


This

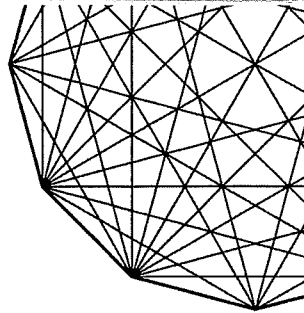
Deep, Beautiful, and
Elegant Theories of
How the World Works

 **Explains**

Edited by
JOHN BROCKMAN

2013

Everything



HARPER PERENNIAL

NEW YORK • LONDON • TORONTO • SYDNEY • NEW DELHI • AUCKLAND

NATURE IS CLEVERER THAN WE ARE

TERRENCE J. SEJNOWSKI

Computational neuroscientist; Francis Crick Professor, the Salk Institute; coauthor (with Patricia S. Churchland), The Computational Brain

We have the clear impression that our deliberative mind makes the most important decisions in our life: what work we do, where we live, whom we marry. But contrary to this belief, the biological evidence points toward a decision process in an ancient brain system called the basal ganglia, brain circuits that consciousness cannot access. Nonetheless, the mind dutifully makes up plausible explanations for the decisions.

The scientific trail that led to this conclusion began with honeybees. Worker bees forage the spring fields for nectar, which they identify with the color, fragrance, and shape of a flower. The learning circuit in the bee brain converges on VUMmx1, a single neuron that receives the sensory input and, a bit later, the value of the nectar, and learns to predict the nectar value of that flower the next time the bee encounters it. The delay is important, because the key is prediction, rather than a simple association. This is also the central core of temporal-difference (TD) learning, which entails learning a sequence of decisions leading to a goal and is particularly effective in uncertain environments, like the world we live in.

Buried deep in your midbrain, there's a small collection of neurons—found in our earliest vertebrate ancestors, and projecting throughout the cortical mantle and basal ganglia—that are important for decision making. These neurons release a neurotransmitter called dopamine, which has a powerful influence on our behavior. Dopamine has been called a “reward molecule,” but

more important than reward itself is the ability of these neurons to predict reward: If I had that job, how happy would I be? Dopamine neurons, which are central to motivation, implement TD learning, just as VUMmx1 does.

TD learning solves the problem of finding the shortest path to a goal. It's an on-line algorithm, because it learns by exploring and discovers the value of intermediate decisions in reaching the goal. It does this by creating an internal value function, which can be used to predict the consequences of actions. Dopamine neurons evaluate the current state of the entire cortex and inform the brain about the best course of action from the current state. In many cases, the best course of action is a guess, but because guesses can be improved, TD learning creates, over time, a value function of oracular powers. Dopamine may be the source of the “gut feeling” you sometime experience, the stuff that intuition is made from.

When you're considering various options, prospective brain circuits evaluate each scenario, and the transient level of dopamine registers the predicted value of each decision. The level of dopamine is also related to your level of motivation, so not only will a high level of dopamine indicate a high expected reward, but you will also have a higher level of motivation to pursue it. This is quite literally the case in the motor system, where a higher tonic dopamine level produces faster movements. The addictive power of cocaine and amphetamines is a consequence of increased dopamine activity, hijacking the brain's internal motivation system. Reduced levels of dopamine lead to anhedonia, an inability to experience pleasure; and the loss of dopamine neurons results in Parkinson's disease, an inability to initiate actions and thoughts.

TD learning is powerful because it combines information about value along many different dimensions, in effect comparing apples and oranges in achieving distant goals. This is important because rational decision making is very difficult when there are

many variables and unknowns. Having an internal system that quickly delivers good guesses is a great advantage, and may make the difference between life and death when a quick decision is needed. TD learning depends on the sum of your life experiences. It extracts what is essential from these experiences long after the details of the individual experiences are no longer remembered.

TD learning also explains many of the experiments performed by psychologists who trained rats and pigeons in simple tasks. Reinforcement learning algorithms have traditionally been considered too weak to explain complex behaviors, because the feedback from the environment is minimal. Nonetheless, reinforcement learning is universal among nearly all species and is responsible for some of the most complex forms of sensorimotor coordination, such as piano playing and speech. Reinforcement learning has been honed by hundreds of millions of years of evolution. It has served countless species well, particularly our own.

How complex a problem can TD learning solve? TD-Gammon is a computer program that learned how to play backgammon by playing itself. The difficulty with this approach is that the reward comes only at the end of the game, so it's not clear which were the good moves that led to the win. TD-Gammon started out with no knowledge of the game, except for the rules. By playing itself many times and applying TD learning to create a value function to evaluate game positions, TD-Gammon climbed from beginner to expert level, along the way picking up subtle strategies similar to ones that humans use. After playing itself a million times, it reached championship level and was discovering new positional play that astonished human experts. Similar approaches to the game of Go have achieved impressive levels of performance and are on track to reach professional levels.

When there's a combinatorial explosion of possible outcomes, selective pruning is helpful. Attention and working memory allow

us to focus on most of the important parts of a problem. Reinforcement learning is also supercharged by our declarative memory system, which tracks unique objects and events. When large brains evolved in primates, the increased memory capacity greatly enhanced their ability to make complex decisions, leading to longer sequences of actions to achieve goals. We are the only species to create an educational system and to consign ourselves to years of instruction and tests. Delayed gratification can extend into the distant future (in some cases, into an imagined afterlife), a tribute to the power of dopamine to control behavior.

At the beginning of the cognitive revolution in the 1960s, the brightest minds could not imagine that reinforcement learning could underlie intelligent behavior. Minds are not reliable. Nature is cleverer than we are.