GAMES AND ECONOMIC BEHAVIOR **11**, 64–89 (1995)

Fictitious Play by Cases

MARCO LICALZI*

Istituto di Matematica "E. Levi," Università di Parma, 43100 Parma, Italy

Received October 22, 1992

Suppose that two persons are engaged in playing a possibly different game at a sequence of dates and that at each date they decide how to play only on the basis of their experience in past games similar to the current one. We model this situation as a fictitious play algorithm for games with random payoffs, and we provide sufficient conditions for its almost sure convergence in the class of 2 × 2 games. © 1995 Academic Press, Inc.

1. INTRODUCTION

People often deal with new situations by relying on similar experiences. Law is administered by recourse to precedents, management is taught by case studies, behavior changes by limitation, philosophers and mathematicians reason by analogy, and almost everybody has used examples to explain a concept. In most of these cases, people are relying on a similar situation to shed light on a problem.

Game theory should be no exception. When confronted with a game which does not have "an obvious way to play," people look at previous experiences in similar games to learn or deduce a way to play. In most cases, actually, the "obvious way to play" arises from the consideration of similar situations. Thus, even when an established solution seems to be in place, the usage of similarities might help to explain how this has emerged.

* I am grateful to Pierpaolo Battigalli, Cristina Molinari, Roy Radner, Lin Zhou, and especially to David Kreps and an anonymous referee for helpful comments. I thank the Department of Economics, Tel Aviv University, for its hospitality while part of this research was being conducted. The financial assistance of Bocconi University and Stanford University is gratefully acknowledged.

0899-8256/95 $12.00

Despite their potential importance, however, similarities have not attracted much attention in economics or game theory. To the best of our knowledge, only four lines of research have so far made explicit use of some notion of similarity. Rubinstein (1988) and Aizpurua *et al.* (1990, 1993) are concerned with the study of preferences consistent with similarity judgements. Gilboa and Schmeidler (1992, 1993), instead, present a more explicit model of decision making based on similarities. Neither of these models has been applied to game theory yet. Cotter (1991) studies some sufficient conditions under which the information structures of economic agents engaged in playing games induce similar behaviors. Only Kreps (1990) and Fudenberg and Kreps (1992), however, have so far raised explicit questions on how to incorporate the usage of similarities into game theory.

This work is inspired by one of these questions: *How does fictitious play extend to situations where players try to extrapolate from past experiences in similar games?* (Fudenberg and Kreps, 1992, p. 25). More precisely, we have in mind the following situation. Two persons are engaged in playing a game at a sequence of dates. The game may change from date to date, but all games which can come up for play belong to a commonly known class **G**. Each player considers the games in **G** sufficiently similar to each other that his or her experience over each game from **G** played in the past is an (equally valuable) guide to how any game in this class should be played. In this sense, all games in **G** are equivalent "cases" of the same class. Therefore, to decide what to play, each player tries to assess what the opponent will do in the current game by using this experience to form a prediction. In this framework, we assume that both assessments and behavior rules conform with the spirit (and almost the letter) of fictitious play.

After briefly recalling the standard algorithm of fictitious play, Section 2 describes our model of fictitious play by "cases." Its most important property is that the games to be played are drawn randomly from **G**. This has two relevant consequences. First, since the laws of motion for the algorithm are stochastic, there are some substantial differences between fictitious play by cases and standard fictitious play: for instance, strict Nash equilibria are not necessarily absorbing. Second, our algorithm includes as a special case the model of a learning process proposed by Fudenberg and Kreps (1993) for an equilibrium in mixed strategies as the purification *à la* Harsanyi (1973) of a game of incomplete information.

It is well known that standard fictitious play converges for any 2×2 game (Miyasawa, 1961). Section 3 analyzes fictious play by cases in the class of 2×2 games and provides a set of sufficient conditions for its almost sure convergence. If we interpret fictitious play by cases as a learning process based on the use of similarities, this is a case where some sort of learning takes place. However, although fictitious play by cases may appear more realistic, it is not likely to solve the difficulties associated with the interpretation of fictitious play as a learning process. For instance, it is not difficult to devise a variation of Shapley's counterexample which shows that convergence in the class of 3×3 games is not guaranteed. Finally, we mention some extensions of the algorithm which appear more promising (but technically very challenging) in Section 4, which closes the paper.

2. Fictitious Play by Cases

Standard Fictitious Play

The algorithm of fictitious play was originally introduced by Brown (1951) as a method for computing equilibria in noncooperative finite two-player games, but has gained popularity as a model of adaptive learning in games. We begin with a description of the algorithm and then recall the essentials of its interpretation as a learning process. For a more extensive discussion, see Fudenberg and Kreps (1992, 1993) which we follow in their notation.

Let G be a finite game in strategic form with two players, indexed by $i = 1, 2$. Call player 1 Row and player 2 Column; Row will be male and Column will be female. Given a player i, denote by $-i = 3 - i$ his or her opponent. Let S^i, $i = 1, 2$, be the finite set of pure strategies (or actions) for player i, let $S = S^1 \times S^2$ be the set of pure strategy profiles, and let u^i: $S \to \mathbf{R}$ give player's i payoffs. Furthermore, for $i = 1, 2$, let Σ^i be the set of mixed strategies for player i, let $\Sigma = \Sigma^1 \times \Sigma^2$ be the set of mixed strategy profiles and extend the domain of u^i from S to Σ by linearity in the probabilities.

Sacrificing efficiency for clarity, the algorithm of fictitious play can be described as follows:

1. *Initialization*. Set the time clock to $t = 1$ and let ζ_1 be the null history. For each player i and strategy s^{-i} in S^{-i}, define an initial weight $\eta_1^i(s^{-i}, \zeta_1) \geq 0$ such that $\Sigma_{s^{-i} \in S^{-i}} \eta_1^i(s^{-i}, \zeta_1) > 0$.
2. *Main Loop*.
   2.1. *Make strategy choices*. For each player i, let his or her current strategy choice s^i be the action selected according to some (possibly mixed) strategy σ which maximizes the weighted utility

$$\sum_{s^{-i} \in S^{-i}} u^i(\cdot, s^{-i}) \eta_t^i(s^{-i}, \zeta_t).$$

(In case of multiple maximizers, there is a fixed tie-breaking rule to select σ.) Let $s = (s^1, s^2)$ be the (pure) strategy profile obtained. Set $t = t + 1$.

2.2. *Update history*. Given t, let ζ_t be the history of play up to time t obtained by concatenation of ζ_{t-1} with the pure strategy profile s just played.

2.3. *Update weights*. For each player i and each strategy s^{-i} in S^{-i}, let the current weight be

$$\eta^i_t(s^{-i}, \zeta_t) = \eta^i_{t-1}(s^{-i}, \zeta_{t-1}) + \begin{cases} 1 & \text{if } -i \text{ has played } s^{-i} \text{ at time } t-1, \\ 0 & \text{otherwise.} \end{cases}$$

2.4. *Restart loop*. Go back to step 2.1.

The interpretation of fictitious play as a learning process leaves the algorithm unaffected but attaches a meaning to the objects defined in this description.

Suppose that Row and Column play G repeatedly at times $t = 1, 2, \ldots$. At each round of play, they observe only the action actually played by their opponent (and not the chosen mixed strategy). This generates a history of past play up to time t, denoted by ζ_t, which is the string $\zeta_t = (s_1, \ldots, s_{t-1})$, where $s_k \in S$ for $k = 1, \ldots, t-1$. Let $\mathbf{Z}_t$ be the set of all histories of play up to time t, let $\mathbf{Z}_1$ be the (singleton) set consisting of the null history, and let $\mathbf{Z}$ be the set of all possibile infinite histories.

At time t, player i looks at the past history of the opponent's play to form an opinion ρ^i_t: $\mathbf{Z}_t \to \Sigma^{-i}$ about what the opponent is going to play in the current game. This opinion depends on the initial conditions of the algorithm (i.e., initial weights and tie-breaking rule). It is obtained by the normalization of the current weights to a sum of 1, and therefore tends to converge to the empirical frequencies of play.

