BOOKS BY JOHN BROCKMAN

As author:
*By the Late John Brockman*
*37*
*Afterwords*
*The Third Culture: Beyond the Scientific Revolution*
*Digerati*

As editor:
*About Bateson*
*Speculations*
*Doing Science*
*Ways of Knowing*
*Creativity*
*The Greatest Inventions of the Past 2,000 Years*
*The Next Fifty Years*
*The New Humanists*
*Curious Minds*
*What We Believe But Cannot Prove*
*My Einstein*
*Intelligent Thought*
*What Is Your Dangerous Idea?*

As coeditor:
*How Things Are*

# WHAT ARE YOU OPTIMISTIC ABOUT?

Today's Leading Thinkers
on Why Things Are Good and Getting Better

*edited by* JOHN BROCKMAN

2007

HARPER PERENNIAL

NEW YORK • LONDON • TORONTO • SYDNEY

too, the new equipment is simply on a par with much of the old.

Finally, empirical science is at last beginning systematically to address the sources and wellsprings of human happiness and human flourishing, and the findings of these studies must themselves be taken as important data points for the design and marketing of (putative) technologies of enhancement.

In sum, I am optimistic that we will soon see the end of those overused, and mostly ad-hoc, appeals to the 'natural.'

## A Breakthrough in Understanding Intelligence Is Around the Corner


### TERRENCE SEJNOWSKI

Computational neuroscientist, Salk Institute; coauthor (with Patricia Churchland) of *The Computational Brain.*


The clinically depressed often have a more realistic view of their problems than those who are optimistic. Without a biological drive for optimism, it might be difficult to motivate humans to take on difficult problems and face long odds. What optimistic view of the future drives string theorists in physics working on theories that are probably hundreds of years ahead of their time? There is always the hope that a breakthrough is just around the corner.

In 1956, a small group of optimists met for a summer conference at Dartmouth, inspired by the recent invention of digital



computers and breakthroughs in writing computer programs that could solve mathematical theorems and play games. Since mathematics was among the highest of human achievements, they thought that engineered intelligence was imminent. Last summer, fifty years later, another meeting was held at Dartmouth that brought together the founders of the field of artificial intelligence and a new generation of researchers. Despite all the evidence to the contrary, the pioneers from the first meeting were still optimistic and chided the younger generation for having given up the goal of achieving human-level intelligence.

Problems that seem easy – such as seeing, hearing, and moving about – are much more difficult to program than theorem-proving and chess. How could this be? It took hundreds of millions of years to evolve efficient ways for animals to find food, avoid danger, and interact with one another, but humans have been developing mathematics for only a few thousand years, probably using bits of our brains that were meant to do something altogether different. We vastly underestimated the complexity of our interactions with the world, because we are unaware of the immense computation our brains perform to make seeing objects and turning doorknobs seem effortless.

The early pioneers of AI sought logical descriptions that were black or white and geometric models with a few parameters, but the world is high-dimensional and comes in shades of gray. The new generation of researchers has made progress by focusing on specific problems in planning, computer vision, and other areas of AI. Intractable problems have yielded to probabilistic analysis of large databases using powerful statistical techniques. The first algorithms that could handle this complexity were neural networks with many thousands of parameters, which learned to categorize input patterns from labeled examples. New algorithms for machine learning have been discovered that can extract hidden statistical structure from large data sets without the need for any labels. Progress is accelerating now that the Internet provides truly large data sets of text and images. Computational linguists, for example, have adopted statistical algorithms for parsing sentences and language translation, having found transformational grammars too impoverished.

One of the most impressive learning systems is TD-Gammon, a computer program that taught itself to play backgammon at the championship level. Built by Gerald Tesauro at IBM Yorktown Heights, TD-Gammon started out with little more than the board position and the rules of the game, and the only feedback was who won. TD-Gammon solved the temporal credit-assignment problem: If, after a long string of choices, you win, how do you know which choices were responsible for the victory? Unlike rule-based game programs, TD-Gammon discovered, on its own, better ways to play positions and developed a surprisingly subtle sense of when to play safely and when to be aggressive. This captures some important aspects of human intelligence.

Neuroscientists have discovered that dopamine neurons, found in the brains of all vertebrates, are central to reward learning. The transient responses of dopamine neurons signal to the brain predictions for future reward, which are used to guide behavior and regulate synaptic plasticity. The dopamine responses have the same properties as the temporal difference learning algorithm used in TD-Gammon. Reinforcement learning was dismissed years ago as too weak a learner to handle the complexity of cognition. This belief needs to be reevaluated in



the light of the successes of TD-Gammon and learning algorithms in other areas of AI.

What would a biological theory of intelligence look like, based on internal brain states derived from experimental studies rather than introspection? I am optimistic that we are finally on the right track – and that before too long, an unexpected breakthrough will occur.

## AI Will Arise

### JORDAN POLLACK

Computer scientist, director of the Dynamical & Evolutionary Machine Organization Laboratory, Brandeis University.

I often attack the original, 'we can program it' direction of the field of artificial intelligence, but I'm still optimistic that our primitive electromechanical and computing machines will one day become intelligent enough to be treated as living creatures. There won't be a way to prove they are intelligent or alive, but just as digitized music and rasterized print appear continuous because the differences are below our perceptual thresholds, electro-mechanical animations of intelligence will appear as alive because the differences will be less than we can detect.

I have a predictive sketch for how such intelligent machines might arrive. My definition of 'robot' is any device controlled by software and interacting with the physical world. An economically viable robot is such a system that earns a consistent