Fictitious Play in Coordination Games[1]

Aner Sela[2] and Dorothea K Herreiner[3]

Projektbereich B
Discussion Paper No B-423
December 1997

Abstract. We study the Fictitious Play process with bounded and unbounded recall in pure coordination games for which failing to coordinate yields a payoff of zero for both players. It is shown that every Fictitious Play player with bounded recall may fail to coordinate against his own type. On the other hand, players with unbounded recall are shown to coordinate (almost surely) against their own type as well as against players with bounded recall. This implies that a FP player's realized average utility is (almost surely) at least as large as his minmax payoff in 2x2 coordination games.

JEL Classification Numbers: C72, D83.

Keywords: Learning, Fictitious Play, Coordination Games, Pure Coordination Games.

---

[1] We would like to thank Karl Schlag and Dov Monderer for helpful comments. Financial support from the Deutsche Forschungsgemeinschaft, Sonderforschungsbereich 504 at the University of Mannheim, and from the Deutsche Forschungsgemeinschaft, Sonderforschungsbereich 303 at the University of Bonn is gratefully acknowledged.

[2]Department of Economics, University of Mannheim, Seminargebaeude A5, D-68131 Mannheim, Germany.

[3] Wirtschaftstheorie III, Department of Economics, University of Bonn, Adenauerallee 24-26, 53113 Bonn, Germany.

1. **Introduction**

Consider 2 players engaged in the repeated play of a finite game in strategic form. In each period, every player observes the history of past moves, and forms a belief about the next move of his opponent. He then chooses an action that is a myopic pure best reply, that is, an action that maximizes his expected payoff in the next period according to his belief. We call such a process a myopic learning process. One of the most common myopic learning models is the *Fictitious Play* (FP) process proposed by Brown in 1951. In the FP process, each player believes that each one of his opponents is using a stationary mixed strategy which is the empirical distribution of this opponent's past actions. A variation of the FP process is the Fictitious Play process with bounded recall or short, the *Bounded Fictitious Play* (BFP). In the BFP process the players' beliefs are based only on a finite number of the most recent periods (players' recall length). When a player has multiple best replies to his belief we assume that he chooses each of them with a "strictly positive probability". Most research on the FP and similar processes has focused on the question whether players' beliefs approach equilibrium (see for example, Robinson (1951), Miyasawa (1961), Monderer/Shapley (1996)). However, this notion of convergence does not seem appropriate for coordination games as was shown by Fudenberg and Kreps (1993): consider a game in which each agent gets 1 if the players coordinate on the same action, and if they fail to coordinate, each gets 0. In this case the players' beliefs may approach the mixed equilibrium strategy, although both players never coordinate, such that both players receive payoff 0 in every period. From the standpoint of players learning how their opponents behave, this sort of example does not seem a very satisfactory notion of approaching equilibrium. This led us to consider the question whether the players coordinate or not in pure coordination games. For illustrative purposes we study the FP in $2x2$ coordination games. We limit our analysis to $2x2$ pure coordination games[4] since every $2x2$ coordination game is best response equivalent to $2x2$ pure coordination game. The results of this paper are easily extended to the class of all NxN pure coordination games.

---

[4] A pure coordination game is a coordination game for which the payoffs off the diagonal are zero.

For 2x2 pure coordination games examples are known where BFP players converge to equilibrium (see Li (1994)[5], proposition 3). We show that in a generic class of games, every BFP player may never coordinate against another BFP player. On the other hand, a FP player, independent of initial beliefs, coordinates almost surely from some stage on against any player with unbounded (FP) or bounded recall (BFP)[6]. With this we can address the common objection to the FP process in 2x2 coordination games (Fudenberg and Kreps (1993), and Young (1993)) namely that a FP player's realized average payoff might be less than his minmax payoff[7]. As our results show, this cannot happen (almost surely) in $2x2$ coordination games.

