Software Agents for Learning Nash Equilibria in Non-Cooperative Games

Alfredo Garro, Marco Iusi

***Abstract*—This paper describes SALENE, a Multi Agent System (MAS) for learning Nash Equilibria in non-cooperative games. SALENE is based on the following assumptions: if agents representing the players act as rational players, i.e. they act to maximise their expected utility in each match of a game, and if such agents play k matches of the game they will converge in playing one of the Nash Equilibria of the game. SALENE can be conceived as a heuristic and efficient method to compute at least one Nash Equilibria in a non-cooperative game represented in its normal form.**

Index Terms—**Multi-Agent Systems, Game Theory, Nash Equilibria.**

I. Introduction

The complexity of NASH [21], the problem consisting in computing Nash equilibria in non-cooperative games, is considered one of the most important open problem in Complexity Theory [22]. In 2005, Daskalakis, Goldbergy, and Papadimitriou showed that the problem of computing a Nash equilibrium in a game with four or more players is complete for the complexity class PPAD[1] [7], moreover, Chen and Deng extended this result for 2-player games [5]. However, even in the two players case, the best algorithm known has an exponential worst-case running time [23]; furthermore, if the computation of equilibria with simple additional properties is required, the problem immediately becomes NP-hard [3, 6, 11, 12].

Motivated by these results, recent studies have dealt with the problem of computing Nash Equilibria by exploiting approaches based on the concepts of learning and evolution [10, 15]. In these approaches the Nash Equilibria of a game are not statically computed but are the result of the evolution of a system composed by agents playing the game. In particular, each agent after different rounds will learn to play a strategy that, under the hypothesis of agent's rationality, will be one of the Nash equilibria of the game [2, 4, 9, 13, 18].

In this paper we present SALENE, a MAS for learning Nash Equilibria in non-cooperative games. In particular, given a static non cooperative game described in its normal form, the agents of the system will play the static game k times; after each match each agent will decide which strategy to play in the next match on the basis of his beliefs about the strategies that the other agents are adopting. More specifically, each agent assumes that his beliefs about the other players' strategies are correct and he plays a strategy that is a best response to his beliefs. By increasing k the agents will converge in playing one of the Nash equilibria of the game.

This paper is structured as follows. In Section 2, a formal definition of the problem will be given and the system requirements detailed. In Section 3 and in Section 4 the design and the implementation of SALENE will be described respectively, then, in Section 5, some experimental results will be shown. Finally, in Section 6, conclusions and future efforts will be addressed.

II. Problem Definition and System Requirements

An n-person strategic game G can be defined as a tuple $G = (N; (A^i)_{i\in N}; (r^i)_{i\in N})$, where $N = \{1, 2, \ldots, n\}$ is the set of players, A^i is a finite set of actions for player $i\in N$, and $r^i : A^1 \times \ldots \times A^n \rightarrow \Re$ is the payoff function of player i. The set A^i is called also the set of pure strategies of player i. The Cartesian product $\times_{i\in N}A^i = A^1 \times \ldots \times A^n$ can be denoted by A and $r : A \rightarrow \Re^N$ can denote the vector valued function whose ith component is r^i, i.e., $r(a) = (r^1(a), \ldots, r^n(a))$, so it is possible to write (N, A, r) for short for $(N; (A^i)_{i\in N}; (r^i)_{i\in N})$.

For any finite set A^i the set of all probability distributions on A^i can be denoted by $\Delta(A^i)$. An element $\sigma^i \in \Delta(A^i)$ is a mixed strategy for player i.

A (Nash) equilibrium of a strategic game $G = (N, A, r)$ is an N-tuple of (mixed) strategies $\sigma = (\sigma^i)_{i\in N}$, $\sigma^i \in \Delta(A^i)$, such that for every $i \in N$ and any other strategy of player i, $\tau^i \in \Delta(A^i)$, $r^i(\tau^i,\sigma^{-i}) \leq r^i(\sigma^i,\sigma^{-i})$, where r^i denotes also the expected payoff to player i in the mixed extension of the game and σ^{-i} represents the mixed strategies in σ of all the other players. Basically, supposing that all the other players do not change their strategies it is not possible for any player i to play a different strategy τ^i able to gain a better payoff of that gained by playing σ^i. σ^i is called a Nash equilibrium strategy for player i.

