Learning to Take Turns

Peter Vanderschraaf

Carnegie Mellon University

&

Brian Skyrms

University of California, Irvine

August, 2001

Technical Report No. CMU-PHIL-119

Philosophy

Methodology

Logic

Pittsburgh, Pennsylvania 15213

Learning to Take Turns

Peter Vanderschraaf

Carnegie Mellon University

Brian Skyrms

University of California, Irvine

§1. Introduction

Jan and Jill have a new computer game. Playing is more fun than watching, but watching is more fun than doing nothing. Only one can play at a time. If both try to play at the same time, they fight and no one gets to play. What should they do? The answer is easy for us: they should take turns. When experimental game theorists put subjects in situations qualitatively similar to that of Jan and Jill, they quickly learn to take turns (Rapoport, Guyer and Gordon 1976, Prisbey 1992). It is not so easy for Jan and Jill. Children sometimes find it difficult to learn to take turns. But learning to take turns is important for sharing. In this situation, it seems the best mechanism for sharing equally. And one might argue that it serves as a prototype for more general forms of reciprocity.

Classical game theory allows for the possibility of talking turns, but it does not explain its emergence. Suppose for simplicity that one play of the game takes a certain amount of time, and there is time for ten plays before bedtime. If Jan and Jill take turns, then neither would prefer to deviate from this pattern of action on her own. In game-theoretic parlance, they follow a *Nash equilibrium* (Nash 1950, 1951). There are two such "turn taking" equilibria of this finitely repeated game, depending on whether Jill or Jan plays first. But if Jill gets all ten plays, they follow another Nash equilibrium which is the most favorable outcome for Jill. And the Nash equilibrium in which Jan gets all ten plays is Jan's favorite outcome. There are also some mixed equilibria where the children flip mental coins. A mixed equilibrium might serve as a fair solution to their problem, in the sense of equal expected utility for Jan and Jill. But at a mixed equilibrium the children expect to end up in conflict a substantial portion of the time. At a "turn taking" equilibrium, Jan and Jill share the opportunities of playing equally without ever fighting. This seems to us a decisive objection to the mixed Nash equilibrium route to fairness.

One might well argue that the adults in game theory experiments take turns because they are already socialized. Taking turns is one of the techniques in their social tool kit. But how can people learn to take turns in the first place? Here we use a simple learning dynamics, *Markov fictitious play*, to investigate the possibility of players spontaneously learning to take turns. Markov fictitious play is perhaps the simplest form of adaptive dynamics based on pattern recognition. Like ordinary fictitious play, it has a Bayesian interpretation. We will find that players starting from randomly chosen initial positions and using Markov fictitious play can learn to take turns quite often, although this result is by no means bound to occur. We investigate how the payoff structure influences the probability of learning to take turns in various games.

§2. Markov Fictitious Play

Fictitious Play and Multinomial Sampling

In ordinary fictitious play, players on each round form a prediction of the play of others based on historical frequency, and then play their best response to the prediction. The method of forming beliefs based on past frequency has a parametric Bayesian interpretation as Bayesian updating for multinomial sampling with a Dirichlet prior.

Suppose that we are sampling from an urn with replacement and that the urn contains balls of k different colors. Let the unknown chance of drawing color $i \in \{1, ..., k\}$ be X_i. The prior density will be on the random vector $\boldsymbol{X} = (X_1, \ldots, X_k)$. The natural conjugate prior is the Dirichlet with parameter $\boldsymbol{\Theta} = (\theta_1, \ldots, \theta_k)$. The θ_i's can be any positive numbers. The prior probability of drawing color i is:

$$\frac{\theta_i}{\sum_{j=1}^{k} \theta_j}.$$

Conditional on evidence of a sample of $T \geq 0$ balls, n_1 of color 1, ... , n_k of color k, the probability of the $T+1$st ball drawn being of color i is the *Generalized Rule of Succession*:

$$X_i^{T+1} = \frac{n_i + \theta_i}{T + \sum_{j=1}^{k} \theta_j}.$$

For an updating rule in this class, if the θ_i's which are the prior weights are fixed, then all the information from the sample that is relevant to the probability that the next item be of color i, is contained in (T, n_i), the ordered pair of the sample size and the frequency of color i. This property can be used to give a subjective Bayesian characterization of beliefs that behave as if they came from a multinomial model with a Dirichlet prior (Johnson 1932, Zabell 1982).

Philosophers may recognize the class of rules that comes from multinomial sampling with Dirichlet priors as the rules in Rudolf Carnap's final system of inductive logic (Carnap 1971, 1980). For multinomial sampling, these rules possess the virtue of *Bayesian consistency* (Diaconis and Freedman 1986): For any values of the *chances* (the parameters of the multinomial model), as more and more data are observed the updating rule will (with chance = 1) yield degrees of belief that, in the limit, concentrate point mass on the true chance probability, that is, $(X_1^{T+1}, \ldots, X_k^{T+1}) = \boldsymbol{X}^{T+1} \rightarrow \boldsymbol{X}$ with chance = 1 as $T \rightarrow \infty$.

Fictitious play models belief updating of players engaged in a game with the Generalized Rule of Succession. Fictitious play gets its name because in its original interpretation, *Bayesian rational* players who maximize expected payoff mentally simulate each other's actions and update their beliefs over a sequence of imaginary plays

(Brown 1951). The interpretation of fictitious play we favor is anything but "fictitious". In our view, fictitious play is best viewed as a process by which players engaged in an actual sequence of plays of a game update their beliefs about each other according to what they actually observe. In standard 2-player fictitious play, n_i is the number of times a player's opponent has followed his ith alternative pure strategy over T plays. Fictitious play does not always converge because in the interactive game-theoretic situation, multinomial sampling is the wrong model. Trials are not independent and identically distributed. But with good luck, they may be so in the limit. There are large interesting classes of games for which fictitious play converges to Nash equilibrium and there are other interesting classes of games for which it does not converge at all.[1]

Fictitious play acts as if only sample frequencies are relevant to prediction and so it cannot detect patterns. The simplest patterns are those one might find in a Markov chain. So we might move from a parametric model of multinomial sampling with unknown composition of the urn, to a model of a Markov chain with unknown transition probabilities. Then we could base our prediction of the next play by the other player (or players) on the current state and the empirical transition counts from that state to other states in the same way as before.

In a Markov chain model, each state j is equipped with its own prior weight, θ_j, which may be different for different states, and for a given state the probability of a transition to state i, is given by the generalized rule of succession where n_i is the count of transitions from the state at issue to state i, and T is the total number of transitions from the state at issue. These are the rules of *Carnapian inductive logic for Markov chains* (Martin 1967, Kuipers 1988, Skyrms 1991). Carnapian inductive logic for Markov chains is *Bayesian consistent* for recurrent Markov chains. (If the true model is a recurrent Markov chain, then with probability one each state is visited infinitely often (Skyrms 1991). A subjective characterization of beliefs that give rise to the Carnapian inductive rules for Markov chains can be given (Zabell 1995).

Below we define Markov fictitious play. For simplicity we only consider two players. For each player, the relevant state is the outcome of the game they follow. Her end of this outcome is a best response to the act she expects from the other player. Each player enters with prior degrees of belief over the acts of the other, and these beliefs must give positive probability to each act. Each player enters with a set of generalized rules of succession (one for each state) for forecasting transitions. That is, for each state, each player has a set of prior weights for use in the generalized rule of succession to calculate transition probabilities from that state. These prior weights must all be positive. At each stage, each player chooses an act that is a best response given her beliefs. (If the current act is a best response and there is a tie for best response, she chooses the current act. If the current act is not a best response, and there is a tie for best response among other acts, she chooses a best response at random.) We now proceed to apply Markov fictitious play to situations like the one described at the beginning of this paper.

Markov Chain Expectations

We begin with a 2×2 game in strategic form, summarized in Figure 1. When they follow the strategy profile (s_i, s_j) (Player 1 chooses s_i and Player 2 chooses s_j), then Player 1's payoff is $u_1(s_i, s_j)$ and Player 2's payoff is $u_2(s_i, s_j)$.

Figure 1. General 2 × 2 Game

		Player 2	
		s_1	s_2
Player 1	s_1	$(u_1(s_1, s_1), u_2(s_1, s_1))$	$(u_1(s_1, s_2), u_2(s_1, s_2))$
	s_2	$(u_1(s_2, s_1), u_2(s_2, s_1))$	$(u_1(s_2, s_2), u_2(s_2, s_2))$

Players 1 and 2 play this game at a particular time or *round* $T + 1$, and possibly have some knowledge of what has occurred in previous rounds $1, 2, ..., T$. (For the time being, we will not specify whether or not at round $T+1$ Players 1 and 2 are the same players who have played the Figure 1 game in the previous rounds.) Let $s_{ij} = (s_i, s_j)$, $i, j \in \{1,2\}$, $s_{ij}^T = \{(s_i, s_j)$ is played at round $T\}$, let $\mu_i^T(\cdot)$ denote Player i's subjective probability distribution at time T, and set

$$\begin{aligned} \alpha_{ij,kl}^{T+1} &= \mu_1^{T+1}(s_{kl}^{T+1} | s_{ij}^T) \\ \beta_{ij,kl}^{T+1} &= \mu_2^{T+1}(s_{kl}^{T+1} | s_{ij}^T) \end{aligned}$$.[2]

That is, the $\alpha_{ij,kl}^{T+1}$'s are Player 1's transition probabilities and the $\beta_{ij,kl}^{T+1}$'s are Player 2's transition probabilities. Then we have two matrices of transition probabilities, or Markov chains:

(2.*a*.1)
$$\boldsymbol{\alpha}^{T+1} = \begin{pmatrix} \alpha_{11,11}^{T+1} & \alpha_{11,12}^{T+1} & \alpha_{11,21}^{T+1} & \alpha_{11,22}^{T+1} \\ \alpha_{12,11}^{T+1} & \alpha_{12,12}^{T+1} & \alpha_{12,21}^{T+1} & \alpha_{12,22}^{T+1} \\ \alpha_{21,11}^{T+1} & \alpha_{21,12}^{T+1} & \alpha_{21,21}^{T+1} & \alpha_{21,22}^{T+1} \\ \alpha_{22,11}^{T+1} & \alpha_{22,12}^{T+1} & \alpha_{22,21}^{T+1} & \alpha_{22,22}^{T+1} \end{pmatrix}$$

