t_1}, s_{t_1+1}, \dots, s_{t_2}$ .

Similarly, we define an extension operator,  $\text{EXT}$ , so that for every  $t' > t$ ,  $\text{EXT}_{1, t'}(p)$  is used to denote the sequence generated by copying the pattern  $p$  repeatedly until we get a sequence of length  $t'$ ; i.e.,

$$\text{EXT}_{1, t'}(\zeta) = \overbrace{\zeta, \zeta, \dots, \zeta}^{t' \text{ div } t \text{ times}}, s_1, \dots, s_{t' \bmod t}$$

When  $X$  is a collection of finite sequences in  $\Gamma$  we define  $\text{EXT}_{1, t'}(X) = \{x' \mid x' = \text{EXT}_{1, t'}(x) \ x \in X\}$ .

We say that a pattern  $p$  has appeared in a given history  $z_t$  if the obvious definition holds; i.e.,  $t - 1 \geq l(p)$  and  $\exists t' \in \{l(p), \dots, t - 1\}$  such that  $\text{PRO}_{t'-l(p)+1, t'}(z_t) = p$ .

## 4. CONFUSED LEARNING LEADING TO PURE EQUILIBRIA

In this section we study the implications of the first behavioral postulate. We show that a confused search for “the right way to play the game” must always reach a pure strategy Nash equilibrium when the stage game is what we define to be “simple.” The basic intuition underlying this reachability result is that random prediction rules that always assign a positive probability to the Cournot prediction on one hand, and tremble independently across players on the other hand, must reach a pure strategy Nash equilibrium in many applications. The following stylized example is used to motivate our definition of “simple games.”

<sup>8</sup> The definition above considers only the case where all the players are involved in playing each strategy along the pattern (i.e.,  $s_l \in S$  for every  $l$ ). In the sequel we call this type of patterns *patterns of full dimensionality*. In Section 8.2 we briefly discuss *patterns of partial dimensionality* that may involve a strict subset of the set of players.

	Col E	Col F	Col G	Col H
Row A	$-10^6, -10^6$	$-10^6, -10^6$	2, 0	0, 2
Row B	$-10^6, -10^6$	$-10^6, -10^6$	0, 2	2, 0
Row C	2, 1	0, 0	$-10^6, -10^6$	$-10^6, -10^6$
Row D	0, 0	1, 2	$-10^6, -10^6$	$-10^6, -10^6$

FIG. 4.1. A nonsimple game.

The game in Fig. 4.1 is a composition of “Matching Pennies” (the subgame  $\{A, B\} \times \{G, H\}$ ), and “Battle of the Sexes” (the subgame  $\{C, D\} \times \{E, F\}$ ). If the players fail to coordinate on one of these subgames, each of them pays a heavy fine.

Note that the Matching Pennies subgame is best-response-closed in the sense that the best response to every strategy in this subgame is, again, in that subgame. We suggest that myopic players that happened to play some strategy in this closed subgame at the first stage of the repeated game might (with a positive probability) keep playing the strategies in that “closed subgame” forever, and never reach the pure strategy Nash equilibria of the complete game. Our reachability result will thus be restricted to games that don’t contain such closed subgames. We proceed with the formal definitions.

**DEFINITION.** Uniform reachability of PN equilibria. Let  $\Gamma$  be a finite strategic-form game with  $\text{PN}(\Gamma) \neq \emptyset$ . A learning model  $(\mu, \phi)$  on  $\Gamma$  satisfies uniform reachability of PN equilibria iff  $\exists \epsilon > 0$ , such that for every  $t$ , and for each  $z_t \in \mathcal{Z}_t$  satisfying  $\mathcal{P}_{(\mu, \phi)}(Z_t = z_t) > 0$ ,  $\mathcal{P}_{(\mu, \phi)}(s_n \in \text{PN}(\Gamma) \text{ for some } n > t | z_t) \geq \epsilon$ .

**DEFINITION.** Strategic-form subgame. A strategic-form subgame  $\Gamma^*$  of the strategic-form game  $\Gamma = (I, \langle S^i \rangle_{i \in I}, \langle \pi^i \rangle_{i \in I})$  is a strategic-form game  $\Gamma^* = (I, \hat{S}^1, \dots, \hat{S}^n, \hat{\pi}^1, \dots, \hat{\pi}^n)$  such that for every  $i \in I$ ,  $\hat{S}^i \subseteq S^i$ , and  $\hat{\pi}^i$  is the restriction of  $\pi^i$  to  $\times_{i \in I} \hat{S}^i$ .

We use  $\Gamma^* \subseteq \Gamma$  to denote a strategic-form subgame of  $\Gamma$ .  $S^*(\Gamma^*)$  is used to denote the pure strategy profiles in  $\Gamma^*$ .

**DEFINITION.** BR-closed subgame. Let  $\Gamma$  be a finite strategic-form game. A strategic-form subgame  $\Gamma^* \subseteq \Gamma$  is BR-closed iff for every  $s \in S^*(\Gamma^*)$ , for every  $s' \in \text{BR}(s)$  (in  $\Gamma$ ),  $s' \in S^*(\Gamma^*)$ .

**DEFINITION.** *Simple strategic-form game.* A strategic form game  $\Gamma$  is BR-simple iff every BR-closed subgame  $\Gamma^* \subseteq \Gamma$  contains a pure strategy Nash equilibrium of  $\Gamma$ .

Consider now a probabilistic learning model on a simple strategic-form game that satisfies the following condition:

There is a fixed positive probability  $\epsilon$  such that for each player  $i$ :

(a) If the last observation at some (advanced enough) stage of the repeated game is  $s$ , then the probability that  $i$  will play a best response to  $s$  in the next stage of the game is at least  $\epsilon$  (we call this condition “ $\epsilon$  compatibility with BR dynamics”), and

(b) If the last observation at some (advanced enough) stage of the repeated game is  $s$ , and somewhere in the observed past play  $s$  has been followed by  $s'$ , then the probability that  $i$  will “recognize the pattern  $s, s'$ ,” in the sense that he will expect his rivals to play  $s'$  once again in the next stage of the game (after playing  $s$  at the current stage), is at least  $\epsilon$  (we call this condition “ $\epsilon$  compatibility with sophisticated BR dynamics”).

The formal definitions that follow only spell out the conditions on the prediction rules  $\mu_i^i(\cdot)$ . Since we are analyzing a myopic learning model on a generic game, the condition  $\mu_i^i(s^{-i}) \geq \epsilon$  immediately implies that the corresponding behavior rule satisfies  $\phi_i^i(\text{BR}^i(s^{-i})) \geq \epsilon$ . We thus omit the latter condition.

**DEFINITION.**  $\epsilon$  compatibility with BR dynamics. A learning model  $(\mu, \phi)$  on a strategic-form game  $\Gamma$  is  $\epsilon$  compatible with BR dynamics iff  $\exists \bar{T}$  and  $\exists \epsilon > 0$  such that  $\forall t \geq \bar{T}, \forall z_t \in \mathcal{Z}_t, \mu_i^i(z_t)(\text{PRO}_{t-1}^{-i}(z_t)) \geq \epsilon$ , for every  $i \in I$ .

**DEFINITION.**  $\epsilon$  compatibility with sophisticated BR dynamics. A learning model  $(\mu, \phi)$  on a strategic-form game  $\Gamma$  is  $\epsilon$  compatible with sophisticated BR dynamics iff  $\exists \bar{T}$  and  $\exists \epsilon > 0$  such that  $\forall t \geq \bar{T}, \forall z_t \in \mathcal{Z}_t$ , and  $\forall l \leq t - 2$  for which  $\text{PRO}_l(z_t) = \text{PRO}_{l-1}(z_t)$ ,  $\mu_i^i(z_t)(\text{PRO}_{l+1}^{-i}(z_t)) > \epsilon$ , for every  $i \in I$ .

In the sequel we say that  $(\mu, \phi)$  is a *confused learning model* on  $\Gamma$  iff  $(\mu, \phi)$  satisfies  $\epsilon$  compatibility with BR dynamics and  $\epsilon$  compatibility with sophisticated BR dynamics. The following proposition asserts that confused learning implies uniform reachability of pure strategy Nash equilibria when the stage game is simple.

**PROPOSITION 4.** *Let  $\Gamma$  be a generic, simple, strategic-form game. Let  $(\mu, \phi)$  be a confused learning model on  $\Gamma$ . Then  $(\mu, \phi)$  satisfies uniform reachability of PN equilibria.*

The proof is left (together with all other proofs) for Appendix 1.

## 5. CONFUSED LEARNING WITH PATTERN RECOGNITION

### 5.1. Introduction

In this section we superimpose a pattern recognition scheme on the confused learning model of Section 4. Note that confused learning is incompatible with strategic pattern recognition; compatibility with BR dynamics, for example, implies that for every BR cycle in the game, for every natural  $n$ , there is a positive probability that the players repeat the cycle  $n$  times successively. Our model will distinguish between two distinct learning modes: *A confused learning mode* where the players find the observed history hard to interpret and many learning behaviors can occur, and a *pattern recognition mode* where the players recognize some strategic pattern in the observed history of play.

The basic idea of the model can be summarized schematically as follows: For large enough  $t$ , for every history  $z_t$ , the model is in exactly one of the following three states.

**State 1:** *Confused learning (as described in Section 4). For every history in this state, many reasonable learning behaviors can occur. The confused search for the “right way to play the game” leads the players to a PN equilibrium (as proved in Proposition 4) unless they are stopped on the way because they recognize some strategic pattern in the extended history of play (State 2).*

**State 2:** *Pattern recognition. If some strategic pattern repeats successively at the end of the observed history  $z_t$ , then it might be the case that learning is not as confused as in State 1. In particular, if the pattern has appeared successively a large number of times, then the agents recognize it with high probability. If the pattern is not a PN pattern, then upon recognizing the pattern, the agents move back to the confused State 1 since expectations are being contradicted. If the pattern is a PN pattern, the agents may eventually move to State 3.*

**State 3:** *Convergence to a PN pattern. History shows that a fixed PN pattern has been played successively without interruptions for a long time. The players use the fixed pattern to form their predictions for the next period’s play. Since the pattern contains only PN strategies, expectations are realized and the players keep playing the pattern repeatedly. (Note that convergence to a PN equilibrium is a special case where the agents converge to a singleton PN pattern.)*

In formally defining a “confused learning model with strategic pattern recognition” we will often refer to the number of times a fixed pattern has appeared successively in a given history of the game. Thus, we would not want to call a sequence like  $A, B, C, A, B, C, A, B, C$  a strategic pattern,

but rather we would consider that sequence as three successive repetitions of the *basic pattern*  $A, B, C$ . Formally, we define recursively the collection of basic strategic patterns of length  $n$  in a fixed game  $\Gamma$ ,  $\text{SP}_n$ , as follows:

Let  $\text{SP}_1 = \{s \mid s \in S(\Gamma)\}$ , and let  $\text{SP}_n = \{s_1, \dots, s_n \mid s_i \in S(\Gamma) \text{ for every } i = 1, \dots, n\} \setminus \{\cup_{(j|n \text{ MOD } j=0)} \text{EXT}_{1,n}(\text{SP}_j)\}$  for every  $n = 2, 3, \dots$ .

We then say that  $p$  is a *basic pattern* in  $\Gamma$  iff  $p$  is an element of  $\cup_{n \in \mathcal{N}} \text{SP}_n$ . We use  $\mathcal{SP}$  to denote the collection of basic patterns in  $\Gamma$  plus a (unique) empty pattern  $\theta$ ; i.e.,  $\mathcal{SP} = \cup_{n \in \mathcal{N}} \text{SP}_n \cup \theta$ . We say that a pattern  $p' = s'_1, \dots, s'_n$  (of length  $n \geq 2$ ) is a  $m$  periods shift of a pattern  $p = s_1, \dots, s_n$  if the usual definition holds; i.e., if  $m \in \{1, \dots, n - 1\}$  and  $s'_1, \dots, s'_n = s_{m+1}, \dots, s_n, s_1, \dots, s_m$ .  $\text{SH}_1(p)$  is used to denote the 1 period shift of  $p$ . When  $p$  is a singleton pattern we let  $\text{SH}_1(p) = p$ . To simplify notation we use  $\text{SH}_{-1}(p)$  to denote the pattern  $\text{SH}_{l(p)-1}(p)$ ; i.e., when  $p = s_1, \dots, s_n$ ,  $\text{SH}_{-1}(p) = s_n, s_1, \dots, s_{n-1}$ .

### 5.2. The Convergence Result

To formally define a confused learning model with strategic pattern recognition on a strategic-form game  $\Gamma$ , we superimpose on the original model described in Section 3, a function  $\psi: \cup_{t=1}^{\infty} \mathcal{Z}_t \mapsto \mathcal{SP}$  that is used to describe the learning mode of the players when observing the history  $z_t$ :

—  $\psi(z_t) = \theta$  describes the case where the players find the history  $z_t$  hard to interpret and the confused learning assumptions of Section 4 hold.

—  $\psi(z_t) = p \neq \theta$  describes the case where the players recognize the pattern  $p$  in  $z_t$ .

In the sequel we call  $\psi$  a *pattern recognition frame* on  $\Gamma$ . We proceed by imposing specific assumptions on the collection  $(\mu, \phi, \psi)$ .

The first two assumptions A.1 and A.2 are very similar to the “confused learning” assumptions of Section 4. The only difference is that now we restrict these assumptions to those cases where the agents are in the confused learning mode.

*Assumption A.1. Conditional compatibility with BR dynamics.*  $\exists \bar{T}, \exists \epsilon > 0$  such that  $\forall t \geq \bar{T}, \forall z_t \in \mathcal{Z}_t$  for which  $\psi(z_t) = \theta, \mu_i^i(z_t)(\text{PRO}_{t-1}^{-i}(z_t)) \geq \epsilon$ , for every  $i \in I$ .