Finally, each player i follows a behavior rule ϕ^i to play the infinitely repeated game. Thus, $\phi^i = (\phi^i_1, \phi^i_2, \ldots)$ with ϕ^i_t: $\mathbf{Z}_t \to \Sigma^i$. In the model of fictitious play, players are presumed to follow the *myopic* behavior rule ϕ^i_t which prescribes to play the current game by choosing a strategy which maximizes the immediate payoffs with respect to the current opinion on how the opponent will act.

This interpretation enriches the fictitious play algorithm with the notions of history, opinions and (myopic) behavior rules. Starting from given initial opinions, players use the history of play to form (or revise) their opinions. These determine their choice of strategies, which in turn becomes a piece of the observed history. Learning takes place if the opinions generated by this process converge to a limit point, which represents what the players learn.

When learning occurs in the fictitious play model, players' opinions converge to the marginal distributions of a (possibly mixed) Nash equilibrium of G. However, although convergence is known to occur in special classes of games, it cannot be guaranteed in general (Shapley, 1964). Note also the apparent naiveté of fictitious play: players try to learn using a misspecified model where the opponents' play is stationary. However, under appropriate assumptions, the algorithm can be justified under the heading of either bounded rationality (Fudenberg and Kreps, 1992) or pseudo-Bayesian rationality (Eichberger, 1992).

The Revised Algorithm

There are certainly many ways in which the fictitious play algorithm can be extended to let players use their experiences in similar games. The extension that we consider here tries to be minimal, in the sense that we leave unchanged as much as possible of the original algorithm.

It is a simple observation that, if players are to use similarities, they must take into consideration several (distinct) games. Thus, we assume that there is a (nonempty) set **G** of $m \times n$ games in strategic form and that games are randomly drawn for play from **G** according to some probability distribution. Without loss of generality, we assume also that each $m \times n$ game is represented by some $m \times n$ bimatrix where the Row and Column players and their strategy sets are unmistakenly identified by their position. This assumption is called *common positioning,* and its main advantage is that we can use the same symbol to identify players or strategies of different games in **G**. Therefore, for instance, we denote by s_1^1 the first strategy of the Row player regardless of the game under consideration; analogously, S^i denotes the strategy set of player i for any game in **G**.

In the absence of similarities, the natural extension of fictitious play from a single game G to a set **G** of games is to initialize and independently run the standard fictitious play algorithm for each game in **G**. Thus, while every game should be played according to the myopic behavior rule, there would be distinct opinions, histories, and initial conditions for each game.

If players use similarities, however, it seems natural that initial conditions, histories, and opinions should interact to some extent. For tractability, we make substantially simplifying assumptions on the extent of these interactions. All games in **G** are judged equally similar by both players and the same initial conditions apply to each game in **G**. Since consistency with the myopic behavior rule suggests that the interaction should depend only on how players relate the current game under play with other games observed in the past, this implies that evidence from the past bears equally on the evolution of the opinions associated with each game in **G**. Therefore, a player holds identical opinions about each game in **G**, even though these might differ from the opponent's.

This leads to the following algorithm, which we call fictitious play by cases.

1. *Initialization*. Set the time clock to $t = 1$ and let ζ_1 be the null history. For each game G in **G**, each player i and strategy s^{-i} in S^{-i}, define an initial weight $\eta_1^i(s^{-i}, \zeta_1, G) \geq 0$ (independent of G) such that $\Sigma_{s^{-i}\in S^{-i}} \eta_1^i(s^{-i}, \zeta_1, G) > 0$.

2. *Main Loop*.

   2.1. *Select game*. Randomly select from **G** the current game H to be played according to some given probability measure.

   2.2. *Make strategy choices*. Given the current game H, for each player i let his or her current strategy choice s^i be the action selected according to some (possibly mixed) strategy σ which maximizes the weighted utility

$$\sum_{s^{-i}\in S^{-i}} u^i(\cdot, s^{-i})\eta_t^i(s^{-i}, \zeta_t, H).$$

   (In case of multiple maximizers, there is a fixed tie-breaking rule to select σ.) Let $s = (s^1, s^2)$ be the (pure) strategy profile obtained. Set $t = t + 1$.

   2.3. *Update history*. Given t, let ζ_t be the history of play up to time t obtained by concatenation of ζ_{t-1} with the game-strategy pair (H, s) formed by the game H just selected for play and by the (pure) strategy profile s just played in it.

   2.4. *Update weights*. For each game G, each player i and each strategy s^{-i} in S^{-i}, let the current weight be

$$\eta_t^i(s^{-i}, \zeta_t, G) = \eta_{t-1}^i(s^{-i}, \zeta_{t-1}, G) + \begin{cases} 1 & \text{if } -i \text{ has played } s^{-i} \text{ at time } t-1, \\ 0 & \text{otherwise.} \end{cases}$$

   2.5. *Restart loop*. Go back to step 2.1.

Informally, the algorithm can be described as follows. At each round, some game H is randomly selected for play. As in standard fictitious play, players follow a myopic behavior rule and make their strategy choices to maximize the immediate payoffs given their opinions (or normalized weights). After this stage, however, something different takes place in fictitious play by cases. The piece of history relevant to form an opinion about a game G is the string of game-strategy pairs (H, s) such that H is judged similar to G.[1] Thus, after H has been played, players update their

[1] In fact, a bit less: each player needs only to consider the opponent's choice and thus (H, s^{-i}) suffices for the task.

opinions for each game by incrementing the weight for the strategy just played by the opponent (and leaving the other weights unchanged).

An Example

Before discussing it, we illustrate the algorithm with a simple example. Let **G** be the class of all 2 × 2 games $G(z_1, z_2)$ which can be obtained by substituting a pair of positive real numbers to (z_1, z_2) in the bimatrix on the left of Fig. 1. For instance, letting $z_1 = 1$ and $z_2 = 7$, we obtain the game $G(1, 7)$ given on the right of the same figure.

In each game, the pure strategies available to Row are North (N) and South (S), while Column can choose either West (W) or East (E). Correspondingly, the pure strategy profiles are NW, NE, SW, and SE. The game $G(z_1, z_2)$ is a variation of the Stag Hunt (see Aumann, 1990) and has three equilibria: NW, SE (both strict) and an equilibrium in mixed strategies which depends on the pair $z = (z_1, z_2)$. Using a simplified notation where p stands for the mixed strategy which gives probability p to N and $(1 - p)$ to S and q has the analogous meaning for W and E, the equilibrium in mixed strategies is given by $(p^*(z_2), q^*(z_1)) = (7/(8 + z_2), 7/(8 + z_1))$.

At each date $t = 1, 2, \ldots$ an independent draw from the two (independent) random variables Z_1, Z_2 determines which game $G(z_1, z_2)$ is to be played. Suppose that the first four realizations of Z_1 and Z_2 are respectively $\{1, 4, 2, 16\}$ and $\{7, 2, 1, 19\}$. Thus, at date $t = 1$, game $G(1, 7)$ is drawn for play. Suppose that Row's initial weights for Column's strategies are 1 for W and 2 for E, while Column's initial weights for Row's strategies are 2 for N and 1 for S. By an easy computation, Row maximizes his payoff in game $G(1, 7)$ by choosing S while Column does so by choosing W. See Fig. 2, where the expected payoffs are computed after normalizing the weights to sum to 1, which is of no consequence.

At date $t = 2$, the current weights (for any game to come) are updated to 2 for both players and both strategies and game $G(4, 2)$ is drawn for play. Now, myopic behavior leads to the strict equilibrium SE which is also played at date $t = 3$, when game $G(2, 1)$ is drawn for play. At date $t = 4$, however, game $G(16, 19)$ is drawn for play and both players switch

	W	E	$(z_1 = 1)$		W	E
N	$(8 + z_1, 8 + z_2)$	$(0, 7)$	$\longrightarrow$	N	$(9, 15)$	$(0, 7)$
S	$(7, 0)$	$(7, 7)$	$(z_2 = 7)$	S	$(7, 0)$	$(7, 7)$

FIG. 1. A class of similar games and one of its representatives.