(2.*a*.2)
$$\boldsymbol{\beta}^{T+1} = \begin{pmatrix} \beta_{11,11}^{T+1} & \beta_{11,12}^{T+1} & \beta_{11,21}^{T+1} & \beta_{11,22}^{T+1} \\ \beta_{12,11}^{T+1} & \beta_{12,12}^{T+1} & \beta_{12,21}^{T+1} & \beta_{12,22}^{T+1} \\ \beta_{21,11}^{T+1} & \beta_{21,12}^{T+1} & \beta_{21,21}^{T+1} & \beta_{21,22}^{T+1} \\ \beta_{22,11}^{T+1} & \beta_{22,12}^{T+1} & \beta_{22,21}^{T+1} & \beta_{22,22}^{T+1} \end{pmatrix}$$

One can view $\boldsymbol{\alpha}^{T+1}$ and $\boldsymbol{\beta}^{T+1}$ as *Markov chains of beliefs*, since the transition probabilities are the players' degrees of beliefs that they will follow a certain strategy profile given that they just followed another profile. Let $s_{i\cdot}$ ($s_{\cdot j}$) denote the pure strategies that range when i (j) is held fixed, and let $E_i^{T+1}(\cdot)$ denote Player i's expectation operator with respect to $\mu_i^{T+1}(\cdot)$. At round $T+1$, players compute expected payoffs for their next strategy choice given the strategy profile just played at round T. That is,

$$\begin{aligned} E_1^{T+1}(u_1(s_{1\cdot})|s_{ij}^T) &= u_1(s_{11}) \cdot \mu_1^{T+1}(s_{\cdot 1}|s_{ij}^T) + u_1(s_{12}) \cdot \mu_1^{T+1}(s_{\cdot 2}|s_{ij}^T) \\ &= u_1(s_{11}) \cdot (\alpha_{ij,11}^{T+1} + \alpha_{ij,21}^{T+1}) + u_1(s_{12}) \cdot (\alpha_{ij,12}^{T+1} + \alpha_{ij,22}^{T+1}) \end{aligned}$$

and similarly

$$E_1^{T+1}(u_1(s_{2\cdot})|s_{ij}^T) = u_1(s_{21})\cdot(\alpha_{ij,11}^{T+1}+\alpha_{ij,21}^{T+1}) + u_1(s_{22})\cdot(\alpha_{ij,12}^{T+1}+\alpha_{ij,22}^{T+1})$$

$$E_2^{T+1}(u_2(s_{\cdot 1})|s_{ij}^T) = u_2(s_{11})\cdot(\beta_{ij,11}^{T+1}+\beta_{ij,12}^{T+1}) + u_2(s_{21})\cdot(\beta_{ij,21}^{T+1}+\beta_{ij,22}^{T+1})$$

$$E_2^{T+1}(u_2(s_{\cdot 2})|s_{ij}^T) = u_2(s_{12})\cdot(\beta_{ij,11}^{T+1}+\beta_{ij,12}^{T+1}) + u_2(s_{22})\cdot(\beta_{ij,21}^{T+1}+\beta_{ij,22}^{T+1})$$

For illustration, consider the Figure 2 game. This game models one of Jan and Jill's opportunities to play their computer game before bedtime. If Jill plays (s_2), then Jan's unique best response is to watch (s_1), for only s_1 maximizes Jan's expected payoff given that Jill chooses s_2. s_2 is also Jill's unique best response to Jan's choosing s_1, so s_{12} is a *strict* Nash equilibrium. s_{21} is also a strict Nash equilibrium. The Figure 2 game is an *impure coordination game* (Lewis 1969), because Jill's and Jan's preferences over the equilibria s_{12} and s_{21} conflict.

Figure 2. Computer Game

		Player 2 (Jill)	
		s_1	s_2
Player 1 (Jan)	s_1	$(1,1)$	$(2,3)$
	s_2	$(3,2)$	$(0,0)$

$s_1 =$ watch, $s_2 =$ play

Suppose the transition matrices satisfy

$$(2.b.1) \qquad \boldsymbol{\alpha}^{T+1} = \begin{pmatrix} \alpha_{11,11}^{T+1} & \alpha_{11,12}^{T+1} & \alpha_{11,21}^{T+1} & \alpha_{11,22}^{T+1} \\ \alpha_{12,11}^{T+1} & \alpha_{12,12}^{T+1} & \alpha_{12,21}^{T+1} & \alpha_{12,22}^{T+1} \\ 0 & 0 & 1 & 0 \\ \alpha_{22,11}^{T+1} & \alpha_{22,12}^{T+1} & \alpha_{22,21}^{T+1} & \alpha_{22,22}^{T+1} \end{pmatrix}$$

$$(2.b.2)\qquad \boldsymbol{\beta}^{T+1} = \begin{pmatrix} \beta^{T+1}_{11,11} & \beta^{T+1}_{11,12} & \beta^{T+1}_{11,21} & \beta^{T+1}_{11,22} \\ \beta^{T+1}_{12,11} & \beta^{T+1}_{12,12} & \beta^{T+1}_{12,21} & \beta^{T+1}_{12,22} \\ 0 & 0 & 1 & 0 \\ \beta^{T+1}_{22,11} & \beta^{T+1}_{22,12} & \beta^{T+1}_{22,21} & \beta^{T+1}_{22,22} \end{pmatrix}$$

that is, if the players play s_{21} at a given round, they believe with probability one that they will play s_{21} at the next round. Then if the players play s_{21} at round T, then

$$\begin{aligned} E_1^{T+1}(u_1(s_{1\cdot})|s_{21}^T) &= u_1(s_{11})\cdot(\alpha^{T+1}_{21,11}+\alpha^{T+1}_{21,21}) + u_1(s_{12})\cdot(\alpha^{T+1}_{21,12}+\alpha^{T+1}_{21,22}) \\ &= 1\cdot(0+1)+2\cdot(0+0) = 1 \\ E_1^{T+1}(u_1(s_{2\cdot})|s_{21}^T) &= u_1(s_{21})\cdot(\alpha^{T+1}_{21,11}+\alpha^{T+1}_{21,21}) + u_1(s_{22})\cdot(\alpha^{T+1}_{21,12}+\alpha^{T+1}_{21,22}) \\ &= 3\cdot(0+1)+0\cdot(0+0) = 3 \\ E_2^{T+1}(u_2(s_{\cdot 1})|s_{21}^T) &= u_2(s_{11})\cdot(\beta^{T+1}_{21,11}+\beta^{T+1}_{21,12}) + u_2(s_{21})\cdot(\beta^{T+1}_{21,21}+\beta^{T+1}_{21,22}) \\ &= 1\cdot(0+0)+2\cdot(1+0) = 2 \\ E_2^{T+1}(u_2(s_{\cdot 2})|s_{21}^T) &= u_2(s_{12})\cdot(\beta^{T+1}_{21,11}+\beta^{T+1}_{21,12}) + u_2(s_{22})\cdot(\beta^{T+1}_{21,21}+\beta^{T+1}_{21,22}) \\ &= 3\cdot(0+0)+0\cdot(1+0) = 0 \end{aligned}$$

that is, at the $T+1$st round of play they maximize their expected payoffs by playing s_{21} again. So the players' transition probabilities for s_{21} characterize a Nash equilibrium. More precisely, their s_{21} transition probabilities are an *equilibrium-in-beliefs* because they define equilibrium play from the perspective of each players' probabilities over the other player's strategies (Aumann 1987). If for some m the transition matrices satisfy (2.b.1) and (2.b.2) for all rounds of play $T \geq m$, then if the players visit the equilibrium s_{21} at the mth or any subsequent round of play, then given their beliefs they will play s_{21} for all successive rounds.

Now suppose that the transition matrices are such that

$$(2.c.1)\qquad \boldsymbol{\alpha}^{T+1} = \begin{pmatrix} \alpha^{T+1}_{11,11} & \alpha^{T+1}_{11,12} & \alpha^{T+1}_{11,21} & \alpha^{T+1}_{11,22} \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ \alpha^{T+1}_{22,11} & \alpha^{T+1}_{22,12} & \alpha^{T+1}_{22,21} & \alpha^{T+1}_{22,22} \end{pmatrix}$$

$$(2.c.2) \qquad \boldsymbol{\beta}^{T+1} = \begin{pmatrix} \beta_{11,11}^{T+1} & \beta_{11,12}^{T+1} & \beta_{11,21}^{T+1} & \beta_{11,22}^{T+1} \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ \beta_{22,11}^{T+1} & \beta_{22,12}^{T+1} & \beta_{22,21}^{T+1} & \beta_{22,22}^{T+1} \end{pmatrix}$$

that is, each player believes with probability 1 that if s_{12} (s_{21}) has just been played, s_{21} (s_{12}) will be played next. Then if the two players play s_{12} at round T, then

$$\begin{aligned} E_1^{T+1}(u_1(s_{1\cdot})|s_{12}^T) &= u_1(s_{11}) \cdot (\alpha_{12,11}^{T+1} + \alpha_{12,21}^{T+1}) + u_1(s_{12}) \cdot (\alpha_{12,12}^{T+1} + \alpha_{12,22}^{T+1}) \\ &= 1 \cdot (0+1) + 2 \cdot (0+0) = 1 \\ E_1^{T+1}(u_1(s_{2\cdot})|s_{12}^T) &= u_1(s_{21}) \cdot (\alpha_{12,11}^{T+1} + \alpha_{12,21}^{T+1}) + u_1(s_{22}) \cdot (\alpha_{12,12}^{T+1} + \alpha_{12,22}^{T+1}) \\ &= 3 \cdot (0+1) + 0 \cdot (0+0) = 3 \\ E_2^{T+1}(u_2(s_{\cdot 1})|s_{12}^T) &= u_2(s_{11}) \cdot (\beta_{12,11}^{T+1} + \beta_{12,12}^{T+1}) + u_2(s_{21}) \cdot (\beta_{12,21}^{T+1} + \beta_{12,22}^{T+1}) \\ &= 1 \cdot (0+0) + 2 \cdot (1+0) = 2 \\ E_2^{T+1}(u_2(s_{\cdot 2})|s_{12}^T) &= u_2(s_{12}) \cdot (\beta_{12,11}^{T+1} + \beta_{12,12}^{T+1}) + u_2(s_{22}) \cdot (\beta_{12,21}^{T+1} + \beta_{12,22}^{T+1}) \\ &= 3 \cdot (0+0) + 0 \cdot (1+0) = 0 \end{aligned}$$