*Assumption A.2. Conditional compatibility with sophisticated BR dynamics.*  $\exists \bar{T}$  and  $\exists \epsilon > 0$  such that  $\forall t \geq \bar{T}, \forall z_t \in \mathcal{Z}_t$  for which  $\psi(z_t) = \theta$ , and for every  $l \leq t - 2$  for which  $\text{PRO}_l(z_t) = \text{PRO}_{t-1}(z_t), \mu_i^i(z_t)(\text{PRO}_{l+1}^{-i}(z_t)) > \epsilon$ , for every  $i \in I$ .



Note that Assumptions A.1 and A.2 impose some independence-across-periods on the random prediction rules; e.g., even if the player has failed to comply with BR dynamics over a long period, there is still an  $\epsilon$  probability that she will comply with this behavior at the next stage of the game. Note further that these  $\epsilon$ -probability assumptions are the only restrictions imposed on the players' behavior when facing the confused-mode histories. Thus, with some probability  $1 - \epsilon'$ , the agents can behave in any arbitrary way (e.g., follow fictitious play) when observing such histories.

We proceed by complementing these confused learning assumptions with a set of assumptions on strategic pattern recognition. Basically, we would want to stipulate that (\*) for every strategic pattern  $p$  there is a uniform bound, say  $T_p$ , such that the players must recognize  $p$  (with probability one) whenever it has appeared almost  $T_p$  times successively at the end of the observed history. For example, if the bound for the pattern  $A, B, T_{A, B}$ , is 4, then whenever  $z_t$  ends with the string  $A, B, A, B, A, B, A$ , the agents recognize the pattern  $A, B$ , and accordingly play a best response to strategy  $B$  on the next stage of the game.

Yet, it is easy to show that any such set of assumptions is internally inconsistent in the sense that for some histories the players will be required to recognize two (or more) contradicting patterns with probability one. For example, if the best response to strategy  $B$  is  $C$ , the bound for the pattern  $A, B$  is 4, and the bound for the pattern  $A, B, A, B, A, B, A, C$  is 3, then if the observed history ends with the string  $A, B, A, B, A, B, A, C, A, B, A, B, A, B, A, B, A, C, A, B, A, B, A, B, A$  the players must simultaneously recognize both patterns with probability one. That is, each player must play a best response to  $B$  with probability one and a best response to  $C$  with probability one at the same time—a possible inconsistency.

To resolve the problem we assume that our bounded rationality players can fully recognize only patterns of length shorter than some fixed bound  $L$ .<sup>9</sup> Assuming then that the bounds  $T_p$  are large enough relative to  $L$  we may reformulate (\*) to get an internally consistent set of assumptions.

*Assumption B.1. Boundedness.* There is an upper bound  $L \geq 1$  such that for every history  $z_t$  for which  $\psi(z_t) = p$  for some basic pattern  $p$ ,  $l(p) \leq L$ .

*Assumption B.2. Compatibility with pattern recognition.* For every basic pattern  $p$  with  $l(p) \leq L$  (where  $L$  is as defined in Assumption B.1), there is a uniform bound  $T_p$  satisfying  $T_p \cdot l(p) \geq 3L$  such that for every

<sup>9</sup> Since we only impose  $\epsilon$ -probability assumptions on the players' behavior in the confused learning mode, the players may still *partially recognize* patterns of length  $> L$  with some high probability  $1 - \epsilon'$  (where  $\epsilon'$  is a function of  $\epsilon$  and the structure of the specific pattern).

history  $z_t$ , if  $\text{PRO}_{t-T_p \cdot l(p)+1, t-1}(z_t) = \text{EXT}_{1, T_p \cdot l(p)-1}(p)$ , then (1)  $\psi(z_t) = p$ , and (2)  $\mu^i(z_t)(\text{PRO}_{l(p)}^{-i}(p)) = 1$ , for every  $i \in I$ .<sup>10</sup> Moreover, if  $p'$  is a shift of  $p$ , then  $T_{p'} = T_p$ .

Assumption B.2 implies that  $T_p \cdot l(p) \geq 2L$  for every pattern  $p$ . This guarantees that the conditions stated in the assumption can only apply to one pattern at a time;<sup>11</sup> i.e., assume (by way of contradiction) that there is a history  $z_t$ , and two distinct basic patterns of length  $\leq L$ ,  $p_1$  and  $p_2$ , such that the conditions of Assumption B.2 are satisfied with respect to both patterns. Assume without loss of generality that  $l(p_1) > l(p_2)$ . Note that if  $l(p_2) = 1$ , then since  $T_{p_2} \cdot l(p_2) \geq 2L > l(p_1)$ ,  $p_1$  consists of  $l(p_2)$  repetitions of  $p_2$  and thus is not a basic pattern. Assume  $l(p_2) > 1$ . Observe (\*\*\*) that since both patterns are recognizable given  $z_t$ , the last  $l(p_2)$  elements of the pattern  $\text{SH}_{-1}(p_1)$  must be equal to the pattern  $\text{SH}_{-1}(p_2)$ . Note also that since  $p_1$  and  $p_2$  are basic patterns  $l(p_1) \bmod l(p_2) \neq 0$  (for if this isn't the case,  $p_1$  must consist of  $l(p_1) \text{div } l(p_2)$  successive repetitions of  $p_2$ , and it is not a basic pattern). Thus, since  $T_{p_2} \cdot l(p_2) \geq 2L \geq 2l(p_1)$  so that the successive appearances of  $p_2$  cover at least the last  $2l(p_1)$  periods of the history  $z_t$ , it follows that the last  $l(p_2)$  elements of the pattern  $\text{SH}_{-1}(p_1)$  must be equal to some shift of the pattern  $\text{SH}_{-1}(p_2)$ , say  $\text{SH}_k(\text{SH}_{-1}(p_2))$ . This, together with (\*\*\*) implies that  $\text{SH}_{-1}(p_2) = \text{SH}_k(\text{SH}_{-1}(p_2))$ , for some  $k \in \{1, \dots, l(p_2) - 1\}$ . It is straightforward to verify that  $p_2$  cannot be a basic pattern in such a case.

In the sequel we thus say that a given history  $z_t$  satisfies the sufficient conditions for convergence (to  $p$ ) iff there is a basic PN pattern  $p$  with  $l(p) \leq L$  such that  $\text{PRO}_{t-T_p \cdot l(p)+1, t-1}(z_t) = \text{EXT}_{1, T_p \cdot l(p)-1}(p)$ . Note that these sufficient conditions are stationary in the sense that the agents always recognize a given pattern after some fixed number of successive repetitions. In the next section we claim that this stationary approach is "too stylized" and thus suggest a generalized model where the agents can modify their pattern recognition behavior along the repeated game. The simplistic scheme suggested above, however, is sufficient for our main convergence result:

**DEFINITION.** Confused learning with stationary pattern recognition. Let  $(\mu, \phi)$  be a myopic learning model on a generic, strategic-form game  $\Gamma$ . We say that  $(\mu, \phi)$  is a confused learning model with stationary pattern

<sup>10</sup> The assumption can be generalized by requiring that for every pattern  $p$  with  $l(p) \leq L$  there is a summable sequence  $\{\beta_n(p)\}_{n \in \mathcal{N}}$ , such that whenever  $p$  appears  $n \geq T_p$  times successively, the probability that each player recognizes the pattern is  $\geq 1 - \beta_n(p)$ . To simplify the exposition we restrict the analysis above to the case where  $\beta_n(p) = 0$  for large  $n$ .

<sup>11</sup> The stronger assumption  $T_p \cdot l(p) \geq 3L$  is used in the Proof of Proposition 5.2.

recognition on  $\Gamma$  iff there is a pattern recognition frame  $\psi$  such that the collection  $(\mu, \phi, \psi)$  satisfies A.1, A.2, B.1, B.2, and  $\psi(z_t) = p$  if and only if the sufficient conditions for convergence to  $p$  hold.<sup>12</sup>

**DEFINITION.** Convergence to a PN pattern. Let  $(\mu, \phi)$  be learning model on a strategic-form game  $\Gamma$  with  $\text{PN}(\Gamma) \neq \emptyset$ . We say that  $(\mu, \phi)$  converges a.s. to a PN pattern iff  $\mathcal{P}_{(\mu, \phi)}(z \text{ for which } \exists T \text{ and a PN pattern } p \text{ such that } \text{PRO}_{T+(k-1) \cdot l(p), T+k \cdot l(p)-1}(z) = p \text{ for every } k = 1, 2, \dots) = 1$ .

**PROPOSITION 5.2.** *Let  $\Gamma$  be a generic, simple strategic-form game. Let  $(\mu, \phi)$  be a confused learning model with stationary pattern recognition on  $\Gamma$ . Then  $(\mu, \phi)$  converges a.s. to a PN pattern.*

The probability of convergence to a specific pattern is not necessarily decreasing in the pattern's length in our model. In a repeated play of the Battle of the Sexes (Fig. 2.2), for example, if the agents basically play the PN pattern  $(A, A), (B, B)$  repeatedly, but "tremble" with some small probability  $\epsilon$  to satisfy the confused learning assumptions A.1 and A.2, then the probability of convergence to  $(A, A), (B, B)$  goes to one as  $\epsilon$  approaches 0. In general note that since we have only imposed  $\epsilon$ -probability conditions on the agents' behavior in the confused mode (and left their complete behavior unspecified), "new" patterns of length  $\geq 2$  may emerge with high probability even when  $\epsilon$  is arbitrarily small. The specific distribution of PN patterns adopted by the agents at the limit might thus take many different forms.

### 5.3. Nonstationary Pattern Recognition

In the Introduction we emphasized the motivation to present an intuitively appealing set of assumptions on learning by agents of bounded rationality. The basic model presented in the previous section, however, seems too stylized to fit any realistic learning effort. The problem follows from the fact that the players in that model always recognize each pattern  $p$  after some fixed, history-independent number of successive repetitions,  $T_p$ . Consider, for example, the case where the players always recognize the singleton pattern  $A$  after 100 successive repetitions. Assume that the best response to  $A$  is  $B$ . Given the rich-support assumptions of the model,

<sup>12</sup> Existence is trivial. Just choose a positive probability  $\epsilon < \text{card}(S)^{-1}$ , an integer  $L$ , and a corresponding collection of  $T_p$ 's that satisfy the conditions of Assumption B.2. Let  $(\mu, \phi)$  be any learning model where the agents recognize a pattern iff the sufficient conditions hold, and where—in the confused learning mode—for every strategy profile  $s$  that has been observed in the past-play, each player plays a best response to  $s$  with probability at least  $\epsilon$ .

realized play might be

$$\overbrace{A, A, \dots, A}^{100 \text{ times}}, B, \overbrace{A, A, \dots, A}^{100 \text{ times}}, B, \overbrace{A, A, \dots, A}^{100 \text{ times}}, B, \overbrace{A, A, \dots, A}^{100 \text{ times}}, \\ B, A, A, A, A \dots^{13}$$

We suggest that this behavior is unreasonable, and that the agents should recognize the “familiar” singleton pattern  $A$  (with probability one) sooner; i.e., after a smaller number of cyclic repetitions, when the observed history is as above. In this section we present a generalized version of the model that accommodates such “dynamic pattern recognition.”

In the generalized model, the agents can change their pattern-recognition patterns as the game evolves. Different agents might have different inclinations for strategic pattern recognition. These individual inclinations might even change along the game. The notation  $\psi(z_t) = p \neq \theta$  will now be used to denote the case where at least one player believes that he has identified a cyclic pattern in the observed path of play;  $\psi(z_t) = \theta$  will still denote the default mode where all the players feel confused when observing  $z_t$ . Assumptions A.1, A.2, B.1, and B.2 still hold in the modified model. However, we now complement the sufficient conditions of Assumption B.2 by the following minimal, necessary conditions:

*Assumption B.3. Necessary conditions for pattern recognition.*

3.1: For every history  $z_t$ , for every basic pattern  $p$  with  $l(p) \geq 3$ , if  $\psi(z_t) = p$ , then  $\text{PRO}_{t-2 \cdot l(p)+1, t-1}(z_t) = \text{EXT}_{1, 2 \cdot l(p)-1}(p)$ .

3.2: For every history  $z_t$ , for every basic pattern  $p$  with  $l(p) = 2$ , if  $\psi(z_t) = p$ , then  $\text{PRO}_{t-4, t-1}(z_t) = \text{EXT}_{2, 5}(p)$ .

3.3: For every history  $z_t$ , for every singleton pattern  $p$ , if  $\psi(z_t) = p$ , then  $\text{PRO}_{t-2, t-1}(z_t) = p, p$ .

For example, a necessary condition for  $\psi(z_t) = A, B, C, D$  is that the last seven observations in  $z_t$  are  $A, B, C, D, A, B, C$ . When the pattern is short (of length 1 or 2) we require a slightly stronger condition; a necessary condition for  $\psi(z_t) = A, B$  is that the last four observations in  $z_t$  are  $B, A, B, A$ . In the sequel we say that a pattern  $p$  is recognizable given the history  $z_t$  iff the necessary conditions for  $\psi(z_t) = p$  (as defined in this assumption) hold. Note that these minimal necessary conditions leave space for dynamic, nonstationary pattern recognition as suggested above.

<sup>13</sup> Assumption B.2 implies that when  $T_A = 100$ ,  $L \leq 33$  so that the agents cannot recognize with probability one the longer pattern

$$\overbrace{A, A, \dots, A, B}^{100 \text{ times}}$$

If, at some stage of the repeated game, some “unfamiliar” pattern starts repeating successively in the observed play, it might take the players a long time, say, 100 successive repetitions, to recognize that pattern. If, however, after a short interruption, the same pattern starts repeating successively once again, the players might recognize it earlier, after 20 cyclic repetitions. If the same pattern starts repeating successively once more—for the third time in a relatively short period of time—the players might recognize it even earlier, after 3 successive repetitions.

