Dendritic solutions to the credit assignment problem
Blake A Richards1,2,3 and Timothy P Lillicrap4

Guaranteeing that synaptic plasticity leads to effective learning requires a means for assigning credit to each neuron for its contribution to behavior. The ‘credit assignment problem’ refers to the fact that credit assignment is non-trivial in hierarchical networks with multiple stages of processing. One difficulty is that if credit signals are integrated with other inputs, then it is hard for synaptic plasticity rules to distinguish credit-related activity from non-credit-related activity. A potential solution is to use the spatial layout and non-linear properties of dendrites to distinguish credit signals from other inputs. In cortical pyramidal neurons, evidence hints that top-down feedback signals are integrated in the distal apical dendrites and have a distinct impact on spike-firing and synaptic plasticity. This suggests that the distal apical dendrites of pyramidal neurons help the brain to solve the credit assignment problem.

Addresses
1 Department of Biological Sciences, University of Toronto Scarborough, Toronto, ON, Canada
2 Department of Cell and Systems Biology, University of Toronto, Toronto, ON, Canada
3 Learning in Machines and Brains Program, Canadian Institute for Advanced Research, Toronto, ON, Canada
4 DeepMind, London, United Kingdom

Introduction: the credit assignment problem
The flexibility of learning in animals indicates that the brain possesses general purpose learning algorithms. A learning algorithm is a set of rules for translating the experiences an animal has into changes in their neural circuits (e.g. synaptic changes). The ultimate goal of a learning algorithm is to alter the behavioral phenotype of the animal, helping it to adapt to the environment. Understanding the brain’s learning algorithms is key to understanding the biological basis of animal intelligence.

The formal study of learning algorithms often utilizes the concept of a loss function (also known as a cost function) [1,2]. Within neuroscience, a loss function provides a metric for the failure of the current phenotype in achieving an animal’s goals (Figure 1a) [3]. For example, a loss function could measure motor slips or sensory prediction errors. Ideally, the brain would have some way of ensuring that changes in a neural circuit reduce a given loss function [3], at least within the environments that the animal is likely to encounter [4]. To do this, it is useful to assign ‘credit’ (or ‘blame’) to each neuron or synapse for its contribution to the loss function [5,6]. However, outside of very simple neural circuits, credit assignment calculations are difficult. In a hierarchical sensorimotor circuit with multiple stages of processing, such as the mammalian neocortex, the credit that a neuron in a sensory area deserves for any motor errors depends on that neuron’s downstream connections to motor circuits (Figure 1b) [7]. The difficulty of assigning credit in the context of hierarchical circuits is known as the credit assignment problem [8].

Typically, solutions to the credit assignment problem have been explored in neural network models that treat each neuron as a single voltage compartment with a single type of output (e.g. a scalar firing-rate or spike train) [7,9**,10**,11–14,15*]. This strategy is reasonable at face value: it fits with the basic properties of neural computation and helps to reduce mathematical complexity. However, there are two reasons that this strategy may have inadvertently made it more difficult to identify the brain’s solution to the credit assignment problem. First, if each neuron is calculating everything using a single voltage value, then any incoming signals about credit (e.g. feedback from another cortical area) must be integrated with other signals about sensory data, or they must arrive at a separate time. The result is that any credit related signals must be carefully timed or they risk becoming entangled with other ongoing calculations (Figure 1c,d). There is some evidence of clock-like phasic activity in various parts of the brain [16], but none of these seem to exhibit the clear segregation between feedforward and feedback activity required for credit assignment. Second, if a neuron only has one type of output, for example, a firing rate, then it is not immediately obvious how neural circuits can disambiguate credit related activity from basic information transmission (Figure 1f).

