Lost in a jungle? Survival may depend on copying others or trial-and-error learning.

CULTURAL EVOLUTION

Conquering by Copying

A computer tournament reveals the benefit of copying someone else's actions over solving a problem solo, a finding that has implications for cultural evolution

Suppose you find yourself in an unfamiliar environment where you don't know how to get food, avoid predators, or travel from A to B. Would you invest time working out what to do on your own, or observe other individuals and copy them? If you copy, who would you copy? The first individual you see? The most common behaviour? Do you always copy, or do so selectively?

What would you do?

With those provocative words, posted on a Web site and included in a flyer sent to colleagues and academic departments around the world in late 2007, Kevin Laland threw down the gauntlet on an international competition that he hoped would accelerate his academic field. Laland, an evolutionary biologist at the University of St. Andrews in the United Kingdom, is part of a consortium that has a €2 million grant to explore how human culture evolves. A crucial part of this issue is how people develop new behaviors. To probe this question, Laland and the rest of the consortium wanted to examine the relative importance of social learning, the acquisition of behaviors from watching other people, versus individual innovation.

The ability to learn from others is central to the evolution and persistence of culture, and it is viewed as part of the reason humans have come to dominate the planet. But, notes Laland, "it has proven quite tricky in a formal mathematical sense to link social learning to our success as a species."

Equally important, it hasn't been clear how people best learn socially. Sometimes individuals copy the behaviors of others seemingly at random; other times they appear to decide who to copy based on the level of prestige of the individual. "We find evidence of different rules, but it was frustrating because we didn't know which of these rules was the best," Laland says.

To address that question, the consortium decided to host a tournament, with a €10,000 prize, in which all comers would pit computer programs incorporating social-learning strategies against each other. "It was a gamble," Laland recalls. "My biggest fear was that nobody would participate."

The gamble paid off. More than 100 teams, including a few from high schools, competed—"far more than we even hoped for," says Laland. The analysis of the competition, reported on page 208, "addresses in a clearly new way some of the questions that have been nagging at the field of evolutionary social learning for more than 2 decades," says Luc-Alain Giraldeau, a behavioral ecologist at the University of Quebec, Montreal, in Canada.

A simple approach, from a surprising entry, won hands down. Most researchers had thought that a mix of learning on one's own and social learning would be the best strategy. Yet a pair of graduate students won the contest with a strategy heavily tilted toward imitation rather than innovation (see sidebar, p. 166). "The main take-home message is that it pays to imitate success, except when there is evidence that what has been successful recently is no longer working well," says Robert Axelrod of the University of Michigan, Ann Arbor.

The tournament's results, say some researchers, have implications for human cultural evolution. "It implies that our success as a species rests heavily on the right social and networking skills of knowing who, what, and when to copy," notes Samuel Bowles, an economist at the Santa Fe Institute.

Getting the game together

As strange as a face-off might seem for science, there's a powerful precedent for a tournament giving the study of human behavior a much-needed jolt. In the 1980s, Axelrod organized a competition to get a handle on why cooperation evolves. He challenged contestants to engage repeatedly in the prisoner's dilemma—in which two prisoners must choose between squealing on one another or cooperating in their refusal to talk to the police. "The strength of a tournament is that it provides the means to achieve insights that can be totally unexpected by both the contestants and even the designers of the tournament," says Axelrod.

In his contest, involving 28 entrants, a simple tit-for-tat strategy—cooperate if your fellow prisoner cooperates, otherwise don't—proved the most effective. The tournament "was enormously influential because before, not many people were studying the evolution of cooperation," says Laland, "It kick-started the field."

Laland felt the study of social learning would benefit from a similar boost. He tapped St. Andrews postdoctoral fellow Luke Rendell to design the tournament and recruited help from experts in social learning, cultural evolution, and game theory. Rendell's first task was to figure out a problem to be solved by contestants. The cooperation field had already accepted the prisoner's dilemma as relevant, but there was no such theoretical tool for assessing the adaptive value of social learning in a complex environment.

After much discussion, Laland's team decided to center the tournament on a learn-

CREDIT: M. TWOMBLY/SCIENCE

Scheming. Daniel Cownden (*foreground*) and Timothy Lillicrap worked out a winning social-learning strategy.

A Winning Combination

The chance to spend 4 months designing a computer program to compete in a social-learning tournament (see main text, p. 165) immediately appealed to Timothy Lillicrap, a graduate student in computational neuroscience at Queen's University in Kingston, Canada. "So much time in science is spent getting half-answers to complicated questions, which often take decades to appreciate fully. The tournament was an opportunity to work on something with a hard and fast goal and a hard and fast answer," he recalls.

