Extending Fictitious Play with Pattern Recognition

November 13, 2013

Author:
Ronald Chu (info@ronaldchu.nl)

Supervisor:
dr. G.A.W. Vreeswijk

Abstract

Fictitious play, an algorithm to predict the opponents next move based on the observed history of play, is one of the oldest simple yet very effective algorithms in game theory. Although using pattern recognition as a more sophisticated way to analyze the history of play seems a logical step, there is little research available on this subject. In this thesis we will examine two different types of pattern recognition, and formulate several algorithms that incorporate these approaches. These algorithms and the basic fictitious play variants they extend are empirically tested in eight tournaments on some well known formal-form games. The results obtained will show that adding pattern recognition to fictitious play improves performance, and demonstrate the general possibilities of applying pattern recognition to agents in game theory.

Contents

Abstract **1**

Contents **2**

1 Introduction **3**

2 Fictitious play **4**
- 2.1 Brown's original algorithm . . . 4
- 2.2 Convergence properties . . . 6
- 2.3 Adaptations . . . 7
  - 2.3.1 Smoothed fictitious play . . . 8
  - 2.3.2 Weighted fictitious play . . . 9

3 Pattern recognition **9**
- 3.1 N-Period fictitious play . . . 10
- 3.2 Cyclic pattern detection . . . 11

4 Experiment Setup **12**

5 Results per game **14**
- 5.1 Asymmetric coordination game . . . 14
- 5.2 Symmetric coordination game . . . 15
- 5.3 Shapley's Game . . . 15
- 5.4 Battle of the Sexes . . . 16
- 5.5 Chicken . . . 17
- 5.6 Matching Pennies . . . 18
- 5.7 Prisoner's Dilemma . . . 18
- 5.8 Random games . . . 18

6 Pattern Recognition Results **19**
- 6.1 Cyclic pattern recognition . . . 19
- 6.2 N-Period fictitious play . . . 20

7 Conclusion **20**

References **21**

A Tournament Results **24**
- A.1 Asymmetric coordination game . . . 24
- A.2 Symmetric coordination game . . . 26
- A.3 Shapley's game . . . 28
- A.4 Battle of the Sexes . . . 30
- A.5 Chicken . . . 32
- A.6 Matching pennies . . . 34
- A.7 Prisoner's dilemma . . . 36
- A.8 Random generated games . . . 37

1 Introduction

The field of game theory studies strategic decision making in so called games where a fixed number of players have a limited number of possible actions to choose from. The combination of the actions of all players, their joint actions, determines the outcome of the game. It is used in numerous scientific fields to either describe or prescribe behavior, and forms a theoretical basis in the field of multi-agent systems. The first know game theoretical discussion occurred in the early 18th century in a letter by James Waldegrave[2] about a gambling situation in the card game piquet and the board game trictrac. Waldegrave describes what we would classify today as a minimax solution to a zero-sum two-person game. In 1928 the german mathematician John von Neumann published the ground-breaking article "Zur theorie der gesellschaftsspiele"[31], and modern game theory as a scientific field truly started with the publication of the book "Theory of Games and Economic Behavior" by von Neumann and economist Oskar Morgenstern[33] in 1944.

The mathematician George W. Brown (who directed the construction of the von Neumann computer at Princeton in 1946) worked at the RAND institute where he applied von Neumann's mathematical game theory in military combat situations[24]. Here he introduced fictitious play (FP) as an algorithm to find the value of a zero-sum game, first in an internal RAND report[5] and later in his 1951 publication[6] where he describes the algorithm:

> "The iterative method in question can be loosely characterized by the fact that it rests on the traditional statistician's philosophy of basing future decisions on the relevant history. Visualize two statisticians, perhaps ignorant of min-max theory, playing many plays of the same discrete zero-sum game. One might naturally expect a statistician to keep track of the opponent's past plays and, in the absence of a more sophisticated calculation, perhaps to choose at each play the optimum pure strategy against the mixture represented by all the opponent's past plays."

In other words, fictitious play uses the observations from the past to build an observed frequency distribution of the actions of its opponent(s) and uses the action(s) with the highest observed frequency as a prediction of the opponent(s) next move. It then chooses a strictly myopic response to maximize its expected payoff. It was created by Brown, and first investigated by Robbinson[25] who proved its validity in 1951, as a heuristic for computing Nash equilibria by playing a fictitious game against itself. Note the term fictitious play may be misleading in its current use, where it is used to model a actual player, since there is nothing fictional about the observed history of play and the algorithm does not "play a fictional game in its head" to make its predictions.

Fictitious play shares some similarities with reinforcement learning, which does not look at the actions of the opponent but uses the (cumulative) payoff of actions played in the past to decide its next move, and no-regret learning, which choses the action with the least amount of regret[1] as its next move. All three algorithms (or common adaptations thereof) use a statistical variable to

[1] Regret is the difference between the received payoff thus far and the payoff it could have received if it had played a pure strategy all along (assuming the opponent had played the exact same actions).

determine their next move and get slower in changing their behavior as the game progresses. Fictitious play, however, does not try to maximize its received payoff directly. Instead of predicting what actions are most profitable it predicts the action(s) of the opponent(s), and then uses a myopic best response strategy to maximize received payoff.

A next logical step in the evolution of the fictitious play algorithm is to perform a more sophisticated analysis of the observed history of play. There are of course many different approaches one can take to perform such an analysis. In this thesis we will focus on performing pattern recognition; looking at sequences of actions in stead of single actions in the history of play. Although pattern recognition is a broad and established field with much ongoing research, there exists, as Spiliopoulos puts it, "a conspicuous gap in the literature regarding learning in games – the absence of empirical verification of learning rules involving pattern recognition" [29].

In the next chapter will we give a thorough overview of the basic fictitious play algorithm and two common adaptations. We will then examine two different approaches to add pattern recognition, and formulate several algorithms that extend fictitious play with the proposed pattern recognition. Finally we will formulate a set of experiments to obtain empirical data on the performance of the pattern recognizing algorithms and the basic fictitious play algorithms they extend. An analysis of the data obtained will show if, how and why pattern recognition influences performance.

2 Fictitious play

2.1 Brown's original algorithm

$$\begin{array}{c c} & \begin{array}{cc} a & b \end{array} \\ \begin{array}{c} A \\ B \end{array} & \begin{pmatrix} 1,1 & 0,0 \\ 0,0 & 2,2 \end{pmatrix} \end{array}$$

Fig. 2.1: A 2-player normal-form game in a typical payoff matrix representation

A normal-form game in game theory is a description of a game in the form of all players' strategy spaces and payoff functions. A n-player game has a finite strategy space $X = \prod_i X_i$. The strategy space X_i for each individual player $i = 1, 2, \ldots, n$ consists of all actions available to player i. In the coordination game displayed in fig. 2.1 the strategy space for the first player (the row player) $X_1 = \{A, B\}$, and the strategy space X consists of the product of the strategy spaces of both players $\{\{A, a\}, \{A, b\}, \{B, a\}, \{B, b\}\}$. There is a payoff (or utility) function $u_i(x)$ for each player i which determines the utility received by player i given the joint actions $x \in X$ played that round. In case of the coordination game the utility $u_i(\{A, a\}) = 1$ for both players can be easily read from the payoff matrix in fig. 2.1.

As mentioned before FP uses the observed history of play to determine its actions. Because the lack of history the first action x_i^1 is specified arbitrarily[2]. The algorithm consists of two components: a forecasting and a response component. The game is played several rounds, and for each time index (or round number) $t \in \mathbb{N}^0$ the n-tuple $x^t = (x_1^t, x_2^t, \ldots, x_n^t)$ where $x^t \in X$, consists of the actions played that round by all players. Let $h^t = (x^1, x^2, \ldots, x^t)$ be the

[2]The superscript t in x^t indicates a time index, not an exponent. When exponentiation is intended the base is a number or between brackets

observed history up to round t. The forecasting function $p_i^t(x_j)$, where $x_j \in X_j$, returns the proportion of the time when x_j was played in the history h^t:

$$p_i^t(x_j) = \frac{\sum_{u=1}^{t} I^{t-u}(x_j)}{t} \tag{2.1}$$

Let $p^t = \prod_i p_i^t$ be the associated product distribution. The indicator function $I^t(x_j)$ returns 1 if $x_j = x_j^t$ (action x_j was played by player j at time t) and returns 0 otherwise. Note that $\sum_{x_i \in X_i} p_i^t(x_i) = 1$ for every i and every t.

The response is an action from the best-reply correspondence $BR_i(p^t)$; the set of all actions of i that maximize expected utility given p^t. The utility of a probability distribution is defined as the expected utility of its outcomes by defining $u_i(p) = \sum_{x \in X} u_i(x)p(x)$. Let Δ_i denote the set of all probability distributions on the set X_i, and let $\Delta = \prod \Delta_i$. Let $p_i^t \in \Delta_i$ and $p^t \in \Delta$ denote the observed probability distribution in round t for the actions of player i and the joint actions of all players, respectively. For every $p \in \Delta$ and every $x \in X$ the probability of x is $p^t(x) = \prod_i p_i^t(x_i)$. To be able to predict the expected utility for each action we need to be able to distinguish between our actions X_i and the actions of the other players $X_{-i} = \prod_{i \neq j} X_j$. In the same fashion let $\Delta_{-i} = \prod_{j \neq i} \Delta_j$ and let $p_{-i}^t \in \Delta_{-i}$. Now we can combine an action x_i and a partial probability distribution p_{-i} to obtain a full probability distribution $(x_i, p_{-i}) \in \Delta$ that places a probability 1 on x_i, and define the best-reply correspondence as:

$$BR_i(p_{-i}^t) = \{x_i \in X_i : u_i(x_i, p_{-i}^t) \geq u_i(x_i', p_{-i}^t) \text{ for all } x_i' \in X_i\}.$$

The algorithm as described by Brown is a process where player update their beliefs and determine their next move alternatingly. In this interpretation an (alternating) fictitious play process is a sequence (i_t, j_t) such that

$$i_{t+1} \in BR_i(p_j^t) \quad \text{and} \quad j_t \in BR_j(p_i^t)$$

if $i_1 \in X_i$. As pointed out by Berger[4] almost all later work on fictitious play employs simultaneous belief updating, which was described by Brown[5] merely as an alternate notion. This is not surprising since normal-form games usually apply simultaneous belief updating (or at least imperfect information; no player has information about the other players' moves in the current round). Games where players choose an action alternatingly and observe the other player's actions as they are played are usually represented in extensive-form. We will not be an exception and use simultaneous belief updating. The definition of a simultaneous fictitious play process is a function $f(h^t) = x^{t+1}$ such that

$$x^{t+1} = f(h^t) \Rightarrow x_i^{t+1} \in BR_i(p^t) \text{ for all } t \geq t_0.$$