so the players will play s_{21} at round $T+1$. Similarly, if the two players play s_{21} at round T, then

$$\begin{aligned} E_1^{T+1}(u_1(s_{1\cdot})|s_{21}^T) &= u_1(s_{11}) \cdot (\alpha_{21,11}^{T+1} + \alpha_{21,21}^{T+1}) + u_1(s_{12}) \cdot (\alpha_{21,12}^{T+1} + \alpha_{21,22}^{T+1}) \\ &= 1 \cdot (0+0) + 2 \cdot (1+0) = 2 \\ E_1^{T+1}(u_1(s_{2\cdot})|s_{21}^T) &= u_1(s_{21}) \cdot (\alpha_{21,11}^{T+1} + \alpha_{21,21}^{T+1}) + u_1(s_{22}) \cdot (\alpha_{21,12}^{T+1} + \alpha_{21,22}^{T+1}) \\ &= 3 \cdot (0+0) + 0 \cdot (1+0) = 0 \\ E_2^{T+1}(u_2(s_{\cdot 1})|s_{21}^T) &= u_2(s_{11}) \cdot (\beta_{21,11}^{T+1} + \beta_{21,12}^{T+1}) + u_2(s_{21}) \cdot (\beta_{21,21}^{T+1} + \beta_{21,22}^{T+1}) \\ &= 1 \cdot (0+1) + 2 \cdot (0+0) = 1 \\ E_2^{T+1}(u_2(s_{\cdot 2})|s_{21}^T) &= u_2(s_{21}) \cdot (\beta_{21,11}^{T+1} + \beta_{21,12}^{T+1}) + u_2(s_{22}) \cdot (\beta_{21,21}^{T+1} + \beta_{21,22}^{T+1}) \\ &= 3 \cdot (0+1) + 0 \cdot (0+0) = 3 \end{aligned}$$

so the players will play s_{12} at round $T+1$. So given these transition matrices, if the players ever play *either* coordination equilibrium s_{12} or s_{21}, they will maximize expected payoffs by alternating between s_{12} and s_{21} on successive plays. If the transition matrices are of the form (2.c.1) and (2.c.2) for all rounds $T \geq m$, then if the players ever play either s_{12} or s_{21} at round m or at any round thereafter, they will alternate successively between s_{12} and s_{21} for all future rounds of play, and this alternation scheme is a "taking turns" equilibrium.

Note that this "taking turns" equilibrium resembles a *correlated equilibrium* (Aumann 1974, 1987) where the players peg their pure strategies upon information they have regarding some random event. Figure 3 summarizes a "coin-flip" correlated equilibrium of the Figure 2 game, in which Jan and Jill observe a coin toss and play s_{12} if the coin lands "heads-up" and s_{21} if the coin lands "tails-up".

Figure 3. "Coin Flip" Correlated Equilibrium of the Computer Game

		Player 2 (Jill)	
		s_1	s_2
Player 1 (Jan)	s_1	$(1, 1)$	$(2, 3)^{\omega_1}$
	s_2	$(3, 2)^{\omega_2}$	$(0, 0)$

ω_1 = "heads", ω_2 = "tails"

If they follow this correlated equilibrium over successive plays, the players alternate randomly between the profiles s_{12} and s_{21} according to the result of the coin flip at each round. If the coin is fair and the players know this, then they achieve a fair expected payoff of $\frac{1}{2} \cdot 3 + \frac{1}{2} \cdot 2 = \frac{5}{2}$ for both at each round by following the "coin-flip" equilibrium. However, the actual average payoff each receives converges to $\frac{5}{2}$ only in the limit. The "taking turns" equilibrium described in the previous paragraph does not require an external experiment. One can view the strategies defined by the "taking turns" equilibrium at the current round as a correlated equilibrium where the players correlate their strategies now with what they did at the previous round. But there is an important difference between the "coin-flip" and the "taking turns" equilibria. At the "taking turns" equilibrium, Jan and Jill alternate *successively* between s_{12} and s_{21}. As a result, the "taking turns" equilibrium *always* results in a fair outcome for any even number of plays. In the case where Jill and Jan play the Computer Game ten times, by following the "coin-flip" equilibrium each gets five plays with probability only $\binom{10}{5}\frac{1}{2}^{10} \approx 0.246$, while by following the "taking turns" equilibrium each is guaranteed five plays.

Carnapian Markov Dynamics

So far, we have looked at how players can use Markov chains of beliefs to compute expected payoffs and possibly follow an equilibrium without saying anything explicit about how this process might influence the beliefs themselves. It is natural to suppose that the players will modify their transition probabilities over time as they learn from experience. One possible way the players might update their transition probabilities follows the Carnapian inductive logic for Markov chains or *Carnapian Markov rule* developed in Skyrms (1991). Player 1 applies this Carnapian Markov rule to update her transition probabilities as follows: Let

$$1_{s_{ij}}(m) = \begin{cases} 1 \text{ if Players 1 and 2 play } s_{ij} \text{ at round } m \\ 0 \text{ otherwise} \end{cases}$$

$$1_{s_{ij,kl}}(m) = \begin{cases} 1 \text{ if Players 1 and 2 play } s_{ij} \text{ at round } m-1 \text{ and } s_{kl} \text{ at round } m \\ 0 \text{ otherwise} \end{cases}$$

and set

$$n_{ij}^{T} = \sum_{m=1}^{T} 1_{s_{ij}}(m)$$

$$n_{ij,kl}^{T} = \sum_{m=1}^{T} 1_{s_{ij},s_{kl}}(m) \ .$$

That is, over the first T rounds, n_{ij}^{T} is the number of rounds at which the players have played s_{ij} and $n_{ij,kl}^{T}$ is the number of transitions from s_{ij} to s_{kl}. $\alpha_{ij,kl}^{0}$ is Player 1's prior probability for the transition from s_{ij} to s_{kl}, and $\lambda_1 > 0$ is a constant which governs the weight which Player 1 places on the prior. At round $T+1$:

(*a*) $1_{s_{ij,kl}}(T+1) = 1$ if

$$E_1^{T+1}(u_1(s_{k\cdot})|s_{ij}^T) > E_1^{T+1}(u_1(s_{p\cdot})|s_{ij}^T) \text{ for } s_p \neq s_k$$
$$E_2^{T+1}(u_2(s_{\cdot l})|s_{ij}^T) > E_2^{T+1}(u_2(s_{\cdot q})|s_{ij}^T) \text{ for } s_q \neq s_l$$

and (*b*) $1_{s_{ij,kl}}(T+1) = 1$ only if

$$E_1^{T+1}(u_1(s_{k\cdot})|s_{ij}^T) \geq E_1^{T+1}(u_1(s_{p\cdot})|s_{ij}^T) \text{ for } s_p \neq s_k$$
$$E_2^{T+1}(u_2(s_{\cdot l})|s_{ij}^T) \geq E_2^{T+1}(u_2(s_{\cdot q})|s_{ij}^T) \text{ for } s_q \neq s_l$$

Player 1's transition probability for s_{kl} from s_{ij} is defined by

$$\mu_1^{T+1}(s_{kl}^{T+1}|s_{ij}^T) = \alpha_{ij,kl}^{T+1} = \frac{n_{ij,kl}^T + \lambda_1\alpha_{ij,kl}^0}{n_{ij}^T + \lambda_1} \,. \tag{2.1}$$

The product $\lambda_1\alpha_{ij,kl}^0$ corresponds to the prior weight of the Generalized Succession Rule. Condition (*b*) is included to allow for alternative rules for handling ties in the conditional expected payoffs, in line with the similar practice in the ordinary fictitious play literature. Player 2's transition probabilities are similarly defined. For the remainder of this paper whatever we discuss about the Markov fictitious play process with respect to Player 1 applies to Player 2, *mutatis mutandis*. (2.1) can also be written as

$$\mu_1^{T+1}(s_{kl}^{T+1}|s_{ij}^T) = \frac{1_{s_{ij,kl}}(T)}{n_{ij}^T + \lambda_1} + \frac{n_{ij}^{T-1} + \lambda_1}{n_{ij}^T + \lambda_1} \cdot \mu_1^T(s_{kl}^T|s_{ij}^{T-1}) \tag{2.1'}$$

which emphasizes the recursive nature of this inductive rule.[3]

In ordinary fictitious play, strict Nash equilibria are fixed points of sequences of updated beliefs. The situation is more general for Markov fictitious play. Both strict Nash equilibria and "taking turns" equilibria are fixed points of Markov fictitious play. To show this, we generalize in a straightforward way the arguments we gave to show that the absorbing state of the Markov Chain pairs (2.*b*.1) and (2.*b*.1) and the periodic set of the Markov chain pairs (2.*c*.1) and (2.*c*.1) characterize equilibria-in-beliefs of the Figure 2 game. The following result is proved in Appendix 2.:

Proposition 1. (*i*) If the transition matrices of Markov fictitious play make a strict Nash equilibrium an absorbing state of the Markov chains, then Markov deliberators who ever visit this state will always follow this equilibrium. (*ii*) If the transition matrices make a "taking turns" equilibrium a periodic set of the Markov chains, then deliberators who enter into this set always follow this equilibrium.

Proposition 1 says that strict Nash and "taking turns" equilibria are equilibria of the Markov fictitious play process, that is, they are *deliberational equilibria.* Proposition 1 also establishes that in games with two distinct strict Nash equilibria, deliberational equilibria of Markov fictitious play exist which characterize both the strict Nash equilibria and a "taking turns" equilibrium.

Another important result, also proved in Appendix 2, is the following:

Proposition 2. (*i*) If Markov deliberators ever follow a strict Nash equilibrium s^* in consecutive rounds, then they will follow s^* all subsequent rounds. (*ii*) If s_1^* and s_2^* are distinct strict Nash equilibria and Markov deliberators ever follow s_1^*, s_2^* and s_1^* in consecutive rounds, then they will subsequently always alternate between s^* and s^{**} in consecutive rounds.