2. Preliminaries

Let Γ denote the set of all $2x2$ coordination games in strategic form. A game G is a *coordination game* if the players have the same number of strategies, which can be indexed so that it is always a strict Nash equilibrium for both players to play strategies having the same index. Note that every $2x2$ coordination game is best response equivalent to $2x2$ pure coordination game for which the payoffs off the diagonal are zero.[8]

There are two players denoted by 1 and 2. For every player i, the other player is denoted by $-i$. The strategy set of player 1 is $S^1 = \{a_1, b_1\}$ and the strategy set of player 2 is $S^2 = \{a_2, b_2\}$. Let $S = S^1 x S^2$. Player i's payoff function is $U^i : S \to R$, where R denotes the set of real numbers. The set of mixed strategies of player i is $\Delta^{\mathrm{i}} = \Delta(\mathrm{S}^{\mathrm{i}})$. Let $\Delta = \Delta^1 x \Delta^2$. We think of $x^i \in S^i$ as being an extreme point in Δ^i.

---

[5] Unfortunately, his results about the convergence to risk-dominated equilibria are flawed. See footnote 12 for further comments.

[6] There is a related literature which analyzes the impact of the finiteness of memory on equilibrium play. Young (1993) studies a different kind of FP process of finite memory and shows that players coordinate after a finite number of periods. Other literature does not use the FP process but relies on finite automata to model memory of players. See Neyman (1985) and Rubinstein (1986) for some earlier work. Ben-Porath (1986), Gilboa/Samet (1989), and Lehrer (1988) analyze the strategic advantage the player with the longer or with the shorter recall length may have.

[7] v^1 is a minmax payoff of player 1 if, $v^1 = \min_{x^2 \in \Delta(S^2)} \max_{x^1 \in \Delta(S^1)} U^1(x^1, x^2)$.

[8] It can be easily verified that in every NxN pure coordination game every myopic learning player discussed in this paper (FP or BFP player) uses from some stage on at most two actions from his

A game $G \in \Gamma$ is *non-degenerate* if for every i, $U^i(\cdot, x^{-i})$ is a one-to-one function for every $x^{-i} \in S^{-i}$. Note that every $2x2$ coordination game is non-degenerate game.

A *Fictitious Play* (FP) *player*, observes at each point of time the history of past moves and forms a belief about his opponent's action in the next period. He believes that his opponent is using a stationary mixed strategy which is the empirical distribution of his opponents' past actions. That is,

$b^i(t) = \frac{1}{t}\left(\sum_{s=1}^{t-1} x^{-i}(s) \; + b^i(0)\right)$, $t > 0$, where $b^i(t) \in \Delta^{-i}$ is the belief of player i about the strategy of player $-i$ at stage t, and $b^i(0)$ is the initial belief of player i. $b^i_m(t),\ m \in S^{-i}$, is player i's belief about strategy m as his opponent's choice at stage t. A FP player chooses at each stage $t \geq 1$ an action $x^i(t)$ that maximizes his payoff in period t, i.e. $x^i(t)$ is a best reply to his belief $b^i(t)$.

A *bounded fictitious play (BFP) player* with recall $k \geq 1$ believes that his opponent is using a stationary mixed strategy which is the empirical distribution of his opponent's past actions over the k most recent periods (k denotes the player's *length of recall*), that is,

$$b^i(t) = \frac{1}{t}\left(\sum_{s=1}^{t-1} x^{-i}(s) \; + b^i(0)\right), \; 1 \leq t \leq k,$$

And,

$$b^i(t) = \frac{1}{k} \sum_{s=t-k}^{t-1} x^{-i}(s) \;, \; t > k.$$

A BFP player chooses in each period $t \geq 1$, an action $x^i(t)$ which is a best reply according to his belief $b^i(t)$.

We assume the following tie-breaking rule: if a player is indifferent among some of his best replies at stage t, he uses each of them with a positive probability.

If at some stage FP players choose a pure equilibrium profile, then they will choose this equilibrium profile from this stage on (see Monderer/Sela (1993)). This result does not necessarily hold for BFP players. If player i has a tie between his best replies when a pure equilibrium profile is chosen, it depends on what action $x^{-i}(t-k+1)$ is

strategy set, such that the players play actually from some stage on a $2x2$ coordination subgame. Hence, all our results hold for all NxN pure coordination games.

dropped from the memory of length k at stage t. If that action $x^{-i}(t-k+1)$ is the same as his opponent's action in the pure equilibrium profile then the beliefs at stage $t+1$ is the same as at stage t and due to the tie, an action different from the pure equilibrium profile might be chosen. If the dropped action $x^{-i}(t-k+1)$ is not the pure equilibrium profile then the belief at stage $t+1$ is different and therefore the pure equilibrium profile will be chosen forever.

