Expert Systems With Applications 46 (2016) 474–484

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

Convergence analysis for pure stationary strategies in repeated potential games: Nash, Lyapunov and correlated equilibria

Julio B. Clempner[a,*], Alexander S. Poznyak[b,c]

[a] Center for Economics, Management and Social Research, National Polytechnic Institute, Lauro Aguirre 120, col. Agricultura, Del. Miguel Hidalgo, 11360 Mexico City, Mexico
[b] Department of Control Automatics, Center for Research and Advanced Studies, Av. IPN 2508, Col. San Pedro Zacatenco, 07360 Mexico City, Mexico
[c] Universidad Autonoma de Ciudad del Carmen, Campeche, Mexico

ARTICLE INFO

Keywords:
Nash equilibrium
Lyapunov equilibrium
Correlated equilibrium
Convergence
Complexity analysis
Repeated Markov games
Best-reply analysis

ABSTRACT

In game theory the interaction among players obligates each player to develop a belief about the possible strategies of the other players, to choose a best-reply given those beliefs, and to look for an adjustment of the best-reply and the beliefs using a learning mechanism until they reach an equilibrium point. Usually, the behavior of an individual cost-function, when such best-reply strategies are applied, turns out to be non-monotonic and concluding that such strategies lead to some equilibrium point is a non-trivial task. Even in repeated games the convergence to a stationary equilibrium is not always guaranteed. The best-reply strategies analyzed in this paper represent the most frequent type of behavior applied in practice in problems of bounded rationality of agents considered within the *Artificial Intelligence* research area. They are naturally related with the, so-called, fixed-local-optimal actions or, in other words, with one step-ahead optimization algorithms widely used in the modern *Intelligent Systems* theory.

This paper shows that for an ergodic class of finite controllable Markov games the best-reply strategies lead necessarily to a Lyapunov/Nash equilibrium point. One of the most interesting properties of this approach is that an expedient (or absolutely expedient) behavior of an ergodic system (repeated game) can be represented by a Lyapunov-like function non-decreasing in time. We present a method for constructing a Lyapunov-like function: the Lyapunov-like function replaces the recursive mechanism with the elements of the ergodic system that model how players are likely to behave in one-shot games. To show our statement, we first propose a non-converging state-value function that fluctuates (increases and decreases) between states of the Markov game. Then, we prove that it is possible to represent that function in a recursive format using a one-step-ahead fixed-local-optimal strategy. As a result, we prove that a Lyapunov-like function can be built using the previous recursive expression for the Markov game, i.e., the resulting Lyapunov-like function is a monotonic function which can only decrease (or remain the same) over time, whatever the initial distribution of probabilities. As a result, a new concept called Lyapunov games is suggested for a class of repeated games. Lyapunov games allow to conclude during the game whether the applied strategy provides the convergence to an equilibrium point (or not). The time for constructing a Potential (Lyapunov-like) function is exponential. Our algorithm tractably computes the Nash, Lyapunov and the correlated equilibria: a Lyapunov equilibrium is a Nash equilibrium, as well it is also a correlated equilibrium. Validity of the proposed method is successfully demonstrated both theoretically and practically by a simulated experiment related to the Duel game.

1. Introduction

1.1. Brief review

An agent blindly takes directives without thinking: it is an entity capable executing particular tasks without explicit instruction. An agent is considered intelligent if it can be autonomous, flexible, and social. To behave intelligently an agent requires decision making.

* Corresponding author. Tel.: +525548775074; fax: +525548775074.
E-mail addresses: julio@clempner.name (J.B. Clempner), apoznyak@ctrl.cinvestav.mx (A.S. Poznyak).

http://dx.doi.org/10.1016/j.eswa.2015.11.006
0957-4174/

In general, artificial intelligence (AI) makes emphasis in methods related to machine learning, knowledge representation and reasoning, decision making under uncertainty, planning, and other well-studied areas. Game theory attempts to model the principles of rational interaction among players. It is the same goal of the modern AI research which is focusing on studying modern multiagent intelligent systems and how to represent intelligent behavior. In repeated games the interaction among players obligates each agent to look for an adjustment of the best-reply strategies and the beliefs using a learning mechanism until they reach an equilibrium point. The best-reply strategy approach is frequently applied, for example, in repeated games related to intelligent systems such as “Nim game” , “Three Cards two persons game” , “Red-Black card” , “Russian Roulette” , “A Pursuit game” , “Fighter-Bomber Duel” , “Simplified 2-Person Poker” , “United Nations Security Council” , “Bargaining game” , “Battle of the Sexes” (see (Jones, 1980)) and “ Duoligopolistic Market” , “Taxation and Provision of Government Service” , “Oil Price Arrangement” , “Capitalist Worker Treatment” , “Consumption Stock Pollution Equilibrium” , “Innovation Production Dilemma (R&D competition)” , “ Nonrenewable Resources: The Doomsday Problem” (see (Dockner, Jorgensen, Van Long, & Sorger, 2000)). The fundamental problem facing best-reply dynamics is that it does not predict how players arrive at an equilibrium point.

The process of finding an equilibrium point can be justified as a mathematical shortcut represented by the result of a learning algorithm (Fudenberg & Levine, 1999; Poznyak, Najim, & Gomez-Ramirez, 2000) or an evolutionary process. But, the learning or evolutionary justifications logically imply that beliefs and choices will not be consistent if players do not have time to learn or evolve.