We argued above that in the modified model, the pattern recognition mode should represent those cases where at least one player believes that his opponents follow some cyclic pattern. Assumption B.2 stipulates that, at the limit, if the pattern has repeated successively a large enough number of times, all the players recognize it with probability one. In general, however, it might be the case that only one player recognizes a pattern in a given history of the game, while the other players are still “confused” when observing the same history. The following Assumption B.4 says that in such cases, each player recognizes the pattern with some positive probability  $\delta$ .

*Assumption B.4. Coordinated pattern recognition.* There is a positive probability  $\delta$  such that for every basic pattern  $p$ , for every history  $z_t$  for which  $\psi(z_t) = p$ ,  $\mu_i^i(z_t)(\text{PRO}_{i(p)}^{-i}(p)) \geq \delta$ , for every  $i \in I$ .

For example, if  $z_t = \dots, A, B, C, A, B, C, A, B, C, A$  and  $\psi(z_t) = C, A, B$ , the assumption implies that with probability at least  $\delta^{\text{card}(I)}$ , the players recognize the pattern  $C, A, B$  by playing a best response to the pattern’s prediction, strategy  $B$ , at the next stage of the game.

One undesirable complication that may arise from using minimal necessary conditions for pattern recognition is that the agents might recognize one pattern, then immediately recognize a different pattern that was generated when the previous pattern was contradicted and so on . . . . If, for example, the best response to some strategy  $B$  is  $C$ , and the agents recognize the pattern  $A, C, A, B$  in some history  $\tilde{z}_t$ , then the extended history  $\tilde{z}_{t+1}$  ends with the sequence  $A, C, A, C$  so that the necessary conditions for recognizing the pattern  $C, A$  hold. The next assumption says that if  $z_t$  is a history that ends with pattern contradiction (i.e., if the players were in the pattern recognition mode given the corresponding history  $z_{t-1}$  and the pattern has been contradicted at date  $t - 1$ ), then the agents will not recognize new patterns for a while. Specifically, a necessary condition for recognizing a “new” pattern  $p'$  at some future date  $t + k$  is that  $k$  is bigger than  $l(p')$ . When applied to the example above, the assumption implies that the agents cannot be in the pattern recognition mode with respect to the pattern  $C, A$  given the extended history  $\tilde{z}_{t+1}$ .

*Assumption B.5. Caution in pattern recognition.*  $\exists \bar{T}$  such that for every basic pattern  $p$ , for every  $t \geq \bar{T}$ , for every history  $z_t$  for which  $\psi(z_t) = p$ , for every extension of  $z_t$ , say  $z_{t+k}$ , satisfying (a)  $\text{PRO}_t(z_{t+k}) \neq \text{PRO}_{l(p)}(p)$ , and (b)  $\psi(z_{t+k}) = p'$  for some basic pattern  $p'$ , it must be the case that  $k > l(p')$ .

Recalling the two postulates concerning the agents' learning behavior, confused learning, and pattern recognition, we may interpret B.5 as fixing the priorities between these basic assumptions. The agents' fundamental strategic confusion is stronger than their belief in strategic pattern recognition in the advanced stages of the game (before convergence).

The following final assumption complements the caution assumption. It says that if the agents are in the pattern recognition mode with respect to some pattern  $p$  given the history  $z_t$  and realized play at date  $t$  conforms to the pattern's prediction, then the agents stay in the pattern recognition mode with respect to (the appropriate shift of) the same pattern.

*Assumption B.6. Continuation.* For every basic pattern  $p$ , for every history  $z_t$  for which  $\psi(z_t) = p$ ,  $\psi(z_t, \text{PRO}_{l(p)}(p)) = \text{SH}_1(p)$ .

For example, if  $z_t = \dots A, B, C, A, B$ ,  $\psi(z_t) = A, B, C$ , and the realized play at date  $t$  is  $s_t = C$ , then the assumption implies that  $\psi(z_t, s_t) = B, C, A$ .

**DEFINITION.** Confused learning with dynamic pattern recognition. Let  $(\mu, \phi)$  be a myopic learning model on a generic, strategic-form game  $\Gamma$ . We say that  $(\mu, \phi)$  is a confused learning model with dynamic pattern recognition on  $\Gamma$  iff there is a pattern recognition frame  $\psi$  such that the collection  $(\mu, \phi, \psi)$  satisfies A.1, A.2, and B.1–B.6.

Note that the stationary-pattern-recognition model of Section 5.2 is a special, degenerated, case of confused learning with dynamic pattern recognition; i.e., let  $(\mu, \phi)$  be a confused learning model with stationary pattern recognition. Observe that the model satisfies the additional assumptions B.3, B.4, and B.6. Assume by way of contradiction that it violates the caution assumption B.5. Then there is a history  $z_t$  and an extension of  $z_t$ ,  $z_{t+j}$ , such that a pattern  $p$  has been contradicted at date  $t$ ,  $\psi(z_{t+j}) = \tilde{p}$  and  $j \leq l(\tilde{p})$ . Since  $T_{\tilde{p}} \cdot l(\tilde{p}) \geq 3L$ , the successive repetitions of  $\tilde{p}$  must then cover the last  $2L$  elements of the history  $z_t$ . Since  $T_p \cdot l(p) \geq 3L$  as well, both patterns  $p$  and  $\tilde{p}$  cover the last  $2L$  elements of  $z_t$ . From the uniqueness proof in Section 5.2, it follows that  $p = \tilde{p}$ . Since  $p$  was contradicted at date  $t$ , we get a contradiction to the assumption  $\psi(z_{t+j}) = \tilde{p}$  which proves that B.5 must hold as well.

**PROPOSITION 5.3.** *Let  $\Gamma$  be a generic, simple strategic-form game. Let  $(\mu, \phi)$  be a confused learning model with dynamic pattern recognition on  $\Gamma$ . Then  $(\mu, \phi)$  converges a.s. to a PN pattern.*

## 6. CONVERGENCE OF BEHAVIOR TO A MIXED EQUILIBRIUM

Can the agents' behavior converge to a mixed strategy Nash equilibrium in our strictly myopic learning environment?

The concept of "convergence of behavior strategies" was proposed by Fudenberg and Kreps (1993) as a more appropriate notion of convergence to a mixed strategy Nash equilibrium than the previously used "convergence of beliefs (or empirical frequencies)" criterion. Formally, Fudenberg and Kreps studied the following definition of local stability:

**DEFINITION.** Local stability.<sup>14</sup> Let  $(\mu, \phi)$  be a myopic learning model on a strategic-form game  $\Gamma$ . A mixed strategy profile  $\sigma^*$  is said to be locally stable with respect to  $(\mu, \phi)$  iff for every  $\epsilon > 0$  there is some history  $z_t$  such that  $\mathcal{P}_{(\mu, \phi)}(\lim_{t' \rightarrow \infty} \phi_{t'}(z_{t'}) = \sigma^* | z_t) > 1 - \epsilon$ .

Fudenberg and Kreps presented two constructions under which local stability of a mixed strategy Nash equilibrium profile is technically possible. In the first construction, the players maintain the equilibrium beliefs unless and until sufficient evidence against these beliefs has accumulated. Moreover, when facing the equilibrium beliefs, the players choose the "right" tie-breaking rules; that is, the equilibrium strategies. In the second construction, the players maintain the empirical beliefs but they are only asymptotically myopic. Thus, they keep playing the equilibrium strategies unless and until the cost of doing so becomes too large. As discussed by Fudenberg and Kreps (1993, p. 346), "in both constructions the players use precisely the equilibrium strategies with no positive reason at all. . . . Thus, although the constructions show that convergence to a mixed strategy Nash equilibrium is possible, neither one convinces us that it would in fact happen, except perhaps for players who have been trained in game theory and therefore know how they are 'expected' to act."

The players in our model, however, are strictly myopic, and in general know nothing about game theory and equilibrium analysis. In this setting, a necessary condition for convergence of intended behavior to a mixed strategy Nash equilibrium is that the predictions reach the equilibrium strategies and stay there forever. This seems "too strong to be possible" in our bounded rationality, incomplete information environment. As noted by Jordan (1993), generically, a mixed strategy is expected-payoff maximizing for some player only if the expected mixed strategies of the other players lie in a subspace of lower dimension of the corresponding strategies' space.

<sup>14</sup> Fudenberg and Kreps studied a deterministic learning model where  $\mu_i^j: \mathcal{Z}_i \mapsto \Delta(S^{-i})$ . Although our probabilistic prediction rules are different, the resulting myopic behavior rules are of the same form in both cases.

Since the number of finite histories of play is countable, it is a priori unlikely that the players' expectations will ever lie in the subspace for which the mixed strategy is a best response.

Moreover, we now show that convergence of behavior to *any* fixed behavior strategy (that is not a pure equilibrium) is incompatible with strategic pattern recognition in the sense that the players don't recognize some patterns even when they appear successively an arbitrarily large number of times.

**DEFINITION.**  $\tilde{\mathcal{E}}$  incompatibility with pattern recognition. Let  $(\mu, \phi)$  be a myopic learning model on a strategic-form game  $\Gamma$ . Let  $\tilde{\mathcal{E}}$  be a collection of basic patterns in  $\Gamma$ . A learning model  $(\mu, \phi)$  is said to be  $\tilde{\mathcal{E}}$  incompatible with pattern recognition iff for every  $\epsilon > 0$  there is some history  $z_t$  such that  $\mathcal{P}_{(\mu, \phi)}(\bigcap_{p \in \tilde{\mathcal{E}}} \bigcap_{n \in \mathcal{N}} p \text{ has appeared } n \text{ times successively} \mid z_t) > 1 - \epsilon$ .

In the following proposition we use  $\tilde{\mathcal{E}}(\sigma^*)$  to denote the family of basic patterns that correspond to the mixed strategy  $\sigma^*$ ; i.e.,  $p = s_1, s_2, \dots, s_m \in \tilde{\mathcal{E}}(\sigma^*)$  iff  $p$  is a basic pattern in  $\Gamma$  and  $s_l \in \text{supp}(\sigma^*)$  for every  $l = 1, \dots, m$ . To smooth the exposition we restrict the formal exposition to games with no pure strategy Nash equilibria.

**PROPOSITION 6.** *Let  $\Gamma$  be a strategic-form game with  $\text{PN}(\Gamma) = \emptyset$ . Let  $(\mu, \phi)$  be a myopic learning model on  $\Gamma$ . If a mixed strategy profile  $\sigma^*$  is locally stable with respect to  $(\mu, \phi)$ , then  $(\mu, \phi)$  is  $\tilde{\mathcal{E}}(\sigma^*)$  incompatible with pattern recognition.*

An important immediate Corollary to Proposition 6 is that convergence of behavior to some mixed strategy profile  $\sigma^*$  is impossible if there exists a uniform bound  $T$  such that each player always plays a best response to some strategy that has been observed in the last  $T$  periods. In this sense, a necessary condition for convergence of behavior to a mixed strategy Nash equilibrium is that the players consider arbitrarily long histories when forming their predictions. Agents with bounded memory can never converge to play a mixed strategy Nash equilibrium.

## 7. A GENERAL CONVERGENCE RESULT

By adding one assumption to the model of Section 5 we now get a general convergence result for finite strategic-form games. The extra assumption,  $*$ -adaptiveness, says that for every subset  $E$  of pure strategies in the game, there is a uniform bound  $T_E$ , such that whenever the realized play at the last  $T_E$  periods of the game is a subset of  $E$ , the probability



that each player will play a best response to some strategy in  $E$  on the next stage of the game is one. In the formal definition that follows we use the expression  $\mu_t^i(z_t)(E)$  to denote the sum,  $\sum_{s \in E} \mu_t^i(z_t)(s^{-i})$ .

*Assumption C. \*-adaptiveness.* For every  $E \subseteq S(\Gamma)$ , there is a bound  $T_E$  such that for every history  $z_t$  with  $t \geq T_E$  for which  $\text{PRO}_{t'}(z_t) \in E$   $\forall t' = t - T_E, t - T_E + 1, \dots, t - 1$ ,  $\mu_t^i(z_t)(E) = 1$ , for every  $i \in I$ .

It is easy to prove that Assumption C doesn't contradict our previous assumptions A.1, A.2, and B.1–B.6 if the bounds  $T_E$  are chosen appropriately.

Recalling that a subgame  $\Gamma^*$  is closed iff the best response to each strategy in the subgame is in the subgame, we say that a closed subgame  $\Gamma^*$  is *minimal* iff it doesn't contain another closed subgame. Note that every strict pure strategy Nash equilibrium is a minimal closed subgame. Note also that in games that don't have strict equilibria, it might be the case that the (unique) minimal closed subgame is the game itself (e.g., Matching Pennies). In the following proposition we use "convergence to MCS with no PN" to denote the event:  $\{s_n \in \Gamma^* \text{ eve., for some minimal closed subgame } \Gamma^* \text{ such that } \text{PN}(\Gamma) \cap S^*(\Gamma^*) = \emptyset\}$ .<sup>15</sup>

**PROPOSITION 7.** *Let  $\Gamma$  be a generic, finite strategic-form game. Let  $(\mu, \phi)$  be a confused learning model with dynamic pattern recognition on  $\Gamma$  that satisfies the additional assumption C. Then,  $\mathcal{P}_{(\mu, \phi)}(\text{convergence to a PN pattern or convergence to MCS with no PN}) = 1$ .*

To demonstrate the implications of Proposition 7, consider again the stylized example of Figure 4.1. The game in that figure has three minimal closed subgames: The two strict Nash equilibria and the Matching Pennies subgame. Our result thus implies that a confused learning model with strategic pattern recognition on that game that satisfies the additional assumption C must either converge to the Matching Pennies subgame, or to some PN pattern in the Battle of the Sexes subgame.

