13

There Are No Killer Apps but Connecting Neural Activity to Behavior through Computation Is Still a Good Idea

P. Read Montague

Abstract

The quest to understand the relationship between neural activity and behavior has been ongoing for well over a hundred years. Although research based on the stimulus-and-response approach to behavior, advocated by behaviorists, flourished during the last century, this view does not, by design, account for unobservable variables (e.g., mental states). Putting aside this approach, modern cognitive science, cognitive neuroscience, neuroeconomics, and behavioral economics have sought to explain this connection computationally. One major hurdle lies in the fact that we lack even a simple model of cognitive function. This chapter sketches an application that connects neuromodulator function to decision making and the valuation that underlies it. The nature of this hypothesized connection offers a fruitful platform to understand some of the informational aspects of dopamine function in the brain and how it exposes many different ways of understanding motivated choice.

Introduction

Let's face it. Computational neuroscience and its fledging product, computational psychiatry, simply do not have a killer app—yet. Certainly nothing like Newton's laws, William Rowan Hamilton's transformative approach to dynamics, the Dirac equation, Darwin's evolution by natural selection and its rendering in the twentieth century modern synthesis (Mayr and Provine 1980/1998), or even Shannon's breakthrough efforts in what is now called information theory. Marshaling such a pantheon isn't quite fair, but it makes a point. We should require a lot from any account that calls itself a killer app,

From "Computational Psychiatry: New Perspectives on Mental Illness,"
A. David Redish and Joshua A. Gordon, eds. 2016. Strüngmann Forum Reports, vol. 20,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03542-2.

especially in an area that purports to connect mind and brain in a meaningful way. In the world of sustaining healthy human mental function and characterizing and treating unhealthy human mental function, the killer app will depend on a much more evolved body of constructs (models) surrounding cognition. The limiting factor is (at least) our woefully simple models of cognitive function. We simply do not yet have an evolved and integrated model of human cognition that can render a human-like model agent in a perceptual problem or learning problem to use such a set up to gain penetrating insight into a psychiatric disorder. Instead, the current situation bears the hallmarks of the early days of any discipline: some very provocative models exist, focused in particular areas and mapped with variable success to experimental data extracted from candidate neural systems.

In this chapter, I will sketch an application that connects neuromodulator function to decision making and the valuation that underlies it. The nature of this hypothesized connection has proved to be a fruitful platform for understanding some of the informational aspects of dopamine function in the brain and how it exposes many different ways of understanding motivated choice.

The Platform of Reinforcement Learning

For well over a 100 hundred years, models of learning have been dominated by psychological concepts about how animals adapt to the changing world around them. The foundational ideas emerged in the early nineteenth century from the work of physiologist Ivan Pavlov and his star student Jerzy Konorski on the conditioned reflex (Pavlov 1927; Konorski 1949). This work developed into an entire behaviorist movement that flourished through the twentieth century with its now familiar collection of names: Thorndike, Hull, Watson, Skinner and so on. One of the strictures of this movement was to remove all mention of variables that could not be observed, especially any mention of unobserved mental states. All behavior was to be rendered as stimulus and response, a view that modern cognitive science, cognitive neuroscience, neuroeconomics, behavioral economics, and their computational expressions toss aside. Apparently, there was something to be feared about positing unobserved states of mind as though such unobserved entities prevent the hard science from taking place. The behaviorists were likely just reacting to Freud’s influence on psychology; however, it is noteworthy that unobserved or unobservable entities and states pervade physics and biophysics despite the fact that both areas are viewed as hard science. Biophysical models of ionic channel function have long and happily accepted unobserved states and state transitions, usually cast mathematically as hidden Markov models (Hille 2007). Latent states, latent variables, unseen fitness or hazard functions, hidden Markov models, and their more exotic congeners are now simply part of the inventory of modern computational approaches to mind and brain.

From “Computational Psychiatry: New Perspectives on Mental Illness,”
A. David Redish and Joshua A. Gordon, eds. 2016. Strüngmann Forum Reports, vol. 20,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03542-2.

