MIT Open Access Articles

Toward an Integration of Deep Learning and Neuroscience

Massachusetts Institute of Technology

Toward an Integration of Deep Learning and Neuroscience

Adam H. Marblestone[1]*, Greg Wayne[2] and Konrad P. Kording[3]

[1] Synthetic Neurobiology Group, Massachusetts Institute of Technology, Media Lab, Cambridge, MA, USA, [2] Google Deepmind, London, UK, [3] Rehabilitation Institute of Chicago, Northwestern University, Chicago, IL, USA

Neuroscience has focused on the detailed implementation of computation, studying neural codes, dynamics and circuits. In machine learning, however, artificial neural networks tend to eschew precisely designed codes, dynamics or circuits in favor of brute force optimization of a cost function, often using simple and relatively uniform initial architectures. Two recent developments have emerged within machine learning that create an opportunity to connect these seemingly divergent perspectives. First, structured architectures are used, including dedicated systems for attention, recursion and various forms of short- and long-term memory storage. Second, cost functions and training procedures have become more complex and are varied across layers and over time. Here we think about the brain in terms of these ideas. We hypothesize that (1) the brain optimizes cost functions, (2) the cost functions are diverse and differ across brain locations and over development, and (3) optimization operates within a pre-structured architecture matched to the computational problems posed by behavior. In support of these hypotheses, we argue that a range of implementations of credit assignment through multiple layers of neurons are compatible with our current knowledge of neural circuitry, and that the brain's specialized systems can be interpreted as enabling efficient optimization for specific problem classes. Such a heterogeneously optimized system, enabled by a series of interacting cost functions, serves to make learning data-efficient and precisely targeted to the needs of the organism. We suggest directions by which neuroscience could seek to refine and test these hypotheses.

Keywords: cost functions, neural networks, neuroscience, cognitive architecture

1. INTRODUCTION

Machine learning and neuroscience speak different languages today. Brain science has discovered a dazzling array of brain areas (Solari and Stoner, 2011), cell types, molecules, cellular states, and mechanisms for computation and information storage. Machine learning, in contrast, has largely focused on instantiations of a single principle: function optimization. It has found that simple optimization objectives, like minimizing classification error, can lead to the formation of rich internal representations and powerful algorithmic capabilities in multilayer and recurrent networks (LeCun et al., 2015; Schmidhuber, 2015). Here we seek to connect these perspectives.

The artificial neural networks now prominent in machine learning were, of course, originally inspired by neuroscience (McCulloch and Pitts, 1943). While neuroscience has continued to play a role (Cox and Dean, 2014), many of the major developments were guided by insights into the

mathematics of efficient optimization, rather than neuroscientific findings (Sutskever and Martens, 2013). The field has advanced from simple linear systems (Minsky and Papert, 1972), to nonlinear networks (Haykin, 1994), to deep and recurrent networks (LeCun et al., 2015; Schmidhuber, 2015). Backpropagation of error (Werbos, 1974, 1982; Rumelhart et al., 1986) enabled neural networks to be trained efficiently, by providing an efficient means to compute the gradient with respect to the weights of a multi-layer network. Methods of training have improved to include momentum terms, better weight initializations, conjugate gradients and so forth, evolving to the current breed of networks optimized using batch-wise stochastic gradient descent. These developments have little obvious connection to neuroscience.

We will argue here, however, that neuroscience and machine learning are again ripe for convergence. Three aspects of machine learning are particularly important in the context of this paper. First, machine learning has focused on the optimization of cost functions (**Figure 1A**).

Second, recent work in machine learning has started to introduce complex cost functions, those that are not uniform across layers and time, and those that arise from interactions between different parts of a network. For example, introducing the objective of temporal coherence for lower layers (non-uniform cost function over space) improves feature learning (Sermanet and Kavukcuoglu, 2013), cost function schedules (non-uniform cost function over time) improve[1] generalization (Saxe et al., 2013; Goodfellow et al., 2014b; Gülçehre and Bengio, 2016) and adversarial networks—an example of a cost function arising from internal interactions— allow gradient-based training of generative models (Goodfellow et al., 2014a)[2]. Networks that are easier to train are being used to provide "hints" to help bootstrap the training of more powerful networks (Romero et al., 2014).

Third, machine learning has also begun to diversify the architectures that are subject to optimization. It has introduced simple memory cells with multiple persistent states (Hochreiter and Schmidhuber, 1997; Chung et al., 2014), more complex elementary units such as "capsules" and other structures (Delalleau and Bengio, 2011; Hinton et al., 2011; Tang et al., 2012; Livni et al., 2013), content addressable (Graves et al., 2014; Weston et al., 2014) and location addressable memories (Graves et al., 2014), as well as pointers (Kurach et al., 2015) and hard-coded arithmetic operations (Neelakantan et al., 2015).

These three ideas have, so far, not received much attention in neuroscience. We thus formulate these ideas as three hypotheses about the brain, examine evidence for them, and sketch how experiments could test them. But first, let us state the hypotheses more precisely.

---

[1]Hyper-parameter optimization shows that complicated schedules of training, which differ across parts of the network, lead to optimal performance (Maclaurin et al., 2015).

[2]In adversarial networks, a generator network is trained to fool a discriminator network into being unable to distinguish generated samples from real data samples, while the discriminator network is trained to prevent the generator network from fooling it in this way.

1.1. *Hypothesis 1* – The Brain Optimizes Cost Functions

The central hypothesis for linking the two fields is that biological systems, like many machine-learning systems, are able to optimize cost functions. The idea of cost functions means that neurons in a brain area can somehow change their properties, e.g., the properties of their synapses, so that they get better at doing whatever the cost function defines as their role. Human behavior sometimes approaches optimality in a domain, e.g., during movement (Körding, 2007), which suggests that the brain may have learned optimal strategies. Subjects minimize energy consumption of their movement system (Taylor and Faisal, 2011), and minimize risk and damage to their body, while maximizing financial and movement gains. Computationally, we now know that optimization of trajectories gives rise to elegant solutions for very complex motor tasks (Harris and Wolpert, 1998; Todorov and Jordan, 2002; Mordatch et al., 2012). We suggest that cost function optimization occurs much more generally in shaping the internal representations and processes used by the brain. Importantly, we also suggest that this requires the brain to have mechanisms for efficient credit assignment in multilayer and recurrent networks.
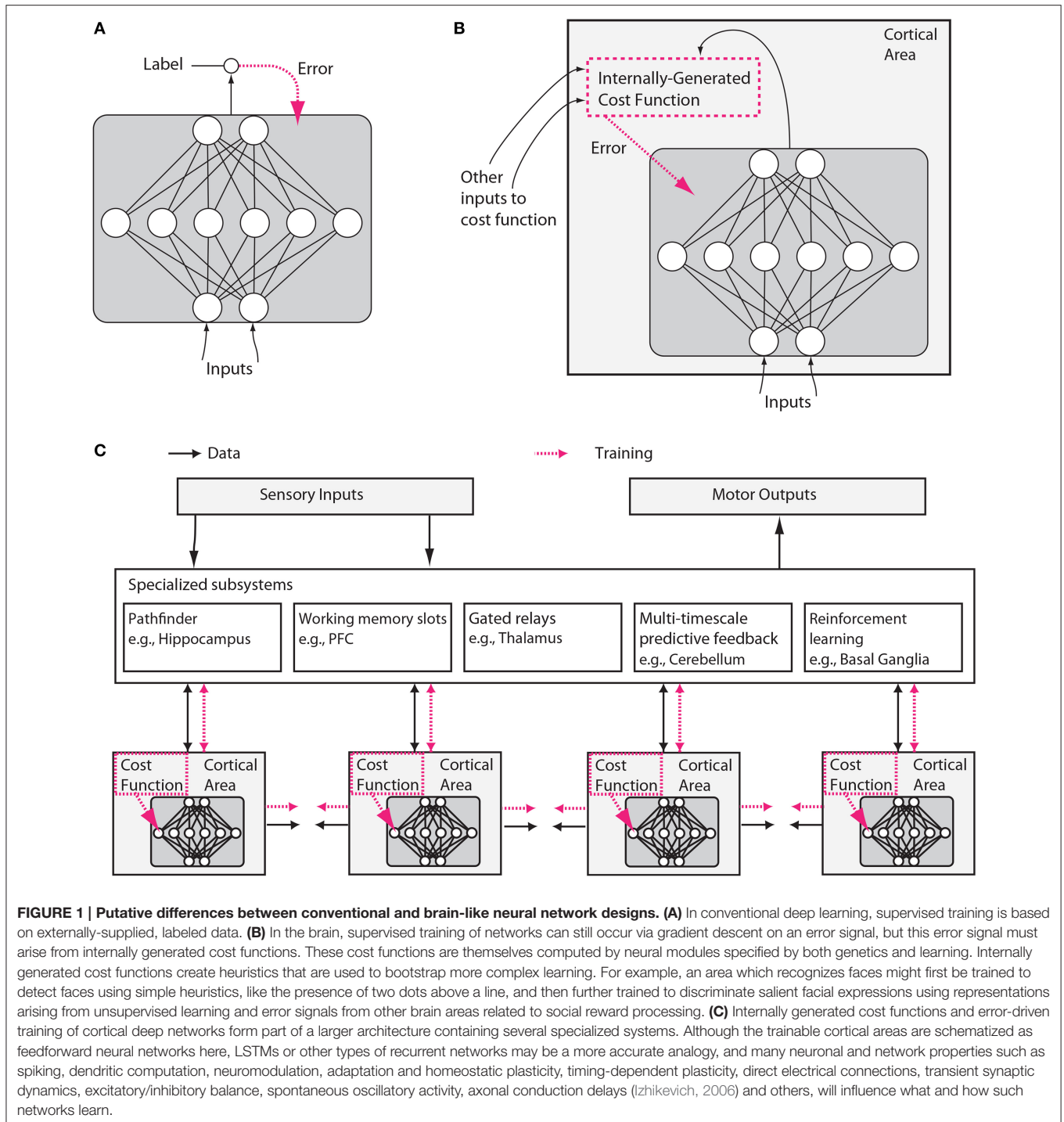
1.2. *Hypothesis 2* – Cost Functions Are Diverse across Areas and Change over Development

A second realization is that cost functions need not be global. Neurons in different brain areas may optimize different things, e.g., the mean squared error of movements, surprise in a visual stimulus, or the allocation of attention. Importantly, such a cost function could be locally generated. For example, neurons could locally evaluate the quality of their statistical model of their inputs (**Figure 1B**). Alternatively, cost functions for one area could be generated by another area. Moreover, cost functions may change over time, e.g., guiding young humans to understanding simple visual contrasts early on, and faces a bit later[3]. This could allow the developing brain to bootstrap more complex knowledge based on simpler knowledge. Cost functions in the brain are likely to be complex and to be arranged to vary across areas and over development.

1.3. *Hypothesis 3* – Specialized Systems Allow Efficient Solution of Key Computational Problems

A third realization is that structure matters. The patterns of information flow seem fundamentally different across brain areas, suggesting that they solve distinct computational problems. Some brain areas are highly recurrent, perhaps making them predestined for short-term memory storage (Wang, 2012). Some areas contain cell types that can switch between qualitatively different states of activation, such as a persistent firing mode vs. a transient firing mode, in response to particular neurotransmitters (Hasselmo, 2006). Other areas, like

---

[3]Psychologists have been quantifying the subtleties of many such developmental stagings, e.g., of our perceptual and motor performance, e.g., Nardini et al. (2010), Dekker and Nardini (2015), and McKone et al. (2009).

FIGURE 1 | Putative differences between conventional and brain-like neural network designs. (A) In conventional deep learning, supervised training is based on externally-supplied, labeled data. **(B)** In the brain, supervised training of networks can still occur via gradient descent on an error signal, but this error signal must arise from internally generated cost functions. These cost functions are themselves computed by neural modules specified by both genetics and learning. Internally generated cost functions create heuristics that are used to bootstrap more complex learning. For example, an area which recognizes faces might first be trained to detect faces using simple heuristics, like the presence of two dots above a line, and then further trained to discriminate salient facial expressions using representations arising from unsupervised learning and error signals from other brain areas related to social reward processing. **(C)** Internally generated cost functions and error-driven training of cortical deep networks form part of a larger architecture containing several specialized systems. Although the trainable cortical areas are schematized as feedforward neural networks here, LSTMs or other types of recurrent networks may be a more accurate analogy, and many neuronal and network properties such as spiking, dendritic computation, neuromodulation, adaptation and homeostatic plasticity, timing-dependent plasticity, direct electrical connections, transient synaptic dynamics, excitatory/inhibitory balance, spontaneous oscillatory activity, axonal conduction delays (Izhikevich, 2006) and others, will influence what and how such networks learn.

the thalamus appear to have the information from other areas flowing through them, perhaps allowing them to determine information routing (Sherman, 2005). Areas like the basal ganglia are involved in reinforcement learning and gating of discrete decisions (Doya, 1999; Sejnowski and Poizner, 2014). As every programmer knows, specialized algorithms matter for efficient solutions to computational problems, and the brain is likely to make good use of such specialization (**Figure 1C**).

These ideas are inspired by recent advances in machine learning, but we also propose that the brain has major differences from any of today's machine learning techniques. In particular, the world gives us a relatively limited amount of information that we could use for supervised learning (Fodor and Crowther, 2002). There is a huge amount of information available for unsupervised learning, but there is no reason to assume that a *generic* unsupervised algorithm, no matter how powerful, would

learn the precise things that humans need to know, in the order that they need to know it. The evolutionary challenge of making unsupervised learning solve the "right" problems is, therefore, to find a sequence of cost functions that will deterministically build circuits and behaviors according to prescribed developmental stages, so that in the end a relatively small amount of information suffices to produce the right behavior. For example, a developing duck imprints (Tinbergen, 1965) a template of its parent, and then uses that template to generate goal-targets that help it develop other skills like foraging.

Generalizing from this and from other studies (Minsky, 1977; Ullman et al., 2012), we propose that many of the brain's cost functions arise from such an internal bootstrapping process. Indeed, we propose that biological development and reinforcement learning can, in effect, program the emergence of a sequence of cost functions that precisely anticipates the future needs faced by the brain's internal subsystems, as well as by the organism as a whole. This type of developmentally programmed bootstrapping generates an internal infrastructure of cost functions which is diverse and complex, while simplifying the learning problems faced by the brain's internal processes. Beyond simple tasks like familial imprinting, this type of bootstrapping could extend to higher cognition, e.g., internally generated cost functions could train a developing brain to properly access its memory or to organize its actions in ways that will prove to be useful later on. The potential bootstrapping mechanisms that we will consider operate in the context of unsupervised and reinforcement learning, and go well beyond the types of curriculum learning ideas used in today's machine learning (Bengio et al., 2009).

In the rest of this paper, we will elaborate on these hypotheses. First, we will argue that both local and multi-layer optimization is, perhaps surprisingly, compatible with what we know about the brain. Second, we will argue that cost functions differ across brain areas and change over time and describe how cost functions interacting in an orchestrated way could allow bootstrapping of complex function. Third, we will list a broad set of specialized problems that need to be solved by neural computation, and the brain areas that have structure that seems to be matched to a particular computational problem. We then discuss some implications of the above hypotheses for research approaches in neuroscience and machine learning, and sketch a set of experiments to test these hypotheses. Finally, we discuss this architecture from the perspective of evolution.

2. THE BRAIN CAN OPTIMIZE COST FUNCTIONS

Much of machine learning is based on efficiently optimizing functions, and, as we will detail below, the ability to use backpropagation of error (Werbos, 1974; Rumelhart et al., 1986) to calculate gradients of arbitrary parametrized functions has been a key breakthrough. In Hypothesis 1, we claim that the brain is also, at least in part[4], an optimization machine.

But what exactly does it mean to say that the brain can optimize cost functions? After all, many processes can be viewed as optimizations. For example, the laws of physics are often viewed as minimizing an action functional, while evolution optimizes the fitness of replicators over a long timescale. To be clear, our main claims are: that (a) the brain has powerful mechanisms for credit assignment during learning that allow it to optimize global functions in multi-layer networks by adjusting the properties of each neuron to contribute to the global outcome, and that (b) the brain has mechanisms to specify exactly which cost functions it subjects its networks to, i.e., that the cost functions are highly tunable, shaped by evolution and matched to the animal's ethological needs. Thus, the brain uses cost functions as a key driving force of its development, much as modern machine learning systems do.

To understand the basis of these claims, we must now delve into the details of how the brain might efficiently perform credit assignment throughout large, multi-layered networks, in order to optimize complex functions. We argue that the brain uses several different types of optimization to solve distinct problems. In some structures, it may use genetic pre-specification of circuits for problems that require only limited learning based on data, or it may exploit local optimization to avoid the need to assign credit through many layers of neurons. It may also use a host of proposed circuit structures that would allow it to actually perform, in effect, backpropagation of errors through a multi-layer network, using biologically realistic mechanisms—a feat that had once been widely believed to be biologically implausible (Crick, 1989; Stork, 1989). Potential such mechanisms include circuits that literally backpropagate error derivatives in the manner of conventional backpropagation, as well as circuits that provide other efficient means of approximating the effects of backpropagation, i.e., of rapidly computing the approximate gradient of a cost function relative to any given connection weight in the network. Lastly, the brain may use algorithms that exploit specific aspects of neurophysiology—such as spike timing dependent plasticity, dendritic computation, local excitatory-inhibitory networks, or other properties—as well as the integrated nature of higher-level brain systems. Such mechanisms promise to allow learning capabilities that go even beyond those of current backpropagation networks.

---

[4]Our point in this section will not be that all learning in the brain can be captured by cost function optimization, but rather, somewhat more narrowly,

our claim is that the algorithms for optimization like backpropagation in deep learning may have correspondences in biological brains. We feel that it is an important task for neuroscience to determine whether and how brains implement these algorithms. The brain may also disclose dynamics that are unlike these algorithms, so we are not disclaiming the possibility of broader theories. In machine learning, many useful algorithms are not explicitly formulated as cost function optimization; for example, many algorithms are based on linear algebra procedures like singular value decomposition, rather than explicit optimization. Such methods can be made nonlinear by using nonlinear kernels—relatedly, some brain circuits run specialized computations using fixed nonlinear basis functions (e.g., in cerebellum). Moreover, while an implicit cost function can be attributed to account for many dynamical processes, as well as many popular learning algorithms, our claim is not merely that the brain uses other learning procedures that lead to solutions which implicitly minimize a cost function, but rather that it actually finds its solutions by performing a powerful form of optimization as such.

2.1. Local Self-organization and Optimization without Multi-layer Credit Assignment

Not all learning requires a general-purpose optimization mechanism like gradient descent[5]. Many theories of cortex (George and Hawkins, 2009; Kappel et al., 2014) emphasize potential self-organizing and unsupervised learning properties that may obviate the need for multi-layer backpropagation as such. Hebbian plasticity, which adjusts weights according to correlations in pre-synaptic and post-synaptic activity, is well established[6]. Various versions of Hebbian plasticity (Miller and MacKay, 1994), e.g., with nonlinearities (Brito and Gerstner, 2016), can give rise to different forms of correlation and competition between neurons, leading to the self-organized formation of ocular dominance columns, self-organizing maps and orientation columns (Miller et al., 1989; Ferster and Miller, 2000). Often these types of local self-organization can also be viewed as optimizing a cost function: for example, certain forms of Hebbian plasticity can be viewed as extracting the principal components of the input, which minimizes a reconstruction error (Pehlevan and Chklovskii, 2015).

To generate complex temporal patterns, the brain may also implement other forms of learning that do not require any equivalent of full backpropagation through a multilayer network. For example, "liquid-" (Maass et al., 2002) or "echo-state machines" (Jaeger and Haas, 2004) are randomly connected recurrent networks that form a basis set (also known as a "reservoir") of random filters, which can be harnessed for learning with tunable readout weights. Variants exhibiting chaotic, spontaneous dynamics can even be trained by feeding back readouts into the network and suppressing the chaotic activity (Sussillo and Abbott, 2009). Learning only the readout layer makes the optimization problem much simpler (indeed, equivalent to regression for supervised learning). Additionally, echo state networks can be trained by reinforcement learning as well as supervised learning (Bush, 2007; Hoerzer et al., 2014). Reservoirs of random nonlinear filters are one interpretation of the diverse, high-dimensional, mixed-selectivity tuning properties of many neurons, e.g., in the prefrontal cortex (Enel et al., 2016). Other variants of learning rules that modify only a fraction of the synapses inside a random network are being

developed as models of biological working memory and sequence generation (Rajan et al., 2016).

2.2. Biological Implementation of Optimization

We argue that the above mechanisms of local self-organization are likely insufficient to account for the brain's powerful learning performance (Brea and Gerstner, 2016). To elaborate on the need for an efficient means of gradient computation in the brain, we will first place backpropagation into its computational context (Hinton, 1989; Baldi and Sadowski, 2015). Then we will explain how the brain could plausibly implement approximations of gradient descent.

2.2.1. The Need for Efficient Gradient Descent in Multi-layer Networks

The simplest mechanism to perform cost function optimization is sometimes known as the "twiddle" algorithm or, more technically, as "serial perturbation." This mechanism works by perturbing (i.e., "twiddling"), with a small increment, a single weight in the network, and verifying improvement by measuring whether the cost function has decreased compared to the network's performance with the weight unperturbed. If improvement is noticeable, the perturbation is used as a direction of change to the weight; otherwise, the weight is changed in the opposite direction (or not changed at all). Serial perturbation is therefore a method of "coordinate descent" on the cost, but it is slow and requires global coordination: each synapse in turn is perturbed while others remain fixed.

Weight perturbation (or parallel perturbation) perturbs all of the weights in the network at once. It is able to optimize small networks to perform tasks but generally suffers from high variance. That is, the measurement of the gradient direction is noisy and changes drastically from perturbation to perturbation because a weight's influence on the cost is masked by the changes of all other weights, and there is only one scalar feedback signal indicating the change in the cost[7]. Weight perturbation is dramatically inefficient for large networks. In fact, parallel and serial perturbation learn at approximately the same rate if the time measure counts the number of times the network propagates information from input to output (Werfel et al., 2005).

Some efficiency gain can be achieved by perturbing neural activities instead of synaptic weights, acknowledging the fact that any long-range effect of a synapse is mediated through a neuron. Like weight perturbation and unlike serial perturbation, minimal global coordination is needed: each neuron only needs to receive a feedback signal indicating the global cost. The variance of node perturbation's gradient estimate is far smaller than that of weight perturbation under the assumptions that either all neurons or all weights, respectively, are perturbed and that they are perturbed at the same frequency. In this case, node perturbation's variance is proportional to the number of cells in the network, not the number of synapses.

---

[5]Of course, some circuits may also be heavily genetically pre-specified to minimize the burden on learning. For instance, particular cell adhesion molecules (Hattori et al., 2007) expressed on particular parts of particular neurons defined by a genetic cell type (Zeisel et al., 2015), and the detailed shapes and placements of neuronal arbors, may constrain connectivity in some cases, though in other cases local connectivity is thought to be only weakly constrained (Kalisman et al., 2005). Genetics is sufficient to specify complex circuits involving hundreds of neurons, such as central pattern generators (Yuste et al., 2005) which create complex self-stabilizing oscillations, or the entire nervous systems of small worms. Genetically guided wiring should not be thought of as fixed "hard-wiring" but rather as a programmatic construction process that can also accept external inputs and interact with learning mechanisms (Marcus, 2004).

[6]Hebbian plasticity even has a well-understood biological basis in the form of the NMDA receptors, which are activated by the simultaneous occurrence of chemical transmitter delivered from the pre-synaptic neuron, and voltage depolarization of the post-synaptic neuron.

[7]The variance can be mitigated by averaging out many perturbations before making a change to the baseline value of the weights, but this would take significant time for a network of non-trivial size as the variance of weight perturbation's estimates scales in proportion to the number of synapses in the network.

All of these approaches are slow either due to the time needed for serial iteration over all weights or the time needed for averaging over low signal-to-noise ratio gradient estimates. To their credit however, none of these approaches requires more than knowledge of local activities and the single global cost signal. Real neural circuits in the brain have mechanisms (e.g., diffusible neuromodulators) that appear to code the signals relevant to implementing those algorithms. In many cases, for example in reinforcement learning, the cost function, which is computed based on interaction with an unknown environment, cannot be differentiated directly, and an agent has no choice but to deploy clever twiddling to explore at some level of the system (Williams, 1992).

Backpropagation, in contrast, works by computing the sensitivity of the cost function to each weight based on the layered structure of the system. The derivatives of the cost function with respect to the last layer can be used to compute the derivatives of the cost function with respect to the penultimate layer, and so on, all the way down to the earliest layers[8]. Backpropagation can be computed rapidly, and for a single input-output pattern, it exhibits no variance in its gradient estimate. The backpropagated gradient has no more noise for a large system than for a small system, so deep and wide architectures with great computational power can be trained efficiently.

2.2.2. Biologically Plausible Approximations of Gradient Descent

To permit biological learning with efficiency approaching that of machine learning methods, some provision for more sophisticated gradient propagation may be suspected. Contrary to what was once a common assumption, there are now many proposed "biologically plausible" mechanisms by which a neural circuit could implement optimization algorithms that, like backpropagation, can efficiently make use of the gradient. These include Generalized Recirculation (O'Reilly, 1996), Contrastive Hebbian Learning (Xie and Seung, 2003), random feedback weights together with synaptic homeostasis (Lillicrap et al., 2014; Liao et al., 2015), spike timing dependent plasticity (STDP) with iterative inference and target propagation (Bengio et al., 2015a; Scellier and Bengio, 2016), complex neurons with backpropagating action-potentials (Körding and König, 2000), and others (Balduzzi et al., 2014). While these mechanisms differ in detail, they all invoke feedback connections that carry error phasically. Learning occurs by comparing a prediction with a target, and the prediction error is used to drive top-down changes in bottom-up activity.

As an example, consider O'Reilly's temporally eXtended Contrastive Attractor Learning (XCAL) algorithm (O'Reilly et al., 2012, 2014b). Suppose we have a multilayer neural network with an input layer, an output layer, and a set of hidden layers in between. O'Reilly showed that the same functionality as backpropagation can be implemented by a bidirectional network with the same weights but symmetric connections. After computing the outputs using the forward connections

---

[8]If the error derivatives of the cost function with respect to the last layer of unit activities are unknown, then they can be replaced with node-perturbation-like correlations, as is common in reinforcement learning.

only, we set the output neurons to the values they should have. The dynamics of the network then cause the hidden layers' activities to evolve toward a stable attractor state linking input to output. The XCAL algorithm performs a type of local modified Hebbian learning at each synapse in the network during this process (O'Reilly et al., 2012). The XCAL Hebbian learning rule compares the local synaptic activity (pre x post) during the early phase of this settling (before the attractor state is reached) to the final phase (once the attractor state has been reached), and adjusts the weights in a way that should make the early phase reflect the later phase more closely. These contrastive Hebbian learning methods even work when the connection weights are not precisely symmetric (O'Reilly, 1996). XCAL has been implemented in biologically plausible conductance-based neurons and basically implements the backpropagation of error approach.

Approximations to backpropagation could also be enabled by the millisecond-scale timing of of neural activities (O'Reilly et al., 2014b). Spike timing dependent plasticity (STDP) (Markram et al., 1997), for example, is a feature of some neurons in which the sign of the synaptic weight change depends on the precise millisecond-scale relative timing of pre-synaptic and post-synaptic spikes. This is conventionally interpreted as Hebbian plasticity that measures the potential for a causal relationship between the pre-synaptic and post-synaptic spikes: a pre-synaptic spike could have contributed to causing a post-synaptic spike only if it occurs shortly beforehand[9]. To enable a backpropagation mechanism, Hinton has suggested an alternative interpretation: that neurons could encode the types of error derivatives needed for backpropagation in the temporal derivatives of their firing rates (Hinton, 2007, 2016). STDP then corresponds to a learning rule that is sensitive to these error derivatives (Xie and Seung, 2000; Bengio et al., 2015b). In other words, in an appropriate network context, STDP learning could give rise to a biological implementation of backpropagation[10].

---

[9]Interestingly, STDP is not a unitary phenomenon, but rather a diverse collection of different rules with different timescales and temporal asymmetries (Sjöström and Gerstner, 2010; Mishra et al., 2016). Effects include STDP with the inverse temporal asymmetry, symmetric STDP and STDP with different temporal window sizes. STDP is also frequency dependent, which can be explained by rules that depend on triplets rather than pairs of spikes (Pfister and Gerstner, 2006). In some cortical neurons, STDP even switches its sign as the synapse moves away from the neuron's soma into the dendritic tree (Letzkus et al., 2006). While STDP is often included explicitly in models, biophysical derivations of STDP from various underlying phenomena are also being attempted, some of which involve the post-synaptic voltage (Clopath and Gerstner, 2010) or a local dendritic voltage (Urbanczik and Senn, 2014). Meanwhile, other theories suggest that STDP may enable the use of precise timing codes based on temporal coincidence of inputs, the generation and unsupervised learning of temporal sequences (Abbott and Blum, 1996; Fiete et al., 2010), enhancements to distal reward processing in reinforcement learning (Izhikevich, 2007), stabilization of neural responses (Kempter et al., 2001), or many other higher-level properties (Nessler et al., 2013; Kappel et al., 2014).
[10]Hinton has suggested (Hinton, 2007, 2016) that this could take place in the context of autoencoders and recirculation (Hinton and McClelland, 1988). Bengio and colleagues have proposed (Bengio, 2014; Bengio and Fischer, 2015; Scellier and Bengio, 2016) another context in which the connection between STDP and plasticity rules that depend on the temporal derivative of the post-synaptic firing rate can be exploited for biologically plausible multilayer credit assignment. This setting relies on clamping of outputs and stochastic relaxation in energy-based models (Ackley et al., 1958), which leads to a continuous network

Another possible mechanism, by which biological neural networks could approximate backpropagation, is "feedback alignment" (Lillicrap et al., 2014; Liao et al., 2015). There, the feedback pathway in backpropagation, by which error derivatives at a layer are computed from error derivatives at the subsequent layer, is replaced by a set of random feedback connections, with no dependence on the forward weights. Subject to the existence of a synaptic normalization mechanism and approximate sign-concordance between the feedforward and feedback connections (Liao et al., 2015), this mechanism of computing error derivatives works nearly as well as backpropagation on a variety of tasks. In effect, the forward weights are able to adapt to bring the network into a regime in which the random backwards weights actually carry the information that is useful for approximating the gradient. This is a remarkable and surprising finding, and is indicative of the fact that our understanding of gradient descent optimization, and specifically of the mechanisms by which backpropagation itself functions, are still incomplete. In neuroscience, meanwhile, we find feedback connections almost wherever we find feedforward connections, and their role is the subject of diverse theories (Callaway, 2004; Maass et al., 2007). It should be noted that feedback alignment as such does not specify exactly how neurons represent and make use of the error signals; it only relaxes a constraint on the transport of the error signals. Thus, feedback alignment is more a primitive that can be used in fully biological (approximate) implementations of backpropagation, than a fully biological implementation in its own right. As such, it may be possible to incorporate it into several of the other schemes discussed here.

The above "biological" implementations of backpropagation still lack some key aspects of biological realism. For example, in the brain, neurons tend to be either excitatory or inhibitory but not both, whereas in artificial neural networks a single neuron may send both excitatory and inhibitory signals to its downstream neurons. Fortunately, this constraint is unlikely to limit the functions that can be learned (Parisien et al., 2008; Tripp and Eliasmith, 2016). Other biological considerations, however, need to be looked at in more detail: the highly recurrent nature of biological neural networks, which show rich dynamics in time, and the fact that most neurons in mammalian brains communicate via spikes. We now consider these two issues in turn.

2.2.2.1. Temporal credit assignment:

The biological implementations of backpropagation proposed above, while applicable to feedforward networks, do not give a natural implementation of "backpropagation through time" (BPTT) (Werbos, 1990) for recurrent networks, which is widely used in machine learning for training recurrent networks on sequential processing tasks. BPTT "unfolds" a recurrent

network across multiple discrete time steps and then runs backpropagation on the unfolded network to assign credit to particular units at particular time steps[11]. While the network unfolding procedure of BPTT itself does not seem biologically plausible, to our intuition, it is unclear to what extent temporal credit assignment is truly needed (Ollivier and Charpiat, 2015) for learning particular temporally extended tasks.

