Game theoretical solution concepts for learning agents with extensive–form games

(Extended Abstract)

Fabio Panozzo
Politecnico di Milano
Piazza Leonardo da Vinci 32
Milano, Italy
panozzo@elet.polimi.it

ABSTRACT

My Ph.D thesis focuses on the study of solution concepts for rational learning agents in extensive–form games in absence of common knowledge; specifically, on the definition of solution concepts, their search, analysis of static and dynamic property, characterization of learning dynamics. Summarily, my work is finalized to better understand how to integrate more thoroughly game theory and machine learning.

Categories and Subject Descriptors

I.2.11 [**Computing Methodologies**]: Distributed Artificial Intelligence

General Terms

Algorithms, Economics

Keywords

Game Theory (cooperative and non-cooperative)

1. INTRODUCTION

Game theory provides the most elegant formal tools to study strategic interaction situations. Many of these describe situations in which the assumption of common knowledge holds (e.g. solution concepts like Nash equilibrium and its refinements), but in real–world situations this assumption is rarely verified. Thus to study these situations we must resort to alternative methods in which each agent has own beliefs and can learn from its observations.

The objective of my Ph.D. thesis is the theoretical and algorithmic study of the behavior of rational agent with extensive–form games. The work is motivated by the lack of a thorough study in literature of these situations that are common in real–world applications.

2. STATE OF THE ART

Appears in: *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2012)*, Conitzer, Winikoff, Padgham, and van der Hoek (eds.), 4-8 June 2012, Valencia, Spain.

2.1 Solution concepts

One of the most important solution concepts in game theory is the Nash equilibrium (NE): it describes a strategy profile in which each agent has no incentive to deviate unilaterally. In extensive–form games this notion of equilibrium is not appropriate because it allows agents to play strategies that in a real situation they would not play. There are some important refinements of NE. The most common is the concept of *subgame perfect equilibrium* (SPE) that constrains a strategy profile to be an NE in every subgame [5], where a subgame is a portion of the game tree well defined. The concept of SPE is satisfactory with perfect–information games, while it is not when information is imperfect. The "natural" extension of the SPE to situations with imperfect information is the *sequential equilibrium* (SE) [10]. An SE is defined as a pair $(\sigma, \hat{\sigma})$, where σ are the strategies of the agents and $\hat{\sigma}$ are the agents' beliefs over the opponents strategies, such that strategies σ are sequentially rational to the beliefs $\hat{\sigma}$ (in the sense of backward induction) and the beliefs $\hat{\sigma}$ are consistent to the strategies σ according to the notion of consistency of Kreps and Wilson. This notion of consistency requires essentially that the beliefs are correct everywhere in the game tree. The concept of SE is a refinement of the concept of SPE, the SPEs containing the SEs.

The previous concepts are based on assumption of common knowledge between agents that rarely is verify in the real world. For this reason the concept of self–confirming equilibrium (SCE) that relaxes the NE, capturing interesting settings in which information is not common, was introduced. Similarly to an SE, an SCE is defined as a pair $(\sigma, \hat{\sigma})$ where σ are best response to the beliefs $\hat{\sigma}$. Differently, beliefs $\hat{\sigma}$ can be wrong off the equilibrium path (instead they must be correct/confirmed on the equilibrium path). Obviously, every NE is an SCE (with correct beliefs), while an NE may be not an SCE.

2.2 Learning dynamics

Without the assumption of common knowledge, agents can only learn to play the game by trials and errors. The most common learning algorithm is Q–learning that can be utilized to study the agents' behavior in a game. In some situations, learning algorithms are proved to converge to some specific solution concept. However, also in these cases learning dynamics may be long and it could be more interesting to characterize the dynamics than the steady state. Evolutionary game theory [9] provides models (like replicator dynamics) to study steady states from a dynamic point of view to understand the features of these states and their basins of attraction. However, the characterization of the learning dynamics when there are multiple agents that repeatedly match is an open issue.

2.3 Equilibrium computation

Game theory and microeconomics provide only models and solution concepts, but they do not provide computational tools to deal with games. The development of these tools is an interesting topic, with the name of *equilibrium computation*, in computer science.

A number of computational results are known on the NE: computing an exact NE [3] and approximating it [2] are $\mathcal{PPAD}$–complete; $\mathcal{PPAD}$ is in $\mathcal{NP}$, it does not include $\mathcal{NP}$–complete problems unless $\mathcal{NP}$ = co–$\mathcal{NP}$, and it is not known whether $\mathcal{PPAD}$ is in $\mathcal{P}$, but it is commonly believed that it is not; instead, the problem to verify whether or not a solution is a NE is in $\mathcal{P}$.

