

Absolutely expedient algorithms for learning Nash equilibria

V V PHANSALKAR, P S SASTRY and M A L THATHACHAR

Department of Electrical Engineering, Indian Institute of Science, Bangalore 560012, India

Dedicated to the memory of Professor K G Ramanathan

Abstract. This paper considers a multi-person discrete game with random payoffs. The distribution of the random payoff is unknown to the players and further none of the players know the strategies or the actual moves of other players. A class of absolutely expedient learning algorithms for the game based on a decentralised team of Learning Automata is presented. These algorithms correspond, in some sense, to rational behaviour on the part of the players. All stable stationary points of the algorithm are shown to be Nash equilibria for the game. It is also shown that under some additional constraints on the game, the team will always converge to a Nash equilibrium.

Keywords. Nash equilibria; Decentralised learning algorithm.

1. Introduction

This paper is concerned with a learning problem in a general multiperson stochastic game with incomplete information. We study a class of decentralised algorithms for learning Nash equilibria. For this purpose, we employ team concepts associated with Learning Automata models [13].

The game we consider is a discrete stochastic game played by N players. Each of the players has finitely many actions one of which he plays at each instant. After each play, the payoffs to individual players are random variables. The objective for each player is to maximise his expected payoff. Further, the game is one of incomplete information [6]. Thus, nothing is known regarding the distributions of the random payoffs. For learning optimal strategies, the game is played repeatedly. We are interested in (asymptotically) learning equilibrium strategies, in the sense of Nash, with respect to the expected value of the payoff. Our interest will be in decentralised learning algorithms. Hence, after each play, each of the players updates his strategy based solely on his current action or move and his payoff. None of the players has any information regarding the other players. As a matter of fact, none of the players need to even know the existence of other players. Thus the game we tackle is also of imperfect information [6].

Such games are useful in tackling problems in many areas such as decentralised control, optimisation, pattern recognition and computer vision. Some of the applications of the game model considered in this paper are discussed in [14]. In many such problems Nash equilibria, in fact, represent the desired solutions. (For a good discussion on the rationality of Nash equilibria see [4, Ch. 2]).

We use a team of learning automata [13] for evolving to the optimal strategies. Games of learning automata have been used as models for adaptive decision making