		Opinion on opponent		Expected payoffs		Choice of action
Round 1	Row	1	2	3	7	S
($z_1 = 1, z_2 = 7$)	Column	2	1	10	7	W
Round 2	Row	2	2	6	7	S
($z_1 = 4, z_2 = 2$)	Column	2	2	5	7	E
Round 3	Row	2	3	4	7	S
($z_1 = 2, z_2 = 1$)	Column	2	3	3.6	7	E
Round 4	Row	2	4	8	7	N
($z_1 = 16, z_2 = 19$)	Column	2	4	9	7	W

FIG. 2. An example of fictitious play by cases.

independently from the strict equilibrium SE to the other strict equilibrium NW.

This example shows that the characteristic property that strict equilibria are absorbing for standard fictitious play is no longer true for fictitious play by cases. In fact, since the selection of the game to be played is stochastic, it is possible that some (similar) game drawn for play has sufficiently different payoffs to override for at least one player the tendency of his or her current opinions to favor the equilibrium strategy profile just played.

More precisely, let the current players' opinions be described by the pair $(\hat{p}, \hat{q})$, where $\hat{p}$ is the probability according to Column that Row will play N and similarly $\hat{q}$ is the probability according to Row that Column will play W. As it is easy to show, given (z_1, z_2) and disregarding ties, Row plays S if and only if $\hat{q} - q^*(z_1) < 0$ while Column plays E if and only if $\hat{p} - p^*(z_2) < 0$. Moreover, $\hat{p}$ decreases if Row plays S and so does $\hat{q}$ if Column plays E. In standard fictitious play, p^* and q^* are constant: thus, once the two inequalities are simultaneously satisfied and SE is played, $\hat{p}$ and $\hat{q}$ change in the right direction to make the inequalities even stronger. Hence, these hold *a fortiori* forever after. On the contrary, in fictitious play by cases $p^*(z_2)$ and $q^*(z_1)$ are (in this example, decreasing) functions. Thus, there might be (in this example, sufficiently high) realizations of z_1 or z_2 which make $p^*(z_2)$ or $q^*(z_1)$ so low to reverse the sign of at least one of these inequalities, moving the process away from the SE equilibrium.

Comparison of the Algorithms

Mathematically, the major difference between standard fictitious play and fictitious play by cases is that the latter is a stochastic algorithm while the former is deterministic. In other words, both the outcome and the evolution of opinions generated by standard fictitious play are determined only by the initial conditions, whereas for fictitious play by cases they depend also (and often, mainly) on the particular realization of the stream of games to be played. The example above, for instance, shows how this can make strict Nash equilibria fail to be absorbing for fictitious play by cases.

In general, the interactions between similar games may carry the evolution of opinions about a game away from any strategy profile, upsetting the dynamics that would emerge if the game were analyzed alone as in standard fictitious play. When the algorithm is interpreted as a learning process, this has two implications: enough contrary evidence from similar games may upset even strict equilibria, and (when learning takes place) what is learned in the end may be path-dependent. Both implications seem particularly appropriate for the learning interpretation.

A related problem is that the question of existence (and possibly uniqueness) of limit points for the opinions generated by this process is much more complicated for fictitious play by cases. In fact, even mild generalizations of the algorithm could make its asymptotic behavior too difficult to study. For instance, if we allow different opinions for any game G in **G**, there would be several learning processes running simultaneously while influencing each other. We conjecture that techniques from statistical mechanics or the theory of differential inclusions could be of help, but so far we have made little progress on this issue.

In any case, it should be apparent that the convergence results known for standard fictitious play are of no help here, where convergence must be attained in a probabilistic sense. However, they suggest where it might be easier to establish some kind of convergence. Positive results in this direction are given in the next section, where sufficient conditions are given for almost sure convergence of opinions in the class of 2×2 games.

Finally, we note that fictitious play by cases may be interpreted as a process by which people try to learn not how to play *a specific game*, but how to play classes of *games*. In the example above, for instance, the current values of $\hat{p}$ and $\hat{q}$ and the assumption of myopic behavior suffice to determine how to play any game in the class. Thus, if the sequence $\{(\hat{p}_t, \hat{q}_t)\}$ converges to a limit point, one might say that some way to play all the games in the class has been learned. Moreover, if the support of Z_1 or Z_2 is uncountable, most of these games will never be drawn for play although the opinion which is learned in the limit applies to them as well. Hence, in some sense, it is possible to learn how to play even a game that has never been played before.

3. Convergence in 2×2 Games

In this section, we study the asymptotic behavior of the algorithm of fictitious play by cases when **G** is an arbitrary set of 2×2 games. We provide a set of sufficient conditions on the sampling distribution over **G** such that, for almost every history, the sequence of opinion pairs $\{(\hat{p}_t, \hat{q}_t)\}$ converges to a limit point as $t \to +\infty$. Moreover, we characterize the set of possible limit points.

The strategy of proof for this result exploits the property that fictitious play by cases falls in a class of algorithms widely studied in the theory of stochastic approximation and generalized urn processes. See for instance Arthur *et al.* 1983, 1986; Benveniste *et al.*, 1990; Nevel'son and Has'minskiĭ, 1973. Several variants of convergence results are known for these algorithms. Under appropriate assumptions, we prove one of these results (which we call the Convergence Lemma) for an algorithm which includes fictitious play by cases as a special case. The lemma is already known in this literature, but we offer our proof in the appendix because we have not been able to find a satisfactory one.

The algorithm that we consider is of the form

$$X_{n+1} = X_n + c_n A(X_n, Z_{n+1}), \qquad n = 0, 1, 2, \ldots, \tag{1}$$

where X_n is an d-vector of estimates, Z_n is a k-vector of random variables, c_n is a positive real number, and A is a vector-valued function. Let D denote an open subset of $\mathbf{R}^d$ where the algorithm is defined and let X_0 denote the initial condition. Assume that the random d-vector X_0 and the sequence of random k-vectors $Z_1, Z_2, \ldots$ are defined on some probability space $(\Omega, \mathcal{A}, P)$ and denote by $\mathcal{A}_n$ the σ-algebra of events generated by $X_0, Z_1, \ldots, Z_n$.

We make the following assumptions on the form of the algorithm:

(A1) There exists a compact set $C \subset D$ such that $X_0 \in C$ implies $X_n \in C$ for all n and all realizations of the sequence $Z_1, Z_2, \ldots$ Moreover, $X_0 \in C$ and $\sup_\kappa |Z_\kappa| \le K_1 < +\infty$ with probability one.

(A2) $\{c_n\}$ is a nonincreasing sequence such that $\Sigma_n\, c_n = +\infty$ and $\Sigma_n\, c_n^2 < +\infty$.

(A3) Given any $x \in C$, there exists a probability distribution $\mu(\cdot; x)$ on $\mathbf{R}^\kappa$ such that $P(Z_{n+1} \in B \mid \mathcal{A}_n) = \int_B \mu(dz; X_n)$ for any Borel subset B of $\mathbf{R}^k$ and any $n = 0, 1, \ldots$.

(A4) There exists a constant K_2 such that $|A(x, z)| \le K_2$ for all $x \in C$ and all $|z| \le K_1$.

(A5) The function $a(x) = \int A(x, z)\mu(dz; x)$ is locally Lipschitz on C; i.e., there exists a constant K_3 such that $|a(x) - a(x')| \le K_3|x - x'|$ for all $x, x' \in C$. For any initial condition $x_0 \in C$, the solution of the vector field $\dot{x} = a(x)$ remains in C.

Given a nonrandom d-vector x, let P_x denote the probability measure P conditional on the initial condition being $X_0 = x$. Unless specified, the usage of the "almost surely" (a.s.) qualifier in the following is to be understood with respect to the P_x distribution induced by an arbitrary initial condition $x \in C$. Similarly, all expected values are taken with respect to P_x, unless explicitly noted otherwise. Given an algorithm of the form (1) which satisfies (A1)–(A5), we call the vector field $\dot{x} = a(x)$ its *associated deterministic system*.

