manuscript No.

Brown's Original Fictitious Play

Ulrich Berger

Vienna University of Economics, Department VW5
Augasse 2-6, A-1090 Vienna, Austria
e-mail: ulrich.berger@wu-wien.ac.at

March 2005

Abstract What modern game theorists describe as "fictitious play" is not the learning process George W. Brown defined in his 1951 paper. His original version differs in a subtle detail, namely the order of belief updating. In this note we revive Brown's original fictitious play process and demonstrate that this seemingly innocent detail allows for an extremely simple and intuitive proof of convergence in an interesting and large class of games: nondegenerate ordinal potential games.

Journal of Economic Literature classification numbers: C72.

Key words Fictitious Play, Learning Process, Ordinal Potential Games.

1 Introduction

Almost every modern textbook on game theory at least mentions the classical learning process known as *fictitious play* or as the *Brown-Robinson learning process*, introduced by Brown (1951) as an algorithm for finding the value of a zero-sum game, and first studied by Robinson (1951). Informal descriptions usually depict two players playing a finite game repeatedly. After arbitrary initial moves in the first round, in every round each player plays a myopic pure best response against the empirical strategy distribution of his opponent. It was once hoped that the players' beliefs, i.e. their empirical strategy distributions, always converge to the set of Nash equilibria in such a process.

While Robinson (1951) showed this to hold for zero-sum games, and Miyazawa (1961) extended it to 2×2 games, Shapley (1964) gave an example of a 3×3 game where the beliefs do not converge. Since then a bulk of literature has tried to identify classes of games where global convergence holds, including Milgrom and Roberts (1991), Krishna (1992), Hofbauer (1995), Monderer and Shapley (1996a, 1996b), Hon-Snir et al. (1998), Hahn (1999), Sela (2000), and Berger (2005). Others have found classes of games where fictitious play need not converge, e.g. Cowan (1992), Jordan (1993), Gaunersdorfer and Hofbauer (1995), Monderer and Sela (1996), Foster and Young (1998), or Krishna and Sjöström (1998). For overviews see Krishna and Sjöström (1997), Fudenberg and Levine (1998), or Hofbauer and Sigmund (2003).

All these authors employ the version of fictitious play informally described above, where players update their beliefs *simultaneously*.[1] It is notable, however, that this is *not* the original version of Brown's learning process. Brown (1951), routinely cited as the origin of fictitious play, clearly states that players update *alternatingly*.

Indeed, fictitious play was "invented" two years earlier by Brown (1949) in an unpublished RAND report. In this report, he also defines his algorithm with players updating their beliefs alternatingly, mentioning only briefly that a *minor variant* (p. 4) of the process is to let players update simul-

[1] Some of them actually deal with continuous-time fictitious play, also first described by Brown (1949), but introduce this process via discrete-time fictitious play with simultaneous updating.

taneously. To distinguish these two versions of the process, I will speak of AFP and SFP (for alternating fictitious play and simultaneous fictitious play), respectively.

AFP, the original version of fictitious play, seems to have gradually disappeared from the literature in the following decades. Robinson (1951) refers to Brown's 1949 report; she starts with a description of SFP, followed by a description of AFP as an *alternate notion* (p. 297) of the process. However, she makes clear that her proofs and results hold for either version of fictitious play. Luce and Raiffa (1957, appendix A6.9) and Karlin (1959, chapter 6.5) exclusively describe Brown's original version, AFP. A few years later, Shapley (1964) remarks that *various conventions can be adopted with regard to* [. . .] *simultaneous vs. alternating moves* (p. 24) and then goes on working with SFP *for simplicity* (p. 25). To my knowledge, all the later literature on fictitious play works with SFP, and does not even mention the possibility of alternate updating.

Indeed, SFP may be considered a simpler object of study than AFP, since players are treated symmetrically under SFP, and this usually enhances analytical convenience. However, as I show below, simultaneous updating may also generate subtle problems which do not arise under AFP. In the following I give an extremely simple and intuitive proof of global convergence of AFP in nondegenerate ordinal potential games — a result which hitherto remains unproven for SFP.

