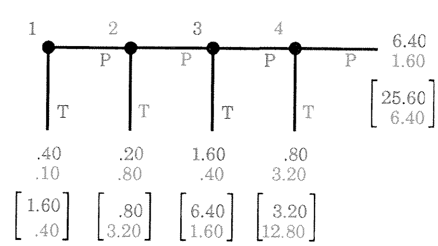
Lab Notes

Table stakes: Rita's and Bruce's possible payoffs during the four-move centipede game. Each turn is symbolized by a numbered dot, and the lines labeled P lead to the next turn. The lines labeled T give the payoffs at any turn if a player takes the money. The top row of numbers (in red) is Rita's pot, and the bottom row (in blue) is Bruce's. (The numbers in brackets are for the high-stakes version.) If both players pass for the entire game, they get the payoffs shown in the rightmost column.

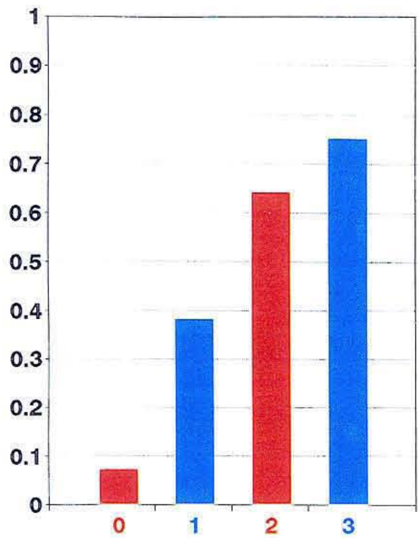


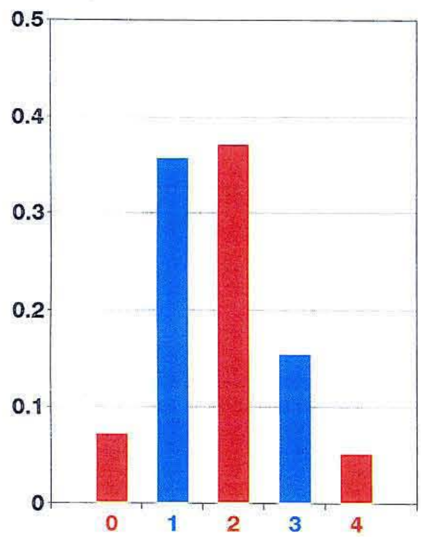
Do Unto Others Before They Do Unto You

"The best judge in a beauty contest is the one who always picks the average score," according to Thomas Palfrey (PhD '81), professor of economics and political science. In other words, if a particular contestant winds up with a score of 8.7 after all the judges' votes are tallied, then the judge who actually wrote that score on his or her own scorecard had the best feel for how the contestants would rank. A game called the centipede is another exercise in second-guessing, and it works like this: Rita and Bruce start with two pots of money, one of which is four times bigger than the other. Rita moves first. If she takes the money, the game's over—she gets the big pot and Bruce gets the small one. If she passes, each pot doubles. Now Bruce gets the opportunity to take the larger one, and so on. If the game survives two such innings—four passes in all—it ends anyway and Rita gets the big bucks. (The original version ran for 100 passes, hence the name.)

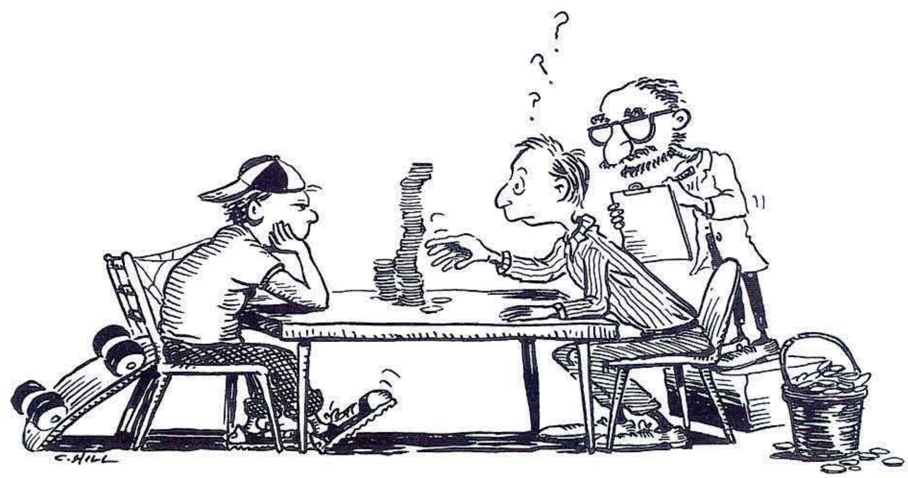
Rita and Bruce both know that the game will end after four passes, and both know how big each pot will be at every step. Game theorists call this a game of "perfect information," since both players can see all the way to the end and plan accordingly. The winning strategy, says game theory, is simplicity itself—take the money at your very first opportunity. Assuming that the small pot began with a dime and the large one with forty cents, as shown in the chart at left, the logic runs as follows: If Bruce passes on turn four, he knows that Rita will get $6.40 and he'll get $1.60; if he takes the money, he'll get $3.20 and leave Rita 80 cents. Thus Bruce should take the money. But Rita knows this, too, so therefore she should freeze Bruce out and take the money on turn three, awarding herself $1.60 and Bruce a lousy 40 cents. And Bruce knows that Rita knows, so he should preemptively grab the dough on turn two, winning 80 cents and sticking Rita with 20 cents. And finally, Rita can see that Bruce will stiff her if she passes, so she should take the 40 cents offered her on the very first turn.

