Nash equilibrium selection by convention

Rasmus Bokrantz, 810819-7919, bokrantz@kth.se

May 19, 2010

Introduction

Social interaction between individuals can be modeled as a repeated game where players are selected at random from a large population at the beginning of each period. A social convention can in this setting be viewed as a collective mechanism for Nash equilibrium selection: the play is said to follow a convention if each player has an expected and customary behaviour which is a best-reply under the belief that all opponents also play according to their customary behavior. Any pure strategy Nash equilibrium is thus a candidate for becoming a convention. This raises the question of how one Nash equlibirum may become prevailing at the cost of the other equilibria in the game. In this report, we will examine a model proposed by Young [1] which adresses this questions and provides a mechanism for how a Nash equilibrium is established as a convention.

The remainder of this report is organized as follows. We will first introduce Young's model of so called *adaptive play* and the class of *weakly acyclic* games to which it is applied. We will then turn to outline Young's main results for two situations: when players never makes mistakes and when the probability of a tremble to a nonoptimal strategy is positive. As it turns out, these two cases yield remarkably different results. We will also examine the special implications of Young's model for 2×2 games. Finally, we will consider a modification of Young's model suggested by Hurkens [2]. Despite being very similar on a technical level, the implications of Hurkens' approach are qualitatively different to those of Young's. In the interest of space, the presentation is kept informal throughout.

Adaptive play

Adaptive play is defined as follows. Consider a repeated n-player where at the beginning of each period, players are drawn at random from n nonempty classes to play each of the n roles in the game. All members of each class are assumed to have identical utility functions and finite set of available strategies. Before choosing strategy, each player forms a belief about the opponents' strategies by observing k plays drawn at random without replacement from the m most recent games, where $1 \leq k \leq m$. The probability that each of the m most recent periods is observed is assumed to be positive, but need not be equal between periods. Each player then chooses strategy maximizing his expected payoff under the resulting subjective belief.

Weakly acyclic games

The dynamics of adaptive play constitutes a finite Markov chain whose states are the histories of play in the m most recent periods. A *convention* is defined as an absorbing state, i.e., a state for which the transition probability to any other state is zero. It is easy to see that necessary and sufficient conditions for a state to be absorbing is that a strict pure strategy Nash equilibrium have been played for m successive periods. Existence of a such an equilibrium is however not sufficient for ensuring that the game will be absorbed in a convention. If there exists cycles of best replies which avoids the Nash equilibrium components, and if these in addition are not present in the initial sequence of m plays, adaptive play may fail to converge. Young constructs a two-player example with a best-reply cycle of length twelve to show that this indeed can occur. To circumvent this limitation, Young restricts his analysis to the class of *weakly acyclic* games: games that cannot be trapped in a best-reply cycle. Necessary and sufficient conditions for a game to be weakly acyclic is that from all possible states, there exists a finite sequence of best replies that ends in pure strategy Nash equilibrium.

Adaptive play without mistakes

Assuming a weakly acyclic game where all players always best-reply, Young states that convergence to a convention is obtained in a finite number of steps from any initial state if the sampling of previous games is sufficiently incomplete. This result is proved by establishing a positive integer M and a positive probability p such that from any state, the probability that adaptive play converges to a convention in at most M periods is at least p. It then follows the probability of reaching a convention after at least rM periods is at least $1-(1-p)^r$, which tends to one as r tends to infinity.

To see that this result does not hold when $k=m$, consider a "Battle of the Sexes" game on the form

	Yield	Not Yield
Yield	0, 0	1, $\sqrt{2}$
Not Yield	$\sqrt{2}$, 1	0, 0

where (Yield, Not Yield) and (Not Yield, Yield) are the two plausible candidates for becoming a convention. If $k=m$, the best-reply for each player is determined on basis of identical sets of historic plays. By symmetry of the game, the strategies chosen by each player in each period of the game will be perfectly correlated. Since the highest possible payoff for each player is an irrational number and the frequency of a strategy profile rational, there can never be a tie which renders a player indifferent with respect to his choice of strategy. Thus, if the game is initiated from a sequence of m games where both plays have failed to coordinate, they will miscoordinate forever.