3. **Results**

First we show that BFP players may (depending on the players' initial beliefs) never coordinate in an open and dense subset of $2x2$ coordination games.

Proposition 1: For every $k \geq 1$, there exists an open and dense subset of $2x2$ coordination games, in which BFP players with recall k, may never coordinate.

Proof: Let $k \geq 1$ and let $\overline{\Gamma}$ be the set of all games $\overline{G}$ of the form, $\begin{array}{ccc} & a_2 & b_2 \\ a_1 & (x,z) & (0,0) \\ b_1 & (0,0) & (y,w) \end{array}$ $x, y, z, w > 0$.[9] The games in $\overline{\Gamma}$ for which either of the following conditions hold is an open and dense subset of $\overline{\Gamma}$:

1. $\dfrac{x}{y} \in (k-1, \infty)$ and $\dfrac{w}{z} \in (k-1, \infty)$.

2. $\dfrac{x}{y} \in (s, s+1)$ and $\dfrac{w}{z} \in (s, s+1)$, for some $\mathrm{s} = 1,2,3,\ldots k-2$,

The players may fail to coordinate in either of the cases 1 or 2. We show just the first case. The other case can be verified by the same reasoning. Assume that $\dfrac{x}{y} \in (k-1, \infty)$ and $\dfrac{\mathrm{w}}{\mathrm{z}} \in (k-1, \infty)$, and $b(0) = (b^1(0), b^2(0)) = ((1,0),(0,1))$.[10] Every $k+1$ periods the same path is generated by the process as follows: During the first k periods the players play (a_1, b_2) since $b^1_{b_2}(t) \leq \dfrac{k-1}{k}$ and $b^2_{a_1}(t) \leq \dfrac{k-1}{k}$.

[9] Note that every 2x2 coordination game is best response equivalent to a game from the set of games $\overline{G}$.

[10] The proof goes through for an open and dense subset of initial beliefs.

In the $(k+1)$-th stage, the players have pure beliefs, that is, $b^1_{b_2}(t) = b^2_{a_1}(t) = 1$, and therefore they play (b_1, a_2) in this stage. Hence, the players never coordinate. ■

The meaning of Proposition 1 is that when BFP players have "similar and opposite preferences", i.e. the payoff ratio between the two Nash equilibria is approximately the same for both players and each player prefers a different Nash equilibrium, then the players never coordinate.

Knowing that all BFP players may fail to coordinate for any (finite) recall length, poses the question: Is there a class of pure coordination games in which a FP player may fail to coordinate as well ?

The following theorem shows that if a FP plays against a BFP player, they coordinate almost surely in every $2x2$ coordination game.

Theorem A: Let player 1 be a FP player and player 2 be a BFP player. Then, in every 2x2 pure coordination game, independent of initial beliefs and actions, the players coordinate almost surely from some stage on.

Proof: Let $\overline{G}$ be a $2x2$ pure coordination game, such that $(p = (p_{a_1}, p_{b_1}), q = (q_{a_2}, q_{b_2}))$ is the mixed strategy equilibrium of $\overline{G}$. We assume that the players never coordinate. Note that if the players coordinate infinitely often, they coordinate (almost surely) from some stage on for ever.

Suppose that along the path generated by the players, the players do face only finitely often a tie. Due to the tie-breaking rule such a path occurs with probability one. Without loss of generality, we can concentrate on the case where players never face a tie[11].

Denote the number of times $(m,n) \in S$ was played up to stage t by $T_{(m,n)}(t)$, $m \in S^1$, $n \in S^2$.

Now, since one of the players has a finite recall, the players generate a repeated cycle of Y stages, such that the empirical distribution of the players' actions converges as follows: $\lim_{t\to\infty} T_{(a_1,b_2)}(t) = q_b Y = z$ and $\lim_{t\to\infty} T_{(a_2,b_1)}(t) = q_a Y = w$ where w and z are

[11] Whether there are many or zero ties does not matter, since the process is deterministic once all ties are resolved.