But that's a pretty low-reward strategy, and it's not what I'd do. It apparently isn't what almost anybody else would do, either. Professor of Political Science Richard McKelvey, Palfrey, and graduate student Mark Fey (BS '90) have





As the game progresses, players are more likely to run with the money. The top graph shows the probability (vertical axis, with "1" being certainty) of a player taking the pot after the number of passes shown on the horizontal axis. (In other words, zero passes is Red's first turn.) The bottom graph shows the relative frequency (vertical axis) with which games ended after the number of passes shown on the horizontal axis. Thus, most games ended after one or two passes.

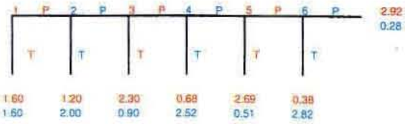


studied the centipede, and only one of their 138 experimental subjects ran with the loot at every opportunity. In fact, most people passed on their first move. From then on, the probability of a player taking increased with every turn. Even so, a significant number of games went the full four moves, and nine subjects even passed on turn four!

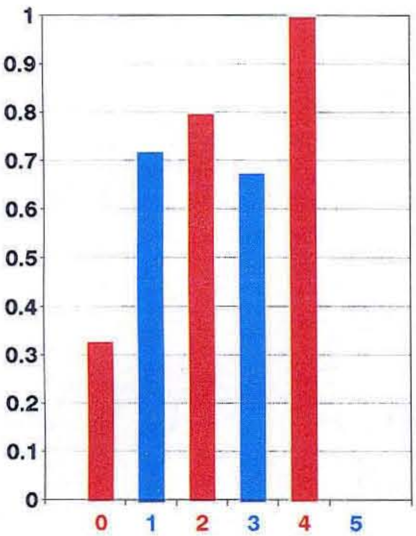
The games were played in Caltech's Laboratory for Experimental Economics and Political Science, by Caltech and Pasadena City College undergrads. The players communicated through a network of personal computers that also recorded their moves and calculated their winnings. At evening's end, the participants got paid—in real cash. Starting with 40 cents in the pot doesn't imply trivial stakes—one player walked away with $75.00 for less than an hour's work. (If this person had known game theory, he or she would only have netted $7.00—40 cents times 10 games, plus $3.00 for showing up.) The players were designated as either Red (the first mover) or Blue at the beginning of the session, and kept their color for the duration. In order to prevent anybody from capitalizing on what they learned about an opponent, each Red played exactly one match with every Blue, and vice versa, and no player participated in more than one of the seven sessions. Some sessions played a six-move centipede, or had a more generous scale of payoffs—variations designed to encourage greed—but 10 to 20 percent of these games still went the distance.

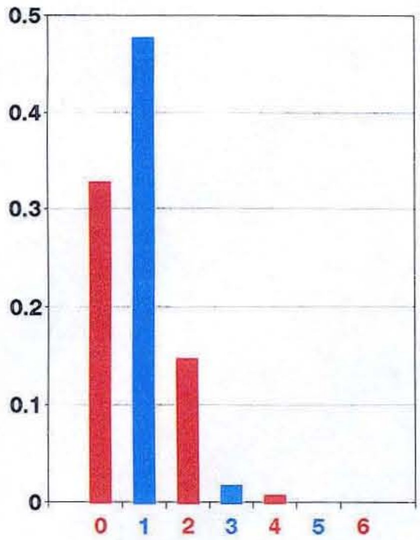
The remaining 128 subjects showed a spread of behavior between the one grabby guy and the nine passive people. Most participants appeared to be learning on the job. The later games in a session tended to be over more quickly, as people who'd been burned before pounced sooner. But some individuals appeared to be playing haphazardly, with no clear pattern emerging from their behavior over the course of a session.