2 Definitions

We start with the notation and some definitions:

Let (A, B) be an $n{\times}m$ bimatrix game, i.e. a finite two-player game, where player 1, the row player, has pure strategies $i \in N = \{1, 2, \ldots, n\}$, and player 2, the column player, has pure strategies $j \in M = \{1, 2, \ldots, m\}$. A and B are the $n{\times}m$ payoff matrices for players 1 and 2. If player 1 chooses i and player 2 chooses j, the payoffs to players 1 and 2 are a_{ij} and b_{ij}, respectively. The sets of mixed strategies of player 1 and 2 are the probability simplices S_n and S_m, respectively, and mixed strategies are written as column vectors. With a little abuse of notation we will not distinguish between a pure strategy and the corresponding mixed strategy representation as a unit vector.

Player 1's expected payoff for playing i against player 2's mixed strategy $\mathbf{y}$ is $(A\mathbf{y})_i$. Analogously, $(B^t\mathbf{x})_j$ (where the superscript t denotes the transpose of a matrix) is the expected payoff for player 2 playing j against the mixed strategy $\mathbf{x}$. If both players use mixed strategies $\mathbf{x}$ and $\mathbf{y}$, respectively, the expected payoffs are $\mathbf{x}{\cdot}A\mathbf{y}$ to player 1 and $\mathbf{y}{\cdot}B^t\mathbf{x}$ to player 2, where the dot denotes the scalar product of two vectors. We denote by $BR_1(.)$ and $BR_2(.)$ the players' pure strategy best response correspondences. A strategy profile $(\mathbf{x}^*, \mathbf{y}^*)$ is a Nash equilibrium if for all i, $x_i^* > 0$ implies $i \in BR_1(\mathbf{y}^*)$ and for all j, $y_j^* > 0$ implies $j \in BR_2(\mathbf{x}^*)$. It is called a pure Nash equilibrium, if $\mathbf{x}^*$ and $\mathbf{y}^*$ are pure strategies.

2.1 Nondegenerate Games

It has been shown by Monderer and Sela (1996) that even $2{\times}2$ games may exhibit SFP processes with nonconvergent beliefs, if the game is degenerate.[2]

[2] Miyazawa's (1961) proof relies on a particular tie-breaking rule.

Their example also holds for AFP. Hence we must restrict ourselves to nondegenerate games if we want to obtain convergence results.

Definition 1 *We call a bimatrix game* (A, B) degenerate*, if for some* $i, i' \in N$*, with* $i \neq i'$*, there exists* $j \in M$ *with* $a_{i'j} = a_{ij}$*, or if for some* $j, j' \in M$*, with* $j \neq j'$*, there exists* $i \in N$ *with* $b_{ij'} = b_{ij}$*. Otherwise, the game is said to be* nondegenerate*.*

2.2 Potential Games

Monderer and Shapley (1996b) define several classes of games with a so-called *potential*. The class of *ordinal potential games* is an important class of games. It contains the class of *weighted potential games*, which include Rosenthal's (1973) *congestion games* and have previously been called *rescaled partnership games* by Hofbauer and Sigmund (1988).

Definition 2 *A bimatrix game* (A, B) *is an* ordinal potential game*, if there exists an ordinal potential function, i.e., a function* $F: N \times M \to \mathbb{R}$*, such that for all* $i, i' \in N$ *and* $j, j' \in M$*,*

$$a_{i'j} - a_{ij} > 0 \quad \Leftrightarrow \quad F(i', j) - F(i, j) > 0,$$

and

$$b_{ij'} - b_{ij} > 0 \quad \Leftrightarrow \quad F(i, j') - F(i, j) > 0.$$

Definition 3 *A bimatrix game* (A, B) *is a* weighted potential game*, if there exist positive weights* w_1 *and* w_2 *and a function* $F: N \times M \to \mathbb{R}$*, such that for all* $i, i' \in N$ *and* $j, j' \in M$*,*

$$a_{i'j} - a_{ij} = w_1 [F(i', j) - F(i, j)]$$

and

$$b_{ij'} - b_{ij} = w_2[F(i,j') - F(i,j)].$$

Clearly, a weighted potential game has an ordinal potential. Note that the former imposes a cardinal condition on the payoffs, while the latter requires only an ordinal condition. Hence the class of ordinal potential games is a "large" class from a measure theoretic viewpoint, while the class of weighted potential games is negligible.[3] For weighted potential games Monderer and Shapley (1996b) proved global convergence of SFP to the equilibrium set. However, for ordinal potential games no such result is known.