Table 1 shows the first rounds of two players playing the coordination game in figure 2.1. Player 1 is using fictitious play, the first moves were predefined in this game and the tie-breaking rule, applied when there is more than one action in the best-reply correspondence, is to pick one at random. In round 2 the observed history consists of one observation, so the probability that the opponent will play action a is $p_1^1(A) = 1.0$. In round 2 the probabilities are equal, but because of the different payoffs $BR = \{B\}$. In rounds 4 and 10 there is more than one action in the best-reply correspondence and one needs to be picked at random.

t	$p_2^t(a)$	$p_2^t(b)$	$u_1(A, p_{-1}^t)$	$u_1(B, p_{-1}^t)$	BR_1	x_1^t	x_2^t
1	N/A	N/A	N/A	N/A	N/A	A	a
2	1.00	0.00	1.00	0.00	{A}	A	b
3	0.50	0.50	0.50	1.00	{B}	B	a
4	0.67	0.33	0.67	0.67	{A,B}	B	a
5	0.75	0.25	0.75	0.50	{A}	A	b
6	0.60	0.40	0.60	0.80	{B}	B	b
7	0.50	0.50	0.50	1.00	{B}	B	a
8	0.57	0.43	0.57	0.86	{B}	B	a
9	0.62	0.38	0.62	0.75	{B}	B	a
10	0.67	0.33	0.67	0.67	{A,B}	B	a
11	0.70	0.30	0.70	0.60	{A}	A	b
12	0.64	0.36	0.64	0.73	{B}	B	b

Table 1: Fictitious play (player 1) in a coordination game

2.2 Convergence properties

Fictitious play was designed and proven to find the value game, which corresponds to a Nash equilibrium (NE) in zero-sum games. A Nash equilibrium is a probability distribution $p^* \in \Delta$ such that for every player i and every distribution $p \in \Delta$ it holds that:

$$u_i(p_i^*, p_{-i}^*) \geq u_i(p_i, p_{-i}^*). \tag{2.2}$$

$$\begin{array}{cc} & \begin{array}{cc} A & B \end{array} \\ \begin{array}{c} A \\ B \end{array} & \begin{pmatrix} 1,1 & 0,0 \\ 0,0 & 1,1 \end{pmatrix} \end{array}$$

Fig. 2.2: A coordination game

A distribution p^* corresponds to a pure-strategy equilibrium if $p^*(x) = 1$ for some joint action set $x \in X$, and corresponds to a strict equilibrium if equation 2.2 holds strictly for every i and every $p \neq p^*$. Note that because in a mixed equilibrium some player has more than one best reply every strict equilibrium must also be a pure-strategy equilibrium. A game has the fictitious play property if every limit point of every sequence generated by such a fictitious play process corresponds to a Nash equilibrium of the game. In other words, if such a sequence converges to the closed set of Nash equilibria, which is weaker than requiring every sequence to converge to a specific NE. Note that if the marginal frequency distribution converges to a mixed equilibrium it is still possible that the period-by-period behavior is nowhere near equilibrium. In the coordination game (fig. 2.2) for example it is possible that two players applying fictitious play miscoordinate every turn alternating between $\{A, B\}$ and $\{B, A\}$, resulting in a marginal frequency distribution that corresponds to the mixed Nash equilibrium. Classes of games that have been proven to have the fictitious play property are:

Finite zero-sum two-player games (Robinson[25])
: Zero-sum games are games where for every joint action set $\{x_1, \ldots, x_n\}$ the utility received by all players always sums up to 0.

Nondegenerate $2 \times n$ games (Miyazawa[21], Berger[3])
: A two-player game is nondegenerate if no mixed strategy with k actions has more than k pure best responses. It was proven for 2×2 games by Miyazawa and recently extended by Berger to $2 \times n$ games.

Weighted potentional games (Monderer and Shapley[22, 23])

The potential is a function $\rho : X \to \mathbb{R}$, which is the same for each player, such that there exist positive real numbers $\lambda_1, \lambda_2, \ldots, \lambda_n$ such that for every player i, every set of opponent's actions $x_{-i} \in X_{-i}$ and every pair of actions $x_i, x_i' \in X_i$ the change in utility can be rescaled to equal the change in potential:

$$\lambda_i u_i(x_i, x_{-i}) - \lambda_i u_i(x_i', x_{-i}) = \rho(x_i, x_{-i}) - \rho(x_i', x_{-i}).$$

Games with an interior ESS (Hofbauer[14])

An evolutionary stable strategy is a strategy that, in an evolutionary environment, once adopted cannot be invaded by another strategy that is initially rare[13].

Some supermodular games (Milgrom and Roberts, Krishna, Hahn).

Supermodular games are those characterized by strategic complementarities, roughly meaning that when one player takes a higher action, the others want to do the same[18]. The fictitious player property has been proven to exist in finite-player games with compact strategy sets and continuous utilities[20] and all 3×3 games with strategic complementarities[3][16][11].

$$\begin{array}{c} \\ A \\ B \\ C \end{array} \begin{array}{c} \begin{array}{ccc} a & b & c \end{array} \\ \begin{pmatrix} 1,0 & 0,0 & 0,1 \\ 0,1 & 1,0 & 0,0 \\ 0,0 & 0,1 & 1,0 \end{pmatrix} \end{array}$$

Fig. 2.3: Shapley's game

Although the list of classes of games which have the fictitious play property has grown since the introduction of fictitious play it soon became apparent that this property does not hold for every class of games. In 1964 Shapley[27] described the 3×3 game in fig. 2.3 where the row player wants to play the same action as his opponent (e.g. $\{B, b\}$) whereas the column player wants to play one action before (e.g. $\{B, a\}$). This game most commonly known as Shapley's game has a unique mixed Nash equilibrium where both players have an equal probability to play each action. A fictitious play process does not converge to this equilibrium but results in a cycle with exponentially more periods in each consecutive cycle. Another example of a class of games that do not possess the fictitious play property is the coordination game. A fictitious play process can fall into a cyclic pattern, depending on the initial move and the tie-breaking rule that is applied when there is more than one action in the best-reply correspondence.

2.3 Adaptations

The original algorithm has a very rigid specification of predicting according to the empirical distribution of play and choosing an action to maximize the immediate expected utility. Fortunately there are possible modifications without losing the properties of convergence mentioned above[9]. We will discuss two adaptations: smoothed fictitious play, which adds random trembles, and weighted fictitious play, where the weight of observations decreases over time and more recent observations thus carry more weight.

[3] The decisions of two or more players are called strategic complements if they mutually reinforce each other.

2.3.1 Smoothed fictitious play

In smoothed fictitious play (SFP) the players' responses are smoothed by small trembles or random shocks. It was developed by Fudenberg and Kreps[9] along the lines of Harsanyi's purification theorem[12], which explains decision making when playing a mixed strategy as the result of unobserved utility perturbations that sometimes lead players to have a strict preference for one action, and sometimes a strict preference for another. They define smooth fictitious play as the family of algorithms that employ a standard FP forecasting rule, and a response rule that maximizes the actual utility U_i:

$$U_i(x_i, p_{-i}) = u_i(x_i, p_{-i}) - \gamma_i w_i(x_i). \tag{2.3}$$

Here the smoothing function $w_i : \Delta_i \to \mathbb{R}$ is any smooth, differentiable and strictly concave function such that $|\nabla w_i(q_i)| \to \infty$ whenever x_i approaches the boundary of Δ_i, and the response parameter $\gamma_i > 0$.

To understand the variant of SFP we will be using, suppose the choice probabilities of player i are described by the logistic function[4] $q_i(x_i|p_{-i})$ with denotes the probability that action x_i will be played next:

$$q_i(x_i|p_{-i}) = \frac{e^{u_i(x_i,p_{-i})/\gamma_i}}{\sum\limits_{x_i' \in X_i} e^{u_i(x_i',p_{-i})/\gamma_i}}. \tag{2.4}$$

If the response parameter γ_i is close to zero this function closely approximates the best-reply correspondence, while with a larger γ the probabilities are more evenly spread over all actions[5]. From an observer's standpoint a player using eq. 2.4 may look like he is using a best-reply function which is perturbed by small utility trembles; each period the actual utility $U_i = u_i + \epsilon_i^t$ is perturbated by the extreme-valued variable ϵ_i^t (whose cumulative distribution is $\ln P(\epsilon_i^t \leq z) = -e^{-z/\gamma_i}$). Another argument for using the logistic function involves the Shannon entropy. The amount of information conveyed by a probability distribution q_i can be represented by the entropy function $q_i \ln(q_i)$:

$$-\sum_{x_i' \in X_i} q_i(x_i', p_{-i}) \ln(q_i(x_i', p_{-i})).$$

It can be shown[10] that the optimal q_i is given by the logistic function if we define the actual utility (eq. 2.3) as a weighted combination of the utility u_i and the information gained from experimenting:

$$U_i(x_i, p_{-i}) = u_i(x_i, p_{-i}) - \gamma_i \sum_{x_i' \in X_i} q_i(x_i', p_{-i}) \ln(q_i(x_i', p_{-i})).$$

With Smoothed fictitious play not only the beliefs, as we have seen with classic FP, but also the behavior converges to equilibrium in 2×2 games. In a finite game if all players use SFP with sufficiently small smoothing parameters then

[4]See McKelvey and Palfrey[19] for details about the quantal response equilibrium. The general idea is that players make errors, but since the probability of an action being chosen is related to its utility it is unlikely that costly errors are made.

[5]When $\gamma \to \infty$ the function $q_i(x_i|p_{-i}) \to \frac{1}{|X_i|}$, thus all actions have an equal chance of being chosen next.

with probability 1 the joint empirical distribution converges to the set of coarse correlated ϵ-equilibria[6]. Hofbauer and Sandholm[15] have shown that the empirical distribution of play in SFP converges pointwise to Nash equilibrium in the classes of games that have the fictitious play property.

2.3.2 Weighted fictitious play

The second adaptation of classic FP we will discuss involves techniques where the observations from the past decay over time, making them less influential than more recent observations. If, as classic fictitious play does, the entire history is considered it is harder to obtain change in beliefs and behavior as the history gets larger thus making it harder for fictitious play to catch up when the opponent changes its strategy. A crude but effective way to resolve this issue is fictitious play with finite memory, where only the last m rounds are considered. Weighted fictitious play (WFP), however, employs a more sophisticated approach where a weight factor $0 \leq \gamma \leq 1$ is applied to the history[7]. It uses the original best-reply response rule and a modification of the forecasting rule (eq. 2.1), where every round every prior observation is multiplied with the weight factor (i.e. at time t an observation from $t' \leq t$ is multiplied with $(\gamma)^{t-t'}$):

$$p_i^t(x_j) = \frac{I^t(x_j) + \sum_{u=1}^{t-1} (\gamma)^u I^{t-u}(x_j)}{1 + \sum_{u=1}^{t-1} (\gamma)^u}. \tag{2.5}$$

The weight factor γ determines the rate of decay. When $\gamma = 1$ there is no decay and weighted fictitious play behaves like classic fictitious play. When $\gamma = 0$ the entire history decays instantly and weighted fictitious play behaves like the learning rule introduced by Cournot[8] in the late 19th century which simply assumes that players play a best response to the most recent observation. Weighted fictitious play, also known as exponential fictitious play, has proven to be a good model for human decision making in an experiment conducted by van Huyck[32] where participants played the coordination game.