In 1951 J. F. Nash proved that a strategic (non-cooperative) game $G = (N, A, r)$ has a (Nash) equilibrium σ [17]; in his honour, the computational problem of finding such equilibria is known as NASH [21].

A. Garro is with the Department of Electronics, Informatics and Systems (DEIS), University of Calabria, Rende (CS), 87036 Italy. (e-mail: alfredo.garro @unical.it).

M. Iusi is with the Department of Electronics, Informatics and Systems (DEIS), University of Calabria, Rende (CS), 87036 Italy. (e-mail: marco.iusi@tin.it).

[1] PPAD (polynomial parity argument, directed version) class was introduced by Papadimitriou in his seminal work in 1991 [20].

In order to exemplify the definitions given above let us consider a game with two players (n=2) and $|A^1|=|A^2|=m$, i.e., the sets of pure strategies have both cardinality equals to m [3]. In this case the set of pure strategies for each player could be identified with the ordered set M = {1, 2, . . . , m} and the game could be represented by two $m \times m$ matrices B and W. The first player is called the row player and the second player is called the column player. If the row player plays strategy i and the column player strategy j, the payoff will be B_{ij} for the first player and W_{ij} for the second player.

A mixed strategy, a probability distribution over pure strategies, is a vector $\chi \in \Re^M$ such that $\sum_{s \in M} \chi_s = 1$ and for every $s \in M$, $\chi_s \geq 0$.

When the row player plays mixed strategy χ and the column player plays mixed strategy γ, their expected payoffs will be, respectively, $\chi^t B\gamma$ and $\chi^t W\gamma$ (χ^t is the transpose of vector χ).

A Nash Equilibrium of the game described by the matrices B and W is a pair of mixed strategies (χ, γ) such that for all mixed strategies $\overline{\chi}$ and $\overline{\gamma}$, of the row and the column player respectively, $\chi^t B\gamma \geq \overline{\chi}^t B\gamma$ and $\chi^t W\gamma \geq \chi^t W\overline{\gamma}$.

Starting from the problem definition discussed above, SALENE was conceived as a system for learning at least one Nash Equilibrium of a non-cooperative game given in the form $G = (N; (A^i)_{i\in N}; (r^i)_{i\in N})$. In particular, the system asks the user for:

- the number n of the players which defines the set of players $N = \{1, 2, \dots, n\}$;
- for each player $i \in N$, the related finite set of pure strategies A^i and his payoff function $r^i : A^1 \times \dots \times A^n \rightarrow \Re$;
- the number k of times the players will play the game.

Then, the system creates n agents, one associated to each player, and a referee. The players and the referee both know that G is the actual game to be played, i.e. there is *complete information* [8, 19]. Each player is a rational player i.e. his goal is to maximise his expected utility/payoff[2]. In particular, in SALENE a rational player acts to maximise his expected utility in each single match without considering the overall utility that he could obtain in a set of matches.

This kind of agents will play the game G k times, after each match, each agent will decide the strategy to play in the next match to maximise his expected utility on the basis of his beliefs about the strategies that the other agents are adopting. By increasing k the agents will converge in playing one of the Nash Equilibria of the game. This conclusion relays on the hypothesis that the agents will act as rational players and derives straightly from the assumptions on which the Nash's theorem is based [8, 17, 19, 25].

[2] Payoffs are numeric representations of the utility obtainable by a player in the different outcomes of a game.

III. SYSTEM DESIGN

On the basis of the requirements highlighted in the previous section the SALENE (Software Agent for LEarning Nash Equilibria) MAS was designed. The class diagram of SALENE is shown in Figure 1.