Proposition 2 is the Markov fictitious play analog of the Absorption Theorem of ordinary fictitious play (Fudenberg and Tirole 1998, Vanderschraaf 2001). The Absorption Theorem says that if players who update beliefs according to the ordinary fictitious play rule ever visit a strict Nash equilibrium, they will remain at this equilibrium for all future rounds of play. The situation is more complex with Markov fictitious play. Proposition 2 says that if Markov deliberators visit a strict Nash equilibrium twice in a row, then their play is "absorbed" into this strict equilibrium for all future rounds of play. But if Markov deliberators visit one strict Nash equilibrium, then visit another and then return to the first, their play is "absorbed" into a "taking turns" equilibrium of consecutive visits between the two strict equilibria.

§3. Learning to Take Turns

Jan and Jill

"Taking turns" equilibria exist, but we would like to know whether they can arise spontaneously, and if so whether this should be an extremely unlikely occurrence or one that we might expect to occur naturally with some regularity. Suppose we just pick some initial beliefs for Jan and Jill at random, and start them out in Markov fictitious play. What should we expect to see?

Proposition 2 suggests the possibility that a "taking turns" equilibrium has a nonnegligible *basin of attraction*, that is, a part of the simplex of possible transition probabilities such that if the players' beliefs ever fall in this part, they will start to follow the taking "turns equilibrium" and continue to follow it.[4] If players' beliefs ever enter into the basin of attraction of the "taking turns" equilibrium, then their beliefs eventually converge to the equilibrium-in-beliefs that defines this equilibrium. When the four inequalities

(3.*a*)
$$\alpha^{T+1}_{12,21} > \frac{1}{2} - \alpha^{T+1}_{12,11}$$
$$\beta^{T+1}_{12,21} > \frac{1}{2} - \beta^{T+1}_{12,22}$$
$$\alpha^{T+1}_{21,12} > \frac{1}{2} - \alpha^{T+1}_{21,22}$$
$$\beta^{T+1}_{21,12} > \frac{1}{2} - \beta^{T+1}_{21,11}$$

are satisfied, Jan and Jill are sure to follow the "taking turns" equilibrium of the Figure 2 game, so this equilibrium indeed has a nonnegligible basin of attraction.[5]

Players can learn to take turns in the Computer Game from initial conditions that do not necessarily satisfy (3.*a*). Suppose that Jan's and Jill's transition priors satisfy

$$\alpha^{0}_{11,12} > \frac{1}{2},\ \alpha^{0}_{12,21} > \frac{1}{2}, \alpha^{0}_{21,12} > \frac{1}{2}, \alpha^{0}_{22,21} > \frac{1}{2}$$

$$\beta^0_{11,12} > \frac{1}{2},\ \beta^0_{12,12} > \frac{1}{2},\ \beta^0_{21,21} > \frac{1}{2},\ \beta^0_{22,21} > \frac{1}{2}$$

that is, Jill initially believes that Jan is more likely than not to try to stay at given strict Nash equilibrium over successive plays, while Jan believes that Jill is more likely than not to alternate between the strict Nash equilibria. So long as all eight stated inequalities are satisfied, play will follow a cycle where the pattern s_{12}, s_{22}, s_{21}, s_{11}, s_{12} repeats. Suppose now that λ_1 is large relative to λ_2, so that Jan requires more transitions to reverse only one of the stated inequalities than Jill requires to change two. That is, relative to each other, Jan is resolute in her initial beliefs while Jill is a "quick learner". Eventually, Jill reverses an inequality while Jan does not. Suppose Jill first reverses her inequality $\beta^0_{12,12} > \frac{1}{2}$ and concludes that $\beta^T_{12,22} > \frac{1}{2}$. Then when they visit s_{12} Jan and Jill now go to s_{21} because Jan did not reverse her inequality $\alpha^0_{12,21} > \frac{1}{2}$. Similarly, if Jill reverses her inequality $\beta^0_{21,21} > \frac{1}{2}$ so that $\beta^T_{21,11} > \frac{1}{2}$, then from s_{21} the players now make the transition to s_{12}. Once both reversals occur Jan and Jill take turns between s_{12} and s_{21}. Suppose the reversal $\beta^T_{12,22} > \frac{1}{2}$ occurs first. Then play enters into the cycle s_{12}, s_{21}, s_{11}, s_{12}, and eventually this leads Jill to make the second reversal by reinforcing the transition from s_{21} to s_{11}. Then play enters the cycle s_{12}, s_{21}, s_{12} and is absorbed into the "taking turns" equilibrium by Proposition 2. Jan has "trained" Jill to take turns!

However, it is by no means a foregone conclusion that if Jan and Jill update their beliefs according to the Carnapian Markov rule that they will learn to take turns. The strict Nash equilibria also have nonnegligible basins of attraction for this dynamics. Other of Markov fictitious play are also possible, including sequences of play where Jan and Jill miscoordinate on s_{11} and s_{22} every time. What are their prospects for learning to take turns as Markov fictitious players?

We explored the properties of Markov fictitious play in Jan and Jill's Computer Game by running a series of 10,000 computer simulations. The results of these and other simulations we ran are summarized in detail in Appendix 1. In each simulation, a pair of initial beliefs over the transitions of the states of the Computer Game were selected at random and then updated according to the Markov fictitious play rule (2.1). Beliefs nearly always converged to either a pure Nash equilibrium or a taking turns equilibrium. In our simulations, players settled into a taking turns equilibrium over a third of the time, and more often than they settled into either strict Nash equilibrium. (Other theoretical possibilities such as a mixed Nash equilibrium were seen less than 2% of the time.) We don't attach much significance to the exact numbers in these simulation results. But we can draw an important qualitative conclusion. From a randomly chosen starting point under fictitious play, Jan and Jill can spontaneously learn to take turns. It is far from guaranteed that they will succeed, but it does not require a miracle for them to do so.

Populations

Suppose that Jan and Jill do not have the computer to themselves, but that other children in the room take their places, and are themselves replaced by others. All the children have seen what has happened, but new children have their own initial weights for transitions. In this social setting we have a far more heterogeneous assortment of learning styles, here represented by different inductive rules. Is it possible to learn to take turns in this more challenging setting?

There is a natural way to generalize the Carnapian inductive rule. If one uses a "representative" interpretation of the updating process in which at each round fresh individuals assume the roles of each player and update according to the observed past history of plays, then the players' priors and weighting constants λ_i's can vary with T, so that (2.1) generalizes to:

(3.1) $$\mu_1^{T+1}(s_{kl}^{T+1}|s_{ij}^T) = \alpha_{ij,kl}^{T+1} = \frac{n_{ij,kl}^T + \lambda_1(T+1)\cdot\alpha_{ij,kl}^0(T+1)}{n_{ij}^T + \lambda_1(T+1)}$$

In this variant of Markov fictitious play, the dynamical system of updated beliefs is constantly "bombarded" by the random fluctuations in individual players' priors and weighting constants. There are no analogs of Propositions 1 and 2 for the (3.1) Markov dynamics, because even if a population settles into a strict Nash or a "taking turns" equilibrium for a very long time, it is always a theoretical possibility that newcomers with their own idiosyncratic weighting constants and priors can disrupt the system and throw the population off the incumbent equilibrium.

Do Markov deliberators in such an inherently "noisy" system have any prospect of ever reaching an equilibrium? As a matter of fact, for the 2×2 case the (3.1) Markov dynamics exhibits remarkable convergence properties. Again we ran computer simulations, this time with the representative Markov fictitious play rule (3.1) with each representative's initial transition probabilities and weighting factor picked at random. We were somewhat surprised to find that the results in this social learning model were quite similar to those of the previous case where Jan and Jill were the only players. In these simulations, taking turns emerges about a quarter of the time, and play *always* converges to either one of the strict Nash equilibria or to a taking turns equilibrium. In these cases the expectation that the players would take turns arose spontaneously and then was passed along to new players in the social context.

§4. Conclusion

Our Markov fictitious play model shows that taking turns can emerge as the result of trial and error learning. At the same time, Markov fictitious play supplies an answer to an old criticism of fictitious play. It is well known that players who update their beliefs according to traditional fictitious play can converge to a Nash equilibrium-in-beliefs and yet *never* successfully coordinate their behavior. For instance, if Jan and Jill update their beliefs according to their history of play of the Figure 2 game with the traditional fictitious play rule, then it is possible for their beliefs to converge to the mixed Nash equilibrium-in-beliefs at which each believes the other plays s_1 with probability $\frac{1}{2}$. Yet if this happens, then their play oscillates between (s_1, s_1) and (s_2, s_2), and neither ever actually gets to play the computer game! This is not what we would expect Jill and Jan to do, and this is not what experimental subjects placed in similar situations do. Faced with impure coordination problems like the Figure 2 game in the laboratory, people tend to quickly settle into "taking turns" equilibria (Rapoport, Guyer and Gordon 1976, Chapters 9, 10, 11). Traditional fictitious play cannot account for this commonplace phenomenon. Traditional fictitious play fails to model even the simplest nontrivial patterns of play, and so many have argued that fictitious play is too crude to be used as a model of learning in game theory. Markov fictitious play extends the traditional model so as to model an especially important part of social interaction, namely, learning to taking turns. Our general approach is also not limited to Markov fictitious play. Other models of pattern recognition learning could allow players to learn to take turns (Sonsino 1997). This suggests that game theorists should study models of fictitious play with the aim of finding the correctly general model, not that fictitious play is the wrong model of learning in game theory. To be sure, we do not claim that the Markov fictitious play rules defined by (2.1) and (3.1) are fully accurate models of human learning. This kind of Markov fictitious play surely oversimplifies the learning process that goes on in human communities. Yet the very simplicity of this model is illuminating. Even a community of players who have no prior experience of interactions and who update their beliefs according to the simple rules (2.1) and (3.1) can learn to take turns. We conjecture that players like Jan and Jill, who are neither so naive nor so ignorant as the deliberators of our Markov learning model, have even better prospects for learning to take turns.

References

Aumann, Robert. (1974) "Subjectivity and Correlation in Randomized Strategies." *Journal of Mathematical Economics*, 1: 67-96.

Aumann, Robert. (1987) "Correlated Equilibrium as an Expression of Bayesian Rationality." *Econometrica*, 55: 1-18.