3 Pattern recognition

In this chapter we examine two distinct implementations of pattern recognition in fictitious play.

Rothelli[26], Lahav[17] and Spiliopoulos[29, 30] have created similar pattern recognizing algorithms to describe human learning behavior. Their models search the entire history of play for possible patterns that match the last few moves played, apply a form of decay and finally use the frequency distribution obtained to predict the next move. This approach is similar to conditional fictitious play proposed by Aoyagi[1], who proved that if two players both apply conditional fictitious play that recognizes patterns of the same length their beliefs converge to the equilibrium in zero-sum games with a unique Nash equilibrium. We will examine Spiliopoulos' model because it is an intuitive extension

[6]A coarse correlated ϵ-equilibrium is a joint distribution $q \in \Delta$ with a small $\epsilon < 0$ such that no player can opt-out and gain more in expectation than ϵ.

of weighted fictitious play and in its simplest form does not apply any additional transformations designed to model how humans form beliefs and make decisions.

Sonsino[28] proposes confused learning with cycle detection at the end of the realized history of play and proves convergence to a fixed pattern of pure Nash equilibria for a large class of games. We will briefly touch confused learning and focus on the cycle detection he proposes.

3.1 N-Period fictitious play

Spiliopoulos[29, 30] proposes n-period fictitious play (FPN), where $n \in \mathbb{N}^+$, as an extension to weighted fictitious play which keeps track of sequences of actions of length n in the observed history of play.

It does not directly search for patterns, but uses the observations of the last $n-1$ rounds as a premise and uses only the part of the history when it forecasts the opponent's next move[7]. It does respond to patterns as they will be visible in the history; with $n = 3$ and a pattern AAB the probability $p(B|AA)$, for example, will be high. Even patterns longer than n can indirectly be derived from the history sometimes. The pattern $AABBA$ will result in higher probabilities $p(B|AB)$, $p(A|BB)$ and $p(A|BA)$, but both $p(A|AA)$ and $p(B|AA)$ will be about equal.

FP1 is exactly the same as WFP, and FP3, for example, keeps track of the observed frequencies of all possible patterns of three, which enables a forecast of the next round given what was played in the last two rounds. First the indicator function I^t is extended to indicate wether a sequence of actions was played. Let $\bar{x}_j$ be a sequence of n actions. The indicator $I^t(\bar{x}_j)$ return 1 if the sequence resembles the actions played in rounds $t-n+1, \ldots, t-1, t$, and 0 otherwise. The forecasting function p_i^t in equation 2.5 is extended to return the proportion of the time the actions $\bar{x}_j$ were played:

$$p_i^t(\bar{x}_j) = \frac{I^t(\bar{x}_j) + \sum_{u=1}^{t-1} (\gamma)^u I^{t-u}(\bar{x}_j)}{1 + \sum_{u=1}^{t-1} (\gamma)^u}$$

which can then be used to calculate the proportion of the time the action x_j was played given that the actions $\bar{x}_j$ were played in the previous $n-1$ rounds:

$$p_i^t(x_j|\bar{x}_j) = \frac{p_i^t(\bar{x}_j + x_j)}{\sum_{x'_j \in X_j} p_i^t(\bar{x}_j + x'_j)}.$$

Spiliopoulos continues by extending the FPN algorithm to allow the players' perceptions to follow commonly ascribed psychophysics principles, which we will not discuss here because human leaning behavior is beyond the scope of our investigation. We will use the n-period fictitious play algorithm in its simplest form in our experiment.

[7] This is similar to Aoyagi's conditional history. An important difference is that conditional FP uses the joint actions of both players as a premise, while we only consider the actions of the opponent.

3.2 Cyclic pattern detection

Sonsino[28] describes a learning model called confused learning where players can choose their learning behavior independently throughout the game and learn to learn; they learn to choose the best learning model and are confused in the sense that they do not a priori know what the best strategy is. An agent may use different prediction rules in different states, may abort using a prediction model that failed to produce good responses, and with a small probability can behave in any reasonable way by following any myopic strategy (Cournot best-reply, fictitious play or any other learning strategy). We will not discuss the inner workings and convergence properties of this model, but rather look at the superimposed ability of strategic pattern recognition of cyclic patterns that repeat successively at the observed path of play Sonsino proposes. Detection is restricted to basic patterns; patterns that do not contain a smaller subpattern[8]. If a player recognizes such a repeated pattern it will assume the other players will continue to follow that pattern and, being strictly myopic, plays a best-reply response against the predicted next joint actions of the pattern.

For every pattern p with length l_p there is a uniform bound T_p such that the model must detect p with probability 1 if it has appeared almost T_p times (if the next round may complete T_p uninterrupted occurrences of p). To eliminate the possible inconsistency that the model detects more than one pattern with probability one at the same time, it is assumed that only patterns with length shorter than some fixed bound L are detected and that $T_p l_p \geq 3L$. This ensures T_p is large enough relative to L. The detection of patterns after a history-independent fixed number of repetitions is to stylized to represent any realistic learning effort. To accommodate any form of non-stationary pattern detection a set of necessary conditions is added to the model. Patterns longer than 2 must appear almost two times (again, if the next round mat complete two uninterrupted occurrences of the pattern) in the past $2l_p - 1$ rounds. Patterns of length 2 must appear almost $2^1/_2$ times in the past 4 rounds and patterns of length 1 must appear 2 times in the past 2 rounds. For the pattern $ABCD$, for example, the last 7 observations must be $ABCDABC$, while a shorter pattern AB requires the last 4 observations to be $BABA$. To enforce continuation a detected pattern p remains detected until it is contradicted, and to prevent the model from immediately detecting a different pattern after a previously detected pattern was contradicted, a pattern p may only be detected l_p rounds after the previous pattern was detected. The existence of these minimal and the necessary conditions define a range in which pattern detection can occur.

The notion of convergence of behavior is more appropriate than the convergence of beliefs or empirical frequencies. Such convergence seems unlikely for this myopic behavior with bounded rationality in an incomplete information environment. Sonsino shows that convergence of behavior to any fixed strategy that is not a pure equilibrium is incompatible with pattern recognition[9], and that convergence of behavior to some mixed strategy profile is only possible if agents consider arbitrarily long histories (and thus impossible for agents with bounded memory). By adding the following assumption he obtains a general convergence result for finite games: for every subset of pure strategies E there

[8] A and AAB are examples of basic patterns, $ABAB$ and AA are not

[9] A detected pattern can only sustain itself if it consists only of pure Nash equilibria, because the player will break other patterns by playing a best-reply action.

is a uniform bound T_E such that if the last T_E rounds of play are in E then the probability that the next move will be a best reply against some move in E is equal to one. With the additional assumption the probability of convergence to either a PN or a minimal closed subgame[10] without PN is equal to one.

4 Experiment Setup

The algorithms used in our experiment (listed in Table 2) are defined by their forecasting and response components so we can quickly inspect not only the performance of the algorithm as a whole but also the performance of the individual components without diving into the internal beliefs. This enables us to quickly identify which behavior to examine in more detail.

Model	**Name**	**Forecast**	**Response**	**Parameters**
Brown's original FP	FP	Simple	Best reply	
Smooth FP	SFP	Simple	Smoothed	$\gamma = 1$
Weighted FP	WFP	Weighted	Best reply	
Smooth weighted FP	SWFP	Weighted	Smoothed	$\gamma = 1$
N-pattern	FPN	N-Pattern	Best reply	$N = 2, 3$
Smoothed n-pattern	SFPN	N-Pattern	Smoothed	$N = 2, 3$, $\gamma = 1$
Weak cycle	FPwCL	Weak cycle	Best reply	$L = 2, 3, 20$
Smoothed weak cycle	SFPwCL	Weak cycle	Smoothed	$L = 2, 3, 20$, $\gamma = 1$
Strong cycle	FPsCL	Strong cycle	Best reply	$L = 2, 3, 20$
Smoothed strong cycle	SFPsCL	Strong cycle	Smoothed	$L = 2, 3, 20$, $\gamma = 1$

Table 2: FP models used in the experiment.

Simple forecasting (section 2.1)
Brown's original forecasting algorithm.

Weighted forecasting (section 2.3.2)
The weighted forecasting algorithm with a weight factor γ of 0.9.

***N*-Pattern forecasting** (section 3.1)
The pattern detecting adaptation of WFP proposed by Spiliopoulos. We will use a pattern length n of both 2 and 3, resulting in four different algorithms: FP2, SFP2, FP3 and SFP3. The weight factor γ is 0.9. Note that SFPN stands for smoothed FPN, and not for subjective FPN proposed by Spiliopoulos. We do not use subjective FPN because our aim is to study the effects of pattern detection, not the effects of subjective beliefs.

Weak cycle forecasting (section 3.2)
Patterns are only detected at the upper bound described by Sonsino, when the next round may complete T_p consecutive occurrences of the pattern. The pattern length L affects how quick a pattern is detected because $T_p = \text{floor}(3L/l_p)$. We will use the same short pattern length L of 2 and 3, and another large L of 20 resulting in six different algorithms. When

[10] A subgame (a subset of the enclosing game) is closed if and only if the best response to each action in the subgame is also in that subgame. A subgame is minimal if it contains no other closed subgame.

a pattern is detected the algorithms returns a probability distribution where the next action in the pattern has probability of one. If no pattern is detected it employs the Simple forcasting algorithm.

Strong cycle forecasting (section 3.2)
Patterns are detected at the lower bound described by Sonsino. The pattern length L influences neither the computational and memory cost nor the speed with which patterns are detected. We will use the same pattern lengths 2, 3 and 20, and this algorithms also employs the Simple forecasting algorithm when no pattern is detected.

Best reply response (section 2.1)
The best reply correspondence. If there is a tie (i.e. the set BR contains more than one action) then one is chosen at random.

Smoothed response (section 2.3.1)
The trembled best reply algorithm with a response parameter γ of 1.

