Common Rationality and the Choice Among Solutions:
an experimental analysis of the Centipede Game

In the two decades since its first appearance, Rosenthal's (1982) centipede game has become a touchstone of the theory of perfect information games. Game theorist Robert Aumann claims that almost every paper on the subject mentions the Centipede game, and in many papers it is the chief object of analysis (1998: 98). The generic N-legged Centipede Game is a two-player game with perfect information and N moves that alternate between players. The Centipede structure, seen in the figure below for players A and B, comes from the characteristic that the set composed of each information set and all its predecessors coincides with the set of decision nodes of the game. The game has a unique subgame-perfect equilibrium in which both players choose T (Take) at every decision node rather than P (Pass). The iterated deletion of weakly dominated strategies, or backwards induction, reduces the game to the subgame-perfect outcome, which is a Pareto inferior outcome. The backwards induction procedure is one of the fundamental techniques of game theory.

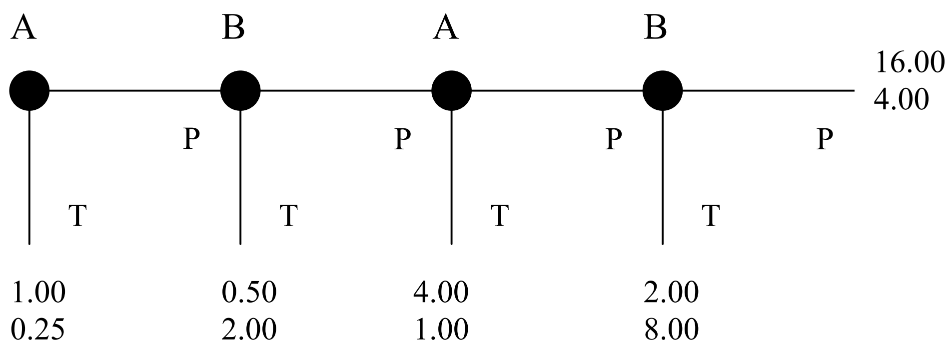

Figure 1 - The Four-move Centipede Game

Although there have been many theoretical analyses of the Centipede game (e.g., Rosenthal 1982; Francesconi 1991; Aumann 1992, 1996, 1998; Binmore 1988, 1996, 1997; Dulleck and Oechssler 1997; Ewerhart 1997; Cressman and Schlag 1998; Droste et al. 1998; Levine 1998; Rabinowicz 1998; Zauner 1999; Spiegler 1999; Battigalli and Bonanno 1999; Ponti 2000), there have been only three empirical studies of human behavior in controlled Centipede experiments. McKelvey and Palfrey (1992) experimented with the game pictured in Figure 1 and Fey et al. (1996) experimented with a constant-sum version of the same game. Nagel and Tang (1998) experimented with a normal-form version of the game in Figure 1 (players simultaneously choose actions for every node, rather than sequentially). In these three empirical studies, observed behaviors were quite different from the theoretical predictions. Subjects rarely choose Take at the first node.

A variety of explanations have been brought forth to explain these empirical observations. McKelvey and Palfrey (1992) argued that the observed patterns of behavior are consistent with both errors in the strategy choices of players and the presence of altruistic "souls of cooperation" who choose Pass rather than Take at every node. Altruistic individuals Pass at the first node as do other subjects who believe there is a sufficiently high positive probability that they may be paired with such an altruist. Individuals learn over time and thus reduce the errors they make and reduce the uncertainty about the proportion of altruists in the population. Zauner (1999) also tries to explain the McKelvey and Palfrey data in terms of an equilibrium model with errors. In his model, the doubts players have about aspects of the game are modeled as random disturbances to each player's payoffs. Zauner considers five different models, each differing in the specification of the error terms, and then selects two "best" models according to a likelihood criterion. Nagel and Tang, however, find that learning is not necessarily related to convergence to equilibrium as predicted by the equilibrium models with error. Their normal-form design with switching partners also suggests that the observed behaviors are not a result of strategies attempting to induce reciprocity across nodes or sequential plays of the game. Others authors, using the Palfrey and McKelvey data, have argued that the data are better explained by a combination of altruistic and spiteful behavior (Levine 1998) or by subject error and the beliefs that one may be paired with an "absent-minded" subject prone to make errors (Dulleck and Oechssler 1997). The empirical analyses to date have not only been unable to elucidate the underlying motivations of individuals who make choices inconsistent with standard theory, they have also not addressed the question of whether or not subjects who do act according to theory (i.e., choose Take at the first node) do so after applying the iterated deletion of weakly dominated strategies.

In order to contribute to our collective understanding of the way in which individuals make decisions in finite dynamic games, I ran a series of extensive-form Centipede experiments in which human subjects play (1) with other humans (i.e., replication of McKelvey and Palfrey), (2) with non-human players that choose Take and Pass with the same frequency as human subjects, and (3) with non-human players that play like the quintessential rational agents from game theory text books. Treatment (3) was run with an additional high-payoff session. Human subjects knew if they were paired with a human or nonhuman player and, if the opposing player was non-human, they were informed of the way in which the opposing player made its decisions. Expectations of the probability that the opposing player would Take or Pass at a given node were also elicited with monetary payoffs for accurate predictions. Subjects were also asked to explain the way in which they made their decision. In the treatment in which humans played with the quintessential rational agent from game theory text books, subjects were also rewarded for correctly explaining, immediately after the strategy of the non-human player (Take at every node) was revealed, the way in which the virtual player arrived at its strategy.