Brown, G. W. (1951) "Iterative solutions of games by fictitious play," in T.C. Koopmans (ed.) *Activity Analysis of Production and Allocation.* New York: Wiley.

Carnap, R. (1971) "A Basic System of Inductive Logic, Part 1" in R. Carnap and R. Jeffrey (eds.) *Studies in Inductive Logic and Probability*, vol. I. University of California Press: Berkeley.

Carnap, R. (1980) "A Basic System of Inductive Logic, Part 2" in R. Jeffrey (ed.) *Studies in Inductive Logic and Probability*, vol. II. University of California Press: Berkeley.

Diaconis, P. and Freedman, D. (1986) "On the Consistency of Bayes Estimates" *Annals of Statistics* 14: 1-26.

Fudenberg, Drew and Levine, David. (1998) *The Theory of Learning in Games.* Cambridge, Massachusetts: MIT Press.

Johnson, W. E. (1932) "Probability: the Deductive and Inductive Problems" *Mind* 49, 409-423.

Kuipers, T. A. F. (1988) "Inductive Logic by Similarity and Proximity" In *Analogical Reasoning* (ed. D. A. Helman). Kluwer: Dordrecht.

Lewis, D. (1969) *Convention: A Philosophical Study*, Harvard University Press, Cambridge, Massachusetts.

Martin, J. J. (1967) *Bayesian Decision Problems and Markov Chains.* Wiley: New York.

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

Nash, J. (1950) "Equilibrium points in n-person games." *Proceedings of the National Academy of Sciences of the United States* 36: 48-49.

Nash, J. (1951) "Non-Cooperative Games." *Annals of Mathematics* 54: 286-295.

Prisbey, J. (1992) "An Experimental Analysis of Two-Person Reciprocity Games", California Institute of Technology Social Science Working Paper 787.

Rapoport, A., Guyer, M. and Gordon, D. (1976) *The 2 × 2 Game*. Ann Arbor: The University of Michigan Press.

Richards, D. (1997) "The Geometry of Inductive Reasoning in Games" *Economic Theory* 10:185-193.

Shapley, L. S. (1964) "Some topics in two-person games", in *Advances in Game Theory*, ed. M. Drescher, L. S. Shapley and A. W. Tucker. Princeton University Press: Princeton.

Skyrms, B. (1990) *The Dynamics of Rational Deliberation*. Harvard University Press: Cambridge, Massachusetts.

Skyrms, B. (1991) "Carnapian Inductive Logic for Markov Chains" *Erkenntnis* 35 439-460.

Sonsino, D. (1997) "Learning to Learn, Pattern Recognition, and Nash Equilibrium" *Games and Economic Behavior* (1997) 18: 286-331.

Vanderschraaf, P. (2001) *Learning and Coordination*. New York: Routledge.

Vanderschraaf, P. and Skyrms, B. (1993) "Deliberational Correlated Equilibria" *Philosophical Topics* 21: 191-227.

Zabell, S. (1982) "W. E. Johnson's 'sufficientness' postulate" *Annals of Statistics* 10 1091-1099.

Zabell, S. (1995) "Characterizing Markov Exchangeable Sequences" *Journal of Theoretical Probability* 8, 175-178.

Appendix 1. Computer Simulations

Each simulation was run using a fixed 2×2 base game with two strict Nash equilibria.[6] This property is a necessary and sufficient condition for the existence of a "taking turns" equilibrium in a 2×2 game. We ran two sets of simulations on 16 different 2×2 games, each of which has important applications in moral and political philosophy and the social sciences. The payoff structures of these 16 games are given in Table 1.

In the first set of simulations, transition probabilities were updated according to rule (2.1) as described above in §2. Our program dynamics simulates a sequence of successive rounds of play of a game, or *run*, in which a pair of players enter into the game with fixed transition priors and weighting constants and who update their transition probabilities according to rule (2.1). At the start of a run, $\alpha^0_{ij,kl}$ and $\beta^0_{ij,kl}$ were sampled from the uniform distribution over the interval $[0, 1]$, and λ_1 and λ_2 were sampled from the random variable $Y = 10 \cdot \text{abs}(X)$, where X is a normally distributed random variable with mean 0 and variance 1. $\alpha^0_{ij,kl}$ and $\beta^0_{ij,kl}$ were uncorrelated, and λ_1 and λ_2 were uncorrelated as well. We put no *a priori* upper bound on the size of the a player's weighting constant λ_i, and sample priors from the uniform distribution so as not to bias the dynamics initially in favor of any particular equilibrium. Each simulation in this set consisted of 10,000 independent runs of 500 rounds each. A run of 500 rounds is typically more than sufficiently long for a pair of deliberators satisfying these parameters to settle into one of the game's equilibrium points. The parameters of the 10,000 runs in each simulation were stochastically independent of each other so as to investigate the tendency of a pair of Markov Carnapian updaters to settle into a certain equilibrium of a game with no prior correlation on any equilibrium they might have derived from the observing other players who have been playing a game with "taking turns" equilibria. Table 2 summarizes the results of this set of simulations. Table 2 records the frequency distribution of convergence into a pure strategy profile, a "taking turns" equilibrium or a mixed Nash equilibrium for each simulation. For instance, in the simulation run on the Computer Game of Figure 2, in 10,000 runs the dynamics converged upon the strict Nash equilibrium (s_1, s_2) 32.29% of the time (3229 times in 10,000 runs), upon the strict Nash equilibrium (s_2, s_1) 31.10% of the time, upon a mixed Nash equilibrium 1.93% of the time, and upon a "taking turns" equilibrium 34.68% of the time.

In the second set of simulations, transition probabilities were updated according to rule (3.1) as described in §3 where at any given round T, $\alpha^0_{ij,kl}(T)$ and $\beta^0_{ij,kl}(T)$ were sampled from the uniform distribution over the interval $[0, 1]$, and $\lambda_1(T)$ and $\lambda_2(T)$ were sampled from the random variable $Y = 10 \cdot \text{abs}(X)$, where X is a normally distributed random variable with mean 0 and variance 1. $\alpha^0_{ij,kl}(T)$ and $\beta^0_{ij,kl}(T)$ were uncorrelated, as were $\lambda_1(T)$ and $\lambda_2(T)$. Setting these conditions on the dynamics simulates a run in which after each round of play the players who have just faced each other are replaced with a fresh pair of players who have their own priors and weighting constants, and who also know the frequencies of transitions over past rounds. We put no *a priori* upper bound on the $\lambda_i(T)$'s, and sampled priors from the uniform distribution for the same reason we followed the similar practice in the simulations of the (2.1) dynamics, namely, so as not to initially bias the (3.1) dynamics towards any of the game's equilibria. Each simulation consisted of 10,000 independent runs of 500 rounds each. Intuitively, we might think of each run in a simulation as mimicking the activities of a community whose members participate in a game for 500 rounds, each round pairing two fresh members who have observed the frequency of transitions made by their predecessors. 500 rounds is typically more than sufficiently many for a system of deliberators satisfying these parameters to settle into one of the game's equilibrium points. (Convergence failure over any finite number of rounds is a theoretical possibility with the rule (3.1) Markov dynamics, but this never occurred in any of the 160,000 runs of our simulations.) The parameters within each run and over the 10,000 runs in each simulation were all stochastically independent so as to investigate the tendency of a community of Markov Carnapian updaters to settle into a certain equilibrium of a game with no prior correlation on any equilibrium they might have derived from the practices of other communities. In Table 3, we record the frequency distribution of convergence into a pure strategy profile and a "taking turns" equilibrium for each simulation. For instance, in 10,000 runs of the Figure 2 Computer Game, the dynamics converged upon the strict Nash equilibrium (s_1, s_2) 37.31% of the time (3731 times in 10,000 runs), upon the strict Nash equilibrium (s_2, s_1) 38.87% of the time, and upon a "taking turns" equilibrium 23.82% of the time.

These simulations show that a "taking turns" equilibrium can indeed emerge spontaneously, even if the belief updating rule is a very simple ordinary Markov updating rule used by players who have no *a priori* tendency to follow any particular equilibrium. In every simulation run with the rule (2.1) dynamics, the system almost always converges either to a strict Nash equilibrium or to a "taking turns" equilibrium. (In less than 2% of the cases, the dynamics converged to a suboptimal sequence of plays corresponding to a mixed Nash equilibrium.) In *every* simulation run with the rule (3.1), the system converges either to a strict Nash equilibrium or to a "taking turns" equilibrium. (We believe this remarkable result stems from the fact that in a system of rule (3.1) deliberators, transitions probabilities are constantly bombarded with the "noise" of priors and $\lambda_i(T)$'s of fresh representatives, which forestalls the system settling into the negligible basin of attraction of a mixed Nash equilibrium.) In most of the games with symmetric payoffs, the "taking turns" equilibrium emerges somewhat less often than either of the strict Nash equilibria. This may be due in part to the simplicity of the Markov updating rules (2.1) and (3.1) and the fact that taking turns is a more complicated equilibrium for players to follow than simply following one strategy profile unconditionally. What is striking is that taking turns emerges *at all* in these simulations. For recall that by construction these players have no prior experience of taking turns, and no encouragement from others who have already learned to take turns to help guide them, as our young friends Jan and Jill would have. These types of players represent a "lower bound" in terms of prior knowledge and sophistication of players who *could* learn to take turns, and they converge to "turn taking" equilibria a significant amount of the time.

Note that we ran simulations on only two games with asymmetric payoffs. This is because games with asymmetric payoffs do not fit that well into our model of representative players playing the game, since we cannot easily explain the asymmetries in payoffs without making arbitrary assumptions about the players in the system who get assigned to the row or the column position. The results of the simulations on Games 15 and 16 are not at all surprising. The fact that the asymmetries of the game seem to favor one player over the other makes one of the strict Nash equilibria an especially powerful attractor of the Markov dynamics, same as in ordinary fictitious play. Yet while we believe that our representative model applies best to games with a symmetric payoff structure, it is interesting to note that in the rule (3.1) simulations run on the two asymmetric games, taking turns emerges more frequently than the strict Nash equilibrium which is the weaker attractor in ordinary fictitious play.