There are 40 algorithms for 2×2 games and 60 for 3×3 games in our experiment[11]. The algorithms are tested with games very common to game theory literature in general and adaptive learning in particular. For each game each algorithms faces all algorithms, themselves included, in a tournament where each game lasts 1000 rounds and is repeated 100 times. Finally they are tested with 100 randomly generated games, one tournament per game. The games used are:

Asymmetric Coordination game
There are two pure strategy equilibria and a mixed strategy equilibrium. Both players prefer the same equilibrium which Pareto dominates the other.

$$\begin{array}{c} \\ A \\ B \end{array} \begin{array}{c} \begin{array}{cc} A & B \end{array} \\ \begin{pmatrix} 5,5 & 0,0 \\ 0,0 & 3,3 \end{pmatrix} \end{array}$$

Symmetric Coordination game
There are two pure strategy equilibria and a mixed strategy equilibrium. Neither equilibrium is preferred or dominated.

$$\begin{array}{c} \\ A \\ B \end{array} \begin{array}{c} \begin{array}{cc} A & B \end{array} \\ \begin{pmatrix} 1,1 & 0,0 \\ 0,0 & 1,1 \end{pmatrix} \end{array}$$

Shapley's game
The only equilibrium is a mixed strategy where each player plays each strategy with equal probability. The game is a non-zero sum variant of rock-paper-scissors designed to cause cyclic behavior in fictitious play algorithms.

$$\begin{array}{c} \\ A \\ B \\ C \end{array} \begin{array}{c} \begin{array}{ccc} A & B & C \end{array} \\ \begin{pmatrix} 1,0 & 0,0 & 0,1 \\ 0,1 & 1,0 & 0,0 \\ 0,0 & 0,1 & 1,0 \end{pmatrix} \end{array}$$

Battle of the sexes
There are two Pareto optimal pure strategy equilibria and a mixed strategy equilibrium. A different pure strategy equilibrium is preferred by each player.

$$\begin{array}{c} \\ A \\ B \end{array} \begin{array}{c} \begin{array}{cc} A & B \end{array} \\ \begin{pmatrix} 3,2 & 0,0 \\ 0,0 & 2,3 \end{pmatrix} \end{array}$$

[11]There are 10 forecasters (FP, WFP, FP2, FP3, FPwC-2, FPwC-3, FPsC-2, FPsC-3 and FPsC-20) times 2 responders (Best reply, Smoothed) times 2 or 3 possible initial moves equals 40 or 60 algorithms.

Chicken
A variant of Battle of the Sexes with the same equilibria, where both players' actions corresponding to their preferred pure equilibria yields the worst outcome for both.

$$\begin{array}{c} \\ A \\ B \end{array} \begin{array}{c} \begin{array}{cc} A & B \end{array} \\ \begin{pmatrix} 0,0 & -1,1 \\ 1,-1 & -10,-10 \end{pmatrix} \end{array}$$

Matching pennies
Zero-sum game where the only equilibrium is a mixed strategy where each player plays each strategy with equal probability.

$$\begin{array}{c} \\ A \\ B \end{array} \begin{array}{c} \begin{array}{cc} A & B \end{array} \\ \begin{pmatrix} -1,1 & 1,-1 \\ 1,-1 & -1,1 \end{pmatrix} \end{array}$$

Prisoner's dilemma
Each player has a dominant strategy resulting in a pure strategy equilibrium. However, the second pure strategy equilibrium is Pareto optimal.

$$\begin{array}{c} \\ A \\ B \end{array} \begin{array}{c} \begin{array}{cc} A & B \end{array} \\ \begin{pmatrix} 2,2 & 0,3 \\ 3,0 & 1,1 \end{pmatrix} \end{array}$$

Random games
We've generated 100 random 2×2 games using GAMUT. All eight payoffs lie between 0 and 100.

5 Results per game

In this section we will first give an in-depth analysis of the behavior of the players per tournament, and then use those findings to analyze the overall performance of the algorithms employing pattern detection to see if, how and why pattern detection is a useful addition to fictitious play.

Since the behavior of similar algorithms often overlaps we will refer to them as one algorithm: FPC denotes both FPwC and FPsC, without specifying the parameter L FPsC and FPwC denotes all used values for L and FPN denotes both FP2 and FP3. We will, unless explicitly mentioned otherwise, refer to both the best-reply and the smoothed response variants when we mention a model (FP denotes both FP and SFP, WFP both WFP and SWFP and so on), because there are no clear positive or negative differences between the two responses in most of the results. A list of all players and their received payoff per tournament, as well as a simplified overview of the results discussed below can be found in appendix A.

5.1 Asymmetric coordination game

All matches converge to an equilibrium joint strategy. If both players have the same initial action they will continue to play that action, thus playing either B, B or the preferred equilibrium strategy A, A. If the initial actions are different all algorithms, with one exception, play the preferred equilibrium strategy onward from round three. The players behave exactly the same because the pattern detection forecasters cannot yet detect patterns in the two rounds before converges occurs and thus behave as either FP or WFP. In round three FP2 always returns the exact same probabilities as WFP even though it is technically performing pattern detection at that stage. If the player's initial moves differ (A, B and B, A) they both play opposite actions in round two (B, A

and A, B, respectively). In round three FP forecasts an equal probability of the opponent playing either action $p(A) = p(B) = 0,5$ and because of the higher payoff of A, A the best-reply correspondence $BR = \{A\}$. WFP, because of the decay factor $\gamma = 0.9$, forecasts 0.5 for the action observed in round two and 0.5×0.9 for the action in round one but since $0.5 \times 0.9 \times 5 > 0.5 \times 3$ still results in $BR = \{A\}$.

The one exception mentioned above is when FP2 with initial action B plays against an opponent with initial action A. After observing the history ABA the forecasted probability $p(B|A) = 1$ and $p(A|A) = 0$ resulting in a $BR = \{B\}$. Round four is the only round that differs from the other matches.

5.2 Symmetric coordination game

The most important difference with the asymmetric coordination game is that when there is an equal probability forecasted for two actions the best-reply correspondence will contain two actions since both equilibria have the same payoffs, and thus the tie-breaking rule chooses one at random. This makes the behavior no longer deterministic, and reduces both the number of games that converge at all and the number of games that converge to the preferred equilibrium. If both players have the same initial action they still will continue to play that strategy. When the initial actions are different the results differ from the previous game:

- WFP in self-play will change its action every round thus will miscoordinate the entire match.
- FP2 in self-play has an equal chance to converge to either A, A or B, B, or to miscoordinate the entire match. In round three the random tie-breaking rule will result in either coordination or miscoordination, behavior which will remain the same for the rest of the match.
- FP3 in self-play behaves similar to FP2. If the players coordinate in round three they will converge to that equilibrium, but some miscoordination occurs in the first 20 to 30 rounds. If the players miscoordinate in round three they too will display some random behavior in the first 20 to 30 rounds, and eventually coordinate in 60% of the observed games.

All other combinations of players will result in convergence to one of the two equilibria in three or four rounds.

5.3 Shapley's Game

The mixed equilibrium strategy where each action has an equal probability to be chosen leads to an average payoff, for both players, of $1/2$ per round when played optimal and $1/3$ when played purely at random. FPN and FPsC score above $1/2$, meaning they have exploited at least some of their opponents. The smoothed response does differ from the best-reply response but has no clear positive or negative effect. FP scores lower than and FPwC-20 equal to the expected payoff $1/3$ of pure random play. There is a clear order in performance of the algorithms:

1. FP2 (avg. 0.65)
2. FP3 (avg. 0.62)

3. FPsC (avg. 0.56)
4. WFP (avg. 0.44)
5. FPwC-2 (avg. 0.43)
6. FPwC-3 (avg. 0.40)
7. FPwC-20 (avg. 0.33)
8. FP (avg. 0.30)

This game is designed to lead fictitious play to cyclic behavior and as a result all matches display cycles of three singleton patterns[12] (because there are three possible actions) $CC \ldots BB \ldots AA \ldots$ in that order but with different repetition lengths. The only exception for FPN in self-play where there is no convergence. The speed with which the players change their own strategy after the opponent has switched determines its performance.

- FP detects changes slower as the game progresses. The actual detection speed depends on the cycle length in that game. In self-play FP plays very long cycles with increasing length, in 1000 rounds FP completes at most two cycles in 63% of the games and at most three cycles in 94% of the games. The exact cycle length in self-play is determined by the random tie-breaking process; a tie-break in favor of the action in the last round increases a cycle length with 1 which results in longer cycles for the rest of the game.
- WFP detects change faster than FP, but like FP depends on the cycle length in that game. In self-play WFP plays cycles of 1 repetition and receives 0 payoff if the initial round yielded 0 payoff for both players and 0.5 payoff otherwise.
- FPN detects change in one round. Because in the history of play the singleton patterns[13] occur the most, the forecaster will predict the same action the opponent played in the previous round. There is no convergence in its behavior in self-play Because both players detect change in one round the forecast often depends on the tie-breaker in the early stages of the game which reduces the length of singleton patterns played, often to lengths shorter than the pattern length N of 2 or 3.
- FPwC detects change in $3L$ rounds and plays cycles of $6L$ repetitions in self-play.
- FPsC detects change in two rounds, that is, the new singleton pattern must occur for two rounds to be detected so it has to fall back to FP for one round. This weakness explains the variation of 1 in cycle length (see appendix A.3) and its behavior in self-play where its behavior comes close to cycles of 4 repetitions, but it does not fully converge to a three singleton pattern cycle.

5.4 Battle of the Sexes

1. FP, FPwC (avg 2.58)
2. FPsC (avg 2.45)
3. FP2 (avg 2.38)
4. WFP (avg 2.31)

[12]A singleton pattern is a cyclic pattern of one action $AAA \ldots$ for any $A \in X_i$

[13]A repeating pattern of one action AA for FP2 and AAA for FP3 for any action $A \in X_i$

5. FP3 (avg 2.27)

If both players have the same initial action they will continue to play that joint action equilibrium for the rest of the game. If the initial actions differ:

- FP converges to each equilibrium 50% of the time in self-play, the tie-breaking determines when en how convergence occurs.
- WFP miscoordinates the entire game in self-play, the game converges to a $BAABA$ cycle. The game converges to the opponents preferred equilibrium against FP and FPwC.
- FP2 converges to the equilibrium corresponding to its own initial action against all opponents except in self-play where the players miscoordinate.
- FP3 converges to the opponent's preferred equilibrium. In self-play and against FP2 the tie-breaker can break the miscoordination causing the game to converge to FP2's preferred equilibrium, or to any of the two equilibria in self-play.
- FPwC behaves like FP, except FPwC-2 against WFP where the strategies converge to each equilibrium 50% of the time.
- FPsC converges to the equilibrium matching the opponents first move, except against FP and FPwC when their first move does not match their preferred equilibrium. In those matches and in self-play the game converges to each equilibrium 50% of the time.

The pattern recognition in FPsC predicts the opponent's strategy fast and accurate which, in this coordination game, unfortunately results in convergence to the opponent's preferred equilibrium. FP and FPwC, on the other hand, have the advantage that their predictions are less strong which makes it possible for the best-reply response to include their preferred equilibrium strategy[14]. In other words, the uncertainty of FP's forecasts allows the best-reply response to let FP display teaching behavior whereas FPsC merely display following behavior. And since almost all matches result in coordination the winning strategy is teaching the opponent to play your own preferred equilibrium.