A number of works deal with the problem to compute a NE with general–sum strategic–form games (especially with two agents), e.g., [1, 11, 13, 14], and with the problem to solve large zero–sum extensive–form games, e.g., [8]. An SPE can be easily found by applying *backward induction* [5].

The computation of an SE (and of an SPE) can be tackled by a simple variation of the Lemke's algorithm by introducing a specific lexicographic perturbation [12], this puts the problem to find an SE in the $\mathcal{PPAD}$ class. The problems of verifying a SE with an arbitrary number of agents and a QPE with two agents are in $\mathcal{P}$ [6].

The unique result on the computability of an SCE is discussed in [7], where the authors show that the problem to compute an SCE can be formulated as an MILP.

3. THE THESIS OBJECTIVE

The objective of my Ph.D. thesis is the theoretical and computational study of solution concepts and learning dynamics with extensive–form games with two or more agents and without the common knowledge assumption.

In particular I focus on:

- the study of the relation between solution concepts and learning dynamics: during the learning phase the agents can play strategy profiles that can be described in terms of solution concepts;
- the study of the complexity of the verification problem, i.e. given a strategy profile, verify whether or not it is a solution concept: in a learning dynamic can be observe strategies for finite time, it is important to know which equilibrium (even temporarily) they tend to;
- the study of the complexity of the search problem, i.e. given a representation of game (in general extensive–form) search a solution concept: during a learning process it is important to know where the agents could move to converge;
- the search for and characterization of the regions that containing the points that condition the learning dynamics;
- the study of the solution concepts under varying of the initial assumptions (e.g. epistemic–based games) or in asymmetric situations in which agents have different information about each other.

4. PROGRESS

I focused on the study of solution concepts in an extensive–form game that is repeatedly played by different individuals without common knowledge [4]. More precisely, for each agent (representing a *role*), there is a population (finite or infinite) of *individuals* and, at each repetition of the game, one individual is drawn from each population and these are matched and then play the game. I derived different mathematical formulations to find *finite* or *infinite heterogeneous self–confirming equilibrium* (HSCE), I proved relations between different types of self–confirming equilibria, I defined the region (in the space of utilities) of HSCE and I started to study equilibria in terms of attractors, saddles and repellers. At the current state I analyzed these features in two–players games, but, as next step of my work, I would like extend these concepts to multi–player (more than two) games and I will deeply study the characterization of equilibria in terms of dynamics with the aid of machine learning algorithms. I planned to study computational complexity to define the classes of problems of search an verify of equilibria.

5. REFERENCES

[1] S. Ceppi, N. Gatti, G. Patrini, and M. Rocco. Local search techniques for computing equilibria in two–player general–sum strategic–form games. In *AAMAS*, pages 1469–1470, 2010.

[2] X. Chen, X. Deng, and S. Teng. Computing Nash equilibria: approximation and smoothed complexity. In *FOCS*, pages 603–612, 2006.

[3] C. Daskalakis, A. Mehta, and C. Papadimitriou. A note on approximate Nash equilibria. *THEOR COMPUT SCI.*, 410(17):1581–1588, 2009.

[4] E. Dekel, D. Fudenberg, and D. K. Levine. Learning to play bayesian games. *GAME ECON BEHAV*, 46:282–303, 2004.

[5] D. Fudenberg and J. Tirole. *Game Theory*. 1991.

[6] N. Gatti and F. Panozzo. New results on the verification of nash refinements for extensive–form games. *AAMAS*, 2012.

[7] N. Gatti, F. Panozzo, and S. Ceppi. Computing a self–confirming equilibrium in two–player extensive–form games. pages 981–988, 2010.

[8] A. Gilpin, T. Sandholm, and T. Sørensen. Potential-aware automated abstraction of sequential games, and holistic equilibrium analysis of texas hold'em poker. In *AAAI*, pages 50–57, 2007.

[9] B. V. Karl Tuyls, Pieter Jan't Hoen. An evolutionary dynamical analysis of multi-agent learning in iterated games. *AAMAS*, 12(1):115–153, 2006.

[10] D. R. Kreps and R. Wilson. Sequential equilibria. *ECONOMETRICA*, 50(4):863–894, 1982.

[11] C. Lemke and J. Howson. Equilibrium points of bimatrix games. *SIAM J APPL MATH*, 12(2):413–423, 1964.

[12] P. Miltersen and T. Sorensen. Computing sequential equilibria for two–player games. In *SODA*, pages 107–116, 2006.

[13] R. Porter, E. Nudelman, and Y. Shoham. Simple search methods for finding a Nash equilibrium. In *AAAI*, pages 664–669, 2004.

[14] T. Sandholm, A. Gilpin, and V. Conitzer. Mixed–integer programming methods for finding Nash equilibria. In *AAAI*, pages 495–501, 2005.