If the system is given access to appropriate memory stores and representations (Buonomano and Merzenich, 1995; Gershman et al., 2012, 2014) of temporal context, this could potentially mitigate the need for temporal credit assignment as such—in effect, memory systems could "spatialize" the problem of temporal credit assignment[12]. For example, memory networks (Weston et al., 2014) store everything by default up to a certain buffer size, eliminating the need to perform credit assignment over the write-to-memory events, such that the network only needs to perform credit assignment over the read-from-memory events. In another example, certain network architectures that are superficially very deep, but which possess particular types of "skip connections," can actually be seen as ensembles of comparatively shallow networks (Veit et al., 2016); applied in the time domain, this could limit the need to propagate errors far backwards in time. Other, similar specializations or higher-levels of structure could, potentially, further ease the burden on credit assignment.

Can generic recurrent networks perform temporal credit assignment in in a way that is more biologically plausible than BPTT? Indeed, new discoveries are being made about the capacity for supervised learning in continuous-time recurrent networks with more realistic synapses and neural integration properties. In internal FORCE learning (Sussillo and Abbott, 2009), internally generated random fluctuations inside a chaotic recurrent network are adjusted to provide feedback signals that drive weight changes internal to the network while the outputs are clamped to desired patterns. This is made possible by a learning procedure that rapidly adjusts the network output to a state where it is close to the clamped values, and exerts continuous control to keep this difference small throughout the learning process[13]. This procedure is able to control and exploit the chaotic dynamical patterns that are spontaneously generated by the network.

Werbos has proposed in his "error critic" that an online approximation to BPTT can be achieved by learning to predict the backward-through-time gradient signal (costate) in a manner analogous to the prediction of value functions in reinforcement

---

dynamics (Hopfield, 1984) in which hidden units are perturbed toward target values (Bengio and Fischer, 2015), loosely similar to that which occurs in XCAL. This dynamics then allows the STDP-based rule to correspond to gradient descent on the energy function with respect to the weights (Scellier and Bengio, 2016). This scheme requires symmetric weights, but in an autoencoder context, Bengio notes that these can arise spontaneously (Arora et al., 2015).

[11]Even BPTT has arguably not been completely successful in recurrent networks. The problems of vanishing and exploding gradients led to long short term memory networks with gated memory units. An alternative is to use optimization methods that go beyond first order derivatives (Martens and Sutskever, 2011). This suggests the need for specialized systems and structures in the brain to mitigate problems of temporal credit assignment.

[12]Interestingly, the hippocampus seems to "time stamp" memories by encoding them into ensembles with cellular compositions and activity patterns that change gradually as a function of time on the scale of days (Rubin et al., 2015; Cai et al., 2016), and may use "time cells" to mark temporal positions within episodes on a timescale of seconds (Kraus et al., 2013).

[13]Control theory concepts also appear to be useful for simplifying optimization problems in certain other settings (Todorov, 2009; Hennequin et al., 2014).

learning (Werbos and Si, 2004). This kind of idea was recently applied in (Jaderberg et al., 2016) to allow decoupling of different parts of a network during training and to facilitate backpropagation through time. Broadly, we are only beginning to understand how neural activity can itself represent the time variable (Xu et al., 2014; Finnerty et al., 2015)[14], and how recurrent networks can learn to generate trajectories of population activity over time (Liu and Buonomano, 2009). Moreover, as we discuss below, a number of cortical models also propose means, other than BPTT, by which networks could be trained on sequential prediction tasks, even in an online fashion (O'Reilly et al., 2014b; Cui et al., 2015; Brea et al., 2016). A broad range of ideas can be used to approximate BPTT in more realistic ways.

2.2.2.2. *Spiking networks:*

It has been difficult to apply gradient descent learning directly to spiking neural networks[15, 16], although there do exist learning rules for doing so in specific representational contexts and network structures (Bekolay et al., 2013). A number of optimization procedures have been used to generate, indirectly, spiking networks which can perform complex tasks, by performing optimization on a continuous representation of the network dynamics and embedding variables into high-dimensional spaces with many spiking neurons representing each variable (Thalmeier et al., 2015; Abbott et al., 2016; DePasquale et al., 2016; Komer and Eliasmith, 2016). The use of recurrent connections with multiple timescales can remove the need for backpropagation in the direct training of spiking recurrent networks (Bourdoukan and Denève, 2015). Fast connections maintain the network in a state where slow connections have local access to a global error signal. While the biological realism of these methods is still unknown, they all allow connection weights to be learned in spiking networks.

These and other novel learning procedures illustrate the fact that we are only beginning to understand the connections between the temporal dynamics of biologically realistic networks, and mechanisms of temporal and spatial credit assignment. Nevertheless, we argue here that existing evidence suggests that biologically plausible neural networks can solve these problems— in other words, it is possible to efficiently optimize complex functions of temporal history in the context of spiking networks of biologically realistic neurons. In any case, there is little doubt that spiking recurrent networks using realistic population coding schemes can, with an appropriate choice of connection weights, compute complicated, cognitively relevant functions[17].

The question is how the developing brain efficiently learns such complex functions.

2.3. Other Principles for Biological Learning

The brain has mechanisms and structures that could support learning mechanisms different from typical gradient-based optimization algorithms employed in artificial neural networks.

2.3.1. Exploiting Biological Neural Mechanisms

The complex physiology of individual biological neurons may not only help explain how some form of efficient gradient descent could be implemented within the brain, but also could provide mechanisms for learning that go beyond backpropagation. This suggests that the brain may have discovered mechanisms of credit assignment quite different from those dreamt up by machine learning.

One such biological primitive is dendritic computation, which could impact prospects for learning algorithms in several ways. First, real neurons are highly nonlinear (Antic et al., 2010), with the dendrites of each *single* neuron implementing[18] something computationally similar to a three-layer neural network (Mel, 1992)[19]. Individual neurons thus should not be regarded as single "nodes" but as multi-component sub-networks. Second, when a neuron spikes, its action potential propagates back from the soma into the dendritic tree. However, it propagates more strongly into the branches of the dendritic tree that have been active (Williams and Stuart, 2000), potentially simplifying the problem of credit assignment (Körding and König, 2000). Third, neurons can have multiple somewhat independent dendritic compartments, as well as a somewhat independent somatic compartment, which means that the neuron should be thought of as storing more than one variable. Thus, there is the possibility for a neuron to store both its activation itself, and the error derivative of a cost function with respect to its activation, as required in backpropagation, and biological implementations of backpropagation based on this principle have been proposed (Körding and König, 2001; Schiess et al., 2016)[20]. Overall, the implications of dendritic computation for credit assignment in deep networks are only beginning to

---

[14]In one intriguing study of interval timing, single neurons exhibited response patterns over time which were scaled to the interval duration, and cooling the brain to slow down neural dynamics led to longer intervals being computed by the brain (Xu et al., 2014).

[15]Analogs of weight perturbation and node perturbation are known for spiking networks (Seung, 2003; Fiete and Seung, 2006). Seung (2003) also discusses implications of gradient based learning algorithms for neuroscience, echoing some of our considerations here.

[16]A related, but more general, question is how to learn over many layers of non-differentiable structures. One option is to perform updates via finite-sized rather than infinitesimal steps, e.g., via target-propagation (Bengio, 2014).

[17]Eliasmith and others have shown (Eliasmith and Anderson, 2004; Eliasmith et al., 2012; Eliasmith, 2013) that complex functions and control systems can be

compiled onto such networks, using nonlinear encoding and linear decoding of high-dimensional vectors.

[18]Dendritic computation may also have other functions, e.g., competitive interactions between dendrites in a single neuron could also allow neurons to contribute to multiple different ensembles (Legenstein and Maass, 2011).

[19]Localized activity in dendrites drives localized plasticity, with inhibitory interneurons, and interactions between inputs at different parts of the dendritic tree, controlling the local sign and spatial distribution of this plasticity (Sjöström and Häusser, 2006; Cichon and Gan, 2015).

[20]In the model of Körding and König (2001), single spikes are used to transmit activations and burst spikes are used to transmit error information. In other models, including the dendritic voltage in a plasticity rule leads to error-driven and predictive learning that can approximate backpropagation *inside* a single complex neuron (in effect backpropagating from the net somatic output, through nonlinearities at the dendritic branch points, all the way back to the individual input synaptic weights) and that generalize to a reinforcement learning context (Urbanczik and Senn, 2014; Schiess et al., 2016). Single neurons with active dendrites and many synapses may also embody learning rules of greater complexity, such as the storage and recall of temporal patterns (Hawkins and Ahmad, 2016).

be considered[21]. But it is clear that the types of bi-directional, non-linear, multi-variate interactions that are possible *inside* a single neuron could support gradient descent learning or other powerful optimization mechanisms.

Beyond dendritic computation, diverse mechanisms (Marblestone and Boyden, 2014) like retrograde (post-synaptic to pre-synaptic) signals using cannabinoids (Wilson and Nicoll, 2001), or rapidly-diffusing gases such as nitric oxide (Arancio et al., 1996), are among many that could enable learning rules that go beyond conventional conceptions of backpropagation. Harris has suggested (Harris, 2008; Lewis and Harris, 2014) how slow, retroaxonal (i.e., from the outgoing synapses back to the parent cell body) transport of molecules like neurotrophins could allow neural networks to implement an analog of an exchangeable currency in economics, allowing networks to self-organize to efficiently provide information to downstream "consumer" neurons that are trained via faster and more direct error signals. The existence of these diverse mechanisms may call into question traditional, intuitive notions of "biological plausibility" for learning algorithms.

Another potentially important biological primitive is neuromodulation. The same neuron or circuit can exhibit different input-output responses and plasticity depending on a global circuit state, as reflected by the concentrations of various *neuromodulators* like dopamine, serotonin, norepinephrine, acetylcholine, and hundreds of different neuropeptides such as opiods (Bargmann, 2012; Bargmann and Marder, 2013). These modulators interact in complex and cell-type-specific ways to influence circuit function. Interactions with glial cells also play a role in neural signaling and neuromodulation, leading to the concept of "tripartite" synapses that include a glial contribution (Perea et al., 2009). Modulation could have many implications for learning. First, modulators can be used to gate synaptic plasticity on and off selectively in different areas and at different times, allowing precise, rapidly updated orchestration of where and when cost functions are applied. Furthermore, it has been argued that a single neural circuit can be thought of as multiple overlapping circuits with modulation switching between them (Bargmann, 2012; Bargmann and Marder, 2013). In a learning context, this could potentially allow sharing of synaptic weight information between overlapping circuits. Dayan (2012) discusses further computational aspects of neuromodulation. Overall, neuromodulation seems to expand the range of possible algorithms that could be used for optimization.

2.3.2. Learning in the Cortical Sheet

A number of models attempt to explain cortical learning on the basis of specific architectural features of the 6-layered cortical sheet. These models generally agree that a primary function of the cortex is some form of unsupervised learning via prediction (O'Reilly et al., 2014b; Brea et al., 2016)[22].

Some cortical learning models are explicit attempts to map cortical structure onto the framework of message-passing algorithms for Bayesian inference (Lee and Mumford, 2003; Dean, 2005; George and Hawkins, 2009), while others start with particular aspects of cortical neurophysiology and seek to explain those in terms of a learning function, or in terms of a computational function, e.g., hierarchical clustering (Rodriguez et al., 2004). For example, the nonlinear and dynamical properties of cortical pyramidal neurons—the principal excitatory neuron type in cortex (Shepherd, 2014)—are of particular interest here, especially because these neurons have multiple dendritic zones that are targeted by different kinds of projections, which may allow the pyramidal neuron to make comparisons of top-down and bottom-up inputs[23].

Other aspects of the laminar cortical architecture could be crucial to how the brain implements learning. Local inhibitory neurons targeting particular dendritic compartments of the L5

---

[22] multi-layer perceptrons trained to minimize a prediction error by gradient descent. Within each unit, predictive autoencoders form a data compression in their middle layers, which is then fed up to higher levels as well as laterally. This system is suggestive of the power of using modular units of intermediate complexity, each of which minimizes a prediction error locally, e.g., in a local few-layer network. The system currently uses a fixed format for transmission of vectors from one unit to another, but ideally the inter-module connections should also be trained by gradient descent as well or by reinforcement learning rather than being fixed. The cortical-column-like modules could also be made more complex and could be organized into higher-order structures like Minsky's semantic networks, frames and K-lines (Minsky, 1988) rather than in simple hierarchies, or such an architecture could self-organize via reinforcement learning or other mechanisms for defining inter-column connections. Such a system also needs connections with specific kinds of memory and long-range information routing systems.

[23] This idea has been used by Hawkins and colleagues to suggest mechanisms for continuous online sequence learning (Cui et al., 2015; Hawkins and Ahmad, 2016) and by Larkum and colleagues for comparison of top-down and bottom-up signals (Larkum, 2013). The Larkum model focuses on the layer 5 (L5) pyramidal neuron type. The cell body of this neuron lies in L5 but extends its "apical" dendritic tree all the way up to a tuft at the top of the cortex in layer 1 (L1), which is a primary target of feedback projections. In the model, interactions between local spiking in these different dendritic zones, which are targeted by different kinds of projections, are crucial to the learning function. The model of Hawkins (Cui et al., 2015; Hawkins and Ahmad, 2016) also focused on the unique dendritic structure of the L5 pyramidal neuron, and distinguishes internal states of the neuron, which impact its responsiveness to other inputs, from activation states, which directly translate into spike rates. Three integration zones in each neuron, and dendritic NMDA spikes (Palmer et al., 2014) acting as local coincidence detectors (Shai et al., 2015), allow temporal patterns of dendritic input to impact the cell's internal state. Intra-column inhibition is also used in this model. Other cortical models pay less attention to the details of dendritic computation, but still provide detailed interpretations of the inter-laminar projection patterns of the neocortex. For example, in O'Reilly et al. (2014b), an architecture is presented for continuous learning based on prediction of the next input. Time is discretized into 100 ms bins via an alpha oscillation, and the deep vs. shallow layers maintain different information during these time bins, with deep layers maintaining a record of the previous time step, and shallow layers representing the current state. The stored information in the deep layers leads to a prediction of the current state, which is then compared with the actual current state. Periodic bursting locked to the oscillation provides a kind of clock that causes the current state to be shifted into the deep layers for maintenance during the subsequent time step, and recurrent loops with the thalamus allow this representation to remain stable for sufficiently long to be used to generate the prediction. Other theories utilize the biophysics of dendritic computation and spike timing dependent plasticity to explain how neurons could learn to make predictions (Brea et al., 2016) on a timescale of seconds using neurons with intrinsic plasticity time constants of a few tens of milliseconds.

---

[21] Interestingly, some connectomic studies are finding more obvious connectivity structure at the level of dendritic organization than at the cellular level (Morgan et al., 2016).

[22] An interesting recent study explored this idea in the context of a model of modular cortical-column-like units (Piekniewski et al., 2016). Local units are

pyramidal could be used to exert precise control over when and how the relevant feedback signals and associative mechanisms are utilized. Notably, local inhibitory networks could also give rise to competition (Petrov et al., 2010) between different representations in the cortex, perhaps allowing one cortical column to suppress others nearby, or perhaps even to send more sophisticated messages to gate the state transitions of its neighbors (Bach and Herger, 2015). Moreover, recurrent connectivity with the thalamus, structured bursts of spiking, and cortical oscillations (not to mention other mechanisms like neuromodulation) could control the storage of information over time, to facilitate learning based on temporal prediction. These concepts begin to suggest preliminary, exploratory models for how the detailed anatomy and physiology of the cortex could be interpreted within a machine-learning framework that goes beyond backpropagation. But these are early days: we still lack detailed structural/molecular and functional maps of even a single local cortical microcircuit.

2.3.3. One-shot Learning

Human learning is often one-shot: it can take just a single exposure to a stimulus to never forget it, as well as to generalize from it to new examples. One way of allowing networks to have such properties is what is described by I-theory, in the context of learning invariant representations for object recognition (Anselmi et al., 2015). Instead of training via gradient descent, image templates are stored in the weights of simple-complex cell networks while objects undergo transformations, similar to the use of stored templates in HMAX (Serre et al., 2007). The theories then aim to show that you can invariantly and discriminatively represent objects using a single sample, even of a new class (Anselmi et al., 2015)[24].

Additionally, the nervous system may have a way of quickly storing and replaying sequences of events. This would allow the brain to move an item from episodic memory into a long-term memory stored in the weights of a cortical network (Ji and Wilson, 2007), by replaying the memory over and over. This solution effectively uses many iterations of weight updating to fully learn a single item, even if one has only been exposed to it once. Alternatively, the brain could rapidly store an episodic memory and then retrieve it later without the need to perform slow gradient updates, which has proven to be useful for fast reinforcement learning in scenarios with limited available data (Blundell et al., 2016).

Finally, higher-level systems in the brain may be able to implement Bayesian learning of sequential programs, which is a powerful means of one-shot learning (Lake et al., 2015). This type of cognition likely relies on an interaction between multiple brain areas such as the prefrontal cortex and basal ganglia.

These potential substrates of one-shot learning rely on mechanisms other than simple gradient descent. It should be noted, though, that recent architectural advances, including specialized spatial attention and feedback mechanisms (Rezende et al., 2016), as well as specialized memory mechanisms (Santoro et al., 2016), do allow some types of one-shot generalization to be driven by backpropagation-based learning.

2.3.4. Active Learning

Human learning is often active and deliberate. It seems likely that, in human learning, actions are chosen so as to generate interesting training examples, and sometimes also to test specific hypotheses. Such ideas of active learning and "child as scientist" go back to Piaget and have been elaborated more recently (Gopnik et al., 2000). We want our learning to be based on maximally informative samples, and active querying of the environment (or of internal subsystems) provides a way route to this.

At some level of organization, of course, it would seem useful for a learning system to develop explicit representations of its uncertainty, since this can be used to guide the system to actively seek the information that would reduce its uncertainty most quickly. Moreover, there are population coding mechanisms that could support explicit probabilistic computations (Zemel and Dayan, 1997; Sahani and Dayan, 2003; Rao, 2004; Ma et al., 2006; Eliasmith and Martens, 2011; Gershman and Beck, 2016). Yet it is unclear to what extent and at what levels the brain uses an explicitly probabilistic framework, or to what extent probabilistic computations are emergent from other learning processes (Orhan and Ma, 2016)[25, 26].

---

[24] I-theory can perhaps be viewed as a generalized alternative paradigm to the online optimization of cost functions via multi-layer gradient descent, as used in deep learning. It exploits similar network architectures as conventional deep learning, e.g., hierarchical convolutional networks for the case of feedforward vision, but rather than backpropagating errors, it uses local circuits and learning rules to store templates against which new inputs are compared. This relies on a theory of generalization in learning based on combinations of tuned units (Poggio and Bizzi, 2004), which has been applied to both vision and motor control. Neurons with the required Gaussian-like tunings to stored templates could be obtained through canonical, local, normalization-based circuits (Kouh and Poggio, 2008), which can also be tweaked to implement other aspects of a vision architecture like softmax operations and pooling.

[25] One alternative picture that contrasts with straightforward cost function optimization emphasizes the types of computation that appear most naturally suited to heterogeneous, stochastic, noisy, continually changing neural circuitry (Maass, 2016). On this view, network plasticity is viewed as a sampling-based approximation to Bayesian inference (Kappel et al., 2015) where transiently changing synapses sample from a posterior distribution of network configurations, rather than as gradient descent on a cost function. This view emphasizes Monte-Carlo sampling procedures, rather than cost function optimization.

[26] Sampling based inference procedures are used widely in Bayesian statistics, and efforts have been made to connect these procedures with circuit-based models of computations (Mansinghka and Jonas, 2014). It currently appears difficult, however, to reconcile generic Marcov Chain Monte Carlo (MCMC) dynamics, which mix slowly, with the fast time scales of human psychophysics. But Bayesian methods are powerful and come with a methodology for model comparison (Ghahramani, 2005). In machine learning, variational Bayesian methods have recently become popular precisely because they are capable of fast though approximate posterior inference (inferring causes from observables), but seem to be powerful enough to create strong models. For example, stochastic gradient descent optimization is beginning to be used for variational Bayesian inference (Kingma and Welling, 2013). Restricted Boltzmann Machines (RBMs) also achieve fast inference in shallow architectures—with only a small number of iterations of mixing required—but they do not mix quickly when stacked into deep hierarchies as deep Boltzmann machines. The greedy, layer-wise pre-training of a deep belief network (Hinton et al., 2006) provides a heuristic way to stack the RBMs by auto-encoding, but these have achieved less competitive results than current variational Bayesian models. The problem of fast inference in MCMC models is the subject of current research, including at the interface with biologically

Standard gradient descent does not incorporate any such adaptive sampling mechanism, e.g., it does not deliberately sample data so as to maximally reduce its uncertainty. Interestingly, however, stochastic gradient descent can be used to generate a system that samples adaptively (Alain et al., 2015; Bouchard et al., 2015). In other words, a system can learn, by gradient descent, how to choose its own input data samples in order to learn most quickly from them by gradient descent.

Ideally, the learner learns to choose actions that will lead to the largest improvements in its prediction or data compression performance (Schmidhuber, 2010). In Schmidhuber (2010), this is done in the framework of reinforcement learning, and incorporates a mechanisms for the system to measure its own rate of learning. In other words, it is possible to reinforcement-learn a policy for selecting the most interesting inputs to drive learning. Adaptive sampling methods are also known in reinforcement learning that can achieve optimal Bayesian exploration of Markov Decision Process environments (Sun et al., 2011; Guez et al., 2012).

These approaches achieve optimality in an arbitrary, abstract environment. But of course, evolution may also encode its implicit knowledge of the organism's natural environment, the behavioral goals of the organism, and the developmental stages and processes which occur inside the organism, as priors or heuristics[27] which would further constrain the types of adaptive sampling that are optimal in practice. For example, simple heuristics like seeking certain perceptual signatures of novelty, or more complex heuristics like monitoring situations that other people seem to find interesting, might be good ways to bias sampling of the environment so as to learn more quickly. Other such heuristics might be used to give internal brain systems the types of training data that will be most useful to those particular systems at any given developmental stage.

We are only beginning to understand how active learning might be implemented in the brain. We speculate that multiple mechanisms, specialized to different brain systems and spatio-temporal scales, could be involved. The above examples suggest that at least some such mechanisms could be understood from the perspective of optimizing cost functions.

2.4. Differing Biological Requirements for Supervised and Reinforcement Learning

We have suggested ways in which the brain could implement learning mechanisms of comparable power to backpropagation. But in many cases, the system may be more limited by the available training signals than by the optimization process itself. In machine learning, one distinguishes supervised learning, reinforcement learning and unsupervised learning, and the training data limitation manifests differently in each case.

Both supervised and reinforcement learning require some form of teaching signal, but the nature of the teaching signal

in supervised learning is different from that in reinforcement learning. In supervised learning, the trainer provides the entire vector of errors for the output layer and these are back-propagated to compute the gradient: a locally optimal direction in which to update all of the weights of a potentially multi-layer and/or recurrent network. In reinforcement learning, however, the trainer provides a scalar evaluation signal, but this is not sufficient to derive a low-variance gradient. Hence, some form of trial and error twiddling must be used to discover how to increase the evaluation signal. Consequently, reinforcement learning is generally much less efficient than supervised learning.

Reinforcement learning in shallow networks is simple to implement biologically. For reinforcement learning of a deep network to be biologically plausible, however, we need a more powerful learning mechanism, since we are learning based on a more limited evaluation signal than in the supervised case: we do not have the full target pattern to train toward. Nevertheless, approximations of gradient descent can be achieved in this case, and there are cases in which the scalar evaluation signal of reinforcement learning can be used to efficiently update a multi-layer network by gradient descent. The "attention-gated reinforcement learning" (AGREL) networks of Stanisor et al. (2013), Brosch et al. (2015), and Roelfsema and van Ooyen (2005), and variants like KickBack (Balduzzi, 2014), give a way to compute an approximation to the full gradient in a reinforcement learning context using a feedback-based attention mechanism for credit assignment within the multi-layer network. The feedback pathway, together with a diffusible reward signal, together gate plasticity. For networks with more than three layers, this gives rise to a model based on columns containing parallel feedforward and feedback pathways (Roelfsema and van Ooyen, 2005), and for recurrent networks that settle into attractor states it gives a reinforcement-trained version (Brosch et al., 2015) of the Almeida/Pineda recurrent backpropagation algorithm (Pineda, 1987). The process is still not as efficient or generic as backpropagation, but it seems that this form of feedback can make reinforcement learning in multi-layer networks more efficient than a naive node perturbation or weight perturbation approach.

The machine-learning field has recently been tackling the question of credit assignment in deep reinforcement learning. Deep Q-learning (Mnih et al., 2015) demonstrates reinforcement learning in a deep network, wherein most of the network is trained via backpropagation. In regular Q learning, we define a function Q, which estimates the best possible sum of future rewards (the return) if we are in a given state and take a given action. In deep Q learning, this function is approximated by a neural network that, in effect, estimates action-dependent returns in a given state. The network is trained using backpropagation of local errors in Q estimation, using the fact that the return decomposes into the current reward plus the discounted estimate of future return at the next moment. During training, as the agent acts in the environment, a series of loss functions is generated at each step, defining target patterns that can be used as the supervision signal for backpropagation. As Q is a highly nonlinear function of the state, tricks are needed to make deep Q learning efficient and stable, including experience replay and

---

plausible models (Bengio et al., 2016). When these models are made to perform fast inference, they actually become somewhat similar to variational Bayesian methods, since they rely on feedforward approximate inference, at least to initialize the system.

[27]Heuristics are widely used to simplify motor planning and control, e.g., McLeod and Dienes (1996).

a particular type of mini-batch training. It is also necessary to store the outputs from the previous iteration (or clone the entire network) in evaluating the loss function for the subsequent iteration[28].

This process for generating learning targets provides a kind of bridge between reinforcement learning and efficient backpropagation-based gradient descent learning[29]. Importantly, only temporally local information is needed making the approach relatively compatible with what we know about the nervous system.

Even given these advances, a key remaining issue in reinforcement learning is the problem of long timescales, e.g., learning the many small steps needed to navigate from London to Chicago. Many of the formal guarantees of reinforcement learning (Williams and Baird, 1993), for example, suggest that the difference between an optimal policy and the learned policy becomes increasingly loose as the discount factor shifts to take into account reward at longer timescales. Although the degree of optimality of human behavior is unknown, people routinely engage in adaptive behaviors that can take hours or longer to carry out, by using specialized processes like *prospective memory* to "remember to remember" relevant variables at the right times, permitting extremely long timescales of coherent action. Machine learning has not yet developed methods to deal with such a wide range of timescales and scopes of hierarchical action. Below we discuss ideas of hierarchical reinforcement learning that may make use of callable procedures and sub-routines, rather than operating explicitly in a time domain.

As we will discuss below, some form of deep reinforcement learning may be used by the brain for purposes beyond optimizing global rewards, including the training of local networks based on diverse internally generated cost functions. Scalar reinforcement-like signals are easy to compute, and easy to deliver to other areas, making them attractive mechanistically. If the brain does employ internally computed scalar reward-like signals as a basis for cost functions, it seems likely that it will have found an efficient means of reinforcement-based training of deep networks, but it is an open question whether an analog of deep Q networks, AGREL, or some other mechanism entirely, is used in the brain for this purpose. Moreover, as we will discuss further below, it is possible that reinforcement-type learning is made more efficient in the context of specialized brain systems like short term memories, replay mechanisms, and hierarchically organized control systems. These specialized systems could reduce reliance on a need for powerful credit assignment

---

[28]Many other reinforcement learning algorithms, including REINFORCE (Williams, 1992), can be implemented as fully online algorithms using "eligibility traces," which accumulate the sensitivity of action distributions to parameters in a temporally local manner (Sutton and Barto, 1998).

[29]Zaremba and Sutskever (2015) also bridges reinforcement learning and backpropagation learning in the same system, in the context of a neural network controlling discrete interfaces, and illustrates some of the challenges of this approach: compared to an end-to-end backpropagation-trained Neural Turing Machine (Graves et al., 2014), reinforcement based training allows training of only relatively simple algorithmic tasks. Special measures need to be taken to make reinforcement efficient, including limiting the number of possible actions, subtracting a baseline reward, and training the network using a curriculum schedule.

mechanisms for reinforcement learning. Finally, if the brain uses a diversity of scalar reward-like signals to implement different cost functions, then it may need to mediate delivery of those signals via a comparable diversity of molecular substrates. The great diversity of neuromodulatory signals, e.g., neuropeptides, in the brain (Bargmann, 2012; Bargmann and Marder, 2013) makes such diversity quite plausible, and moreover, the brain may have found other, as yet unknown, mechanisms of diversifying reward-like signaling pathways and enabling them to act independently of one another.

3. THE COST FUNCTIONS ARE DIVERSE ACROSS BRAIN AREAS AND TIME

In the last section, we argued that the brain can optimize functions. This raises the question of what functions it optimizes. Of course, in the brain, a cost function will itself be created (explicitly or implicitly) by a neural network shaped by the genome. Thus, the cost function used to train a given sub-network in the brain is a key innate property that can be built into the system by evolution. It may be much cheaper in biological terms to specify a cost function that allows the rapid learning of the solution to a problem than to specify the solution itself.

In Hypothesis 2, we proposed that the brain optimizes not a single "end-to-end" cost function, but rather a diversity of internally generated cost functions specific to particular brain functions[30]. To understand how and why the brain may use a diversity of cost functions, it is important to distinguish the differing types of cost functions that would be needed for supervised, unsupervised and reinforcement learning. We can also seek to identify types of cost functions that the brain may need to generate from a functional perspective, and how each may be implemented as supervised, unsupervised, reinforcement-based or hybrid systems.

3.1. How Cost Functions May Be Represented and Applied

What additional circuitry is required to actually impose a cost function on an optimizing network? In the most familiar case, supervised learning may rely on computing a vector of errors at the output of a network, which will rely on some comparator circuitry[31] to compute the difference between the network outputs and the target values. This difference could then be backpropagated to earlier layers. An alternative way to impose a cost function is to "clamp" the output of the network, forcing it to occupy a desired target state. Such clamping is actually assumed in some of the putative biological implementations of backpropagation described above, such as XCAL and target propagation. Alternatively, as described above, scalar reinforcement signals are attractive as internally-computed cost functions, but using them in deep networks requires special mechanisms for credit assignment.

---

[30]This is distinct from a game-theoretic scenario in which multiple actors can achieve an equilibrium, e.g., Gemp and Mahadevan (2015).

[31]Single neurons act as comparators in the motor system, e.g., Brownstone et al. (2015), and networks in the retina adapt so as to report local differences in space or time rather than absolute values, a form of predictive coding (Hosoya et al., 2005).

In unsupervised learning, cost functions may not take the form of externally supplied training or error signals, but rather can be built into the dynamics inherent to the network itself, i.e., there may be no need for a *separate* circuit to compute and impose a cost function on the network. For example, specific spike-timing-dependent and homeostatic plasticity rules have been shown to give rise to gradient descent on a prediction error in recurrent neural networks (Galtier and Wainrib, 2013). Thus, specific unsupervised objectives could be implemented implicitly through specific local network dynamics[32] and plasticity rules inside a network without explicit computation of cost function, nor explicit propagation of error derivatives.

Alternatively, explicit cost functions could be computed, delivered to an optimizing network, and used for unsupervised learning, following a variety of principles being discovered in machine learning (e.g., Radford et al., 2015; Lotter et al., 2015). These networks rely on backpropagation as the sole learning rule, and typically find a way to encode the desired cost function into the error derivatives which are backpropagated. For example, prediction errors naturally give rise to error signals for unsupervised learning, as do reconstruction errors in autoencoders, and these error signals can also be augmented with additional penalty or regularization terms that enforce objectives like sparsity or continuity, as described below. Then these error derivatives can be propagated throughout the network via standard backpropagation. In such systems, the objective function and the optimization mechanism can thus be mixed and matched modularly. In the next sections, we elaborate on these and other means of specifying and delivering cost functions in different learning contexts.