Adaptive play with mistakes

Now consider the situation where the probability that a player by mistake deviates to a nonoptimal strategy is positive. Convergence to an absorbing state now becomes impossible, as mistakes constantly perturb the process away from the equilibrium. Instead, Young studies the asymptotic behavior in the limit when the probability of mistakes approaches zero. The states tha have positive probability in this limit are denoted *stochastically stable*. Young constructs an algorithm for computing these states, and then proceeds to prove that every stochastically stable state is a convention. Thus, when the probability of mistakes is small but nonvanishing, the set of possible conventions is restricted to the set of stochastically stable equilibrium. This should be contrasted to the previous results that any pure strategy Nash equilibrium constitutes a candidate for a convention when the probability of mistakes is identically zero.

Implications for 2×2 games

Young model holds special implications for 2×2 games. Consider adaptive play of a "Stag-hunt" game on the form

	Stag	Hare
Stag	4, 4	0, 3
Hare	3, 0	2, 2

This game has two pure strategy Nash equilibria: the payoff dominant equilibrium (Stag, Stag) and the risk dominant equilibrium (Hare, Hare)[1]. In the case that players make no mistakes, Young's analysis gives that the game will settle in a play of either (Stag, Stag) or (Hare, Hare) in a finite number of steps. Which of the two equilibrium that becomes established as a convention is determined by the frequency of (Hare) and (Stag) plays in the historic plays sampled at the beginning of each period. An increase in the number of (Stag) plays in the initial

[1] A Nash equilibrium is considered payoff dominant if it Pareto dominates the other equilibria in the game, and risk dominant if it has the greates product between the players deviation losses.

sequence of m plays will thus lead to an increase of the probability that the outcome of the game is (Stag, Stag). We can thus conclude that the outcome of the game is history dependent.

Now assume that each player having decided to play best-reply by mistakes deviates to the other pure strategy with positive probability. It then follows by Young's analysis that the game will settle in a stochastically stable convention. Since (Hare, Hare) is risk dominant, both players are more likely to play this equilibrium when the uncertainty about the opponents' strategy increases. This implies that the number of mistakes needed to disrupt the game from (Stag, Stag) to (Hare, Hare) is smaller than the number of mistakes needed for the converse transition. As the probability of mistakes tends to zero, the games will thus settle in the risk dominant (Hare, Hare) equilibrium, irrespective of the play in the initial sequence of m periods. The outcome of the game is thus said to be ergodic, as opposed to history dependent. Young formalizes this result by showing that for any 2×2 game with two strict Nash equilibria in pure strategies, only weakly risk dominant equilibria are stochastically stable. This result does not extend to more general classes of games.

Generalization by Hurkens

In [2], a model of learning closely related to Young's model is proposed. The key difference between Hurkens' and Young's analysis is that sampling of previous plays occurs with replacement in Hurkens' model, as opposed to without replacement in Young's model. This has the implication that a single mistake can have larger impact on the outcome of the game, as this mistake may be sampled multiple times. With respect to the Stag-hunt example above, a single mistake may thus be sufficient to move the system from (Stag, Stag) to (Hare, Hare) or vice versa. Since the probability of a transition from one equilibrium to another is now of the same order, both equilibrium will have positive probability in the limit when the mistake probability goes to zero. The connection between risk dominance and stochastic stability is thus not valid in Hurkens' framework. Another difference to the model of Young is that Hurkens makes no restriction to the class of weakly acyclic games. Instead, Hurkens proves that the play will be absorbed in a so called *minimal curb set*: a set of strategy profiles that are closed under rational behavior, i.e., it contains all its best replies, that contains no strict subsets which also possess this property. This result is not in contradiction with Young's result as a pure strategy Nash equilibrium is a minimal curb set as a singleton.

Conclusion

The general theme that carries over from Young's work to Hurkens' is the notion that conventions are created through positive feedback during the course of the game. A Nash equilibrium thus evolves into a convention through the dynamics of the game, rather than being selected for on basis of some intrinsic properties.

References

[1] Young, H.P. (1993) The evolution of conventions, *Econometrica*, **61:1**, 57–84.

[2] Hurkens, S. (1995) Learning by forgetful players, *Games and economic behaviour*, **11**, 304–329.