Capturing the Regularities of Learning: From Bush–Mosteller to Sutton–Barto

One key area where a rigid stimulus-response framing was very useful was learning, since it is here that experimental psychology first began to identify so-called learning rules—statistically lawful mappings between input, internal state, and the output of the entire creature. All mobile creatures need to learn because they move; movement ensures that the contingencies for survival change, and do so on multiple time and space scales. Sessile creatures (e.g., sea cucumber) have developed some very peculiar strategies for adapting to environmental threat (they partially eviscerate themselves as a defense), but a moving creature is where real learning action takes place. At the minimum, mobile creatures must deal with the environmental changes that result from their own movement. In this sense, movement and learning have always been partner processes, so behaviorist paradigms provide very nice and structured ways to probe simple learning and capture the results in equally simple laws.

The rules that characterize learning in mobile animals start with the work of Ivan Pavlov, who originated the modern interpretation of the conditioned reflex: ring the bell, feed the dog, rinse-and-repeat. Through this regular training, the originally neutral bell comes to elicit the features (orientating, salivating, secretion of digestive enzymes, and so on) of the unconditioned response to food. This is classical conditioning, and its "cousin," instrumental conditioning, contains the same regularities but requires an action on the part of the animal. Pavlov generated a tradition around this idea, and it was certainly the foundation for the behaviorism movement, as mentioned above. However, for modern computationalists, the important steps were taken just after World War II with the emergence of work by Robert Bush and Frederick Mosteller (e.g., Bush and Mosteller 1951a, b, 1953, 1955). At that time, these investigators were considered part of a new breed of mathematical psychologists—perhaps the first generation (were it not for Hermann Helmholtz's work in the late nineteenth century). They originated the idea of prediction learning and introduced the first rigorous account of the kinds of learning described by the behaviorists. Rendering learning as a problem of learning-to-predict was a departure from the correlational theories of Konorski (1949) and Hebb (1949). The problem with such correlational accounts are manifold, but the main impediment is that they do not provide a natural way for a "correlation-based" learner to learn chains of events. Both correlational accounts and prediction accounts for learning, however, viewed the animal as a statistical learner whose "learning job" was to extract regularities latent in the statistics of their experience.

In the Bush and Mosteller account, the conditioning described by Pavlov was rendered as a trial-based prediction of the unconditioned response and also provided a simple way to update that prediction from trial to trial:

$$p_{t+1} = p_t + \alpha\left(R_o - p_t\right). \tag{13.1}$$

From "Computational Psychiatry: New Perspectives on Mental Illness,"
A. David Redish and Joshua A. Gordon, eds. 2016. Strüngmann Forum Reports, vol. 20,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03542-2.

The goal here is to associate stimuli with actions, and the Bush–Mosteller model updates the probability p that the action (salivation) occurs on trial $t+1$ as a function of its value in the previous trial t and the value of the observed reward R_o. This is the first good account of prediction learning to explain the learning associated with behavioral conditioning paradigms. But Bush and Mosteller went further and modeled the animal as a collection of probabilistic processes. In an obituary for the late Robert Bush, Mosteller (1974:170) aptly describes the modern flavor of their approach:

> In the models for learning that Bush and I developed, the fundamental representation was that prior to a trial an organism was a vector of response probabilities. A stimulus corresponded to a mathematical operator that replaced the organism's current vector by a new probability vector. In the models of Bush and Mosteller (1955), the effects of previous responses were summed up in the current vector, independent of the path to the present state. The operators had a linear form, so that if p is a vector of probabilities ($p_1, p_2, ..., p_k$) and Q is applied to p the new vector is:
>
> $$Qp = \alpha p + (1 - \alpha)\,\lambda,$$
>
> where λ is also a probability vector ($\lambda_1, \lambda_2, ..., \lambda_k$) and α is a scalar, $0 \leq \alpha \leq 1$. If Q is repeatedly applied, the limiting vector is λ, when $\alpha \neq 1$.