3.2. Cost Functions for Unsupervised Learning

There are many objectives that can be optimized in an unsupervised context, to accomplish different kinds of functions or guide a network to form particular kinds of representations.

3.2.1. Matching the Statistics of the Input Data Using Generative Models

In one common form of unsupervised learning, higher brain areas attempt to produce samples that are statistically similar to those actually seen in lower layers. For example, the wake-sleep algorithm (Hinton et al., 1995) requires the sleep mode to sample potential data points whose distribution should then match the observed distribution. Unsupervised pre-training of deep networks is an instance of this (Erhan

and Manzagol, 2009), typically making use of a stacked auto-encoder framework. Similarly, in target propagation (Bengio, 2014), a top-down circuit, together with lateral information, has to produce data that directs the local learning of a bottom-up circuit and vice-versa. Ladder autoencoders make use of lateral connections and local noise injection to introduce an unsupervised cost function, based on internal reconstructions, that can be readily combined with supervised cost functions defined on the networks top layer outputs (Valpola, 2015). Compositional generative models generate a scene from discrete combinations of template parts and their transformations (Wang and Yuille, 2014), in effect performing a rendering of a scene based on its structural description. Hinton and colleagues have also proposed cortical "capsules" (Hinton et al., 2011; Tang et al., 2012, 2013) for compositional inverse rendering. The network can thus implement a statistical goal that embodies some understanding of the way that the world produces samples[33].

Learning rules for generative models have historically involved local message passing of a form quite different from backpropagation, e.g., in a multi-stage process that first learns one layer at a time and then fine-tunes via the wake-sleep algorithm (Hinton et al., 2006). Message-passing implementations of probabilistic inference have also been proposed as an explanation and generalization of deep convolutional networks (Chen et al., 2014; Patel et al., 2015). Various mappings of such processes onto neural circuitry have been attempted (George and Hawkins, 2009; Lee and Yuille, 2011; Sountsov and Miller, 2015), and related models (Makin et al., 2013, 2016) have been used to account for optimal multi-sensory integration in the brain. Feedback connections tend to terminate in distinct layers of cortex relative to the feedforward ones (Felleman and Van Essen, 1991; Callaway, 2004) making the idea of separate but interacting networks for recognition and generation potentially attractive[34]. Interestingly,

---

[32]Beginning with Hopfield's definition of an energy function for inference in certain classes of symmetric network (Hopfield, 1982), researchers have discovered networks with inherent dynamics that implicitly optimizes certain objectives even while the connection weights are fixed, such as statistical reconstruction of the input via stochastic relaxation in Boltzmann machines (Ackley et al., 1958). Fast approximations of some of these inference procedures are perhaps biologically plausible and could rely on dendritic computation (Bengio et al., 2016). Iterative local Hebbian-like *learning* rules are often used to train the *weights* of such networks, without explicitly propagating error derivatives in the manner of backpropagation. In an appropriate network context, many other combinations of network dynamics and plasticity rules can give rise to inference and learning procedures that implicitly descend cost functions in activity space and/or weight space.

[33]Dreams arguably illustrate that the brain uses generative models which also involve selective recall and recombination of episodic memories.

[34]Much is known about the architecture of cortical feedback vs. feedforward connections. For example, canonically, feedforward connections project from superficial cortical layers to layer 4 of the recipient layer, while feedback connections terminate outside layer 4 and often originate in deeper layers. These types of relationships can be used anatomically to define the hierarchical organization of visual areas, as in Felleman and Van Essen (1991), although the original studies were performed in primates and the precise generalization to rodent cortex is not fully clear (Berezovskii et al., 2011), and there may be various alternate or overlapping anatomical pathways (Callaway, 2004), e.g., with some pathways involved in specific functions like gain control, others routed through specific gating mechanisms, and so forth. Advances in connectomics should allow this architecture to be studied more directly. The study of receptive field properties in the visual cortical hierarchy has led to many insights into this hierarchical system. For example, while each neuron in V1 has a classical local receptive field, neural responses at a given location in V1 also depend on visual locations far from the classical receptive field, e.g., through various forms of surround suppression. These studies have allowed an understanding of the spatial scales over which feedback connections operate in the early visual system (Angelucci et al., 2002). In particular, feedback connections are invoked to account for longer-range receptive field interactions, whereas horizontal connections are invoked to account for shorter-range receptive field interactions (Schwabe et al., 2006). Feedforward and feedback pathways are also distinguished dynamically, e.g., by propagating different oscillatory frequencies (Van Kerkoerle et al., 2014; Bastos et al., 2015), and molecular, e.g., with NMDA receptors playing an important role in feedback processing.

such sub-networks might even be part of the same neuron and map onto "apical" vs. "basal" parts of the dendritic tree (Körding and König, 2001; Urbanczik and Senn, 2014).

Generative models can also be trained via backpropagation. Recent advances have shown how to perform variational approximations to Bayesian inference inside backpropagation-based neural networks (Kingma and Welling, 2013), and how to exploit this to create generative models (Goodfellow et al., 2014a; Gregor et al., 2015; Radford et al., 2015; Eslami et al., 2016). Through either explicitly statistical or gradient descent based learning, the brain can thus obtain a probabilistic model that simulates features of the world.

3.2.2. Cost Functions That Approximate Properties of the World

A perceiving system should exploit statistical regularities in the world that are not present in an arbitrary dataset or input distribution. For example, objects are sparse, at least in certain representations: there are far fewer objects than there are potential places in the world, and of all possible objects there is only a small subset visible at any given time. As such, we know that the output of an object recognition system must have sparse activations. Building the assumption of sparseness into simulated systems replicates a number of representational properties of the early visual system (Olshausen and Field, 1997; Rozell et al., 2008), and indeed the original paper on sparse coding obtained sparsity by gradient descent optimization of a cost function (Olshausen and Field, 1996). A range of unsupervised machine learning techniques, such as the sparse autoencoders (Le et al., 2012) used to discover cats in YouTube videos, build sparseness into neural networks. Building in such spatio-temporal sparseness priors should serve as an "inductive bias" (Mitchell, 1980) that can accelerate learning.

But we know much more about the regularities of objects. As young babies, we already know (Bremner et al., 2015) that objects tend to persist over time. The emergence or disappearance of an object from a region of space is a rare event. Moreover, object locations and configurations tend to be coherent in time. We can formulate this prior knowledge as a cost function, for example by penalizing representations which are not temporally continuous. This idea of continuity is used in a great number of artificial neural networks and related models (Wiskott and Sejnowski, 2002; Földiák, 2008; Mobahi et al., 2009). Imposing continuity within certain models gives rise to aspects of the visual system including complex cells (Körding et al., 2004), specific properties of visual invariance (Isik et al., 2012), and even other representational properties such as the existence of place cells (Wyss et al., 2006; Franzius et al., 2007). Unsupervised learning mechanisms that maximize temporal coherence or slowness are increasingly used in machine learning[35].

We also know that objects tend to undergo predictable sequences of transformations, and it is possible to build this assumption into unsupervised neural learning systems (George and Hawkins, 2009). The minimization of prediction error explains a number of properties of the nervous system (Friston and Stephan, 2007; Huang and Rao, 2011), and biologically plausible theories are available for how cortex could learn using prediction errors by exploiting temporal differences (O'Reilly et al., 2014b) or top-down feedback (George and Hawkins, 2009). In one implementation, a system can simply predict the next input delivered to the system and can then use the difference between the actual next input and the predicted next input as a full vectorial error signal for supervised gradient descent. Thus, rather than optimization of prediction error being implicitly implemented by the network dynamics, the prediction error is used as an explicit cost function in the manner of supervised learning, leading to error derivatives which can be back-propagated. Then, no special learning rules beyond simple backpropagation are needed. This approach has recently been advanced within machine learning (Lotter et al., 2015, 2016). Recently, combining such prediction-based learning with a specific gating mechanism has been shown to lead to unsupervised learning of disentangled representations (Whitney et al., 2016). Neural networks can also be designed to learn to invert spatial transformations (Jaderberg et al., 2015). Statistically describing transformations or sequences is thus an unsupervised way of learning representations.

Furthermore, there are multiple modalities of input to the brain. Each sensory modality is primarily connected to one part of the brain[36]. But higher levels of cortex in each modality are heavily connected to the other modalities. This can enable forms of self-supervised learning: with a developing visual understanding of the world we can predict its sounds, and then test those predictions with the auditory input, and vice versa. The same is true about multiple parts of the same modality: if we understand the left half of the visual field, it tells us an awful lot about the right. Indeed, we can use observations of one part of a visual scene to predict the contents of other parts (Noroozi and Favaro, 2016; van den Oord et al., 2016), and optimize a cost function that reflects the discrepancy. Maximizing mutual information is a natural way of improving learning (Becker and Hinton, 1992; Mohamed and Rezende, 2015), and there are many other ways in which multiple modalities or processing streams could mutually train one

---

[35]Temporal continuity is exploited in Poggio (2015), which analyzes many properties of deep convolutional networks with respect to their biological plausibility, including their apparent need for large amounts of supervised training data, and concludes that the environment may in fact provide a sufficient number of "implicitly," though not explicitly, labeled examples to train a deep convolutional network for object recognition. Implicit labeling of object identity, in this case, arises from temporal continuity: successive frames of a video are likely to have the same objects in similar places and orientations. This allows the brain to derive

an invariant signature of object identity which is independent of transformations like translations and rotations, but which does not yet associate the object with a specific name or label. Once such an invariant signature is established, however, it becomes basically trivial to associate the signature with a label for classification (Anselmi et al., 2015). Poggio (2015) also suggests specific means, in the context of I-theory (Anselmi et al., 2015), by which this training could occur via the storage of image templates using Hebbian mechanisms among simple and complex cells in the visual cortex. Thus, in this model, the brain has used its implicit knowledge of the temporal continuity of object motion to provide a kind of minimal labeling that is sufficient to bootstrap object recognition. Although not formulated as a cost function, this shows how usefully the assumption of temporal continuity could be exploited by the brain.
[36]Although, some multi-sensory integration appears to occur even in the early sensory cortices (Cappe et al., 2012).

another. This way, each modality effectively produces training signals for the others[37]. Evidence from psychophysics suggests that some kind of training via detection of sensory conflicts may be occurring in children (Nardini et al., 2010).

3.3. Cost Functions for Supervised Learning

In what cases might the brain use supervised learning, given that it requires the system to "already know" the exact target pattern to train toward? One possibility is that the brain can store records of states that led to good outcomes. For example, if a baby reaches for a target and misses, and then tries again and successfully hits the target, then the difference in the neural representations of these two tries reflects the direction in which the system should change. The brain could potentially use a comparator circuit to directly compute this vectorial difference in the neural population codes and then apply this difference vector as an error signal.

Another possibility is that the brain uses supervised learning to implement a form of "chunking," i.e., a consolidation of something the brain already knows how to do: routines that are initially learned as multi-step, deliberative procedures could be compiled down to more rapid and automatic functions by using supervised learning to train a network to mimic the overall input-output behavior of the original multi-step process. Such a process is assumed to occur in cognitive models like ACT-R (Servan-Schreiber and Anderson, 1990), and methods for compressing the knowledge in neural networks into smaller networks are also being developed (Ba and Caruana, 2014). Thus supervised learning can be used to train a network to do in "one step" what would otherwise require long-range routing and sequential recruitment of multiple systems.

3.4. Repurposing Reinforcement Learning for Diverse Internal Cost Functions

Certain generalized forms of reinforcement learning may be ubiquitous throughout the brain. Such reinforcement signals may be repurposed to optimize diverse internal cost functions. These internal cost functions could be specified at least in part by genetics.

Some brain systems such as in the striatum appear to learn via some form of temporal difference reinforcement learning (Tesauro, 1995; Foster et al., 2000). This is reinforcement learning based on a global value function (O'Reilly et al., 2014a) that predicts total future reward or utility for the agent. Reward-driven signaling is not restricted to the striatum, and is present even in primary visual cortex (Chubykin et al., 2013; Stanisor et al., 2013). Remarkably, the reward signaling in primary visual cortex is mediated in part by glial cells (Takata et al., 2011), rather than neurons, and involves the neurotransmitter

acetylcholine (Chubykin et al., 2013; Hangya et al., 2015). On the other hand, some studies have suggested that visual cortex learns the basics of invariant object recognition in the absence of reward (Li and Dicarlo, 2012), perhaps using reinforcement only for more refined perceptual learning (Roelfsema et al., 2010).

But beyond these well-known global reward signals, we argue that the basic mechanisms of reinforcement learning may be widely re-purposed to train local networks using a variety of internally generated error signals. These internally generated signals may allow a learning system to go beyond what can be learned via standard unsupervised methods, effectively guiding or steering the system to learn specific features or computations (Ullman et al., 2012).

3.4.1. Cost Functions for Bootstrapping Learning in the Human Environment

Special, internally-generated signals are needed specifically for learning problems where standard unsupervised methods—based purely on matching the statistics of the world, or on optimizing simple mathematical objectives like temporal continuity or sparsity—will fail to discover properties of the world which are statistically weak in an objective sense but nevertheless have special significance to the organism (Ullman et al., 2012). Indigo bunting birds, for example, learn a template for the constellations of the night sky long before ever leaving the nest to engage in navigation-dependent tasks (Emlen, 1967). This memory template is directly used to determine the direction of flight during migratory periods, a process that is modulated hormonally so that winter and summer flights are reversed. Learning is therefore a multi-phase process in which navigational cues are memorized prior to the acquisition of motor control.

In humans, we suspect that similar multi-stage bootstrapping processes are arranged to occur. Humans have innate specializations for social learning. We need to be able to read one another's expressions as indicated with hands and faces. Hands are important because they allow us to learn about the set of actions that can be produced by agents (Ullman et al., 2012). Faces are important because they give us insight into what others are thinking. People have intentions and personalities that differ from one another, and their feelings are important. How could we hack together cost functions, built on simple genetically specifiable mechanisms, to make it easier for a learning system to discover such behaviorally relevant variables?

Some preliminary studies are beginning to suggest specific mechanisms and heuristics that humans may be using to bootstrap more sophisticated knowledge. In a groundbreaking study, Ullman et al. (2012) asked how could we explain hands, to a system that does not already know about them, in a cheap way, without the need for labeled training examples? Hands are common in our visual space and have special roles in the scene: they move objects, collect objects, and caress babies. Building these biases into an area specialized to detect hands could guide the right kind of learning, by providing a downstream learning system with many likely positive examples of hands on the basis of innately-stored, heuristic signatures about how hands tend to look or behave (Ullman et al., 2012). Indeed, an internally supervised learning algorithm containing specialized,

---

[37]Other brain-inspired unsupervised objectives are being developed for unsupervised visual learning. One recent paper (Higgins et al., 2016) uses an objective function that seeks representations of statistically independent factors in images, by introducing a regularization term that pushes the distribution of latent factors learned in a generative model to be close to a unit Gaussian. This is based on a theory that the ventral visual stream is optimized to disentangle factors of variation in images.

hard-coded biases to detect hands, on the basis of their typical motion properties, can be used to bootstrap the training of an image recognition module that learns to recognize hands based on their appearance. Thus, a simple, hard-coded module bootstraps the training of a much more complex algorithm for visual recognition of hands.

Ullman et al. (2012) then further exploits a combination of hand and face detection to bootstrap a predictor for gaze direction, based on the heuristic that faces tend to be looking toward hands. Of course, given a hand detector, it also becomes much easier to train a system for reaching, crawling, and so forth. Efforts are underway in psychology to determine whether the heuristics discovered to be useful computationally are, in fact, being used by human children during learning (Yu and Smith, 2013; Fausey et al., 2016).

Ullman refers to such primitive, inbuilt detectors as innate "proto-concepts" (Ullman et al., 2012). Their broader claim is that such pre-specification of mutual supervision signals can make learning the relevant features of the world far easier, by giving an otherwise unsupervised learner the right kinds of hints or heuristic biases at the right times. Here we call these approximate, heuristic cost functions "bootstrap cost functions." The purpose of the bootstrap cost functions is to reduce the amount of data required to learn a specific feature or task, but at the same time to avoid a need for fully unsupervised learning.

Could the neural circuitry for such a bootstrap hand-detector be pre-specified genetically? The precedent from other organisms is strong: for example, it is famously known that the frog retina contains circuitry sufficient to implement a kind of "bug detector" (Lettvin et al., 1959). Ullman's hand detector, in fact, operates via a simple local optical flow calculation to detect "mover" events. This type of simple, local calculation could potentially be implemented in genetically-specified and/or spontaneously self-organized neural circuitry in the retina or early dorsal visual areas (Bülthoff et al., 1989), perhaps similarly to the frog's "bug detector."

How could we explain faces without any training data? Faces tend to have two dark dots in their upper half, a line in the lower half and tend to be symmetric about a vertical axis. Indeed, we know that babies are very much attracted to things with these generic features of upright faces starting from birth, and that they will acquire face-specific cortical areas[38] in their first few years of life if not earlier (McKone et al., 2009). It is easy to define a local rule that produces a kind of crude face detector

(e.g., detecting two dots on top of a horizontal line), and indeed some evidence suggests that the brain can rapidly detect faces without even a single feed-forward pass through the ventral visual stream (Crouzet and Thorpe, 2011). The crude detection of human faces used together with statistical learning should be analogous to semi-supervised learning (Sukhbaatar et al., 2014) and could allow identifying faces with high certainty.

Humans have areas devoted to emotional processing, and the brain seems to embody prior knowledge about the structure of emotional expressions and how they relate to causes in the world: emotions should have specific types of strong couplings to various other higher-level variables such as goal-satisfaction, should be expressed through the face, and so on (Phillips et al., 2002; Skerry and Spelke, 2014; Baillargeon et al., 2016; Lyons and Cheries, 2016). What about agency? It makes sense to describe, when dealing with high-level thinking, other beings as optimizers of their own goal functions. It appears that heuristically specified notions of goals and agency are infused into human psychological development from early infancy and that notions of agency are used to bootstrap heuristics for ethical evaluation (Hamlin et al., 2007; Skerry and Spelke, 2014). Algorithms for establishing more complex, innately-important social relationships such as joint attention are under study (Gao et al., 2014), building upon more primitive proto-concepts like face detectors and Ullman's hand detectors (Ullman et al., 2012). The brain can thus use innate detectors to create cost functions and training procedures to train the next stages of learning. This prior knowledge, encoded into brain structure via evolution, could allow learning signals to come from the right places and to appear developmentally at the right times.

It is intuitive to ask whether this type of bootstrapping poses a kind of "chicken and egg" problem: if the brain already has an inbuilt heuristic hand detector, how can it be used to train a detector that performs any better than those heuristics? After all, isn't a trained system only as good as its training data? The work of Ullman et al. (2012) illustrates why this is not the case. First, the "innate detector" can be used to train a downstream detector that operates based on different cues: for example, based on the spatial and body context of the hand, rather than its motion. Second, once multiple such pathways of detection come into existence, they can be used to improve each other. In Ullman et al. (2012), appearance, body context, and mover motion are all used to bootstrap off of one another, creating a detector that is better than any of its training heuristics. In effect, the innate detectors are used not as supervision signals *per se*, but rather to guide or steer the learning process, enabling it to discover features that would otherwise be difficult. If such affordances can be found in other domains, it seems likely that the brain would make extensive use of them to ensure that developing animals learn the precise patterns of perception and behavior needed to ensure their later survival and reproduction.

Thus, generalizing previous ideas (Ullman et al., 2012; Poggio, 2015), we suggest that the brain uses optimization with respect to internally generated heuristic[39] detection signals

---

[38]In the visual system, it is still unknown why a clustered spatial pattern of representational categories arises, e.g., a physically localized "area" that seems to correspond to representations of faces (Kanwisher et al., 1997), another area for representations of visual word forms (McCandliss et al., 2003), and so on. It is also unknown why this spatial pattern seems to be largely reproducible across individuals. Some theories are based on bottom-up correlation-based clustering or neuronal competition mechanisms, which generate category-selective regions as a byproduct. Other theories suggest a computational reason for this organization, in the context of I-theory (Anselmi et al., 2015), involving the limited ability to generalize transformation-invariances learned for one class of objects to other classes (Leibo et al., 2015b). Areas for abstract culture-dependent concepts, like the visual word form area, suggest that the decomposition cannot be "purely genetic." But it is conceivable that these areas could at least in part reflect different local cost functions.

[39]Psychologists have postulated other innate heuristics, e.g., in the context of object tracking (Franconeri et al., 2012). That infant object concepts are trainable but only

to bootstrap learning of biologically relevant features which would otherwise be missed by an unsupervised learner. In one possible implementation, such bootstrapping may occur via reinforcement learning, using the outputs of the innate detectors as local reinforcement signals, and perhaps using mechanisms similar to Stanisor et al. (2013), Rombouts et al. (2015), Brosch et al. (2015), and Roelfsema and van Ooyen (2005) to perform reinforcement learning through a multi-layer network. It is also possible that the brain could use such internally generated heuristic detectors in other ways, for example to bias the inputs delivered to an unsupervised learning network toward entities of interest to humans via an attentional process (Joscha Bach, personal communication), to bias hippocampal replay (Kumaran et al., 2016) or other aspects of memory access, or to directly train simple classifiers (Ullman et al., 2012).

3.4.2. Cost Functions for Learning by Imitation and through Social Feedback

It has been widely observed that the capacity for imitation and social learning may be a feature that is uniquely human, and that enables other human traits (Ramachandran, 2000). Humans need to learn more from the environment by than trial and error can provide for, and more than genetically orchestrated internal bootstrapping signals can effectively guide. Hence, babies spend a long time watching adults, especially adults they are attached to Meltzoff (1999), and later use specific kinds of social cues from their parents to shape their development. Babies and children learn about cause and effect through models based on goals, outcomes and agents, not just pure statistical inference. For example, young children make inferences about causality selectively in situations where a human is trying to achieve an outcome (Meltzoff et al., 2012, 2013). Minsky (2006) discusses how we derive not just skills but also goals from our attachment figures, through socially induced emotions like pride and shame. To do all this requires a powerful infrastructure of mental abilities: we must attribute social feedback to particular aspects of our goals or actions, and hence we need to signal to each other positively and negatively, to draw attention to these aspects. Minsky speculates (Minsky, 2006) that the development of such "learning by being told" led to language by selecting for the development of increasingly precise parsing of synatic structures in relation to our representations of agents and action-plans.

How does this connect with cost functions? The idea of goals is central here, as we need to be able to identify the goals of others, update our own goals based on feedback, and measure the success of actions relative to goals. It has been proposed that human intrinsically use a model based on abstract goal and costs to underpin learning about the social world (Jara-Ettinger et al., 2016). Perhaps we even learn about our "selves" by inferring a model of our own goals and cost functions. Relatedly, machine learning in some settings can infer their cost functions from samples of behavior (Ho and Ermon, 2016).

along certain dimensions (Scholl, 2004) also suggests the notion of a heuristically "guided" or "bootstrapped" learning process in this context.

3.4.3. Cost Functions for Story Generation and Understanding

It has been widely noticed in cognitive science and AI that the generation and understanding of stories are crucial to human cognition. Researchers such as Winston have framed story understanding as the key to human-like intelligence (Winston, 2011). Stories consist of a linear sequence of episodes, in which one episode refers to another through cause and effect relationships, with these relationships often involving the implicit goals of agents. Many other cognitive faculties, such as conceptual grounding of language, could conceivably emerge from an underlying internal representation in terms of stories.

Perhaps the ultimate series of bootstrap cost functions would be those which would direct the brain to utilize its learning networks and specialized systems so as to construct representations that are specifically useful as components of stories, to spontaneously chain these representations together, and to update them through experience and communication. How could such cost functions arise? One possibility is that they are bootstrapped through imitation and communication, where a child learns to mimic the story-telling behavior of others. Another possibility is that useful representations and primitives for stories emerge spontaneously from mechanisms for learning state and action chunking in hierarchical reinforcement learning and planning. Yet another is that stories emerge from learned patterns of saliency-directed memory storage and recall (e.g., Xiong et al., 2016). In addition, priors that direct the developing child's brain to learn about and attend to social agency seem to be important for stories.

In this section, we have seen how cost functions can be specified that could lead to the learning of increasingly sophisticated mental abilities in a biologically plausible manner. Importantly, however, cost functions and optimization are not the whole story. To achieve more complex forms of optimization, e.g., for learning to understand complex patterns of cause and effect over long timescales, to plan and reason prospectively, or to effectively coordinate many widely distributed brain resources, the brain seems to invoke specialized, pre-constructed data structures, algorithms and communication systems, which in turn facilitate specific kinds of optimization. Moreover, optimization occurs in a tightly orchestrated multi-stage process, and specialized, pre-structured brain systems need to be invoked to account for this meta-level of control over when, where and how each optimization problem is set up. We now turn to how these pre-specialized systems may orchestrate and facilitate optimization.

4. OPTIMIZATION OCCURS IN THE CONTEXT OF SPECIALIZED STRUCTURES

Optimization of initially unstructured "blank slate" networks is not sufficient to generate complex cognition in the brain, we argue, even given a diversity of powerful genetically-specified cost functions and local learning rules, as we have posited above. Instead, in Hypothesis 3, we suggest that specialized, pre-structured architectures are needed for at least two purposes.

First, pre-structured architectures are needed to allow the brain to find efficient solutions to certain types of problems. When we write computer code, there are a broad range of algorithms and data structures employed for different purposes: we may use dynamic programming to solve planning problems, trees to efficiently implement nearest neighbor search, or stacks to implement recursion. Having the right kind of algorithm and data structure in place to solve a problem allows it to be solved efficiently, robustly and with a minimum amount of learning or optimization needed. This observation is concordant with the increasing use of pre-specialized architectures and specialized computational components in machine learning (Graves et al., 2014; Weston et al., 2014; Neelakantan et al., 2015). In particular, to enable the learning of efficient computational solutions, the brain may need pre-specialized systems for planning and executing sequential multi-step processes, for accessing memories, and for forming and manipulating compositional and recursive structures[40].

Second, the training of optimization modules may need to be coordinated in a complex and dynamic fashion, including delivering the right training signals and activating the right learning rules in the right places and at the right times. To allow this, the brain may need specialized systems for storing and routing data, and for flexibly routing training signals such as target patterns, training data, reinforcement signals, attention signals, and modulatory signals. These mechanisms may need to be at least partially in place in advance of learning.

Looking at the brain, we indeed seem to find highly conserved structures, e.g., cortex, where it is theorized that a similar type of learning and/or computation is happening in multiple places (Braitenberg and Schutz, 1991; Douglas and Martin, 2004). But we also see a large number of specialized structures, including thalamus, hippocampus, basal ganglia and cerebellum (Solari and Stoner, 2011). These structures evolutionarily pre-date (Lee et al., 2015) the cortex, and hence the cortex may have evolved to work in the context of such specialized mechanisms. For example, the cortex may have evolved as a trainable module for which the training is orchestrated by these older structures.

Even within the cortex itself, microcircuitry within different areas may be specialized: tinkered variations on a common ancestral microcircuit scaffold could potentially allow different cortical areas, such as sensory areas vs. prefrontal areas, to be configured to adopt a number of qualitatively distinct computational and learning configurations (Yuste et al., 2005; Marcus et al., 2014a,b), even while sharing a common gross physical layout and communication interface. Within cortex, over forty distinct cell types—differing in such aspects as dendritic organization, distribution throughout the six cortical layers, connectivity pattern, gene expression, and electrophysiological properties—have already been found (Markram et al., 2015; Zeisel et al., 2015). Central

pattern generator circuits provide an example of the kinds of architectures that can be pre-wired into neural microcircuitry, and may have evolutionary relationships with cortical circuits (Yuste et al., 2005). Thus, while the precise degree of architectural specificity of particular cortical regions is still under debate (Marcus et al., 2014a,b), various mechanisms could offer pre-specified heterogeneity.

In this section, we explore the kinds of computational problems for which specialized structures may be useful, and attempt to map these to putative elements within the brain. Our preliminary sketch of a functional decomposition can be viewed as a summary of suggestions for specialized functions that have been made throughout the computational neuroscience literature, and is influenced strongly by the models of O'Reilly, Eliasmith, Grossberg, Marcus, Hayworth and others (Marcus, 2001; O'Reilly, 2006; Eliasmith et al., 2012; Hayworth, 2012; Grossberg, 2013). The correspondence between these models and actual neural circuitry is, of course, still the subject of extensive debate.

Many of the computational and neural concepts sketched here are preliminary and will need to be made more rigorous through future study. Our knowledge of the functions of particular brain areas, and thus our proposed mappings of certain computations onto neuroanatomy, also remains tentative. Finally, it is still far from established which processes in the brain emerge from optimization of cost functions, which emerge from other forms of self-organization, which are pre-structured through genetics and development, and which rely on an interplay of all these mechanisms[41]. Our discussion here should therefore be viewed as a sketch of potential directions for further study.

4.1. Structured Forms of Memory

One of the central elements of computation is memory. Importantly, multiple different kinds of memory are needed (Squire, 2004). For example, we need memory that is stored for a long period of time and that can be retrieved in a number of ways, such as in situations similar to the time when the memory was first stored (content addressable memory). We also need memory that we can keep for a short period of time and that we can rapidly rewrite (working memory). Lastly, we need the kind of implicit memory that we cannot explicitly recall, similar to the kind of memory that is classically learned using

---

[40]Of course, specialized architecture also enters the picture at the level of the pre-structuring of trainable/optimizable modules themselves. Just as in deep learning, convolutional networks, LSTMs, residual networks and other specific architectures are used to make learning efficient and fast, even though more generic architectures like multilayer perceptrons or generally RNNs are universal function approximators.

[41]It is interesting to consider how standard neural network models of vision would fit into this categorization. Consider convolutional neural networks, for example, with the convolutional filters optimized via supervised backpropagation. This is by no means a completely unstructured prior to backpropagation-based training. Indeed, these networks typically contain max-pooling and normalization layers with fixed computations that are not altered during learning, as well as fixed architectural features such as number and arrangement of layers, size and stride of the sliding window, and so forth. Likewise "hierarchical max-pooling" (HMAX) models (Serre et al., 2007) of the ventral stream are so-named because of these fixed architectural aspects. Thus, in a hypothetical biological implementation of such systems, these aspects would be pre-structured by genetics even if the convolutional weights would be trained via some kind of gradient descent optimization. There are some plausible neural circuits that would implement these standardized normalization and max pooling operations (Kouh and Poggio, 2008). Moreover, in a biological implementation, the machinery necessary to carry out the optimization itself would need to be embodied by appropriate, genetically structured circuitry.

gradient descent on errors, i.e., sculpted into the weight matrix of a neural network.

4.1.1. Content Addressable Memories

Content addressable memories[42] are classic models in neuroscience (Hopfield, 1982). Most simply, they allow us to recognize a situation similar to one that we have seen before, and to "fill in" stored patterns based on partial or noisy information, but they may also be put to use as sub-components of many other functions. Recent research has shown that including such memories allows deep networks to learn to solve problems that previously were out of reach, even of LSTM networks that already have a simpler form of local memory and are already capable of learning long-term dependencies (Graves et al., 2014; Weston et al., 2014). Hippocampal area CA3 may act as an auto-associative memory[43] capable of content-addressable pattern completion, with pattern separation occurring in the dentate gyrus (Rolls, 2013). If no similar pattern is available, an unfamiliar input will be stored as a new memory (Kumaran et al., 2016). Such systems could permit the retrieval of complete memories from partial cues, enabling networks to perform operations similar to database retrieval or to instantiate lookup tables of historical stimulus-response mappings, among numerous other possibilities.

Of course, memory systems may be organized—through cost function optimization or other mechanisms—into higher-order structures. Cost functions might be used to bias memory representations to adopt particular structures, e.g., to be organized into data structures like like Minskys frames and trans-frames (Minsky, 2006).

4.1.2. Working Memory Buffers

Cognitive science has long characterized properties of the working memory. Its capacity is somewhat limited, with the old idea being that verbal working memory has a capacity of "seven plus or minus two" (Miller, 1956), while visual working memory has a capacity of four (Luck and Vogel, 1997) (or, other authors defend, one). There are many models of working memory (O'Reilly and Frank, 2006; Singh and Eliasmith, 2006;

---

[42]Attractor models of memory in neuroscience tend to have the property that only one memory can be accessed at a time (although a brain can have many such memories that can be accessed in parallel). Recent machine learning systems, however, have constructed differentiable addressable memory (Graves et al., 2014) and gating (Whitney et al., 2016) systems by allowing weighted superpositions of memory registers or gates to be queried—it is unclear whether the brain uses such mechanisms.