Lillicrap teamed up with Daniel Cownden, a fellow graduate student in mathematics, and together they obsessively spent hundreds of hours perfecting their social-learning strategy. Although they lacked experience in social-learning research, Lillicrap knew about figuring out the most efficient way to accomplish a task and estimating outcomes based on what known data is available, whereas Cownden understood evolutionary game theory. This mix of knowledge proved a winning combination, as their program easily beat about 100 others. "Certainly a reason for our success was the balance between Tim's concrete, computer science, 'code it and see' approach and my abstract, mathematical, 'hem and haw' approach," says Cownden.

Early on, the pair set up their own in-house competition according to the tournament's rules. The duo met weekly to face off with the latest iterations of their computer programs. The loser had to make the winner dessert but got to see the winner's program and crib from it if desired. They quickly homed in on potentially good approaches. "We were one of the few entries which noticed that it is virtually always better to observe others rather than gather information through innovation," says Lillicrap.

The two students created a player that had access to all the hidden variables, allowing it to "cheat" and make nearly perfect decisions. By recording how this superplayer worked the game, they obtained information needed to train a neural network that would underpin their final strategy. The cheater figured out, for example, how to evaluate whether to learn something that might be useful later on or just use a behavior already in its repertoire.

Cownden and Lillicrap called the strategy they entered "discountmachine" because it discounted less certain future rewards for more guaranteed immediate gains. That involved weighing how fast the environment will change—if it's changing fast, then past social learning gets outdated quickly—and how good, and reliable, the payoffs for an action were. Based on that information, discountmachine decides whether to do what it already knows how to do, getting an immediate reward, or whether to see what someone else is doing and learn from them, in the hope of getting what might be a bigger reward later.

Cownden and Lillicrap were definitely dark-horse candidates going into the contest. "We were quite surprised nobody from a social-learning lab had gone on to win," says Luke Rendell, an organizer of the tournament from the University of St. Andrews in the United Kingdom. "You have to take your hat off to them."

For Cownden, the tournament has made a big difference in his graduate studies. He wants to develop a better method for comparing the value of information with the value of actions in making decisions and plans to combine decision theory with evolutionary game theory to do so. In this way, he hopes to come up with a better way to approach problems such as that offered by the tournament. "My research goals were vague and unformed when I encountered the tournament. Now they are focused and clear."

–E.P.

ing problem called the restless multiarmed bandit. The name is inspired by the "one-armed bandit"—the slot machine—in which pulling a lever sets off a game of chance. In the tournament's case, the multiple "arms" represent different abstract "behaviors"—100 possible ones in all, each associated with a different payoff. The challenge is to maximize one's payoff over the long run. "I often envisage it as being dropped onto a jungle island, where there are a number of possible ways of getting food," says Rendell. One might learn to gather fruit, hunt, fish, grub for bugs, dig out tubers, et cetera, with varying degrees of success.

In a new environment like a jungle, there are two ways to pick up a new skill: figure out what to do by trial and error or copy what others do after observing or interacting with them. Both learning approaches take time and may or may not lead to a good payoff. But eventually, one builds up a repertoire of useful behaviors.

Rendell created a computer program for running the multiarmed bandit problem under varying conditions. In each scenario, 100 individuals have three options every round: "observe," to learn a behavior by watching another individual; "innovate," to develop one of the behaviors on one's own; or "exploit," which is the equivalent of pulling one of the bandit's arms by engaging in one of the behaviors that have been acquired. Each round, only exploiters get points, so there is always a tradeoff between using an existing behavior and getting a payoff right away or learning a new, and potentially better, behavior that could earn points later.

The simulated individuals in Rendell's artificial world can follow different strategies programmed into them, such as observing more and innovating less, and each round some randomly "die." They are replaced, at times by copies of individuals with more effective learning strategies, as determined by earned points; the process approximates natural selection's survival of the fittest.

Rendell also made it so that he could vary the conditions between rounds. Learning from others might become more or less error-prone, for example. The payoff points associated with each behavior could also increase or decrease over the course of a simulation, adding another element to the changing environment. This twist introduces the possibility that the simulated individuals are acting on outdated information, which changes the dynamic between social learning and innovating.

Ready. Set. Go.

Designing, testing, and refining the tournament's rules took 18 months, and during that time no one in the consortium knew if it would attract entrants. To their relief, the interest was overwhelming; 104 entries from 16 countries

CREDIT: BRONWYN MCLEAN

Published by AAAS

and more than a dozen disciplines accepted the challenge, each providing a computer program encoding a strategy to guide the behavior of individuals in Rendell's tournament.

