Feedback Control for Learning in Games

Gurdal Arslan and Jeff S. Shamma*

Department of Mechanical and Aerospace Engineering

University of California–Los Angeles

37–146 Engineering IV

Los Angeles, CA 90095

{garslan,shamma}@seas.ucla.edu

This talk presents an overview of recent results that adopt a "feedback control" perspective in repeated matrix games. In particular, it is shown how the introduction of feedback control principles can result in convergence to a mixed strategy Nash equilibrium even for previously nonconvergent examples.

We consider the usual set-up in repeated matrix games, e.g., (Fudenberg & Levine, 1998), in which multiple decision-making players adjust their strategies according to observations of each other's actions. The game is noncooperative in that each player may have its own objective/utility function, and these objectives are not shared among players.

One challenge in such systems from a "learning" perspective is that the environment, as seen through the eyes of a single player, consists of other evolving players, and as such, is non-stationary. This non-stationarity limits the degree to which the environment can be learned through repeated interactions, especially since the drivers behind the changing environment—other player utility functions—are unknown.

Many strategy adaptation methods in this setting have been analyzed and a variety of convergence and non-convergence results have been obtained. In particular, the recent work (Hart & Mas-Colell, 2003) analyzes a three player game, namely the Jordan anti-coordination game (Jordan, 1993), where opponent utility functions are unknown, but player actions can be observed. The paper shows that a large class of strategy adaptation methods cannot converge to a Nash equilibrium. This class of strategies includes a variety of well known adaptation and learning methods, such as fictitious play and replicator dynamics. Similar non-convergence results may be traced back to works such as Crawford (1985).

Interacting players collectively form a *feedback* system, i.e., the influence an player has on other players is ultimately reflected back to the originating player. This viewpoint leads to a strong connection to the concept of feedback control. Feedback control has a long tradition in the theory and practice of engineered systems (Basar, 2000). Feedback control theory is concerned with the derivation of decision making policies to enhance a specified performance metric for a given dynamic system using limited information. The term "feedback" reflects that decision making policies are based on actual behavior, which is "fed back" to the controller (decision maker) and compared to desired behavior.

Our recent work (Shamma & Arslan, 2003a; Shamma & Arslan, 2003b; Arslan & Shamma, 2004) has shown that the introduction of feedback control methods can overcome obstacles to convergence to Nash equilibrium. The results apply to both the Jordan game (Jordan, 1993) and Shapley game (Shapley, 1964).

The particular notion used from feedback control models is "derivative action". Derivative action is a very commonplace tool in the design of feedback controllers, and a discussion may be found in any undergraduate textbook. A typical feedback controller responds to observations of an "error signal", i.e., a discrepancy between desired behavior and actual measured (feedback) behavior. In derivative action, the error signal and its derivative are used to formulate a controller response. The usual interpretation is that derivative action introduces an anticipatory aspect to a controller's response.

The implementation to learning in games requires that a player uses available observations *and their derivatives* in adaptively updating player strategies. Indeed, similar notions have been explored in the learning in games literature for two-player zero-sum games, cf., (Conlisk, 1993).

*Corresponding author.

This talk will present an overview of the feedback control approach to learning in games. Points to be covered include:

- An brief introduction to feedback control and its relevance to learning in games.
- "Derivative action" versions of fictitious play and gradient play in continuous-time, and associated analytical convergence results for pure and mixed strategy Nash equilbria.
- Discrete-time implementation of derivative action methods, and convergence in probability results using the ODE method of stochastic approximation (Benaim & Hirsch, 1996).
- Extensions to "utility measurement" scenarios, where player actions are not observable, but only individual stage-by-stage rewards are available.

References

Arslan, G., & Shamma, J.S. 2004. Distributed convergence to Nash equilibria with local utility measurements. *In: 43rd IEEE Conference on Decision and Control.* submitted for publication.

Basar, T. (ed). 2000. *Control Theory: Twenty-Five Seminal Papers (1932–1981).* Wiley.

Benaim, M., & Hirsch, M.W. 1996. A dynamical systems approach to stochastic approximation. *SIAM Journal on Control and Optimization*, **34**, 437–472.

Conlisk, J. 1993. Adaptation in games: Two solutions to the Crawford puzzle. *Journal of Economic Behavior and Organization*, **22**, 25–50.

Fudenberg, D., & Levine, D.K. 1998. *The Theory of Learning in Games.* Cambridge, MA: MIT Press.

Hart, S., & Mas-Colell, A. 2003. Uncoupled dynamics do not lead to Nash equilibrium. *American Economic Review*, **93**(5), 1830–1836.

Jordan, J. 1993. Three problems in game theory. *Games and Economic Behavior*, **5**, 368–386.

Shamma, J.S., & Arslan, G. 2003a. *Dynamic fictitious play, dynamic gradient play, and distributed convergence to Nash equilibria.* submitted to *IEEE Transactions on Automatic Control*, http://www.seas.ucla.edu/~shamma/papers.html.

Shamma, J.S., & Arslan, G. 2003b. A feedback stabilization approach to fictitious play. *Pages 4140–4145 of: Proceedings of the 42nd IEEE Conference on Decision and Control.*

Shapley, L.S. 1964. Some topics in two-person games. *Pages 1–29 of:* M. Dresher, L.S. Shapley, & Tucker, A.W. (eds), *Advances in Game Theory.* Princeton, NJ: University Press.