This is an extremely rich model of the processes putatively at work inside the learner. Notice that even the history-independent assumption (the Markovian assumption) is present in their papers as of the early 1950s, "…independent of the path to the present state" (see also Rescorla and Wagner 1972). As forward looking as the Bush–Mosteller approach was, it still missed some important aspects of learning, such as the detailed dependence on timing of stimuli and other well-known conditioning effects such as secondary conditioning: If A predicts reward, and B is trained to predict A, then B will also predict reward. From the psychological and computer science literature there emerged another approach to this problem offered up by Richard Sutton and Andrew Barto (1981, 1987, 1998; for complete references, see Sutton 1988). Their work focused on an incremental learning algorithm, called the method of temporal differences, which exploited differences between successive predictions rather than simply the difference between a prediction and an outcome. This difference is crucial, as it framed the "goal of learning" as the problem of *learning to value the future of the states* of the agent. This agent is portrayed as moving about in some kind of high-dimensional state-space, making transitions from one state S_t at time t to another state S_{t+1} at time $t + 1$. There were two basic assumptions to this approach. First, the *goal of learning* is to learn the value of states taken as the discounted amount of future reward expected from that state forward into the distant future. The other assumption was that it did not matter how the state was reached:

From "Computational Psychiatry: New Perspectives on Mental Illness,"
A. David Redish and Joshua A. Gordon, eds. 2016. Strüngmann Forum Reports, vol. 20,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03542-2.

$$V\left(S_t\right)=E\left(r_t+\gamma r_{t+1}+\gamma^2 r_{t+2}+\cdots\right) \textit{for } 0<\gamma\leq 1. \qquad (13.2)$$

Here the expected value operation "*E*" is slightly bad notation. The expectation *E* is taken for each "tic" forward and so we should read it as expressing that each *r* is a separate expected value of reward at each step to the future of . So the value of the current state of the agent depends on its future. The second bit of ambiguity in Equation 13.2 is that the expectation implicitly includes the rule that the agent uses to transition from state S_t to state S_{t+1} The single *E* symbol does not specify this clearly in Equation 12.2, but these details do not matter here.

Overall, the Sutton–Barto account appears to be a small change from the Bush–Mosteller approach; moreover, it mimicked approaches from the late 1950s by Samuel, who made automatic checker-playing programs (Samuel 1959). The Sutton–Barto effort did account, however, for secondary conditioning, as well as the way that an agent learns to chain events together. It also connected to animal conditioning (Sutton and Barto 1987) and to an independently developing area of optimal control called dynamic programming (Bellman 1957). As I review below, it also reached down to important biological observations. This multidimensional reach, which crossed levels of description, is and was what makes this work so important. Sutton and Barto understood the connection of their work to previous approaches but they also understood that they had added crucial insights. Quoting from Sutton (1988:9):

> This article introduces a class of incremental learning procedures specialized for prediction – that is, for using past experience with an incompletely known system to predict its future behavior. Whereas conventional prediction-learning methods assign credit by means of the difference between predicted and actual outcomes, the new methods assign credit by means of the difference between *temporally successive predictions*. Although such *temporal-difference methods* have been used in Samuel's checker player, Holland's bucket brigade, and the author's Adaptive Heuristic Critic, they have remained poorly understood. Here we prove their convergence and optimality for special cases and relate them to supervised-learning methods.

The Valuation of the Future

Let us turn back to the central idea of Sutton–Barto: the value of a state scales according to the value of the discounted future it portends. Why should a mobile creature need to value the future? One word: uncertainty. It appears that in our world a vast amount of uncertainty lies in the future with the more important bits of uncertainty rolling out in the near future. Therefore, the valuation of a state based on its expected future, when combined with the assumption that the system is Markovian (history independent), animates the power of this simple approach. Let us take Equation 13.2 and step forward one tic to a new state S_{t+1} and value this new state in exactly the same way: as the expected value of reward from that state forward:

$$V\left(S_{t+1}\right) = E\left(r_{t+1} + \gamma r_{t+2} + \gamma^2 r_{t+3} + \cdots\right). \tag{13.3}$$

In principle, Equations 13.2 and 13.3 would require a system to run through a state infinitely often and sum a long (infinite) series of numbers to estimate the value of the state. Herein lies the very nice way that this valuation function is formulated. If we scale Equation 13.3 by the discount factor γ (the rate that the future becomes less valuable at each "tic") then:

$$\gamma V\left(S_{t+1}\right) = E\left(\gamma r_{t+1} + \gamma^2 r_{t+2} + \gamma^3 r_{t+3} + \cdots\right). \tag{13.4}$$

