Available online at www.sciencedirect.com
SCIENCE DIRECT®

Physica A 355 (2005) 119–130

PHYSICA A

www.elsevier.com/locate/physa

Convergence of reinforcement learning to Nash equilibrium: A search-market experiment

Eric Darmon[a,*], Roger Waldeck[b]

[a]GREDEG-CNRS and Université de Nice Sophia-Antipolis, 250 Avenue Albert Einstein, F-06560 Valbonne, France
[b]GET ENST-Bretagne Department LUSSI and ICI, Technopole Brest Iroise CS 83818, F-29238 Brest Cedex 3, France

Received 31 October 2004; received in revised form 1 February 2005
Available online 2 May 2005

Abstract

Since the introduction of Reinforcement Learning (RL) in Game Theory, a growing literature is concerned with the theoretical convergence of RL-driven outcomes towards Nash equilibrium. In this paper, we apply this issue to a search-theoretic framework (posted-price market) where sellers are confronted with a population of imperfectly informed buyers and take one decision per period (posted prices) with no direct interactions between sellers. We focus on three different scenarios with varying buyers' characteristics. For each of these scenarios, we quantitatively and qualitatively test whether the learned variable (price strategy) converges to the Nash equilibrium. We also study the impact of the temperature parameter (defining the exploitation/exploration trade off) on these results.

Keywords: Reinforcement learning; Nash equilibrium; Search market; Agent-based modeling

1. Introduction

Since the seminal paper of Erev and Roth (see Ref. [1]) introducing Reinforcement Learning (RL) as an efficient modeling tool to approach human actual behavior,

*Corresponding author.
E-mail address: darmon@idefi.cnrs.fr (E. Darmon).

0378-4371/$ - see front matter
doi:10.1016/j.physa.2005.02.074

a growing literature tried to compare theoretically the properties of large multi-agent systems driven by RL to that of the Nash outcome. One key issue is to determine in which conditions these two may coincide. For example, Ref. [2] considers this issue in a congestion game analogous to a market entry game. Comparing two specifications of the RL algorithm, the author sketches two situations: in the first one, she considers one isolated RL agent that plays against the $n-1$ other agents endowed with fixed mixed strategies and so do not react on the RL-agent's decisions. In this setting, it is shown that initial conditions play a crucial role on the final strategy played by the RL agent. The case of n RL agents is then examined. The main conclusion is that this second case exhibits more rapidly a stable aggregate behavior and that the relative performance of one RL algorithm depends on the type of considered environment (endogenous evolving versus constant).

This paper provides an illustration of the convergence issue on a simple decentralized market (see Refs. [3,4] for examples of RL applications to decentralized markets). Such markets have been used to analyze price formation when the market is not ruled by an auctionner. Recently, these models have been applied to e-commerce in order to explain the persistant price dispersion despite consumers' lower search costs (see Refs. [5,6]). In this respect, Ref. [7] shows that the link between information and pricing is often misleading in the context of search theory models. Notably, better information may lead to higher price dispersion and a more intensive search by shoppers may lead to higher prices in symmetric mixed strategy equilibrium of the game. We here consider a simple posted-price market of that type: imperfectly informed buyers wish to buy an homogeneous item at the best price on one hand, and sellers try to obtain the maximum profit by setting discrete prices within a bounded range of potential prices on the other hand. There are two types of buyers: (i) uninformed (visit randomly one seller period and shop if the proposed price is less than their reservation price) and (ii) informed buyers (visit k sellers per period and buy at the firm setting the lowest of the k prices if this price is less than their reservation price). The repartition between the two types is governed by an exogenous parameter a. Buyers' behavior is then characterized by the couple (a, k). We can identify two polar cases namely Case I ("competitive setting") where there is only informed buyers and where k is equal to the number of sellers and Case II ("monopolistic competition") where there are only uninformed buyers. For Case I, the Nash equilibrium is the Bertrand competitive outcome, while in Case II, the Nash equilibrium is the monopoly outcome. For intermediate cases, Ref. [7] already established the Nash equilibrium in mixed strategies. For these three cases, we test whether the distribution of learned prices converges qualitatively and quantitatively to the Nash mixed equilibrium.

The remainder of this paper is divided as follows: Section 2 presents the simulation model and the implementation of the RL algorithm. Section 3 summarizes the results for the two polar cases (Cases I and II). Section 4 presents the results for one representative intermediate case. Section 5 concludes.