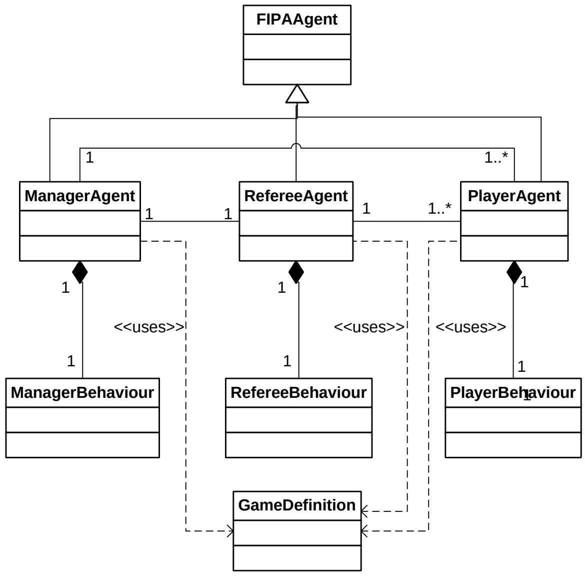


Fig. 1. Class diagram of SALENE

The Manager Agent interacts with the user and it is responsible for the global behaviour of the system. In particular, after having obtained from the user the input parameters G and k (see section II), the Manager Agent creates both n Player Agents, one associated to each player, and a Referee Agent that coordinates and monitors the behaviours of the players. The Manager Agent sends to all the agents the definition G of the game then he asks the Referee Agent to orchestrate k matches of the game G. In each match, the Referee Agent asks each Player Agent which pure strategy he has decided to play, then, after having acquired the strategies from all players, the Referee Agent communicates to each Player Agent both the strategies played and the payoffs gained by all players. After playing k matches of the game G the Referee Agent communicates all the matches' data to the Manager Agent which analyses it and properly presents the obtained results to the user.

A Player Agent is a rational player that, given the game definition G, acts to maximise his expected utility in each single match of G. In particular the behaviour of the Player Agent i can be described by the following main steps:

1. In the first match the Player Agent i chooses to play a pure strategy randomly generated considering all the pure strategies playable with the same probability: if $|A^i|=m$ the probability of choosing a pure strategy $s \in A^i$ is $1/m$.
2. The Player Agent i waits for the Referee Agent to ask him which strategy he wants to play, then he communicates to the Referee Agent the chosen pure strategy as computed in step 1 if he is playing his first match or in step 4 otherwise;

3. The Player Agent waits for the Referee Agent to communicate him both the pure strategies played and the payoffs gained by all players;
4. The Player Agent decides the mixed strategy to play in the next match. In particular, the Player Agent updates the beliefs about the mixed strategies currently adopted by the other players and consequently recalculate the strategy able to maximise his expected utility. Basically, the Player Agent i tries to find the strategy $\sigma^i \in \Delta(A^i)$, such that for any other strategy $\tau^i \in \Delta(A^i)$, $r^i(\tau^i,\sigma^{-i}) \leq r^i(\sigma^i,\sigma^{-i})$ where r^i denotes his expected payoff and σ^{-i} represents his beliefs about the mixed strategies currently adopted by all the other players, i.e. $\sigma^{-i}=(\sigma^j)_{j\in N, j\neq i}$, $\sigma^j \in \Delta(A^j)$. In order to evaluate σ^j for each other player $j\neq i$ the Player Agent i considers the pure strategies played by the player j in all the previous matches and computes the frequency of each pure strategy, this frequency distribution will be the estimate for σ^j. If there is at least an element in the actually computed set $\sigma^{-i}=(\sigma^j)_{j\in N, j\neq i}$ that differs from the set σ^{-i} as computed in the previous match, the Player Agent i solves the inequality $r^i(\tau^i,\sigma^{-i}) \leq r^i(\sigma^i,\sigma^{-i})$ that is equivalent to solve the optimization problem $P=\{\max(r^i(\sigma^i,\sigma^{-i})), \sigma^i\in\Delta(A^i)\}$. It is worth noting that P is a linear optimization problem, actually, given the set σ^{-i}, $r^i(\sigma^i,\sigma^{-i})$ is a linear objective function in σ^i (see the two players example reported in Section II)**,** and with $|A^i|=m$ $\sigma^i\in\Delta(A^i)$ is a vector $\chi \in \Re^M$ such that $\sum_{s\in M} \chi_s = 1$ and for every $s\in$M $\chi_s\geq 0$, so the constraint $\sigma^i\in\Delta(A^i)$ is a set of $m+1$ linear inequalities. P is solved by the Player Agent by using an efficient method for solving problems in linear programming [14, 16], in particular the predictor-corrector method of Mehrotra [16], whose complexity is polynomial for both average and worst case. The obtained solution for σ^i is a pure strategy because it is one of the vertices of the polytope which defines the feasibility region for P. The obtained strategy σ^i will be played by the Player Agent i in the next match; $r^i(\sigma^i,\sigma^{-i})$ represents the expected payoff to player i in the next match;
5. back to step 2.