## 8. GENERALIZATIONS AND DISCUSSION OF ASSUMPTIONS

### 8.1. *Relaxing Strict Myopia*

The strict myopia assumption was used in the Introduction to motivate our confused learning and strategic pattern recognition postulates. Indeed, these assumptions seem to suit best a learning environment where the agents are strictly myopic. In this section, however, we informally suggest

<sup>15</sup>  $\{E_n \text{ eve.}\}$  is used to denote the event  $\liminf E_n$  (Billingsley, 1986).

	Col C	Col D
Row C	4,4	0,5
Row D	5,0	1,1

FIG. 8.1. Prisoner's Dilemma.

that the myopia assumption can be relaxed to some degree. Our agents may try to take account of the more distant future when searching for the "best" move for the next stage of the game.<sup>16</sup> The players are of bounded rationality, however, and thus they don't follow rigorously the Bayesian learning scheme (Kalai and Lehrer, 1993). Moreover, we assume that: (1) learning is basically confused so that many different behaviors (e.g., those *corresponding to* Cournot dynamics and sophisticated Cournot dynamics) may occur in most advanced stages of the game and (2) the agents recognize cyclic patterns in the observed path of play. Any learning model that satisfies these assumptions must converge to a PN pattern when the stage game is simple.

Note that convergence to a PN pattern can be given a nonmyopic justification in many applications, by assuming that whenever a player detects a recurrent pattern in the path of play, he believes that his rivals will keep following the pattern if he follows the pattern, while any deviation on his part will trigger a painful retaliation by his opponents. Yet, "Compatibility with Strategic Pattern Recognition" must be given some myopic justification in certain applications. In a repeated play of the Prisoner's Dilemma (Fig. 8.1), for example, compatibility with singleton pattern recognition implies that any string of repeated Cooperation must come to an end at some stage of the repeated game. It thus precludes the possibility that both players adopt some "tit for tat" (that is, "she will cooperate as long as I cooperate, but she will fiercely retaliate if I deviate") type of conjectures concerning their opponents' behavior in the repeated game.

To justify myopia, the literature typically suggests that the players are impatient and thus discount the future heavily, or that the players ignore the impact of their current strategic behavior on the future play of their opponents because of their bounded rationality. To get a more palatable justification of the assumption, Fudenberg and Kreps (1993) invoked a

<sup>16</sup> Technically, this can be done by reformulating our basic model so that the assumptions only refer to the behavior rules, bypassing the (not necessarily myopic) prediction process from which the agents arrive at these behaviors.

large population story assuming that players are repeatedly drawn from some large populations, so that: (1) repeated meetings between a given set of players are rare and (2) each player is unaware of how his specific opponents at some given stage of the game have acted in the previous rounds in which they were called to play. Alternatively, we may assume that players have a finite life span, and that once a player dies, he is immediately replaced by an identical offspring that observes the entire history upon joining the game. Assuming that the players are of bounded rationality and their life span is random and relatively short, we hypothesize that strictly myopic behavior will emerge.

## 8.2. Other Forms of Pattern Recognition

The strategic pattern recognition scheme suggested in the previous sections has the following restrictive features: (1) It only considers *patterns of full dimensionality* that are built from pure strategy profiles in the game and specify the behavior of all the players simultaneously. (2) Players may fully recognize a pattern only if it has appeared successively with no interruptions at the end of the observed history of play. This section is a brief discussion of some of the “other cases” where strategic pattern recognition might seem appropriate.

Consider first the case where a pattern appears successively with non-contradicting interruptions. For example, take  $z_t = A, B, C, A, A, B, C, D, A, B, C, A, D, A, B, C, D, C, C, A, B$ , where the pattern  $A, B, C$  has appeared four times but the successive appearances of the pattern have been interrupted by some random sequence of strategies that don't include a *contradicting pattern* of the type  $A, B, \tilde{s}$ , where  $\tilde{s}$  is some strategy profile different from  $C$ .

Formally, we may say that a strategic pattern  $p$  of length  $n \geq 2$  in a strategic-form game  $\Gamma$  has appeared  $N$  times successively (with interruptions) in a given history  $z_t$  iff there is a sequence of  $N$  indexes in  $\{n, n + 1, \dots, t - 1\}$ , say  $i_1 \leq i_2 \leq \dots \leq i_N$ , such that (1)  $p$  has appeared at  $i_k$  for every  $k = 1, 2, \dots, N$ , and (2) for each  $\tilde{t} \in \{i_1, \dots, t - 2\}$ , if  $\text{PRO}_{\tilde{t}-n+2, \tilde{t}}(z_{\tilde{t}}) = \text{PRO}_{1, n-1}(p)$ , then  $\text{PRO}_{\tilde{t}+1}(z_{\tilde{t}}) = \text{PRO}_n(p)$ .<sup>17</sup>

The agents of Section 5 cannot fully recognize patterns that have appeared successively *with interruptions*; e.g., given any history ending with  $A, B$ , in which the pattern  $A, B, C$  has appeared  $n$  times successively with interruptions (so that the necessary conditions for recognizing the pattern  $A, B, C$  as defined in Section 5 don't hold), there is a positive probability  $\epsilon$  that the players will follow Cournot dynamics and play a best response to

<sup>17</sup> In the case of singleton patterns the only definition of successive appearances that makes sense is the one requiring no interruptions.

the strategy  $B$  on the next stage of the game. If the best response to  $B$  is different from the best response to  $C$ , the probability of recognizing the pattern  $A, B, C$  given such a history is bounded above by  $1 - \epsilon$ , independently of  $n$ .

Similarly, the agents of Section 5 cannot recognize with probability one patterns of partial dimensionality that involve a proper subset of the set of players; our model ignores such patterns even when “all the other players except player  $i$  successively repeat a given pattern” although one may claim that—as far as player  $i$  is concerned—these patterns are almost equivalent to patterns of full dimensionality.

We also ignore the following types of patterns:

—Stochastic patterns where, for example, the strategy profile  $A$  is “always” followed by  $B$  or by  $C$ , but in no specific order.

—“Patterns of patterns” where, for example, the patterns  $A, B$  and  $A, C$  appear repeatedly, one after the other, but in between the appearance of  $A, B$  and the following appearance of  $A, C$  (and vice versa) there appears a sequence of strategies of random length and random structure that doesn’t include the strategy  $A$ .

—Time related patterns where, for example, agent  $i$  plays strategy  $\tilde{s}^i$  on Friday.

To defend the narrow scope of our pattern recognition scheme, we may first argue that these other patterns are somewhat more difficult to observe, and thus there is “always”—no matter how many times such a pattern appears successively—a positive probability  $\epsilon'$  that the agents will fail to recognize it. An additional excuse for ignoring patterns of partial dimensionality may be based on the notion of sophisticated learning (Milgrom and Roberts, 1991). A sophisticated player observing a recognizable pattern of partial dimensionality might expect that the players involved in playing the pattern will respond on the next stage of the game to the latest changes that have occurred in the play of those agents that are not following the pattern (these may include the player himself). He thus may be uncertain whether these players will keep following the pattern on the next stage of the repeated game, and feel, overall, “confused” about the future course of the game.

At the same time we claim that the main convergence result should hold even if the agents are able to recognize some of the other patterns mentioned above. The basic intuition still applies: The confused learning forces should lead the players to an equilibrium pattern when the stage game is simple. The players may adopt such a pattern as a fixed behavioral convention on playing the game. In Appendix 2 we briefly outline a variation on the original model in which the agents can recognize patterns that appear successively with interruptions.

### 8.3. *Justifying the Uniform Assumptions*

The following example demonstrates that the uniform component of uniform reachability is essential to the main convergence result; i.e., the weaker confused learning assumptions: “For every confused-mode-history there is a positive probability that the players follow the BR dynamics and each of the behaviors associated with the sophisticated BR dynamics” are not sufficient for the main convergence result.

Fix a strategic-form game  $\Gamma$  where  $s_1$  and  $s_2 \neq s_1$  are non-Nash pure strategies in  $\Gamma$ , and  $s^*$  is a PN equilibrium such that  $s^* = \text{BR}(s_1) = \text{BR}(s_2)$ .

Let  $\{\alpha_t\}$  be a summable sequence of numbers in  $(0, 1)$ , and consider a “learning model”  $(\mu, \phi)$  on  $\Gamma$  defined by the following conditions:

- At date 0,  $s_1$  is played with probability 1.
- At every date  $t = 1, 3, 5, \dots$  if  $s_{t-1} \neq s^*$  then  $s_t = s_2$  with probability  $1 - \alpha_t$ , and  $s_t = s^*$  with the complementary probability.
- At every date  $t = 2, 4, 6, \dots$  if  $s_{t-1} \neq s^*$  then  $s_t = s_1$  with probability  $1 - \alpha_t$ , and  $s_t = s^*$  with the complementary probability.
- At every date  $t \geq 1$  if  $s_{t-1} = s^*$  then  $s_t = s^*$  with probability 1.

It is straightforward to verify that the model described above is a “confused learning model with strategic pattern recognition” if the bound  $L$  is set equal to 1, and if we relax the uniform part of the confused learning assumptions A.1 and A.2 (as discussed above). Yet,  $\mathcal{P}_{(\mu, \phi)}$ (convergence to a PN pattern)  $= 1 - \prod_{t \in \mathcal{J}} (1 - \alpha_t) < 1$ , since the product  $\prod_{t \in \mathcal{J}} (1 - \alpha_t)$  converges whenever the sequence  $\alpha_t$  is summable.

## 9. DISCUSSION

Herbert Simon (1992) writes on bounded rationality and the theory of decision making: “We know today that human reasoning, the product of bounded rationality, can be characterized as selective search through large spaces of possibilities. The selectivity of the search, hence its feasibility, is obtained by applying rules of thumb, or heuristics to determine what paths should be traced and what ones can be ignored. The search halts when a satisfactory solution has been found, almost always long before all alternatives have been examined.” With this spirit our main result demonstrates that a satisfactory solution in a myopic search for the right way to play an infinitely repeated simple game can be an equilibrium pattern of the stage game.

Some of the informal interpretations of Nash equilibria in the fundamental literature on game theory are “a steady state of an environment in which the players act repeatedly” (Osborne and Rubinstein, 1993), and “some convention...concerning how to play the game. ...a learned behavior” (Kreps, 1990). Our results support these interpretations—as far as pure strategy Nash equilibria of the stage game are concerned. When the stage game is simple, agents of bounded rationality that know nothing about game theory and equilibrium analysis may adopt a fixed pattern of such equilibria as a “behavioral convention” or a “steady state solution” to the game.

We were unable to derive any positive results concerning the general concept of mixed strategy Nash equilibrium. We have moreover shown that adaptive agents with bounded memory can never converge to play a mixed strategy Nash equilibrium. In this respect, we follow Fudenberg and Kreps (1994, 1995), Jordan (1993), and Kalai and Lehrer (1995) in questioning the adequacy of Nash equilibrium as an attainable solution to strategic games.

Our basic postulates, confused learning and strategic pattern recognition, are intuitively appealing, and we believe they would survive empirical testing. The model also provides a natural channel through which the context within which the repeated game is played can influence the equilibrium selection. In a repeated play of the Battle of the Sexes, for example, if one player is more aggressive than the other in the sense that she is less willing to settle on (say) equilibrium  $B$  that gives her the lower payoff, then our agents may converge to an equilibrium pattern like  $A, A, A, B$  or  $A, A, B, A, B$  in which the aggressive player gets a higher average payoff than her opponent.

The work is a first attempt to model strategic pattern recognition. We hope it will help to initiate more research on this important subject.

## APPENDIX 1: PROOFS OMITTED IN THE TEXT

In all the following proofs we assume w.l.g. that the  $\epsilon$  of the compatibility with BR dynamics assumption and the  $\epsilon$  of the compatibility with sophisticated BR dynamics assumption are equal.

*Proof of Proposition 4 (and a reachability lemma).* While proving the proposition, we will also prove a reachability lemma (Lemma 4.1) that is essential for proving the convergence results of Propositions 5.2, 5.3, and 7. Our specific plan is as follows: Recalling the compatibility with (sophisticated) BR dynamics assumptions of Proposition 4, we define a binary relation on the pure strategies' set,  $C$  reachability (in  $n$  steps). The

assumptions of Proposition 4 immediately imply that if a strategy  $s^*$  is C reachable from strategy  $s$  in no more than  $\text{COMP} = [\text{card}(S \setminus \text{PN}(\Gamma))]^2 - [\text{card}(S \setminus \text{PN}(\Gamma))] + 1$  steps, and the last observation in some history  $z_t$  is  $s$ , then the conditional probability (given  $z_t$ ) that  $s_n = s^*$  for some  $n \geq t$  is at least  $\epsilon^{\text{COMP} \cdot \text{card}(I)}$ . We then prove that in simple games, for every strategy  $s$ , there exists a PN strategy that is C reachable from  $s$  in no more than COMP steps (this is stated in Lemma 4.1). It follows that a model satisfying the conditions of Proposition 4 satisfies uniform reachability of PN equilibria.

**DEFINITION. C reachability.** Fix a generic, simple, strategic-form game  $\Gamma$ , and let  $s_0 \in S(\Gamma)$ . A strategy  $s' \in S(\Gamma)$  is C reachable from  $s_0$  (in  $n$  steps) iff there is a finite sequence of strategies in  $S(\Gamma)$ , say  $s_0, s_1, \dots, s_{n-1}, s_n = s'$  such that for every  $m = 1, 2, \dots, n$ , for every  $i \in I$ , one of the following two conditions hold:

- (1)  $s_m^i \in \text{BR}^i(s_{m-1})$ , or
- (2)  $m \geq 2$ , and  $\exists l \in \{0, \dots, m-2\}$  such that  $s_l = s_{m-1}$ , and  $s_m^i \in \text{BR}^i(s_{l+1})$ .