Since game theory's prediction was a colossal failure, some other factor was obviously at work. The theory assumes that humanity has the predatory instincts of a leveraged-buyout artist—one maximizes one's own rewards and the heck with the other guy. This is called the "rational" strategy. But someone whose moral development has progressed beyond that of the shark (or who "just wants to bankrupt the Social Sciences department," as Palfrey remarks) would realize that by passing, both players are better off because both pots get bigger. And if these people exist (even as an endangered species) an intelligent opportunist would realize that the way to beat the game is to make like an altruist and pass in the first inning, then



Above: The six-move constant-sum centipede game's payoff structure.
Below: Conditional probabilities (top) and frequencies (bottom) for the six-move, constant-sum game ending, again plotted versus the number of passes. This data is from PCC students—Techers "took" so quickly that there are very few late-round data points for them.





move in for the kill and claim a twice-doubled pot in the second. Thus, if you believe your opponent may be altruistic, it's in your selfish best interest to pretend that you are too, at least for a little while. This will even work if your opponent is selfish—if, by passing, you can make the other person believe that *you* are the altruist, that person may decide to play you for a sucker and pass the pot back once, in order to burn you on the following turn.

McKelvey and Palfrey developed a mathematical model for the game in which they assumed that a small percentage of the players were altruists. The rest were not, but knew that there was a smattering of angels in their number and played accordingly. The model also included a random-error function to mimic the small probability that a player might hit the wrong key, forget what turn it was, or otherwise mess up. Computer simulations based on this model, and run on the Caltech/JPL Cray X-MP supercomputer, plotted the probability of the game ending in a "take" at any given turn. The predictions agreed with the actual games very nicely.

Unfortunately, this explanation didn't survive a second set of experiments designed to test it. This time, the two pots started out equal, and if Red passed, one-fourth of the money in one pot was moved to the other. After each succeeding pass, the larger pot absorbed one-fourth of the smaller one, as in the chart at left. Here, the logic of the game dictates that saints as well as swine should take the money on the first move. The combined pots don't grow, and the initial 50-50 distribution is certainly the most equitable outcome. But again, over half of the games played went beyond the first move. The rate at which people chose to take thereafter, however, increased very rapidly, and all the games ended early. Nine sessions of this game (including three at the University of Iowa) were played, using six- and ten-move centipedes in order to let the smaller pot *really* dwindle.

With altruism joining capitalism in the dustbin of history, how can one construct a model in which people often pass, even when it doesn't seem to be in their selfish best interest to do so? Fey, McKelvey, and Palfrey have come up with one, dubbed "quantal response." (The term "quantal" comes to the social sciences by way of biology, where it describes a yes-or-no, all-or-nothing response—a skin test for allergies, for instance. Such models are widely applied to discrete-choice situations—what kind of car to buy, for example—but had not been combined with experimental game theory before.) Folks make mistakes, says quantal-response equilibrium, but folks know that everyone else makes mistakes too. Everyone intends to take the money, but every so often, someone does something dumb, like the "Wheel of Fortune" contestants who pick letters that have already been used. (Of course, the more costly a mistake—passing in a late inning, in this case—the less likely someone is to make it.) And since some people are more error-prone than others, a player may elect to pass at first, in hopes of having drawn an inept opponent. But as the session progresses, the odds of passing drop—the klutzes begin to wise up, so there are fewer of them to prey on.

When this model was run, it agreed with both sets of experiments. "You don't have to hypothesize altruists or other extraneous factors," says Palfrey. "Different levels of skill and a bit of noise will give the same results." These models are of more than academic interest—speculative bubbles, such as occurred in real estate in the 1980s, are real-world examples of the centipede game. If all investors behaved with the perfect rationality of game theory, they'd anticipate that the bubble would eventually burst. The bubble would never grow in the first place—nobody would buy in, for fear of being left holding the bag. In real life, of course, the winners are the ones who guess best what the average person is guessing, and bail out just before everyone else does. "A lot of game theory has been built on introspection by some very smart people, but introspection only gets you so far. These sorts of adjustments to make the model more realistic wouldn't happen if there weren't experiments. Now we know that a little bit of error goes a long way." □ —*DS*