Notice now that there is a way to relate the value function of time *t* to the value at time *t*+1:

$$V\left(S_t\right) = E\left\{r_t\right\} + \gamma V\left(S_{t+1}\right). \tag{13.5}$$

If the agent (the learner) had perfect estimates of the value of all its states, then this expression would be exact. The real world is not exact so this condition never holds perfectly; however, it does give a natural way to define an error signal. Simply subtract the left-hand side from the right-hand side:

$$\text{“0”} = E\left\{r_t\right\} + \gamma V\left(S_{t+1}\right) - V\left(S_t\right). \tag{13.6}$$

The quotes here indicate that this difference is never really 0 in the real world. However, Sutton and Barto exploited this temporal difference (TD) to build an approach to learning that has wide-reaching applications and implications (Sutton and Barto 1998). An error signal of just this type was first hypothesized to be a general mechanism for biological systems to learn how to value states, store predictions, and link value to action based on predictions (Montague et al. 1993, 1995, 1996, 2004; Schultz et al. 1997; Dayan et al. 2000; Dayan and Abbot 2001; Dayan and Daw 2008; Niv and Montague 2008; Dayan 2012).

Connecting Levels: Computation, Behavior, and Neuronal Activity

An early sign of a strong connection of a TD error signal to a biological system was the connection of octopaminergic neuron activity to odorant conditioning in honeybees (Real 1992; for physiology, see Hammer 1993). Behaviorally, honeybees show many sophisticated computations in the way they sample flowers yielding variable returns. For two flowers yielding the same mean return (let's say in nectar units), bees will sample more frequently from the flower with lower variance in the nectar return despite the matched average return (Real 1992; Hammer and Menzel 1995). All things being equal, bees avoid the flower color that predicts the more variable return. This is just one of many sophisticated computations available to the honeybee for adjusting its behavior in pursuit of maximizing its returns on nectar. What is remarkable is the presence and functional role of aminergic neurons in the bee's subesophageal ganglion that project to widespread targets throughout the bee brain and deliver the neuromodulator octopamine (a close chemical cousin to dopamine). Without diving into too much detail, let me summarize by saying: (a) it is known now that octopamine action at target neural sites is crucial for conditioning, (b) the physiological behavior of these neurons is consistent with a TD error signal, and (c) this error signal can be mapped simply onto action choice by the bee in a manner (somewhat artificial though it may be) that can account straightforwardly for how the bee trades off mean and variance of nectar returns (see also Douglas 1995; Montague et al. 1995). The TD account was provocative because of the way it linked levels of description of the bee: the physiology of the octopamine neurons under a rigorous behavioral challenge, the behavioral consequences of the TD error computation putatively performed by these aminergic neurons, and the computations that reached from the level of the neurons to the choice behavior of the bee.

A richer connection between the TD algorithm (the computation), behavioral choice, and detailed recordings of neuronal behavior arose in the early 1990s from the work of Wolfram Schultz and his colleagues. In their work, nonhuman primates were trained on simple conditioning tasks where a light would indicate which of two levers were to be pushed to receive a juice reward. Early in the training, dopamine neurons give phasic responses only to the delivery of reward; these responses disappear, however, with training after which the neurons give transient responses only to the earliest consistent predictor of future reward (here the initial cue light). These data are now over twenty years old, but the guidance of the Sutton–Barto TD algorithm in understanding the features of these kinds of data is made clear by the then novel interpretation that the algorithm provided. These features include:

1. Temporal consistency was key and clearly being encoded into the response of the dopamine neurons. For the nearly identical behavioral paradigms, the neurons lose their initial response to reward delivery, develop a transient response to the earliest predictive cue, but differ in the way they respond at the trigger cue. When the timing of the trigger cue is completely predictable, neurons give no response; when the trigger cue has temporal uncertainty, neurons continue to modulate at the trigger. This is to be expected quite naturally from a TD error signal-based account; however, a trial-based account such as the Bush–Mosteller rule (and also of course the Rescorla–Wagner rule) would have to add extra detail to explain these data.
2. Sensory-reward prediction and sensory-sensory prediction are unified in a TD error-based account. The sensory-sensory prediction piece is exemplified by the response (or lack thereof) of the dopamine neurons to the trigger cue.