The Manager Agent, receives from the Referee Agent all the data about the k matches of the game G and computes an estimate of a Nash Equilibrium of G, i.e. an N-tuple $\sigma=(\sigma^i)_{i\in N}$, $\sigma^i\in\Delta(A^i)$. In particular, in order to estimate σ^i (the Nash equilibrium strategy of the player i), the Manager Agent computes, on the basis of the pure strategies played by the player i in each of the k match, the frequency of each pure strategy: this frequency distribution will be the estimate for σ^i. The so computed set $\sigma=(\sigma^i)_{i\in N}$, $\sigma^i\in\Delta(A^i)$ will be then properly proposed to the user together with the data exploited for its estimation.

IV. System Implementation

The JADE-based classes of SALENE were straightforwardly derived from the class diagram reported in Figure 1. In particular:

– *ManagerAgent*, *RefereeAgent* and *PlayerAgent* extend the *Agent* class of JADE [1];
– *ManagerBehaviour*,*RefereeBehaviour* and *PlayerBehaviour* extend *FSMBehaviour* class of JADE which models a complex task whose sub-tasks correspond to the activities performed in the states of a finite state machine. In particular, the behaviours of both the Referee and the Player Agent are also cyclic.

The interactions among SALENE Agents are appositely defined through sequences of ACL messages instances of the *ACLMessage* class of JADE.

V. Experimental Results

SALENE was tested on different games that differ from each other both in the number and in the kind of Nash Equilibria. This section presents the results obtained for three popular games: (1) *The Prisoner's Dilemma* which has one pure Nash Equilibrium (that is an equilibrium in which all the players play a pure strategy); (2) *Matching Pennies* which has one mixed Nash Equilibrium (that is an equilibrium in which at least one player plays a mixed strategy); (2) *Battle of the Sexes* which has three Nash Equilibria (two Pure Equilibria and one Mixed Equilibrium).

A. The Prisoner's Dilemma

An informal description of the *Prisoner's Dilemma* can be found in [24]. Formally, in a game G of *Prisoner's Dilemma* (PD), two players (n=2) simultaneously choose a move, either cooperate (c) or defect (d), so $A^1=A^2=\{c,d\}$ and $|A^1|=|A^2|=m=2$. There are thus four possible outcomes for each encounter: both cooperate (cc), the first player cooperates, while the second defects (cd), vice versa (dc), and both players defect (dd). Each player receives a payoff after each encounter as reported in Table Ia-b. Table I semantic derives straightly from the *bimatrix* representation of a two-player game as discussed in Section II. In particular, the move of Player 1 determines the row, the move of Player 2 determines the column, and the pair (X,Y) in the corresponding cell indicates that payoff of Player 1 is X and the payoff of Player 2 is Y. Regarding the payoffs reported in Table Ia the following order must hold: T>R>P>L. Table Ib shows a valid assignment for the payoffs.

TABLE I
(A) Payoffs For Prisoner's Dilemma

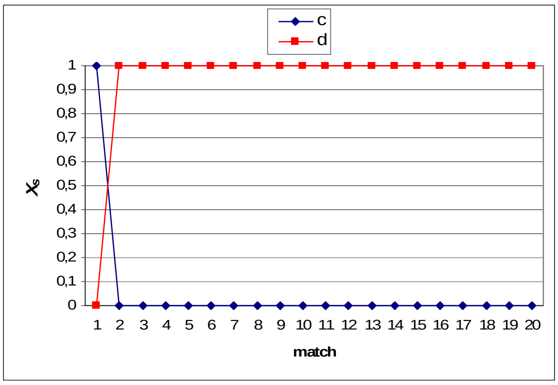
		Player 2	
		c	d
Player 1	c	R,R	L,T
	d	T,L	P,P

(B) A Valid Assignment for the Payoffs (T>R>P>L)

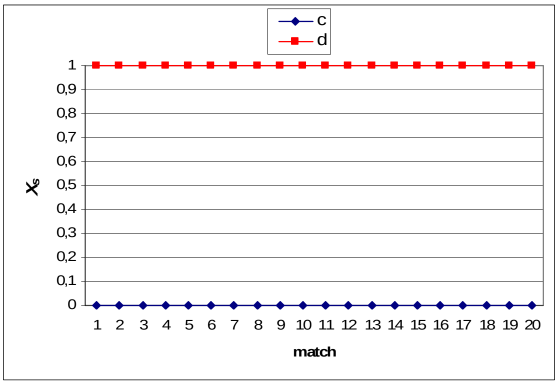
		Player 2	
		c	d
Player 1	c	6,6	0,10
	d	10,0	3,3

By looking at Table Ib, it is possible to note that for Player 1 d is the best response to c (10>6) and d is also the best response to d (3>0). The same is true for Player 2, so both players rationally will play their pure strategy d that is their dominant strategy. Formally, the *Prisoner's Dilemma* has one Nash Equilibrium $\sigma=\{\sigma^1,\sigma^2\}=\{\chi,\gamma\}$, $\sigma^1\in\Delta(A^1)$, $\sigma^2\in\Delta(A^2)$, $\chi\in\Re^2$, $\gamma\in\Re^2$, where $\chi=[0, 1]$ and $\gamma=[0, 1]$.