5.5 Chicken

The smoothed response results in either exactly the same or a lower payoff compared to the best-reply response The effect negative effect is much stronger on FP, FPsC and FPwC-20. If the initial actions differ players will continue to play those actions resulting in an average payoff of 1 per round for the player that plays action B and -1 for the other player. As a result, all players with initial action A have a negative overall score. Once an equilibrium is played the players will continue to play that equilibrium strategy. When both players change strategy simultaneously the game converges to $AA \ldots B$ cycles. This cycle is 10 actions long for FP because the payoff for B, B is ten times worse than that of the preferred equilibrium (either A, B or B, A). When one player changes strategy before the other, thus is quicker to play B after several rounds of A, A, the game immediately converges to that players preferred equilibrium.

[14] When a pattern is recognized the forecast for the next action always has probability 1, whereas FP forecasts are between 0 en 1. Iff the probability $p(X) \leq 1.5 \times p(Y)$, where $X, Y \in X_i$ and Y corresponds to the preferred equilibrium, then $Y \in BR$.

From the table in appendix A.5 it is clear that apart from FPN there are no clear stronger strategies or equal results for different opponents. FP2 and FP3 are winners because they detect the singleton cycles faster than it takes the others to switch actions. FPsC scores high for the same reason, but in self-play the fast detection of singleton cycles leads to short AAB cycles resulting in a very low score in those matches, and as a result a lower overall score.

5.6 Matching Pennies

There is a clear order of performance, where every algorithm wins against all below it and loses against all above it in the ranking:

1. FPN (avg. 0.22, 0.23 for $N = 2, 3$, respectively)
2. FPsC (avg. 0.03)
3. WFP (avg. -0.02)
4. FPwC (avg. -0.10, -0.13 for $L = 2, 3$, respectively)
5. FP, FPwC-20 (avg. -0.12)

All matches result in $A, B - B, B - B, A - A, A$ cycles, the length varies from match to match and never converges fully, except for FPsC in self-play which detects changes in the opponents actions in two rounds thus resulting in cycles with a total length of 8. The variation in cycle length remain small with exception of FP in self-play which displays longer cycles as the game progresses. The equilibrium strategy is played in self-play by all but FP3, although FP3 plays very close to equilibrium in self-play. FP versus both FPwC-3 and FPwC-20, which all behave like FP because they are unable to detect any patterns, and WFP versus FPwC-2 and FPsC also play the equilibrium strategy.

5.7 Prisoner's Dilemma

Action B is the dominant strategy and the best-reply correspondence will therefor only contain that action, resulting in all players playing B after their first turn in all games.

5.8 Random games

This experiment was performed with 100 different random games, so any in-depth analysis requires an overview of all individual games and the properties they possess. Since the intention of this experiment is to expose our algorithms to random situations, and not to analyze the random games themselves, the overall ranking of the tournament is the only information relevant to our experiment:

1. FPN
2. FPsC
3. FPwC ($L = 2, 3$)
4. WFP
5. FPwC ($L = 20$)
6. FP

The results raised the question whether running a tournament on random games is a useful test at all, because most random games are unfair towards one player or strategy. In 84 of the 200 games we used the initial move determined the winner of the match, regardless of the algorithm used.

The tournament ranking (see appendix A.8) clearly shows that all algorithms receive a higher total payoff with initial action A than with initial action B. An analysis of each individual random game shows that all algorithm performs better with initial action A in 41 of the 100 games and with initial action B in 43 of the 100 games. Initial action A is more profitable because the utility received per game in those 41 games is higher than the utility per game players with initial action B receive in the 43 games where B is the better strategy.

6 Pattern Recognition Results

With the results of the per-game analyses of the tournaments we are now able to assess the effectiveness of the pattern recognizing algorithms used. We will compare FPC with FP and FPN with WFP to see the difference with the basic fictitious play variants they are based on.

6.1 Cyclic pattern recognition

Both strong and weak pattern recognition perform better than FP in the random games, where only FPN scores higher. Exception is FPwC-20 which performs only marginally better than FP. This is to be expected since FPwC-20 detects patterns very slowly, in $3L = 60$ rounds, and thus falls back to FP very often. FPC scores higher than FP in Shapley's game and matching pennies, there is no performance improvement in the rest of the games. The behavior is exactly the same as FP in both coordination games and the prisoner's dilemma, almost the same with no clear advantage or disadvantage in chicken, and slightly disadvantageous in the battle of the sexes. FPsC performs better than FPwC in matching pennies and Shapley's game, there is no clear performance difference between strong and weak pattern recognition in all other games.

The pattern length L does not influence the performance of strong pattern recognition. Shorter patterns are allowed to be detected sooner, and with only two actions $\{A, B\}$ a pattern longer than three actions always contains either ABAB or AA making it impossible to detect longer patterns. Detection speed is independent of L and the ability to detect longer patterns is not an advantage because the detection of shorter pattern always blocks the detection of longer pattern.

The difference between weak pattern recognition with a different L responsible for the observed performance differences is simply that FPwC-3 is slightly slower in detecting patterns because of the higher pattern length L. In comparison with FPwC-2 and FPwC-3, FPwC-20 is much slower in detecting patterns and very likely to not detect any patterns at all because patterns need ${}^{60}/_{l_p}$ repetitions to be detected; short patterns are likely to be broken before they can be detected and long patterns are unlikely to occur at all. FPwC-20 performs exactly like FP in matching pennies and Shapley's game, and better than FP but still worse than FPwC-2 and FPwC-3 in the battle of the sexes, matching pennies, Shapley's game and in the random games.

6.2 N-Period fictitious play

In the random generated games, Shapley's game and matching pennies FPN outperforms all other algorithms tested in our experiment. It performs better than WFP in the symmetric coordination game. In the battle of the sexes WFP scores higher than FP3 but lower than FP2, and in the asymmetric coordination game WFP scores equal to FP3 and lower than FP2.

FPN quickly detects change if the opponent switches from one singleton pattern to another. Because AA and BB will occur more often than AB and BA in the observed history of play FPN will forecast the same action the opponent played in the previous round. This forecast is only wrong for one round when the opponent will switch actions.

The differences between $N = 2$ and $N = 3$ are marginal. FP3 performs slightly better in the random games and asymmetric coordination game, where FP2's forecast is wrong in one round. FP2 performs slightly better in matching pennies. There is no difference in performance between FP2 and FP3 in the remaining games.

7 Conclusion

In this thesis we have studied two distictly different approaches to perform pattern recognition on the observerd history of play, obtained empirical data on the performance of those algorithms and the fictitious play variants they extend, and found the new pattern-recognition based FP variants to be significantly more effective than the traditional FP variants.
N-period fictitious play is an extension of weighted fictitious play inspired by pattern recognition, which has proved itself to be significantly more effective than WFP. FP3 performs only slightly better than FP2 in two, and slightly worse in one of the games tested in our experiment. We cannot at this stage, however, recommend not to use a pattern length $N > 2$. The costs are higher, but the effectiveness of a higher pattern length against other opponents, which are not related to FP, remains to be tested.

In contrast to FPN, the strong and weak pattern recognition capabilities of PFC have no relation at all to the fictitious play forecaster algorithm, which in our experiments only serves as a fallback forecaster in case no patterns are detected. It can easily be replaced by any other adaptive algorithm, as Sonsino describes. Furthermore, the strong and weak detection algorithms are merely implementations of the lower and upper bounds of the area where, in Sonsono's model, pattern recognition *may* occur. This does not mean that they are bad algorithms per se, nor does it mean that good or bad results necessarily imply cyclic pattern recognition at the end of the observed history of play is a good or bad strategy. The results of these two approaches do, however, give valuable insights that should be taken into account when designing a cyclic pattern recognizing algorithm.

The results we obtained are encouraging and illustrate the possibilities and merits of applying pattern recognition in machine learning and game theory. Adding a layer of pattern recognition on top of FP improves performance significantly in comparison to FP on its own and FPN nearly always outperforms WFP.

Fictitious play is a basic algorithm which has its limitations and weaknesses yet is very effective for its simplicity. The pattern recognition algorithms we studied share some of those strengths and weaknesses because they are either and extension of FP or use FP as a fallback strategy, and the matches show similar behavior because there were only FP-based opponents participating in the tournament. But FPN and FPC do show different behavior, overcoming FP's vulnerably to cyclic behavior and introducing new problems coordinating in self-play, for example.

The field of pattern recognition, however, is vast and offers many idea's and possibilities for (advanced) pattern recognition that have no relation at all to fictitious play. Confining further research to fictitious play is an unnecessary limitation on the broad spectrum of possibilities that pattern recognition has to offer. Despite the effectiveness of the pattern-detection based fictitious play algorithms studied, pattern detection should be seen as a stand-alone approach to machine learning. The best reply response works well in combination with the tested pattern detecting forecasts, but we should not ignore possibilities where pattern recognition provides both components or a complete learning algorithm where such a distinction is not applicable at all.

References

[1] M. Aoyagi. Evolution of beliefs and the nash equilibrium of normal form games. *Journal of Economic Theory*, 70(2):444–469, 1996.

[2] D. Bellhouse. The problem of waldegrave. *Electronic Journal for the History of Probability and Statistics*, 3:1–12, 2007.

[3] U. Berger. Fictitious play in $2 \times n$ games. *Journal of Economic Theory*, 120(2):139–154, 2005.

[4] U. Berger. Brown's original fictitious play. *Journal of Economic Theory*, 135(1):572–578, 2007.

[5] G.W. Brown. Some notes on computation of games solutions. Technical report, DTIC Document, 1949.

[6] G.W. Brown. Iterative solution of games by fictitious play. *Activity analysis of production and allocation*, 13(1):374–376, 1951.

[7] Y.W. Cheung and D. Friedman. Individual learning in normal form games: Some laboratory results. *Games and Economic Behavior*, 19(1):46–76, 1997.

[8] A. Cournot. Research into the mathematical principles of the theory of wealth (1838), trans. *BACON. New York*, 2, 1897.

[9] D. Fudenberg and D. Kreps. *Learning mixed equilibria*. MIT Press, 1992.

[10] D. Fudenberg and D.K. Levine. *The theory of learning in games*, volume 2. The MIT press, 1998.

[11] S. Hahn. The convergence of fictitious play in 3×3 games with strategic complementarities. *Economics Letters*, 64(1):57–60, 1999.

[12] J.C. Harsanyi. Games with randomly disturbed payoffs: A new rationale for mixed-strategy equilibrium points. *International Journal of Game Theory*, 2(1):1–23, 1973.

[13] WGS Hines. Evolutionary stable strategies: a review of basic theory. *Theoretical Population Biology*, 31(2):195–272, 1987.

[14] J. Hofbauer. Stability for the best response dynamics. Unpublished, 1995.

