

The following is submitted as a proposal for presentation at the Economic Science Association meetings, June 19 – 22, 2003, in Pittsburgh, PA.

Author:

Randall E. Waldron
Associate Professor of Economics
University of South Dakota

Contact Information:

Mailing Address:

School of Business
414 E. Clark St.
Vermillion, SD 57069

Email: rwaldron@usd.edu

Phone: (605) 677-6643

Fax: (605) 677-5427

Paper Title:

**Learning Optimal Strategies in Rapidly Repeated Games:
Lessons from “Rock/Paper/Scissors”**

Extended Abstract:

A fundamental proposition of elementary non-cooperative game theory is that rational players will choose strategies consistent with Nash equilibria. These strategies may require randomization, or the mixing of strategies. In this paper I present the results of a series of experiments designed to investigate whether players do in fact play Nash strategies in simple but non-trivial finitely repeated games with a unique equilibrium in mixed strategies. The games all involve variations on the familiar childhood game, “Rock/Paper/Scissors.” The variations include scoring changes to alter the strategies from the traditional game, incentives which create zero-sum and non-zero-sum games, human-versus-human competition, and human-versus-computer competition. In particular, I investigate whether players move towards the Nash strategies as the games progress for each variation.

A series of games was held with incentives to promote optimal play, and a number of college students with no specific training in game theory were invited to participate. One group, serving as a control, was asked to play the traditional version of the game, scoring one point for each win. Another group was asked to play a more complex version of the game, where winning with "rock" yielded more points than winning with "scissors," which in turn was worth more than winning with "paper." In terms of the Nash equilibrium, this variation altered the optimal mix of strategies. Players were paired and completed 50 rounds of the game in rapid succession. Rewards were based on the difference between the two players' scores within each pair, effectively making the game zero-sum. Later games were extended to 200 rounds to test for learning over a longer interval.

Analysis of the traditional game's results suggests that the control group did in fact play strategies consistent with Nash equilibrium probabilities, although the evidence suggests that players switch actions too often to be consistent with independence among rounds. However, in the more complex game, it appears that players began playing strategies more consistent with the traditional game's equilibrium than the complex game's equilibrium. However, as play

proceeded, players moved away from Nash equilibrium strategies for the simple game. It was not possible, however, to conclude that participants were playing optimal strategies by the end of the game, raising the possibility that learning did not occur quickly, or that some players were able to exploit the poor play of the rival.

Human-versus-human competition proved difficult to evaluate because of the challenge of distinguishing the exploitation of poor play from poor play itself: neither appears consistent with the null hypothesis, which is the prescription of the Nash concept. To counter this difficulty, additional experiments were designed and conducted using a spreadsheet-based computer program to control one player. The program was designed to play randomly when it appeared the human was playing randomly with the prescribed equilibrium strategies (computed using a chi-square test), and to play “intelligently” otherwise by calculating its expected payoffs based on what the human had done in preceding rounds, looking back as far as 50 rounds and using “pair-wise” play, or in other words evaluating how often the human followed each particular action with each available action. Thus the computer program was intended to exploit both the use of wrong probabilities by the human as well as simple patterns played by the human, even if they resulted in apparent equilibrium probabilities.

Two types of human-versus-computer experiments were conducted, one for the zero-sum form of the game and the other for a non-zero-sum variation in which the subject was rewarded based solely on the number of points accumulated without regard to the computer’s score. Although the scoring descriptions of these two games appear similar (e.g., 3 points awarded for a win with rock, 2 for a win with scissors, 1 for a win with paper), the reward mechanism alters the equilibrium substantially. The purpose of the two experiments was in part to determine if participants could distinguish the implications of the different rewards systems and to evaluate learning over 200 rounds in each game.

Preliminary evidence from the computer based experiments suggests that players did not initially play Nash equilibrium strategies, but did move closer to them over the course of the game. The full analysis of the results will be provided in the paper and presented at the conference. However, it is also apparent from the preliminary results that players were often able to defeat the “intelligent” play of the computer in the zero sum variation, even without playing Nash strategies. While a full analysis of the results is pending, it appears that the “adaptive expectations” formulated by the computer in an attempt to exploit the poor play of the human regularly proved futile. Human subjects, while failing to adopt Nash strategies, also avoided the simple patterns and consistently poor play that would have been easy to exploit, and in fact these subjects were themselves able to exploit the computer’s slow learning as it deviated from purely random actions. Thus, an interesting insight and avenue for further exploration is the challenge of designing mechanisms to exploit non-optimal play.