

# A Verification Framework for Fictitious Play Based Learning Algorithms

M. Smyrnakis, H. Y. Qu, S. Veres

*Department of Automatic Control and Systems Engineering, Faculty of Engineering, University of Sheffield.*

## Abstract

Distributed optimisation techniques have gained increasing attention due to fast development of autonomous robots. Many algorithms have been proposed to make optimisation more efficient. In this paper we propose a framework, which is based on probabilistic verification techniques, in order to compare the performance of various game-theoretic algorithms, in particular, fictitious play and its variants, after a finite number of iterations. To demonstrate the effectiveness of the framework, we apply the framework to a game which is inspired by wireless communication network problems, on five variations of fictitious play algorithms.

**Keywords** Probabilistic Verification; Fictitious Play; Game Semantics

## 1. INTRODUCTION

Distributed constrained optimisation (DCOP) tasks have many applications, e.g., wireless sensor networks and smart grid, and numerous algorithms have been used to solve DCOP tasks among a set of agents. In many cases, agents have limited coordination steps due to the nature of the task or communication limitations. Fictitious play is of particular interest as the canonical example of game-theoretic learning algorithms. It is known that it converges to Nash equilibrium in many classes of games. However, this convergence can be very slow. It is also possible to become trapped in cycles which correspond to sub-optimal rewards.

The classical way to study these algorithms for a given game is to run simulations over the game. If a DCOP learning algorithm allows probabilistic strategies, then only one possible outcome can be checked in one simulation run. Thus a large number of simulations are needed to reveal the possible outcomes of a learning algorithm and to gain confidence in it. Also there is no rule to specify the number of repetitions needed, and to guarantee that after the simulations every possible outcome will be discovered. Moreover, it is difficult to compare the general behaviour of two or more algorithms, e.g. how often they converge to an equilibrium point or to compare possible rewards of the players, since we do not know all the possible outcomes and their probability distribution.

To attack these problems and to provide solutions, we propose a verification framework to analyse the behaviour of game-theoretic DCOP algorithms. The techniques can examine all possible joint actions of the agents, states of the world after a finite number of coordination steps between the agents, and how likely a joint action is observed. In addition, this framework can be used as a comparison tool for the short term performance between various DCOP algorithms in tasks of interest.

## 2. FICTITIOUS PLAY BASED LEARNING ALGORITHMS

### 2.1. FICTITIOUS PLAY (FP) [1]

FP is a widely used learning technique in game theory. In FP, each player chooses his action according to the best response to his beliefs about his opponents' joint mixed strategy. The players, after each iteration, update their

beliefs about their opponents' strategies, and play again the best response according to their new beliefs.

### 2.2. VARIANTS OF THAT DISCOUNT THE IMPORTANCE OF OLD ACTIONS

FP is based on the implicit assumption that players use the same strategy in every iteration of the game. Therefore all the previously observed actions have the same importance. Geometric Fictitious Play (GFP) [2] addresses this problem by exponentially decaying the importance of the historical observations by a constant factor. Adaptive Forgetting Factor Fictitious Play (AFFFP) [3] instead of using a constant discount factor updates the discount factor in each iteration of the game, by using gradient accent and the likelihood of the previously observed action.

### 2.3. VARIANTS OF FP THAT MODEL OPPONENTS' STRATEGIES

An alternative to discount the importance of the old actions is to predict opponents' strategies. Two variants of FP, where players predict their opponents' strategies and then use these predictions to choose an action, are Extended Kalman Filters Fictitious Play (EKFFP) [4] and Particle Filters Fictitious Play (PFFP) [5].

## 3. VERIFICATION FRAMEWORK

Our framework was inspired by probabilistic model checking, in particular, verification of discrete-time Markov chains (DTMCs) [6]. It explores every possible execution trace of a system to verify the probabilistic behavioural properties a system has. It is the first time that it is employed to facilitate the analysis of DCOP using game theory and fictitious play in particular. Our framework constructs a model of a DCOP game under a learning algorithm, and computes the probability of reaching a Nash equilibrium in the model. The model is composed of a set of *states*, and a set of *transitions* between states. The state is updated at each step of the game and it contains all information that the algorithm needs to determine the next move. A transition between two states is then the result of the joint action of the players and the environment. Each transition is associated with a probability determined by the algorithm, expressing the likelihood of choosing the corresponding joint action in the next move.