[15] J. Hofbauer and W.H. Sandholm. On the global convergence of stochastic fictitious play. *Econometrica*, 70(6):2265–2294, 2002.

[16] V. Krishna. Learning in games with strategic complementarities. Unpublished, 1992.

[17] Y. Lahav. Behavioral pattern learning models for decision making in games. *Journal of Pattern Recognition Research*, 4(1):133–151, 2009.

[18] J. Levin. Supermodular games. *Lectures Notes, Department of Economics, Stanford University*, 2003.

[19] R.D. McKelvey and T.R. Palfrey. Quantal response equilibria for normal form games. *Games and economic behavior*, 10(1):6–38, 1995.

[20] P. Milgrom and J. Roberts. Adaptive and sophisticated learning in normal form games. *Games and economic Behavior*, 3(1):82–100, 1991.

[21] K. Miyazawa. On the convergence of the learning process in a 2×2 non-zero-sum two-person game. econometric research program, princeton university. *Research Memorandum*, 25(33), 1961.

[22] D. Monderer and L.S. Shapley. Fictitious play property for games with identical interests. *Journal of Economic Theory*, 68(1):258–265, 1996.

[23] D. Monderer and L.S. Shapley. Potential games. *Games and economic behavior*, 14:124–143, 1996.

[24] Alex M. Mood and Lyman Porter. In memoriam george w. brown. `http://www.universityofcalifornia.edu/senate/inmemoriam/georgewbrown.htm`.

[25] J. Robinson. An iterative method of solving a game. *The Annals of Mathematics*, 54(2):296–301, 1951.

[26] T.F. Rötheli. Pattern recognition and procedurally rational expectations. *Journal of economic behavior & organization*, 37(1):71–90, 1998.

[27] L.S. Shapley et al. Some topics in two-person games. *Advances in game theory*, 52:1–29, 1964.

[28] D. Sonsino. Learning to learn, pattern recognition, and nash equilibrium. *Games and Economic Behavior*, 18(2):286–331, 1997.

[29] L. Spiliopoulos. Pattern recognition and subjective belief learning in repeated mixed strategy games. 2009.

[30] L. Spiliopoulos. Pattern recognition and subjective belief learning in a repeated constant-sum game. *Games and Economic Behavior*, 2012.

[31] J. v. Neumann. Zur theorie der gesellschaftsspiele. *Mathematische Annalen*, 100(1):295–320, 1928.

[32] J.B. Van Huyck, R.C. Battalio, and F.W. Rankin. On the origin of convention: Evidence from coordination games. *The Economic Journal*, 107(442):576–596, 1997.

[33] J. Von Neumann and O. Morgenstern. *Theory of Games and Economic Behavior (Commemorative Edition)*. Princeton university press, 2007.

[34] H. Peyton Young. Learning. In *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions*, chapter 2, pages 25–43. Princeton University Press, 1998.

A Tournament Results

This appendix contains the results of each tournament in two different tables. The first table is a simplified overview of the behavior of the different combination of algorithms grouped by forecaster. Because of the limited space available the descriptions of the behavior are very brief. Section 5 describes the behavior in detail. The second table is a ranking of all algorithms ordered by average received payoff per round of all matches played in the tournament. For each algorithm the forecaster and responder (see section 4), the first move and the pattern length parameters N for FPN and L for FPC are displayed separately, to makes it easier to see the performance of the individual components.

A.1 Asymmetric coordination game

	FP	WFP	FPN	FPwC	FPsC
FP	$t \geq 3 : A, A$	$t \geq 3 : A, A$	$t \geq 3 : A, A$ $t = 4$: FPN-2 Miscoordinate	$t \geq 3 : A, A$	$t \geq 3 : A, A$
WFP	$t \geq 3 : A, A$	$t \geq 3 : A, A$	$t \geq 3 : A, A$ $t = 4$: FPN-2 Miscoordinate	$t \geq 3 : A, A$	$t \geq 3 : A, A$
FPN	$t \geq 3 : A, A$	$t \geq 3 : A, A$	$t \geq 3 : A, A$ $t = 4$: FPN-2 Miscoordinate	$t \geq 3 : A, A$	$t \geq 3 : A, A$
FPwC	$t \geq 3 : A, A$	$t \geq 3 : A, A$	$t \geq 3 : A, A$ $t = 4$: FPN-2 Miscoordinate	$t \geq 3 : A, A$	$t \geq 3 : A, A$
FPsC	$t \geq 3 : A, A$	$t \geq 3 : A, A$	$t \geq 3 : A, A$ $t = 4$: FPN-2 Miscoordinate	$t \geq 3 : A, A$	$t \geq 3 : A, A$

Behavior when initial actions differ.

#	Payoff/round	Name	Param	Forecaster	Responder	1st move
1	4.99465625	SFPN	2	N-Pattern	Smooth	A
2	4.9946375	FPwC	2	Weak cycle	BR	A
3	4.99463125	SFPsC	2	Strong cycle	Smooth	A
	4.99463125	FP		Simple	BR	A
4	4.994625	WFP		Weighted	BR	A
	4.994625	FPwC	20	Weak cycle	BR	A
	4.994625	SFPwC	2	Weak cycle	Smooth	A
5	4.99461875	SFPsC	3	Strong cycle	Smooth	A
	4.99461875	FPsC	2	Strong cycle	BR	A
	4.99461875	FPN	2	N-Pattern	BR	A
	4.99461875	SWFP		Weighted	Smooth	A
	4.99461875	SFPwC	3	Weak cycle	Smooth	A
6	4.9946125	SFPwC	20	Weak cycle	Smooth	A
	4.9946125	SFP		Simple	Smooth	A
7	4.99460625	FPsC	3	Strong cycle	BR	A
8	4.99459375	FPsC	20	Strong cycle	BR	A
	4.99459375	SFPsC	20	Strong cycle	Smooth	A
9	4.9945875	FPwC	3	Weak cycle	BR	A
10	4.994055	FPN	3	N-Pattern	BR	A
11	4.99389875	SFPN	3	N-Pattern	Smooth	A
12	3.995	SWFP		Weighted	Smooth	B
	3.995	FP		Simple	BR	B
	3.995	FPsC	20	Strong cycle	BR	B
	3.995	FPwC	2	Weak cycle	BR	B
	3.995	SFPwC	3	Weak cycle	Smooth	B
	3.995	FPwC	3	Weak cycle	BR	B
	3.995	FPsC	3	Strong cycle	BR	B
	3.995	SFPsC	20	Strong cycle	Smooth	B
	3.995	SFP		Simple	Smooth	B
	3.995	FPsC	2	Strong cycle	BR	B
	3.995	SFPwC	20	Weak cycle	Smooth	B
	3.995	WFP		Weighted	BR	B
	3.995	SFPsC	3	Strong cycle	Smooth	B
	3.995	SFPwC	2	Weak cycle	Smooth	B
	3.995	FPwC	20	Weak cycle	BR	B
	3.995	SFPsC	2	Strong cycle	Smooth	B
13	3.99363125	FPN	3	N-Pattern	BR	B
14	3.99356875	SFPN	3	N-Pattern	Smooth	B
15	3.99198	FPN	2	N-Pattern	BR	B
16	3.991905	SFPN	2	N-Pattern	Smooth	B

A.2 Symmetric coordination game

	FP	WFP	FPN	FPwC	FPsC
FP	Equilibrium at $t = 3$ or 4	Equilibrium at $t = 3$ or 4	Equilibrium at $t = 3$ or 4	Equilibrium at $t = 3$ or 4	Equilibrium at $t = 3$ or 4
WFP	Equilibrium at $t = 3$ or 4	Miscoordinate	Equilibrium at $t = 3$ or 4	Equilibrium at $t = 3$ or 4	Equilibrium at $t = 3$ or 4
FPN	Equilibrium at $t = 3$ or 4	Equilibrium at $t = 3$ or 4	Equilibrium or miscoordinate	Equilibrium at $t = 3$ or 4	Equilibrium at $t = 3$ or 4
FPwC	Equilibrium at $t = 3$ or 4	Equilibrium at $t = 3$ or 4	Equilibrium at $t = 3$ or 4	Equilibrium at $t = 3$ or 4	Equilibrium at $t = 3$ or 4
FPsC	Equilibrium at $t = 3$ or 4	Equilibrium at $t = 3$ or 4	Equilibrium at $t = 3$ or 4	Equilibrium at $t = 3$ or 4	Equilibrium at $t = 3$ or 4

Behavior when initial actions differ.

#	Payoff/round	Name	Param	Forecaster	Responder	1st move
1	0.998135	SFPsC	2	Strong cycle	Smooth	B
2	0.99806625	FPwC	2	Weak cycle	BR	A
3	0.99805125	FPwC	3	Weak cycle	BR	A
4	0.9980425	SFPwC	3	Weak cycle	Smooth	B
5	0.997985	FPwC	20	Weak cycle	BR	A
6	0.99798375	SFPwC	20	Weak cycle	Smooth	A
7	0.99797375	FP		Simple	BR	B
8	0.99797	SFP		Simple	Smooth	A
9	0.99796625	FPwC	2	Weak cycle	BR	B
10	0.99793875	FPsC	20	Strong cycle	BR	A
11	0.99793125	SFPsC	3	Strong cycle	Smooth	B
12	0.9979275	FPsC	2	Strong cycle	BR	B
13	0.997925	FPsC	20	Strong cycle	BR	B
14	0.99792125	FPwC	3	Weak cycle	BR	B
15	0.99791625	FPwC	20	Weak cycle	BR	B
16	0.997915	SFPwC	20	Weak cycle	Smooth	B
17	0.99791	FPsC	3	Strong cycle	BR	A
	0.99791	SFPwC	3	Weak cycle	Smooth	A
18	0.99790625	SFPsC	20	Strong cycle	Smooth	B
19	0.9978975	SFPsC	2	Strong cycle	Smooth	A
20	0.99789625	SFPsC	3	Strong cycle	Smooth	A
21	0.9978925	FPsC	3	Strong cycle	BR	B
22	0.99789	FP		Simple	BR	A
	0.99789	FPsC	2	Strong cycle	BR	A
23	0.99786375	SFPsC	20	Strong cycle	Smooth	A
24	0.99782625	SFPwC	2	Weak cycle	Smooth	B
	0.99782625	SFPwC	2	Weak cycle	Smooth	A
25	0.99782125	SFP		Simple	Smooth	B
26	0.9885375	FPN	3	N-Pattern	BR	A
27	0.98735125	FPN	3	N-Pattern	BR	B
28	0.9861425	SFPN	3	N-Pattern	Smooth	B
29	0.98479375	SFPN	3	N-Pattern	Smooth	A
30	0.97404625	SFPN	2	N-Pattern	Smooth	A
31	0.97272375	FPN	2	N-Pattern	BR	B
32	0.9715475	SFPN	2	N-Pattern	Smooth	B
33	0.97036	FPN	2	N-Pattern	BR	A
34	0.94820375	SWFP		Weighted	Smooth	B
35	0.94816125	WFP		Weighted	BR	A
36	0.94814625	SWFP		Weighted	Smooth	A
37	0.948055	WFP		Weighted	BR	B