We say that  $s_0, s_1, \dots, s_n$  is a C sequence iff for every  $m = 1, 2, \dots, n$ , for every  $i \in I$ , one of the two conditions stated in the definition of C reachability holds. Note that C reachability is a transitive relation, and that every BR sequence is a C sequence. Note also that (since we restrict the analysis to generic games) once a C sequence hits a PN equilibrium it must stay there forever; i.e., if  $s_0, s_1, \dots, s_n$  is a C sequence, and  $s_m \in \text{PN}(\Gamma)$  for some  $m < n$ , then  $s_l = s_m \quad \forall l \geq m$ . We define a C extension of a C sequence  $s_0, s_1, \dots, s_m$  to be a C sequence  $\hat{s}_0, \hat{s}_1, \dots, \hat{s}_k$  such that  $k \geq m$ , and  $\hat{s}_j = s_j$  for  $j = 0, 1, \dots, m$ .

We now show that in simple games, pure strategy Nash equilibria are “always” C reachable in no more than COMP steps. Before stating and proving this result formally we have to introduce some additional terminology:

Let  $z_0 = s_0, s_1, \dots, s_k$  be a C sequence. We say that strategy  $s'$  has appeared in the sequence  $z_0$  (at date  $t$ ) iff the singleton pattern  $s'$  has appeared in that sequence (at date  $t$ ). We say that the pair  $s', s''$  has appeared in the sequence  $z_0$  (at date  $t$ ) iff the pattern  $s', s''$  has appeared in that sequence (at date  $t$ ). We say that the sequence  $z_0$  is *pair-wise disjoint* iff no pair has appeared in the sequence in two different dates. We say that the sequence contains *no repetitions* if no element in the sequence has appeared in both dates  $t$  and  $t+1$  for some  $t$ .

For any sequence of strategy profiles in  $\Gamma$ , say  $s_1, \dots, s_n$ , we use  $\{s_1, \dots, s_n\}$  to denote the set corresponding to that sequence; i.e.,  $\tilde{s} \in \{s_1, \dots, s_n\}$  iff  $\tilde{s} = s_i$  for some  $i \in \{1, \dots, n\}$ . For any set of pure strategies

$X$ , we use  $SG(X)$  to denote the subgame generated by the strategies in  $X$ ; i.e.,  $SG(X)$  is the subgame defined by  $\hat{S}^i = \{s^i \mid \exists \tilde{s} \in X \text{ s.t. } \tilde{s}^i = s^i\}$  for every player  $i$ .

LEMMA 4.1. *Fix a generic, simple, strategic-form game  $\Gamma$ . For every  $s \in S \setminus PN(\Gamma)$ ,  $\exists s^* \in PN(\Gamma)$  that is C reachable from  $s$  in no more than  $[card(S \setminus PN(\Gamma))]^2 - [card(S \setminus PN(\Gamma))] + 1$  steps.*

*Proof.* Let  $s_0 \in S \setminus PN(\Gamma)$ . We prove the lemma by proving the existence of a pair-wise disjoint, no repetitions C sequence leading from  $s_0$  to a PN. Note that the number of distinct strategy pairs  $s', s''$ , such that  $s'$  and  $s'' \in S \setminus PN(\Gamma)$ , and  $s' \neq s''$ , is  $[card(S \setminus PN(\Gamma))]^2 - [card(S \setminus PN(\Gamma))]$ . Thus, if  $s_0, s_1, \dots, s_n = s^*$  is a pair-wise disjoint, no repetitions C sequence leading from  $s_0$  to a PN  $s^*$ , then  $n \leq [card(S \setminus PN(\Gamma))]^2 - [card(S \setminus PN(\Gamma))] + 1$ , which is the bound given in the lemma.

Let  $X$  denote the space of pair-wise disjoint, no repetitions C sequences starting from  $s_0$ . Note that  $X$  is nonempty since it includes the one element sequence  $s_0$ . Let  $\succ$  denote the set inclusion partial ordering on  $X$ ; i.e., for every two sequences in  $X$ , say  $z_1 = s_0, s_1, \dots, s_n$ , and  $z_2 = s_0, \tilde{s}_1, \dots, \tilde{s}_m$ , we say that  $z_1 \succ z_2$  iff  $\{s_0, s_1, \dots, s_n\} \supset \{s_0, \tilde{s}_1, \dots, \tilde{s}_m\}$ .

By Hausdorff maximal principle there exists a pair-wise disjoint, no repetitions C sequence in  $X$  that is maximal under the set inclusion partial ordering. That is, there is a pair-wise disjoint, no repetitions C sequence, say  $z_0 = s_0, s_1, \dots, s_n$  such that no sequence in  $X$  is an extension of  $z_0$ . We now show that  $s_n$  must be a PN in  $\Gamma$ .

Assume by way of contradiction that  $s_n$  is not a PN. Note that if  $s_n$  has not appeared at some date  $t < n$  in the sequence  $z_0$ , then  $s_0, s_1, \dots, s_n, BR(s_n)$  is a pair-wise disjoint, no repetitions C extension of  $z_0$ , which contradicts the maximality of  $z_0$ .

Thus, if  $s_n$  is not a PN, it must have appeared at some date  $t < n$  in the sequence  $z_0$ . Let  $t_1, t_2, \dots, t_m$  be the subsequence of indexes (in  $\{0, 1, 2, \dots, n - 1\}$ ) in which the strategy  $s_n$  has previously appeared in the sequence  $z_0$ . The strategies  $s_{t_1+1}, s_{t_2+1}, \dots, s_{t_m+1}$  will accordingly denote the previous successors of  $s_n$ . Consider the subgame spanned by these strategies and the best response to these strategies; i.e.,  $SG(\{s_{t_1+1}, s_{t_2+1}, \dots, s_{t_m+1}, BR(s_{t_1+1}), BR(s_{t_2+1}), \dots, BR(s_{t_m+1})\})$ . In the sequel we use SGA to denote this subgame.

Note that by definition of C reachability (most importantly, by the fact that the conditions in the definition of C reachability hold independently across players),

(\*) for every strategy  $\tilde{s}$  in SGA, the sequence  $z_0, \tilde{s}$  is a C extension of  $z_0$ .<sup>18</sup>

<sup>18</sup> This "subgame spanning" idea has been demonstrated in Example 2.1.



Since  $\Gamma$  is simple, if SGA is closed, it must contain a PN equilibrium, say  $s^*$ . But then it follows from (\*) that  $z_0, s^*$  is a pair-wise disjoint, no repetitions C extension of  $z_0$ , which contradicts the maximality of  $z_0$ . Thus, SGA cannot be closed and,

(\*\*\*) there must be some strategy  $s' \in \text{SGA}$  such that  $s'' = \text{BR}(s') \notin \text{SGA}$ .

Now note that if  $s_n = s'$  then, by definition of C reachability,  $s_{t+1} = s''$ . But then  $s'' \in \text{SGA}$  which contradicts (\*\*). Thus,

(1)  $s_n \neq s'$ .

Note also that if the pair  $s_n, s'$  has appeared in the sequence  $z_0$ , then (by definition of SGA)  $s'' = \text{BR}(s') \in \text{SGA}$ , which again contradicts (\*\*). Thus,

(2) the pair  $s_n, s'$  has not appeared in the sequence  $z_0$ .

From (\*), (\*\*), (1), and (2) it finally follows that the sequence  $z_0, s'$  is a pair-wise disjoint, no repetitions C extension of  $z_0$ . Since this contradicts the maximality of  $z_0$ , it follows that  $s_n$  must be a pure strategy Nash equilibrium. ■

Given the definition of C reachability and the result of Lemma 4.1, it is now easy to observe that if a learning model satisfies  $\epsilon$  compatibility with BR dynamics and  $\epsilon$  compatibility with sophisticated BR dynamics, then it satisfies uniform reachability of PN equilibria and the proposition holds. For the sake of a complete presentation we write down the form argument: Let  $s_0 \in S \setminus \text{PN}(\Gamma)$ .

By the lemma  $\exists s^* \in \text{PN}(\Gamma)$ ,  $\exists q \leq [\text{card}(S \setminus \text{PN}(\Gamma))]^2 - [\text{card}(S \setminus \text{PN}(\Gamma))] + 1$ , and a sequence, say,  $s_0 = x_0, x_1, \dots, x_q = s^*$ , such that for every  $m = 1, \dots, q$ , for every  $i \in I$ , one of the following conditions hold:

(1)  $s_m^i = \text{BR}^i(s_{m-1})$ , or

(2)  $m \geq 2$ , and  $\exists l \in \{0, \dots, m-2\}$  such that  $s_l = s_{m-1}$ , and  $s_m^i = \text{BR}^i(s_{l+1})$ .

Thus, for every  $m = 1, \dots, q$ , for every  $i \in I$ ,  $\exists y_m(i) \in \{x_0, \dots, x_{m-1}\}$  such that  $x_m^i = \text{BR}^i(y_m(i))$ . Let  $t \geq \bar{T}$ , and let  $z_t \in \mathcal{Z}_t$  be a history ending with  $s$  (i.e.,  $\text{PRO}_{t-1}(z_t) = s = x_0$ .) Denote  $z_{t+m} = \{\text{PRO}_{1,t-2}(z_t), x_0, \dots, x_m\}$  for  $m = 0, 1 \dots q-1$ . Since  $\mu$  is  $\epsilon$  compatible with BR dynamics, and  $\epsilon$  compatible with sophisticated BR dynamics, we get that

$$\mu_{t+m}^i(z_{t+m}) \left( (y_{m+1}(i))^{-i} \right) \geq \epsilon \quad \text{and} \quad \phi_{t+m}^i(z_{t+m}) (x_{m+1}^i) \geq \epsilon,$$

for every  $i \in I$ , for every  $m = 0, \dots, q-1$ . This implies that  $\mathcal{P}_{(\mu, \phi)}(s_{t+m} = x_{m+1} \mid z_{t+m}) \geq \epsilon^{\text{card}(I)}$  for every  $m = 0, \dots, q-1$ , so that  $\mathcal{P}_{(\mu, \phi)}(s_{t+q-1} = s^* \mid z_t) \geq \epsilon^{(\text{card}(I) \cdot q)} \geq \epsilon^{(\text{card}(I) \cdot [\text{card}(S)]^2)}$ , whenever the conditional probability is well defined.

Thus, for every  $t \geq \bar{T}$ , for every  $z_t$  such that  $\text{PRO}_{t-1}(z_t) \notin \text{PN}(\Gamma)$ ,  $\mathcal{P}(s_n = s^* \text{ for some } n \geq t \mid z_t) \geq \epsilon^{(\text{card}(T) \cdot [\text{card}(S)]^2)}$ , whenever the conditional probability is well defined, and  $(\mu, \phi)$  satisfies uniform reachability of PN equilibria. ■

In the following proofs, we sometimes use conditional probability statements of the form  $\mathcal{P}_{(\mu, \phi)}(E \mid z_t)$  without writing down the  $\mathcal{P}_{(\mu, \phi)}(z_t) > 0$  qualification. The condition is assumed implicitly.

*Proof of Proposition 5.2.* Fix a generic, simple, strategic-form game  $\Gamma$ . Let  $(\mu, \phi)$  be a confused learning model with stationary pattern recognition on  $\Gamma$ . We start by proving three straightforward lemmas. Lemma 5.2.1 says that if the conditional probability of convergence to a PN pattern given the history  $z_t$  is uniformly bounded below across histories by some positive  $\epsilon$ , then the model converges to a PN pattern with probability one. Lemma 5.2.2 states an obvious implication of Assumption B.2: If the players recognize a PN pattern at some stage of the repeated game, then they must converge to that pattern. Similarly, Lemma 5.2.3 shows that if the players recognize a pattern that is not a PN pattern at some stage of the repeated game, then the pattern must be contradicted within the next  $L$  stages of the game.

LEMMA 5.2.1. *Assume there is some  $\epsilon > 0$  and an integer  $T$  such that for every  $t \geq T$ , for each  $z_t \in \mathcal{Z}_t$  satisfying  $\mathcal{P}_{(\mu, \phi)}(Z_t = z_t) > 0$ ,  $\mathcal{P}_{(\mu, \phi)}(\text{convergence to a PN pattern} \mid z_t) \geq \epsilon$ . Then,  $(\mu, \phi)$  converges a.s. to a PN pattern.*

*Proof.* Assume, by way of contradiction, that

(\*)  $\exists \delta > 0, \exists z_t$  such that  $\mathcal{P}_{(\mu, \phi)}(Z_t = z_t) > 0$ , and  $\mathcal{P}_{(\mu, \phi)}(\text{convergence to a PN pattern} \mid z_t) \leq 1 - \delta$ .

Note that (by standard arguments) the event  $\Lambda =$  “convergence to a PN pattern” is in the  $\sigma$  field generated by the  $s_t$ ’s, and thus is  $\mathcal{F}$  measurable. By the Paul Levy zero-one law (Chung, p. 341) the probability of  $\Lambda$  conditional on  $z_{t'}$  approaches the indicator function of  $\Lambda$  as  $t'$  approaches infinity. Since, by (\*),  $\Lambda$  has a probability less than 1 given some history  $z_t$ , there exist continuations of  $z_t$  on which the conditional probability of  $\Lambda$  can be made arbitrarily small. That is, for every  $\delta' > 0, \exists z_{t'}$  such that  $\mathcal{P}_{(\mu, \phi)}(\text{convergence to a PN pattern} \mid z_{t'}) \leq \delta'$ .

Choose  $\delta' < \epsilon$  to get a contradiction to the supposition. ■

LEMMA 5.2.2. *Let  $z_t$  be any history such that  $\psi(z_t) = p$  for some basic PN pattern  $p$ . Then,  $\mathcal{P}_{(\mu, \phi)}(\text{convergence to } p \mid z_t) = 1$ .*

*Proof.* The proof follows directly from successive applications of Assumption B.2.

First note that since  $p$  is a basic PN pattern, B.2 implies that (in a generic game)  $\mathcal{P}_{(\mu, \phi)}(s_t = \text{BR}(\text{PRO}_{l(p)}(p))) = \text{PRO}_{l(p)}(p) | z_t = 1$ .