integers. Otherwise the FP player would prefer from some stage on one of his actions, which is the best reply to his belief such that the players will coordinate for ever. Suppose that the players switch from (a_1,b_2) to (b_1,a_2) at stage t_0, and without loss of generality suppose that the players played k times (a_1,b_2) up to stage t_0. Thus by the FP's behavior rules:

$$\frac{k}{t_0} < \frac{z}{Y} \text{ and } \frac{k+1}{t_0+1} > \frac{z}{Y}.$$

This implies that ,

$$\frac{k+1}{t_0+1} - \frac{z}{Y} < \frac{k+1}{t_0+1} - \frac{k}{t_0} < \frac{k+1}{t_0+1} - \frac{k}{t_0+1} < \frac{1}{t_0+1}.$$

Note that if the players switch from (a_1,b_2) to (b_1,a_2) at stage t, Player 1's best response in the next period is necessarily a_1. Thus, the players switch back at stage $t+1$ from (b_1,a_2) to (a_1,b_2), otherwise they would coordinate at $t+1$. Hence, by induction on t and the assumption that the limit distribution $(\frac{z}{Y},\frac{w}{Y})$ is never obtained in a finite stage, we obtain that for every $t > t_0$,

$\frac{T_{(a_1,b_2)}(t)}{t} - \frac{z}{Y} < \frac{1}{t}$. By the same argument, for every $t > t_0$, $\frac{z}{Y} - \frac{T_{(a_1,b_2)}(t)}{t} < \frac{1}{t}$.

Therefore, for every $t > t_0$, $\left|\frac{T_{(a_1,b_2)}(t)}{t} - \frac{z}{Y}\right| < \frac{1}{t}$.

Let $t = rY$ for some large integer r. Clearly $\frac{T_{(a_1,b_2)}(t)}{rY} = \frac{v}{rY}$ for some integer v. By the last inequality,

$$\left|\frac{T_{(a_1,b_2)}(t)}{rY} - \frac{z}{Y}\right| = \left|\frac{v}{rY} - \frac{rz}{rY}\right| = \frac{|v-rz|}{rY} < \frac{1}{rY}.$$

Since both v and rz are integers, $v - rz$ must be equal to zero, and therefore the limit distribution $(\frac{z}{Y},\frac{w}{Y})$ is obtained at a finite step. A contradiction to our assumption. ■

The following result shows that a FP player coordinates not only against every BFP player, but also against his own type in almost every $2x2$ pure coordination game.

Theorem B: FP players, independent of initial beliefs, coordinate almost surely from some stage on in almost every 2x2 pure coordination game.[12]

Proof: Let $\overline{G}$ be a $2x2$ pure coordination game, where the mixed strategy equilibrium of the game is $(p=(p_{a_1},p_{b_1}),q=(q_{a_2},q_{b_2}))\in\Delta^1 x\Delta^2$, where $p_{a_1}\neq q_{b_2}$. Let player 1 and 2 be FP players. Assume that the players never coordinate. Note that if FP players coordinate once at some stage, they coordinate always from this stage on.

Now, suppose that along the path generated by the players, the players do face only finitely often a tie. Due to the tie-breaking rule such a path occurs with probability one. Without loss of generality, we can concentrate on the case where players never face a tie. Hence, the players play at every stage either (a_1,b_2) or (b_1,a_2). By Monderer and Shapley (1996), a FP player's belief converges to his opponent's Nash equilibrium strategy in every 2x2 game[13] with the diagonal property[14]. Since every 2x2 coordination game is a game with the diagonal property, Monderer/Shapley's result holds for 2x2 coordination games. By our assumption of failure of coordination and Monderer/Shapley's result, we obtain that the belief of player i converges to player j's mixed strategy equilibrium, i.e. $\lim_{t\to\infty} b^1_{b_2}(t)=q_b$, and $\lim_{t\to\infty} b^2_{a_1}(t)=p_a$. However, $\lim_{t\to\infty}\left|b^1_{b_2}(t)-b^2_{a_1}(t)\right|=0$, and $p_a\neq q_b$. Hence, we obtain a contradiction to our assumption about failure of coordination. ■

[12] Li (1994) claims that FP players always coordinate on the risk-dominant equilibrium. He first claims to show that BFP players cannot coordinate on the risk-dominated equilibrium (his proof erroneously relies on beliefs which imply coordination although he considers failure of coordination). Combining this with the result of Miyasawa (1961) he deduces that FP players must coordinate on the risk dominant equilibrium. He completely ignores the possibility of convergence to (the mixed) equilibrium with failure of coordination. We show in Theorem A that almost surely there is no convergence to equilibrium with failure of coordination.