Table 1. Payoff Structures of 2 × 2 Games Used in Simulations

1. Computer Game

Player 2

Player 1	s_1	s_2
s_1	(1, 1)	(2, 3)
s_2	(3, 2)	(0, 0)

2. Battle of the Sexes *I*

Player 2

Player 1	s_1	s_2
s_1	(2, 1)	(0, 0)
s_2	(0, 0)	(1, 2)

3. Battle of the Sexes *II*

Player 2

Player 1	s_1	s_2
s_1	(3, 1)	(0, 0)
s_2	(0, 0)	(1, 3)

4. Battle of the Sexes *III*

Player 2

Player 1	s_1	s_2
s_1	(5, 1)	(0, 0)
s_2	(0, 0)	(1, 5)

5. Telephone Tag *I*

Player 2

Player 1	s_1	s_2
s_1	(0, 0)	(2, 1)
s_2	(1, 2)	(0, 0)

6. Telephone Tag *II*

Player 2

Player 1	s_1	s_2
s_1	(0, 0)	(3, 1)
s_2	(1, 3)	(0, 0)

7. Telephone Tag *III*

Player 2

Player 1	s_1	s_2
s_1	(0, 0)	(5, 1)
s_2	(1, 5)	(0, 0)

8. Chicken *I*

Player 2

Player 1	s_1	s_2
s_1	(6, 6)	(2, 7)
s_2	(7, 2)	(0, 0)

9. Chicken *II*

Player 1 (rows), Player 2 (columns)

	s_1	s_2
s_1	(3, 3)	(2, 4)
s_2	(4, 2)	(0, 0)

10. Chicken *III*

Player 1 (rows), Player 2 (columns)

	s_1	s_2
s_1	(0, 0)	(2, 3)
s_2	(3, 2)	(− 10, − 10)

11. Pickup

Player 1 (rows), Player 2 (columns)

	s_1	s_2
s_1	(4, 4)	(4, 5)
s_2	(5, 4)	(0, 0)

12. Stag Hunt

Player 1 (rows), Player 2 (columns)

	s_1	s_2
s_1	(2, 2)	(0, 1)
s_2	(1, 0)	(1, 1)

13. Winding Road

Player 1 (rows), Player 2 (columns)

	s_1	s_2
s_1	(1, 1)	(0, 0)
s_2	(0, 0)	(1, 1)

14. Intersection

Player 1 (rows), Player 2 (columns)

	s_1	s_2
s_1	(0, 0)	(1, 1)
s_2	(1, 1)	(0, 0)

15. Battle of the Sexes *IV*

Player 1 (rows), Player 2 (columns)

	s_1	s_2
s_1	(3, 1)	(0, 0)
s_2	(0, 0)	(1, 4)

16. Chicken *IV*

Player 1 (rows), Player 2 (columns)

	s_1	s_2
s_1	(6, 6)	(2, 7)
s_2	(8, 2)	(0, 0)

Table 2. Simulation Results of the Rule (2.1) Markov Dynamics						
Game	(s_1, s_1)	(s_1, s_2)	(s_2, s_1)	(s_2, s_2)	"take turns"	mixed
1. Computer Game	0.00%	32.29%	31.10%	0.00%	34.68%	1.93%
2. Battle of the Sexes *I*	37.48%	0.00%	0.00%	36.52%	24.63%	1.37%
3. Battle of the Sexes *II*	39.67%	0.00%	0.00%	40.22%	19.50%	0.61%
4. Battle of the Sexes *III*	42.34%	0.00%	0.00%	42.60%	14.75%	0.31%
5. Telephone Tag *I*	0.00%	36.21%	37.57%	0.00%	25.02%	1.20%
6. Telephone Tag *II*	0.00%	39.97%	40.36%	0.00%	19.09%	0.58%
7. Telephone Tag *III*	0.00%	42.60%	42.81%	0.00%	14.26%	0.33%
8. Chicken *I*	0.00%	36.68%	36.98%	0.00%	24.99%	1.35%
9. Chicken *II*	0.00%	36.33%	36.92%	0.00%	25.40%	1.35%
10. Chicken *III*	0.00%	41.97%	41.67%	0.00%	15.95%	0.41%
11. Pickup	0.00%	41.36%	41.31%	0.00%	16.62%	0.71%
12. Stag Hunt	31.37%	0.00%	0.00%	31.81%	34.88%	1.94%
13. Winding Road	30.39%	0.00%	0.00%	32.14%	35.79%	1.68%
14. Intersection	0.00%	31.51%	31.21%	0.00%	35.31%	1.97%
15. Battle of the Sexes *IV*	30.53%	0.00%	0.00%	55.64%	13.83%	0.00%
16. Chicken *IV*	0.00%	7.21%	79.02%	0.00%	13.77%	0.00%

Table 3. Simulation Results of the Rule (3.1) Markov Dynamics

Game	(s_1, s_1)	(s_1, s_2)	(s_2, s_1)	(s_2, s_2)	"take turns"	mixed
1. Computer Game	0.00%	37.31%	38.87%	0.00%	23.82%	0.00%
2. Battle of the Sexes *I*	37.75%	0.00%	0.00%	38.02%	24.23%	0.00%
3. Battle of the Sexes *II*	36.86%	0.00%	0.00%	38.14%	25.00%	0.00%
4. Battle of the Sexes *III*	37.38%	0.00%	0.00%	37.22%	25.40%	0.00%
5. Telephone Tag *I*	0.00%	37.44%	37.37%	0.00%	25.19%	0.00%
6. Telephone Tag *II*	0.00%	37.55%	37.73%	0.00%	24.72%	0.00%
7. Telephone Tag *III*	0.00%	37.79%	37.57%	0.00%	24.64%	0.00%
8. Chicken *I*	0.00%	38.68%	37.53%	0.00%	23.79%	0.00%
9. Chicken *II*	0.00%	38.75%	36.48%	0.00%	24.77%	0.00%
10. Chicken *III*	0.00%	38.32%	37.03%	0.00%	24.65%	0.00%
11. Pickup	0.00%	36.91%	38.54%	0.00%	24.55%	0.00%
12. Stag Hunt	37.68%	0.00%	0.00%	37.07%	25.25%	0.00%
13. Winding Road	36.95%	0.00%	0.00%	37.71%	25.34%	0.00%
14. Intersection	0.00%	37.72%	36.78%	0.00%	25.50%	0.00%
15. Battle of the Sexes *IV*	12.83%	0.00%	0.00%	67.66%	19.51%	0.00%
16. Chicken *IV*	0.00%	0.19%	97.13%	0.00%	2.68%	0.00%

Appendix 2. Proofs of §2 Results

Proposition 1. (*i*) If the transition matrices of Markov fictitious play make a strict Nash equilibrium an absorbing state of the Markov chains, then Markov deliberators who ever visit this state will always follow this equilibrium. (*ii*) If the transition matrices make a "taking turns" equilibrium a periodic set of the Markov chains, then deliberators who enter into this set always follow this equilibrium.

PROOF. (*i*) Let $s^*_{ij} = (s_{i^*}, s_{j^*})$ be a strict Nash equilibrium. Assume that $\alpha^T_{i^*j^*,i^*j^*} = \beta^T_{i^*j^*,i^*j^*} = 1$, that is, s^*_{ij} is an absorbing state of the Markov chains α^T and β^T defined by (2.*a*.1) and (2.*a*.2). Suppose that $s^T_{ij} = s^*_{ij}$, that is, Markov deliberators follow s^*_{ij} at round T. Then at round $T+1$, by (2.1′)

$$\begin{aligned}\alpha^{T+1}_{i^*j^*,i^*j^*} &= \frac{1_{s_{i^*j^*,i^*j^*}}(T)}{n^T_{i^*j^*} + \lambda_1} + \frac{n^{T-1}_{i^*j^*} + \lambda_1}{n^T_{i^*j^*} + \lambda_1} \cdot \alpha^T_{i^*j^*,i^*j^*} \\ &= \frac{1_{s_{i^*j^*,i^*j^*}}(T)}{n^T_{i^*j^*} + \lambda_1} + \frac{n^{T-1}_{i^*j^*} + \lambda_1}{n^T_{i^*j^*} + \lambda_1} \cdot 1 \\ &= \frac{n^T_{i^*j^*} + \lambda_1}{n^T_{i^*j^*} + \lambda_1} = 1\end{aligned}$$

and similarly

$$\alpha^{T+1}_{i^*j^*,kl} = \alpha^T_{i^*j^*,kl} = 0 \text{ for } s_{kl} \neq s_{i^*j^*} \,.$$

So at round $T+1$, Player 1's expected payoff if she follows the pure strategy s_l is

$$E^{T+1}_1(u_1(s_{l\cdot})|s^*_{ij}) = \sum_j u_1(s_{lj}) \cdot \mu^{T+1}_1(s_{\cdot j}|s^*_{ij}) = u_1(s_l, s_{j^*}) \cdot 1$$

and so

$$E^{T+1}_1(u_1(s^*_{i\cdot})|s^*_{ij}) = u_1(s^*_{ij}) > u_1(s_l, s_{j^*}) = E^{T+1}_1(u_1(s_{l\cdot})|s^*_{ij}) \text{ for } l \neq i$$

because s^*_{ij} is a strict equilibrium. Hence Player 1 will follow s_i at round $T+1$ and by a symmetric argument Player 2 will follow s_j at $T+1$. By induction, the players will follow s^*_{ij} for all rounds after T.