Let  $z_{t+1}$  be the corresponding extended history  $z_t$ ,  $\text{PRO}_{l(p)}(p)$ , and observe that the sufficient conditions for recognizing the pattern  $\text{SH}_1(p)$  hold in  $z_{t+1}$  (this follows from the fact that the sufficient conditions hold in  $z_t$ , from the fact that realized play at date  $t$  conforms to the pattern's prediction and from the assumption that  $T_p = T_{\text{SH}_1(p)}$ ). This implies that  $\mathcal{P}_{(\mu, \phi)}(s_{t+1} = \text{BR}(\text{PRO}_{l(p)}(\text{SH}_1(p)))) = \text{PRO}_{l(p)}(\text{SH}_1(p)) | z_{t+1} = 1$ .

Repeating this argument a large enough number of times we may conclude that  $\mathcal{P}_{(\mu, \phi)}(p$  has appeared  $k$  times successively at the dates  $t + l(p), t + 2l(p), \dots, t + kl(p) | z_t) = 1$ .

Since this equality holds for every natural  $k$  the result follows. ■

**LEMMA 5.2.3.** *Let  $z_t$  be any history such that  $\psi(z_t) = p$  for some basic pattern  $p$  that is not a PN pattern. Then,  $\mathcal{P}_{(\mu, \phi)}(p$  is contradicted within  $L$  periods of the repeated play  $| z_t) = 1$ .*

*Proof.* The proof is a straightforward modification of the Proof of Lemma 5.2.2:

If the last element of the pattern  $p$  is not a PN (i.e.,  $\text{PRO}_{l(p)}(p) \notin \text{PN}(\Gamma)$ ) then the pattern must be contradicted at the next stage of the game (this follows immediately from Assumption B.2).

If, however, the last element of  $p$  is a PN, then realized play at date  $t$  must conform with the pattern's prediction (as in Lemma 5.2.2), but then—since  $T_p = T_{\text{SH}_1(p)}$ —the players must recognize the pattern  $\text{SH}_1(p)$  at the corresponding extended history  $z_{t+1}$ . If the last element of  $\text{SH}_1(p)$  is not a PN, then that pattern must be contradicted at date  $t + 1$ . If the last element of  $\text{SH}_1(p)$  is a PN, then realized play at date  $t + 1$  will conform to the pattern's prediction and the players must recognize the pattern  $\text{SH}_2(p)$  at the corresponding extended history  $z_{t+2} \dots$

Since  $p$  is not a PN pattern, it must contain at least one element that is not a PN. Since  $l(p) \leq L$ , the pattern  $p$  must be contradicted within  $L$  repetitions of the process described above. ■

The Proof of Proposition 5.2 also relies on the fact that under the assumptions of the model, any history of the repeated game that ends with pattern-contradiction, has an extension that converges to a (singleton) PN pattern. We present this result separately in the following lemma:

**LEMMA 5.2.4.**  $\exists \delta > 0$  such that whenever  $z_t$  satisfies  $\psi(z_{t-1}) = p$  for some basic pattern  $p$  and  $\text{PRO}_{t-1}(z_t) \neq \text{PRO}_{l(p)}(p)$  (i.e., the pattern  $p$  has been contradicted at date  $t - 1$ ),  $\mathcal{P}_{(\mu, \phi)}(\text{convergence to a PN pattern} | z_t) \geq \delta$ .

*Proof.* We prove the lemma by extending  $z_t$  into a history  $z_{t+k(z_t)}$  such that:

(1)  $\psi(z_{t+k(z_t)}) = p$  for some singleton PN pattern  $p$ .

(2) The extension of  $z_t$  into  $z_{t+k(z_t)}$  is compatible with the assumptions of the model in the sense that  $\mathcal{P}_{(\mu, \phi)}(z_{t+k(z_t)} | z_t) > \epsilon^{k(z_t)}$  where  $\epsilon$  is as defined in Assumptions A.1 and A.2.

Moreover, we show that  $k(z_t)$  is uniformly bounded above across histories by some finite  $K$ . It then follows from (1) above and from Lemma 5.2.2 that  $\mathcal{P}_{(\mu, \phi)}(\text{convergence to a PN pattern} | z_t) \geq \mathcal{P}_{(\mu, \phi)}(z_{t+k(z_t)} | z_t) \geq \epsilon^{k(z_t)} \geq \epsilon^K$ , and the lemma holds.

In extending  $z_t$  we use a collection of finite sequences (one sequence for each pure strategy profile in  $\Gamma$ ),  $\langle L_s \rangle_{s \in S(\Gamma)}$ , defined as follows:

If  $s \in \text{PN}(\Gamma)$ , then  $L_s$  is a sequence of  $T_s - 1$  repetitions of  $s$  (where  $T_s$  is the bound defined in Assumption B.2 for the singleton pattern  $s$ ).

If  $s \notin \text{PN}(\Gamma)$ , then recall that—by Lemma 4.1—there is a pair-wise disjoint no repetitions C sequence leading from  $s$  to a PN strategy, say  $s^*(s)$ , in no more than COMP steps. In this case we let  $L_s$  be the extension of that C sequence in which the PN strategy  $s^*(s)$  is repeated  $T_{s^*(s)} - 1$  times exactly at the end of the sequence; e.g., if the C sequence leading from  $s$  to  $s^*(s)$  is  $s, s_1, \dots, s_{n-1}, s_n = s^*$ , and  $T_{s^*} = 5$ , then  $L_s = s, s_1, \dots, s_{n-1}, s^*, s^*, s^*, s^*$ .

Note that every history that ends with such a sequence  $L_s$  satisfies the sufficient conditions for convergence to a singleton PN pattern. Note further that the length of the sequences  $L_s$  is (uniformly) bounded above by  $K = \text{COMP} + \max_{s^* \in \text{PN}(\Gamma)} T_{s^*} - 1$ .

Let  $s$  denote the last element in  $z_t$  and let  $z_{t+k(z_t)}$  be the extension of  $z_t$  generated by appending the sequence  $L_s$  to  $z_t$ ; i.e.,  $s_t$  will be the second element in  $L_s$ ,  $s_{t+1}$  will be the third element in that sequence and so on... Let  $s^*$  denote the PN strategy to which the extended sequence converges. Note that the process of appending the sequence  $L_s$  to  $z_t$  is “compatible” with  $\epsilon$  compatibility with BR dynamics and  $\epsilon$  compatibility with sophisticated BR dynamics (in the sense of Proposition 4). Thus, to complete the proof of the lemma it is enough to show that the process of extending  $z_t$  to  $z_{t+k(z_t)}$  cannot be disrupted by “unwarranted” pattern recognition; i.e., it cannot be the case that  $\psi(z_{t+j}) = \tilde{p}$  for some nonempty basic pattern  $\tilde{p} \neq s^*$ , for some index  $j \leq k(z_t)$ .

So assume by way of contradiction that  $\psi(z_{t+j}) = \tilde{p} \neq s^*$  for some basic pattern  $\tilde{p}$ , for some index  $j \leq k(z_t)$ . From the definition of the sequences  $\langle L_s \rangle_{s \in S}$  and the condition  $T_p \cdot l(p) \geq 3L$  it follows directly that  $\tilde{p}$  cannot be a singleton pattern (different from  $s^*$ ). Thus, there must be a pair  $s', s''$  where  $s' \neq s''$  that appears in the pattern  $\tilde{p}$ . From the definition of the

sequences  $L_s$ , it then follows that  $j < l(\tilde{p})$ . Since  $T_{\tilde{p}} \cdot l(\tilde{p}) \geq 3L$ , we conclude that the successive repetitions of  $\tilde{p}$  must cover the last  $2L$  elements of the history  $z_{t-1}$ .

But then both patterns  $p$  (the one contradicted at date  $t - 1$ ) and  $\tilde{p}$  cover the last  $2L$  elements of the history  $z_{t-1}$ . From the uniqueness proof in Section 5.2 (after the statement of Assumption B.2) it follows that  $\tilde{p} = p$ . Since  $p$  has been contradicted at date  $t - 1$ , we finally conclude that  $\psi(z_{t+j}) = \tilde{p} \neq s^*$  is impossible. ■

With these four lemmas the Proof of Proposition 5.2 is straightforward:

Consider an arbitrary history  $z_t$  ending with the strategy profile  $s$ . Let  $z_{t+k(z_t)}$  denote the extension of  $z_t$  ending with the sequence  $L_s$  (as defined in the Proof of Lemma 5.2.4).

Let  $\epsilon' = \delta \cdot \epsilon^K$  (where  $\delta$  is as in Lemma 5.2.4, the  $\epsilon$  is from Assumptions A.1 and A.2, and  $K$  is the complexity bound defined in the proof of Lemma 5.2.4). We now show that

$$(*) \mathcal{P}_{(\mu, \phi)}(\text{convergence to a PN pattern} \mid z_t) \geq \epsilon'.$$

If  $\mathcal{P}_{(\mu, \phi)}(z_{t+k(z_t)} \mid z_t) > \epsilon^{k(z_t)}$ , then  $(*)$  holds since  $k(z_t) \leq K$ . If not, it must be the case that the process of concatenating the sequence  $L_s$  to  $z_t$  was interrupted by pattern recognition; i.e.,  $\exists j \in \{0, 1, \dots, k(z_t) - 1\}$  and a pattern  $p$  such that  $\psi(z_{t+j}) = p \neq \theta$ .

Letting  $j$  denote the first date of such an interruption, we have

$$(**) \mathcal{P}_{(\mu, \phi)}(z_{t+j} \mid z_t) \geq \epsilon^K.$$

If  $p$  is a PN pattern, then  $(**)$  and Lemma 5.2.2 together imply that  $\mathcal{P}_{(\mu, \phi)}(\text{convergence to a PN pattern} \mid z_t) \geq \epsilon^K > \epsilon'$  and  $(*)$  holds.

If  $p$  is not a PN pattern, then Lemma 5.2.3 implies that  $p$  must be contradicted within  $L$  periods. Applying Lemma 5.2.4 to the corresponding extended history we get that  $\mathcal{P}_{(\mu, \phi)}(\text{convergence to a PN pattern} \mid z_t) > \delta \cdot \epsilon^K$ , which completes the proof of  $(*)$ . The proposition then follows immediately from Lemma 5.2.1. ■

*Proof of Proposition 5.3.* Fix a generic, simple, strategic-form game  $\Gamma$ . Let  $(\mu, \phi)$  be a confused learning model with dynamic pattern recognition on  $\Gamma$ . The proof of the proposition is a straightforward generalization of the Proof of Proposition 5.2.<sup>19</sup> First we have to generalize Lemmas 5.2.2, 5.2.3, and 5.2.4. We call the modified lemmas 5.3.2, 5.3.3, and 5.3.4 correspondingly.

<sup>19</sup> We have presented the proof of the stationary degenerated case first in order to simplify the exposition.

LEMMA 5.3.2.  $\exists \xi > 0$  such that whenever  $z_t$  is a history such that  $\psi(z_t) = p$  for some basic PN pattern  $p$ ,  $\mathcal{P}_{(\mu, \phi)}(\text{convergence to } p | z_t) \geq \xi$ .

*Proof.* The proof follows directly from successive applications of Assumptions B.4 and B.6, and from Lemma 5.2.2.

First note that since  $p$  is a basic PN pattern, Assumption B.4 implies that (in a generic game)  $\mathcal{P}_{(\mu, \phi)}(s_t = \text{BR}(\text{PRO}_{l(p)}(p)) = \text{PRO}_{l(p)}(p) | z_t) \geq \delta^{\text{card}(I)}$ .

Let  $z_{t+1}$  be the corresponding extended history  $z_t, \text{PRO}_{l(p)}(p)$ , and observe that by Assumption B.6,  $\psi(z_{t+1}) = \text{SH}_1(p)$ . Assumption B.4 then implies that

$$\begin{aligned} \mathcal{P}_{(\mu, \phi)}(s_{t+1} = \text{BR}(\text{PRO}_{l(p)}(\text{SH}_1(p)))) &= \text{PRO}_{l(p)}(\text{SH}_1(p)) | z_{t+1}) \\ &\geq \delta^{\text{card}(I)}. \end{aligned}$$

Repeating these argument a large enough number of times we may conclude that  $\mathcal{P}_{(\mu, \phi)}(p$  has appeared  $T_p$  times successively at the dates  $t + l(p), t + 2l(p), \dots, t + T_p \cdot l(p) | z_t) \geq \delta^{(T_p \cdot l(p) + 1) \cdot \text{card}(I)}$ .

Let  $z_{\tilde{t}}$  be the corresponding extension of  $z_t$  that ends with  $T_p$  successive repetitions of the pattern  $p$ . Note that the sufficient conditions for convergence to  $p$  hold in  $z_{\tilde{t}}$ . By Lemma 5.2.2, the probability of convergence to  $p$  given  $z_{\tilde{t}}$  is one. Thus, the probability of convergence to a PN pattern given  $z_t$  is at least  $\delta^{(T_p \cdot l(p) + 1) \cdot \text{card}(I)}$ . Since, by Assumption B.1,  $T_p \cdot l(p)$  is bounded above by some integer  $D$ , we may set  $\xi = \delta^{(D+1) \cdot \text{card}(I)}$  to complete the proof. ■

LEMMA 5.3.3.  $\exists \rho > 0$  such that whenever  $z_t$  is a history such that  $\psi(z_t) = p$  for some basic pattern  $p$  that is not a PN pattern,  $\mathcal{P}_{(\mu, \phi)}(p$  is contradicted within  $L$  periods of the repeated play  $| z_t) \geq \rho$ .

*Proof.* The proof is a straightforward modification of the Proof of Lemma 5.2.3:

If the last element of  $p$  is not a PN (i.e.,  $\text{PRO}_{l(p)}(p) \notin \text{PN}(\Gamma)$ ) then Assumption B.4 implies that with probability at least  $\delta^{\text{card}(I)}$ , the pattern will be contradicted at the next stage of the game.