[43]Computational analogies have also been drawn between associative memory storage and object recognition (Leibo et al., 2015a), suggesting the possibility of closely related computations occurring in parts of neocortex and hippocampus. Indeed, the hippocampus and olfactory cortex (a more ancient and simpler structure than the neocortex Shepherd, 2014; Fournier et al., 2015) are few-layer structures described in comparative anatomy as "allocortex," as opposed to the six-layered "neocortex," and both types of cortex have some anatomical similarities (particularly for CA1 and subiculum, though less so for CA3 and dentate gyrus) such as the presence of pyramidal neurons. It has been suggested that the hippocampus can be thought of as the top of the cortical hierarchy (Hawkins and Blakeslee, 2007), responsible for handling and remembering information that could not be fully explained by lower levels of the hierarchy. These computational connections are still tentative.

Warden and Miller, 2007; Wang, 2012; Buschman and Miller, 2014), some of which attribute it to persistent, self-reinforcing patterns of neural activation (Goldman et al., 2003) in the recurrent networks of the prefrontal cortex. Prefrontal working memory appears to be made up of multiple functionally distinct subsystems (Markowitz et al., 2015). Neural models of working memory can store not only scalar variables (Seung, 1998), but also high-dimensional vectors (Eliasmith and Anderson, 2004; Eliasmith et al., 2012) or sequences of vectors (Choo and Eliasmith, 2010). Working memory buffers seem crucial for human-like cognition, e.g., reasoning, as they allow short-term storage while also—in conjunction with other mechanisms—enabling generalization of operations across anything that can fill the buffer.

4.1.3. Storing State in Association with Saliency

Saliency, or interestingness, measures can be used to tag the importance of a memory (Gonzalez Andino and Grave de Peralta Menendez, 2012). This can allow removal of the boring data from the training set, allowing a mechanism that is more like optimal experimentation. Moreover, saliency can guide memory replay or sampling from generative models, to generate more training data drawn from a distribution useful for learning (Ji and Wilson, 2007; Mnih et al., 2015). Conceivably, hippocampal replay could allow a batch-like training process, similar to how most machine learning systems are trained, rather than requiring all training to occur in an online fashion. Plasticity mechanisms in memory systems which are gated by saliency are starting to be uncovered in neuroscience (Dudman et al., 2007). Importantly, the notions of "saliency" computed by the brain could be quite intricate and multi-faceted, potentially leading to complex schemes by which specific kinds of memories would be tagged for later context-dependent retrieval. As a hypothetical example, representations of both timing and importance associated with memories could perhaps allow retrieval only of important memories that happened within a certain window of time (MacDonald et al., 2011; Kraus et al., 2013; Rubin et al., 2015). Storing and retrieving information selectively based on specific properties of the information itself, or of "tags" appended to that information, is a powerful computational primitive that could enable learning of more complex tasks. Relatedly, we know that certain pathways become associated with certain kinds of memories, e.g., specific pathways for fear-related memory in mice.

4.2. Structured Routing Systems

To use its information flexibly, the brain needs structured systems for routing data. Such systems need to address multiple temporal and spatial scales, and multiple modalities of control. Thus, there are several different kinds of information routing systems in the brain which operate by different mechanisms and under different constraints.

4.2.1. Attention

If we can focus on one thing at a time, we may be able to allocate more computational resources to processing it, make better use of scarce data to learn about it, and more easily

store and retrieve it from memory[44]. Notably in this context, attention allows improvements in learning: if we can focus on just a single object, instead of an entire scene, we can learn about it more easily using limited data. Formal accounts in a Bayesian framework talk about attention reducing the sample complexity of learning (Chikkerur et al., 2010). Likewise, in models, the processes of applying attention, and of effectively making use of incoming attentional signals to appropriately modulate local circuit activity, can themselves be learned by optimizing cost functions (Jaramillo and Pearlmutter, 2004; Mnih et al., 2014). The right kinds of attention make processing and learning more efficient, and also allow for a kind of programmatic control over multi-step perceptual tasks.

How does the brain determine where to allocate attention, and how is the attentional signal physically mediated? Answering this question is still an active area of neuroscience. Higher-level cortical areas may be specialized in allocating attention. The problem is made complex by the fact that there seem to be many different types of attention—such as object-based, feature-based and spatial attention in vision—that may be mediated by interactions between different brain areas. The frontal eye fields (area FEF), for example, are important in visual attention, specifically for controlling saccades of the eyes to attended locations. Area FEF contains "retinotopic" spatial maps whose activation determines the saccade targets in the visual field. Other prefrontal areas such as the dorsolateral prefrontal cortex and inferior frontal junction are also involved in maintaining representations that specify the targets of certain types of attention. Certain forms of attention may require a complex interaction between brain areas, e.g., to determine targets of attention based on higher-level properties that are represented across multiple areas, like the identity and spatial location of a specific face (Baldauf and Desimone, 2014).

There are many proposed neural mechanisms of attention, including the idea that synchrony plays a role (Baldauf and Desimone, 2014), perhaps by creating resonances that facilitate the transfer of information between synchronously oscillating neural populations in different areas[45]. Other proposed mechanisms include specific circuits for attention-dependent signal routing (Anderson and Van Essen, 1987; Olshausen et al., 1993). Various forms of attention also have specific neurophysiological signatures, such as enhancements in synchrony among neural spikes and with the ambient local field potential, changes in the sharpness of neural tuning curves, and other properties. These diverse effects and signatures of attention may be consequences of underlying pathways that wire up to

particular elements of cortical microcircuits to mediate different attentional effects (Bobier et al., 2014).

4.2.2. Buffers
One possibility is that the brain uses distinct groups of neurons, which we can call "buffers," to store distinct variables, such as the subject or object in a sentence (Frankland and Greene, 2015). Having memory buffers allows the abstraction of a variable.

Once we establish that the brain has a number of memory buffers, we need ways for those buffers to interact. We need to be able to take a buffer, do a computation on its contents and store the output into another buffer. But if the representations in each of two groups of neurons are learned, and hence are coded differently, how can the brain "copy and paste" information between these groups of neurons? Malsburg argued that such a system of separate buffers is impossible because the neural pattern for "chair" in buffer 1 has nothing in common with the neural pattern for "chair" in buffer 2—any learning that occurs for the contents of buffer 1 would not automatically be transferable to buffer 2. Various mechanisms have been proposed to allow such transferability, which focus on ways in which all buffers could be trained jointly and then later separated so that they can work independently when they need to[46].

4.2.3. Discrete Gating of Information Flow between Buffers
Dense connectivity is only achieved locally, but it would be desirable to have a way for any two cortical units to talk to one another, if needed, regardless of their distance from one another, and without introducing crosstalk[47]. It is therefore critical to be able to dynamically turn on and off the transfer of information between different source and destination regions, in much the manner of a switchboard. Together with attention, such dedicated routing systems can make sure that a brain area receives exactly the information it needs. Such a discrete routing system is, of course, central to cognitive architectures like ACT-R (Anderson, 2007). The key feature of ACT-R is the ability to evaluate the IF clauses of tens of thousands of symbolic rules

---

[44]Attention also arguably solves certain types of perceptual binding problem (Reynolds and Desimone, 1999).

[45]The precise roles of synchrony in information routing and other processes, and when it should be viewed as a causal factor vs. as an epiphenomenon of other mechanisms, is still being worked out. In some theories, oscillations occur as consequences of certain recurrent processing loops, e.g., thalamo-cortico-striatal loops (Eliasmith et al., 2012). In other models, so-called "dynamic circuit motifs," involving specific combinations of cellular and synaptic sub-types, both generate synchronies (e.g., in part via intrinsically rhythmic pacemaker neurons) and exploit them for specific computational roles, particularly in the rapid dynamic formation of communication networks (Womelsdorf et al., 2014).

[46]One idea for achieving such transferability is that of a partitionable (Hayworth, 2012) or annexable (Bostrom, 1996) network. These models posit that a large associative memory network links all the different buffers. This large associative memory network has a number of stable attractor states. These are called "global" attractor states since they link across all the buffers. Forcing a given buffer into an activity pattern resembling that of its corresponding "piece" of an attractor state will cause the entire global network to enter that global attractor state. During training, all of the connections between buffers are turned on, so that their learned contents, though not identical, are kept in correspondence by being part of the same attractor. Later, the connections between specific buffers can be turned off to allow them to store different information. Copy and paste is then implemented by turning on the connections between a source buffer and a destination buffer (Hayworth, 2012). Copying between a source and destination buffer can also be implemented, i.e., learned, in a deep learning system using methods similar to the addressing mechanisms of the Neural Turing Machine (Graves et al., 2014).

[47]Micro-stimulation experiments, in which an animal learns to behaviorally report stimulation of electrode channels located in diverse cortical regions, suggest that many areas can be routed or otherwise linked to behavioral "outputs" (Histed et al., 2013), although the mechanisms behind this—e.g., whether this stimulation gives rise to a high-level percept that the animal then uses to make a decision—are unclear. Likewise, it is possible to reinforcement-train an animal to control the activity of individual neurons (Fetz, 1969, 2007).

(called "productions"), in parallel, approximately every 50 ms. Each rule requires equality comparisons between the contents of many constant and variable memory buffers, and the execution of a rule leads to the conditional routing of information from one buffer to another.

What controls which long-range routing operations occur when, i.e., where is the switchboad and what controls it? Several models, including ACT-R, have attributed such parallel rule-based control of routing to the action selection circuitry (Gurney et al., 2001; Terrence Stewart, 2010) of the basal ganglia (BG) (O'Reilly and Frank, 2006; Stocco et al., 2010), and its interaction with working memory buffers in the prefrontal cortex. In conventional models of thalamo-cortico-striatal loops, competing actions of the direct and indirect pathways through the basal ganglia can inhibit or disinhibit an area of motor cortex, thereby gating a motor action[48]. Models like (O'Reilly and Frank, 2006; Stocco et al., 2010; Terrence Stewart, 2010) propose further that the basal ganglia can gate not just the transfer of information from motor cortex to downstream actuators, but also the transfer of information between cortical areas. To do so, the basal ganglia would dis-inhibit a thalamic relay (Sherman, 2005, 2007) linking two cortical areas. Dopamine-related activity is thought to lead to temporal difference reinforcement learning of such gating policies in the basal ganglia (Frank and Badre, 2012). Beyond the basal ganglia, there are also other, separate pathways involved in action selection, e.g., in the prefrontal cortex (Daw et al., 2006). Thus, multiple systems including basal ganglia and cortex could control the gating of long-range information transfer between cortical areas, with the thalamus perhaps largely constituting the switchboard itself.

How is such routing put to use in a learning context? One possibility is that the basal ganglia acts to orchestrate the training of the cortex. The basal ganglia may exert tight control[49] over the cortex, helping to determine when and how it is trained. Indeed, because the basal ganglia pre-dates the cortex evolutionarily, it is possible that the cortex evolved as a flexible, trainable resource that could be harnessed by existing basal ganglia circuitry. All of the main regions and circuits of the basal ganglia are conserved from our common ancestor with

the lamprey more than five hundred million years ago. The major part of the basal ganglia even seems to be conserved from our common ancestor with insects (Strausfeld and Hirth, 2013). Thus, in addition to its real-time action selection and routing functions, the basal ganglia may sculpt how the cortex learns.

4.3. Structured State Representations to Enable Efficient Algorithms

Certain algorithmic problems benefit greatly from particular types of representation and transformation, such as a grid-like representation of space. In some cases, rather than just waiting for them to emerge via gradient descent optimization of appropriate cost functions, the brain may be pre-structured to facilitate their creation.

4.3.1. Continuous Predictive Control

We often have to plan and execute complicated sequences of actions on the fly, in response to a new situation. At the lowest level, that of motor control, our body and our immediate environment change all the time. As such, it is important for us to maintain knowledge about this environment in a continuous way. The deviations between our planned movements and those movements that we actually execute continuously provide information about the properties of the environment. Therefore, it seems important to have a specialized system, optimized for high-speed continuous processing, that takes all our motor errors and uses them to update a dynamical model of our body and our immediate environment that can predict the delayed sensory results of our motor actions (McKinstry et al., 2006).

It appears that the cerebellum is such a structure, and lesions to it abolish our way of dealing successfully with a changing body. Incidentally, the cerebellum has more connections than the rest of the brain taken together, apparently in a largely feedforward architecture, and the tiny cerebellar granule cells, which may form a randomized high-dimensional input representation (Marr, 1969; Jacobson and Friedrich, 2013), outnumber all other neurons. The brain clearly needs a dedicated way of quickly and continuously correcting movements to minimize errors, without needing to rely on slow and complex association learning in the neocortex in order to do so.

Newer research shows that the cerebellum is involved in a broad range of cognitive problems (Moberget et al., 2014) as well, potentially because they share computational problems with motor control. For example, when subjects estimate time intervals, which are naturally important for movement, it appears that the brain uses the cerebellum even if no movements are involved (Gooch et al., 2010). Even individual cerebellar Purkinjie cells may learn to generate precise timings of their outputs (Johansson et al., 2014). The brain also appears to use inverse models to rapidly predict motor activity that would give rise to a given sensory target (Hanuschkin et al., 2013; Giret et al., 2014). Such mechanisms could be put to use far beyond motor control, in bootstrapping the training of a larger architecture by exploiting continuously changing error signals to update a real-time model of the system state.

---

[48]Conventionally, models of the basal ganglia involve all or none gating of an action, but recent evidence suggests that the basal ganglia may also have continuous, analog outputs (Yttri and Dudman, 2016).

[49]It has been suggested that the basic role of the BG is to provide tonic inhibition to other circuits (Grillner et al., 2005). Release of this inhibition can then activate a "discrete" action, such as a motor command. A core function of the BG is thus to choose, based on patterns detected in its input, which of a finite set of actions to initiate via such release of inhibition. In many models of the basal ganglia's role in cognitive control, the targets of inhibition are thalamic relays (Sherman, 2005), which are set in a default "off" state by tonic inhibition from the basal ganglia. Upon disinhibition of a relay, information is transferred from one cortical location to another—a form of conditional "gating" of information transfer. For example, the BG might be able to selectively "clamp" particular groups of cortical neurons in a fixed state, while leaving others free to learn and adapt. It could thereby enforce complex training routines, perhaps similar to those used to force the emergence of disentangled representations in (Kulkarni et al., 2015). The idea that the basal ganglia can train the cortex is not new, and already appears to have considerable experimental and anatomical support (Pasupathy and Miller, 2005; Ashby et al., 2007, 2010; Turner and Desmurget, 2010).

4.3.2. Hierarchical Control

Importantly, many of the control problems we appear to be solving are hierarchical. We have a spinal cord, which deals with the fast signals coming from our muscles and proprioception. Within neuroscience, it is generally assumed that this system deals with fast feedback loops and that this behavior is learned to optimize its own cost function. The nature of cost functions in motor control is still under debate. In particular, the timescale over which cost functions operate remains unclear: motor optimization may occur via real-time responses to a cost function that is computed and optimized online, or via policy choices that change over time more slowly in response to the cost function (Körding, 2007). Nevertheless, the effect is that central processing in the brain has an effectively simplified physical system to control, e.g., one that is far more linear. So the spinal cord itself already suggests the existence of two levels of a hierarchy, each trained using different cost functions.

However, within the computational motor control literature (see e.g., DeWolf and Eliasmith, 2011), this idea can be pushed far further, e.g., with a hierarchy including spinal cord, M1, PMd, frontal, prefrontal areas. A low level may deal with muscles, the next level may deal with getting our limbs to places or moving objects, a next layer may deal with solving simple local problems (e.g., navigating across a room) while the highest levels may deal with us planning our path through life. This factorization of the problem comes with multiple aspects: First, each level can be solved with its own cost functions, and second, every layer has a characteristic timescale. Some levels, e.g., the spinal cord, must run at a high speed. Other levels, e.g., high-level planning, only need to be touched much more rarely. Converting the computationally hard optimal control problem into a hierarchical approximation promises to make it dramatically easier.

Does the brain solve control problems hierarchically? There is evidence that the brain uses such a strategy (Botvinick et al., 2009; Botvinick and Weinstein, 2014), beside neural network demonstrations (Wayne and Abbott, 2014). The brain may use specialized structures at each hierarchical level to ensure that each operates efficiently given the nature of its problem space and available training signals. At higher levels, these systems may use an abstract syntax for combining sequences of actions in pursuit of goals (Allen et al., 2010). Subroutines in such processes could be derived by a process of chunking sequences of actions into single actions (Graybiel, 1998; Botvinick and Weinstein, 2014). Some brain areas like Broca's area, known for its involvement in language, also appear to be specifically involved in processing the hierarchical structure of behavior, as such, as opposed to its detailed temporal structure (Koechlin and Jubault, 2006).

At the highest level of the decision making and control hierarchy, human reward systems reflect changing goals and subgoals, and we are only beginning to understand how goals are actually coded in the brain, how we switch between goals, and how the cost functions used in learning depend on goal state (Buschman and Miller, 2014; O'Reilly et al., 2014b; Pezzulo et al., 2014). Goal hierarchies are beginning to be incorporated into deep learning (Kulkarni et al., 2016).

Given this hierarchical structure, the optimization algorithms can be fine-tuned. For the low levels, there is sheer unlimited training data. For the high levels, a simulation of the world may be simple, with a tractable number of high-level actions to choose from. Finally, each area needs to give reinforcement to other areas, e.g., high levels need to punish lower levels for making planning complicated. Thus this type of architecture can simplify the learning of control problems.

Progress is being made in both neuroscience and machine learning on finding potential mechanisms for this type of hierarchical planning and goal-seeking. This is beginning to reveal mechanisms for chunking goals and actions and for searching and pruning decision trees (O'Reilly et al., 2014a; Huys et al., 2015; Balaguer et al., 2016; Krishnamurthy et al., 2016; Tamar et al., 2016). The study of model-based hierarchical reinforcement learning and prospective optimization (Sejnowski and Poizner, 2014), which concerns the planning and evaluation of nested sequences of actions, implicates a network coupling the dorsolateral prefontral and orbitofrontal cortex, and the ventral and dorsolateral striatum (Botvinick et al., 2009). Hierarchical RL relies on a hierarchical representation of state and action spaces, and it has been suggested that error-driven learning of an optimal such representation in the hippocampus[50] gives rise to place and grid cell properties (Stachenfeld, 2014), with goal representations themselves emerging in the amygdala, prefrontal cortex and other areas (O'Reilly et al., 2014a).

The question of how control problems can be successfully divided into component problems remains one of the central questions in neuroscience (Wolpert and Flanagan, 2016) and machine learning (Kulkarni et al., 2016), and the cost functions involved in learning to create such decompositions are still unknown. These considerations may begin to make plausible, however, how the brain could not only achieve its remarkable feats of motor learning—such as generating complex "innate" motor programs, like walking in the newborn gazelle almost immediately after birth—but also the kind of planning that allows a human to prepare a meal or travel from London to Chicago.

4.3.3. Spatial Planning

Spatial planning requires solving shortest-path problems subject to constraints. If we want to get from one location to another, there are an arbitrarily large number of simple paths that could be taken. Most naive implementations of such shortest paths problems are grossly inefficient. It appears that, in animals, the hippocampus aids—at least in part through "place cell" and "grid cell" systems—in efficient learning about new environments and in targeted navigation in such environments (Brown et al., 2016). Interestingly, once an environment becomes familiar, it appears that areas of the neocortex can take over the role of navigation (Hasselmo and Stern, 2015).

In some simple models, targeted navigation in the hippocampus is achieved via the dynamics of "bump attractors" or propagating waves in a place cell network with Hebbian plasticity and adaptation (Hopfield, 2009; Buzsáki and Moser, 2013; Ponulak and Hopfield, 2013), which allows the network to effectively chart out a path in the space of place cell

---

[50]Like many brain areas, the hippocampus is richly innervated by a variety of reward-related and other neuromodulatory systems (Verney et al., 1985; Colino and Halliwell, 1987; Hasselmo and Wyble, 1997).

representations. Other navigation models make use of the grid cell system. The place cell network may[51] take input from a grid cell network that computes precise distances and directions, perhaps by integrating head direction and velocity signals—grid cells fire when the animal is on any node of a regularly spaced hexagonal grid. Different parts of the entorhinal cortex contain grid cells with different grid spacings, and place cells may combine information from multiple such grids in order to build up responses to particular single positions. These systems are highly structured temporally, e.g., containing nested gamma and theta oscillation structures that are phased locked to sequences of place-cell responses, interfering oscillators frequency-shifted by the animal's motion velocity (Zilli and Hasselmo, 2010), tuned cellular resonances (Giocomo et al., 2007; Buzsáki, 2010), and other neural phenomena that lie far outside a conventional artificial neural network description. It seems that an intricate interplay of spatial and temporal network structures may be essential for encoding sequences of spatiotemporal events across multiple scales, and using them to drive multiple forms of learning, e.g., supporting forward and reverse sequence replay with various temporal compression factors (Buzsáki, 2010).

Higher-level cognitive tasks such as prospective planning appear to share computational sub-problems with path-finding (Hassabis and Maguire, 2009)[52]. Interaction between hippocampus and prefrontal cortex could perhaps support a more abstract notion of "navigation" in a space of goals and sub-goals. Interestingly, there is preliminary evidence from fMRI that abstract concepts are also represented according to grid-cell-like hexagonal grid structures in humans (Constantinescu et al., 2016), as well as preliminary evidence that social relationships may also be represented through a hippocampal map (Tavares et al., 2015). Having specialized structures for path-finding could thus simplify a variety of computational problems at different levels of abstraction.

4.3.4. Variable Binding
Language and reasoning appear to present a problem for neural networks (Minsky, 1991; Marcus, 2001; Hadley, 2009): we seem to be able to apply common grammatical rules to sentences regardless of the content of those sentences, and regardless of whether we have ever seen even remotely similar sentences in the training data. While this is achieved automatically in a computer with fixed registers, location addressable memories, and hard-coded operations, how it could be achieved in a biological brain, or emerge from an optimization algorithm, has been under debate for decades.

As the putative key capability underlying such operations, variable binding has been defined as "the transitory or permanent tying together of two bits of information: a variable (such as an X or Y in algebra, or a placeholder like subject or verb

in a sentence) and an arbitrary instantiation of that variable (say, a single number, symbol, vector, or word)" (Marcus et al., 2014a,b). A number of potential biologically plausible binding mechanisms (Eliasmith et al., 2012; Hayworth, 2012; Kriete et al., 2013; Goertzel, 2014) are reviewed in Marcus et al. (2014a) and Marcus et al. (2014b). Some, such as vector symbolic architectures[53], which were proposed in cognitive science (Plate, 1995; Stewart and Eliasmith, 2009; Eliasmith, 2013), are also being considered in the context of efficiently-trainable artificial neural networks (Danihelka et al., 2016)—in effect, these systems learn how to use variable binding.

Variable binding could potentially emerge from simpler memory systems. For example, the Scrub-Jay can remember the place and time of last visit for hundreds of different locations, e.g., to determine whether high-quality food is currently buried at any given location (Clayton and Dickinson, 1998). It is conceivable that such spatially-grounded memory systems enabled a more general binding mechanism to emerge during evolution, perhaps through integration with routing systems or other content-addressable or working memory systems.

4.3.5. Hierarchical Syntax
Fixed, static hierarchies (e.g., the hierarchical organization of cortical areas Felleman and Van Essen, 1991) only take us so far: to deal with long chains of arbitrary nested references, we need *dynamic* hierarchies that can implement recursion on the fly. Human language syntax has a hierarchical structure, which Berwick et al described as "composition of smaller forms like words and phrases into larger ones" (Berwick et al., 2012; Miyagawa et al., 2013). The extent of recursion in human language and thought may be captured by a class of automata known as higher-order pushdown automata, which can be implemented via finite state machines with access to nested stacks (Rodriguez and Granger, 2016). Specific fronto-temporal networks may be involved in representing and generating such hierarchies (Dehaene et al., 2015), e.g., with the hippocampal system playing a key role in implementing some analog of a pushdown stack (Rodriguez and Granger, 2016)[54].

Little is known about the underlying circuit mechanisms for such dynamic hierarchies, but it is clear that specific affordances for representing such hierarchies in an efficient way would be beneficial. This may be closely connected with the issue of variable binding, and it is possible that operations similar to pointers could be useful in this context, in both the brain and artificial neural networks (Kriete et al., 2013; Kurach et al., 2015). Augmenting neural networks with a differentiable analog of a

---

[51]It remains unclear whether place cells take input from the grid cell system or vice versa (Hasselmo, 2015).

[52]Other spatial problems such as mental rotation may require learning architectures specialized for geometric coordinate transformations (Hinton et al., 2011; Jaderberg et al., 2015) or binding mechanisms that support structural, compositional, parametric descriptions of a scene (Hayworth et al., 2011).

[53]There is some direct fMRI evidence for anatomically separate registers representing the contents of different sentence roles in the human brain (Frankland and Greene, 2015), which is suggestive of a possible anatomical binding mechanism, but also consistent with other mechanisms like vector symbolic architectures. More generally, the substrates of symbolic processing in the brain may bear an intimate connection with the representation of objects in working memory in the prefrontal cortex, and specifically with the question of how the PFC represents multiple objects in working memory simultaneously. This question is undergoing extensive study in primates (Warden and Miller, 2007, 2010; Siegel et al., 2009; Rigotti et al., 2013).

[54]There is controversy around claims that recursive syntax is also present in songbirds (Van Heijningen et al., 2009).

push-down stack is another such affordance being pursued in machine learning (Joulin and Mikolov, 2015).

4.3.6. Mental Programs and Imagination

Humans excel at stitching together sub-actions to form larger actions (Verwey, 1996; Acuna et al., 2014; Sejnowski and Poizner, 2014). Structured, serial, hierarchical probabilistic programs have recently been shown to model aspects of human conceptual representation and compositional learning (Lake et al., 2015). In particular, sequential programs were found to enable one-shot learning of new geometric/visual concepts (Lake et al., 2015). Generative programs have also been proposed in the context of scene understanding (Battaglia et al., 2013). The ability to deal with problems in terms of sub-problems is central both in human thought and in many successful algorithms.

One possibility is that the hippocampus supports the rapid construction and learning of sequential programs, e.g., in multi-step planning. An influential idea—known as the "complementary learning systems hypothesis"—is that the hippocampus plays a key role in certain processes where learning must occur quickly on the basis of single episodes, whereas the cortex learns more slowly by aggregating and integrating patterns across large amounts of data (Herd et al., 2013; Leibo et al., 2015a; Blundell et al., 2016; Kumaran et al., 2016). The hippocampus appears to explore, in simulation, possible future trajectories to a goal, even those involving previously unvisited locations (Ólafsdóttir et al., 2015). Hippocampal-prefrontal interaction has been suggested to allow rapid, subconscious evaluation of potential action sequences during decision-making, with the hippocampus in effect simulating the expected outcomes of potential actions that are generated and evaluated in the prefrontal (Mushiake et al., 2006; Wang et al., 2015). The role of the hippocampus in imagination, concept generation (Kumaran et al., 2009), scene construction (Hassabis and Maguire, 2007), mental exploration and goal-directed path planning (Hopfield, 2009; Ólafsdóttir et al., 2015; Brown et al., 2016) suggests that it could help to create generative models to underpin more complex inference such as program induction (Lake et al., 2015) or common-sense world simulation (Battaglia et al., 2013). For example, a sequential, programmatic process, mediated jointly by the basal ganglia, hippocampus and prefrontal cortex might allow one-shot learning of a new concept, as in the sequential computations underlying a process like Bayesian Program Learning (Lake et al., 2015).

Another related possibility is that the cortex itself intrinsically supports the construction and learning of sequential programs (Bach and Herger, 2015). Recurrent neural networks have been used for image generation through a sequential, attention-based process (Gregor et al., 2015), although their correspondence with the brain is unclear[55].

[55]The above mechanisms are spontaneous and subconscious. In conscious thought, too, the brain can clearly visit the multiple layers of a program one after the other. We make high-level plans that we fill with lower-level plans. Humans also have memory for their own thought processes. We have some ability to put "on hold" our current state of mind, start a new train of thought, and then come back to our original thought. We also are able to ask, introspectively, whether we have had a given thought before. The neural basis of these processes is unclear, although one may speculate that the hippocampus is involved.

4.4. Other Specialized Structures

Importantly, there are many other specialized structures known in neuroscience, which arguably receive less attention than they deserve, even for those interested in higher cognition. In the above, in addition to the hippocampus, basal ganglia and cortex, we emphasized the key roles of the thalamus in routing, of the cerebellum as a fast and rapidly trainable control and modeling system, of the amygdala and other areas as a potential source of utility functions, of the retina or early visual areas as a means to generate detectors for motion and other features to bootstrap more complex visual learning, and of the frontal eye fields and other areas as a possible source of attention control. We ignored other structures entirely, whose functions are only beginning to be uncovered, such as the claustrum (Crick and Koch, 2005), which has been speculated to be important for rapidly binding together information from many modalities. Our overall understanding of the functional decomposition of brain circuitry still seems very preliminary.

4.5. Relationships with Other Cognitive Frameworks Involving Specialized Systems

A recent analysis (Lake et al., 2016) suggested directions by which to modify and enhance existing neural-net-based machine learning toward more powerful and human-like cognitive capabilities, particularly by introducing new structures and systems which go beyond data-driven optimization. This analysis emphasized that systems should construct generative models of the world that incorporate compositionality (discrete construction from re-usable parts), inductive biases reflecting causality, intuitive physics and intuitive psychology, and the capacity for probabilistic inference over discrete structured models (e.g., structured as graphs, trees, or programs) (Tervo et al., 2016) to harness abstractions and enable transfer learning.

We view these ideas as consistent with and complementary to the framework of cost functions, optimization and specialized systems discussed here. One might seek to understand how optimization and specialized systems could be used to implement some of the mechanisms proposed in Lake et al. (2016) inside neural networks. Lake et al. (2016) emphasize how incorporating additional structure into trainable neural networks can potentially give rise to systems that use compositional, causal and intuitive inductive biases and that "learn to learn" using structured models and shared data structures. For example, sub-dividing networks into units that can be modularly and dynamically combined, where representations can be copied and routed, may present a path toward improved compositionality and transfer learning (Andreas et al., 2015). The control flow for recombining pre-existing modules and representations could be learned via reinforcement learning (Andreas et al., 2016). How to implement the broad set of mechanisms discussed in Lake et al. (2016) is a key computational problem, and it remains open at which levels (e.g., cost functions and training procedures vs. specialized computational structures vs. underlying neural primitives) architectural innovations will need to be introduced to capture these phenomena.

Primitives that are more complex than those used in conventional neural networks—for instance, primitives that act as state machines with complex message passing (Bach and Herger, 2015) or networks that intrinsically implement Bayesian inference (George and Hawkins, 2009)—could potentially be useful, and it is plausible that some of these may be found in the brain. Recent findings on the power of generic optimization also do not rule out the idea that the brain may explicitly generate and use particular types of structured representations to constrain its inferences; indeed, the specialized brain systems discussed here might provide a means to enforce such constraints. It might be possible to further map the concepts of Lake et al. (2016) onto neuroscience via an infrastructure of interacting cost functions and specialized brain systems under rich genetic control, coupled to a powerful and generic neurally implemented capacity for optimization. For example, it was recently shown that complex probabilistic population coding and inference can arise automatically from backpropagation-based training of simple neural networks (Orhan and Ma, 2016), without needing to be built in by hand. The nature of the underlying primitives in the brain, on top of which learning can operate, is a key question for neuroscience.

5. MACHINE LEARNING INSPIRED NEUROSCIENCE

Hypotheses are primarily useful if they lead to concrete, experimentally testable predictions. As such, we now want to go through the hypotheses and see to which level they can be directly tested, as well as refined, through neuroscience.