Of course, real neurons are not single compartments — they possess complex dendritic trees that integrate different signals in different locations [17–27], often in non-linear manners that have important functional implications [28–44]. Moreover, active channels in dendrites can drive spiking behavior that is different from regular
Loss functions and credit assignment. (a) Illustration of a loss function. A loss function provides a metric for the performance of an agent on some learning task. In a neural circuit, loss functions are functions of synaptic strength. The goal of learning is to find synaptic strengths that minimize the loss function. Here, an arbitrary loss function is plotted for a network with only two synapses. (b) Illustration of the credit assignment problem. A multilayer neural network with two neurons per layer is shown. Circles indicate neurons, with green circles indicating highly active neurons. Arrows indicate synaptic connections and the width of the arrows indicates synaptic strength. If an input arrives at the left-hand neuron, its activity causes strong activation in the downstream left-hand neurons, due to strong synaptic connections. However, if the loss function specifies that the target was to give an output at the right-hand, then an error is generated. To make it more likely that the right-hand output neuron would be activated, it would help to increase the feedforward activity of the right-hand middle neuron, X. In other words, this neuron deserves some 'credit' for the incorrect output. Credit assignment can be achieved if the error signal at the top-level is sent back to the middle-layer. (c) However, if the middle-layer neuron is a single compartment, this error signal, E, would be integrated with the ongoing activity, X, thereby altering the 'forward' computation being performed by this neuron. (d) A possible solution is to have carefully timed phases where feedforward and feedback signals are received at distinct times. (e) An alternative is to integrate the credit assignment signal in a separate dendritic compartment. (f) and (g) Illustration of the use of specialized spike-waveforms for credit assignment. (f) If incoming inputs and credit signals both produce the same type of spiking in a neuron (indicated by ‘X’ and ‘E’, respectively), it is difficult to differentiate credit assignment from ongoing processing. (g) In contrast, if credit signals drive dendritic non-linearities that produce unique spike-waveforms (e.g. a complex spike or high-frequency burst), then it is easy to differentiate credit assignment from other processes.

spiking [45,46]. One possibility, then, is to segregate credit signals into dendritic compartments, where (i) they can be kept separate from other ongoing calculations (Figure 1e), and (ii) they can drive unique spike-waveforms that signal credit information (Figure 1g). Thus, there has been a growing interest in understanding whether one of the solutions to the credit assignment problem lies in dendritic computation [47,48**,*49**,50] (and see also IMN Sacramento et al. arXiv: 1801.00062).

What counts as evidence for credit assignment?
The ideal experiment for understanding credit assignment in the brain would be to measure a loss function explicitly, then demonstrate that a given synaptic plasticity mechanism was responsible for ensuring reductions in that loss function during learning. Such experiments are currently outside of our technical reach, though, because it is often unclear how we can identify a loss function in the brain and track its progress over time [3]. Furthermore, there is no reason to assume that the brain explicitly represents any of the loss functions it may be reducing. Indeed, at the neural level, it is possible to reduce a loss function without there being any direct neural correlate of said loss function to find [51,52].

Given these realities, the best strategy for scientists to study credit assignment depends on the level of analysis. For example, if the desire is to examine whether credit assignment actually shapes activity in the brain based on the extent to which different neurons contribute to a task [53], then it is possible to use tetrode recordings and similar approaches [54]. In contrast, if the desire is to understand the cellular mechanisms by which credit is
assigned in a hierarchy, then studies of synaptic plasticity are key. Historically, the study of plasticity rules has focused on two-factor Hebbian updates [55–57], which emphasize correlations in pre and postsynaptic activity [58]. However, the cumulative evidence from computational modeling and machine learning suggests that a simple Hebbian learning algorithm based solely on two factors — pre and postsynaptic activity — is insufficient for credit assignment in difficult tasks where the loss function depends on downstream circuits and delayed outcomes [55–57]. A starting place for coming to grips with this issue is to consider learning rules wherein pre and postsynaptic activity determine an ‘eligibility trace’ that indicates which synapses are eligible for updates, but a third (or possibly fourth) factor that depends on feedback or neuromodulation determines whether long-term potentiation (LTP) or long-term depression (LTD) occur [7,14,47,48**,59–62,63*,64]. Accordingly, these models predict that LTP/LTD should depend not only on pre and postsynaptic activity, but also on additional ‘credit signals’ carrying information about things like action outcomes, prediction errors, rewards/punishments, and attention [55–57,65]. Thus, a practical, experimental framework for studying credit assignment is to examine Hebbian synaptic plasticity rules in a circuit and determine whether additional feedback signals carrying credit-related information can regulate the synaptic changes that occur.

There are several lines of experimental evidence supporting a role for neuromodulators in credit assignment in various neural circuits, including the hippocampus, neocortex and striatum [66–69]. Indeed, neuromodulators have been shown to have modulating effects on Hebbian-like synaptic plasticity in these circuits [66–68]. However, we also know that neuromodulator systems tend to transmit widely to a volume of tissue, and thus are not usually neuron specific, let alone dendrite specific. The most interesting role that dendrites could play in credit assignment would be to provide a site for fine-grained credit assignment calculations, since effective credit assignment in deep networks typically requires some form of neuron-by-neuron credit signal [10**,57]. Thus, while neuromodulators undoubtedly play a crucial role in credit assignment systems, we here focus on neuron-by-neuron credit assignment mechanisms, which are more likely to be linked to dendritic processing.

