CCIT Report #572

February 2006

Regret Minimization in Repeated Matrix Games with Variable Stage Duration

Shie Mannor
Department of Electrical and Computer Engineering, McGill University
3480 University Street, Montreal
Quebec, Canada
shie@ece.mcgill.ca

Nahum Shimkin
Department of Electrical Engineering
Technion, Israel Institute of Technology
Haifa 32000, Israel
shimkin@ee.technion.ac.il

January 25, 2006

Abstract

In a game against an arbitrary opponent, a player faces an opponent whose actions he cannot predict. Repeated games offer the opportunity for adaptive play against such an opponent, in the sense that the minimax payoff may be improved upon by reacting to the observed deviations of the opponent from a worst-case strategy. For repeated matrix games, in particular, well known results establish the existence of no-regret strategies; such strategies secure a long-term average payoff that comes close to the maximal payoff that could be obtained, in hindsight, by playing any fixed action against the observed actions of the opponent. This paper considers the extension of these ideas to repeated games with variable stage duration, where the duration of each stage of the game may depend on the actions of both players, and where the performance measure of interest is the average payoff per unit time. We start the analysis of this model by showing that no-regret strategies, in the above sense, do not exist in general. Consequently, we consider two classes of adaptive strategies, one based on Blackwell's approachability theorem and the other on calibrated forecasts, and examine their performance guarantees. In either case we show that the long-term average payoff is higher than a certain function of the empirical distribution of the opponent's actions, and in particular is strictly higher than the minimax value of the repeated game whenever that empirical distribution deviates from a minimax strategy in the stage game. Along the way, we provide sufficient conditions for existence of no-regret strategies in our model.

Keywords: no-regret strategies, regret minimization, Hannan consistency, best-response envelope, repeated matrix games, variable duration games, approachability, calibrated play.