5.1. *Hypothesis 1*– Existence of Cost Functions

There are multiple general strategies for addressing whether and how the brain optimizes cost functions. A first strategy is based on observing the endpoint of learning. If the brain uses a cost function, and we can guess its identity, then the final state of the brain should be close to optimal for the cost function. We could thus compare (Güçlü and van Gerven, 2015) receptive fields that are optimized in a simulation, according to a particular cost function, with the measured receptive fields. Various techniques exist to carry out such comparisons in fRMI studies, including population receptive field estimation (Dumoulin and Wandell, 2008; Güçlü and van Gerven, 2015) and representational dissimilarity matrices (Kriegeskorte et al., 2008; Khaligh-Razavi and Kriegeskorte, 2014). This strategy is only beginning to be used at the moment, perhaps because it has been difficult to measure the receptive fields or other representational properties across a large population of *individual* neurons (fMRI operates at a much coarser level), but this situation is beginning to improve technologically with the emergence of large-scale recording methods (Hasselmo, 2015).

A second strategy could directly quantify how well a cost function describes learning. If the dynamics of learning minimize a cost function then the underlying vector field should have a strong gradient descent type component and a weak rotational component, i.e., weight changes will primarily move down the gradient rather than drifting in the nullspace. If we could somehow continuously monitor the synaptic strengths, while externally manipulating them, then we could, in principle, measure the vector field in the space of synaptic weights, and calculate its divergence as well as its rotation. For at least the subset of synapses that are being trained via some approximation to gradient descent, the divergence component should be strong relative to the rotational component. This strategy has not been developed yet due to experimental difficulties with monitoring large numbers of synaptic weights[56].

A third strategy is based on perturbations: cost function based learning should undo the effects of perturbations which disrupt optimality, i.e., the system should return to local minima after a perturbation, and indeed perhaps to the same local minimum after a sufficiently small perturbation. If we change synaptic connections, e.g., in the context of a brain machine interface, we should be able to produce a reorganization that can be predicted based on a guess of the relevant cost function. This strategy is starting to be feasible in motor areas.

Lastly, if we knew structurally which cell types and connections mediated the delivery of error signals vs. input data or other types of connections, then we could stimulate specific connections so as to impose a user-defined cost function. In effect, we would use the brain's own networks as a trainable deep learning substrate, and then study how the network responds to training. Brain machine interfaces can be used to set up specific local learning problems, in which the brain is asked to create certain user-specified representations, and the dynamics of this process can be monitored (Sadtler et al., 2014). Likewise, brain machine interfaces can be used to give the brain access to new datastreams, and to investigate how those datastreams are incorporated into task performance, and whether such incorporation is governed by optimality principles (Dadarlat et al., 2015). In order to do this kind of experiment fully and optimally, we must first understand more about how the system is wired to deliver cost signals. Much of the structure that would be found in connectomic circuit maps, for example, would not just be relevant for short-timescale computing, but also for creating the infrastructure that supports cost functions and their optimization.

Many of the learning mechanisms that we have discussed in this paper make specific predictions about connectivity or dynamics. For example, the "feedback alignment" approach to biological backpropagation suggests that cortical feedback connections should, at some level of neuronal grouping, be largely sign-concordant with the corresponding feedforward connections, although not necessarily of concordant weight (Liao et al., 2015), and feedback alignment also makes predictions for synaptic normalization mechanisms (Liao et al., 2015). The Kickback model for biologically plausible backpropagation has a specific role for NMDA receptors (Balduzzi et al., 2014). Some models that incorporate dendritic coincidence detection for learning temporal sequences predict that a given axon should make only a small number of synapses on a given dendritic

---

[56]Fluorescent techniques like (Hayashi-Takagi et al., 2015) might be helpful.

segment (Hawkins and Ahmad, 2016). Models that involve STDP learning will make predictions about the dynamics of changing firing rates (Hinton, 2007, 2016; Bengio et al., 2015a,b; Bengio and Fischer, 2015), as well as about the particular network structures, such as those based on autoencoders or recirculation, in which STDP can give rise to a form of backpropagation.

It is critical to establish the unit of optimization. We want to know the scale of the modules that are trainable by some approximation of gradient descent optimization. How large are the networks which share a given error signal or cost function? On what scales can appropriate training signals be delivered? It could be that the whole brain is optimized end-to-end, in principle. In this case we would expect to find connections that carry training signals from each layer to the preceding ones. On successively smaller scales, optimization could be within a brain area, a microcircuit[57], or an individual neuron (Mel, 1992; Körding and König, 2000, 2001; Hawkins and Ahmad, 2016). Importantly, optimization may co-exist across these scales. There may be some slow optimization end-to-end, with stronger optimization within a local area and very efficient algorithms within each cell. Careful experiments should be able to identify the scale of optimization, e.g., by quantifying the extent of learning induced by a local perturbation.

The tightness of the structure-function relationship is the hallmark of molecular and to some extent cellular biology, but in large connectionist learning systems, this relationship can become difficult to extract: the same initial network can be driven to compute many different functions by subjecting it to different training[58, 59]. It can be hard to understand the way a neural network solves its problems.

---

[57]The use of structured microcircuits rather than individual neurons as the units of learning can ease the burden on the learning rules possessed by individual neurons, as exemplified by a study implementing Helmholtz machine learning in a network of spiking neurons using conventional plasticity rules (Roudi and Taylor, 2015; Sountsov and Miller, 2015). As a simpler example, the classical problem of how neurons with only one output axon could communicate both activation and error derivatives for backpropagation ceases to be a problem if the unit of optimization is not a single neuron. Similar considerations hold for the issue of weight symmetry, or approximate sign-concordance in the case of feedback alignment (Liao et al., 2015).

[58]Within this framework, networks that adhere to the basic statistics of neural connectivity, electrophysiology and morphology, such as the initial cortical column models from the Blue Brain Project (Markram et al., 2015), would recapitulate some properties of the cortex, but—just like untrained neural networks—would not spontaneously generate complex functional computation without being subjected to a multi-stage training process, naturalistic sensory data, signals arising from other brain areas and action-driven reinforcement signals.

[59]Not only in applied machine learning, but also in today's most advanced neuro-cognitive models such as SPAUN (Eliasmith et al., 2012; Eliasmith, 2013), the detailed local circuit connectivity is obtained through an optimization process of some kind to achieve a particular functionality. In the case of modern machine learning, training is often done via end-to-end backpropagation through an architecture that is only structured at the level of higher-level "blocks" of units, whereas in SPAUN each block is optimized (Eliasmith and Anderson, 2004) separately according to a procedure that allows the blocks to subsequently be stitched together in a coherent way. Technically, the Neural Engineering Framework (Eliasmith and Anderson, 2004) used in SPAUN uses singular value decomposition, rather than gradient descent, to compute the connections weights as optimal linear decoders. This is possible because of a nonlinear mapping into a high-dimensional space, in which approximating any desired function can be done via a hyperplane regression (Tapson and van Schaik, 2013).

How could one tell the difference, then, between a gradient-descent trained network vs. untrained or random networks vs. a network that has been trained against a different kind of task? One possibility would be to train artificial neural networks against various candidate cost functions, study the resulting neural tuning properties (Todorov, 2002), and compare them with those found in the circuit of interest (Zipser and Andersen, 1988). This has already been done to aid the interpretation of the neural dynamics underlying decision making in the PFC (Sussillo, 2014), working memory in the posterior parietal cortex (Rajan et al., 2016) and object or action representation in the visual system (Tacchetti et al., 2016; Yamins and DiCarlo, 2016a,b). Some have gone on to suggest a direct correspondence between cortical circuits and optimized, appropriately regularized (Sussillo et al., 2015), recurrent neural networks (Liao and Poggio, 2016). In any case, effective analytical methods to reverse engineer complex machine learning systems (Jonas and Kording, 2016), and methods to reverse engineer biological brains, may have some commonalities.

Does this emphasis on function optimization and trainable substrates mean that we should give up on reverse engineering the brain based on detailed measurements and models of its specific connectivity and dynamics? On the contrary: we should use large-scale brain maps to try to better understand (a) how the brain implements optimization, (b) where the training signals come from and what cost functions they embody, and (c) what structures exist, at different levels of organization, to constrain this optimization to efficiently find solutions to specific kinds of problems. The answers may be influenced by diverse local properties of neurons and networks, such as homeostatic rules of neural structure, gene expression and function (Marder and Goaillard, 2006), the diversity of synapse types, cell-type-specific connectivity (Jiang et al., 2015), patterns of inter-laminar projection, distributions of inhibitory neuron types, dendritic targeting and local dendritic physiology and plasticity (Markram et al., 2015; Bloss et al., 2016; Morgan et al., 2016; Sandler et al., 2016) or local glial networks (Perea et al., 2009). They may also be influenced by the integrated nature of higher-level brain systems, including mechanisms for developmental bootstrapping (Ullman et al., 2012), information routing (Gurney et al., 2001; Stocco et al., 2010), attention (Buschman and Miller, 2010) and hierarchical decision making (Lee et al., 2015). Mapping these systems in detail is of paramount importance to understanding how the brain works, down to the nanoscale dendritic organization of ion channels and up to the real-time global coordination of cortex, striatum and hippocampus, all of which are computationally relevant in the framework we have explicated here. We thus expect that large-scale, multi-resolution brain maps would be useful in testing these framework-level ideas, in inspiring their refinements, and in using them to guide more detailed analysis.

5.2. *Hypothesis 2– Biological Fine-Structure of Cost Functions*

Clearly, we can map differences in structure, dynamics and representation across brain areas. When we find such differences,

the question remains as to whether we can interpret these as resulting from differences in the internally-generated cost functions, as opposed to differences in the input data, or from differences that reflect other constraints unrelated to cost functions. If we can directly measure aspects of the cost function in different areas, then we can also compare them across areas. For example, methods from inverse reinforcement learning[60] might allow backing out the cost function from observed plasticity (Ng and Russell, 2000).

Moreover, as we begin to understand the "neural correlates" of particular cost functions—perhaps encoded in particular synaptic or neuromodulatory learning rules, genetically-guided local wiring patterns, or patterns of interaction between brain areas—we can also begin to understand when differences in observed neural circuit architecture reflect differences in cost functions.

We expect that, for each distinct learning rule or cost function, there may be specific molecularly identifiable types of cells and/or synapses. Moreover, for each specialized system there may be specific molecularly identifiable developmental programs that tune it or otherwise set its parameters. This would make sense if evolution has needed to tune the parameters of one cost function without impacting others.

How many different types of internal training signals does the brain generate? When thinking about error signals, we are not just talking about dopamine and serotonin, or other classical reward-related pathways. The error signals that may be used to train specific sub-networks in the brain, via some approximation of gradient descent or otherwise, are not necessarily equivalent to reward signals. It is important to distinguish between cost functions that may be used to drive optimization of specific sub-circuits in the brain, and what are referred to as "value functions" or "utility functions," i.e., functions that predict the agent's aggregate future reward. In both cases, similar reinforcement learning mechanisms may be used, but the interpretation of the cost functions is different. We have not emphasized global utility functions for the animal here, since they are extensively studied elsewhere (e.g., O'Reilly et al., 2014a; Bach, 2015), and since we argue that, though important, they are only a part of the picture, i.e., that the brain is not solely an end-to-end reinforcement trained system.

Progress in brain mapping could soon allow us to classify the types of reward signals in the brain, follow the detailed anatomy and connectivity of reward pathways throughout the brain, and map in detail how reward pathways are integrated into striatal, cortical, hippocampal and cerebellar microcircuits. This program is beginning to be carried out in the fly brain, in which twenty specific types of dopamine neuron project to distinct anatomical compartments of the mushroom body to train distinct odor classifiers operating on a set of high-dimensional

odor representations (Caron et al., 2013; Aso et al., 2014a,b; Cohn et al., 2015). It is known that, even within the same system, such as the fly olfactory pathway, some neuronal wiring is highly specific and molecularly programmed (Hattori et al., 2007; Hong and Luo, 2014), while other wiring is effectively random (Caron et al., 2013), and yet other wiring is learned (Aso et al., 2014a). The interplay between such design principles could give rise to many forms of "division of labor" between genetics and learning. Likewise, it is believed that birdsong learning is driven by reinforcement learning using a specialized cost function that relies on comparison with a memorized version of a tutor's song (Fiete et al., 2007), and also that it involves specialized structures for controlling song variability during learning (Aronov et al., 2011). These detailed pathways underlying the construction of cost functions for vocal learning are beginning to be mapped (Mandelblat-Cerf et al., 2014). Starting with simple systems, it should become possible to map the reward pathways and how they evolved and diversified, which would be a step on the way to understanding how the system learns.

These types of mapping efforts would be a first step toward the ability to create a concrete model of the brain's optimization architecture. Our discussion here has focused on trying to anticipate, based on known neuroscience knowledge and on approaches becoming successful in machine learning, the *kinds* of local cost functions that the brain may rely on, and how specialized brain systems may enable efficient solutions to optimization problems. However, this framework-level discussion is not a formal specification, either of the architecture, or of a notion of biologically applied cost function that could be directly measured based on neural data. In order to move toward a more formal specification of the kind of model we are proposing here, it would be useful to map the architecture of the brain's reward systems and to identify other biological pathways that may mediate the generation and delivery of error signals. Based on such maps, one could identify regions which are proposed to be subject to a single cost function. Otherwise, the problem of inference of the cost function, e.g., based on neural dynamics becomes ill-posed: one can define a local cost function for an *arbitrary* dynamics by integrating the trajectory of the system, but this approach in general lacks explanatory power and also, crucially, lacks any circuit-level relationship with the brain's actual neural mechanisms of optimization, i.e., such a defined cost function does not necessarily correspond to the cost functions that the biological machinery is actually organized to optimize. Notably, some of the relevant biological pathways mediating cost functions and error signals may involve key biomolecular or gene expression aspects, not just real-time patterns of neural activity.

Another related consideration, in trying to formalize this type approach and to infer cost functions from neural measurements, is that not all neurons in the circuit may be subject to optimization: after all, some neurons may be needed to generate the error signals themselves, or to mediate the optimization process for other neurons, or to perform other unrelated functions. Furthermore, within a given region, there may be multiple sub-circuits subject to different optimization pressures.

---

[60]There is a rich tradition of trying to estimate the cost function used by human beings (Ng and Russell, 2000; Finn et al., 2016; Ho and Ermon, 2016). The idea is that we observe (by stipulation) behavior that is optimal for the human's cost function. We can then search for the cost function that makes the observed behavior most probable and simultaneously makes the behaviors that could have been observed, but were not, least probable. Extensions of such approaches could perhaps be used to ask which cost functions the brain is optimizing.

It is the claim that the brain actually has structured biological machinery to generate, route and apply specific cost functions that gives substance to our proposal, over and above the trivial claim that many kinds of dynamics can be viewed as optimizations, but our knowledge of this machinery is still limited. This is not to mention the difficulties involved in inferring cost functions in the presence of noise or constraints on the dynamics. Thus, one cannot blindly collect the neurons in an arbitrary region, measure their dynamics, and hope to infer their cost function by solving an inverse problem—instead, a rich interplay between structural mapping, dynamic mapping, hypothesis generation, modeling and perturbation is likely to be necessary in order to gain a detailed knowledge of which cost functions the brain uses and how it does so.

5.3. *Hypothesis 3*– Embedding within a Pre-structured Architecture

If different brain structures are performing distinct types of computations with a shared goal, then optimization of a joint cost function will take place with different dynamics in each area. If we focus on a higher level task, e.g., maximizing the probability of correctly detecting something, then we should find that basic feature detection circuits should learn when the features were insufficient for detection, that attentional routing structures should learn when a different allocation of attention would have improved detection and that memory structures should learn when items that matter for detection were not remembered. If we assume that multiple structures are participating in a joint computation, which optimizes an overall cost function (but see Hypothesis 2), then an understanding of the computational function of each area leads to a prediction of the measurable plasticity rules.

6. NEUROSCIENCE INSPIRED MACHINE LEARNING

Machine learning may be equally transformed by neuroscience. Within the brain, a myriad of subsystems and layers work together to produce an agent that exhibits general intelligence. The brain is able to show intelligent behavior across a broad range of problems using only relatively small amounts of data. As such, progress at understanding the brain promises to improve machine learning. In this section, we review our three hypotheses about the brain and discuss how their elaboration might contribute to more powerful machine learning systems.

6.1. *Hypothesis 1*– Existence of Cost Functions

A good practitioner of machine learning should have a broad range of optimization methods at their disposal as different problems ask for different approaches. The brain, we have argued, is an implicit machine learning mechanism which has been evolved over millions of years. Consequently, we should expect the brain to be able to optimize cost functions efficiently, across many domains and kinds of data. Indeed, across different animal phyla, we even see *convergent* evolution of certain brain

structures (Shimizu and Karten, 2013; Güntürkün and Bugnyar, 2016), e.g., the bird brain has no cortex yet has developed homologous structures which—as the linguistic feats of the African Gray Parrot demonstrate—can give rise to quite complex intelligence. It seems reasonable to hope to learn how to do truly general-purpose optimization by looking at the brain.

Indeed, there are multiple kinds of optimization that we may expect to discover by looking at the brain. At the hardware level, the brain clearly manages to optimize functions efficiently despite having slow hardware subject to molecular fluctuations, suggesting directions for improving the hardware of machine learning to be more energy efficient. At the level of learning rules, the brain solves an optimization problem in a highly nonlinear, non-differentiable, temporally stochastic, spiking system with massive numbers of feedback connections, a problem that we arguably still do not know how to efficiently solve for neural networks. At the architectural level, the brain can optimize certain kinds of functions based on very few stimulus presentations, operates over diverse timescales, and clearly uses advanced forms of active learning to infer causal structure in the world.

While we have discussed a range of theories (O'Reilly, 1996; Körding and König, 2001; Hinton, 2007, 2016; Roelfsema et al., 2010; Balduzzi et al., 2014; Lillicrap et al., 2014; O'Reilly et al., 2014a; Bengio et al., 2015a) for how the brain can carry out optimization, these theories are still preliminary. Thus, the first step is to understand whether the brain indeed performs multi-layer credit assignment in a manner that approximates full gradient descent, and if so, how it does this. Either way, we can expect that answer to impact machine learning. If the brain does *not* do some form of backpropagation, this suggests that machine learning may benefit from understanding the tricks that the brain uses to avoid having to do so. If, on the other hand, the brain does do backpropagation, then the underlying mechanisms clearly can support a very wide range of efficient optimization processes across many domains, including learning from rich temporal data-streams and via unsupervised mechanisms, and the architectures behind this will likely be of long-term value to machine learning[61]. Moreover, the search for biologically

---

[61] Successes of deep learning are already being used, speculatively, to rationalize features of the brain. It has been suggested that large networks, with many more neurons available than are strictly needed for the target computation, make learning easier (Goodfellow et al., 2014b). In concordance with this, visual cortex appears to be a 100-fold over-complete representation of the retinal output (Lewicki and Sejnowski, 2000). Likewise, it has been suggested that biological neurons stabilized (Turrigiano, 2012) to operate far below their saturating firing rates mirror the successful use of rectified linear units in facilitating the training of artificial neural networks (Roudi and Taylor, 2015). Hinton and others have also suggested a biological motivation (Roudi and Taylor, 2015) for "dropout" regularization (Srivastava et al., 2014), in which a fraction of hidden units is stochastically set to zero during each round of training: such a procedure may correspond to the noisiness of neural spike trains, although other theories interpret spikes as sampling in probabilistic inference (Buesing et al., 2011), or in many other ways. Randomness of spiking has some support in neuroscience (Softky and Koch, 1993), although recent experiments suggest that spike trains in certain areas may be less noisy than previously thought (Hires et al., 2015). The key role of proper initialization in enabling effective gradient descent is an important recent finding (Saxe et al., 2013; Sutskever and Martens, 2013) which may also be reflected by biological mechanisms of neural homeostasis or self-organization that would enforce appropriate initial conditions for learning. Retinal

plausible forms of backpropagation has already led to interesting insights, such as the possibility of using random feedback weights (feedback alignment) in backpropagation (Lillicrap et al., 2014), or the unexpected power of internal FORCE learning in chaotic, spontaneously active recurrent networks (Sussillo and Abbott, 2009). This and other findings discussed here suggest that there are still fundamental things we don't understand about backpropagation—which could potentially lead not only to more biologically plausible ways to train recurrent neural networks, but also to fundamentally simpler and more powerful ones.

6.2. *Hypothesis 2*– Biological Fine-structure of Cost Functions

A good practitioner of machine learning has access to a broad range of learning techniques and thus implicitly is able to use many different cost functions. Some problems ask for clustering, others for extracting sparse variables, and yet others for prediction quality to be maximized. The brain also needs to be able to deal with many different kinds of datasets. As such, it makes sense for the brain to use a broad range of cost functions appropriate for the diverse set of tasks it has to solve to thrive in this world.

Many of the most notable successes of deep learning, from language modeling (Sutskever et al., 2011), to vision (Krizhevsky et al., 2012), to motor control (Levine et al., 2015), have been driven by end-to-end optimization of single task objectives. We have highlighted cases where machine learning has opened the door to multiplicities of cost functions that shape network modules into specialized roles. We expect that machine learning will increasingly adopt these practices in the future.

In computer vision, we have begun to see researchers re-appropriate neural networks trained for one task (e.g., ImageNet classification) and then deploy them on new tasks other than the ones they were trained for or for which more limited training data is available (Oquab et al., 2014; Yosinski et al., 2014; Noroozi and Favaro, 2016). We imagine this procedure will be generalized, whereby, in series and in parallel, diverse training problems, each with an associated cost function, are used to shape visual representations. For example, visual data streams can be segmented into elements like foreground vs. background, objects that can move of their own accord vs. those that cannot, all using diverse unsupervised criteria (Ullman et al., 2012; Poggio, 2015). Networks so trained can then be shared, augmented, and retrained on new tasks. They can be introduced as front-ends for systems that perform more complex objectives or even serve to produce cost functions for training other circuits (Watter et al., 2015). As a simple example, a network that can discriminate between images of different kinds of architectural structures (pyramid, staircase, etc.) could act as a critic for a building-construction network.

Scientifically, determining the order in which cost functions are engaged in the biological brain will inform machine

fixation has been tentatively connected with robustness of convolutional networks to adversarial perturbations in images (Luo et al., 2015). But making these speculative claims of biological relevance more rigorous will require researchers to first evaluate *whether* biological neural circuits are performing multi-layer optimization of cost functions in the first place.

learning about how to construct systems with intricate and hierarchical behaviors via divide-and-conquer approaches to learning problems, active learning, and more.

6.3. *Hypothesis 3*– Embedding within a Pre-structured Architecture

A good practitioner of machine learning should have a broad range of algorithms at their disposal. Some problems are efficiently solved through dynamic programming, others through hashing, and yet others through multi-layer backpropagation. The brain needs to be able to solve a broad range of learning problems without the luxury of being reprogrammed. As such, it makes sense for the brain to have specialized structures that allow it to rapidly learn to approximate a broad range of algorithms.

The first neural networks were simple single-layer systems, either linear or with limited non-linearities (Rashevsky, 1939). The explosion of neural network research in the 1980s (Rumelhart et al., 1986) saw the advent of multilayer networks, followed by networks with layer-wise specializations as in convolutional networks (Fukushima, 1980; LeCun and Bengio, 1995). In the last two decades, architectures with specializations for holding variables stable in memory like the LSTM (Hochreiter and Schmidhuber, 1997), the control of content-addressable memory (Graves et al., 2014; Weston et al., 2014), and game playing by reinforcement learning (Mnih et al., 2015) have been developed. These networks, though formerly exotic, are now becoming mainstream algorithms in the toolbox of any deep learning practitioner. There is no sign that progress in developing new varieties of structured architectures is halting, and the heterogeneity and modularity of the brain's circuitry suggests that diverse, specialized architectures are needed to solve the diverse challenges that confront a behaving animal.

The brain combines a jumble of specialized structures in a way that works. Solving this problem *de novo* in machine learning promises to be very difficult, making it attractive to be inspired by observations about how the brain does it. An understanding of the breadth of specialized structures, as well as the architecture that combines them, should be quite useful.

7. DID EVOLUTION SEPARATE COST FUNCTIONS FROM OPTIMIZATION ALGORITHMS?

Deep learning methods have taken the field of machine learning by storm. Driving the success is the separation of the problem of learning into two pieces: (1) An algorithm, backpropagation, that allows efficient distributed optimization, and (2) Approaches to turn any given problem into an optimization problem, by designing a cost function and training procedure which will result in the desired computation. If we want to apply deep learning to a new domain, e.g., playing Jeopardy, we do not need to change the optimization algorithm—we just need to cleverly set up the right cost function. A lot of work in deep learning, perhaps the majority, is now focused on setting up the right cost functions.

We hypothesize that the brain also acquired such a separation between optimization mechanisms and cost functions. If neural

circuits, such as in cortex, implement a general-purpose optimization algorithm, then any improvement to that algorithm will improve function across the cortex. At the same time, different cortical areas solve different problems, so tinkering with each area's cost function is likely to improve its performance. As such, functionally and evolutionarily separating the problems of optimization and cost function generation could allow evolution to produce better computations, faster. For example, common unsupervised mechanisms could be combined with area-specific reinforcement-based or supervised mechanisms and error signals, much as recent advances in machine learning have found natural ways to combine supervised and unsupervised objectives in a single system (Rasmus and Berglund, 2015).

This suggests interesting questions[62]: When did the division between cost functions and optimization algorithms occur? How is this separation implemented? How did innovations in cost functions and optimization algorithms evolve? And how do our own cost functions and learning algorithms differ from those of other animals?

There are many possibilities for how such a separation might be achieved in the brain. Perhaps the six-layered cortex represents a common optimization algorithm, which in different cortical areas is supplied with different cost functions. This claim is different from the claim that all cortical areas use a single unsupervised learning algorithm and achieve functional specificity by tuning the inputs to that algorithm. In that case, both the optimization mechanism and the implicit unsupervised cost function would be the same across areas (e.g., minimization of prediction error), with only the training data differing between areas, whereas in our suggestion, the optimization mechanism would be the same across areas but the cost function, *as well as* the training data, would differ. Thus the cost function itself would be like an ancillary input to a cortical area, in addition to its input and output data. Some cortical microcircuits

could then, perhaps, compute the cost functions that are to be delivered to other cortical microcircuits. Another possibility is that, within the same circuitry, certain aspects of the wiring and learning rules specify an optimization mechanism and are relatively fixed across areas, while others specify the cost function and are more variable. This latter possibility would be similar to the notion of cortical microcircuits as molecularly and structurally configurable elements, akin to the cells in a field-programmable gate array (FPGA) (Marcus et al., 2014a,b), rather than a homogenous substrate. The biological nature of such a separation, if any exists, remains an open question. For example, individual parts of a neuron may separately deal with optimization and with the specification of the cost function, or different parts of a microcircuit may specialize in this way, or there may be specialized types of cells, some of which deal with signal processing and others with cost functions.

8. CONCLUSIONS

Due to the complexity and variability of the brain, pure "bottom up" analysis of neural data faces potential challenges of interpretation (Robinson, 1992; Jonas and Kording, 2016). Theoretical frameworks can potentially be used to constrain the space of hypotheses being evaluated, allowing researchers to first address higher-level principles and structures in the system, and then "zoom in" to address the details. Proposed "top down" frameworks for understanding neural computation include entropy maximization, efficient encoding, faithful approximation of Bayesian inference, minimization of prediction error, attractor dynamics, modularity, the ability to subserve symbolic operations, and many others (Pinker, 1999; Marcus, 2001; Bialek, 2002; Knill and Pouget, 2004; Bialek et al., 2006; Friston, 2010). Interestingly, many of the "top down" frameworks boil down to assuming that the brain simply optimizes a single, given cost function for a single computational architecture. We generalize these proposals assuming both a heterogeneous combination of cost functions unfolding over development, and a diversity of specialized sub-systems.

Much of neuroscience has focused on the search for "the neural code," i.e., it has asked which stimuli are good at driving activity in individual neurons, regions, or brain areas. But, if the brain is capable of generic optimization of cost functions, then we need to be aware that rather simple cost functions can give rise to complicated stimulus responses. This potentially leads to a different set of questions. Are differing cost functions indeed a useful way to think about the differing functions of brain areas? How does the optimization of cost functions in the brain actually occur, and how is this different from the implementations of gradient descent in artificial neural networks? What additional constraints are present in the circuitry that remain fixed while optimization occurs? How does optimization interact with a structured architecture, and is this architecture similar to what we have sketched? Which computations are wired into the architecture, which emerge through optimization, and which arise from a mixture of those two extremes? To what extent are cost functions explicitly computed in the brain, vs. implicit in its local learning rules? Did the brain evolve to separate the

---

[62]It would be interesting to study these questions in specific brain systems. The primary visual cortex, for example, is still only understood very incompletely (Olshausen and Field, 2004). It serves as a key input modality to both the ventral and dorsal visual pathways, one of which seems to specialize in object identity and the other in motion and manipulation. Higher-level areas like STP draw on both streams to perform tasks like complex action recognition. In some models (e.g., Jhuang et al., 2007), both ventral and dorsal streams are structured hierarchically, but the ventral stream primarily makes use of the spatial filtering properties of V1, whereas the dorsal stream primarily makes use of its spatio-*temporal* filtering properties, e.g., temporal frequency filtering by the space-time receptive fields of V1 neurons. Given this, we can ask interesting questions about V1. Within a framework of multilayer optimization, do both dorsal and ventral pathways impose cost functions that help to shape V1's response properties? Or is V1 largely pre-structured by genetics and local self-organization, with different optimization principles in the ventral and dorsal streams only having effects at higher levels of the hierarchy? Or, more likely, is there some interplay between pre-structuring of the V1 circuitry and optimization according to multiple cost functions? Relatedly, what establishes the differing roles of the downstream ventral vs. dorsal cortical areas, and can their differences be attributed to differing cost functions? This relates to ongoing questions about the basic nature of cortical circuitry. For example, DiCarlo et al. (2012) suggests that visual cortical regions containing on the order of 10000 neurons are locally optimized to perform disentangling of the manifolds corresponding to their local views of the transformations of an object, allowing these manifolds to be linearly separated by readout areas. Yet, DiCarlo et al. (2012) also emphasizes the possibility that certain computations such as normalization are pre-initialized in the circuitry prior to learning-based optimization.

mechanisms involved in cost function generation from those involved in the optimization of cost functions, and if so how? What kinds of meta-level learning might the brain apply, to learn when and how to invoke different cost functions or specialized systems, among the diverse options available, to solve a given task? What crucial mechanisms are left out of this framework? A more in-depth dialog between neuroscience and machine learning could help elucidate some of these questions.

Much of machine learning has focused on finding ever faster ways of doing end-to-end gradient descent in neural networks. Neuroscience may inform machine learning at multiple levels. The optimization algorithms in the brain have undergone a couple of hundred million years of evolution. Moreover, the brain may have found ways of using heterogeneous cost functions that interact over development so as to simplify learning problems by guiding and shaping the outcomes of unsupervised learning. Lastly, the specialized structures evolved in the brain may inform us about ways of making learning efficient in a world that requires a broad range of computational problems to be solved over multiple timescales. Looking at the insights from neuroscience may help machine learning move toward general intelligence in a structured heterogeneous world with access to only small amounts of supervised data.

In some ways our proposal is opposite to many popular theories of neural computation. There is not one mechanism of optimization but (potentially) many, not one cost function but a host of them, not one kind of a representation but a representation of whatever is useful, and not one homogeneous structure but a large number of them. All these elements are held together by the optimization of internally generated cost functions, which allows these systems to make good use of one another. Rejecting simple unifying theories is in line with a broad range of previous approaches in AI. For example, Minsky and Papert's work on the Society of Mind (Minsky, 1988)— and more broadly on ideas of genetically staged and internally bootstrapped development in connectionist systems (Minsky,

1977)—emphasizes the need for a system of internal monitors and critics, specialized communication and storage mechanisms, and a hierarchical organization of simple control systems.

At the time these early works were written, it was not yet clear that gradient-based optimization could give rise to powerful feature representations and behavioral policies. One can view our proposal as a renewed argument against simple end-to-end training and in favor of a heterogeneous approach. In other words, this framework could be viewed as proposing a kind of "society" of cost functions and trainable networks, permitting internal bootstrapping processes reminiscent of the Society of Mind (Minsky, 1988). In this view, intelligence is enabled by many computationally specialized structures, each trained with its own developmentally regulated cost function, where both the structures and the cost functions are themselves optimized by evolution like the hyperparameters in neural networks.