In order to compare the analytical result with the result obtainable in SALENE we ran 30 experiments each consisting of 100 matches (k=100) of the *Prisoner's Dilemma*. In the case of the *Prisoner's Dilemma* the expected result was that as soon as a Player Agent played his dominant strategy *d*, he would never change his choice, so after few matches the Player Agents played both their dominant strategy *d* so converging in playing the Nash Equilibrium of the game. The experiments confirm the expected result, as an example Figure 2a-b reports one of the experiments carried out. In particular, Figure 2a(2b) shows the strategy played by Player 1(Player 2) in each of the *k* match of an experiment. As showed in Figure 2a-b, after few matches both the Player Agents play their pure strategy *d* as required by the Nash Equilibrium of the game.



(a) Strategy played by Player 1



(b) Strategy played by Player 2

Fig. 2. The Prisoner's Dilemma: experimental results

B. Matching Pennies

An informal description of the *Matching Pennies* game follows: the game is played between two players, each player has a penny and must secretly turn it to heads or tails, the players then reveal their choices simultaneously; if the pennies match (both heads or both tails), Player 1 receives S dollars from Player 2. If the pennies do not match (one heads and one tails), Player 2 receives S dollars from Player 1. This is an example of a zero-sum game, where one player's gain is exactly equal to the other player's loss. Formally, in a game *G* of *Matching Pennies* (MP), two players (n=2) simultaneously choose a move, either heads (h) or tails (t), so $A^1=A^2=\{h,t\}$ and $|A^1|=|A^2|=m=2$. Each player receives a payoff after each encounter as reported in Table IIa-b.

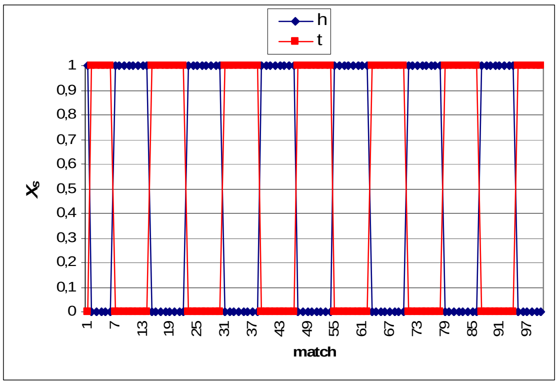
TABLE II

(A) PAYOFFS FOR MATCHING PENNIES

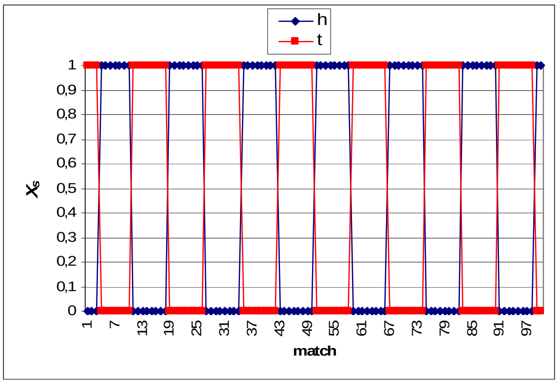
		Player 2	
		h	t
Player 1	h	L,-L	-L,L
	t	-L,L	L,-L

(B) A VALID ASSIGNMENT FOR THE PAYOFFS (S=1)

		Player 2	
		h	t
Player 1	h	1,-1	-1,1
	t	-1,1	1,-1



(a) Strategy played by Player 1



(b) Strategy played by Player 2

Fig. 3. Matching Pennies: experimental results

Matching Pennies has no pure strategy Nash equilibrium since there is no pure strategy (heads or tails) that is a best response to a best response, i.e. a dominant strategy. Alternatively, there is not a pure strategy that a player would ever change when told the pure strategy played by the other player. Instead, the unique Nash equilibrium of *Matching Pennies* is in mixed strategies: each player chooses heads or tails with equal probability. In this way, each player makes the other indifferent in choosing heads or tails, so neither player has an incentive to try another strategy. Formally, *Matching Pennies* has one mixed Nash Equilibrium $\sigma=\{\sigma^1,\sigma^2\}=\{\chi,\gamma\}$, $\sigma^1\in\Delta(A^1)$, $\sigma^2\in\Delta(A^2)$, $\chi\in\Re^2$, $\gamma\in\Re^2$, where $\chi=[0.5, 0.5]$ and $\gamma=[0.5, 0.5]$.