The Convergence Lemma associated with this algorithm is a stochastic analog of Liapunov's stability theorem.

THEOREM 1 (Convergence Lemma). *Consider an algorithm of the form* (1) *and suppose that it satisfies* (A1)–(A5). *Let U be an open subregion of D such that $U \cap C \neq \emptyset$ and suppose that on $U \cap C$ there exists a nonnegative C^2 Liapunov function V for $\dot{x} = a(x)$ such that $\dot{V}(x) = \nabla V(x) \cdot a(x) \leq 0$ for all $x \in U \cap C$. Given any $k_0 \geq 0$ such that the set $\{x \in U \cap C: V(x) \leq k_0\}$ contains the set $\{x \in U \cap C: \dot{V}(x) = 0\}$, let*

$$\Gamma = \{x \in U \cap C: V(x) \leq k_0\}.$$

If X_n visits infinitely often a compact neighborhood L of Γ such that $L \subseteq U \cap C$, then X_n converges a.s. to a random variable with support contained in Γ.

Fictitious play by cases is a special case of the algorithm just described. The following assumption simplifies its (stochastic) laws of motion and makes it more amenable to analysis.

ASSUMPTION 1 (Independence). *The sampling distribution on* **G** *induces two distribution functions over Row's and Column's 4-tuples of payoffs which are (stochastically) mutually and serially independent.*

We state this assumption in a form which is convenient but somewhat implicit, because it involves two distributions induced by the sampling distribution on **G**. A sufficient condition for this assumption to hold is that each payoff in the bimatrix describing a game is an independent draw from some random variable. Thus, for instance, the example discussed above in connection with Fig. 2 satisfies Assumption 1 if Z_1 and Z_2 are (stochastically) independent.

Consider the laws of motion for players' opinions in fictitious play by cases. Let $\hat{p}_t$ and $\hat{q}_t$ be the players' opinions at date t. Since the change in opinions at date $t + 1$ is determined by the players' choices at date t, we compute the probability with which players choose their actions at this date.

Consider Row's strategy choice under the myopic behavior rule for a

2×2 game with arbitrary payoffs like in Fig. 3. Of course, analogous conclusions hold about Column's choice. Assume that Row's opinion is that Column will play W with probability $\hat{q}$ and E with probability $1 - \hat{q}$. Depending on the sign (positive, negative, or zero) of the two quantities $a_1 - a_3$ and $a_2 - a_4$, there are nine possible events. We combine them in four cases by assuming (without loss of generality) that the tie-breaking rule prescribes playing N for Row and W for Column.

If $a_1 - a_3 \geq 0$ and $a_2 - a_4 \geq 0$ (event E_1^R), Row's behavior rule (and the tie-breaking rule) prescribe him to play N, regardless of his opinion. Symmetrically, if $a_1 - a_3 \leq 0$ and $a_2 - a_4 \leq 0$ with at least one inequality holding strictly (event E_2^R), Row should play S, regardless of his opinion. On the contrary, Row's opinion determines his choice in the other two events, which are characterized by the inequality $(a_1 - a_3)(a_2 - a_4) < 0$. When this holds, we can define the real-valued quantities

$$p^* = \frac{b_2 - b_4}{b_2 - b_4 + b_3 - b_1}, \qquad q^* = \frac{a_2 - a_4}{a_2 - a_4 + a_3 - a_1} \tag{2}$$

and check that, when $(a_1 - a_3) > 0$ (event E_3^R), Row should play N if (and, disregarding ties, only if) $\hat{q} > q^*$ while, when $(a_1 - a_3) < 0$ (event E_4^R), he should play N if (and, disregarding ties, only if) $\hat{q} < q^*$. If $\hat{q} = q^*$ and a tie occurs (which is a null event in fictitious play by cases), the tie-breaking rule dictates that Row plays N and, for convenience, we ascribe this possibility to E_1^R.

When a game G is drawn according to the sampling distribution on **G**, one (and only one) of these four exhaustive and mutually exclusive events occurs, respectively with probability $\pi(E_i^R) = \pi_i$, for $i = 1, \ldots, 4$. Simultaneously and independently, one (and only one) of the corresponding events E_i^C occurs for Column's choice, respectively with probability $\beta(E_i^C) = \beta_i$, for $i = 1, \ldots, 4$. Note that π and β do not depend on $\hat{p}$ and $\hat{q}$. Finally, let $F_i^R(q)$ be the distribution function for q^* conditional on E_i^R for $i = 3, 4$; when E_i^R is a null event, we set it to be the zero function for convenience. The distribution function $F_i^C(p)$ for p^* conditional on E_i^C is analogously defined.

$G =$

	W	E
N	(a_1, b_1)	(a_2, b_3)
S	(a_3, b_2)	(a_4, b_4)

FIG. 3. An arbitrary 2×2 game.

At date t, then, Row plays N with probability

$$F^R(\hat{q}_t) = \pi_1 + \pi_3 F_3^R(\hat{q}_t) + \pi_4 [1 - F_4^R(\hat{q}_t)] \tag{3}$$

and S with probability $1 - F^R(\hat{q}_t)$. At date $t + 1$, Column accordingly adjusts her opinion $\hat{q}_{t+1}$. Similarly, at date t Column plays W with probability

$$F^C(\hat{p}_t) = \beta_1 + \beta_3 F_3^C(\hat{P}_t) + \beta_4 [1 - F_4^C(\hat{p}_t)]$$

and E with probability $1 - F^C(\hat{p}_t)$, and at date $t + 1$ Row accordingly adjusts $\hat{q}_{t+1}$.

If we denote by η^1 and η^2 the sum of the initial weights respectively of Row and Column, the stochastic laws of motion for $\hat{q}_t$ and $\hat{p}_t$ are then

$$\hat{p}_{t+1} = \begin{cases} \hat{p}_t + \dfrac{1 - \hat{p}_t}{t + \eta^2} & \text{with probability } F^R(\hat{q}_t) \\ \hat{p}_t - \dfrac{\hat{p}_t}{t + \eta^2} & \text{with probability } 1 - F^R(\hat{q}_t) \end{cases} \tag{5}$$

and

$$\hat{q}_{t+1} = \begin{cases} \hat{q}_t + \dfrac{1 - \hat{q}_t}{t + \eta^1} & \text{with probability } F^C(\hat{p}_t) \\ \hat{q}_t - \dfrac{\hat{q}_t}{t + \eta^1} & \text{with probability } 1 - F^C(\hat{p}_t) \end{cases} \tag{6}$$

which determine an algorithm of the form (1) satisfying (A1)–(A5). The associated deterministic system is obtained by taking expected values in (5) and (6), which gives the system

$$\hat{p}^e_{t+1} = \hat{p}_t + \frac{F^R(\hat{q}_t) - \hat{p}_t}{t + \eta^2}$$

$$\hat{q}^e_{t+1} = \hat{q}_t + \frac{F^C(\hat{p}_t) - \hat{q}_t}{t + \eta^1},$$

where we use the superscript e to denote expectation, and then by moving from the discrete to the continuous-time formulation to obtain

$$\begin{aligned} \dot{p} &= F^R(q) - p \\ \dot{q} &= F^C(p) - q, \end{aligned} \tag{7}$$

where we have dropped any additional notation from p and q for simplicity.

To impose more structure on this system, we introduce a smoothness assumption.

ASSUMPTION 2 (Smoothness). *The functions $F^R(q^*)$ and $F^C(p^*)$ are C^2 functions on the closed unit interval, with first-order derivatives f^R and f^C.*

A necessary and sufficient condition for Assumption 2 to hold is that the conditional distribution functions F_i^R in (3) and F_i^C in (4) are twice continuously differentiable on [0, 1], for $i = 3, 4$.