I find that:

(1) consistent with previous results only about 20% of subjects chose Take at the first node and there was no significant difference in the proportion of subjects who chose Take in any of the sessions except in the high-payoff, rational non-human agent session (see below);

(2) the majority of subjects who chose Take at the first node arrived at their decision through the iterated deletion of weakly dominated strategies;

(3) altruism does not explain the high proportion of individuals choosing Pass at the first node;

(4) error does explain the choices of some of the individuals who chose Pass at the first node but not the majority;

(5) subjects' expectations of the frequency with which other humans would Take or Pass in the Centipede game were correct on average;

(6) the majority of subjects playing against a non-human, unerring rational agent correctly identified the backwards induction solution to the game but also identified the cooperative solution;

(7) the majority of subjects who chose Pass at the first node when playing against a non-human, unerring rational agent were able to articulate the backwards induction argument. Most argued, however, that they could not be sure that the non-human agent would forgo the cooperative solution for the subgame-perfect solution (almost 80% of the subjects believed there was a positive probability that the non-human agent would pass, including a couple of subjects who chose Take at the first node); and

(8) when the payoffs of the low-payoff game with rational non-human players were multiplied by ten, we observed no significant difference in the proportion of subjects correctly identifying the subgame-perfect solution to the game, but the number of subjects choosing Take at the first node was more than twice that observed when humans played with other humans (44%).

Our results taken together suggest that human subjects can identify game-theoretic solutions in simple games, but they also can recognize when such solutions are Pareto inferior. When the opportunity costs of attempting to achieve a Pareto-superior outcome are low, most subjects will, in the words of one subject, "go for it" even when the probability of the other player behaving in a similar way is small.

References

Aumann, R. J. 1992 . ‘‘Irrationality in Game Theory,’’ in Economic Analysis of Markets and Games: Essays in Honor of Frank Hahn. P. Dasgupta, D. Gale, O. Hart, and E. Maskin, Eds. Cambridge, MA: MIT Press, pp. 214-227.

Aumann, R. J. 1996 . ‘‘Reply to Binmore,’’ Games and Economic Behavior **17**: 138-146.

Aumann, R.J. 1998. On the Centipede Game. Games and Economic Behavior 23: 97-105.

Battigalli, P. and G. Bonanno. 1999. Recent Results on Belief, Knowledge and the Epistemic Foundations of Game Theory. Research in Economics 53: 149-225.

Binmore, K. 1988. Modeling rational players: I, II. Economics and Philosophy, 3: 179-214; 4, 9-55.

Binmore, K. 1996 . ‘A Note on Backward Induction. Games and Economic Behavior **17**: 135-137.

Binmore, K. 1997. Rationality and Backward Induction. Journal of Economic Methodology 4(1): 23-41.

Cressman, R. and K. H. Schlag. 1998. The Dynamic (In)Stability of Backwards Induction. Journal of Economic Theory 83: 260-285.

Droste, E., M. Kosfeld, and M. Voorneveld. 1998. Regret Equilibria in Games. Tilburg CentER for Economic Research Discussion Paper: 9819.

Dulleck, U. and J. Oechssler. 1997. The Absent-Minded Centipede. Economics Letters 55(3): 309-15.

Ewerhart, C. 1997. Undecidable Statements in Game Theory. Universitat Bonn Sonderforschungsbereich 303 Discussion Paper: A553.

Fey, M., R.D. McKelvey, and T.R. Palfrey. 1996. An Experimental Study of Constant-Sum Centipede Games. International Journal of Game Theory 25(3): 269-87.

Levine, D.K. 1998. Modeling Altruism and Spitefulness in Experiments. Review of Economic Dynamics 1(3):

Mckelvey, R. D., and T.R. Palfrey. 1992. An experimental study of the centipede game. Econometrica 60: 803-836.

Nagel, R. and F.F. Tang. 1998. Experimental Results on the Centipede Game in Normal Form: an investigation on learning. Journal of Mathematical Psychology 42: 356-384.

Ponti, G. 2000. Cycles of Learning in the Centipede Game. Games and Economic Behavior 20: 115-141.

Rabinowicz, W. 1998. Grappling with the Centipede: Defence of Backward Induction for Bi-terminating Games. Economics and Philosophy 14(1): 95-126.

Rosenthal, R. 1982 . ‘‘Games of Perfect Information, Predatory Pricing, and the Chain Store Paradox,’’ Journal of Economic Theory **25**, 92-100.

Spiegler, R. 1999. Reason-Based Choice and Justifiability in Extensive Form Games. Tel Aviv Foerder Institute for Economic Research and Sackler Institute for Economic Research Working Paper: 19/99.

Zauner, K.G. 1999. A Payoff Uncertainty Explanation of Results in Experimental Centipede Games. Games and Economic Behavior 26: 157–185.