In order to compare the analytical result with the result obtainable in SALENE we ran 30 experiments each consisting of 100 matches (k=100) of *Matching Pennies*. The expected result was that, analyzing the pure strategies played by each player in each of the k match, their frequency distribution would asymptotically converge to the mixed Nash Equilibrium of the game. The experiments confirm the expected result: by increasing k the computed frequency distributions asymptotically converge to the mixed Nash Equilibria of the game. As an example Figure 3a-b reports one of the experiments carried out. In this case the computed frequency distributions were: $\sigma^1=\chi=[0.49, 0.51]$ and $\sigma^2=\gamma=[0.49, 0.51]$.

C. Battle of Sexes

An informal description of the *Battle of Sexes* game follows: a man and a woman plan to meet after work to attend an event: an opera or a football match, but they can not communicate so they have to choose separately where to go. The woman prefers the opera to the football match, whereas the man prefers the football match to the opera, but both prefer to be together at either event than alone at either one. More formally in a game G of *Battle of Sexes* (BS), two players (n=2) simultaneously choose a move, either opera (o) or football (f), so $A^1=A^2=\{o,f\}$ and $|A^1|=|A^2|=m=2$. Each player receives a payoff after each encounter as reported in Table IIIa-b. Regarding the payoffs reported in Table IIIa the following order must hold: T>R>L. Table IIIb shows a valid assignment for the payoffs.

TABLE III
(A) PAYOFFS FOR BATTLE OF SEXES

		Player 2	
		o	f
Player 1	o	T,R	L,L
	f	L,L	R,T

(B) A VALID ASSIGNMENT FOR THE PAYOFFS (T>R>L)

		Player 2	
		o	f
Player 1	o	3,2	0,0
	f	0,0	2,3

Battle of Sexes has two pure strategy Nash equilibria, one where both go to the opera and another where both go to the football game; there is also a Nash equilibrium in mixed strategies, where, given the payoffs listed in Table IIIb, each player attends their preferred event with probability 2/3. Formally, *Battle of Sexes* has three Nash Equilibria:

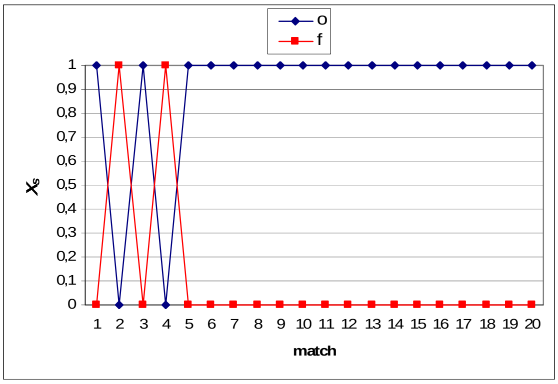
$\sigma_I=\{\sigma_I^1, \sigma_I^2\}=\{\chi_I,\gamma_I\}$, $\sigma_I^1 \in \Delta(A^1)$, $\sigma_I^2 \in \Delta(A^2)$, $\chi_I \in \Re^2$, $\gamma_I \in \Re^2$, where $\chi_I=[1, 0]$ and $\gamma_I=[1, 0]$;

$\sigma_{II}=\{\sigma_{II}^1, \sigma_{II}^2\}=\{\chi_{II},\gamma_{II}\}$, $\sigma_{II}^1 \in \Delta(A^1)$, $\sigma_{II}^2 \in \Delta(A^2)$, $\chi_{II} \in \Re^2$, $\gamma_{II} \in \Re^2$, where $\chi_{II}=[0, 1]$ and $\gamma_{II}=[0, 1]$;

