



Biologically plausible learning in neural networks with modulatory feedback



W. Shane Grant^{a,*}, James Tanner^a, Laurent Itti^{a,b}

^a Department of Computer Science, University of Southern California, Los Angeles, CA, 90089, USA

^b Department of Psychology and Neuroscience Graduate Program, University of Southern California, Los Angeles, CA, 90089, USA

ARTICLE INFO

Article history:

Received 17 June 2016

Received in revised form 6 December 2016

Accepted 17 January 2017

Available online 28 January 2017

Keywords:

Border ownership
Computational modeling
Feedback
Modulatory
Plasticity
Self-organization

ABSTRACT

Although Hebbian learning has long been a key component in understanding neural plasticity, it has not yet been successful in modeling modulatory feedback connections, which make up a significant portion of connections in the brain. We develop a new learning rule designed around the complications of learning modulatory feedback and composed of three simple concepts grounded in physiologically plausible evidence. Using border ownership as a prototypical example, we show that a Hebbian learning rule fails to properly learn modulatory connections, while our proposed rule correctly learns a stimulus-driven model. To the authors' knowledge, this is the first time a border ownership network has been learned. Additionally, we show that the rule can be used as a drop-in replacement for a Hebbian learning rule to learn a biologically consistent model of orientation selectivity, a network which lacks any modulatory connections. Our results predict that the mechanisms we use are integral for learning modulatory connections in the brain and furthermore that modulatory connections have a strong dependence on inhibition.

© 2017 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

The brain has a remarkable ability to learn to process complicated input through self-organization, and since the studies of Hubel and Wiesel (1963) it has been known that the development of early visual processes is dependent on experience. In the decades since, models of visual development have focused on feedforward pathways, with little attention given to the learning of modulatory connections. Modulatory connections, which adjust existing neuron activations instead of directly driving them, dominate feedback pathways, which themselves constitute a majority of the connections in the brain (Markov et al., 2014). Hebbian-based models have come a long way in explaining potential mechanisms of learning (Clopath, Büsing, Vasilaki & Gerstner, 2010; Hebb, 1949; Widloski & Fiete, 2014), especially in feedforward models of V1 (Stevens, Law, Antolik & Bednar, 2013), but an increasing amount of literature suggests that more comprehensively explaining plasticity requires novel approaches (Lim et al., 2015; Zenke, Agnes & Gerstner, 2015). We will argue that the principles of Hebbian learning, known colloquially as fire together, wire together, cannot be used alone to learn correctly or maintain stability in the context of modulatory connections.

The primary contributions of this work are twofold: the development of a new learning rule that handles modulatory connections, and showing that a stimulus driven feedback model of border ownership can be learned in a biologically plausible way as a result of the new learning rule. The new learning rule, which we call conflict learning, is composed of three conceptually simple, physiologically plausible mechanisms: adjusting plasticity based on the activation of strongly learned connections, using inhibition as an error signal to explicitly unlearn connections, and exploiting several timescales. With border ownership as our prototypical example, we show that a Hebbian learning rule fails to properly learn modulatory connections, while the components of our proposed rule enable it to learn the required connections. Border ownership, which involves the assignment of edges to owning objects, is perhaps one of the earliest and simplest visual processes dependent upon modulatory feedback (Kogo & van Ee, 2014), appearing in V1, V2, and V4 (Zhou, Friedman & Von Der Heydt, 2000). Although many models of its function exist (e.g., lateral models: Sakai and Nishimura (2006); Zhaoping (2005), feedforward: Supér, Romeo and Keil (2010), and feedback: Craft, Schütze, Niebur and Von Der Heydt (2007)) those incorporating feedback are especially promising, integrating well with models of attention (Mihalas, Dong, von der Heydt & Niebur, 2011; Qiu, Sugihara & von der Heydt, 2007) and concepts of grouping (Martin & von der Heydt, 2015). However, until now, all of these models have used fixed, hand-crafted weights, with no demonstration of how the connection patterns for border ownership might be learned.

* Corresponding author.

E-mail addresses: wgrant@usc.edu (W.S. Grant), jettanner@usc.edu (J. Tanner), itti@pollux.usc.edu (L. Itti).

With our new learning rule, we demonstrate that inhibitory modulation of plasticity, in conjunction with competition, is likely a crucial mechanism for learning modulatory connections. Additionally, we show that the rule can be used as a drop-in replacement for a Hebbian learning rule even in networks lacking any modulatory connections, such as an orientation selective model of primary visual cortex. Conflict learning is compared against a recent Hebbian learning based rule (GCAL; Stevens et al. (2013)), which is a good baseline rule for comparison because its weight updates are governed purely by Hebbian logic and it operates at a level of abstraction that captures important physiological behaviors while still being usable in large scale neural network models (e.g., orientation selectivity) and being adaptable for use in new network architectures (e.g., border ownership). We demonstrate that conflict learning, like a Hebbian rule such as GCAL, can be used to learn a biologically consistent model of orientation selectivity. Our results further suggest that networks learned with conflict learning have improved noise and stability responses.