3. Dopamine neurons emit information even when they do not modulate their activity. This means, for example, that the neurons are emitting information throughout the duration of the trial. This is a new idea, certainly for the conditioning literature (though the weakly electric fish literature may be a good counterexample), and the lack of modulation during the interval between cues and reward outcome were vexing for the experimentalists who uncovered these data. In fact, this problem blocked their understanding of what the dopaminergic modulation could mean. As Schultz et al. (1993:900) stated: "None of the dopamine neurons showed sustained activity in the delay between the instruction and trigger stimuli that would resemble the activity of neurons in dopamine terminal areas, such as the striatum and frontal cortex.…The lack of sustained activity suggests that dopamine neurons do not encode representational processes, such as working memory, expectation of external stimuli or reward, or preparation of movement. Rather, dopamine neurons are involved with transient changes of impulse activity in basic attentional and motivational processes underlying learning and cognitive behavior."

Such delay period nonmodulation is to be expected from a TD error-based account of the results. However the conclusion is understandable. At the time (1992–1993), results using working memory tasks in nonhuman primates showed delay period activity in dopaminergic terminal regions of cortex (e.g., Brodman area 46; Goldman-Rakic et al. 2004) that depended on intact dopamine transmission through identified receptor types. It seems likely that this inspired Schultz and colleagues to look for delay period activity at the level of parent dopamine neurons in the midbrain that gave rise to this cortical input. The computational model (even the simplest Sutton–Barto TD error account) was essential to see these data in a different light. Furthermore, while trial-based accounts, such as the Rescorla–Wagner rule, have been offered to account for these physiological data, they cannot account for the signature temporal features. This point has been made clearly in a review of the dopamine prediction error hypothesis by Glimcher (2011). Glimcher also makes a very nice social case for the precedence and farsightedness of the Bush–Mosteller approach to modeling animal conditioning, when set beside the very popularly quoted Rescorla–Wagner rule.

Schultz and colleagues have now provided strong evidence for the TD error model and have shown (among a variety of new findings) that midbrain dopaminergic responses encode the expected value of the future predicted reward in similar experiments (see Schultz et al. 1997; Montague et al. 2004; Tobler et al. 2005; and the sober warnings of Dayan and Niv 2008). This summary suggests that mammals possess an efficient prediction system that can be deployed in a variety of behavioral settings to learn to value the near-term future and act reasonably based on those valuations. Here, I have avoided all discussion of

From "Computational Psychiatry: New Perspectives on Mental Illness,"
A. David Redish and Joshua A. Gordon, eds. 2016. Strüngmann Forum Reports, vol. 20,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03542-2.

the interesting complexities that arise when mapping such valuations to action choice and the further connections that can be made to optimizing control models (e.g., Kalman-filter models and models that require more complex representations). Let us finish here with some forward-looking pointers to the way that reinforcement-learning models can be applied to disease states like addiction (McClure et al. 2003a; Redish 2004), human neuroimaging data through model-based approaches to reinforcement learning (McClure et al. 2003b; O'Doherty et al. 2003, 2004), and to "exceptions" where evidence in rodents and nonhuman primates suggests that such models are incomplete (of course they are) or misleadingly wrong (Dayan and Niv 2008; Niv and Schoenbaum 2008). In my opinion, the big questions for reinforcement-learning models involve the nature of the representations used to control midbrain dopamine neurons and the use of these representations in cognitive control (Carter et al. 1998; Botvinick et al. 1999, 2001, 2009).

Reaching toward Humans

Thus far, this account has focused on the valuation aspect of the TD model and ignored the action-selection piece, except for simple conditioning paradigms which help highlight the main points. With the advent of functional MRI, many functional questions can now be asked that would probe directly the claims or extensions of the TD model. In the first direct test of the dopamine prediction error hypothesis in humans, McClure et al. (2003b) and O'Doherty et al. (2003) found that in terms of BOLD signals, the striatum shows activation and deactivation in accord with the TD model. These results are comforting but not definitive. There are many signals that may combine at the level of the striatum to elicit BOLD responses consistent with prediction error signals. Direct measures of dopamine and other neuromodulator delivery during similar behavioral probes will be required to expose the exact contributions of dopamine delivery to such BOLD measurements. One of the novelties opened up by the O'Doherty et al. and McClure et al. work is the possibility of using the computational models to define a computational process encoded throughout some behavioral challenge, and then use estimates of this process to seek its physical correlates. This approach is now called model-based fMRI.