If, however, the last element of  $p$  is a PN, then the assumption implies that with probability at least  $\delta^{\text{card}(I)}$ , realized play at date  $t$  will conform with the pattern's prediction (as in Lemma 5.3.2), but then Assumption B.6 and B.4 together imply that with probability at least  $\delta^{\text{card}(I)}$ , the players will recognize the pattern  $\text{SH}_1(p)$  at the corresponding extended history  $z_{t+1}$ . If the last element of  $\text{SH}_1(p)$  is not a PN, the pattern will be contradicted at that date. If the last element of  $\text{SH}_1(p)$  is a PN, then realized play at date  $t + 1$  will conform to the pattern's prediction and, by one more application of Assumptions B.6 and B.4, there exists a positive

probability  $\delta^{\text{card}(I)}$  that the players will recognize the pattern  $\text{SH}_2(p)$  at the corresponding extended history  $z_{t+2} \dots$ . Since  $p$  must have at least one element that is not a PN, the probability that it will be contradicted within  $L$  repetitions of the process described above is at least  $\delta^{L \cdot \text{card}(I)}$ . ■

The statement of Lemma 5.3.4 is identical to that made in Lemma 5.2.4 and we will not repeat it. The proof however is different since now the conditions under which the model can get into the pattern recognition mode are weaker than before. However, the additional caution assumption B.5 implies that the basic construction used to prove Lemma 5.2.4, can be applied in the generalized model as well.

*Proof of Lemma 5.3.4.* Let  $z_{t+k(z_t)}$  be the extension of  $z_t$  generated by appending the sequence  $L_s$  to  $z_t$  as defined in the Proof of Lemma 5.2.4. As in the proof of that lemma, it is enough to show that the process of extending  $z_t$  to  $z_{t+k(z_t)}$  cannot be disrupted by “unwarranted” pattern recognition; i.e., it cannot be the case that  $\psi(z_{t+j}) = \tilde{p}$  for some nonempty basic pattern  $\tilde{p} \neq s^*$ , for some index  $j \leq k(z_t)$ .

So assume by way of contradiction that  $\psi(z_{t+j}) = \tilde{p} \neq s^*$  for some basic pattern  $\tilde{p}$ , for some index  $j \leq k(z_t)$ . Assumption B.5 implies that  $j \geq l(\tilde{p})$  and this together with the definition of the sequences  $\langle L_s \rangle_{s \in S}$  implies that there is a pair  $s', s''$  that appears in the pattern  $\tilde{p}$  and appears at least twice in the sequence  $L_s$ . From the definition of the sequences  $\langle L_s \rangle_{s \in S}$  it follows that  $\tilde{p} = s^*$ , a contradiction. ■

Proposition 5.3 follows immediately from the three lemmas proved above and from Lemma 4.2.1. The basic argument is identical to the one used to prove Proposition 5.2 from Lemmas 5.2.1–5.2.4. We will not repeat it.

*Proof of Proposition 6.* First note that in order to prove the proposition it is enough to prove that there is some history  $z_t$  such that  $\mathcal{P}_{(\mu, \phi)}(\bigcap_{p \in \tilde{\mathcal{E}}(\sigma^*)} \bigcap_{n \in \mathcal{N}} P$  has appeared  $n$  times successively  $|z_t) > 0$ .

The proof of this assertion follows directly from the fact that (by standard arguments) the event “ $\bigcap_{p \in \tilde{\mathcal{E}}(\sigma^*)} \bigcap_{n \in \mathcal{N}} P$  has appeared  $n$  times successively” is  $\mathcal{F}$  measurable, and from Paul-Levy’s zero or one law (Chung, 1974) as used in the Proof of Lemma 5.1.

Let  $\sigma^*$  be a mixed strategy profile that is locally stable with respect to  $(\mu, \phi)$ .

Let  $\delta = 1/2 \cdot (\min_{s \in \text{supp}(\sigma^*)} \sigma^*(s))$ .

Choose (arbitrarily)  $\epsilon > 0$ , and recall that local stability of  $\sigma^*$  implies that there is some history  $z_t$  such that  $\mathcal{P}_{(\mu, \phi)}(\lim_{t' \rightarrow \infty} \phi_{t'}(z_{t'}) = \sigma^* | z_t) > 1 - \epsilon$ .

Using  $\mathcal{P}_{(\mu, \phi)}^{z_t}(\cdot)$  to denote the conditional probability  $\mathcal{P}_{(\mu, \phi)}(\cdot | z_t)$  we get that  $1 - \epsilon < \mathcal{P}_{(\mu, \phi)}^{z_t}(\bigcup_{t' \geq t} \bigcap_{t'' \geq t'} \bigcap_{s \in \text{supp}(\sigma^*)} |\phi_{t''}(z_{t''})(s) - \sigma^*(s)| \leq \delta) \leq \sum_{t' \geq t} \mathcal{P}_{(\mu, \phi)}^{z_{t'}}(\bigcap_{t'' \geq t'} \bigcap_{s \in \text{supp}(\sigma^*)} |\phi_{t''}(z_{t''})(s) - \sigma^*(s)| \leq \delta)$ .

Thus,  $\exists \tilde{t} \geq t$ , and there is a continuation of  $z_t$ , say  $z_{\tilde{t}}$ , satisfying  $\mathcal{P}_{(\mu, \phi)}(z_{\tilde{t}}) > 0$  such that (\*)  $\mathcal{P}_{(\mu, \phi)}^{z_{\tilde{t}}}(\bigcap_{t'' \geq \tilde{t}} \bigcap_{s \in \text{supp}(\sigma^*)} |\phi_{t''}(z_{t''})(s) - \sigma^*(s)| \leq \delta) > 0$ .

To complete the proof we show that  $\mathcal{P}_{(\mu, \phi)}^{z_{\tilde{t}}}(\bigcap_{p \in \tilde{\mathcal{E}}(\sigma^*)} \bigcap_{n \in \mathbb{N}} p$  has appeared  $n$  times successively)  $> 0$ .

Let  $\Lambda = \{z \in \mathcal{Z}_\infty \mid \bigcap_{t'' \geq \tilde{t}} \bigcap_{s \in \text{supp}(\sigma^*)} |\phi_{t''}(z_{t''})(s) - \sigma^*(s)| \leq \delta\}$ .

Note that (by standard arguments)  $\Lambda$  is a measurable event, and that by inequality (\*),  $\mathcal{P}_{(\mu, \phi)}^{z_{\tilde{t}}}(\Lambda) > 0$ .

Let  $p$  be a strategic pattern in  $\tilde{\mathcal{E}}(\sigma^*)$ , say  $p = \tilde{s}_1, \tilde{s}_2, \dots, \tilde{s}_m$ .

We now prove that for any natural number  $n$ ,  $\mathcal{P}_{(\mu, \phi)}^\Lambda(p \text{ has appeared } n \text{ times successively}) = 1$ .<sup>20</sup>

Fix a natural number  $n$ , and for every  $k = 1, 2, \dots$  let  $E_k$  be the event “ $p$  has appeared  $n$  times successively starting at date  $\tilde{t} + nm(k - 1)$ ”; i.e.,  $E_k = \bigcap_{l=0}^{n-1} \bigcap_{j=0}^{m-1} s_{\tilde{t} + nm(k-1) + ml + j} = \tilde{s}_{j+1}$ .

Note that by definition of  $\Lambda$  and  $\delta$ ,  $\mathcal{P}_{(\mu, \phi)}^\Lambda(E_k) \geq [(\sigma^*(\tilde{s}_1) - \delta) \cdot (\sigma^*(\tilde{s}_2) - \delta) \cdot \dots \cdot (\sigma^*(\tilde{s}_m) - \delta)]^n \geq \delta^{nm} > 0$  for every  $k = 1, 2, \dots$ .

We now claim that for any  $t \geq \tilde{t}$ ,  $\mathcal{P}_{(\mu, \phi)}^\Lambda(\bigcap_{k \geq t} E_k^C) = 0$  (where  $E_k^C$  denotes the complementary event to  $E_k$ ); i.e., by definition of  $\Lambda$  and by inequality (\*) it follows that for every  $j \geq t$ ,  $\mathcal{P}_{(\mu, \phi)}^\Lambda(\bigcap_{k=t}^{t+j} E_k^C) \leq \prod_{k=t}^{t+j} (1 - \tilde{\sigma})$ , where  $\tilde{\sigma} = \delta^{n \cdot m}$ .

Using the inequality  $1 - x \leq e^{-x}$  we get that  $\mathcal{P}_{(\mu, \phi)}^\Lambda(\bigcap_{k=t}^{t+j} E_k^C) \leq e^{-j \cdot \tilde{\sigma}}$ .

Since  $\tilde{\sigma} > 0$ , we finally have

$$\mathcal{P}_{(\mu, \phi)}^\Lambda(\bigcap_{k \geq t} E_k^C) = \lim_{j \rightarrow \infty} \mathcal{P}_{(\mu, \phi)}^\Lambda(\bigcap_{k=t}^{t+j} E_k^C) = 0.$$

Thus,  $\mathcal{P}_{(\mu, \phi)}^\Lambda(\bigcup_{t \geq \tilde{t}} \bigcap_{k \geq t} E_k^C) = 0$ , and  $\mathcal{P}_{(\mu, \phi)}^\Lambda(\bigcap_{t \geq \tilde{t}} \bigcup_{k \geq t} E_k) = 1$ ; i.e.,  $\mathcal{P}_{(\mu, \phi)}^\Lambda(E_k \text{ i.o.}) = 1$ . In particular,  $\mathcal{P}_{(\mu, \phi)}^\Lambda(p \text{ has appeared } n \text{ times successively}) \geq \mathcal{P}_{(\mu, \phi)}^\Lambda(E_k \text{ for some } k) = 1$ .

Since this argument holds for every natural  $n$ , and since there are countably many strategic patterns,  $\mathcal{P}_{(\mu, \phi)}^\Lambda(\bigcap_{p \in \tilde{\mathcal{E}}(\sigma^*)} \bigcap_{n \in \mathbb{N}} p$  has appeared  $n$  times successively)  $= 1$ .

Since  $\mathcal{P}_{(\mu, \phi)}^{z_{\tilde{t}}}(\Lambda) > 0$ , it follows that  $\mathcal{P}_{(\mu, \phi)}^{z_{\tilde{t}}}(\bigcap_{p \in \tilde{\mathcal{E}}(\sigma^*)} \bigcap_{n \in \mathbb{N}} p$  has appeared  $n$  times successively)  $> 0$ , so that  $(\mu, \phi)$  is  $\tilde{\mathcal{E}}(\sigma^*)$  incompatible with pattern recognition. ■

<sup>20</sup> The proof of this assertion is a generalization of the proof of Borel–Cantelli Lemma 2.



*Proof of Proposition 7.* The proof is a direct extension of the Proof of Proposition 5.3. We present the “new elements” in the proof in three simple lemmas (7.1–7.3). First, we claim that the conditions stated in the  $*$ -adaptiveness assumption are sufficient for convergence to a MCS; i.e., whenever a minimal closed subgame  $E$  has been played at for  $T_E$  periods successively, the players must keep playing strategies in  $E$  forever.

**LEMMA 7.1.** (Sufficient Conditions for Convergence to a MCS). *Let  $E$  be a MCS in a strategic-form game  $\Gamma$ . Let  $z_t$  be any history such that  $\text{PRO}_t^i(z_t) \in E \quad \forall t' = t - T_E, t - T_E + 1, \dots, t - 1$ . Then,  $\mathcal{P}_{(\mu, \phi)}(\text{convergence to } E \mid z_t) = 1$ .*

The proof follows immediately from successive applications of Assumption C (just as the proof of Lemma 5.2.2 follows from Assumption B.2).

The next lemma says that whenever a C sequence reaches a BR-closed subgame of  $\Gamma$  it can never “leave” this subgame.

**LEMMA 7.2.** *Let  $\Gamma$  be a strategic-form game. Let  $s_1, s_2, \dots, s_n$  be a C sequence in  $\Gamma$ . If  $\Gamma^*$  is a MCS of  $\Gamma$ , and  $s_m \in S^*(\Gamma^*)$  for some  $m < n$ , then  $s_l \in S^*(\Gamma^*) \quad \forall l = m + 1, \dots, n$ .*

The proof is straightforward: Assume w.l.g. that  $m$  is the first date at which some strategy in  $S^*(\Gamma^*)$  has been played. It then follows from the definition of C reachability and from the assumption that  $\Gamma^*$  is a MCS, that  $s_{m+1} = \text{BR}(s_m) \in S^*(\Gamma^*)$ . Proceeding by induction assume that  $s_{m+k'} \in S^*(\Gamma^*)$  for every  $k' = 0, 1, \dots, k$ , for some  $k < n - m$ . It again follows immediately from the definition of C reachability that for each player  $i$  there is some strategy  $s_l(i)$  in  $\{s_m, s_{m+1}, \dots, s_k\}$ , such that  $s_{k+1}^i \in \text{BR}^i(s_l(i))$ . From the induction hypothesis and the assumption that  $\Gamma^*$  is a MCS it then follows that  $s_{k+1}^i$  is in the subgame  $\Gamma^*$  (in the sense that there exists a strategy, say  $\tilde{s}$ , in that subgame such that  $\tilde{s}^i = s_{k+1}^i$ ), so that  $s_{k+1} \in S^*(\Gamma^*)$ . ■

With Lemmas 7.2 in stock we are ready to prove the following “reachability lemma”:

**LEMMA 7.3.** *Let  $\Gamma$  be a generic, finite, strategic-form game. For every  $s \in S(\Gamma)$  there is a pair-wise disjoint, no repetitions C sequence, say  $s, s_1, \dots, s_n$ , such that one of the following two conditions hold: (1)  $s_n$  is contained in some MCS with no PN, or (2)  $s_n$  is a PN.*

The proof is an immediate extension of the Proof of Lemma 4.1. In Lemma 4.1 we have assumed that the game  $\Gamma$  is simple. To prove Lemma 7.3, remove that assumption and consider the maximal C sequence  $s_0, s_1, \dots, s_n$  as defined in the Proof of Lemma 4.1. Assume that  $s_n$  is not a

PN and repeat the arguments given in the Proof of Lemma 4.1 with the following small change. When considering the case where SGA is closed distinguish between two subcases:

(A) SGA contains a PN equilibrium, say  $s^*$ , in which case argue (as in the original proof) that  $z_0, s^*$  is a pair-wise disjoint, no repetitions, C extension of  $z_0$ , to get a contradiction to the maximality of  $z_0$ .