Monderer and Shapley (1996b) also define *improvement paths* and games with the *finite improvement property*. We extend this definition by introducing *improvement steps*.

Definition 4 *For a bimatrix game (A, B), define the following binary relation on $N \times M$:*
$(i,j) \rightarrow (i',j') \Leftrightarrow$ *($i = i'$ and $b_{ij'} > b_{ij}$) or ($j = j'$ and $a_{i'j} > a_{ij}$). If $(i,j) \rightarrow (i',j')$, we say that this is an* improvement step. *An* improvement path *is a (finite or infinite) sequence of improvement steps $(i_1,j_1) \rightarrow (i_2,j_2) \rightarrow (i_3,j_3) \rightarrow \cdots$ in $N \times M$. An improvement path $(i_1,j_1) \rightarrow \cdots \rightarrow (i_k,j_k)$ is called an* improvement cycle*, if $(i_k,j_k) = (i_1,j_1)$. A bimatrix game is said to have the* finite improvement property *(FIP), if every improvement path is finite, i.e., if there are no improvement cycles.*

[3] The set of $n \times m$ bimatrix games can be identified with the Euclidean space $\mathbb{R}^{2nm}$. It can be shown that within this space, the set of ordinal potential games contains an open set, while the set of weighted potential games is a null set (has Lebesgue–measure zero), if $n \geq 2$ and $m \geq 3$.

It is clear that every nondegenerate game with an ordinal potential has the FIP. Monderer and Shapley (1996b) show that also the opposite direction holds:

Lemma 1 *A nondegenerate bimatrix game has the FIP if and only if it is an ordinal potential game.*

2.3 Fictitious Play

The following definition corresponds to Brown's original version of fictitious play, where players update alternatingly:

Definition 5 *For the* $n \times m$ *bimatrix game* (A, B)*, the sequence* $(i_t, j_t)_{t \in \mathbb{N}}$ *is an* alternating fictitious play process (AFP process)*, if* $i_1 \in N$ *and for all* $t \in \mathbb{N}$*,*

$$i_{t+1} \in BR_1(\mathbf{y}(t)) \;\; \textit{and} \;\; j_t \in BR_2(\mathbf{x}(t)),$$

where the beliefs $\mathbf{x}(t)$ *and* $\mathbf{y}(t)$ *are given by*

$$\mathbf{x}(t) = \frac{1}{t}\sum_{s=1}^{t} i_s \;\; \textit{and} \;\; \mathbf{y}(t) = \frac{1}{t}\sum_{s=1}^{t} j_s.$$

The definition of SFP differs from that of AFP only in the order of updating:

Definition 6 *For the* $n \times m$ *bimatrix game* (A, B)*, the sequence* $(i_t, j_t)_{t \in \mathbb{N}}$ *is a* simultaneous fictitious play process (SFP process)*, if* $(i_1, j_1) \in N \times M$ *and for all* $t \in \mathbb{N}$*,*

$$i_{t+1} \in BR_1(\mathbf{y}(t)) \;\; \textit{and} \;\; j_{t+1} \in BR_2(\mathbf{x}(t)),$$

where $\mathbf{x}(t)$ *and* $\mathbf{y}(t)$ *are defined as above.*

If a fictitious play process (AFP or SFP) converges, it must be constant from some stage on, implying that the limit is a pure Nash equilibrium. Even if the process does not converge, it is easily established that if the beliefs converge, then the limit must be a Nash equilibrium (which need not be pure, however).