AUTHOR CONTRIBUTION

All authors contributed ideas and co-wrote the paper.

REFERENCES

Abbott, L., DePasquale, B., and Memmesheimer, R. (2016). *Building Functional Networks of Spiking Model Neurons*. Available online at: neurotheory.columbia.edu.

Abbott, L. F., and Blum, K. I. (1996). Functional significance of long-term potentiation for sequence learning and prediction. *Cereb. Cortex* 6, 406–416.

Ackley, D., Hinton, G., and Sejnowski, T. (1958) A learning algorithm for Boltzmann machines. *Cogn. Sci.* 9, 147–169.

Acuna, D. E., Wymbs, N. F., Reynolds, C. A., Picard, N., Turner, R. S., Strick, P. L., et al. (2014). Multifaceted aspects of chunking enable robust algorithms. *J. Neurophysiol.* 112, 1849–1856. doi: 10.1152/jn.00028.2014

Alain, G., Lamb, A., Sankar, C., Courville, A., and Bengio, Y. (2015). Variance reduction in SGD by distributed importance sampling. arXiv:1511.06481.

Allen, K., Ibara, S., and Seymour, A. (2010). Abstract structural representations of goal-directed behavior. *Psychol. Sci.* 21, 1518–1524. doi: 10.1177/0956797610383434

Anderson, C. H., and Van Essen, D. C. (1987). Shifter circuits: a computational strategy for dynamic aspects of visual processing. *Proc. Natl. Acad. Sci. U.S.A.* 84, 6297–6301.

Anderson, J. R. (2007). *How Can the Human Mind Occur in the Physical Universe?* Oxford, UK: Oxford University Press.

Andreas, J., Rohrbach, M., Darrell, T., and Klein, D. (2015). Deep compositional question answering with neural module networks. arXiv:1511.02799.

Andreas, J., Rohrbach, M., Darrell, T., and Klein, D. (2016). Learning to compose neural networks for question answering. arXiv:1601.01705.

Angelucci, A., Levitt, J. B., Walton, E. J., Hupe, J.-M., Bullier, J., and Lund, J. S. (2002). Circuits for local and global signal integration in primary visual cortex. *J. Neurosci.* 22, 8633–8646. Available online at: http://www.jneurosci.org/content/22/19/8633.long

Anselmi, F., Leibo, J. Z., Rosasco, L., Mutch, J., Tacchetti, A., and Poggio, T. (2015). Unsupervised learning of invariant representations. *Theor. Comput. Sci.* 633, 112–121. doi: 10.1016/j.tcs.2015.06.048

Antic, S. D., Zhou, W.-L., Moore, A. R., Short, S. M., and Ikonomu, K. D. (2010). The decade of the dendritic nmda spike. *J. Neurosci. Res.* 88, 2991–3001. doi: 10.1002/jnr.22444

Arancio, O., Kiebler, M., Lee, C. J., Lev-Ram, V., Tsien, R. Y., Kandel, E. R., et al. (1996). Nitric oxide acts directly in the presynaptic neuron to produce long-term potentiation in cultured hippocampal neurons. *Cell* 87, 1025–1035.

Aronov, D., Veit, L., Goldberg, J. H., and Fee, M. S. (2011). Two distinct modes of forebrain circuit dynamics underlie temporal patterning in the vocalizations of young songbirds. *J. Neurosci.* 31, 16353–16368. doi: 10.1523/JNEUROSCI.3009-11.2011

Arora, S., Liang, Y., and Ma, T. (2015). Why are deep nets reversible: a simple theory, with implications for training. arXiv:1511.05653.

Ashby, F. G., Ennis, J. M., and Spiering, B. J. (2007). A neurobiological theory of automaticity in perceptual categorization. *Psychol. Rev.* 114:632. doi: 10.1037/0033-295X.114.3.632

Ashby, F. G., Turner, B. O., and Horvitz, J. C. (2010). Cortical and basal ganglia contributions to habit learning and automaticity. *Trends Cogn. Sci.* 14, 208–215. doi: 10.1016/j.tics.2010.02.001

Aso, Y., Hattori, D., Yu, Y., Johnston, R. M., Iyer, N. A., Ngo, T.-T. B., et al. (2014a). The neuronal architecture of the mushroom body provides a logic for associative learning. *eLife* 3:e04577. doi: 10.7554/eLife.04577

Aso, Y., Sitaraman, D., Ichinose, T., Kaun, K. R., Vogt, K., Belliart-Guérin, G., et al. (2014b). Mushroom body output neurons encode valence and guide memory-based action selection in Drosophila. *eLife* 3:e04580. doi: 10.7554/eLife.04580

Bach, J. (2015). "Modeling motivation in MicroPsi 2," in *8th International Conference, AGI 2015, AGI 2015, Berlin, Germany, Proceedings*, Vol. 9205 (Cham: Springer International Publishing), 3–13.

Bach, J., and Herger, P. (2015). "Request confirmation networks for neuro-symbolic script execution," in *Workshop on Cognitive Computation: Integrating Neural and Symbolic Approaches at NIPS*, eds T. Besold, A. Avila Garcez, G. Marcus, and R. Miikkulainen (Montreal, QC).

Baillargeon, R., Scott, R. M., and Bian, L. (2016). Psychological reasoning in infancy. *Annu. Rev. Psychol.* 67, 159–186. doi: 10.1146/annurev-psych-010213-115033

Ba, J., and Caruana, R. (2014). Do deep nets really need to be deep? *Adv. Neural Inform. Process.* 27, 2654–2662. Available online at: https://papers.nips.cc/paper/5484-do-deep-nets-really-need-to-be-deep

Balaguer, J., Spiers, H., Hassabis, D., and Summerfield, C. (2016). Neural mechanisms of hierarchical planning in a virtual subway network. *Neuron* 90, 893–903. doi: 10.1016/j.neuron.2016.03.037

Baldauf, D., and Desimone, R. (2014). Neural mechanisms of object-based attention. *Science* 344, 424–427. doi: 10.1126/science.1247003

Baldi, P., and Sadowski, P. (2015). The Ebb and flow of deep learning: a theory of local learning. arXiv:1506.06472.

Balduzzi, D. (2014). "Cortical prediction markets," in *Proceedings of the 2014 International Conference on Autonomous AgentsMultiagent Systems (AAMAS)* (Paris).

Balduzzi, D., Vanchinathan, H., and Buhmann, J. (2014). Kickback cuts Backprop's red-tape: biologically plausible credit assignment in neural networks. arXiv:1411.6191.

Bargmann, C. I. (2012). Beyond the connectome: how neuromodulators shape neural circuits. *Bioessays* 34, 458–465. doi: 10.1002/bies.201100185.

Bargmann, C. I., and Marder, E. (2013). From the connectome to brain function. *Nat. Methods* 10, 483–490. doi: 10.1038/nmeth.2451

Bastos, A. M., Vezoli, J., Bosman, C. A., Schoffelen, J.-M., Oostenveld, R., Dowdall, J. R., et al. (2015). Visual areas exert feedforward and feedback influences through distinct frequency channels. *Neuron* 85, 390–401. doi: 10.1016/j.neuron.2014.12.018.

Battaglia, P. W., Hamrick, J. B., and Tenenbaum, J. B. (2013). Simulation as an engine of physical scene understanding. *Proc. Natl. Acad. Sci. U.S.A.* 110, 18327–18332. doi: 10.1073/pnas.1306572110

Becker, S., and Hinton, G. E. (1992). Self-organizing neural network that discovers surfaces in random-dot stereograms. *Nature* 355, 161–163.

Bekolay, T., Kolbeck, C., and Eliasmith, C. (2013). "Simultaneous unsupervised and supervised learning of cognitive functions in biologically plausible spiking neural networks," in *Proceedings of the 35th Annual Conference of the Cognitive Science Society* (Berlin), 169–174.

Bengio, Y. (2014). How auto-encoders could provide credit assignment in deep networks via target propagation. arXiv:1407.7906.

Bengio, Y., and Fischer, A. (2015). Early inference in energy-based models approximates back-propagation. arXiv:1510.02777.

Bengio, Y., Lee, D.-H., Bornschein, J., and Lin, Z. (2015a). Towards biologically plausible deep learning. arXiv:1502.04156.

Bengio, Y., Louradour, J., Collobert, R., and Weston, J. (2009). "Curriculum learning," in *Proceedings of the 26th Annual International Conference on Machine Learning* (Montreal, QC), 41–48.

Bengio, Y., Mesnard, T., Fischer, A., Zhang, S., and Wu, Y. (2015b). STDP as presynaptic activity times rate of change of postsynaptic activity. arXiv:1509.05936.

Bengio, Y., Scellier, B., Bilaniuk, O., Sacramento, J., and Senn, W. (2016). Feedforward initialization for fast inference of deep generative networks is biologically plausible. arXiv:1606.01651.

Berezovskii, V. K., Nassi, J. J., and Born, R. T. (2011). Segregation of feedforward and feedback projections in mouse visual cortex. *J. Comp. Neurol.* 519, 3672–3683. doi: 10.1002/cne.22675

Berwick, R. C., Beckers, G. J. L., Okanoya, K., and Bolhuis, J. J. (2012). A bird's eye view of human language evolution. *Front. Evol. Neurosci.* 4:5. doi: 10.3389/fnevo.2012.00005

Bialek, W. (2002). "Thinking about the brain," in *Physics of Bio-Molecules and Cells*, Vol. 75, eds F. Flyvbjerg, F. Jülicher, P. Ormos, and F. David (Berlin; Heidelberg: Springer), 485–578.

Bialek, W., De Ruyter Van Steveninck, R., and Tishby, N. (2006). "Efficient representation as a design principle for neural coding and computation," in *2006 IEEE International Symposium on Information Theory* (Los Alamitos: IEEE), 659–663.

Bloss, E. B., Cembrowski, M. S., Karsh, B., Colonell, J., Fetter, R. D., and Spruston, N. (2016). Structured dendritic inhibition supports branch-selective integration in CA1 pyramidal cells. *Neuron* 89, 1016–1030. doi: 10.1016/j.neuron.2016.01.029

Blundell, C., Uria, B., Pritzel, A., Li, Y., Ruderman, A., Leibo, J. Z., et al. (2016). Model-free episodic control. arXiv:1606.04460.

Bobier, B., Stewart, T. C., and Eliasmith, C. (2014). A unifying mechanistic model of selective attention in spiking neurons. *PLoS Comput. Biol.* 10:e1003577. doi: 10.1371/journal.pcbi.1003577

Bostrom, N. (1996). Cortical integration: possible solutions to the binding and linking problems in perception, reasoning and long term memory. Available online at: http://www.nickbostrom.com/old/cortical.html

Botvinick, M. M., Niv, Y., and Barto, A. C. (2009). Hierarchically organized behavior and its neural foundations: a reinforcement learning perspective. *Cognition* 113, 262–280. doi: 10.1016/j.cognition.2008.08.011

Botvinick, M., and Weinstein, A. (2014). Model-based hierarchical reinforcement learning and human action control. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 369:20130480. doi: 10.1098/rstb.2013.0480

Bouchard, G., Trouillon, T., Perez, J., and Gaidon, A. (2015). Accelerating stochastic gradient descent via online learning to sample. arXiv:1506.09016.

Bourdoukan, R., and Denève, S. (2015). "Enforcing balance allows local supervised learning in spiking recurrent networks," in *Advances in Neural Information Processing Systems* (Montreal, QC), 982–990.

Braitenberg, V., and Schutz, A. (1991). *Anatomy of the Cortex: Studies of Brain Function.* Berlin: Springer.

Brea, J., Gaál, A. T., Urbanczik, R., and Senn, W. (2016). Prospective coding by spiking neurons. *PLoS Comput. Biol.* 12:e1005003. doi: 10.1371/journal.pcbi.1005003

Brea, J., and Gerstner, W. (2016). Does computational neuroscience need new synaptic learning paradigms? *Curr. Opin. Behav. Sci.* 11, 61–66. doi: 10.1016/j.cobeha.2016.05.012

Bremner, J. G., Slater, A. M., and Johnson, S. P. (2015). Perception of object persistence: the origins of object permanence in infancy. *Child Dev. Perspect.* 9, 7–13. doi: 10.1111/cdep.12098

Brito, C. S., and Gerstner, W. (2016). Nonlinear hebbian learning as a unifying principle in receptive field formation. arXiv:1601.00701.

Brosch, T., Neumann, H., and Roelfsema, P. R. (2015). Reinforcement learning of linking and tracing contours in recurrent neural networks. *PLoS Comput. Biol.* 11:e1004489. doi: 10.1371/journal.pcbi.1004489

Brownstone, R. M., Bui, T. V., and Stifani, N. (2015). Spinal circuits for motor learning. *Curr. Opin. Neurobiol.* 33, 166–173. doi: 10.1016/j.conb.2015.04.007

Brown, T. I., Carr, V. A., LaRocque, K. F., Favila, S. E., Gordon, A. M., Bowles, B., et al. (2016). Prospective representation of navigational goals in the human hippocampus. *Science* 352, 1323–1326. doi: 10.1126/science.aaf0784.

Buesing, L., Bill, J., Nessler, B., and Maass, W. (2011). Neural dynamics as sampling: a model for stochastic computation in recurrent networks of spiking neurons. *PLoS Comput. Biol.* 7:e1002211. doi: 10.1371/journal.pcbi.1002211

Buonomano, D. V., and Merzenich, M. M. (1995). Temporal information transformed into a spatial code by a neural network with realistic properties. *Science* 267, 1028–1030.

Bülthoff, H., Little, J., and Poggio, T. (1989). A parallel algorithm for real-time computation of optical flow. *Nature* 337, 549–553.

Buschman, T. J., and Miller, E. K. (2010). Shifting the spotlight of attention: evidence for discrete computations in cognition. *Front. Hum. Neurosci.* 4:194. doi: 10.3389/fnhum.2010.00194

Buschman, T. J., and Miller, E. K. (2014). Goal-direction and top-down control. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 369:20130471. doi: 10.1098/rstb.2013.0471

Bush, K. A. (2007). *An Echo State Model of Non-markovian Reinforcement Learning*. Doctoral Dissertation. Colorado State University.

Buzsáki, G. (2010). Neural syntax: cell assemblies, synapsembles, and readers. *Neuron* 68, 362–385. doi: 10.1016/j.neuron.2010.09.023

Buzsáki, G., and Moser, E. I. (2013). Memory, navigation and theta rhythm in the hippocampal-entorhinal system. *Nat. Neurosci.* 16, 130–138. doi: 10.1038/nn.3304

Cai, D. J., Aharoni, D., Shuman, T., Shobe, J., Biane, J., Song, W., et al. (2016). A shared neural ensemble links distinct contextual memories encoded close in time. *Nature* 534, 115–118. doi: 10.1038/nature17955

Callaway, E. (2004). Feedforward, feedback and inhibitory connections in primate visual cortex. *Neural Netw.* 17, 625–632. doi: 10.1016/j.neunet.2004.04.004

Cappe, C., Rouiller, E. M., and Barone, P. (2012). "The neural bases of multisensory processes," in *Cortical and Thalamic Pathways for Multisensory and Sensorimotor Interplay*, eds M. M. Murray and M. T. Wallace (Boca Raton, FL: CRC Press/Taylor & Francis).

Caron, S. J. C., Ruta, V., Abbott, L. F., and Axel, R. (2013). Random convergence of olfactory inputs in the Drosophila mushroom body. *Nature* 497, 113–117. doi: 10.1038/nature12063

Chen, L.-C., Schwing, A. G., Yuille, A. L., and Urtasun, R. (2014). Learning deep structured models. arXiv:1407.2538.

Chikkerur, S., Serre, T., Tan, C., and Poggio, T. (2010). What and where: a Bayesian inference theory of attention. *Vis. Res.* 50, 2233–2247. doi: 10.1016/j.visres.2010.05.013

Choo, F. X., and Eliasmith, C. (2010). "A spiking neuron model of serial-order recall," in *32nd Annual Conference of the Cognitive Science Society* (Portland).

Chubykin, A. A., Roach, E. B., Bear, M. F., and Shuler, M. G. H. (2013). A cholinergic mechanism for reward timing within primary visual cortex. *Neuron* 77, 723–735. doi: 10.1016/j.neuron.2012.12.039

Chung, J., Gulcehre, C., Cho, K., and Bengio, Y. (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. arXiv:1412.3555.

Cichon, J., and Gan, W.-B. (2015). Branch-specific dendritic $ca2^+$ spikes cause persistent synaptic plasticity. *Nature* 520, 180–185. doi: 10.1038/nature14251

Clayton, N. S., and Dickinson, A. (1998). Episodic-like memory during cache recovery by scrub jays. *Nature* 395, 272–274.

Clopath, C., and Gerstner, W. (2010). Voltage and spike timing interact in STDP–a unified model. *Front. Synaptic Neurosci.* 2:25. doi: 10.3389/fnsyn.2010.00025

Cohn, R., Morantte, I., and Ruta, V. (2015). Coordinated and compartmentalized neuromodulation shapes sensory processing in drosophila. *Cell* 163, 1742–1755. doi: 10.1016/j.cell.2015.11.019

Colino, A., and Halliwell, J. (1987). Differential modulation of three separate K-conductances in hippocampal CA1 neurons by serotonin. *Nature* 328, 73–77. doi: 10.1038/328073a0

Constantinescu, A. O., O'Reilly, J. X., and Behrens, T. E. (2016). Organizing conceptual knowledge in humans with a gridlike code. *Science* 352, 1464–1468. doi: 10.1126/science.aaf0941

Cox, D. D., and Dean, T. (2014). Neural networks and neuroscience-inspired computer vision. *Curr. Biol.* 24, R921–R929. doi: 10.1016/j.cub.2014.08.026

Crick, F. (1989). The recent excitement about neural networks. *Nature* 337, 129–132.

Crick, F. C., and Koch, C. (2005). What is the function of the claustrum? *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 360, 1271–1279. doi: 10.1098/rstb.2005.1661

Crouzet, S. M., and Thorpe, S. J. (2011). Low-level cues and ultra-fast face detection. *Front. Psychol.* 2:342. doi: 10.3389/fpsyg.2011.00342

Cui, Y., Surpur, C., Ahmad, S., and Hawkins, J. (2015). Continuous online sequence learning with an unsupervised neural network model. arXiv:1512.05463.

Dadarlat, M. C., O'Doherty, J. E., and Sabes, P. N. (2015). A learning-based approach to artificial sensory feedback leads to optimal integration. *Nat. Neurosci.* 18, 138–144. doi: 10.1038/nn.3883

Danihelka, I., Wayne, G., Uria, B., Kalchbrenner, N., and Graves, A. (2016). Associative long short-term memory. arXiv:1602.03032.

Daw, N. D., Niv, Y., and Dayan, P. (2006). "Actions, policies, values and the basal ganglia," in *Recent Breakthroughs in Basal Ganglia Research*, ed E. Bezard (Hauppauge, NY: Nova Science), 91–106.

Dayan, P. (2012). Twenty-five lessons from computational neuromodulation. *Neuron* 76, 240–256. doi: 10.1016/j.neuron.2012.09.027

Dean, T. (2005). "A computational model of the cerebral cortex," in *Proceedings of the 20th National Conference on Artificial Intelligence* (Pittsburg, PA).

Dehaene, S., Meyniel, F., Wacongne, C., Wang, L., and Pallier, C. (2015). The neural representation of sequences: from transition probabilities to algebraic patterns and linguistic trees. *Neuron* 88, 2–19. doi: 10.1016/j.neuron.2015.09.019

Dekker, T. M., and Nardini, M. (2015). Risky visuomotor choices during rapid reaching in childhood. *Dev. Sci.* 19, 427–439. doi: 10.1111/desc.12322

Delalleau, O., and Bengio, Y. (2011). "Shallow vs. deep sum-product networks," in *Advances in Neural Information Processing Systems* (Grenada), 666–674.

DePasquale, B., Churchland, M., and Abbott, L. (2016). Using firing-rate dynamics to train recurrent networks of spiking model neurons. arXiv:1601.07620.

DeWolf, T., and Eliasmith, C. (2011). The neural optimal control hierarchy for motor control. *J. Neural Eng.* 8:065009. doi: 10.1088/1741-2560/8/6/065009

DiCarlo, J. J., Zoccolan, D., and Rust, N. C. (2012). How does the brain solve visual object recognition? *Neuron* 73, 415–434. doi: 10.1016/j.neuron.2012.01.010

Douglas, R. J., and Martin, K. A. C. (2004). Neuronal circuits of the neocortex. *Annu. Rev. Neurosci.* 27, 419–451. doi: 10.1146/annurev.neuro.27.070203.144152

Doya, K. (1999). What are the computations of the cerebellum, the basal ganglia and the cerebral cortex? *Neural Netw.* 12, 961–974.

Dudman, J. T., Tsay, D., and Siegelbaum, S. A. (2007). A role for synaptic inputs at distal dendrites: instructive signals for hippocampal long-term plasticity. *Neuron* 56, 866–879. doi: 10.1016/j.neuron.2007.10.020

Dumoulin, S. O., and Wandell, B. A. (2008). Population receptive field estimates in human visual cortex. *Neuroimage* 39, 647–660. doi: 10.1016/j.neuroimage.2007.09.034

Eliasmith, C. (2013). *How to Build a Brain: A Neural Architecture for Biological Cognition*. Oxford, UK: Oxford University Press.

Eliasmith, C., and Anderson, C. H. (2004). *Neural Engineering: Computation, Representation, and Dynamics in Neurobiological Systems*. Cambridge, MA: MIT Press.

Eliasmith, C., and Martens, J. (2011). Normalization for probabilistic inference with neurons. *Biol. Cybern.* 104, 251–262. doi: 10.1007/s00422-011-0433-y

Eliasmith, C., Stewart, T. C., Choo, X., Bekolay, T., DeWolf, T., Tang, Y., et al. (2012). A large-scale model of the functioning brain. *Science* 338, 1202–1205.

Emlen, S. T. (1967). Migratory orientation in the indigo bunting, passerina cyanea: part i: evidence for use of celestial cues. *Auk* 84, 309–342.

Enel, P., Procyk, E., Quilodran, R., and Dominey, P. F. (2016). Reservoir computing properties of neural dynamics in prefrontal cortex. *PLoS Comput. Biol.* 12:e1004967. doi: 10.1371/journal.pcbi.1004967

Erhan, D., and Manzagol, P. (2009). The difficulty of training deep architectures and the effect of unsupervised pre-training. *Twelfth International Conference on Artificial Intelligence and Statistics (AISTATS), JMLR Workshop and Conference Procedings* (Clearwater Beach, FL), 153–160.

Eslami, S., Heess, N., and Weber, T. (2016). Attend, infer, repeat: fast scene understanding with generative models. arXiv:1603.08575.

Fausey, C. M., Jayaraman, S., and Smith, L. B. (2016). From faces to hands: changing visual input in the first two years. *Cognition* 152, 101–107. doi: 10.1016/j.cognition.2016.03.005

Felleman, D. J., and Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cereb. cortex* 1, 1–47.

Ferster, D., and Miller, K. D. (2000). Neural mechanisms of orientation selectivity in the visual cortex. *Annu. Rev. Neurosci.* 23, 441–471. doi: 10.1146/annurev.neuro.23.1.441

Fetz, E. E. (1969). Operant conditioning of cortical unit activity. *Science* 163, 955–958.

Fetz, E. E. (2007). Volitional control of neural activity: implications for brain–computer interfaces. *J. Physiol.* 579, 571–579. doi: 10.1113/jphysiol.2006.127142

Fiete, I. R., Fee, M. S., and Seung, H. S. (2007). Model of birdsong learning based on gradient estimation by dynamic perturbation of neural conductances. *J. Neurophysiol.* 98, 2038–2057. doi: 10.1152/jn.01311.2006

Fiete, I. R., Senn, W., Wang, C. Z. H., and Hahnloser, R. H. R. (2010). Spike-time-dependent plasticity and heterosynaptic competition organize networks to produce long scale-free sequences of neural activity. *Neuron* 65, 563–576.

Fiete, I. R., and Seung, H. S. (2006). Gradient learning in spiking neural networks by dynamic perturbation of conductances. *Phys. Rev. Lett.* 97:048104. doi: 10.1103/PhysRevLett.97.048104

Finnerty, G., and Shadlen, M., Jazayeri M., Nobre A. C., Buonomano D. V. (2015). Time in Cortical Circuits. *J. Neurosci.* 35, 13912–13916. doi: 10.1523/JNEUROSCI.2654-15.2015

Finn, C., Levine, S., and Abbeel, P. (2016). Guided cost learning: deep inverse optimal control via policy optimization. arXiv:1603.00448.

Fodor, J. D., and Crowther, C. (2002). Understanding stimulus poverty arguments. *Ling. Rev.* 18, 105–145. doi: 10.1515/tlir.19.1-2.105

Földiák, P. (2008). Learning invariance from transformation sequences. *J. Neural Comput.* 3, 194–200. doi: 10.1162/neco.1991.3.2.194

Foster, D. J., Morris, R. G. M., Dayan, P. (2000). Models of hippocampally dependent navigation, using the temporal difference learning rule. *Hippocampus* 10, 1–16. doi: 10.1002/(SICI)1098-1063(2000)10:1<1::AID-HIPO1>3.0.CO;2-1

Fournier, J., Müller, C. M., and Laurent, G. (2015). Looking for the roots of cortical sensory computation in three-layered cortices. *Curr. Opin. Neurobiol.* 31, 119–126. doi: 10.1016/j.conb.2014.09.006

Franconeri, S. L., Pylyshyn, Z. W., and Scholl, B. J. (2012). A simple proximity heuristic allows tracking of multiple objects through occlusion. *Atten. Percept. Psychophys.* 74, 691–702. doi: 10.3758/s13414-011-0265-9

Frankland, S. M., and Greene, J. D. (2015). An architecture for encoding sentence meaning in left mid-superior temporal cortex. *Proc. Natl. Acad. Sci. U.S.A.* 112, 11732–11737. doi: 10.1073/pnas.1421236112

Frank, M. J., and Badre, D. (2012). Mechanisms of hierarchical reinforcement learning in corticostriatal circuits 1: computational analysis. *Cereb. Cortex* 22, 509–526. doi: 10.1093/cercor/bhr114

Franzius, M., Sprekeler, H., and Wiskott, L. (2007). Slowness and sparseness lead to place, head-direction, and spatial-view cells. *PLoS Comput. Biol.* 3:e166. doi: 10.1371/journal.pcbi.0030166

Friston, K. (2010). The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* 11, 127–138. doi: 10.1038/nrn2787

Friston, K. J., and Stephan, K. E. (2007). Free-energy and the brain. *Synthese* 159, 417–458. doi: 10.1007/s11229-007-9237-y

Fukushima, K. (1980). Neocognitron: a self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biol. Cybern.* 36, 193–202. doi: 10.1007/BF00344251

Galtier, M. N., and Wainrib, G. (2013). A biological gradient descent for prediction through a combination of stdp and homeostatic plasticity. *Neural Comput.* 25, 2815–2832. doi: 10.1162/NECO_a_00512

Gao, T., Harari, D., Tenenbaum, J., and Ullman, S. (2014). When computer vision gazes at cognition. arXiv:1412.2672.

Gemp, I., and Mahadevan, S. (2015). "Modeling context in cognition using variational inequalities," in *Modeling Changing Perspectives—Reconceptualizing Sensorimotor Experiences: Papers from the 2014 AAAI Fall Symposium* (Arlington, TX).

George, D., and Hawkins, J. (2009). Towards a mathematical theory of cortical micro-circuits. *PLoS Comput. Biol.* 5:e1000532. doi: 10.1371/journal.pcbi.1000532

Gershman, S. J., and Beck, J. M. (2016). "Complex probabilistic inference: from cognition to neural computation," in *Computational Models of Brain and Behavior*, ed A. Moustafa (Hoboken, NJ: Wiley-Blackwell).

Gershman, S. J., Moore, C. D., Todd, M. T., Norman, K. A., and Sederberg, P. B. (2012). The successor representation and temporal context. *Neural Comput.* 24, 1553–1568. doi: 10.1162/NECO_a_00282

Gershman, S. J., Moustafa, A. A., and Ludvig, E. A. (2014). Time representation in reinforcement learning models of the basal ganglia. *Front. Comput. Neurosci.* 7:194. doi: 10.3389/fncom.2013.00194

Ghahramani, I. M. Z. (2005). *A Note On the Evidence and Bayesian Occam's Razor*. Gatsby Unit Technical Report GCNU-TR 2005–003.

Giocomo, L. M., Zilli, E. A., Fransén, E., and Hasselmo, M. E. (2007). Temporal frequency of subthreshold oscillations scales with entorhinal grid cell field spacing. *Science* 315, 1719–1722. doi: 10.1126/science.1139207

Giret, N., Kornfeld, J., Ganguli, S., and Hahnloser, R. H. R. (2014). Evidence for a causal inverse model in an avian cortico-basal ganglia circuit. *Proc. Natl. Acad. Sci. U.S.A.* 111, 6063–6068. doi: 10.1073/pnas.1317087111

Goertzel, B. (2014). "How might the brain represent complex symbolic knowledge?" in *International Joint Conference on Neural Networks (IJCNN)* (Beijing).

Goldman, M. S., Levine, J. H., Major, G., Tank, D. W., and Seung, H. (2003). Robust persistent neural activity in a model integrator with multiple hysteretic dendrites per neuron. *Cereb. Cortex* 13, 1185–1195. doi: 10.1093/cercor/bhg095

Gonzalez Andino, S. L. and Grave de Peralta Menendez, R. (2012). Coding of saliency by ensemble bursting in the amygdala of primates. *Front. Behav. Neurosci.* 6:38. doi: 10.3389/fnbeh.2012.00038

Gooch, C. M., Wiener, M., Wencil, E. B., and Coslett, H. B. (2010). Interval timing disruptions in subjects with cerebellar lesions. *Neuropsychologia* 48, 1022–1031.

Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., et al. (2014a). Generative adversarial networks. arXiv:1406.2661.

Goodfellow, I. J., Vinyals, O., and Saxe, A. M. (2014b). Qualitatively characterizing neural network optimization problems. arXiv:1412.6544.

Gopnik, A., Meltzoff, A. N., and Kuhl, P. K. (2000). *The Scientist in the Crib: What Early Learning Tells us About the Mind*. New York, NY: Harper Paperbacks.

Graves, A., Wayne, G., and Danihelka, I. (2014). Neural Turing Machines. arXiv:1410.5401.

Graybiel, A. M. (1998). The basal ganglia and chunking of action repertoires. *Neurobiol. Learn. Mem.* 70, 119–136.

Gregor, K., Danihelka, I., Graves, A., Rezende, D. J., and Wierstra, D. (2015). DRAW: a recurrent neural network for image generation. arXiv:1502.04623.

Grillner, S., Hellgren, J., Ménard, A., Saitoh, K., and Wikström, M. A. (2005). Mechanisms for selection of basic motor programs–roles for the striatum and pallidum. *Trends Neurosci.* 28, 364–370. doi: 10.1016/j.tins.2005.05.004

Grossberg, S. (2013). Adaptive resonance theory: how a brain learns to consciously attend, learn, and recognize a changing world. *Neural Netw.* 37, 1–47. doi: 10.1016/j.neunet.2012.09.017

Güçlü, U., and van Gerven, M. A. (2015). Deep neural networks reveal a gradient in the complexity of neural representations across the ventral stream. *J. Neurosci.* 35, 10005–10014. doi: 10.1523/JNEUROSCI.5023-14.2015

Guez, A., Silver, D., and Dayan, P. (2012). Efficient bayes-adaptive reinforcement learning using sample-based search. arXiv:1205.3109.