(*ii*) The proof is essentially the Case (*i*) proof applied twice. Let $s^*_{ij} = (s_{i^*}, s_{j^*})$ and $s^*_{kl} = (s_{k^*}, s_{l^*})$ both be strict Nash equilibria. Assume that

$$\alpha^T_{i^*j^*,k^*l^*} = \beta^T_{i^*j^*,k^*l^*} = \alpha^T_{k^*l^*,i^*j^*} = \beta^T_{k^*l^*,i^*j^*} = 1\,,$$

that is, s_{ij} and s^*_{kl} together form a periodic set of the Markov chains α^T and β^T. Suppose that $s^T_{ij} = s^*_{ij}$. Then at round $T+1$, by (2.1′)

$$\begin{aligned}\alpha^{T+1}_{i^*j^*,k^*l^*} &= \frac{1_{s_{i^*j^*,k^*l^*}}(T)}{n^T_{i^*j^*}+\lambda_1} + \frac{n^{T-1}_{i^*j^*}+\lambda_1}{n^T_{i^*j^*}+\lambda_1}\cdot\alpha^T_{i^*j^*,k^*l^*}\\ &= \frac{1_{s_{i^*j^*,k^*l^*}}(T)}{n^T_{i^*j^*}+\lambda_1} + \frac{n^{T-1}_{i^*j^*}+\lambda_1}{n^T_{i^*j^*}+\lambda_1}\cdot 1\\ &= \frac{n^T_{i^*j^*}+\lambda_1}{n^T_{i^*j^*}+\lambda_1} = 1\end{aligned}$$

and similarly

$$\alpha^{T+1}_{i^*j^*,kl} = \alpha^T_{i^*j^*,kl} = 0 \text{ for } s_{kl} \neq s_{k^*l^*}\,.$$

So at round $T+1$, Player 1's expected payoff if she follows the pure strategy s_l is

$$E^{T+1}_1(u_1(s_{l\cdot})|s^*_{ij}) = \sum_j u_1(s_{lj})\cdot\mu^{T+1}_1(s_{\cdot j}|s^*_{ij}) = u_1(s_l, s_{l^*})\cdot 1$$

and so

$$E^{T+1}_1(u_1(s^*_{k\cdot})|s^*_{ij}) = u_1(s^*_{kl}) > u_1(s_l, s_{l^*}) = E^{T+1}_1(u_1(s_{l\cdot})|s^*_{ij}) \text{ for } l \neq i$$

because s^*_{kl} is a strict equilibrium. Hence Player 1 will follow s^*_i at round $T+1$ and by a symmetric argument Player 2 will follow s^*_l at $T+1$.

This gives us $s^{T+1}_{ij} = s^*_{kl}$. But then at round $T+2$, by (2.*i*′),

$$\begin{aligned}\alpha^{T+2}_{k^*l^*,i^*j^*} &= \frac{1_{s_{k^*l^*,i^*j^*}}(T+1)}{n^{T+1}_{k^*l^*}+\lambda_1} + \frac{n^{T}_{k^*l^*}+\lambda_1}{n^{T+1}_{k^*l^*}+\lambda_1}\cdot\alpha^{T+1}_{k^*l^*,i^*j^*} \\ &= \frac{1_{s_{k^*l^*,i^*j^*}}(T+1)}{n^{T+1}_{k^*l^*}+\lambda_1} + \frac{n^{T}_{k^*l^*}+\lambda_1}{n^{T+1}_{k^*l^*}+\lambda_1}\cdot 1 \\ &= \frac{n^{T+1}_{k^*l^*}+\lambda_1}{n^{T+1}_{k^*l^*}+\lambda_1} = 1\end{aligned}$$

and similarly

$$\alpha^{T+1}_{k^*l^*,kl} = \alpha^{T}_{k^*l^*,kl} = 0 \text{ for } s_{kl} \neq s_{i^*j^*}\,.$$

So at round $T+2$, Player 1's expected payoff if she follows the pure strategy s_l is

$$E^{T+2}_1(u_1(s_{l\cdot})|s^*_{kl}) = \sum_j u_1(s_{lj})\cdot\mu^{T+2}_1(s_{\cdot j}|s^*_{kl}) = u_1(s_l, s_{j^*})\cdot 1$$

and so

$$E^{T+2}_1(u_1(s^*_{i\cdot})|s^*_{kl}) = u_1(s^*_{ij}) > u_1(s_l, s_{j^*}) = E^{T+2}_1(u_1(s_{l\cdot})|s^*_{kl}) \text{ for } l \neq i$$

because s^*_{ij} is a strict equilibrium. Hence Player 1 will follow s^*_i at round $T+2$ and by a symmetric argument Player 2 will follow s^*_j at $T+2$. By induction, the players will take turns between s^*_{ij} and s^*_{kl} for all rounds after T.

Note that a symmetric proof yields this case if the players visit s^*_{kl} at round T, so this periodic set is a fixed point of Markov fictitious play no matter where Markov deliberators enter into the set. □

Proposition 2. (*i*) If Markov deliberators ever follow a strict Nash equilibrium s^* in consecutive rounds, then they will follow s^* all subsequent rounds. (*ii*) If s_1^* and s_2^* are distinct strict Nash equilibria and Markov deliberators ever follow s_1^*, s_2^* and s_1^* in consecutive rounds, then they will subsequently always alternate between s^* and s^{**} in consecutive rounds.

PROOF. The notations used in Proposition 1 remain in force here. If Markov deliberators follow the pure strategy profile $s^m = (s_1^m, s_2^m)$ at round m of deliberation, then

(1)
$$E_1^m(u_1(s_{1\cdot}^m)|s^{m-1}) \geq E_1^m(u_1(s_{l\cdot})|s^{m-1}) \text{ for } s_l \neq s_1^m$$
$$E_2^m(u_2(s_{\cdot 2}^m)|s^{m-1}) \geq E_2^m(u_2(s_{\cdot k})|s^{m-1}) \text{ for } s_k \neq s_2^m$$

which implies that

(1′)
$$\sum_{s_j}[u_1(s_1^m,s_j) - u_1(s_l, s_j)]\mu_1^m(s_{\cdot j}|s^{m-1}) \geq 0$$
$$\sum_{s_i}[u_2(s_i, s_2^m) - u_2(s_i, s_k)]\mu_2^m(s_{i\cdot}|s^{m-1}) \geq 0$$

The proofs of (*i*) and (*ii*) are quite similar:

(*i*) Suppose that s_{ij}^* is a strict Nash equilibrium. Assume $s^T = s^{T-1} = s_{ij}^*$. By (1) we have

(*i*.1)
$$E_1^T(u_1(s_{i^*\cdot})|s_{ij}^*) \geq E_1^T(u_1(s_{l\cdot})|s_{ij}^*) \text{ for } s_l \neq s_{k^*} .$$

We want to show that at round $T+1$,

(*i*.2)
$$E_1^{T+1}(u_1(s_{i^*\cdot})|s_{ij}^*) > E_1^{T+1}(u_1(s_{l\cdot})|s_{ij}^*) \text{ for all } s_l \neq s_{i^*}.$$

We have

$$E_1^{T+1}(u_1(s_{i^*\cdot})|s^*_{ij}) - E_1^{T+1}(u_1(s_{l\cdot})|s^*_{ij}) = \sum_{s_j}[u_1(s_{i^*},s_j) - u_1(s_l, s_j)]\mu_1^{T+1}(s_{\cdot j}|s^*_{ij})$$

$$= \sum_{s_j}[u_1(s_{i^*},s_j) - u_1(s_l, s_j)]\Big(\frac{1_{s_{i^*j^*},\cdot}(T)}{n^T_{i^*j^*} + \lambda_1} + \frac{n^{T-1}_{i^*j^*} + \lambda_1}{n^T_{i^*j^*} + \lambda_1} \cdot \mu_1^T(s_{\cdot j}|s^*_{ij})\Big)$$

$$= [u_1(s_{i^*},s_{j^*}) - u_1(s_l, s_{j^*})] \cdot \frac{1 + n^{T-1}_{i^*j^*} + \lambda_1}{n^T_{i^*j^*} + \lambda_1}$$

$$+ \frac{n^{T-1}_{i^*j^*} + \lambda_1}{n^T_{i^*j^*} + \lambda_1} \sum_{s_j \neq s_{j^*}} [u_1(s_{i^*},s_j) - u_1(s_l, s_j)] \cdot \mu_1^T(s_{\cdot j}|s^*_{ij})$$

because by hypothesis at $T-1$ the deliberators visited s^*_{ij}, so

$$1_{s_{i^*j^*},s_{ij}}(T-1) = \begin{cases} 1 \text{ if } s_{ij} = s_{i^*j^*} \\ 0 \text{ otherwise} \end{cases} .$$

Hence,

$$E_1^{T+1}(u_1(s_{i^*\cdot})|s^*_{ij}) - E_1^{T+1}(u_1(s_{l\cdot})|s^*_{ij}) = \frac{1}{n^T_{i^*j^*} + \lambda_1}[u_1(s_{i^*},s_{j^*}) - u_1(s_l, s_{j^*})]$$

$$+ \frac{n^{T-1}_{i^*j^*} + \lambda_1}{n^T_{i^*j^*} + \lambda_1}\sum_{s_j}[u_1(s_{i^*},s_j) - u_1(s_l, s_j)] \cdot \mu_1^T(s_{\cdot j}|s^*_{ij})$$

In the right member this equation, the first term is positive because s^*_{ij} is a strict Nash equilibrium, and the second term is nonnegative by (1′) and (*i*.1). This establishes (*i*.2).

By a similar argument,

$$E_2^{T+1}(u_2(s_{\cdot j^*})|s^*_{ij}) > E_2^{T+1}(u_2(s_{\cdot k})|s^*_{ij}) \text{ for } s_k \neq s_{j^*} ,$$

So at round $T+1$, Players 1 and 2 will follow $s^{T+1} = s^*_{ij}$. (*i*) follows by complete induction on T.