Perhaps the clearest example of experimental evidence for neuron-by-neuron credit assignment is provided by learning in the cerebellum. In the cerebellum, granule cells project to Purkinje cells via parallel fibers, carrying information about input from the spinal cord, the cortex, and the vestibular system. The Purkinje cells carry signals that are considered the output of the cerebellum [70]. Various forms of motor learning may rely on the plasticity of parallel fiber synapses onto Purkinje cells [71–74] (though see [75]). However, this plasticity depends not only on parallel fiber and Purkinje cell activity, but a third term delivered via climbing fibers from the inferior olivary nucleus [70]. Evidence suggests that diverse error signals from downstream motor systems are communicated by the climbing fibers [76,77], and interestingly, climbing fibers synapse onto Purkinje cells in a one-to-one correspondence. This makes one-to-one mapping of error signals possible (Figure 2a). Plasticity at the parallel fiber → Purkinje cell synapses is mediated by a distinct spike-waveform generated by climbing fiber activity and the calcium currents they control, referred to as a ‘complex spike’ (Figure 2b) [45,78,79]. This allows the climbing fibers to control whether positive (LTP) or negative (LTD) changes in the parallel fiber → Purkinje cell connections occur [80–83] (Figure 2c). Moreover, the specific timing of climbing fiber activation controls parallel fiber → Purkinje cell plasticity in a manner that matches the temporal delay between cerebellar activity and error signal receipt [84*]. Thus, this can be modeled as a three-factor learning rule, where pre and/or postsynaptic activities interact with a credit assignment factor provided by the climbing fibers [70,85].

There is a potentially important difference between Purkinje cell credit assignment and credit assignment in other neurons/circuits, though: because Purkinje cells represent the output of the cerebellum, and because there appears to be a one-to-one mapping between climbing fibers and Purkinje cells [45], the credit assignment problem in Purkinje cells is much less difficult. That is, credit assignment in Purkinje cells may be relatively straightforward, since error signals are not being integrated backwards through a complex hierarchy, but instead are directly communicated to each neuron on a one-to-one basis. Therefore, credit calculations in Purkinje cells may not require a dendritic compartment that is segregated from the parallel fiber inputs. In contrast, in pyramidal neurons buried deep in a cortical network in the forebrain, the credit assignment problem is much more daunting and dendritic segregation may be crucial for enabling detailed credit assignment.

Credit assignment in cortical pyramidal neurons

In the neocortex and hippocampus, pyramidal neurons are part of a hierarchical pathway with multiple sources of potential credit-related feedback. Thus, assigning credit in cortical pyramidal neurons may require more involved calculations than in the output layer of the cerebellum with one-to-one climbing fiber → Purkinje cell error signals. Where might these credit calculations take place? To date, direct experimental evidence for credit assignment calculations in neocortical neurons is limited. But, there are converging lines of evidence that led us to propose in a recent computational modeling study that
the distal apical dendrites of pyramidal neurons are involved in credit assignment [48**].

We begin by highlighting the properties of distal apical dendrites that make them suitable for credit assignment calculations. First, the distal apical dendrites in neocortex are a major recipient of higher-order cortico-cortical and thalamo-cortical feedback signals [21,86,18–20,86], and in the CA1 region of the hippocampus they receive long-range information back from entorhinal cortex [26,25]. This is notable because one of the key features of credit assignment in a number of computational models is the use of downstream feedback to control upstream synaptic plasticity [7,9**,10**,14,15*,48**]. Second, distal apical dendrites are electrotonically distant from the soma and the basal and oblique dendrites [34,37,87,88], which receive much of the local feedforward and recurrent inputs to pyramidal neurons [22,23]. As such, distal apical dendrites both receive feedback signals that are required for credit assignment, and they are sufficiently segregated to permit credit assignment calculations in isolation from ongoing sensory integration (Figure 3a).