The first stage of the tournament, which took place in 2008, pitted pairs of strategies head to head, in more than 5000 contests. Each contest started with a simulation in which one strategy guided the actions of all 100 individuals for 100 rounds. Because of elements of chance built into the simulation—it might take one round or many to acquire a particular behavior by observing someone, for example—each individual accumulated a different point tally over that starting period. Then individuals whose actions are controlled by an opponent's program start to "invade," dropping into the "world" to replace some of the ones that "died." The two strategies would then battle for 10,000 rounds. Over that time, individuals accumulating more points reproduced more, representing an ever-greater proportion of the population. The strategy having the most individuals in the final 2500 rounds was declared the winner of that simulation. The overall winner in each head-to-head contest was determined by tallying 20 such simulations.

This first round, requiring more than 100,000 simulations, winnowed the field to 24 top learning strategies, which then further competed in a round-robin runoff to determine the top 10.

In early 2009, the second stage, the "melee," pitted all 10 finalists together in simulations. They played against each other under 15,000 scenarios, each with 10,000 rounds. The whole tournament was only able to be held thanks to more than 65,000 hours of computer time provided by the U.K. National Grid Service.

The winning team, two Canadian graduate students from outside the small world of social-learning research, was a shocker to many. "To be honest, I was quite surprised and annoyed. We really thought we had winning entries because we had run several tournaments using our own strategies before submitting the final versions," says Giraldeau.

The lack of formal training in social learning didn't prevent the winners from creating a strategy that depended more heavily on that process than others. "Very few other strategies realized that it *never* paid to innovate and that observation was the only choice," says Daniel Cowden, one of the two students from Queen's University in Kingston, Canada.

The pair's strategy was "exceptionally clever," says Robert Boyd, a biological anthropologist at the University of California, Los Angeles, who is a consortium member and an author of the tournament report. "In [an] environment where the world is changing, the best strategy is a lot of imitation."

Before the tournament, many researchers had considered imitation limited in value because one might unknowingly waste time copying an outdated or inappropriate behavior. The contest demonstrated that copying was a good approach because those being imitated had likely already decided on—and were enacting—the best strategy they could. "It's kind of parasitizing good ideas that other strategies are generating," says Laland. New, possibly better, behaviors still arise because copying is typically imperfect and errors in imitating a behavior at times yielded improvements.

Like father. Evolutionary biologist Kevin Laland being copied by his son, a tendency that has helped make humans so successful.

CREDIT: GILLIAN BROWN

In general, the learning strategies that best succeeded in the tournament shared several features. They emphasized social learning but spent as much time as possible enacting known behaviors and earning payoff points. The results "say clearly that if you spend too much time learning, life will pass you by," says Rendell. Given the conditions set up in the tournament, devoting just 10% of one's time to learning proved to be optimal.

The competition also showed that it is important to assess changes in the environment, such as shifting payoffs for behaviors, and adjust accordingly, even in the middle of a run. Players benefited, too, if they could keep track of when something was learned, because in a changing environment older behaviors were more likely to become outdated.

Bowles suggests that the tournament's results will reorient thinking about what drives human progress. "Most people, when they think about where new ideas come from, think about some eccentric tinkering in a garage or some shy geek playing around with a computer. We think that's how progress gets made," he explains. "What this group of authors [is] suggesting is that this does go on, but what really is decisive is spreading these ideas."

Next step

Richard McElreath, an evolutionary ecologist at the University of California, Davis, agrees that the social-learning tournament was valuable, but he has some reservations. "Simulations and tournaments can be solidly criticized for being both difficult to interpret and potentially misleading," he says.

Consider the strategy ranked 95th in the tournament. Graduate students Shane Gero of Dalhousie University in Halifax and Marianne Marcoux of McGill University in Montreal, both in Canada, based their entry, "higherlearning," on the idea that the education of graduate students often depends on a lot of innovation and individual learning rather than social learning. In the tournament's simulations "being a graduate student was very maladaptive," Gero notes. In real-world academia, however, such behavior arguably does quite well.

Alan Rogers, an anthropologist at the University of Utah in Salt Lake City, suggests that the options in the tournament were also not as clear-cut as they appeared. He proposes that the very act of doing something, or exploiting, involves some subtle elements of individual learning that are not seen as learning per se.

Two more tournaments testing learning strategies are in the works. One will look at what happens when information flow is restricted: for example, when players can observe only a subset of the behaviors. In the other, the players will know something about the age and success rate of players whose behavior they can copy. Things like the perception of experience and prestige "may be very important in human culture," says Laland.

No matter what happens in any subsequent contests, this one's place in history seems assured. "This tournament will invigorate the field by attracting new scholars," says McElreath. "I expect it to become a classic."

–ELIZABETH PENNISI