Note that F^R and F^C are arbitrary C^2 functions taking values in [0, 1]. In particular, they do not need to be distribution functions, which makes the analysis of the asymptotic behavior of (7) more complicated. General statements, however, are still possible. First, we note that, by Bendixson's criterion, there are no periodic orbits. Second, the system is dissipative in (i.e., never leaves) $C = [0, 1]^2$. Third, by Brouwer's fixed point theorem, the set Φ of fixed points is not empty. Moreover, since F^R and F^C have continuous derivatives, Φ is a finite set of connected components. We will assume that $f^R(q)f^C(p) \neq 1$ for all points $\varphi = (q, p)$ in Φ. In this case, every fixed point is hyperbolic and the connected components of Φ reduce to singletons so that Φ is a finite set of points.[2]

Partition the set Φ into the set Φ_u of unstable points with generic element v and the set Φ_a of asymptotically stable points with generic element α. Denote by $W(\varphi)$ the global stable manifold of a point φ in Φ: obviously, $\varphi \in W(\varphi)$. It is apparent that the global stable manifolds for the fixed points in Φ partition the unit square. The partition is such that the stable manifolds of unstable points are the one-dimensional separatrices of the open and connected stable manifolds of the asymptotically stable points.

THEOREM 2. *Under Assumptions 1 and 2, if $f^1(p)f^2(q) \neq 1$ for all (p, q) in Φ, then the opinions generated under fictitious play by cases converge a.s. to a point in Φ for any assignment of initial weights.*

[2] See Hale and Koçak (1991) for the statement of Bendixson's criterion (p. 373) and the definitions of hyperbolic equilibrium (p. 19), dissipative system (p. 394), and separatrix (p. 398).

Proof. Let $\{(p_t, q_t)\}$ be the sequence of opinions generated by (5) and (6). Since the systems is dissipative in $C = [0, 1]^2$, the set of its cluster points is not empty. For any $\varepsilon > 0$ it is possible to partition C into a compact subset A_ε and its complement A_ε^c such that $\Phi_a \subset A_\varepsilon$, $\Phi_u \subset A_\varepsilon^c$, and A_ε has Lebesgue measure $\lambda(A_\varepsilon) > 1 - \varepsilon$. Each cluster point, then, belongs to one and only one of A_ε, $A_\varepsilon^c \backslash \Phi_u$, or Φ_u.

The rest of the proof is in three steps. First, we show that if $\{(p_t, q_t)\}$ has a cluster point in A_ε then $\{(p_t, q_t)\}$ a.s. has a limit point in Φ_a. Second, we prove that $\{(p_t, q_t)\}$ cannot have a cluster point in $A_\varepsilon^c \backslash \Phi_u$ with positive probability. Hence, with probability one only the points in Φ can be cluster points. Third, we show that if a point in Φ is a cluster point, it must also be a limit point. Since cluster points exist for all ω, the result follows.

For each asymptotically stable point α in Φ, let $K(\alpha)$ denote an arbitrary compact neighborhood of α contained in its domain of attraction. By the converse to Liapunov's stability theorem (see Corollary 5.1 in Krasovskiĭ, 1963, p. 28), there is an open neighborhood $U(\alpha)$ of $K(\alpha)$, where it is possible to construct a Liapunov function which satisfies the assumptions of the Convergence Lemma for $\Gamma = \{\alpha\}$. Since Φ_a is finite, if $\{(p_t, q_t)\}$ visits infinitely often the set $A = \bigcup\{K(\alpha) | \alpha \in \Phi_a\}$, it must also visit infinitely often some $K(\alpha)$, say $K(\alpha')$. Applying the Convergence Lemma to $U(\alpha')$ with $\Gamma = \{\alpha'\}$, it follows that the algorithm converges a.s. to the limit point $\alpha' \in \Phi$. Given any $\varepsilon > 0$, for an appropriate choice of sufficiently large $K(\alpha)$'s, $A = A_\varepsilon$.

By contradiction, suppose now that $\{x_t\} = \{(p_t, q_t)\}$ has a cluster point $\hat{x}$ in $A_\varepsilon^c \backslash \Phi_u$ with positive probability. There are two possible cases: either $\hat{x}$ is in the global stable manifold $W(v)$ of some unstable fixed point v in Φ_u or it is not. If it is not, it must be in the global stable manifold of some asymptotically stable fixed point α; therefore, for $K(\alpha)$ sufficiently large, $\hat{x} \in K(\alpha)$. Hence, x_t visits infinitely often $K(\alpha)$ so that convergence to α occurs almost surely and $\hat{x}$ cannot be a cluster point with positive probability. Suppose then that $\hat{x}$ is in the global stable manifold $W(v)$ of v. Again, there are two possible cases: either $\hat{x}$ is a limit point or it is not (and then there must be another cluster point). We show that neither case is possible.

Since $\hat{x} \in A_\varepsilon^c \backslash \Phi_u$, $a(\hat{x}) \neq 0$. Therefore, by the Rectifiability Theorem (see Theorem 2.1 in Arnold and Il'yashenko, 1988, p. 14), in a sufficiently small neighborhood U of $\hat{x}$, there exists a C^2 diffeomorphism $y = y(x)$ such that in the new coordinates the associated deterministic system $\dot{x} = a(x)$ becomes $\dot{y}_1 = 1$, $\dot{y}_2 = 0$. Informally, in the new coordinates, $y(U)$ is a box-like neighborhood such that the orbits of the system (and in particular those in $W(v)$) enter at one end of the box and flow through the other end. As the motion on $W(v)$ is towards v, the exiting end is the one closer to v.

Define the C^2 Liapunov function $V(x) = y_1(x)$ on U and suppose that

$V(\hat{x}) = k$. For sufficiently small $k_1 < k < k_2$, then, Lemma 6 (see the Appendix) shows that x_t must a.s. leave U towards v conditionally on not exiting from the other side of the box. Since $\hat{x}$ is a cluster point for $\{x_t\}$, however, x_t must visit U i.o. and therefore, by Lemma 8 (see the Appendix and the discussion following it) must leave U towards v with probability one. It follows that $\hat{x}$ cannot be a limit point. Suppose then that there is another cluster point $\bar{x} \neq \hat{x}$. Since x_t must travel from $\hat{x}$ to $\bar{x}$ i.o., if $\bar{x} \notin W(v)$ then x_t must visit A i.o.; thus, by the above, a.s. convergence to a limit point in Φ_a takes place and $\hat{x}$ cannot be a cluster point. On the other hand, if $\bar{x} \in W(v)$, this same reasoning applies to show that x_t must go from one cluster point to another along paths which cannot stay bounded away from $W(v)$. Hence, all the points between $\hat{x}$ and $\bar{x}$ on $W(v)$ must be cluster points. Repeating the rectifying construction for an appropriate neighborhood of each of them (except the fixed point v), we obtain that eventually x_t must leave a.s. each neighborhood towards v and therefore it cannot come back to a cluster point that was left before. It follows that $\hat{x}$ and, in general, any point on $W(v)\backslash v$ cannot be a cluster point.

The last step is easily obtained. If a cluster point $x \in \Phi_a$, x_t visits i.o. a compact neighborhood of x in its domain of attraction and therefore converges a.s. to x by the Convergence Lemma. If a cluster point x is in Φ_u, then there cannot be another cluster point in $W(v)$ because of the argument above, and there cannot be another cluster point in $C\backslash W(v)$ because otherwise x_t should visit A i.o. and thus would converge to a limit point in Φ_a, contradicting the assumption that x is a cluster point. ■

Note the different status of the asymptotically stable points in Φ_a and the unstable points in Φ_u in the proof: we show that there exist sufficient conditions under which the former can be limit points of the algorithm, while for the latter we simply cannot exclude that they might be so. In fact, since the union of the global stable manifolds for asymptotically stable points has Lebesgue measure 1, under the stochastic perturbations of the algorithm defined by (5) and (6), the associated deterministic system spends most of its time in this union. Thus, we conjecture (but are unable to prove) that Theorem 2 can be strengthened as follows.