Gülçehre, Ç., and Bengio, Y. (2016). Knowledge matters: importance of prior information for optimization. *J. Mach. Learn. Res.* 17, 1–32. Available online at: http://jmlr.org/papers/v17/gulchere16a.html

Güntürkün, O., and Bugnyar, T. (2016). Cognition without cortex. *Trends Cogn. Sci.* 20, 291–303. doi: 10.1016/j.tics.2016.02.001

Gurney, K., Prescott, T. J., and Redgrave, P. (2001). A computational model of action selection in the basal ganglia. I. A new functional anatomy. *Biol. Cybern.* 84, 401–410. doi: 10.1007/PL00007984

Hadley, R. F. (2009). The problem of rapid variable creation. *Neural Comput.* 21, 510–532. doi: 10.1162/neco.2008.07-07-572

Hamlin, J. K., Wynn, K., and Bloom, P. (2007). Social evaluation by preverbal infants. *Nature* 450, 557–559. doi: 10.1038/nature06288

Hangya, B., Ranade, S., Lorenc, M., and Kepecs, A. (2015). Central cholinergic neurons are rapidly recruited by reinforcement feedback. *Cell* 162, 1155–1168. doi: 10.1016/j.cell.2015.07.057

Hanuschkin, A., Ganguli, S., and Hahnloser, R. H. R. (2013). A hebbian learning rule gives rise to mirror neurons and links them to control theoretic inverse models. *Front. Neural Circ.* 7:106. doi: 10.3389/fncir.2013.00106

Harris, C. M., and Wolpert, D. M. (1998). Signal-dependent noise determines motor planning. *Nature* 394, 780–784.

Harris, K. (2008). Stability of the fittest: organizing learning through retroaxonal signals. *Trends Neurosci.* 31, 130–136. doi: 10.1016/j.tins.2007.12.002

Hassabis, D., and Maguire, E. (2009). The construction system of the brain. *Philos. Trans. R. Soc. B.* 364, 1263–1271. doi: 10.1098/rstb.2008.0296

Hassabis, D., and Maguire, E. A. (2007). Deconstructing episodic memory with construction. *Trends Cogn. Sci.* 11, 299–306. doi: 10.1016/j.tics.2007.05.001

Hasselmo, M. E. (2006). The role of acetylcholine in learning and memory. *Curr. Opin. Neurobiol.* 16, 710–715. doi: 10.1016/j.conb.2006.09.002

Hasselmo, M. E. (2015). If i had a million neurons: Potential tests of cortico-hippocampal theories. *Progr. Brain Res.* 219, 1–19. doi: 10.1016/bs.pbr.2015.03.009

Hasselmo, M. E., and Stern, C. E. (2015). Current questions on space and time encoding. *Hippocampus* 25, 744–752. doi: 10.1002/hipo.22454

Hasselmo, M. E., and Wyble, B. P. (1997). Free recall and recognition in a network model of the hippocampus: simulating effects of scopolamine on human memory function. *Behav. Brain Res.* 89, 1–34.

Hattori, D., Demir, E., Kim, H. W., Viragh, E., Zipursky, S. L., and Dickson, B. J. (2007). Dscam diversity is essential for neuronal wiring and self-recognition. *Nature* 449, 223–227. doi: 10.1038/nature06099

Hawkins, J., and Ahmad, S. (2016). Why neurons have thousands of synapses, a theory of sequence memory in neocortex. *Front. Neural Circ.* 10:23. doi: 10.3389/fncir.2016.00023

Hawkins, J., and Blakeslee, S. (2007). *On Intelligence.* New York, NY: Henry Holt and Company.

Hayashi-Takagi, A., Yagishita, S., Nakamura, M., Shirai, F., Wu, Y. I., Loshbaugh, A. L., et al. (2015). Labelling and optical erasure of synaptic memory traces in the motor cortex. *Nature* 525, 333–338. doi: 10.1038/nature15257

Haykin, S. S. (1994). *Neural Networks: A Comprehensive Foundation.* New York, NY: Macmillan.

Hayworth, K. J. (2012). Dynamically partitionable autoassociative networks as a solution to the neural binding problem. *Front. Comput. Neurosci.* 6:73. doi: 10.3389/fncom.2012.00073

Hayworth, K. J., Lescroart, M. D., and Biederman, I. (2011). Neural encoding of relative position. *J. Exp. Psychol. Hum. Percept. Perform.* 37, 1032–1050. doi: 10.1037/a0022338

Hennequin, G., Vogels, T. P., and Gerstner, W. (2014). Optimal control of transient dynamics in balanced networks supports generation of complex movements. *Neuron* 82, 1394–1406. doi: 10.1016/j.neuron.2014.04.045

Herd, S. A., Krueger, K. A., Kriete, T. E., Huang T. R., Hazy T. E., and O'Reilly R. C. (2013). Strategic cognitive sequencing: a computational cognitive neuroscience approach. *Comput. Intell. Neurosci.* 2013:149329. doi: 10.1155/2013/149329

Higgins, I., Matthey, L., Glorot, X., Pal, A., Uria, B., Blundell, C., et al. (2016). Early visual concept learning with unsupervised deep learning. arXiv:1606.05579 .

Hinton, G. (1989). Connectionist learning procedures. *Artif. Intell.* 40, 185–234. doi: 10.1016/0004-3702(89)90049-0

Hinton, G. (2007). "How to do backpropagation in a brain," in *Invited Talk at the NIPS'2007 Deep Learning Workshop* (Vancouver, BC).

Hinton, G. (2016). "Can the brain do back-propagation?," in *Invited talk at Stanford University Colloquium on Computer Systems* (Stanford, CA).

Hinton, G. E., Dayan, P., Frey, B. J., and Neal, R. M. (1995). The "wake-sleep" algorithm for unsupervised neural networks. *Science* 268, 1158–1161.

Hinton, G. E., Osindero, S., and Teh, Y.-W. (2006). A fast learning algorithm for deep belief nets. *Neural Comput.* 18, 1527–1554. doi: 10.1162/neco.2006.18.7.1527

Hinton, G., and McClelland, J. (1988). Learning representations by recirculation. *Neural information processing*.

Hinton, G. E., Krizhevsky, A., and Wang, S. (2011). "Transforming auto-encoders," in *Artificial Neural Networks and Machine Leaning*, eds T. Honkela, W. Duch, M. Girolami, and S. Kaski (Helsinki), 44–51.

Hires, S. A., Gutnisky, D. A., Yu, J., O'Connor, D. H., and Svoboda, K. (2015). Low-noise encoding of active touch by layer 4 in the somatosensory cortex. *eLife* 4. doi: 10.7554/eLife.06619

Histed, M. H., Ni, A. M., and Maunsell, J. H. (2013). Insights into cortical mechanisms of behavior from microstimulation experiments. *Progr. Neurobiol.* 103, 115–130. doi: 10.1016/j.pneurobio.2012.01.006

Hochreiter, S., and Schmidhuber, J. (1997). Long short-term memory. *Neural Comput.* 9, 1735–1780.

Hoerzer, G. M., Legenstein, R., and Maass, W. (2014). Emergence of complex computational structures from chaotic neural networks through reward-modulated Hebbian learning. *Cereb. Cortex* 24, 677–690. doi: 10.1093/cercor/bhs348

Ho, J., and Ermon, S. (2016). Generative adversarial imitation learning. arXiv:1606.03476.

Hong, W., and Luo, L. (2014). Genetic control of wiring specificity in the fly olfactory system. *Genetics* 196, 17–29. doi: 10.1534/genetics.113.154336

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci. U.S.A.* 79, 2554–2558.

Hopfield, J. J. (1984). Neurons with graded response have collective computational properties like those of two-state neurons. *Proc. Natl. Acad. Sci. U.S.A.* 81, 3088–3092.

Hopfield, J. J. (2009). Neurodynamics of mental exploration. *Proc. Natl. Acad. Sci. U.S.A.* 107, 1648–1653. doi: 10.1073/pnas.0913991107

Hosoya, T., Baccus, S. A., and Meister, M. (2005). Dynamic predictive coding by the retina. *Nature* 436, 71–77. doi: 10.1038/nature03689

Huang, Y., and Rao, R. (2011). Predictive coding. *Wiley Interdiscip. Rev. Cogn. Sci.* 2, 580–593. doi: 10.1002/wcs.142

Huys, Q. J., Lally, N., Faulkner, P., Eshel, N., Seifritz, E., Gershman, S. J.,et al. (2015). Interplay of approximate planning strategies. *Proc. Natl. Acad. Sci. U.S.A.* 112, 3098–3103. doi: 10.1073/pnas.1414219112.

Isik, L., Leibo, J. Z., and Poggio, T. (2012). Learning and disrupting invariance in visual recognition with a temporal association rule. *Front. Comput. Neurosci.* 6:37. doi: 10.3389/fncom.2012.00037

Izhikevich, E. M. (2006). Polychronization: computation with spikes. *Neural Comput.* 18, 245–282. doi: 10.1162/089976606775093882

Izhikevich, E. M. (2007). Solving the distal reward problem through linkage of STDP and dopamine signaling. *Cereb. Cortex* 17, 2443–2452. doi: 10.1093/cercor/bhl152

Jacobson, G. A., and Friedrich, R. W. (2013). Neural circuits: random design of a higher-order olfactory projection. *Curr. Biol.* 23, R448–R451. doi: 10.1016/j.cub.2013.04.016

Jaderberg, M., Czarnecki, W. M., Osindero, S., Vinyals, O., Graves, A., and Kavukcuoglu, K. (2016). Decoupled neural interfaces using synthetic gradients. arXiv:1608.05343.

Jaderberg, M., Simonyan, K., Zisserman, A., and Kavukcuoglu, K. (2015). "Spatial transformer networks," in *Advances in Neural Information Processing Systems 28 (NIPS 2015).* arXiv:1506.02025.

Jaeger, H., and Haas, H. (2004). Harnessing nonlinearity: predicting chaotic systems and saving energy in wireless communication. *Science* 304, 78–80. doi: 10.1126/science.1091277

Jara-Ettinger, J., Gweon, H., Schulz, L. E., and Tenenbaum, J. B. (2016). The naïve utility calculus: computational principles underlying commonsense psychology. *Trends Cogn. Sci.* 20, 589–604. doi: 10.1016/j.tics.2016.05.011

Jaramillo, S., and Pearlmutter, B. A. (2004). A normative model of attention: receptive field modulation. *Neurocomputing* 58, 613–618. doi: 10.1016/j.neucom.2004.01.103

Jhuang, H., Serre, T., Wolf, L., and Poggio, T. (2007). "A biologically inspired system for action recognition," in *IEEE 11th International Conference on Computer Vision, 2007* (Rio de Janeiro: ICCV), 2007, 1–8.

Jiang, X., Shen, S., Cadwell, C. R., Berens, P., Sinz, F., Ecker, A. S., et al. (2015). Principles of connectivity among morphologically defined cell types in adult neocortex. *Science* 350, aac9462. doi: 10.1126/science.aac9462

Ji, D., and Wilson, M. A. (2007). Coordinated memory replay in the visual cortex and hippocampus during sleep. *Nat. Neurosci.* 10, 100–107. doi: 10.1038/nn1825

Johansson, F., Jirenhed, D.-A., Rasmussen, A., Zucca, R., and Hesslow, G. (2014). Memory trace and timing mechanism localized to cerebellar Purkinje cells. *Proc. Natl. Acad. Sci. U.S.A.* 111, 14930–14934. doi: 10.1073/pnas.1415371111

Jonas, E., and Kording, K. (2016). Could a neuroscientist understand a microprocessor? *bioRxiv*. doi: 10.1101/055624

Joulin, A., and Mikolov, T. (2015). Inferring algorithmic patterns with stack-augmented recurrent nets. arXiv:1503.01007.

Kalisman, N., Silberberg, G., and Markram, H. (2005). The neocortical microcircuit as a tabula rasa. *Proc. Natl. Acad. Sci. U.S.A.* 102, 880–885. doi: 10.1073/pnas.0407088102

Kanwisher, N., McDermott, J., and Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J. Neurosci.* 17, 4302–4311.

Kappel, D., Habenschuss, S., Legenstein, R., and Maass, W. (2015). Network plasticity as bayesian inference. *PLoS Comput. Biol.* 11:e1004485. doi: 10.1371/journal.pcbi.1004485

Kappel, D., Nessler, B., and Maass, W. (2014). STDP installs in Winner-Take-All circuits an online approximation to hidden Markov model learning. *PLoS Comput. Biol.* 10:e1003511. doi: 10.1371/journal.pcbi.1003511

Kempter, R., Gerstner, W., and van Hemmen, J. L. (2001). Intrinsic stabilization of output rates by spike-based Hebbian learning. *Neural Comput.* 13, 2709–2741. doi: 10.1162/089976601317098501

Khaligh-Razavi, S.-M., and Kriegeskorte, N. (2014). Deep supervised, but not unsupervised, models may explain it cortical representation. *PLoS Comput. Biol.* 10:e1003915. doi: 10.1371/journal.pcbi.1003915

Kingma, D. P., and Welling, M. (2013). Auto-Encoding Variational Bayes. arXiv:1312.6114.

Knill, D., and Pouget, A. (2004). The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends Neurosci.* 27, 712–719. doi: 10.1016/j.tins.2004.10.007

Koechlin, E., and Jubault, T. (2006). Broca's area and the hierarchical organization of human behavior. *Neuron* 50, 963–974. doi: 10.1016/j.neuron.2006.05.017

Komer, B., and Eliasmith, C. (2016). A unified theoretical approach for biological cognition and learning. *Curr. Opin. Behav. Sci.* 11, 14–20. doi: 10.1016/j.cobeha.2016.03.006

Körding, K. (2007). Decision theory: what "should" the nervous system do? *Science* 318, 606–610. doi: 10.1126/science.1142998

Körding, K., and König, P. (2000). A learning rule for dynamic recruitment and decorrelation. *Neural Netw.* 13, 1–9. doi: 10.1016/S0893-6080(99)00088-X

Körding, K. P., and König, P. (2001). Supervised and unsupervised learning with two sites of synaptic integration. *J. Comput. Neurosci.* 11, 207–215. doi: 10.1023/A:1013776130161

Körding, K. P., Kayser, C., Einhäuser, W., and König, P. (2004). How are complex cell properties adapted to the statistics of natural stimuli? *J. Neurophysiol.* 91, 206–212. doi: 10.1152/jn.00149.2003

Kouh, M., and Poggio, T. (2008). A canonical neural circuit for cortical nonlinear operations. *Neural Comput.* 20, 1427–1451. doi: 10.1162/neco.2008.02-07-466

Kraus, B. J., Robinson, R. J. II, White, J. A., Eichenbaum, H., and Hasselmo, M. E. (2013). Hippocampal time cells: time versus path integration. *Neuron* 78, 1090–1101. doi: 10.1016/j.neuron.2013.04.015

Kriegeskorte, N., Mur, M., and Bandettini, P. A. (2008). Representational similarity analysis-connecting the branches of systems neuroscience. *Front. Syst. Neurosci.* 2:4. doi: 10.3389/neuro.06.004.2008

Kriete, T., Noelle, D. C., Cohen, J. D., and O'Reilly, R. C. (2013). Indirection and symbol-like processing in the prefrontal cortex and basal ganglia. *Proc. Natl. Acad. Sci. U.S.A.* 110, 16390–16935. doi: 10.1073/pnas.1303547110

Krishnamurthy, R., Lakshminarayanan, A. S., Kumar, P., and Ravindran, B. (2016). Hierarchical reinforcement learning using spatio-temporal abstractions and deep neural networks. arXiv:1605.05359.

Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems* (Lake Tahoe, CL), 1097–1105.

Kulkarni, T. D., Narasimhan, K. R., Saeedi, A., and Tenenbaum, J. B. (2016). Hierarchical deep reinforcement learning: integrating temporal abstraction and intrinsic motivation. arXiv:1604.06057.

Kulkarni, T. D., Whitney, W., Kohli, P., and Tenenbaum, J. B. (2015). Deep Convolutional Inverse Graphics Network. arXiv:1503.03167.

Kumaran, D., Hassabis, D., and McClelland, J. L. (2016). What learning systems do intelligent agents need? complementary learning systems theory updated. *Trends Cogn. Sci.* 20, 512–534. doi: 10.1016/j.tics.2016.05.004

Kumaran, D., Summerfield, J. J., Hassabis, D., and Maguire, E. A. (2009). Tracking the emergence of conceptual knowledge during human decision making. *Neuron* 63, 889–901. doi: 10.1016/j.neuron.2009.07.030

Kurach, K., Andrychowicz, M., and Sutskever, I. (2015). Neural Random-Access Machines. arXiv:1511.06392.

Lake, B. M., Salakhutdinov, R., and Tenenbaum, J. B. (2015). Human-level concept learning through probabilistic program induction. *Science* 350, 1332–1338. doi: 10.1126/science.aab3050

Lake, B. M., Ullman, T. D., Tenenbaum, J. B., and Gershman, S. J. (2016). Building machines that learn and think like people. arXiv:1604.00289.

Larkum, M. (2013). A cellular mechanism for cortical associations: an organizing principle for the cerebral cortex. *Trends Neurosci.* 36, 141–151. doi: 10.1016/j.tins.2012.11.006

LeCun, Y., and Bengio, Y. (1995). "Convolutional networks for images, speech, and time series," in *The Handbook of Brain Theory and Neural Networks*, ed M. A. Arbib (Cambridge, MA: MIT Press), 3361.

LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature* 521, 436–444. doi: 10.1038/nature14539

Lee, A. M., Tai, L.-H., Zador, A., and Wilbrecht, L. (2015). Between the primate and 'reptilian' brain: rodent models demonstrate the role of corticostriatal circuits in decision making. *Neuroscience* 296, 66–74. doi: 10.1016/j.neuroscience.2014.12.042

Lee, T., and Yuille, A. (2011). Efficient coding of visual scenes by grouping and segmentation: theoretical predictions and biological evidence. *Department of Statistics, UCLA.*

Lee, T. S., and Mumford, D. (2003). Hierarchical Bayesian inference in the visual cortex. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* 20, 1434–1448. doi: 10.1364/JOSAA.20.001434

Legenstein, R., and Maass, W. (2011). Branch-specific plasticity enables self-organization of nonlinear computation in single neurons. *J. Neurosci.* 31, 10787–10802. doi: 10.1523/JNEUROSCI.5684-10.2011

Leibo, J. Z., Cornebise, J., Gómez, S., and Hassabis, D. (2015a). Approximate hubel-wiesel modules and the data structures of neural computation. arXiv:1512.08457v1.

Leibo, J. Z., Liao, Q., Anselmi, F., and Poggio, T. (2015b). The invariance hypothesis implies domain-specific regions in visual cortex. *PLoS Comput. Biol.* 11:e1004390. doi: 10.1371/journal.pcbi.1004390

Le, Q., Ranzato, M., Monga, R., Devin, M., Chen, K., Corrado, G., et al. (2012). "Building high-level features using large scale unsupervised learning," in *International Conference in Machine Learning* (Edinburg).

Lettvin, J., Maturana, H., McCulloch, W., and Pitts, W. (1959). What the frog's eye tells the frog's brain. *Proc. IRE* 47, 1940–1951.

Letzkus, J. J., Kampa, B. M., and Stuart, G. J. (2006). Learning rules for spike timing-dependent plasticity depend on dendritic synapse location. *J. Neurosci.* 26, 10420–10429. doi: 10.1523/JNEUROSCI.2650-06.2006

Levine, S., Finn, C., Darrell, T., and Abbeel, P. (2015). End-to-end training of deep visuomotor policies. arXiv:1504.00702.

Lewicki, M. S., and Sejnowski, T. J. (2000). Learning overcomplete representations. *Neural Comput.* 12, 337–365. doi: 10.1162/089976600300015826

Lewis, S. N., and Harris, K. D. (2014). *The Neural Marketplace: I. General Formalism and Linear Theory.* Technical Report. bioRxiv:013185.

Liao, Q., and Poggio, T. (2016). Bridging the gaps between residual learning, recurrent neural networks and visual cortex. arXiv:1604.03640.

Liao, Q., Leibo, J. Z., and Poggio, T. (2015). How important is weight symmetry in backpropagation? arXiv:1510.05067.

Lillicrap, T. P., Cownden, D., Tweed, D. B., and Akerman, C. J. (2014). Random feedback weights support learning in deep neural networks. arXiv:1411.0247.

Li, N., and Dicarlo, J. J. (2012). Neuronal learning of invariant object representation in the ventral visual stream is not dependent on reward. *J. Neurosci.* 32, 6611–6620. doi: 10.1523/JNEUROSCI.3786-11.2012

Liu, J. K., and Buonomano, D. V. (2009). Embedding multiple trajectories in simulated recurrent neural networks in a self-organizing manner. *J. Neurosci.* 29, 13172–13181. doi: 10.1523/JNEUROSCI.2358-09.2009

Livni, R., Shalev-Shwartz, S., and Shamir, O. (2013). An algorithm for training polynomial networks. arXiv:1304.7045.

Lotter, W., Kreiman, G., and Cox, D. (2015). Unsupervised learning of visual structure using predictive generative networks. arXiv:1511.06380.

Lotter, W., Kreiman, G., and Cox, D. (2016). Deep predictive coding networks for video prediction and unsupervised learning. arXiv:1605.08104.

Luck, S. J., and Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature* 390, 279–281.

Luo, Y., Boix, X., Roig, G., Poggio, T., and Zhao, Q. (2015). Foveation-based mechanisms alleviate adversarial examples. arXiv:1511.06292.

Lyons, A. B., and Cheries, E. W. (2016). Inferring social disposition by sound and surface appearance in infancy. *J. Cogn. Dev.* doi: 10.1080/15248372.2016.1200048

Ma, W. J., Beck, J. M., Latham, P. E., and Pouget, A. (2006). Bayesian inference with probabilistic population codes. *Nat. Neurosci.* 9, 1432–1438. doi: 10.1038/nn1790

Maass, W. (2016) Searching for principles of brain computation. *Curr. Opin. Behav. Sci.* 11, 81–92. doi: 10.1016/j.cobeha.2016.06.003

Maass, W., Joshi, P., and Sontag, E. D. (2007). Computational aspects of feedback in neural circuits. *PLoS Comput. Biol.* 3:e165. doi: 10.1371/journal.pcbi.0020165

Maass, W., Natschläger, T., and Markram, H. (2002). Real-time computing without stable states: a new framework for neural computation based on perturbations. *Neural Comput.* 14, 2531–2560. doi: 10.1162/089976602760407955

MacDonald, C. J., Lepage, K. Q., Eden, U. T., and Eichenbaum, H. (2011). Hippocampal time cells bridge the gap in memory for discontiguous events. *Neuron* 71, 737–749. doi: 10.1016/j.neuron.2011.07.012

Maclaurin, D., Duvenaud, D., and Adams, R. (2015). Gradient-based hyperparameter optimization through reversible learning. arXiv:1502.03492.

Makin, J. G., Dichter, B. K., and Sabes, P. N. (2016). Recurrent exponential-family harmoniums without backprop-through-time. arXiv:1605.05799.

Makin, J. G., Fellows, M. R., and Sabes, P. N. (2013). Learning multisensory integration and coordinate transformation via density estimation. *PLoS Comput. Biol.* 9:e1003035. doi: 10.1371/journal.pcbi.1003035

Mandelblat-Cerf, Y., Las, L., Denisenko, N., and Fee, M. S. (2014). A role for descending auditory cortical projections in songbird vocal learning. *eLife* 3:e02152. doi: 10.7554/eLife.02152

Mansinghka, V., and Jonas, E. (2014). Building fast bayesian computing machines out of intentionally stochastic, digital parts. arXiv:1402.4914.

Marblestone, A. H., and Boyden, E. S. (2014). Designing tools for assumption-proof brain mapping. *Neuron* 83, 1239–1241. doi: 10.1016/j.neuron.2014.09.004

Marcus, G. (2001). *The Algebraic Mind: Integrating Connectionism and Cognitive Science*. Cambridge, MA: MIT Press.

Marcus, G. (2004). *The Birth of the Mind: How a Tiny Number of Genes Creates the Complexities of Human Thought*. New York, NY: Basic Books.

Marcus, G., Marblestone, A., and Dean, T. (2014a). Frequently asked question for: the atoms of neural computation. arXiv:1410.8826.

Marcus, G., Marblestone, A., and Dean, T. (2014b). The atoms of neural computation. *Science* 346, 551–552. doi: 10.1126/science.1261661

Marder, E., and Goaillard, J.-M. (2006). Variability, compensation and homeostasis in neuron and network function. *Nat. Rev. Neurosci.* 7, 563–574. doi: 10.1038/nrn1949

Markowitz, D. A., Curtis, C. E., and Pesaran, B. (2015). Multiple component networks support working memory in prefrontal cortex. *Proc. Natl. Acad. Sci. U.S.A.* 112, 11084–11089. doi: 10.1073/pnas.1504172112

Markram, H., Lübke, J., Frotscher, M., and Sakmann, B. (1997). Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. *Science* 275, 213–215.

Markram, H., Muller, E., Ramaswamy, S., Reimann, M. W., Abdellah, M., Sanchez, C. A., et al. (2015). Reconstruction and simulation of neocortical microcircuitry. *Cell* 163, 456–492. doi: 10.1016/j.cell.2015.09.029

Marr, D. (1969). A theory of cerebellar cortex. *J. Physiol.* 202, 437–470.

Martens, J., and Sutskever, I. (2011). "Learning recurrent neural networks with hessian-free optimization," in *Proceedings of the 28th International Conference on Machine Learning* (Bellevue, WA).

McCandliss, B. D., Cohen, L., and Dehaene, S. (2003). The visual word form area: expertise for reading in the fusiform gyrus. *Trends Cogn. Sci.* 7, 293–299. doi: 10.1016/S1364-6613(03)00134-7

McCulloch, W. S., and Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *Bull. Math. Biophys.* 5, 115–133. doi: 10.1007/BF02478259

McKinstry, J. L., Edelman, G. M., and Krichmar, J. L. (2006). A cerebellar model for predictive motor control tested in a brain-based device. *Proc. Natl. Acad. Sci. U.S.A.* 103, 3387–3392. doi: 10.1073/pnas.0511281103

McKone, E., Crookes, K., and Kanwisher, N. (2009). "The cognitive and neural development of face recognition in humans," in *The Cognitive Neurosciences, 4th Edn*, eds S. G. Michael (Cambridge, MA: MIT Press), 4, 467–482.

McLeod, P., and Dienes, Z. (1996). Do fielders know where to go to catch the ball or only how to get there? *J. Exp. Psychol. Hum. Percept. Perform.* 22, 531.

Mel, B. (1992). The clusteron: toward a simple abstraction for a complex neuron. *Adv. Neural Inf. Process. Syst.* 4, 35–42.

Meltzoff, A. N. (1999). Born to learn: what infants learn from watching us. *Role Early Exp. Infant Dev.* 145–164.

Meltzoff, A. N., Waismeyer, A., and Gopnik, A. (2012). Learning about causes from people: observational causal learning in 24-month-old infants. *Dev. Psychol.* 48, 1215–1258. doi: 10.1037/a0027440

Meltzoff, A. N., Williamson, R. A., and Marshall, P. J. (2013). "11 developmental perspectives on action science: lessons from infant imitation and cognitive neuroscience," in *Action Science: Foundations of an Emerging Discipline*, eds W. Prinz, M. Beisert, and A. Herwig (Cambridge, MA: MIT Press), 281–306.

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychol. Rev.* 63, 81–97.

Miller, K. D., Keller, J. B., and Stryker, M. P. (1989). Ocular dominance column development: analysis and simulation. *Science*, 245, 605–615.

Miller, K. D., and MacKay, D. J. C. (1994). The role of constraints in hebbian learning. *Neural Comput.* 6, 100–126.

Minsky, M. (1977). "Plain talk about neurodevelopmental epistemology," in *IJCAI'77 Proceedings of the 5th International Joint Conference on Artificial Intelligence* (San Francisco, CA: Morgan Kaufmann Publishers Inc.), 1083–1092. Available online at: http://hdl.handle.net/1721.1/5763

Minsky, M. (1988). *Society of Mind*. New York, NY: Simon and Schuster.

Minsky, M. (2006). The Emotion Machine. New York, NY: Pantheon.

Minsky, M. L. (1991). Logical versus analogical or symbolic versus connectionist or neat versus scruffy. *AI magazine* 12, 34–51.

Minsky, M. L., and Papert, S. (1972). *Perceptrons: An Introduction to Computational Geometry*. Cambridge, MA: MIT Press.

Mishra, R. K., Kim, S., Guzman, S. J., and Jonas, P. (2016). Symmetric spike timing-dependent plasticity at ca3-ca3 synapses optimizes storage and recall in autoassociative networks. *Nat. Commun.* 7:1552. doi: 10.1038/ncomms11552

Mitchell, T. M. (1980). "The need for biases in learning generalizations," in *Readings in Machine Learning*, eds J. W. Shavlik and T. G. Dietterich (Morgan Kauffman), 184–191. Available online at: http://www.cs.nott.ac.uk/~bsl/G52HPA/articles/Mitchell:80a.pdf

Miyagawa, S., Berwick, R. C., and Okanoya, K. (2013). The emergence of hierarchical structure in human language. *Front. Psychol.* 4:71. doi: 10.3389/fpsyg.2013.00071

Mnih, V., Heess, N., Graves, A., and Kavukcuoglu, K. (2014). "Recurrent models of visual attention," in *Advances in Neural Information Processing Systems* (Montreal, QC), 2204–2212.

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., et al. (2015). Human-level control through deep reinforcement learning. *Nature* 518, 529–533. doi: 10.1038/nature14236

Mobahi, H., Collobert, R., and Weston, J. (2009). "Deep learning from temporal coherence in video," in *Proceedings of the 26th Annual International Conference on Machine Learning - ICML '09* (New York, NY: ACM Press), 1–8.

Moberget, T., Gullesen, E. H., Andersson, S., Ivry, R. B., and Endestad, T. (2014). Generalized role for the cerebellum in encoding internal models: evidence from semantic processing. *J. Neurosci.* 34, 2871–2878. doi: 10.1523/JNEUROSCI.2264-13.2014

Mohamed, S., and Rezende, D. J. (2015). Variational information maximisation for intrinsically motivated reinforcement learning. arXiv:1509.08731

Mordatch, I., Todorov, E., and Popović, Z. (2012). Discovery of complex behaviors through contact-invariant optimization. *ACM Trans. Graph.* 31:43. doi: 10.1145/2185520.2185539

Morgan, J. L., Berger, D. R., Wetzel, A. W., and Lichtman, J. W. (2016). The fuzzy logic of network connectivity in mouse visual thalamus. *Cell* 165, 192–206. doi: 10.1016/j.cell.2016.02.033

Mushiake, H., Saito, N., Sakamoto, K., Itoyama, Y., and Tanji, J. (2006). Activity in the lateral prefrontal cortex reflects multiple steps of future events in action plans. *Neuron* 50, 631–641. doi: 10.1016/j.neuron.2006.03.045

Nardini, M., Bedford, R., and Mareschal, D. (2010). Fusion of visual cues is not mandatory in children. *Proc. Natl. Acad. Sci. U.S.A.* 107, 17041–17046. doi: 10.1073/pnas.1001699107

Neelakantan, A., Le, Q. V., and Sutskever, I. (2015). Neural programmer: inducing latent programs with gradient descent. arXiv:1511.04834.

Nessler, B., Pfeiffer, M., Buesing, L., and Maass, W. (2013). Bayesian computation emerges in generic cortical microcircuits through spike-timing-dependent

plasticity. *PLoS Comput. Biol.* 9:e1003037. doi: 10.1371/journal.pcbi.
1003037

Ng, A., and Russell, S. (2000). "Algorithms for inverse reinforcement learning," in
*ICML '00 Proceedings of the Seventeenth International Conference on Machine
Learning* (San Francisco, CA).

Noroozi, M., and Favaro, P. (2016). Unsupervised learning of visual
representations by solving jigsaw puzzles. arXiv:1603.09246.

Ólafsdóttir, H. F., Barry, C., Saleem, A. B., Hassabis, D., and Spiers, H. J.
(2015). Hippocampal place cells construct reward related sequences through
unexplored space. *eLife* 4:e06063. doi: 10.7554/eLife.06063

Ollivier, Y., and Charpiat, G. (2015). Training recurrent networks online without
backtracking. arXiv:1507.07680.