A.3 Shapley's game

	FP	WFP	FPN	FPwC	FPsC
FP	$0.50, 0.50$ growing cycles	$0.31, 0.69$ growing cycles	$0.22, 0.77$ 4-cycles	$0.33, 0.67$ 16..18-cycles	$0.27, 0.73$ 7,8-cycles
WFP	$0.69, 0.31$ growing cycles	$0, 0$ or $0.50, 0.50$ 1-cycles or 5-cycles	$0.25, 0.74$ 4-cycles	$0.59, 0.41$ 10-cycles	$0.36, 0.64$ 5,6-cycles
FPN	$0.77, 0.22$ 4-cycles	$0.74, 0.25$ 4-cycles	No convergance	$0.77, 0.22$ 4-cycles	$0.59, 0.32$ 3-cycles
FPwC	$0.67, 0.33$ 16..18-cycles	$0.41, 0.59$ 10-cycles	$0.22, 0.77$ 4-cycles	$0.49, 0.49$ 6L-cycles	$0.28, 0.72$ 7,8
FPsC	$0.73, 0.27$ 7,8-cycles	$0.64, 0.36$ 5,6-cycles	$0.32, 0.59$ 3-cycles	$0.72, 0.28$ 7,8-cycles	$0.45, 0.45$ 4-cycles No convergance

Average payoff per round and number of action repetitions in a cycle. Here a 3-cycle denotes $CCCBBBAAACCCBBBAAA\dots$ cycles.

#	Payoff/round	Name	Param	Forecaster	Responder	1st move
1	1.28415015	FPN	2	Spilio	BR	C
2	1.28385551	FPN	2	Spilio	BR	A
3	1.28384567	FPN	2	Spilio	BR	B
4	1.28353152	FPN	2	Spilio	Smooth	B
5	1.28323452	FPN	2	Spilio	Smooth	C
6	1.28294652	FPN	2	Spilio	Smooth	A
7	1.22038351	FPN	3	Spilio	BR	B
8	1.22018747	FPN	3	Spilio	BR	C
9	1.21990467	FPN	3	Spilio	Smooth	B
10	1.21989481	FPN	3	Spilio	Smooth	A
11	1.21984902	FPN	3	Spilio	BR	A
12	1.21964352	FPN	3	Spilio	Smooth	C
13	1.09863281	FPsC	2	Strong	BR	C
14	1.098529	FPsC	2	Strong	Smooth	B
15	1.09841395	FPsC	2	Strong	Smooth	C
16	1.098391	FPsC	3	Strong	BR	C
17	1.09835735	FPsC	3	Strong	Smooth	C
18	1.09832181	FPsC	2	Strong	BR	B
19	1.09831534	FPsC	2	Strong	BR	B

#	Payoff/round	Name	Param	Forecaster	Responder	1st move
20	1.09824796	FPsC	2	Strong	BR	A
21	1.0982338	FPsC	3	Strong	Smooth	B
22	1.09811895	FPsC	2	Strong	Smooth	C
23	1.09809484	FPsC	2	Strong	Smooth	A
24	1.09808231	FPsC	2	Strong	Smooth	B
25	1.09801604	FPsC	2	Strong	BR	A
26	1.09799778	FPsC	3	Strong	BR	B
27	1.09791261	FPsC	3	Strong	Smooth	A
28	1.09789401	FPsC	2	Strong	Smooth	A
29	1.09785101	FPsC	2	Strong	BR	C
30	1.0978163	FPsC	3	Strong	BR	A
31	0.86807015	SWFP		Weighted	Smooth	A
32	0.86769352	WFP		Weighted	BR	C
33	0.86762982	WFP		Weighted	BR	A
34	0.867438	WFP		Weighted	BR	B
35	0.86713719	SWFP		Weighted	Smooth	B
36	0.8669537	SWFP		Weighted	Smooth	C
37	0.85240366	FPwC	2	Weak	Smooth	C
38	0.85239001	FPwC	2	Weak	BR	A
39	0.85238735	FPwC	2	Weak	Smooth	A
40	0.85237397	FPwC	2	Weak	Smooth	B
41	0.85236928	FPwC	2	Weak	BR	B
42	0.8522255	FPwC	2	Weak	BR	C
43	0.78305387	FPwC	3	Weak	Smooth	A
44	0.78300501	FPwC	3	Weak	BR	A
45	0.78287116	FPwC	3	Weak	Smooth	C
46	0.78282251	FPwC	3	Weak	BR	C
47	0.78276984	FPwC	3	Weak	BR	B
48	0.782709	FPwC	3	Weak	Smooth	B
49	0.65509435	FPwC	2	Weak	BR	A
50	0.65505096	FPwC	2	Weak	BR	B
51	0.6550005	FPwC	2	Weak	BR	C
52	0.65500017	FPwC	2	Weak	Smooth	C
53	0.65498452	FPwC	2	Weak	Smooth	A
54	0.6547965	FPwC	2	Weak	Smooth	B
55	0.59945667	SFP		Simple	Smooth	B
56	0.59941118	FP		Simple	BR	C
57	0.59913686	SFP		Simple	Smooth	C
58	0.59891855	FP		Simple	BR	A
59	0.59876281	FP		Simple	BR	B
60	0.59864999	SFP		Simple	Smooth	A

A.4 Battle of the Sexes

	FP	WFP	FP2	FP3	FPwC	FPsC
FP	50/50	FP pref	FPN 1st	FP pref	50/50	FP 1st 50/50
WFP	FP pref	Mis- coordinate	FPN 1st	WFP pref	FPwC pref L=2: 50/50	WFP 1st
FPN	FPN 1st	FPN 1st	Mis- coordinate FP2 pref	FP2 pref 50/50	FPN 1st	FPN 1st
FPwC	50/50	FPwC pref L=2: 50/50	FPN 1st	FPwC pref	50/50	FPwC 1st 50/50
FPsC	FP 1st 50/50	WFP 1st	FPN 1st	FPsC pref	FPwC 1st 50/50	50/50

Behavior when initial actions differ. Pref denotes the preferred equilibrium of the player, 50/50 denotes each equilibrium 50% of the time and 1st denotes the equilibrium corresponding to that player's first action.

#	Payoff/round	Name	Param	Forecaster	Responder	1st move
1	2.58782620253164	FP		Simple	BR	A
2	2.5851617721519	SFPwC	20	Weak	Smooth	A
3	2.58270113924051	FPwC	20	Weak	BR	A
4	2.58207683544304	SFP		Simple	Smooth	A
5	2.57770126582278	FPwC	3	Weak	BR	A
6	2.57670810126582	SFPwC	3	Weak	Smooth	A
7	2.57451835443038	SFP		Simple	Smooth	B
8	2.57213151898734	FP		Simple	BR	B
9	2.57159569620253	SFPwC	20	Weak	Smooth	B
10	2.57092607594937	FPwC	20	Weak	BR	B
11	2.56172556962025	SFPwC	3	Weak	Smooth	B
12	2.56069898734177	FPwC	3	Weak	BR	B
13	2.55778607594937	FPwC	2	Weak	BR	A
14	2.55503493670886	SFPwC	2	Weak	Smooth	A
15	2.54383227848101	SFPwC	2	Weak	Smooth	B
16	2.54223772151899	FPwC	2	Weak	BR	B
17	2.45669962025317	FPsC	2	Strong	BR	A
18	2.45633379746836	FPsC	20	Strong	BR	A
19	2.45561835443038	SFPsC	3	Strong	Smooth	A
20	2.4543682278481	SFPsC	20	Strong	Smooth	A
21	2.45392670886076	SFPsC	2	Strong	Smooth	A
22	2.45266075949367	FPsC	3	Strong	BR	A
23	2.45083949367089	SFPsC	2	Strong	Smooth	B
24	2.44608037974684	FPsC	20	Strong	BR	B
25	2.4441570886076	SFPsC	20	Strong	Smooth	B
26	2.44285658227848	FPsC	2	Strong	BR	B
27	2.44238683544304	SFPsC	3	Strong	Smooth	B
28	2.44075810126582	FPsC	3	Strong	BR	B
29	2.38624202531646	FPN	2	Spilio	BR	A
30	2.38565367088608	SFPN	2	Spilio	Smooth	A
31	2.37449139240506	SFPN	2	Spilio	Smooth	B
32	2.37308253164557	FPN	2	Spilio	BR	B
33	2.3099382278481	WFP		Weighted	BR	A
34	2.30847126582278	SWFP		Weighted	Smooth	A
35	2.29647708860759	SWFP		Weighted	Smooth	B
36	2.29607050632911	WFP		Weighted	BR	B
37	2.27402417721519	FPN	3	Spilio	BR	A
38	2.27257443037975	SFPN	3	Spilio	Smooth	A
39	2.26103696202532	SFPN	3	Spilio	Smooth	B
40	2.25922278481013	FPN	3	Spilio	BR	B

A.5 Chicken

	FP	WFP	FP2 *A*	FP2 *B*	FP3 *A*	FP3 *B*	FPwC 2	FPwC 3	FPwC 20	FPsC
FP	10	9	10	4	5,6	5	8	10	62	4, 5
WFP	9	7	17	4	5,7	5	8	9	9	4, 5
FPN-2 A	10	17					15	18	10	9
B	4	4			8			4		12
FPN-3 A	5,6	5,7						5,6		18, 23
B	5	5						5		4, 11, 18
2	8	8	15				6	7,8	6,8	4, 5
FPwC 3	10	9	18	4	5,6	5		8	10	
20	62	9	10						10	
FPsC	4, 5	4,5	9	12	18, 23	4, 11, 18		4,5		3

A blue number and red number means convergence to A, B or B, A, respectively, in that round number. Two numbers $5, 6$ means in either of those rounds. A green number means convergence to a (number-1) $\times A, A + B, B$ cycle.