CONJECTURE. *Under the assumptions of Theorem 2, the opinions generated under fictitious play by cases converge a.s. to a point in* $\Phi_a \subseteq \Phi$ *for any assignment of initial weights.*

Note also that, although standard fictitious play is a limit case of fictitious play by cases, it does not satisfy Assumption 2 because F^R and F^C are degenerate distributions. Therefore, Miyasawa's (1961) result about convergence of standard fictitious play in any 2×2 game cannot be obtained as a special case of this theorem.

Special Cases

We introduce an additional assumption which appears to be of interest. See LiCalzi (1992) for a more detailed analysis. Recall the events E_i^R and E_i^C defined above for $i = 1, \ldots, 4$. Each pair of events E_i^R and E_j^C corresponds to a different partial order for the payoffs of Row and Column, which we call of type T_{ij}.

In particular, for $i, j = 3, 4$, games of the same type T_{ij} are similar in that players have the same best replies to the opponent's pure strategies. (The same statement holds if we adequately refine E_i^R and E_i^C for $i = 1, 2$: we omit the obvious details.) Thus, it seems an appropriate assumption to require that all games in **G** are of the same type.

ASSUMPTION 3 (Best response similarity). *All the games in* **G** *are of the same type; i.e., players' best replies to pure strategies for each game in* **G** *are the same.*

We can use this assumption to illustrate the intuition behind the dynamics that drives convergence of opinions in fictitious play by cases. Let $(\hat{p}, \hat{q})$ denote the current opinion and assume that all games in **G** are of type T_{33}. Then each game in **G** has three equilibria, two of which are in pure strategies (NW and SE) and one is in mixed strategies. Moreover, this gives $\pi_1 = \pi_4 = 0$ and $\pi_3 = 1$ in (3), and $\beta_1 = \beta_4 = 0$ and $\beta_3 = 1$ in (4), so that both F^R and F^C are increasing (distribution) functions with density functions f^R and f^C. It is easy to check that (at least) both NW and SE are fixed points for (7), so by Theorem 2 both can be "learned" by fictitious play by cases.

As described above, the sampling distribution over G determines what is going to be played via the induced distribution on the pair (p^*, q^*). We call *basin of attraction* of a pure strategy profile s the set of all points (p^*, q^*) which, given the current opinions, lead players to play s. The *attraction map* of a game is the map which partitions the unit square into basins of attraction for all the strategy profiles.

In Fig. 4, we have drawn one such attraction map. The unit square is partitioned by the perpendiculars in $\hat{p}$ and $\hat{q}$ into four square regions; each one of them is a basin of attraction for the pure strategy profile that we have marked. Note that the basins of attraction lie opposite their natural position, because we are taking a different viewpoint; instead of looking at which beliefs $(\hat{p}, q^*)$ lead to the strategy profile s, given the equilibrium in mixed strategies (p^*, q^*), we consider which pairs (p^*, q^*) lead to s given beliefs $(\hat{p}, \hat{q})$.

Using the attraction map, consider for instance what happens when the density functions f^R and f^C are decreasing, in which case the only fixed points for (7) are NW (asymptotically stable) and SE (unstable). Then low

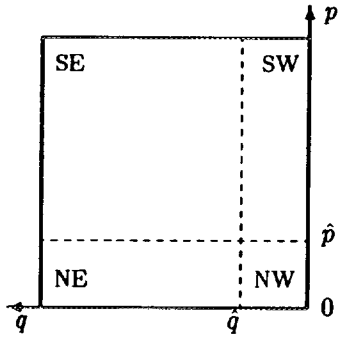

FIG. 4. The attraction map for a 2 × 2 game.

values for p^* and q^* are more likely, which makes (p^*, q^*) likely to lie in the basin of attraction for NW. Whenever this occurs, players choose to play respectively N and W. In turn, this increases the values of $\hat{p}$ and $\hat{q}$, making it even more likely to play NW in the next round. As this process feeds (on average) on itself, we might expect NW to eventually emerge as the only possible equilibrium. This is the content of our conjecture, although Theorem 2 proves only that convergence takes place to either NW or SE.

For a different example, look at the case where f^R and f^C are such that the most likely values for p^* and q^* lie somewhere in the interior of the square. If by random fluctuations one of the two attraction basins for NW or SE happens to include most of this area, then a dynamics similar to the above will start again, leading the system to one of the two equilibria in pure strategies which are also the only asyptotically stable fixed points. However, if influences from these two areas stay balanced, the process might end up somewhere in between as in a tug-of-war between equally strong opponents. Note that this kind of heuristic reasoning holds only on average, so that the process might wander endlessly if likely outcomes are sufficiently distributed. This explains the difficulty of proving convergence results with probability one.

Of course, there are also situations in which fictitious play by cases behaves very much like standard fictitious play. Consider, for instance, the class of games associated with Fig. 1, but assume that the support of both Z_1 and Z_2 is the bounded interval (0, 4) rather than the positive reals. For the same initial weights of Fig. 2, the evolution of fictitious play by cases is entirely deterministic and the process converges to the SE equilibrium for any realization. It is easy to see that this happens because, regardless of the realizations of Z_1 and Z_2, after the second period $(p^*(z_2),$

$q^*(z_1)$) always falls in the basin of attraction of SE. Here, the random microfluctuations of $(p^*(z_2), q^*(z_1))$ cannot affect the deterministic macro-tendency towards SE, as they are never sufficiently large to make the process leave the basin of attraction of SE.

Under Assumption 3, we can improve Theorem 2 when **G** is a class of games of type T_{34} (or T_{43}). Any game in this class has a unique equilibrium in (proper) mixed strategies, so that we can think of **G** as a set of games with similar strategic characteristics. Alternatively, we can view **G** as a set of augmented games associated with the purification *à la* Harsanyi (1973) of a game with a unique equilibrium in mixed strategies, along the lines given in Fudenberg and Kreps (1993).

Since all games in **G** are of type T_{34}, let $\pi_1 = \pi_4 = 0$ and $\pi_3 = 1$ in (3); similarly, let $\beta_1 = \beta_3 = 0$ and $\beta_4 = 1$ in (4). As it is easy to check, this makes F^R an increasing function and F^C a decreasing function. (In fact, a bit more: F^R is a distribution function and F^C is the complement to one of a distribution function.) Hence, the fixed point is unique and $\Phi = \{\varphi\}$ is a singleton. Applying Theorem 2, this gives the following corollary.

COROLLARY 3. *Under Assumptions 1–3, if all the games in* **G** *are of type* T_{34} *(or* T_{43}*) the opinions generated under fictitious play by cases converge a.s. to the unique fixed point of* (7) *for any assignment of initial weights.*

Compare this with Proposition 8.1 in Fudenberg and Kreps (1993), which makes a weaker smoothness assumption on F^R and F^C but requires the positivity of their density functions in a neighborhood of the fixed point.

Unfortunately, an analogous improvement is not possible when **G** is a set of games of type T_{33} (or T_{44}), which have three equilibria. In this case, the discussion above shows that the only additional conclusion we can draw is that (at least) both equilibria in pure strategies are fixed points. Thus, convergence to either one is possible, although not necessarily equally likely.

4. Concluding Remarks

The major message of this paper is that players can learn from broader classes of games. However, there are many issues and some modelling possibilities which we have left unexplored. Our favorites include the following.

First, it not necessary that the class of "cases" be commonly known. A more realistic assumption would be to have players look at similar classes of similar games, allowing play over past games to bear perhaps on Row's opinions but not on Column's. In turn, this suggests that players may refine their similarity classes over time to make them more effective in predicting the opponent's behavior. In this case, we might expect that the supports of F^R and F^C would become smaller over time, making the limit points (if they exist) closer to the Nash equilibria of the games still in the same class. Both of these possibilities can be modeled, but appear difficult to analyze.