One of the key issues in the development of the framework is how to restrain the size of a model to keep computational

complexity acceptable and without losing accuracy of coverage of the original DCOP. In principle, such a restrained model can be unbounded because a game can be played infinitely long even if a Nash equilibrium is reached. To avoid this, we propose *behaviour similarity relation* between states. If two states have a similar behaviour, then only one of them is expanded. It can be proved that the game will demonstrate the same behaviour from these two states, and thus, there is no need to explore both states, hence the name of the similarity relation. Using the similarity relation, a compact model can be obtained in which a Nash equilibrium is captured by a strongly connect component (SCC) in the model. In addition, no transition starting from any state in the SCC could go to states outside the SCC. The sum of transition probabilities for any state is one, which indicates that the model is indeed a DTMC. Therefore, the probabilistic model checking technique for computing reachability probabilities on DTMCs [6] can be adopted to find out the probability of reaching a given Nash equilibrium. The proposed framework is used to make comparisons between the classic fictitious play algorithm (FP) and some of its variations: EKFFP, GFP, AFFFFP and PFFP. It is worth emphasising that although the proposed verification framework is only used in a game theoretic context in this paper, it can be applied to any distributed optimisation algorithms with functions of interest that have not been cast as games, though this is beyond the scope of this paper.

#### 4. EXPERIMENTS

The usefulness of the proposed framework can be seen even in the case of the simple game which is depicted in Table 1, where player 1 can execute action  $A$  or  $B$ , and player 2 can execute  $C$  or  $D$ . Joint actions  $(A, C)$  and  $(B, D)$  lead to positive rewards  $(\alpha, \alpha)$ . This game can be seen as a collision avoidance game, where two UAVs are the players and they gain some reward if they fly at different altitudes. It can also be seen as a channel selection process in wireless communication networks, where two transmitters have to choose between two channels to transmit their data.

Table 1. A simple coordination game.

|     |                    |                    |
|-----|--------------------|--------------------|
|     | $C$                | $D$                |
| $A$ | $[\alpha, \alpha]$ | $[0, 0]$           |
| $B$ | $[0, 0]$           | $[\alpha, \alpha]$ |

An example of the possible states that is generated from our framework for the FP algorithm is depicted in Fig. 1. The matrix in each state illustrates the possibility of choosing each joint action in Table 1. Fig. 2 depicts the probability that each of the tested algorithms has in order to converge a Nash equilibrium with positive rewards. This figure shows that EKFFP is the best learning algorithm among the ones we tested, for the game in Table 1. On the other hand AFFFFP

the worst algorithm since it has the smaller chance to converge in one of the two pure Nash equilibria.

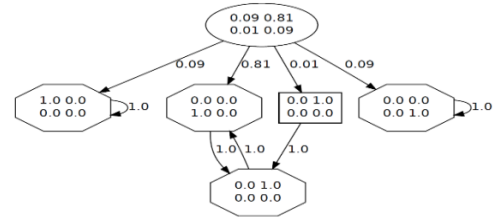


Figure 1. The model under the FP algorithm.

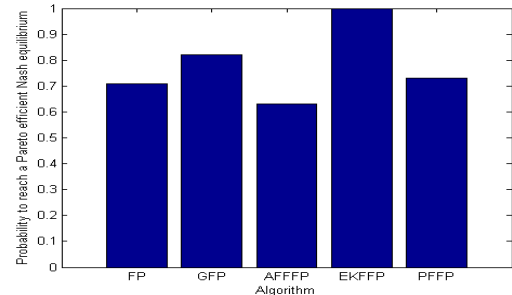


Figure 2. The experimental result.

#### 5. CONCLUSIONS

We developed a probabilistic model checking based framework to study the behaviour of various game-theoretic learning algorithms that are based on fictitious play, and compared their performance. This framework can also be used to explore possible Nash equilibria for robot cooperation in a finite number of iterations, with various initial conditions. The case study on a smart grid model shows that FP, GFP, and AFFFFP are sensitive to the initial conditions and they can easily be trapped in non-Pareto efficient solutions, while EKFFP and PFFP converge more often to the Pareto efficient equilibrium of the game. The future work includes comparisons in games with various numbers of players, and test under our framework non-game-theoretic algorithms for distributed optimisation.

#### REFERENCES

1. Brown G. W., Iterative Solutions of Games by Fictitious Play, 1951; In Activity Analysis of Production and Allocation, p. 374-376, Wiley.
2. Fudenberg D., Levine D. S., The Theory of Learning in Games, 1998, The MIT Press.
3. Smyrnakis M., Leslie, D. S., [Adaptive Forgetting Factor Fictitious Play](#).
4. Smyrnakis M., Veres, S., Coordination of Control in Robot Teams Using Game-theoretic Learning, IFAC, 2014.
5. Smyrnakis M., Leslie D. S., [Dynamic Opponent Modelling in Fictitious Play, 2010; The Computer Journal 53\(9\):1344-1359](#).
6. Courcoubetis C., Yannakakis M., [Verifying Temporal Properties of Finite State Probabilistic Programs, 1988; In proceedings of FOCS'88, p. 338-345, IEEE Computer Society Press](#).