If distal apical dendrites are electrotonically distant, though, how could their computations act as a third-factor to control plasticity in oblique and/or basal dendrites? Synaptic plasticity in pyramidal neurons ultimately depends on local depolarization and non-linear active potentials driven by N-methyl-d-aspartate (NMDA) receptors (‘NMDA spikes’) [49**,89–94]. Thus, what matters for the induction of synaptic plasticity is not postsynaptic action potential firing, per se, but the manner in which postsynaptic activity affects local depolarization in dendritic compartments [95,96]. Interestingly, distal apical dendrites may actually be well-placed to control depolarization in the basal/oblique dendrites. High-frequency burst-firing is a well-established mechanism for inducing dendritic depolarization, NMDA spikes and synaptic plasticity [97,35,98], and burst-firing can alter the sign of synaptic plasticity in proximal synapses in cortical pyramidal neurons [99]. It is notable, then, that the distal apical dendrites are well-placed to drive burst-firing [46]. Specifically, the apical dendrites of pyramidal neurons have a region that is rich in voltage-gated calcium channels [31,33,100,101,43,36], which can induce ‘plateau potentials’ when the apical dendrites are sufficiently depolarized, or if there is coincident activation of distal apical inputs and somatic spiking (Figure 3b) [34,37,102,43]. These plateau potentials induce high-frequency (>100 Hz) bursts of action potentials [34,37,102]. As such, a switch from regular spiking to burst-firing in pyramidal neurons carries a signal indicating that inputs were received at the distal apical dendrites [102,46,103]. Given that burst-firing can regulate local depolarization in the basal/oblique dendrites, it has been proposed that bursts driven by distal apical inputs provide a third-factor that regulates plasticity in other dendrites based on current higher-order feedback [48**,104]. This proposed role for apical dendrites in synaptic plasticity has been used by computational models to implement credit-assignment calculations in simulated pyramidal neurons [48**,104] (IMN Sacramento et al. arXiv: 1801.00062), and thereby reduce high-level loss functions with local synaptic plasticity rules.

As stated above, the properties of apical dendrites make them suitable for credit assignment calculations, but experimental evidence for this proposal is limited. Nonetheless, recent findings support the conclusion that apical dendrites control plasticity in pyramidal neurons. In layer
2/3 pyramidal neurons in somatosensory cortex, apical dendrites receiving associative thalamic input can induce synaptic plasticity of sensory inputs without spiking [105], and these same inputs gate synaptic plasticity when spiking does occur (IMN Williams & Holtmaat bioRxiv: 10.1101/281477). In the CA1 region of the hippocampus, two important studies recently demonstrated that apical-driven plateau potentials control the formation of place-cells and determine whether synaptic plasticity in basal dendrites occurs, even with seconds between the occurrence of basal synaptic input and plateau potentials [43**,106**]. Moreover, a series of studies examining visual cortex in the last few years have provided convincing demonstrations that layer 1 inputs from both higher-order cortex and associative thalamus can carry predictive and error feedback signals [107**,108,109*,110,111], which could be used to calculate prediction errors for credit assignment [9**,63**]. Altogether, the new data coming out in the field of dendritic processing, synaptic plasticity, and cortical coding is consistent with the hypothesized role for apical dendrites in credit assignment. This may help to explain the mysterious architecture of pyramidal neurons, wherein a substantial proportion of long-range inputs arrive at an electronically distant site [46]. However, much more experimental data needs to be collected to understand how apical dendrites might be involved in credit assignment calculations in pyramidal neurons. Specifically, more studies are required to determine: first, how apical inputs modify basal/oblique plasticity rules (Figure 3c), similar to how we know the manner in which climbing fiber inputs modulate parallel fiber inputs to Purkinje neurons [82], and second, whether different pyramidal neurons in different regions use apical signals for credit assignment in different ways (e.g. are there differences in apical credit assignment between hippocampal versus neocortical, or layer 2/3 versus layer 5 pyramidal neurons?). Furthermore, it is possible that the distal compartments of basal dendrites provide another site for credit assignment calculations in cortical pyramidal neurons [49**,112**].

Conclusion

A major goal for researchers in coming years should be a better link between the theory of credit assignment in neural networks [7,9**,10**,11–14], and our growing knowledge of the biophysics of dendrites and dendritic computation [28–44]. Clearly, there is much more to understand about dendritic computations in pyramidal neurons, how they may signal credit information, and how they contribute to learning, in-turn. Three issues that deserve focused attention in our opinion are: first, how do dendritic mechanisms map onto three-factor synaptic update rules [55–57]? Second, what is the role of inhibitory interneuron microcircuits in credit assignment [27,50*,113–115]? Third, how does plasticity of feedback inputs to distal apical dendrites factor into credit assignment [9**,94,99,116–118]? Each of these questions are ripe for extensive investigations.

The credit assignment problem has rarely been an explicit focus in experimental studies of synaptic plasticity. But, arguably, that was for two good reasons. First, the major theoretical advances in credit assignment were developed for artificial neural networks that made few
Conflict of interest statement
Nothing declared.

Acknowledgements
This work was supported by the Natural Sciences and Engineering Research Council of Canada (CHE: RGPIN-2014-04947) and the Canadian Institute for Advanced Research (BAR: Learning in Machine and Brains Program).