A condition which may be easier to analyze is that players usually allow for degrees in the strength of their judgments of similarity. Borrowing from Gilboa and Schmeidler's (1992) case-based decision theory, we could posit a similarity function ξ: $G \times G \to [0, 1]$ such that $\xi(G, H)$ measures the "strength" of the similarity judgment between two games G and H. Assuming for simplicity that both players have the same similarity function, if we change step 2.4 in the algorithm of fictitious play by cases to

2.4. *Update weights.* For each game G, each player i and each strategy s^{-i} in S^{-i}, let the current weight be

$$\eta_t^i(s^{-i}, \zeta_t, G) = \eta_{t-1}^i(s^{-i}, \zeta_{t-1}, G) + \begin{cases} \xi(H, G) & \text{if } -i \text{ has played } s^{-i} \text{ at time } t-1, \\ 0 & \text{otherwise,} \end{cases}$$

the revised algorithm would take into account the different weight that evidence from more or less similar games bears to a particular game. Much more difficult, instead, is the issue of how similarity with games that have different strategic (or worse, extensive) forms might be modelled.

Appendix

Proof of the Convergence Lemma

In this section, we prove the Convergence Lemma stated in Section 3. We start by studying the cumulative random fluctuation for the more general case, where

$$\varepsilon_n(\phi) = \phi(X_{n+1}) - \phi(X_n) - c_n \nabla\phi(X_n) \cdot a(X_n) \tag{8}$$

and ϕ: $C \to \mathbf{R}$ is a C^2 function. Let

$$M_1 = \sup\{|\phi_i'(x)|: x \in C, 1 \le i \le d\}$$

$$M_2 = \sup\{|\phi_{ij}''(x)|: x \in C, 1 \le i, j \le d\}.$$

By Taylor's formula, for any $x, y \in C$ there exists a d-vector $O_\phi(|x - y|^2)$ such that

$$|O_\phi(|x - y|^2)| \leq M_2|x - y|^2 \tag{9}$$

and

$$\phi(x) - \phi(y) = \nabla\phi(x) \cdot (x - y) + O_\phi(|x - y|^2).$$

Substituting in (8), this gives

$$\begin{aligned} \varepsilon_n(\phi) &= \nabla\phi(X_n) \cdot [(X_{n+1} - X_n) - c_n a(X_n)] + |O_\phi(|X_{n+1} - X_n|^2)| \\ &= c_n \nabla\phi(X_n) \cdot [A(X_n, Z_{n+1}) - a(X_n)] + |O_\phi(|X_{n+1} - X_n|^2)|. \end{aligned} \tag{10}$$

Our first lemma establishes the convergence of the cumulative random fluctuation $\sum \varepsilon_n(\phi)$ and the existence of an L^2 upper bound for the maximal fluctuation. For lack of space, we omit its proof which can be found in LiCalzi (1992).

LEMMA 4. *For all m, let*

$$R_m = \sum_{n=0}^{m} \varepsilon_n(\phi)$$

Then R_m converges almost surely and in L^2 as $m \to +\infty$. Moreover, there exists a constant K_4 such that

$$E\{\sup_m |R_m|\}^2 \leq K_4 \sum_{n=0}^{+\infty} c_n^2.$$

We next state the lemma which guarantees that the asymptotic behavior of (1) is almost surely the same of the associated deterministic vector field $\dot{x} = a(x)$. This justifies the use of continuous-time Liapunov techniques on the associated deterministic system for the asymptotic analysis of the discrete-time algorithm (1). Let $t_0 = 0$ and $t_n = \sum_{i=1}^{n} c_i$ for all $n \geq 1$. The trick used to compare the sequence $\{X_n\}$ generated by the algorithm (1) with the trajectory of the vector field $\dot{x} = a(x)$ is to let $X(t) = \sum_k \mathbf{I}(t_k \leq t \leq t_{k+1})X_k$. For $m(n, T) = \inf\{k \geq n: c_{n+1} + \cdots + c_{k+1} \geq T\}$, then, the study of the behavior of X_k between integers n and $m(n, T)$ is equivalent to the study of the behavior of $X(t)$ between times t_n and $t_n + T$. The following lemma (see Theorem 9, in Benveniste *et al.*, 1990, p. 232) states this precisely.

LEMMA 5. *Let $X(t;\, t_n,\, x_n)$ denote the solution of*

$$\dot{x} = a(x) \tag{11}$$

such that $x(t_n) = x_n$. Given arbitrary positive T, δ, and ε, there exists $\nu = \nu(T, \delta, \varepsilon)$, such that

$$P\left\{ \sup_{n \le k \le m(n,T)} |X_k - X(t_k;\, t_n,\, x_n| \ge \delta \right\} \le \varepsilon \qquad \textit{for all } n \ge \nu.$$

We are now ready to study the asymptotic behavior of the algorithm using the Liapunov function V. Let us introduce some notation. Denote by $K = \sup_{x\in U\cap C} V(x)$ the supremum of V on the set $U \cap C$; by $L(k) = \{x \in U \cap C : V(x) \le k\}$ its k-level set; by $\vartheta(k) = \inf\{n\colon X_n \in L(k)\}$ the first time at which the algorithm enters the level set $L(k)$ and by $\tau(k) = \inf\{n\colon X_n \notin L(k)\}$ the first time of exit from $L(k)$.

We prove first that, for $k_2 > k_1 > k_0$, if the value of the Liapunov function V on $U \cap C$ does not get higher than k_2, then it must eventually hit k_1.

LEMMA 6. *Suppose that $k_0 < k_1 < k_2 < K$. Then, for all $x \in L(k_2)$, $\vartheta(k_1) < +\infty$ P_x-a.s. on $\{\tau(k_2) = +\infty\}$.*

Proof. It suffices to show that $P_x\{\tau(k_2) = \vartheta(k_1) = +\infty\} = 0$. By contradiction, suppose that this probability is positive. Let $L = \{x \in U \cap C\colon k_1 \le V(x) \le k_2\}$. By (A6), $\dot{V}(x)$ is bounded away from 0 on L so that there exists some $\alpha > 0$ such that $\dot{V}(x) \le \alpha$ on L. Let φ be a C^2 function on U which coincides with V on $L(k_2)$ and such that $\inf_{x\notin L(k_2)} \varphi(x) = k_2$. Then $\nabla\phi(x) \cdot a(x) \le \alpha$ on L. For arbitrary positive n and sufficiently large T on the set $\{\tau(k_2) = \vartheta(k_1) = +\infty\}$, we have

$$\phi(X_n) - \phi(X_{m(n,T)}) \le k_2 - k_1$$

whence, by (8),

$$\begin{aligned}
\sum_{i=n}^{m(n,T)-1} \varepsilon_i(\phi) &= \phi(X_{m(n,T)}) - \phi(X_n) - \sum_{i=n}^{m(n,T)-1} c_i \nabla\phi(X_i) \cdot a(X_i) \\
&\ge \alpha \sum_{i=n}^{m(n,T)-1} c_i - (k_2 - k_1) \\
&\ge \alpha(T-1) - (k_2 - k_1) \ge 1
\end{aligned}$$

which contradicts Lemma 4 and establishes the result. ∎

The next result establishes a.s. convergence conditionally on the event that X_n never exits $L(k_2)$ for some $k_2 < K$.

LEMMA 7. *Suppose that* $k_0 < k < k_2 < K$. *Then, for all* $x \in L(k)$, X_n converges P_x-*a.s. to* Γ *on* $\{\tau(k_2) = +\infty\}$.