Olshausen, B. A., Anderson, C. H., and Van Essen, D. C. (1993). A neurobiological
model of visual attention and invariant pattern recognition based on dynamic
routing of information. *J. Neurosci.* 13, 4700–4719.

Olshausen, B. A., and Field, D. J. (1996). Emergence of simple-cell receptive field
properties by learning a sparse code for natural images. *Nature* 381, 607–609.

Olshausen, B. A., and Field, D. J. (1997). Sparse coding with an overcomplete basis
set: a strategy employed by V1? *Vis. Res.* 37, 3311–3325.

Olshausen, B. A., and Field, D. J. (2004). What is the other 85% of v1 doing.
Prob. Syst. Neurosci. 4, 182–211. doi: 10.1093/acprof:oso/9780195148220.0
03.0010

Oquab, M., Bottou, L., Laptev, I., and Sivic, J. (2014). "Learning and transferring
mid-level image representations using convolutional neural networks," in
Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition
(Columbus, OH), 1717–1724.

O'Reilly, R. C. (1996). Biologically plausible error-driven learning using local
activation differences: the generalized recirculation algorithm. *Neural Comput.*
8, 895–938.

O'Reilly, R. C. (2006). Biologically based computational models of high-level
cognition. *Science* 314, 91–94. doi: 10.1126/science.1127242

O'Reilly, R. C., and Frank, M. J. (2006). Making working memory work: a
computational model of learning in the prefrontal cortex and basal ganglia.
Neural Comput. 18, 283–328. doi: 10.1162/089976606775093909

O'Reilly, R. C., Hazy, T. E., Mollick, J., Mackie, P., and Herd, S. (2014a).
Goal-driven cognition in the brain: a computational framework. arXiv:14
04.7591.

O'Reilly, R. C., Munakata, Y., Frank, M. J., and Hazy, T. E. (2012).
Computational Cognitive Neuroscience, 1st Edn. Wiki Book. Available online at:
http://ccnbook.colorado.edu

O'Reilly, R. C., Wyatte, D., and Rohrlich, J. (2014b). Learning through time in the
thalamocortical loops. arXiv:1407.3432, 37.

Orhan, A. E., and Ma, W. J. (2016). The inevitability of probability: probabilistic
inference in generic neural networks trained with non-probabilistic feedback
arXiv:1601.03060.

Palmer, L. M., Shai, A. S., Reeve, J. E., Anderson, H. L., Paulsen, O., and Larkum,
M. E. (2014). NMDA spikes enhance action potential generation during sensory
input. *Nat. Neurosci.* 17, 383–390. doi: 10.1038/nn.3646

Parisien, C., Anderson, C. H., and Eliasmith, C. (2008). Solving the problem of
negative synaptic weights in cortical models. *Neural Comput.* 20, 1473–1494.
doi: 10.1162/neco.2008.07-06-295

Pasupathy, A., and Miller, E. K. (2005). Different time courses of learning-
related activity in the prefrontal cortex and striatum. *Nature* 433, 873–876. doi:
10.1038/nature03287

Patel, A., Nguyen, T., and Baraniuk, R. (2015). A probabilistic theory of deep
learning. arXiv:1504.00641.

Pehlevan, C., and Chklovskii, D. B. (2015). "Optimization theory of hebbian/anti-
hebbian networks for pca and whitening," in *53rd Annual Allerton Conference
on Communication, Control, and Computing* (Monticello, IL), 1458–1465.

Perea, G., Navarrete, M., and Araque, A. (2009). Tripartite synapses: astrocytes
process and control synaptic information. *Trends Neurosci.* 32, 421–431. doi:
10.1016/j.tins.2009.05.001

Petrov, A. A., Jilk, D. J., and O'Reilly, R. C. (2010). The Leabra architecture:
specialization without modularity. *Behav. Brain Sci.* 33, 286–287. doi:
10.1017/S0140525X10001160

Pezzulo, G., Verschure, P. F. M. J., Balkenius, C., and Pennartz, C. M. A. (2014).
The principles of goal-directed decision-making: from neural mechanisms

to computation and robotics. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 369,
20130470. doi: 10.1098/rstb.2013.0470

Pfister, J.-P., and Gerstner, W. (2006). Triplets of spikes in a model
of spike timing-dependent plasticity. *J. Neurosci.* 26, 9673–9682. doi:
10.1523/JNEUROSCI.1425-06.2006

Phillips, A. T., Wellman, H. M., and Spelke, E. S. (2002). Infants' ability to connect
gaze and emotional expression to intentional action. *Cognition* 85, 53–78. doi:
10.1016/S0010-0277(02)00073-2

Piekniewski, F., Laurent, P., Petre, C., Richert, M., Fisher, D., and Hylton, T. (2016).
Unsupervised learning from continuous video in a scalable predictive recurrent
network. arXiv:1607.06854.

Pineda, F. J. (1987). Generalization of back-propagation to recurrent neural
networks. *Phys. Rev. Lett.* 59:2229. doi: 10.1103/PhysRevLett.59.2229

Pinker, S. (1999). How the mind works. *Ann. N.Y. Acad. Sci.* 882, 119–127.

Plate, T. A. (1995). Holographic reduced representations. *IEEE Trans. Neural
Netw.* 6, 623–641.

Poggio, T. (2015). *What if...*, MIT Center for Brains Minds and Machines Memo.

Poggio, T., and Bizzi, E. (2004). Generalization in vision and motor control. *Nature*
431, 768–774. doi:10.1038/nature03014

Ponulak, F., and Hopfield, J. J. (2013). Rapid, parallel path planning by propagating
wavefronts of spiking neural activity. *Front. Comput. Neurosci.* 7:98. doi:
10.3389/fncom.2013.00098

Radford, A., Metz, L., and Chintala, S. (2015). Unsupervised representation
learning with deep convolutional generative adversarial networks.
arXiv:1511.06434.

Rajan, K., Harvey, C. D., and Tank, D. W. (2016). Recurrent network
models of sequence generation and memory. *Neuron* 90, 128–142. doi:
10.1016/j.neuron.2016.02.009

Ramachandran, V. S. (2000). *Mirror Neurons and Imitation Learning as the
Driving Force Behind "the Great Leap Forward" in Human Evolution*. Available
online at: https://www.edge.org/conversation/mirror-neurons-and-imitation-
learning-as-the-driving-force-behind-the-great-leap-forward-in-human-
evolution

Rao, R. P. (2004). Bayesian computation in recurrent neural circuits. *Neural
Comput.* 16, 1–38. doi: 10.1162/08997660460733976

Rashevsky, N. (1939). Mathematical biophysics: physico-mathematical
foundations of biology. *Bull. Amer. Math. Soc.* 45, 223–224. doi:
10.1090/S0002-9904-1939-06963-2

Rasmus, A., and Berglund, M. (2015). Semi-supervised learning with ladder
networks. arXiv:1507.02672.

Reynolds, J. H., and Desimone, R. (1999). The role of neural mechanisms of
attention in solving the binding problem. *Neuron* 24, 19–29.

Rezende, D. J., Mohamed, S., Danihelka, I., Gregor, K., and Wierstra, D. (2016).
One-shot generalization in deep generative models. arXiv:1603.05106.

Rigotti, M., Barak, O., Warden, M. R., Wang, X.-J., Daw, N. D., Miller, E. K., et al.
(2013). The importance of mixed selectivity in complex cognitive tasks. *Nature*
497, 585–590. doi: 10.1038/nature12160

Robinson, D. (1992). Implications of neural networks for how we think about brain
function. *Behav. Brain Sci.* 15, 644–655.

Rodriguez, A., and Granger, R. (2016). The grammar of mammalian brain capacity.
Theor. Comput. Sci. 633, 100–111. doi: 10.1016/j.tcs.2016.03.021

Rodriguez, A., Whitson, J., and Granger, R. (2004). Derivation and analysis of
basic computational operations of thalamocortical circuits. *J. Cogn. Neurosci.*
16, 856–877. doi: 10.1162/089892904970690

Roelfsema, P. R., van Ooyen, A., and Watanabe, T. (2010). Perceptual learning
rules based on reinforcers and attention. *Trends Cogn. Sci.* 14, 64–71. doi:
10.1016/j.tics.2009.11.005

Roelfsema, P. R., and van Ooyen, A. (2005). Attention-gated reinforcement
learning of internal representations for classification. *Neural Comput.* 17,
2176–2214. doi: 10.1162/0899766054615699

Rolls, E. T. (2013). The mechanisms for pattern completion and pattern separation
in the hippocampus. *Front. Syst. Neurosci.* 7:74. doi: 10.3389/fnsys.2013.00074

Rombouts, J. O., Bohte, S. M., and Roelfsema, P. R. (2015). How attention can
create synaptic tags for the learning of working memories in sequential tasks.
PLoS Comput. Biol. 11:e1004060. doi: 10.1371/journal.pcbi.1004060

Romero, A., Ballas, N., Kahou, S. E., Chassang, A., Gatta, C., and Bengio, Y. (2014).
Fitnets: hints for thin deep nets. *arXiv* arXiv:1412.6550.

Roudi, Y., and Taylor, G. (2015). Learning with hidden variables. *Curr. Opin. Neurobiol.* 35, 110–118. doi: 10.1016/j.conb.2015.07.006

Rozell, C. J., Johnson, D. H., Baraniuk, R. G., and Olshausen, B. A. (2008). Sparse coding via thresholding and local competition in neural circuits. *Neural Comput.* 20, 2526–2563. doi: 10.1162/neco.2008.03-07-486

Rubin, A., Geva, N., Sheintuch, L., and Ziv, Y. (2015). Hippocampal ensemble dynamics timestamp events in long-term memory. *eLife* 4:e12247. doi: 10.7554/eLife.12247

Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature* 323, 533–536. doi: 10.1038/323533a0

Rumelhart, D. E., and Zipser, D. (1986). "Feature discovery by competitive learning," in *Parallel Distributed Processing*, Vol. 1, eds D. E Rumel hart and J. L. McClelland (Cambridge, MA: MIT Press), 151–163.

Sadtler, P. T., Quick, K. M., Golub, M. D., Chase, S. M., Ryu, S. I., Tyler-Kabara, E. C., et al. (2014). Neural constraints on learning. *Nature* 512, 423–426. doi: 10.1038/nature13665

Sahani, M., and Dayan, P. (2003). Doubly distributional population codes: simultaneous representation of uncertainty and multiplicity. *Neural Comput.* 15, 2255–2279. doi: 10.1162/089976603322362356

Sandler, M., Shulman, Y., and Schiller, J. (2016). A novel form of local plasticity in tuft dendrites of neocortical somatosensory layer 5 pyramidal neurons. *Neuron* 90, 1028–1042. doi: 10.1016/j.neuron.2016.04.032

Santoro, A., Bartunov, S., Botvinick, M., Wierstra, D., and Lillicrap, T. (2016). One-shot learning with memory-augmented neural networks. 13. arXiv:1605.06065. Available online at: https://arxiv.org/abs/1605.06065

Saxe, A. M., McClelland, J. L., and Ganguli, S. (2013). Exact solutions to the nonlinear dynamics of learning in deep linear neural networks. arXiv:1312.6120.

Scellier, B., and Bengio, Y. (2016). Towards a biologically plausible backprop. arXiv:1602.05179.

Schiess, M., Urbanczik, R., and Senn, W. (2016). Somato-dendritic synaptic plasticity and error-backpropagation in active dendrites. *PLoS Comput. Biol.* 12:e1004638. doi: 10.1371/journal.pcbi.1004638

Schmidhuber, J. (2010). Formal theory of creativity, fun, and intrinsic motivation (19902010). *Auton. Ment. Dev. IEEE.* 2, 230–247. doi: 10.1109/TAMD.2010.2056368

Schmidhuber, J. (2015). Deep learning in neural networks: an overview. *Neural Netw.* 61, 85–117. doi: 10.1016/j.neunet.2014.09.003

Scholl, B. J. (2004). Can infants' object concepts be trained? *Trends Cogn. Sci.* 8, 49–51. doi: 10.1016/j.tics.2003.12.006

Schwabe, L., Obermayer, K., Angelucci, A., and Bressloff, P. C. (2006). The role of feedback in shaping the extra-classical receptive field of cortical neurons: a recurrent network model. *J. Neurosci.* 26, 9117–9129. doi: 10.1523/JNEUROSCI.1253-06.2006

Sejnowski, T., and Poizner, H. (2014). Prospective optimization. *Proc. IEEE.* 102, 799–811. doi: 10.1109/jproc.2014.2314297

Sermanet, P., and Kavukcuoglu, K. (2013). "Pedestrian detection with unsupervised multi-stage feature learning," in *Proceedings of Computer Vision and Pattern Recognition (CVPR)* (Portland, OR).

Serre, T., Oliva, A., and Poggio, T. (2007). A feedforward architecture accounts for rapid categorization. *Proc. Natl. Acad. Sci. U.S.A.* 104, 6424–6429. doi: 10.1073/pnas.0700622104

Servan-Schreiber, E., and Anderson, J. (1990). Chunking as a mechanism of implicit learning. *J. Exp. Psychol.* 16, 592–608.

Seung, H. S. (1998). Continuous attractors and oculomotor control. *Neural Netw.* 11, 1253–1258. doi: 10.1016/S0893-6080(98)00064-1

Seung, H. S. (2003). Learning in spiking neural networks by reinforcement of stochastic synaptic transmission. *Neuron* 40, 1063–1073. doi: 10.1016/S0896-6273(03)00761-X

Shai, A. S., Anastassiou, C. A., Larkum, M. E., and Koch, C. (2015). Physiology of layer 5 pyramidal neurons in mouse primary visual cortex: coincidence detection through bursting. *PLoS Comput. Biol.* 11:e1004090. doi: 10.1371/journal.pcbi.1004090

Shepherd, G. M. (2014). The microcircuit concept applied to cortical evolution: from three-layer to six-layer cortex. *Front. Neuroanat.* 5:30. doi: 10.3389/fnana.2011.00030

Sherman, S. M. (2005). Thalamic relays and cortical functioning. *Prog. Brain Res.* 149, 107–126. doi: 10.1016/S0079-6123(05)49009-3

Sherman, S. M. (2007). The thalamus is more than just a relay. *Curr. Opin. Neurobiol.* 17, 417–422. doi: 10.1016/j.conb.2007.07.003

Shimizu, T., and Karten, H. J. (2013). Multiple origins of neocortex: contributions of the dorsal. *Neocortex* 200:75. doi: 10.1007/978-1-4899-0652-6_8

Siegel, M., Warden, M. R., and Miller, E. K. (2009). Phase-dependent neuronal coding of objects in short-term memory. *Proc. Natl. Acad. Sci. U.S.A.* 106, 21341–21346. doi: 10.1073/pnas.0908193106

Singh, R., and Eliasmith, C. (2006). Higher-dimensional neurons explain the tuning and dynamics of working memory cells. *J. Neurosci.* 26, 3667–3678. doi: 10.1523/JNEUROSCI.4864-05.2006

Sjöström, J., and Gerstner, W. (2010). Spike-timing dependent plasticity. *Scholarpedia* 5:1362. doi: 10.4249/scholarpedia.1362

Sjöström, P. J., and Häusser, M. (2006). A cooperative switch determines the sign of synaptic plasticity in distal dendrites of neocortical pyramidal neurons. *Neuron* 51, 227–238. doi: 10.1016/j.neuron.2006.06.017

Skerry, A. E., and Spelke, E. S. (2014). Preverbal infants identify emotional reactions that are incongruent with goal outcomes. *Cognition* 130, 204–216. doi: 10.1016/j.cognition.2013.11.002

Softky, W., and Koch, C. (1993). The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs. *J. Neurosci.* 13, 334–350.

Solari, S. V. H., and Stoner, R. (2011). Cognitive consilience: primate non-primary neuroanatomical circuits underlying cognition. *Front. Neuroanat.* 5:65. doi: 10.3389/fnana.2011.00065

Sountsov, P., and Miller, P. (2015). Spiking neuron network Helmholtz machine. *Front. Comput. Neurosci.* 9:46. doi: 10.3389/fncom.2015.00046

Squire, L. R. (2004). Memory systems of the brain: a brief history and current perspective. *Neurobiol. Learn. Memory* 82, 171–177. doi: 10.1016/j.nlm.2004.06.005

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* 15, 1929–1958. Available online at: http://www.jmlr.org/papers/v15/srivastava14a.html

Stachenfeld, K. (2014). "Design principles of the hippocampal cognitive map," in *Advances in Neural Information Processing Systems* (Montreal, QC).

Stanisor, L., van der Togt, C., Pennartz, C. M. A., and Roelfsema, P. R. (2013). A unified selection signal for attention and reward in primary visual cortex. *Proc. Natl. Acad. Sci. U.S.A.* 110, 9136–9141. doi: 10.1073/pnas.1300117110

Stewart, T., and Eliasmith, C. (2009). "Compositionality and biologically plausible models," in *Oxford Handbook of Compositionality*, eds W. Hinzen, E. Machery, and M. Werning (Oxford University Press). doi: 10.1093/oxfordhb/9780199541072.013.0029

Stocco, A., Lebiere, C., and Anderson, J. R. (2010). Conditional routing of information to the cortex: a model of the basal ganglia's role in cognitive coordination. *Psychol. Rev.* 117, 541–574. doi: 10.1037/a0019077

Stork, D. G. (1989). "Is backpropagation biologically plausible?," in *International Joint Conference on Neural Networks*, Vol. 2 (Washington, DC: IEEE), 241–246.

Strausfeld, N. J., and Hirth, F. (2013). Deep homology of arthropod central complex and vertebrate basal ganglia. *Science (New York, N.Y.)* 340, 157–161. doi: 10.1126/science.1231828

Sukhbaatar, S., Bruna, J., Paluri, M., Bourdev, L., and Fergus, R. (2014). Training convolutional networks with noisy labels. *arXiv preprint arXiv:1406.2080*.

Sun, Y., Gomez, F., and Schmidhuber, J. (2011). "Planning to be surprised: optimal bayesian exploration in dynamic environments," in *Artificial General Intelligence* (Mountain View, CA: Springer), 41–51. doi: 10.1007/978-3-642-22887-2_5

Sussillo, D. (2014). Neural circuits as computational dynamical systems. *Curr. Opin. Neurobiol.* 25, 156–163. doi: 10.1016/j.conb.2014.01.008

Sussillo, D., and Abbott, L. (2009). Generating coherent patterns of activity from chaotic neural networks. *Neuron* 63, 544–557. doi: 10.1016/j.neuron.2009.07.018

Sussillo, D., Churchland, M. M., Kaufman, M. T., and Shenoy, K. V. (2015). A neural network that finds a naturalistic solution for the production of muscle activity. *Nat. Neurosci.* 18, 1025–1033. doi: 10.1038/nn.4042

Sutskever, I., and Martens, J. (2013). "On the importance of initialization and momentum in deep learning," in *Proceedings of the 30 th International Conference on Machine Learning* (Atlanta: JMLR:W&CP).

Sutskever, I., Martens, J., and Hinton, G. E. (2011). "Generating text with recurrent neural networks," in *Proceedings of the 28th International Conference on Machine Learning (ICML-11)* (Bellevue), 1017–1024.

Sutton, R. S., and Barto, A. G. (1998). *Reinforcement Learning: An Introduction.* Cambridge, MA: MIT Press.

Tacchetti, A., Isik, L., and Poggio, T. (2016). Spatio-temporal convolutional neural networks explain human neural representations of action recognition. *arXiv preprint arXiv:1606.04698.*

Takata, N., Mishima, T., Hisatsune, C., Nagai, T., Ebisui, E., Mikoshiba, K., and Hirase, H. (2011). Astrocyte calcium signaling transforms cholinergic modulation to cortical plasticity *in vivo. J. Neurosci.* 31, 18155–18165. doi: 10.1523/JNEUROSCI.5289-11.2011

Tamar, A., Levine, S., and Abbeel, P. (2016). Value iteration networks. *arXiv preprint arXiv:1602.02867.*

Tang, Y., Salakhutdinov, R., and Hinton, G. (2012). Deep mixtures of factor analysers. arXiv:1206.4635

Tang, Y., Salakhutdinov, R., and Hinton, G. (2013). "Tensor analyzers," in *Proceedings of the 30th International Conference on Machine Learning (ICML-13)* (Atlanta, GA).

Tapson, J., and van Schaik, A. (2013). Learning the pseudoinverse solution to network weights. *Neural Netw.* 45, 94–100. doi: 10.1016/j.neunet.2013.02.008

Tavares, R. M., Mendelsohn, A., Grossman, Y., Williams, C. H., Shapiro, M., Trope, Y., et al. (2015). A map for social navigation in the human brain. *Neuron* 87, 231–243. doi: 10.1016/j.neuron.2015.06.011

Taylor, S. V., and Faisal, A. A. (2011). Does the cost function of human motor control depend on the internal metabolic state? *BMC Neurosci.* 12(Suppl. 1):P99. doi: 10.1186/1471-2202-12-S1-P99

Terrence Stewart, C. E., Choo, X., and Eliasmith, C. (2010). "Symbolic reasoning in spiking neurons: a model of the cortex/basal ganglia/thalamus loop," in *32nd Annual Meeting of the Cognitive Science Society* (Portland, OR).

Tervo, D. G. R., Tenenbaum, J. B., and Gershman, S. J. (2016). Toward the neural implementation of structure learning. *Curr. Opin. Neurobiol.* 37, 99–105. doi: 10.1016/j.conb.2016.01.014

Tesauro, G. (1995). Temporal difference learning and TD-Gammon. *Commun. ACM.* 38, 58–68.

Thalmeier, D., Uhlmann, M., Kappen, H. J., and Memmesheimer, R.-M. (2015). Learning universal computations with spikes. arXiv:1505.07866.

Tinbergen, N. (1965). "Behavior and natural selection," in *Ideas in Modern Biology: proceedings of the 16th International Zoological Congress,* ed J. A. Moore (Washington, DC), 521–542.

Todorov, E. (2002). Cosine tuning minimizes motor errors. *Neural Comput.* 14, 1233–1260. doi: 10.1016/j.conb.2016.01.014

Todorov, E. (2009). Efficient computation of optimal actions. *Proc. Natl. Acad. Sci. U.S.A.* 106, 11478–11483. doi: 10.1073/pnas.0710743106

Todorov, E., and Jordan, M. I. (2002). Optimal feedback control as a theory of motor coordination. *Nat. Neurosci.* 5, 1226–1235. doi: 10.1038/nn963

Tripp, B., and Eliasmith, C. (2016). Function approximation in inhibitory networks. *Neural Netw.* 77, 95–106. doi: 10.1016/j.neunet.2016.01.010

Turner, R. S., and Desmurget, M. (2010). Basal ganglia contributions to motor control: a vigorous tutor. *Curr. Opin. Neurobiol.* 20, 704–716. doi: 10.1016/j.conb.2010.08.022

Turrigiano, G. (2012). Homeostatic synaptic plasticity: local and global mechanisms for stabilizing neuronal function. *Cold Spring Harb. Perspect. Biol.* 4:a005736. doi: 10.1101/cshperspect.a005736

Ullman, S., Harari, D., and Dorfman, N. (2012). From simple innate biases to complex visual concepts. *Proc. Natl. Acad. Sci. U.S.A.* 109, 18215–18220. doi: 10.1073/pnas.1207690109

Urbanczik, R., and Senn, W. (2014). Learning by the dendritic prediction of somatic spiking. *Neuron* 81, 521–528. doi: 10.1016/j.neuron.2013.11.030

Valpola, H. (2015). From neural PCA to deep unsupervised learning. arXiv:1411.7783.

van den Oord, A., Kalchbrenner, N., and Kavukcuoglu, K. (2016). Pixel recurrent neural networks. arXiv:1601.06759.

Van Heijningen, C. A., De Visser, J., Zuidema, W., and Ten Cate, C. (2009). Simple rules can explain discrimination of putative recursive syntactic structures

by a songbird species. *Proc. Natl. Acad. Sci. U.S.A.* 106, 20538–20543. doi: 10.1073/pnas.0908113106

Van Kerkoerle, T., Self, M. W., Dagnino, B., Gariel-Mathis, M.-A., Poort, J., Van Der Togt, C., et al. (2014). Alpha and gamma oscillations characterize feedback and feedforward processing in monkey visual cortex. *Proc. Natl. Acad. Sci. U.S.A.* 111, 14332–14341. doi: 10.1073/pnas.1402773111

Veit, A., Wilber, M., and Belongie, S. (2016). Residual networks are exponential ensembles of relatively shallow networks. arXiv:1605.06431.

Verney, C., Baulac, M., Berger, B., Alvarez, C., Vigny, A., and Helle, K. (1985). Morphological evidence for a dopaminergic terminal field in the hippocampal formation of young and adult rat. *Neuroscience* 14, 1039–1052. doi: 10.1016/0306-4522(85)90275-1

Verwey, W. B. (1996). Buffer loading and chunking in sequential keypressing. *J. Exp. Psychol.* 22:544.

Wang, J. X., Cohen, N. J., and Voss, J. L. (2015). Covert rapid action-memory simulation (CRAMS): a hypothesis of hippocampal-prefrontal interactions for adaptive behavior. *Neurobiol. Learn. Memory* 117, 22–33. doi: 10.1016/j.nlm.2014.04.003

Wang, J., and Yuille, A. (2014). Semantic part segmentation using compositional model combining shape and appearance. arXiv:1412.6124.

Wang, X.-J. (2012). "The prefrontal cortex as a quintessential "cognitive-type" neural circuit," in *Principles of Frontal Lobe Function,* Edited by D. T. Stuss and R. T. Knight (Oxford University Press), 226–248.

Warden, M. R., and Miller, E. K. (2007). The representation of multiple objects in prefrontal neuronal delay activity. *Cereb. Cortex* (New York, N.Y.: 1991) 17(Suppl. 1):i41–i50. doi: 10.1093/cercor/bhm070

Warden, M. R., and Miller, E. K. (2010). Task-dependent changes in short-term memory in the prefrontal cortex. *J. Neurosci.* 30, 15801–15810. doi: 10.1523/JNEUROSCI.1569-10.2010

Watter, M., Springenberg, J., Boedecker, J., and Riedmiller, M. (2015). "Embed to control: a locally linear latent dynamics model for control from raw images," in *Advances in Neural Information Processing Systems* (Montreal, QC), 2728–2736.

Wayne, G., and Abbott, L. F. (2014). Hierarchical control using networks trained with higher-level forward models. *Neural Comput.* 26, 2163–2193. doi: 10.1162/NECO_a_00639

Werbos, P. (1974). *Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences.* Doctoral Dissertation, Harvard University, Harvard.

Werbos, P. (1982). Applications of advances in nonlinear sensitivity analysis. *Syst. Model. Optim.* 38, 762–770. doi: 10.1007/bfb0006203

Werbos, P. (1990). Backpropagation through time: what it does and how to do it. *Proc. IEEE.* 78, 1550–1560. doi: 10.1109/5.58337

Werbos, P. J., and Si, J. (eds.). (2004). *Handbook of Learning and Approximate Dynamic Programming,* Vol. 2. Playa del Carmen: John Wiley & Sons.

Werfel, J., Xie, X., and Seung, H. S. (2005). Learning curves for stochastic gradient descent in linear feedforward networks. *Neural Comput.* 17, 2699–2718. doi: 10.1162/089976605774320539

Weston, J., Chopra, S., and Bordes, A. (2014). Memory networks. arXiv:1410.3916.

Whitney, W. F., Chang, M., Kulkarni, T., and Tenenbaum, J. B. (2016). Understanding visual concepts with continuation learning. arXiv:1602.06822.

Williams, R. J. (1992). Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Mach. Learn.* 8, 229–256. doi: 10.1007/BF00992696

Williams, R. J., and Baird, L. C. (1993). *Tight Performance Bounds on Greedy Policies based on Imperfect Value Functions.* Technical Report, Citeseer.

Williams, S. R., and Stuart, G. J. (2000). Backpropagation of physiological spike trains in neocortical pyramidal neurons: implications for temporal coding in dendrites. *J. Neurosci.* 20, 8238–8246.

Wilson, R. I., and Nicoll, R. A. (2001). Endogenous cannabinoids mediate retrograde signalling at hippocampal synapses. *Nature* 410, 588–592. doi: 10.1038/35069076

Winston, P. (2011). "The strong story hypothesis and the directed perception hypothesis," in *AAAI Fall Symposium Series* (Association for the Advancement of Artificial Intelligence).

Wiskott, L., and Sejnowski, T. J. (2002). Slow feature analysis: unsupervised learning of invariances. *Neural Comput.* 14, 715–770. doi: 10.1162/089976602317318938

Wolpert, D. M., and Flanagan, J. R. (2016). Computations underlying sensorimotor learning. *Curr. Opin. Neurobiol.* 37, 7–11. doi: 10.1016/j.conb.2015.12.003

Womelsdorf, T., Valiante, T. A., Sahin, N. T., Miller, K. J., and Tiesinga, P. (2014). Dynamic circuit motifs underlying rhythmic gain control, gating and integration. *Nat. Neurosci.* 17, 1031–1039. doi: 10.1038/nn.3764

Wyss, R., König, P., and Verschure, P. F. M. J. (2006). A model of the ventral visual system based on temporal stability and local memory. *PLoS Biol.* 4:e120. doi: 10.1371/journal.pbio.0040120

Xie, X., and Seung, H. (2000). "Spike-based learning rules and stabilization of persistent neural activity," in *Advances in Neural Information Processing System* (Denver).

Xie, X., and Seung, H. S. (2003). Equivalence of backpropagation and contrastive Hebbian learning in a layered network. *Neural Comput.* 15, 441–454. doi: 10.1162/089976603762552988

Xiong, C., Merity, S., and Socher, R. (2016). Dynamic memory networks for visual and textual question answering. *arXiv preprint arXiv:160 3.01417.*

Xu, M., Zhang, S.-Y., Dan, Y., and Poo, M.-M. (2014). Representation of interval timing by temporally scalable firing patterns in rat prefrontal cortex. *Proc. Natl. Acad. Sci. U.S.A.* 111, 480–485. doi: 10.1073/pnas.1321314111

Yamins, D. L., and DiCarlo, J. J. (2016a). Using goal-driven deep learning models to understand sensory cortex. *Nat. Neurosci.* 19, 356–365. doi: 10.1038/nn.4244

Yamins, D. L., and DiCarlo, J. J. (2016b). Eight open questions in the computational modeling of higher sensory cortex. *Curr. Opin. Neurobiol.* 37, 114–120. doi: 10.1016/j.conb.2016.02.001

Yosinski, J., Clune, J., Bengio, Y., and Lipson, H. (2014). "How transferable are features in deep neural networks?," in *Advances in Neural Information Processing Systems* (Montreal, QC), 3320–3328.

Yttri, E. A., and Dudman, J. T. (2016). Opponent and bidirectional control of movement velocity in the basal ganglia. *Nature* 533, 402–406. doi: 10.1038/nature17639

Yu, C., and Smith, L. B. (2013). Joint attention without gaze following: human infants and their parents coordinate visual attention to objects through eye-hand coordination. *PLoS ONE* 8:e79659. doi: 10.1371/journal.pone.0079659

Yuste, R., MacLean, J. N., Smith, J., and Lansner, A. (2005). The cortex as a central pattern generator. *Nat. Rev. Neurosci.* 6, 477–483. doi: 10.1038/nrn1686

Zaremba, W., and Sutskever, I. (2015). Reinforcement learning neural turing machines. *arXiv preprint arXiv:1505.00521.*

Zeisel, A., Manchado, A. B. M., Codeluppi, S., Lönnerberg, P., La Manno, G., Juréus, A., et al. (2015). Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science* 347, 1138–1142. doi: 10.1126/science.aaa1934

Zemel, R. S., and Dayan, P. (1997). "Combining probabilistic population codes," in *International Joint Conference on Artificial Intelligence* (Nagoya), 1114–1119.

Zilli, E. A., and Hasselmo, M. E. (2010). Coupled noisy spiking neurons as velocity-controlled oscillators in a model of grid cell spatial firing. *J. Neurosci.* 30, 13850–13860. doi: 10.1523/JNEUROSCI.0547-10.2010

Zipser, D., and Andersen, R. (1988). A back-propagation programmed network that simulates response properties of a subset of posterior parietal neurons. *Nature* 331, 679–684. doi: 10.1038/331679a0