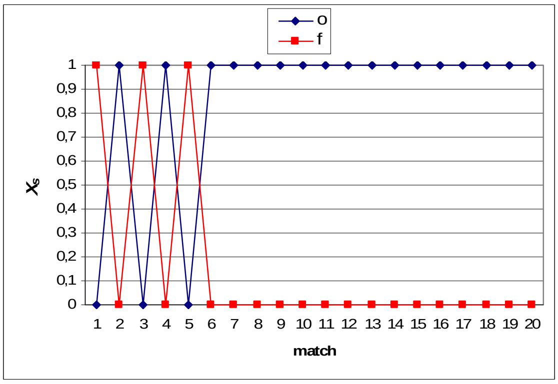
$\sigma_{III}=\{\sigma_{III}^1, \sigma_{III}^2\}=\{\chi_{III},\gamma_{III}\}$, $\sigma_{III}^1 \in \Delta(A^1)$, $\sigma_{III}^2 \in \Delta(A^2)$, $\chi_{III} \in \Re^2$, $\gamma_{III} \in \Re^2$, where $\chi_{III}=[2/3, 1/3]$ and $\gamma_{III}=[1/3, 2/3]$;

In order to compare the analytical result with the result obtainable in SALENE, we ran 30 experiments each consisting of 100 matches (k=100) of the *Battle of Sexes*. *Battle of Sexes* presents an interesting case for games theory since each of the Nash Equilibria is deficient in some way. The two pure strategy Nash Equilibria are unfair, one player consistently does better than the other. In the mixed strategy Nash Equilibrium the players will be together at the same event with probability 4/9 and will be alone with probability 5/9, leaving each player with an expected payoff of 10/9 that is very low if compared with the expected payoff of the two pure Nash Equilibria.

The expected result was that as soon as both the Player Agents played the same pure strategy (o or f), i.e. one of the Pure Nash Equilibria of the game, the Agents would never change their choices: the Player who plays his favorite strategy will not have incentive to change it, the player who does not play his favorite strategy will not change it because in this case his expected payoff will get worse in the next match. In particular, after 1 or h*(T+R) matches, $k \geq h \geq 1$, there is a probability of 50% that from this match on the Player Agents will converge in playing one of the Pure Nash Equilibria of the game.



(a) Strategy played by Player 1



(b) Strategy played by Player 2
Fig. 4. Battle of Sexes: experimental results

The experiments confirm the expected result: in all the experiments after 1 or h*(T+R) matches the Player Agents play one of the two pure Nash Equilibria of the game. As an example Figure 4a-b reports one of the experiments carried out, in this case the played Nash Equilibrium was σ_I.

VI. CONCLUSIONS

The complexity of NASH, the problem consisting in computing Nash equilibria in non-cooperative games, is still debated, but even in the two players case, the best algorithm known has an exponential worst-case running time. Starting from these considerations SALENE, a MAS for learning Nash Equilibria in non cooperative games, was developed.

SALENE is based on the assumptions that if agents representing the players act as rational players, i.e. if each player acts to maximise his expected utility in each match of a game *G*, and if such agents play *k* matches of *G* they will converge in playing one of the Nash Equilibria of the game. In particular, after each match each agent decides the strategy to play in the next match on the basis of his beliefs about the strategies that the other agents are adopting. More specifically, each agent assumes that his beliefs about the other players' strategies are correct and plays a strategy that is a best response to his beliefs. Analyzing the behaviour of each agent in all the *k* matches of *G*, SALENE presents to the user an estimate of a Nash Equilibrium of the game.

A set of experiments was carried out on different games that differ from each other both in the number and in the kind of Nash Equilibria. The experiments demonstrated that:

- if the game has one Pure Nash Equilibrium the agents converge in playing this equilibrium;
- if the game has one Mixed Nash Equilibrium, the frequency distributions of the pure strategies played by each player asymptotically converge to the mixed Nash Equilibrium of the game;
- if the game has $p>1$ Pure Nash Equilibria and $s>1$ Mixed Nash Equilibria the agents converge in playing one of the *p* Pure Nash Equilibria.

SALENE can be conceived as a heuristic and efficient method for computing at least one Nash Equilibria in a non-cooperative game represented in its normal form; actually, the learning algorithm adopted by the Player Agents has a polynomial running time [14, 16] for both average and worst case.

Efforts are currently underway to: (i) evaluate different learning algorithms and extensively testing them on complex games; (ii) let the user ask for the computation of equilibria with simple additional properties.