Proof. It suffices to show that lim sup $V(X_n) \leq k$ for any $k \in (k_0, k_2]$. Thus, given k, choose k_1 such that $k_0 < k_1 < k$. Consider the set $B = \{\tau(k_2) = +\infty$ and lim sup $V(X_n) > c\}$; by contradiction, suppose that $P_x(B) > 0$. Define recursively the entering and exit times on L

$$\vartheta_1 = \inf\{n: X_n \in L(k_1)\}, \qquad \tau_1 = \inf\{n > \vartheta_1: X_n \notin L(k)\}$$

and

$$\vartheta_i = \inf\{n > \tau_{i-1}: X_n \in L(k_1)\}, \qquad \tau_i = \inf\{n > \vartheta_{i-1}: X_n \notin L(k)\}$$

for $i \geq 2$. By Lemma 6, we know that on B the sequence of random variables $\{X_n\}$ must enter $L(k_1)$ and exit $L(k)$ infinitely often. Therefore, all these entering and exit times are finite. Define the function ϕ as in Lemma 6. Then

$$\begin{aligned}\sum_{i=\vartheta_n}^{\tau_n - 1} \varepsilon_i(\phi) &\geq \sum_{i=\vartheta_n}^{\tau_n - 1} c_i \nabla\phi(X_i) \cdot a(X_i) + \sum_{i=\vartheta_n}^{\tau_n - 1} \varepsilon_i(\phi) \\ &= \phi(X_{\tau_n}) - \phi(X_{\vartheta_n}) \\ &\geq k - k_1 > 0\end{aligned}$$

which contradicts Lemma 4 and establishes the result. ∎

Given Lemma 7, it is clear that a sufficient condition for the a.s. convergence of the algorithm (1) may be obtained by estimating $P\{\tau(k_2) = +\infty\}$. This is the object of the next lemma.

LEMMA 8 *Suppose that* $k_0 < k_1 < k_2 < K$. *Then, for all* $x \in L(k_1)$, *there exists a constant* K_5 *such that*

$$P_x\{\tau(k_2) = +\infty\} \geq 1 - K_5 \sum_{n=0}^{+\infty} c_n^2.$$

Proof. Define the function ϕ as in Lemma 6. As in the proof of Lemma 7, on the set $\{\tau(k_2) < +\infty\}$ we have

$$\sum_{n=0}^{\tau(k_2)-1} \varepsilon_n(\phi) \geq \phi(X_{\tau(k_2)}) - \phi(X_0) \geq k_2 - k_1 .$$

Multiplying both sides of the inequality by the indicator of the event $\{\tau(k_2) = +\infty\}$, we have

$$\begin{aligned}(k_2 - k_1)\mathbf{I}(\{\tau(k_2) = +\infty\}) &\leq \mathbf{I}(\{\tau(k_2) = +\infty\}) \left| \sum_{n=0}^{\tau(k_2)-1} \varepsilon_n(\phi) \right| \\ &\leq \sup_i \mathbf{I}(\{i \leq \tau(k_2)\}) \left| \sum_{n=0}^{i-1} \varepsilon_n(\phi) \right| \\ &\leq \sup_i \left| \sum_{n=0}^{i-1} \varepsilon_n(\phi) \right| .\end{aligned}$$

Taking the expected value of the square of the two sides of the inequality and applying Lemma 4, we obtain

$$P_x\{\tau(k_2) < +\infty\} \leq (k_2 - k_1)^{-2} K_4 \sum_{n=0}^{+\infty} c_n^2$$

whence the result follows. ■

Note that the proof of this lemma applies also to the "shifted" algorithm

$$X_{n+s+1} = X_{n+s} + c_{n+s} A(X_{n+s}, Z_{n+s+1})$$

with initial condition $X_s = x$, yielding

$$P\{\tau(k_2) = +\infty | X_s \in L(k_1)\} \geq 1 - K_5 \sum_{n=s}^{+\infty} c_n^2 .$$

Thus, if X_n visits $L(k_1)$ infinitely often, the probability of never exiting $L(k_2)$ is zero. With this in hand, we can finally prove the Convergence Lemma.

Proof of the Convergence Lemma. By (A6), for some $k_1 < k_2$, $L \subseteq L(k_1) \subset L(k_2)$. Hence, by Lemma 7, X_n converges a.s. on $\{\tau(k_2) = +\infty\}$. By the observation following Lemma 8, since X_n visits L infinitely often, the probability that eventually X_n never exits the set $L(k_2)$ is 1 and the result follows. ■

References

Aizpurua, J. M., Nieto, J., and Uriarte, J. R. (1990). "Choice Procedure Consistent with Similarity Relations," *Theory and Decision* **29,** 235–254.

Aizpurua, J. M., Ichiishi, T., Nieto, J., and Uriarte, J. R. (1993). "Similarity and Preference in the Space of Simple Lotteries," *J. Risk Uncertainty* **6,** 289–297.

Arnold, V. I., and Il'Yashenko, Y. S. (1988). "Ordinary Differential Equations," in *Dynamical Systems I* (D. V. Anosov and V. I. Arnold, Eds.), pp. 1–48. Berlin: Springer-Verlag.

Arthur, W. B., Ermoliev, Y. M., and Kaniovski, Y. (1983). "On Generalized Urn Schemes of the Polya Kind," *Cybernetics* **19,** 61–71 (originally published in *Kybernetika*).

Arthur, W. B., Ermoliev, Y. M., and Kaniovski, Y. (1986). "Strong Laws for a Class of Path-Dependent Stochastic Processes with Applications," in *Stochastic Optimization* (V. Arkin, A. Shiraev, and R. Wets, Eds.), pp. 287–300. Berlin: Springer-Verlag.

Aumann, R. J. (1990). "Nash Equilibria Are Not Self-Enforcing," in *Economic Decision Making: Games, Econometrics and Optimization* (J. J. Gabszewicz, J.-F. Richard, and L. A. Wolsey, Eds.), pp. 201–206. Amsterdam: North-Holland.

Benveniste, A., Métivier, M., and Priouret, P. (1990). *Adaptive Algorithms and Stochastic Approximations.* Berlin: Springer-Verlag.

Brown, G. W. (1951). "Iterative Solution of Games by Fictitious Play," in *Activity Analysis of Production and Allocation* (T. C. Koopmans, Ed.), pp. 374–376. New York: Wiley.

Cotter, K. D. (1991). "Similarity of Games with Incomplete Information," *J. Math. Econ.* **20,** 501–520.

Eichberger, J. (1992). "Nash Equilibrium and Learning," in *Recent Developments in Game Theory* (J. Borland, J. Creedy, and J. Eichberger, Eds.), pp. 94–124. Hants, England: Edward Elgar.

Fudenberg, D., and Kreps, D. M. (1992). "Learning and Equilibrium in Games," partial draft 0.64, MIT and Stanford University, January 1992.

Fudenberg, D., and Kreps, D. M. (1993). "Learning Mixed Equilibria," *Games Econ. Behav.* **5,** 320–367.

Gilboa, I., and Schmeidler, D. (1992). "Case-Based Decision Theory," Discussion paper 994, The Center for Mathematical Studies in Economics and Management Science, Northwestern University, May 1992.

Gilboa, I., and Schmeidler, D. (1993). "Case-Based Optimization," Discussion paper 1039, The Center for Mathematical Studies in Economics and Management Science, Northwestern University, April 1993.

Hale, J. and Koçak, H. (1991). *Dynamics and Bifurcations.* Berlin: Springer-Verlag.

Harsanyi, J. (1973). "Games with Randomly Disturbed Payoffs: A New Rationale for Mixed-Strategy Equilibrium Points," *Int. J. Game Theory* **2,** 1–23.

Krasovskiĭ, N. N. (1963). *Stability of Motion.* Stanford, CA: Stanford Univ. Press.

Kreps, D. M. (1990). *Game Theory and Economic Modelling.* Oxford: Clarendon Press.

Li Calzi, M. (1992). "Playing Games by Similarities," Ph.D. dissertation. Stanford University.

MIYASAWA, K. (1961). "On the Convergence of the Learning Process in a 2 × 2 Non-zero-Sum Two-Person Game," Research Memorandum 33. Princeton University.

NEVEL'SON, M. B., AND HAS'MINSKIĬ, R. Z. (1973). *Stochastic Approximation and Recursive Estimation*. Providence, R. I.: Amer. Math. Soc.

RUBINSTEIN, A. (1988). "Similarity and Decision-Making under Risk (Is There a Utility Theory Resolution to the Allais Paradox?)," *J. Econ. Theory* **46,** 145–153.

SHAPLEY, L. S. (1964). "Some Topics in Two-Person Games," in *Advances in Game Theory* (M. Dresher, L. Shapley, and A. Tucker, Eds.), pp. 1–28. Princeton, NJ: Princeton Univ. Press.