Note that the beliefs can be updated recursively. The belief of a player in round $t+1$ is a convex combination of his belief in round t and his opponent's move in round $t+1$:

$$\mathbf{x}(t+1) = \frac{t}{t+1}\mathbf{x}(t) + \frac{1}{t+1} i_{t+1}, \tag{1}$$

$$\mathbf{y}(t+1) = \frac{t}{t+1}\mathbf{y}(t) + \frac{1}{t+1} j_{t+1}. \tag{2}$$

When a player plays some pure strategy at time t and some different pure strategy at time $t+1$, we say that he *switches* at time t.

Definition 7 *Let $(i,j) = (i_t, j_t)$ and $(k,l) = (i_{t+1}, j_{t+1})$. We say that player 1* switches *from i to k at time t, if $i \neq k$. Analogously, player 2* switches *from j to l at time t, if $j \neq l$.*

3 The Result

Assume there is a switch from (i,j) to (k,l) at time $t > 1$, i.e. $i \neq k$ or $j \neq l$. Then i is a best response to $\mathbf{y}(t-1)$ and k is a best response to $\mathbf{y}(t)$, implying $i \cdot A\mathbf{y}(t-1) - k \cdot A\mathbf{y}(t-1) \geq 0$ and $i \cdot A\mathbf{y}(t) - k \cdot A\mathbf{y}(t) \leq 0$. Hence $(i-k) \cdot A\mathbf{y}(t-1) \geq 0$ and $(i-k) \cdot A\mathbf{y}(t) \leq 0$. Left–multiplication with equation (2) shows that we have $(i-k) \cdot Aj \leq 0$, which simply means $a_{ij} \leq a_{kj}$. The same argument applied to player 2 shows that $b_{kj} \leq b_{kl}$.

This simple observation is nothing but the AFP analog of Monderer and Sela's (1997) *Improvement Principle* for SFP. For nondegenerate games we can state it as follows.

Lemma 2 *If an AFP process for a nondegenerate game contains a switch from* (i, j) *to* (k, l)*, then there is an improvement path from* (i, j) *to* (k, l)*.*

Proof Given our observation above, this follows directly from the definitions of nondegeneracy and improvement paths. □

The main result is an immediate consequence of Lemma 2

Theorem 1 *Let* (A, B) *be a nondegenerate ordinal potential game. Then every AFP process converges to a pure Nash equilibrium.*

Proof Assume it does not. In a nonconvergent AFP process there are infinitely many switches. Since there are only finitely many pure strategy pairs, however, at least two such pairs are played infinitely often. Hence there must be a sequence of switches leading from one of these pairs to the other and back again. By Lemma 2, this means that there is an improvement cycle. By Lemma 1 then, the game cannot have an ordinal potential, which contradicts the assumption. □

Intuitively, in an AFP process in a nondegenerate game a switch from one pure-strategy pair to another pure-strategy pair implies that there is an improvement path from the former to the latter. This improvement path consists of one or two improvement steps, depending on whether only one player switches, or both. Hence AFP processes essentially follow improvement paths in nondegenerate games. If a game has the FIP, as nondegenerate ordinal potential games do, then the process cannot involve cycles, and must terminate in a pure Nash equilibrium.

4 Discussion

In nondegenerate ordinal potential games, AFP processes cannot cycle, because there are no improvement cycles. Why has this simple fact remained unnoticed in all the studies of fictitious play? The obvious reason seems to be that these studies work with SFP, and things are more difficult when players update simultaneously. Indeed, if both players switch at the same time in an SFP process, there need *not* be an improvement path from the old to the new pure-strategy profile. To see this, imagine a 2×2 coordination game and two SFP players starting from a pure-strategy profile where they miscoordinate. In the second round they will again miscoordinate, since both switch to their other strategy. Clearly, however, there is no improvement path connecting one miscoordination profile to the other.

It is easy to extend our theorem to SFP processes where from some time on the players never switch simultaneously.[4] The problem is to rule out processes where players continue to do this infinitely often. An example of such a nonconvergent SFP process was constructed by Foster and Young (1998). This is not a counterexample, however, since their game does not have an ordinal potential. While I conjecture our theorem to hold for SFP generically, this question remains open.