References and recommended reading
Papers of particular interest, published within the period of review, have been highlighted as

9. Lee D-H, Zhang S, Fischer A, Bengio Y: Difference target propagation. Springer; 2015:498-515. This paper introduces difference target propagation, an alternative to the backpropagation-of-error algorithm that avoids the requirement for symmetric feedback weights. The authors demonstrate that the algorithm can engage in efficient credit assignment in multilayer networks.
This paper demonstrates that feedback pathways need not be symmetric to feedforward pathways for credit assignment to work. It provided one of the first possible solutions to the weight symmetry problem faced by backpropagation.
This paper provides a learning algorithm for credit assignment with precisely timed spike trains. The authors demonstrate that the algorithm can be used to solve the credit assignment problem in multilayer networks.

www.sciencedirect.com

Dendritic solutions to the credit assignment problem Richards and Lillicrap 33


concrete experimental predictions [6,12,13]. Second, it is difficult/impossible to experimentally explore all of the potential input patterns that may drive synaptic plasticity. For example, spike-timing-dependent plasticity may not actually be the true synaptic update rule, but may instead be what emerges from a learning algorithm that uses feedback for credit assignment when studied with highly constrained spike-timing patterns [119]. Thus, theoretical insights are required to guide synaptic plasticity experiments and provide practical limits on the inputs and spike patterns that need to be tested. One aspect of the current lack of predictions from neural network models is an absence of dendrites and their active properties. The assumption that all neurons are single, linear compartments with just one form of non-linear spiking output has made some mathematical analyses easier. However, whereas in machine learning the circuitry required for learning can be built outside the network being trained and dispensed with when it is not needed, in the real brain the circuitry for learning must fit into existing pathways and is always present. Recent computational work has highlighted the potential importance of dendrites with separate compartments and non-linear properties for solving the credit assignment problem in a biologically realistic manner [47,48**,49**]. Given the success of deep learning in artificial intelligence [2], and the emergence of sophisticated optical tools for studying dendritic computation [120], now is the ideal time for modelers and experimentalists to work together, and unify our understanding of dendritic computation with our theories of learning in hierarchical neural networks.


42. Suzuki M, Larkum ME: Dendritic calcium spikes are clearly detectable at the cortical surface. Nat Commun 2017, 8 http://dx.doi.org/10.1038/s41467-017-00282-4.


48. Guerguev J, Lillippat CP, Richards BA: Towards deep learning • with segregated dendrites. eLife 2017, e22901 http://dx.doi.org/10.7554/eLife.22901. This paper provides a computational model of credit assignment with electronically segregated dendrites. It illustrates how the electronic segregation of credit signals can be used to make credit calculations easier in a biologically realistic framework.

49. Bono J, Clopath C: Modeling somatic and dendritic spike • mediated plasticity at the single neuron and network level. Nat Commun 2017, 8:706 http://dx.doi.org/10.1038/s41467-017-00740-z.

This paper uses computational modeling to explore synaptic plasticity mediated by dendritic mechanisms. It finds that dendrites permit multiple plasticity rules to co-exist in the same cell. One notable finding is that inputs to the distal basal dendrites can be used to switch the sign of synaptic plasticity in other basal compartments.

50. Wilmes KA, Sprekel H, Schreiber S: Inhibition as a binary • switch for excitatory plasticity in pyramidal neurons. PLOS Comput Biol 2016, 12:e1004768 http://dx.doi.org/10.1371/journal. pcbi.1004768. This paper uses multi-compartment modeling of pyramidal neurons to examine the potential for inhibition to gate synaptic plasticity in dendrites. The authors find that with precisely timed inhibition it is possible to gate synaptic plasticity in an all-or-none manner in the dendrites without impacting information processing in the soma.


De Koekkoek Ito Schultz Urbanczik Fremaux


Suvarthan A, Payne HL, Raymond JL: Timing rules for synaptic plasticity matched to behavioral function. Neuron 2016, 92:959-967 http://dx.doi.org/10.1016/j.neuron.2016.10.022. This paper demonstrates that the spike-time-dependent-plasticity rules in different regions of the cerebellum are adapted to the specific delay at which error signals from the climbing fibers would normally arrive during motor learning.


This paper demonstrated the existence of a feedback pathway from premotor cortex to primary visual cortex in mice. The signals communicated by this pathway could potentially be used to solve the credit assignment problem.


This paper explores thalamic inputs to layer 1 of primary visual cortex. One notable finding in this paper is the observation that feedback from associative thalamus contains information about visuomotor mismatches, which are the type of error signal used for credit assignment in some computational models.


This paper demonstrates that a spike-timing-dependent plasticity rule can emerge from a rate-based synaptic plasticity rule that solves the credit assignment problem. The algorithm uses backpropagated gradients contained in the temporal derivatives of the firing rates.