(B) SGA does not contain a PN equilibrium, in which case it follows from Lemma 7.2 that  $s_n$  is an element of a MCS with no PN of  $\Gamma$ .

Thus, if  $s_n$  is not a PN equilibrium of  $\Gamma$ , then it must be an element of a "MCS with no PN," as stated in Lemma 7.3.

We are now ready to prove Proposition 7. The basic idea of the proof is the same as that underlying the Proof of Proposition 5.3: We show that every history  $z_t$  can be extended to a history  $z_{t+k(z_t)}$  such that the extended history either satisfies the sufficient conditions for convergence to a PN pattern or the sufficient conditions for convergence to a MCS with no PN. We further show that  $k(z_t)$  is uniformly bounded across  $z_t$ .

In extending  $z_t$  we used a collection of sequences,  $\langle L_s \rangle_{s \in S(\Gamma)}$ , defined as follows:

—For every  $s \in S(\Gamma)$  for which there is a pair-wise disjoint, no repetitions, C sequence leading from  $s$  to a PN, say  $s^*(s)$ , let  $L_s$  be the extension of that C sequence where  $s^*(s)$  is repeated  $T_{s^*(s)} - 1$  times exactly at the end of that sequence (just as defined in the Proof of Proposition 5.2).

—For each other strategy  $s \in S(\Gamma)$ , by Lemma 7.3, there exists a pair-wise disjoint, no repetitions C sequence, say  $s, s_1, \dots, s_n$ , leading from  $s$  to a MCS with no PN. Let  $L_s$  be an extension of that sequence, say  $s, s_1, \dots, s_n, s_{n+1}, \dots, s_{n+\hat{T}}$ , where (1)  $\hat{T} = \max_{E \in S(\Gamma)} T_E$ , and (2)  $s_{m+1} = BR(s_m)$  for every  $m = n, \dots, n + \hat{T} - 1$ . (That is, we copy the BR cycle starting from  $s_n$  successively until we get a history of length  $n + \hat{T}$ .)

A key element in the Proof of the Proposition is the following version of Lemma 5.3.4:<sup>21</sup>

LEMMA 7.4.  $\exists \delta > 0$  such that whenever  $z_t$  satisfies  $\psi(z_{t-1}) = p$  for some basic pattern  $p$  and  $PRO_{t-1}(z_t) \neq PRO_{l(p)}(p)$  (i.e., the pattern  $p$  has been contradicted at date  $t - 1$ ),  $\mathcal{P}_{(\mu, \phi)}(\text{Convergence to a PN pattern or convergence to MCS with no PN} | z_t) \geq \delta$ .

<sup>21</sup> Since the arguments from this point on are very similar to those used in previous proofs, we outline the basic ideas briefly. The details can be found in Sonsino (1995).

The proof is very similar to the Proof of Lemma 5.3.4: Let  $z_{t+k(z_t)}$  be the extension of  $z_t$  generated by appending the sequence  $L_s$  to  $z_t$ . As in the Proof of Lemma 5.3.4, it is enough to show that (\*) the process of extending  $z_t$  to  $z_{t+k(z_t)}$  cannot be disrupted by “unwarranted” pattern recognition.

We distinguish between two cases: (1) If  $L_s$  converges to a PN, then the proof of (\*) is just as in Lemma 5.3.4. (2) If however  $L_s$  converges to a MCS with no PN, say  $\Gamma^*$ , then—since  $L_s$  may end with successive repetitions of some BR cycle in  $\Gamma^*$ —the process to appending  $L_s$  to  $z_t$  might be disrupted by (\*\*\*) recognizing a pattern that consists only of strategies in  $\Gamma^*$ . But since  $\Gamma^*$  is BR closed, such pattern recognition will not kick the players out of the subgame. From Lemma 7.2 and the fact that  $\Gamma^*$  is BR closed, it follows that we may keep extending  $z_t$  until its last  $\hat{T}$  elements are in  $\Gamma^*$ . This shows that pattern recognition of type (\*\*\*) does not constitute a problem, and we only have to worry about the possibility that the process of extending  $z_t$  to  $z_{t+k(z_t)}$  was disrupted by recognizing some pattern that includes at least one strategy  $\bar{s}$  that is not in the subgame  $\Gamma^*$ . This caution assumption B.5, however, implies (as in the Proof of Lemma 5.3.4) that this cannot happen and Lemma 7.4 holds.

Proposition 7 follows directly from the last lemma and from the Paul-Levy 0–1 law as used in the Proof of Lemma 5.2.1. ■

## APPENDIX 2: A MODEL WITH TWO MODES OF PATTERN RECOGNITION

In this appendix we briefly outline a variation on the model of Section 5 in which the agents can also recognize (with probability one) strategic patterns that appear successively with interruptions.

The modified model will distinguish between two different modes of pattern recognition, a strong pattern recognition mode and a weak pattern recognition mode. Formally, we redefine the pattern recognition frame  $\psi$  so that  $\psi: \bigcup_{t=1}^{\infty} \mathcal{Z}_t \rightarrow (\bigcup_{n \in \mathcal{N}} SP_n \times \{W, S\}) \cup \theta$ . As before,  $\psi(z_t) = \theta$  denotes the case where the agents are in the confused learning mode given  $z_t$ ;  $\psi(z_t) = p, S$  denotes the case where the agents are in the strong pattern recognition mode with respect to  $p$  given the history  $z_t$ ;  $\psi(z_t) = p, W$  denotes the case where the agents are in the weak pattern recognition mode with respect to  $p$  given that history.

The boundedness assumption B.1 and the coordination assumption B.4 will now apply to both types of pattern recognition; i.e., we will restate these assumption so that they apply to any history satisfying  $\psi(z_t) = p$  for some basic pattern  $p$ . The compatibility with pattern recognition

assumption B.2 can be restricted to the strong pattern recognition mode and we prefer not to extend it to cases of weak pattern recognition since the notion of  $n$  successive repetitions with interruptions is general enough to accommodate many different histories including some in which the interruptions between the successive repetitions of the pattern are very short, and others in which these interruptions are quite long and of a confusing nature; we thus do not wish to impose a uniform bound on the number of successive repetitions with interruptions after which the agents recognize patterns with probability 1.

The necessary conditions for strong pattern recognition will be the same as those stated in Assumption B.3 of Section 5; i.e., we keep Assumption B.3 but restrict it to the cases where  $\psi(z_t) = p, S$  for some basic pattern  $p$ . The necessary conditions for weak pattern recognition will be defined accordingly but they will refer to the weaker notion of successive appearances with interruptions.

*Assumption B.3'.* For every history  $z_t$ , for every basic pattern  $p$  with  $l(p) \geq 3$ , if  $\psi(z_t) = p, W$ , then the following three conditions hold:

- (a)  $p$  has appeared successively at least once in  $z_t$ ,
- (b)  $\text{PRO}_{t-l(p)+1, t-1}(z_t) = \text{PRO}_{1, l(p)-1}(p)$ , and
- (c) the necessary conditions for  $\psi(z_t) = p, S$  don't hold.

The necessary conditions for being in the weak pattern recognition mode with respect to a basic pattern of length 2 are defined accordingly with the small change that  $p$  must have appeared successively at least twice in  $z_t$ .

The caution in pattern recognition assumption B.5 of the original model will now apply to both modes of pattern recognition, but we will add another caution assumption that deals with the case where patterns that appear successively "fade away". For example, consider the history  $z_t = \dots A, B, C, D, F, A, B, C, F, A, B, C, B, F, A, B$  where  $C$  is a pure strategy Nash equilibrium. Assuming that the agents weakly recognize the pattern  $A, B, C$  given the history  $z_t$  so that realized play at date  $t$  is  $C$ , it seems reasonable to assume that  $\psi(z_t, C) = \theta$ . Assumption B.5' says that a necessary condition for recognizing a new pattern  $p'$  at any future date  $t+k$  is that  $k > l(p')$ .

*Assumption B.5'.*  $\exists \bar{T}$  such that for every basic pattern  $p$ , for every  $t \geq \bar{T}$ , for every history  $z_t$  for which  $\psi(z_t) = p, W$ , for every extension of  $z_t$ , say  $z_{t+k}$ , satisfying (a)  $\psi(\text{PRO}_{1, t}(z_{t+k})) = \theta$ , and (b)  $\psi(z_{t+k}) = p'$ , for some basic pattern  $p'$ , it must be the case that  $k > l(p')$ .

The continuation assumption B.6 will now be restricted to the strong pattern recognition mode since, for example, if history is  $z_t = \dots A, B, D, A, B, C, A$  (so that the pattern  $A, B$  appears successively with interruptions) where  $B$  is a pure strategy Nash equilibrium, and the players recognize the pattern  $A, B$  at date  $t$ , so that realized play at date  $t$  is  $B$ , there is still no reason for the agents to recognize the pattern  $B, A$  in period  $t + 1$ . Yet, to have some control over the dynamics we impose the following additional assumption:

*Assumption B.6'*. For every basic pattern  $p$ , for every history  $z_t$  for which  $\psi(z_t) = p, W$ , if  $\psi(z_t, \text{PRO}_{I(p)}(p)) = \tilde{p}$ , for some basic pattern  $\tilde{p}$ , then it is either the case that  $\tilde{p}$  is a one period extension of  $p$  (i.e.,  $\tilde{p} = p, s$  for some strategy  $s$ ), or that  $\tilde{p}$  is a one period forward shift of  $p$ .

Note that, in particular, Assumption B.6' enables the players to weakly recognize long patterns before the necessary conditions for strong pattern recognition have been realized; e.g., the agents may recognize the pattern  $A, B, C, \dots, X, Y, Z$  given the history  $A, B, C, D, \dots, Z, A, B, C$ . Under the original formulation the players could have recognized this pattern with probability 1 only given a history that ends with  $A, B, C, D, \dots, Z, A, B, \dots, Y$ .

We claim that a myopic learning model with a two modes-pattern recognition frame that satisfies the modified assumptions A.1, A.2, B.1–B.6, B.3', B.5', and B.6', converges with probability one to a PN pattern when the stage game is simple. The proof is a straightforward extension of the original proof, and thus we omit the details.

## REFERENCES

- Billingsley, P. (1986). *Probability and Measure*, 2nd ed. New York: Wiley.
- Brown, G. W. (1951). "Iterative Solutions of Games by Fictitious Play." in *Activity Analysis of Production and Allocation*. New York: Wiley.
- Chung, K. L. (1974). *A Course in Probability Theory*, 2nd ed. New York: Academic Press.
- Cournot, A. (1838). *Recherches sur les Principes Mathématiques de la Théorie des Richesses*. Translated by N. T. Bacon (1960) as *Researches into the Mathematical Principles of the Theory of Wealth*. London: Hafner.
- Fudenberg, D., and Kreps, M. D. (1993). "Learning Mixed Equilibria." *Games Econ. Behav.* **5**, 320–367.
- Fudenberg, D., and Kreps, M. D. (1994). "Learning in Extensive Form Games, 2: Experimentation and Nash Equilibrium." *Games Econ. Behav.* in press.
- Fudenberg, D., and Kreps, M. D. (1995). "Learning in Extensive Form Games, 1: Self-Confirming Equilibrium." *Games Econ. Behav.* **8**, 20–55.
- Jordan, J. (1993). "Three Problems in Learning Mixed-Strategy Nash equilibria." *Games Econ. Behav.* **5**, 368–386.

- Kalai, E., and Lehrer, E. (1993). "Rational Learning Leads to Nash equilibrium." *Econometrica* **61**(5), 1019–1045.
- Kalai, E., and Lehrer, E. (1995). "Subjective Games and Equilibrium." *Games Econ. Behav.* **8**, 123–163.
- Kreps, M. D. (1990). *A Course in Microeconomic Theory*. Princeton, NJ: Princeton Univ. Press.
- Krishna, V. (1992). "Learning in Games with Strategic Complementarities." Mimeo. Harvard Business School.
- Laird, P. (1994). "Discrete Sequence Prediction and its Applications." Mimeo. NASA Ames Research Center, Moffett Field, CA.
- Laird, P., and Ronald, S. (1994). "Sequence Extrapolation." Mimeo. NASA Ames Research Center, Moffett Field, CA.
- Milgrom, P., and Roberts, J. (1991). "Adaptive and Sophisticated Learning in Repeated Normal Form Games." *Games Econ. Behav.* **3**, 82–100.
- Monderer, D. and Shapley, L. S. (1993). "Fictitious Play Property for Games with Identical Interests." Mimeo.
- Nash, J. F. (1950). "Equilibrium Points in  $n$ -person Games." *Proc. Nat. Acad. Sci. USA* **36**, 48–49.
- Rapoport, A., Melvin, J. G., and Gordon, D. G. (1976). *The  $2 \times 2$  Game*. Ann Arbor: Univ. of Michigan Press.
- Rubinstein, A., and Osborne, J. M. (1993). "A Course in Game Theory." Forthcoming.
- Sanchirico, C. (1993). "Strategic Intent and the Salience of Past Play." Mimeo. Department of Economics, Yale University.
- Simon, H. (1992). *Economics, Bounded Rationality, and the Cognitive Science*. Edward Elgar.
- Sonsino, D. (1995). *Learning to learn, Pattern Recognition, and Nash equilibria*. Ph.D dissertation, Stanford University.