Hypervaluation Disease

Another reach toward humans can be seen in the work of Redish (2004), who used TD-type models to address addiction as (in part) a valuation disease. The temporal-difference reinforcement-learning (TDRL) model uses an error, the TD error, putatively encoded by transient changes in dopaminergic activity (and presumably dopamine delivery) to adjust available parameters to estimate the expected value of discounted future rewards predicated on the current state of the animal. As expressed by Redish (2004):

$$V(t) = \int_t^\infty ds\, \gamma^{s-t} E[R(s)]. \tag{13.7}$$

This is just a continuous version of Equation 13.2, where the dependence on state is replaced with a time variable (in simple settings such an equivalence is fine but potentially confusing). As the animal learns to associate sensory cues with receipt of rewards, the dopamine systems adjusts its value function using the TD error and does so until the error is driven to 0. What if one could induce an error without going through the entire cue-predicts-future-reward machinery? Wouldn't such an "outside" error induce the system to learn the wrong valuation function? The answer is yes. The idea proposed by Redish is that addictive drugs, such as cocaine and methamphetamine, produce "bumps" in dopamine release (bumps in the TD error term) through mechanisms that escape the cue-predicts-reward setting captured by the TDRL model. In doing so, the system cannot learn a value function; temporal differences in this value function cancel the impact of cues associated with drug taking. In short, the proposal is a value function "run away" where uncompensable changes in the value function, induced by pharmacologically mimicked error signals, create a kind of valuation disease condition. In his analysis, Redish is careful to point out that this feature is only one of many aspects of addiction, but the entire proposal centers around the model and its mapping on physical substrates in the brain and aberrant behaviors that can ensue. This way of thinking has opened up many new questions in the area of addiction, and the models here have expanded immensely in recent years. Computationalizing addiction will certainly lead to new insight and likely better models of it.

Flat Valuation Diseases and Rational Freezing Responses

TDRL models also provide a new way to understand some aspects of Parkinson disease, a neurodegenerative disorder associated with a profound loss of midbrain dopamine neurons. By the time symptoms appear to warrant diagnosis, dopamine neuron loss ranges from 70–90%. There is virtually nothing known about how either dopamine systems or downstream targets adapt their dynamics as this loss occurs. However, one possible consequence of dropping the number of neurons is to increase the "dopamine noise" at the target structures. If dopamine delivery fluctuations are to communicate reward prediction errors in their rapid (subsecond) transients and possibly other important computations in their tonic (mesoscale) averages, then decreasing the number of neurons increases dopamine noise. A downstream target may not be able to distinguish the value of one state from another because the dopamine noise level is high. Ultimately this could look like a relative "flat" value function to these downstream targets.

Let us use this caricature and hypothesize that the best response to a flat value function is to commit no new resources—do nothing. In fact, why not just freeze? Perhaps one part of the syndrome of Parkinson disease is a kind of rational freezing response to a flat value function. In this sense, the small dopamine fluctuations are buried in noise. Therapies that would raise baseline dopamine (e.g., taking L-DOPA or perhaps putting engineered dopamine-secreting cells in the striatum) would make those fluctuations significant and perhaps "readable" by the otherwise confused downstream processes. This is, of course, wild speculation, but its possibilities are suggested by the model and thus support modern efforts in computational neuroscience.

Breaking the Reinforcement-Learning Hegemony?

This brief summary has been, in large part, rearward-looking and very focused on the valuation part of the simplest reinforcement-learning models. I close with apologies to those investigators whose work is not mentioned here: reinforcement-learning approaches are now so vast that it was not possible to include all relevant species of animal and explanation in this succinct summary. However, the best outcome for any class of model is to be very, very wrong in some productive way. There are already indications that reinforcement-learning models have guided work to some of these creaky zones (Dayan and Niv 2008; Gershman et al. 2009; Dayan 2012).

From "Computational Psychiatry: New Perspectives on Mental Illness,"
A. David Redish and Joshua A. Gordon, eds. 2016. Strüngmann Forum Reports, vol. 20,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03542-2.