A realization of any rational (expedient) strategy in a conflict situation (or game) is naturally related with its execution by a computer algorithm which is the heart of the Artificial Intelligence area. What is Artificial Intelligence? It is the search for a way to map intelligence into mechanical hardware and enable a structure into that intelligent system to formalize thought. Following Russell and Norvig (1995) Artificial Intelligence is the study of human intelligence and actions replicated artificially, such that the resultant bears to its design a reasonable level of rationality. The best-reply strategies analyzed in this paper represent the most frequent type of an artificial intelligence algorithm applied in practice and realized within bounded rationality.

The best-reply dynamics results in a natural implementation of the behavior of a Lyapunov-like function. The dynamics begins by choosing an arbitrary strategy profile of the players (Myerson, 1978; Nash, 1951; 1996; 2002; Selten, 1975). Then, in each step of the process some player exchanges his strategy to be the best-reply to the current strategies of the other players. A Lyapunov-like function monotonically decreases and it results in the elimination of a strictly-dominated strategy from the strategy space. As a consequence, the complexity of the problem is reduced. In the next step, are eliminated the strategies that survived the first elimination round and are not best-reply to some strategy profile, and so forth. This process ends when the best-reply (Lyapunov-like function) converges to a Lyapunov equilibrium point. Therefore, a Lyapunov game has also the benefit that it is common knowledge of the players that only best-reply are chosen. By the natural evolution of a Lyapunov-like function a strategy played once is not played again, no matter what.

The best-known solution concept of the best-reply dynamics is the Nash equilibrium (Nash, 1951; 1996; 2002), in which each player chooses a randomized strategy, and each player is not able to increase her/his expected utility by unilaterally deviating to a different strategy. The correlated equilibrium (Aumann, 1974; 1987) is an alternative solution concept. While in a Nash equilibrium players randomize independently, in a correlated equilibrium players are allowed to coordinate their behavior “based on signals” from an intermediary. Applied mathematicians, operation researchers, electrical engineers and mathematical economists have studied the computation of solution concepts since the early days of game theory (Goldberg & Papadimitriou, 2006; Govindan & Wilson, 2003; Jiang & Leyton-Brown, 2015; Jiang, Leyton-Brown, & Bhat, 2011; van der Laan, Talman, & van der Heyden, 1987; Lemke & Howson, 1964; von Neumann & Morgenstern, 1944; Papadimitriou, 2005; Papadimitriou & Roughgarden, 2008; 2011; Scarf, 1967) .

1.2. Related work

Potential games were introduced by Monderer and Shapley (1996). However, several definitions of potential games have been introduced in the literature. Voorneveld (2000) suggested the best-reply potential games allowing infinite improvement paths by imposing restrictions only on paths in which players that can improve actually deviate to a best-reply. Dubey, Haimanko, and Zapechelnyuk (2006) presented the notions of pseudo-potential games. All these classes of potential games start with an arbitrary strategy profile, and using a single real-valued function on the strategy space a player that can improve deviate to a better strategy. The iteration process converges to a Nash equilibrium point. Potential games embrace many practical application domains including dominance-solvable games, routing games and shortest-path games (Engelberg & Schapira, 2011; Fabrikant, Jaggard, & Schapira, 2013; Fabrikant & Papadimitriou, 2008). In general, all the classes of potential games reported in the literature are contained into the definition of Lyapunov games.

In this paper we show that for a ergodic class of finite controllable Markov chains games the best-reply strategies lead to one of the Lyapunov/Nash equilibrium points obligatory. As well, we show that the Lyapunov/Nash equilibrium point solution is a correlated equilibrium. This conclusion is done by the *Lyapunov Games concept* which is based on the designing of an individual Lyapunov function (related with an individual cost function) which monotonically decreases (non-increases) during the game.

In Lyapunov games (Clempner, 2006; Clempner & Poznyak, 2011; 2015) a natural existence of the equilibrium point is ensured by definition. Clempner (2015) suggested that the stability conditions and the equilibrium point properties of Cournot and Lyapunov meet in potential games. In general, convergence to an equilibrium point is also guaranteed to exist. A Lyapunov-like function monotonically decreases and converges to a Lyapunov equilibrium point tracking the state-space in a forward direction. The best-reply dynamics result in a natural implementation of the behavior of a Lyapunov-like function. As a result, a Lyapunov game has also the benefit that it is common knowledge of the players that only best-reply is chosen. In addition, a Lyapunov equilibrium point presents properties of stability that are not necessarily presented in a Nash equilibrium point.

A game is said to be stable with respect to a set of strategies if the iterated process of strategies (Guesnerie, 1996; Hofbauer & Sandholm, 2009; Pearce, 1984; Tan & Costa Da Werlang, 1988) (in our case, the best-reply dynamics) selection converges to an equilibrium point, without considering what are the initial strategies the players start with. To converge to an equilibrium point every player selects his/her strategies by optimizing his/her individual cost function looking at the available strategies of other players (Brgers, 1993; Hilas, Jansen, Potters, & Vermeulen, 2003; Osborne & Rubinstein, 1994). Any deviation from such an equilibrium point would return back to the same equilibrium point. This is because the natural evolution of the iterated process of strategies selection that tries to follow the optimal strategies and rectifies the trajectory to reach a stable equilibrium point (this is the case when the equilibrium point is unique) (Bernheim, 1984; Moulin, 1984; Osborne & Rubinstein, 1994; Pearce, 1984). In this sense, we can state that a *Lyapunov equilibrium* point is a strategy once being in the stable state of the strategies choices it is no player's interest to unilaterally change strategy. An important advantage of the Lyapunov games is that every ergodic system can be represented