References

1. Berger, U. (2004). Two more classes of games with the fictitious play property. Mimeo, Vienna University of Economics.
2. Berger, U. (2005). Fictitious play in $2\times n$ games. *Journal of Economic Theory* **120**, 139-154.

[4] It can also be shown that the theorem holds for a continuous-time fictitious play process that does not involve mixed strategies, see Berger (2004).

3. Brown, G. W. (1949). Some notes on computation of games solutions. Report P-78, The Rand Corporation.
4. 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, Wiley: New York.
5. Cowan, S. (1992). *Dynamical systems arising from game theory.* Dissertation, UC Berkeley.
6. Foster, D. P. and Young, H. P. (1998). On the nonconvergence of Fictitious Play in coordination games. *Games and Economic Behavior* **25**, 79-96.
7. Fudenberg, D. and Levine, D. K. (1998). *The theory of learning in games.* MIT Press, Cambridge, MA.
8. Gaunersdorfer, A. and Hofbauer, J. (1995). Fictitious play, Shapley polygons, and the replicator equation. *Games and Economic Behavior* **11**, 279-303.
9. Hahn, S. (1999). The convergence of Fictitious Play in 3×3 games with strategic complementarities. *Economics Letters* **64**, 57-60.
10. Hofbauer, J. (1995). Stability for the best response dynamics. Mimeo, University of Vienna.
11. Hofbauer, J. and Sigmund, K. (1988). *The theory of evolution and dynamical systems.* Cambridge University Press, Cambridge, UK.
12. Hofbauer, J. and Sigmund, K. (2003). Evolutionary game dynamics. *Bulletin of the American Mathematical Society* **40**, 479-519.
13. Hon-Snir, S., Monderer, D., and Sela, A. (1998). A learning approach to auctions. *Journal of Economic Theory* **82**, 65-88.
14. Jordan, J. S. (1993). Three problems in learning mixed-strategy Nash equilibria. *Games and Economic Behavior* **5**, 368-386.
15. Karlin, S. (1959). *Mathematical methods and theory in games, programming, and economics.* Volume 1, Addison-Wesley, Reading, MA.
16. Krishna, V. (1992). Learning in games with strategic complementarities. Mimeo, Harvard University.
17. Krishna, V. and Sjöström, T. (1997). Learning in games: Fictitious play dynamics. In: Hart, S. and Mas-Colell, A. (eds.). Cooperation: Game-theoretic approaches. Springer Verlag, Berlin, Heidelberg.
18. Krishna, V. and Sjöström, T. (1998). On the convergence of fictitious play. *Mathematics of Operations Research* **23**, 479-511.

19. Luce, R. D., and Raiffa, H. (1957). *Games and decisions.* Wiley, New York.
20. Milgrom, P. and Roberts, J. (1991). Adaptive and sophisticated learning in normal form games. *Games and Economic Behavior* **3**, 82-100.
21. Miyazawa, K. (1961). On the convergence of the learning process in a 2×2 non-zero-sum two-person game. Econometric Research Program, Princeton University, Research Memorandum No. 33.
22. Monderer, D. and Sela, A. (1996). A 2×2 game without the fictitious play property. *Games and Economic Behavior* **14**, 144-148.
23. Monderer, D. and Sela, A. (1997). Fictitious play and no-cycling conditions. Mimeo, The Technion.
24. Monderer, D. and Shapley, L. S. (1996a). Fictitious play property for games with identical interests. *Journal of Economic Theory* **68**, 258-265.
25. Monderer, D. and Shapley, L. S. (1996b). Potential games. *Games and Economic Behavior* **14**, 124-143.
26. Robinson, J. (1951). An iterative method of solving a game. *Annals of Mathematics* **54**, 296-301.
27. Rosenthal, R. W. (1973). A class of games possessing pure-strategy Nash equilibria. *International Journal of Game Theory* **2**, 65-67.
28. Sela, A. (2000). Fictitious play in 2×3 games. *Games and Economic Behavior* **31**, 152-162.
29. Shapley, L. S. (1964). Some topics in two-person games. *Annals of Mathematical Studies* **5**, 1-28.