(*ii*) Suppose that $s^*_1 = s^*_{ij}$ and $s^*_2 = s^*_{kl}$ are both strict Nash equilibria. Assume $s^T = s^{T-2} \neq s^{T-1}$, and $s^{T-1} = s^*_{ij}$ and $s^{T-2} = s^T = s^*_{kl}$. By (1) we have

(*ii*.1) $$E_1^{T-1}(u_1(s_{i^*\cdot})|s^*_{kl}) \geq E_1^{T-1}(u_1(s_{l\cdot})|s^*_{kl}) \text{ for } s_l \neq s_{i^*} \text{ , and}$$
$$E_1^{T}(u_1(s_{k^*\cdot})|s^*_{ij}) \geq E_1^{T}(u_1(s_{l\cdot})|s^*_{ij}) \text{ for } s_l \neq s_{k^*} .$$

We want to show that at round $T+1$,

(*ii*.2) $$E_1^{T+1}(u_1(s_{i^*\cdot})|s^*_{kl}) > E_1^{T+1}(u_1(s_{l\cdot})|s^*_{kl}) \text{ for all } s_l \neq s_{i^*}.$$

We have

$$E_1^{T+1}(u_1(s_{i^*\cdot})|s_{kl}^*) - E_1^{T+1}(u_1(s_{l\cdot})|s_{kl}^*) = \sum_{s_j}[u_1(s_{i^*},s_j) - u_1(s_l, s_j)]\mu_1^{T+1}(s_{\cdot j}|s_{kl}^*)$$

$$= \sum_{s_j}[u_1(s_{i^*},s_j) - u_1(s_l, s_j)]\Big(\frac{1_{s_{k^*l^*},\cdot}(T)}{n_{k^*l^*}^T + \lambda_1} + \frac{n_{k^*l^*}^{T-1} + \lambda_1}{n_{k^*l^*}^T + \lambda_1} \cdot \mu_1^T(s_{\cdot j}|s_{kl}^*)\Big)$$

$$= \frac{n_{k^*l^*}^{T-1} + \lambda_1}{n_{k^*l^*}^T + \lambda_1}\sum_{s_j}[u_1(s_{i^*},s_j) - u_1(s_l, s_j)] \cdot \mu_1^T(s_{\cdot j}|s_{kl}^*)$$

because by hypothesis at $T-1$ the deliberators visited s_{ij}^*, so at T, $1_{s_{k^*l^*},s_{kl}}(T) = 0$ for all s_{kl}. Now note that

$$(ii.3)\quad \sum_{s_j}[u_1(s_{i^*},s_j) - u_1(s_l, s_j)]\mu_1^T(s_{\cdot j}|s_{kl}^*)$$

$$= \sum_{s_j}[u_1(s_{i^*},s_j) - u_1(s_l, s_j)]\Big(\frac{1_{k^*l^*,ij}(T-1) + n_{k^*l^*}^{T-2} + \lambda_1}{n_{k^*l^*}^{T-1} + \lambda_1}\Big)\mu_1^{T-1}(s_{\cdot j}|s_{kl}^*)$$

$$= [u_1(s_{i^*},s_{j^*}) - u_1(s_l, s_{j^*})] \cdot \frac{1 + n_{k^*l^*}^{T-2} + \lambda_1}{n_{k^*l^*}^{T-1} + \lambda_1}$$

$$+ \frac{n_{k^*l^*}^{T-2} + \lambda_1}{n_{k^*l^*}^{T-1} + \lambda_1}\sum_{s_j \neq s_{j^*}}[u_1(s_{i^*},s_j) - u_1(s_l, s_j)]\mu_1^{T-1}(s_{\cdot j}|s_{kl}^*)$$

$$= \frac{1}{n_{k^*l^*}^{T-1} + \lambda_1}[u_1(s_{i^*},s_{j^*}) - u_1(s_l, s_{j^*})]$$

$$+ \frac{n_{k^*l^*}^{T-2} + \lambda_1}{n_{k^*l^*}^{T-1} + \lambda_1}\sum_{s_j}[u_1(s_{i^*},s_j) - u_1(s_l, s_j)]\mu_1^{T-1}(s_{\cdot j}|s_{kl}^*)$$

because by hypothesis at $T-2$ the deliberators visited s_{kl}^* and at $T-1$ the deliberators visited s_{ij}^*, so

$$1_{s_{k^*l^*},s_{ij}}(T-1) = \begin{cases} 1 \text{ if } s_{ij} = s_{i^*j^*} \\ 0 \text{ otherwise} \end{cases}.$$

In the rightmost member of equation (*ii*.3), the first term is positive because s_{ij}^* is a strict Nash equilibrium, and the second term is nonnegative by (1′) and (*ii*.1). This establishes (*ii*.2).

By similar arguments,

$$E_2^{T+1}(u_2(s_{\cdot j^*})|s_{kl}^*) > E_2^{T+1}(u_2(s_{\cdot k})|s_{kl}^*) \text{ for } s_k \neq s_{j^*} \,,$$
$$E_1^{T+2}(u_1(s_{k^*\cdot})|s_{ij}^*) > E_1^{T+2}(u_1(s_{l\cdot})|s_{ij}^*) \text{ for } s_l \neq s_{k^*} \text{, and}$$
$$E_2^{T+2}(u_2(s_{\cdot l^*})|s_{ij}^*) > E_1^{T+2}(u_2(s_{\cdot k})|s_{ij}^*) \text{ for } s_k \neq s_{l^*} \,.$$

So at round $T+1$, Players 1 and 2 will follow $s^{T+1} = s_{ij}^*$ and at round $T+2$ they will follow $s^{T+2} = s_{kl}^*$. *(ii)* follows by complete induction on T. □

NOTES

[1]Fudenberg and Levine (1998, Chapter 2) review the best-known convergence theorems for fictitious play. Shapley (1964) was the first to present an example of a game for which the distribution of fictitious play cycles rather than converges whenever the priors are not initially at Nash equilibrium. Richards (1997) gives an example of a game in which the distribution of fictitious play is *chaotic*.

[2]Player i's conditional probability $\mu_i^T(\cdot|B)$ given the event B is defined as

$$\mu_i^T(A|B) = \frac{\mu_i^T(A\cap B)}{\mu_i^T(B)} \text{ if } \mu_i^T(B) > 0 \text{ , and}$$

$$\mu_i^T(A|B) = 0 \text{ if } \mu_i^T(B) = 0$$

for any event A.

[3]This identity is easy to derive. We have

$$\mu_1^T(s_{kl}^T|s_{ij}^{T-1}) = \frac{n_{ij,kl}^{T-1} + \lambda_1\alpha_{ij,kl}^0}{n_{ij}^{T-1} + \lambda_1}$$

and

$$\begin{aligned}\mu_1^{T+1}(s_{kl}^{T+1}|s_{ij}^T) &= \frac{n_{ij,kl}^T + \lambda_1\alpha_{ij,kl}^0}{n_{ij}^T + \lambda_1} = \frac{n_{ij,kl}^{T-1} + 1_{s_{ij,kl}}(T) + \lambda_1\alpha_{ij,kl}^0}{n_{ij}^T + \lambda_1} \\ &= \frac{1_{s_{ij,kl}}(T)}{n_{ij}^T + \lambda_1} + \frac{1}{n_{ij}^T + \lambda_1}\cdot\left(n_{ij,kl}^{T-1} + \lambda_1\alpha_{ij,kl}^0\right) \\ &= \frac{1_{s_{ij,kl}}(T)}{n_{ij}^T + \lambda_1} + \frac{n_{ij}^{T-1} + \lambda_1}{n_{ij}^T + \lambda_1}\cdot\left(\frac{1}{n_{ij}^{T-1} + \lambda_1}\left(n_{ij,kl}^{T-1} + \lambda_1\alpha_{ij,kl}^0\right)\right) \\ &= \frac{1_{s_{ij,kl}}(T)}{n_{ij}^T + \lambda_1} + \frac{n_{ij}^{T-1} + \lambda_1}{n_{ij}^T + \lambda_1}\cdot\mu_1^T(s_{kl}^T|s_{ij}^{T-1})\ .\end{aligned}$$

[4]By *nonnegligible set* we mean a set in the simplex of positive Lebesgue measure.

[5]At the "taking turns" equilibrium of the Figure 2 Computer Game where Jan and Jill alternate between s_{12} and s_{21}, the inequalities

---

(3.*i*) $$E_1^{T+1}(u_1(s_{1\cdot})|s_{12}^T) < E_1^{T+1}(u_1(s_{2\cdot})|s_{12}^T)$$

(3.*ii*) $$E_2^{T+1}(u_2(s_{\cdot 1})|s_{12}^T) > E_2^{T+1}(u_2(s_{\cdot 2})|s_{12}^T)$$

(3.*iii*) $$E_1^{T+1}(u_1(s_{1\cdot})|s_{21}^T) > E_1^{T+1}(u_1(s_{2\cdot})|s_{21}^T)$$

(3.*iv*) $$E_2^{T+1}(u_2(s_{\cdot 1})|s_{21}^T) < E_2^{T+1}(u_2(s_{\cdot 2})|s_{21}^T)$$

must all be satisfied. Inequality (3.*i*) is satisfied when

$$\begin{aligned}
1\cdot(\alpha_{12,11}^{T+1}+\alpha_{12,21}^{T+1})+2\cdot(\alpha_{12,12}^{T+1}+\alpha_{12,22}^{T+1}) &< 3\cdot(\alpha_{12,11}^{T+1}+\alpha_{12,21}^{T+1})+0\cdot(\alpha_{12,12}^{T+1}+\alpha_{12,22}^{T+1})\\
2(\alpha_{12,12}^{T+1}+\alpha_{12,22}^{T+1}) &< 2(\alpha_{12,11}^{T+1}+\alpha_{12,21}^{T+1})\\
\alpha_{12,12}^{T+1}+\alpha_{12,22}^{T+1} &< \alpha_{12,11}^{T+1}+\alpha_{12,21}^{T+1}\\
\alpha_{12,12}^{T+1}+1-\alpha_{12,11}^{T+1}-\alpha_{12,21}^{T+1}-\alpha_{12,12}^{T+1} &< \alpha_{12,11}^{T+1}+\alpha_{12,21}^{T+1}\\
1-\alpha_{12,11}^{T+1}-\alpha_{12,21}^{T+1} &< \alpha_{12,11}^{T+1}+\alpha_{12,21}^{T+1}\\
1-2\alpha_{12,11}^{T+1} &< 2\alpha_{12,21}^{T+1}\\
\frac{1}{2}(1-2\alpha_{12,11}^{T+1}) &< \alpha_{12,21}^{T+1}
\end{aligned}$$

Similarly, inequalities (3.*ii*), (3.*iii*) and (3.*iv*) are satisfied when

(3.*ii*′) $$\beta_{12,21}^{T+1} > \frac{1}{2} - \beta_{12,22}^{T+1}$$

(3.*iii*′) $$\alpha_{21,12}^{T+1} > \frac{1}{2} - \alpha_{21,22}^{T+1}$$

(3.*iv*′) $$\beta_{21,12}^{T+1} > \frac{1}{2} - \beta_{21,11}^{T+1}$$

so (3.*ii*′), (3.*iii*′) and (3.*iv*′) together with

(3.*i*′) $$\alpha_{12,21}^{T+1} > \frac{1}{2} - \alpha_{12,11}^{T+1}$$

show that the "taking turns" equilibrium has a nonnegligible basin of attraction.

[6]The simulations were programmed and run in GAUSS.