Available online at www.sciencedirect.com
SCIENCE DIRECT
ACADEMIC PRESS
Games and Economic Behavior 45 (2003) 73–96
GAMES and Economic Behavior
www.elsevier.com/locate/geb

Learning, hypothesis testing, and Nash equilibrium

Dean P. Foster [a] and H. Peyton Young [b,c,*]

[a] *Department of Statistics, Wharton School, University of Pennsylvania, Philadelphia, PA, USA*
[b] *Department of Economics, Johns Hopkins University, 3400 North Charles Street, Baltimore, MD 21218-2685, USA*
[c] *The Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA*

Received 27 March 2001

Abstract

Consider a finite stage game G that is repeated infinitely often. At each time, the players have hypotheses about their opponents' repeated game strategies. They frequently test their hypotheses against the opponents' recent actions. When a hypothesis fails a test, a new one is adopted. Play is almost rational in the sense that, at each point in time, the players' strategies are ϵ-best replies to their beliefs. We show that, at least $1 - \epsilon$ of the time t these hypothesis testing strategies constitute an ϵ-equilibrium of the repeated game from t on; in fact the strategies are close to being subgame perfect for long stretches of time. This approach solves the problem of learning to play equilibrium with no prior knowledge (even probabilistic knowledge) of the opponents' strategies or their payoffs.

JEL classification: C72; C12

Keywords: Repeated game; Nash equilibrium; Subgame perfect equilibrium; Hypothesis test

1. Statement of the problem

Consider a group of players who are engaged in a repeated game and are trying to learn the behavior of their opponents. At every time they play optimally, or almost optimally, given their beliefs. Their beliefs are generated by a *learning process*, that is, a procedure that maps past play to predictions about future play. Are there learning strategies that come close to equilibrium play of the repeated game *without* assuming any prior knowledge of the opponents' strategies or payoffs?

* Corresponding author.
E-mail address: pyoung@brook.edu (H.P. Young).

0899-8256/$ – see front matter
doi:10.1016/S0899-8256(03)00025-3

There have been several lines of attack on this learning problem, but none of them is robust in the above sense. The oldest branch of the literature is built on fictitious play (Brown, 1951). This simple learning process converges to Nash equilibrium for special classes of games, such as zero-sum games, dominance solvable games, games with strategic complementarities and diminishing returns, and potential games (Robinson, 1951; Milgrom and Roberts, 1991; Krishna, 1992; Monderer and Shapley, 1996). However, there are many examples in which it does not converge to Nash equilibrium (Shapley, 1964; Jordan, 1993; Foster and Young, 1998a). Further, even in some situations where it does converge, such as zero-sum games with only mixed equilibria, the convergence is not of the type we desire: namely, the players' forecasts are not close to the actual next round probabilities of play, nor are their strategies close to being in equilibrium.

A second branch of the literature explores conditions under which Bayesian rational players can learn to play Nash when the distribution of payoff-types is common knowledge. One answer is provided by Jordan (1991), who shows that if the players' prior beliefs about the others' strategies constitutes a sophisticated Bayesian equilibrium, then every accumulation point of the posterior beliefs is, with probability one, a Nash equilibrium in beliefs (see also Jordan, 1992, 1995). In our view this merely pushes the problem of learning an equilibrium onto another level. Moreover, these results do not solve our version of the problem because the learning process does not necessarily result in equilibrium or good prediction for the *realized* payoff-types.

A third approach to the learning problem was pioneered by Kalai and Lehrer (1993). They show that if the players' prior beliefs attach positive probability to all events that have positive probability under the players' actual strategies, then with probability one, play will eventually come ϵ-close to an ϵ-equilibrium of the repeated game. Their result would solve our problem *if* one could identify a robust procedure for constructing priors on the opponents' strategies such that the players' best-response strategies are absolutely continuous with respect to their beliefs. They do not give an example of such a rule; moreover, the subsequent literature suggests that it may be very difficult to do so (Nachbar, 1997, 1999, 2001; Miller and Sanchirico, 1997, 1999; Foster and Young, 2001).

A fourth branch of the literature focuses on learning procedures that are based on backward rather than forward-looking criteria of performance. For example, given a complete history of play, one might ask whether it *would have been* better, on average, to play x instead of y in all of those situations where one actually did play y. If no such substitution would have resulted in a higher average payoff, the history *minimizes conditional regret*. There exist a variety of quite simple learning rules that have this property. Moreover, when every player minimizes conditional regret, the empirical distribution of play converges with probability one to the set of correlated equilibria of the game (Foster and Vohra, 1997, 1998, 1999; Foster, 1999; Fudenberg and Levine, 1995, 1999a, 1999b; Hart and Mas-Colell, 2000, 2001).

Like fictitious play, these learning rules are quite simple computationally, and presume nothing about the players' prior information about the opponents' payoffs. Indeed, they are more satisfactory than fictitious play in the sense that they work for all finite games. But they are less satisfactory in the sense that they rely on backward-looking criteria of