#	Payoff/round	Name	Param	Forecaster	Responder	1st move
1	0.59038475	FPN	3	N-Pattern	BR	B
2	0.58769075	SFPN	3	N-Pattern	Smooth	B
3	0.490706	SFPN	2	N-Pattern	Smooth	B
4	0.485213	FPN	2	N-Pattern	BR	B
5	0.3616	SFPwC	2	Weak cycle	Smooth	B
	0.3616	FPwC	2	Weak cycle	BR	B
6	0.3421025	FPsC	20	Strong cycle	BR	B
7	0.3420765	SFPsC	20	Strong cycle	Smooth	B
8	0.34205325	SFPsC	2	Strong cycle	Smooth	B
9	0.3420335	FPsC	3	Strong cycle	BR	B
10	0.34202125	SFPsC	3	Strong cycle	Smooth	B
11	0.34192475	FPsC	2	Strong cycle	BR	B
12	0.2769	SWFP		Weighted	Smooth	B
	0.2769	WFP		Weighted	BR	B
13	0.1836	SFPwC	3	Weak cycle	Smooth	B
	0.1836	FPwC	3	Weak cycle	BR	B
14	0.142655	FP		Simple	BR	B
15	0.09507475	FPwC	20	Weak cycle	BR	B
16	0.05327025	SFP		Simple	Smooth	B
17	0.04151	SFPwC	20	Weak cycle	Smooth	B
18	-0.384351	SFPN	3	N-Pattern	Smooth	A
19	-0.39252625	FPN	3	N-Pattern	BR	A
20	-0.539725	SFPwC	2	Weak cycle	Smooth	A
21	-0.53972725	FPwC	2	Weak cycle	BR	A
22	-0.624273749999999	SWFP		Weighted	Smooth	A
23	-0.62427625	WFP		Weighted	BR	A
24	-0.65706	FPsC	2	Strong cycle	BR	A
25	-0.657073	FPsC	20	Strong cycle	BR	A
26	-0.6570965	SFPsC	3	Strong cycle	Smooth	A
27	-0.65710325	SFPsC	20	Strong cycle	Smooth	A
28	-0.65710375	SFPsC	2	Strong cycle	Smooth	A
29	-0.65712525	FPsC	3	Strong cycle	BR	A
30	-0.737225500000001	FPN	2	N-Pattern	BR	A
31	-0.737226750000001	SFPN	2	N-Pattern	Smooth	A
32	-0.816023500000001	SFPwC	3	Weak cycle	Smooth	A
33	-0.816024250000001	FPwC	3	Weak cycle	BR	A
34	-0.85673425	FPwC	20	Weak cycle	BR	A
35	-0.90402525	FP		Simple	BR	A
36	-0.9535545	SFPwC	20	Weak cycle	Smooth	A
37	-0.954971	SFP		Simple	Smooth	A

A.6 Matching pennies

	FP	WFP	FPN	FPwC	FPsC
FP	79.5 Equilibrium	17.4 C	6.3 C	30.1 C Equilibrium ($L = 3, 20$)	9.9 C
WFP	17.4 R	7.9 Equilibrium	6.4 C	17.1 R Equilibrium ($L = 2$)	7.9 Equilibrium
FPN	6.3 R	6.4 R	4.5 Equilibrium (FPN-3 near)	5.4 R	4.7 R
FPwC	30.1 R Equilibrium ($L = 3, 20$)	17.1 C Equilibrium ($L = 2$)	5.4 C	24.5 Equilibrium	9.9 C
FPsC	9.9 R	7.9 C Equilibrium	4.7 C	9.9 R	7.9 Equilibrium

Average cycle length, the winner (R = row player, C = column player) if there is one, and whether there is convergence to equilibrium.

#	Payoff/round	Name	Param	Forecaster	Responder	1st move
1	0.2269815	FPN	3	N-Pattern	BR	B
2	0.2264505	SFPN	3	N-Pattern	Smooth	A
3	0.2253945	FPN	3	N-Pattern	BR	A
4	0.2243325	SFPN	3	N-Pattern	Smooth	B
5	0.2235215	FPN	2	N-Pattern	BR	A
6	0.222963	SFPN	2	N-Pattern	Smooth	A
7	0.220386	FPN	2	N-Pattern	BR	B
8	0.220124	SFPN	2	N-Pattern	Smooth	B
9	0.034935	FPsC	2	Strong cycle	BR	B
10	0.0348335	FPsC	20	Strong cycle	BR	A
11	0.0348115	SFPsC	20	Strong cycle	Smooth	B
12	0.0347775	FPsC	2	Strong cycle	BR	A
13	0.0347285	SFPsC	2	Strong cycle	Smooth	B
14	0.034722	FPsC	3	Strong cycle	BR	B
15	0.0346995	SFPsC	2	Strong cycle	Smooth	A
16	0.034698	FPsC	3	Strong cycle	BR	A
17	0.0346715	FPsC	20	Strong cycle	BR	B
18	0.0345835	SFPsC	20	Strong cycle	Smooth	A
19	0.034467	SFPsC	3	Strong cycle	Smooth	B
20	0.034359	SFPsC	3	Strong cycle	Smooth	A
21	-0.021424	SWFP		Weighted	Smooth	A
22	-0.021561	SWFP		Weighted	Smooth	B
23	-0.021564	WFP		Weighted	BR	A
24	-0.0217785	WFP		Weighted	BR	B
25	-0.102832	SFPwC	2	Weak cycle	Smooth	B
26	-0.102865	SFPwC	2	Weak cycle	Smooth	A
27	-0.102896	FPwC	2	Weak cycle	BR	B
28	-0.102921	FPwC	2	Weak cycle	BR	A
29	-0.1281645	SFPwC	3	Weak cycle	Smooth	A
30	-0.1281875	FPwC	3	Weak cycle	BR	B
31	-0.12827	FPwC	3	Weak cycle	BR	A
32	-0.128351	SFPwC	3	Weak cycle	Smooth	B
33	-0.149412	FPwC	20	Weak cycle	BR	B
34	-0.149533	FPwC	20	Weak cycle	BR	A
35	-0.149666	SFPwC	20	Weak cycle	Smooth	A
36	-0.149685	FP		Simple	BR	A
37	-0.149804	SFP		Simple	Smooth	A
38	-0.1498105	SFP		Simple	Smooth	B
39	-0.14988	FP		Simple	BR	B
40	-0.14992	SFPwC	20	Weak cycle	Smooth	B

A.7 Prisoner's dilemma

#	Payoff/round	Name	Param	Forecaster	Responder	1st move
1	1.002	FPN	3	N-Pattern	BR	B
	1.002	FPN	2	N-Pattern	BR	B
	1.002	SFPsC	3	Strong cycle	Smooth	B
	1.002	SFPwC	20	Weak cycle	Smooth	B
	1.002	FPsC	20	Strong cycle	BR	B
	1.002	SFP		Simple	Smooth	B
	1.002	FPwC	20	Weak cycle	BR	B
	1.002	SFPsC	20	Strong cycle	Smooth	B
	1.002	WFP		Weighted	BR	B
	1.002	SFPwC	2	Weak cycle	Smooth	B
	1.002	SFPwC	3	Weak cycle	Smooth	B
	1.002	SFPsC	2	Strong cycle	Smooth	B
	1.002	SFPN	2	N-Pattern	Smooth	B
	1.002	SFPN	3	N-Pattern	Smooth	B
	1.002	FPsC	2	Strong cycle	BR	B
	1.002	FPsC	3	Strong cycle	BR	B
	1.002	FP		Simple	BR	B
	1.002	FPwC	3	Weak cycle	BR	B
	1.002	FPwC	2	Weak cycle	BR	B
	1.002	SWFP		Weighted	Smooth	B
2	1.0005	SFPsC	3	Strong cycle	Smooth	A
	1.0005	SFPwC	20	Weak cycle	Smooth	A
	1.0005	SFPwC	3	Weak cycle	Smooth	A
	1.0005	SFPN	2	N-Pattern	Smooth	A
	1.0005	FPsC	20	Strong cycle	BR	A
	1.0005	FPsC	3	Strong cycle	BR	A
	1.0005	FPN	2	N-Pattern	BR	A
	1.0005	SFPsC	20	Strong cycle	Smooth	A
	1.0005	FP		Simple	BR	A
	1.0005	SWFP		Weighted	Smooth	A
	1.0005	FPwC	3	Weak cycle	BR	A
	1.0005	FPwC	2	Weak cycle	BR	A
	1.0005	SFP		Simple	Smooth	A
	1.0005	WFP		Weighted	BR	A
	1.0005	SFPwC	2	Weak cycle	Smooth	A
	1.0005	FPN	3	N-Pattern	BR	A
	1.0005	SFPN	3	N-Pattern	Smooth	A
	1.0005	FPsC	2	Strong cycle	BR	A
	1.0005	FPwC	20	Weak cycle	BR	A
	1.0005	SFPsC	2	Strong cycle	Smooth	A

A.8 Random generated games

#	Payoff/round	Name	Param	Forecaster	Responder	1st move
1	71.2130641772152	FPN	3	Spilio	BR	A
2	71.1550930734177	FPN	3	Spilio	Smooth	A
3	70.9770168202532	FPN	2	Spilio	Smooth	A
4	70.970158121519	FPN	2	Spilio	BR	A
5	70.6858878278481	FPsC	2	Strong	Smooth	A
6	70.6858404607595	FPsC	2	Strong	BR	A
7	70.6858271848101	FPsC	3	Strong	Smooth	A
8	70.6856380101266	FPsC	2	Strong	BR	A
9	70.6855036101266	FPsC	2	Strong	Smooth	A
10	70.6799129367089	FPsC	3	Strong	BR	A
11	70.3210242177215	FPwC	2	Weak	Smooth	A
12	70.3123301468355	FPwC	2	Weak	BR	A
13	70.2079162734177	FPwC	3	Weak	BR	A
14	70.2077754683544	FPwC	3	Weak	Smooth	A
15	70.177066243038	SWFP		Weighted	Smooth	A
16	70.1684587594937	WFP		Weighted	BR	A
17	70.0846400860759	FPN	3	Spilio	Smooth	B
18	70.0799790936709	FPN	3	Spilio	BR	B
19	69.8905717113924	FPwC	2	Weak	Smooth	A
20	69.8733425670886	FPwC	2	Weak	BR	A
21	69.845581235443	FP		Simple	BR	A
22	69.8368328759494	SFP		Simple	Smooth	A
23	69.6590389468354	FPN	2	Spilio	Smooth	B
24	69.6466559493671	FPN	2	Spilio	BR	B
25	69.593964521519	FPsC	3	Strong	Smooth	B
26	69.5938256202532	FPsC	2	Strong	BR	B
27	69.5936946075949	FPsC	2	Strong	Smooth	B
28	69.5931461518987	FPsC	2	Strong	BR	B
29	69.5930190278481	FPsC	2	Strong	Smooth	B
30	69.5925494683544	FPsC	3	Strong	BR	B
31	69.2230773873418	FPwC	2	Weak	BR	B
32	69.2226707848101	FPwC	2	Weak	Smooth	B
33	69.1129916253165	FPwC	3	Weak	Smooth	B
34	69.1129722582278	FPwC	3	Weak	BR	B
35	69.0921469113924	SWFP		Weighted	Smooth	B
36	69.0836002632911	WFP		Weighted	BR	B
37	68.7978376253164	FPwC	2	Weak	BR	B
38	68.797764243038	FPwC	2	Weak	Smooth	B
39	68.744983686076	SFP		Simple	Smooth	B
40	68.7448962987342	FP		Simple	BR	B