Quantifying consciousness

Back to index

Posted October 22, 2009

Angry red trapezoids kill and eat the carcasses of their green kin. Critters turn blue to signal their desire to mate. Adults gobble their young. Welcome to Polyworld, the purported answer to one of the biggest unsolved problems in science: a theory of consciousness.

By Janelle Weaver

Angry red trapezoids kill and eat the carcasses of their green kin. Critters turn blue to signal their desire to mate. Adults gobble their young. Welcome to Polyworld, the purported answer to one of the biggest unsolved problems in science: a theory of consciousness.

"The objective of Polyworld is to survive any way you can," said Virgil Griffith, a graduate student at CalTech, who created a program to simulate in two dimensions how we evolve awareness of the outside world.

At first, the green shapes don't act very intelligently. They shoot through their rectangular world in straight lines, hoping to accidentally run into food. They run out of energy and die during childbirth. Then the offspring eat their carcasses.

"They didn't know any better," Griffith explained, addressing attendees of the recent CASW New Horizons in Science Briefing in Austin.

These critters had a lot to learn. They developed one ingenious strategy in response to a programming mistake that made propagation effortless. "They were constantly eating their own children," Griffith said. "It was a nice case of an immensely successful solution of exploiting a bug."

Next, they learned to recognize other green trapezoids as food, blue shapes as potential mates, and red shapes as enemies. "Through sheer evolutionary forces, we can achieve

conscious critters," Griffith said.

Distinguishing red, green and blue requires at least three information states, according to the CalTech researcher. In terms of consciousness, the trapezoids' evolutionary exploit places them above a photodiode, which recognizes only black and "not black."

A camera can recognize an impressive number of colors. Does this mean it's as conscious as you? No, because it is a collection of disjointed parts. "You may cut down the middle of the camera, or any kind of cut you want, and the functionality does not degrade significantly," Griffith said.

Consciousness only arises from integration — when the entire system is more than the sum of its parts. The phi value is a measure of integration. (The Greek letter phi, f, looks like a line vertically dissecting a circle. The "I" is for information and the "O" is for integration.)

"We can't calculate the phi value of the human brain, but we do know roughly how the areas are connected," Griffith said. In his simulations, the cerebrum — the brain structure crucial for consciousness — earns a higher phi value than the cerebellum, which controls motor responses.

"We can informationally describe how your green is different from your red, and how your red is different from your blue, and all possible states of your brain," Griffith said. "The idea is that this set of all the differences uniquely describes the informational state that you're brain is in."

Polyworld proved that evolution maximizes phi values. "Evolution can make phis; it's reasonable that it can make brains," Griffith said.

But some question whether it's possible to develop a scientific theory of consciousness. "I don't even know if it's correct, but it's currently the best game in town," Griffith said.

There's just one catch.

"It's really computationally unpleasant. This is why we just do this on paper," Griffith said. He can only calculate phi values for a network of 20 nodes, or neurons. "The smallest interesting brain I know of is C. elegans, which is 302 neurons," Griffith said.

Nonetheless, Griffith remains optimistic. "You just do a half-assed approximation. That's still pretty cool. Hey, you could say within a certain margin of error that your red is probably like your dog's red. I'll take that."

Janelle Weaver is a graduate student in the science communication program at UC Santa Cruz.

Last revised: October 23, 2009
The National Association of Science Writers, Inc.
P.O. Box 7905, Berkeley, CA 94707 | (510) 647-9500