REFERENCES

[1] F. Bellifemine, A. Poggi, and G. Rimassa. *Developing multi-agent systems with a FIPA-compliant agent framework*. In Software Practice and Experience, 31, pp. 103-128, 2001.

[2] M. Benaim and M.W. Hirsch. *Learning process, mixed equilibria and dynamic system arising for repeated games*. In Games Econ. Behav., 29, pp. 36-72, 1999.

[3] V. Bonifaci, U. Di Iorio, and L. Laura. *On the complexity of uniformly mixed Nash equilibria and related regular subgraph problems*. In Proceedings of the 15th International Symposium on Fundamentals of Computation Theory (FCT) 2005.

[4] D. Carmel and S. Markovitch. *Learning Models of Intelligent Agents*. In Proceedings of the 13th National Conference on Artificial Intelligence and the Eighth Innovative Applications of Artificial Intelligence Conference, Vol. 2, pp 62-67, AAAI Press, Menlo Park, California, 1996.

[5] X. Chen and X. Deng. *Settling the Complexity of 2-Player Nash-Equilibrium*. Electronic Colloquium on Computational Complexity, Report No. 140 (2005).

[6] V. Conitzer and T. Sandholm. *Complexity results about Nash equilibria*. In Proceedings of the 18th Int. Joint Conf. on Artificial Intelligence, pages 765–771, 2003.

[7] C. Daskalakis, P. W. Goldbergy, and C. H. Papadimitriou. *The Complexity of Computing a Nash Equilibrium*. Electronic Colloquium on Computational Complexity, Report No. 115 (2005).

[8] A.K. Dixit and S. Skeath. *Games of Strategy (2nd edition)*. W.W. Norton & Company, 2004.

[9] D.B. Fogel. *Evolving Behaviours in the Iterated Prisoner's Dilemma*, Evolutionary Computation, Vol. 1:1, pp 77-97, 1993.

[10] D. Fudenberg and D. Levine. *The Theory of Learning in Games*. Cambridge, MA, MIT Press, 1998.

[11] I. Gilboa and E. Zemel. *Nash and correlated equilibria: some complexity considerations*. Games and Economic Behavior, 1(1):80–93, 1989.

[12] G. Gottlob, G. Greco, and F. Scarcello. *Pure Nash Equilibria: hard and easy games*. In Proceedings of the 9th Conference on Theorethical Aspects of Rationality and Knowledge (TARK-2003), Bloomington, Indiana, USA, June 20-22, 2003.

[13] P. Jehiel and D. Samet. *Learning to Play Games in Extensive Form by Valuation*, NAJ Economics, Peer Reviews of Economics Publications, 3, 2001.

[14] N. Karmarkar. *A New Polynomial-time Algorithm for Linear Programming*. In Combinatorica 4, 373-395, 1984.

[15] J. Maynard Smith. *Evolution and the Theory of Games*. Cambridge University Press, 1982.

[16] S. Mehrotra. *On the Implementation of a Primal-dual Interior Point Method*. In SIAM J. Optimization 2, 575-601, 1992.

[17] J. F. Nash. *Non-cooperative games*. In Annals of Mathematics, volume 54, pages 289–295, 1951.

[18] A. Neyman. *Finitely Repeated Games With Finite Automata*. In Mathematics Of Operations Research, Vol. 23, No. 3, August 1998.

[19] M. J. Osborne. *An Introduction to Game Theory*. Oxford University Press, 2002.

[20] C. H. Papadimitriou. *On inefficient proofs of existence and complexity classes*. In Proceedings of the 4th Czechoslovakian Symposium on Combinatorics, 1991.

[21] C. H. Papadimitriou. *On the complexity of the parity argument and other inefficient proofs of existence*. Journal of Computer and Systems Sciences, 48(3):498–532, 1994.

[22] C. H. Papadimitriou. *Algorithms, Games and the Internet*. In Proceedings of the 33rd Annual ACM Symposium on the Theory of Computing (STOC), pages 749–753, 2001.

[23] R. Savani and B. von Stengel. *Exponentially many steps for finding a Nash equilibrium in a bimatrix game*. In Proceedings of the 45th Symp. Foundations of Computer Science, pages 258–267, 2004.

[24] G. Tsebelis. *Nested Games: rational choice in comparative politics*. University of California Press, 1990.

[25] J. Von Neumann and O. Morgenstern. *Theory of Games and economic Behaviour*, Princeton University Press, 1944.