[13] Miyasawa was the first who proved that FP players' beliefs converge to a Nash equilibrium profile in every 2x2 game, but he proved it under the assumption of a specific tie breaking rule. Monderer/Sela (1996) showed that if we assume another tie-breaking rule, then there is a $2x2$ game in which a belief sequence of FP players does not converge to equilibrium profile. However, Monderer/Sela's result holds for a degenerate game only.

[14] We say that a game $G=(D,E)=(d(m,n),e(m,n))^2_{m,n=1,2}$ has the diagonal property if $f\neq 0$ and $g\neq 0$, where $f=e(1,1)-e(2,1)-e(1,2)+e(2,2)$ and $g=d(1,1)-d(2,1)-d(1,2)+d(2,2)$.

By Theorem A if a FP player plays against a BFP player, they coordinate in every $2x2$ game. By Theorem B if a FP player plays against another FP player, they coordinate in almost every game. Theorem B excludes a class of games with measure 0, which is the class of $2x2$ coordination games where players have identical mixed strategy equilibrium profiles, i.e. $(p,q) \in \Delta^1 x \Delta^2$, and $p = q$. This class of $2x2$ games with measure zero includes the symmetric example of Fudenberg and Kreps (1993) in which each player gets 1 if they coordinate and 0 if the fail to coordinate. To prevent convergence in Fudenberg and Kreps (1993) and Young (1993) either players' beliefs or payoffs have to be irrational numbers. If beliefs and payoffs are rational numbers, then, by using the idea of the proof of Theorem A, it can be shown that FP players coordinate almost surely from some stage on in every 2x2 coordination game, even if $p = q$.

4. References

Ben-Porath, E. (1993): "Repeated Games with Finite Automata". in *Journal of Economic Theory,* 59, 17-32.

Brown, G. W. (1951): "Iterative Solution of Games by Fictitious Play". in *Activity Analysis of Production,* Wiley, New York.

Fudenberg, D. and D. Kreps (1993): "learning Mixed Equilibria". *Games and Economic Behavior,* 5, 320-367.

Fudenberg, D. and D. Levine (1997): Theory of Learning in Games. *Electronic document*. UCLA: http://levine.sscnet.ucla.edu, accessed on 27 August 1997.

Gilboa, I., Samet, D. (1989): "Bounded versus Unbounded Rationality: The Tyranny of the Weak". *Games and Economic Behavior,* 1, 213-221.

Lehrer, E. (1988): "Repeated Games with Stationary Bounded Recall Strategies". *Journal of Economic Theory,* 46, 130-144.

Li, S. (1994). "Dynamic Stability and Learning Processes in 2x2 Coordination Games". *Economics Letters,* 46(2), 105-111.

Miyasawa, K. (1961): "On the Convergence of the Learning Process in 2x2 Non-Zero-Sum Two-Person Game". *Economic Research Program,* Princeton University, Research memorandum No. 33.

Monderer, D., and A. Sela (1993): “Fictitious Play and No-Cycling Conditions“. Mimeo.

Monderer, D., and A. Sela (1996): “A 2x2 Game Without the Fictitious Play Property”. *Games and Economic Behavior,* 14, 144-148.

Monderer, D. and L. S. Shapley (1996): “Fictitious Play Property for Games with Identical Interests”. *Journal of Economic Theory*, 68, 258-265.

Neyman, A. (1985): “Bounded Complexity Justifies Cooperation in the Finitely Repeated Prisoner’s Dilemma”: *Economics Letters,* 19, 227-229.

Robinson, J. (1951): “An Iterative Method of Solving a Game”. *Annals of Mathematics,* 54, 296-301.

Rubinstein, A. (1986): “Finite Automata Play the Repeated Prisoner’s Dilemma”. *Journal of Economic Theory,* 39, 83-96.

Shapley, L. S. (1964): “Some Topics in Two-Person Games. in *Advances in Game Theory* eds. By Dresher, L.S. Shapley, and A.W Tucker, 1-28, Princeton U. Press.

Young, P. (1993): “The Evolution of Conventions“. *Econometrica,* 61, 57-83.