Conflict learning works in a fundamentally different way to previous learning rules by leveraging inhibition as an error signal to dynamically adjust plasticity. Though many existing techniques built upon Hebbian learning, such as those derived from STDP (spike timing-dependent plasticity, Song, Miller and Abbott, 2000) or BCM learning (Bienenstock, Cooper & Munro, 1982), have some method to explicitly control synaptic weakening (e.g., based on signal timing for STDP or comparisons to long term activation averages for BCM), inhibition only indirectly affects learning by lowering activation. Our successful application of the rule to learning models of orientation selectivity as well as border ownership serves as a prediction that modulatory connections in the brain require inhibition and competition to play a bigger role in the dynamics of neural plasticity and activation.

2. Modulatory connections

Modulatory connections are the primary motivation for the development of conflict learning. They are found extensively in feedback projections related to visual processing, for example from visual cortex to the thalamus (Cudeiro & Sillito, 2006; Jones et al., 2012, 2015), from higher visual areas to primary visual cortex (Callaway, 2004; Hupe, James, Girard, Lomber, Payne et al., 2001), as well as from posterior parietal cortex to V5/MT (Friston & Büchel, 2000). Top-down modulatory influences also play a role in phenomena such as attention (Baluch & Itti, 2011; Beuth & Hamker, 2015; Yantis, 2008), object segmentation (Roelfsema, Lamme, Spekreijse & Bosch, 2002), and object recognition (Bar et al., 2006). Attention is a modulatory effect and has the greatest impact on already active representations (Buschman & Kastner, 2015). Modulatory feedback, used in much the same way as in our border ownership experiment, has been used to construct a model of attention that replicates numerous observed attentional effects on both firing rates and receptive field structure (Miconi & VanRullen, 2016).

Modulatory connections can alter the existing activation of a neuron, but cannot cause activity in isolation; they must work in conjunction with driving inputs (Brosch & Neumann, 2014b). We can observe this distinction mathematically by first looking at the activation function for an artificial neuron, which is typically modeled by some function of its weighted inputs:

$$x_j = f\left(\sum_{i \in \text{input}} x_i w_{ij}\right) \quad (1)$$

where w_{ij} is the weight between neurons i and j and x_i is the activation of neuron i .

However, as modulatory connections are defined as those that do not directly drive the activation of a neuron, their effect must be distinguished from driving connections, which, in similar fashion to Brosch and Neumann (2014b), we formalize as:

$$x_j = f(D_j + g(D_j, M_j)) \quad (2)$$

where $D_j = \sum_{i \in \text{driving}} x_i w_{ij}$ and $M_j = \sum_{i \in \text{modulatory}} x_i w_{ij}$. g is a monotonically increasing function with respect to D_j and $D_j = 0$ implies that $g(D_j, M_j) = 0$. Typically, g is a simple product between D_j and M_j (e.g., Bayerl and Neumann, 2004; Brosch and Neumann, 2014a; Roelfsema et al., 2002), hypothesized to be implemented biologically by backpropagation-activated coupling (Larkum, 2013).

When feedforward inputs are taken to be driving and feedback to be modulatory, it can be said that feedback is gated by feedforward, an effect noted by Larkum (2013). Roelfsema et al. (2002) discuss the idea of gating in detail and use it to support a model of figure-ground segregation. This gating allows networks to integrate feedback without struggling to balance it against feedforward input or incurring spurious top-down-driven activation. The physiological mechanics of modulation have been best studied in relation to the thalamus, with a recent review by Varela (2014) showing that modulatory input is extensive and heterogeneous in regard to origin, neurotransmitter, and function. Brosch and Neumann (2014b) discuss the evidence for the potential physiological implementation of modulatory feedback while developing a network-level circuit model for feedforward and feedback interaction.

2.1. Hebbian learning and modulatory connections

Traditional Hebbian based learning rules adapt weights based on some function of the coincidental firing of pre and postsynaptic neurons:

$$\Delta w_{ij} = f(w_{ij}, x_i * g(x_j)) \quad (3)$$

Hebbian learning in its most basic formulation has no mechanism to bound weight growth, making it trivially unstable. For our purposes we use a formulation of Hebbian learning that includes a normalization component for stability, adapted from Stevens et al. (2013):

$$\Delta w_{ij} = \frac{w_{ij} + \eta x_i x_j}{\sum_k (w_{kj} + \eta x_k x_j)} - w_{ij} \quad (4)$$

where η is the learning rate. This weight update, and its normalization, are applied independently to driving and modulatory connections (i.e. all w_{ij} are the same connection type).

To better understand why such a Hebbian rule is not suitable for learning modulatory connections, let us look at the dynamics of a minimal network with two competitive neurons, illustrated in Fig. 1. In this context, competitive means that the neurons are connected such that more active neurons inhibit the activation of those less active through lateral connections. The desired state of this network is to have each competing neuron develop a strong connection to a unique source of modulatory input. It should be noted that this end state is considered desired due to its computational usefulness as a source of top-down information rather than a direct extrapolation from biology.

We can imagine this network as, for example, a simple attention network concerned with detecting apples or oranges in its input. The modulatory connections act as attentional biases towards either apples (M_1) or oranges (M_2). Though one fruit may be desired over the other (e.g., searching for a specific fruit; M_1 active versus M_2), the network has no control over what is present in its input. Features related more to apples (N_1) or to oranges (N_2) may be active regardless of the bias signal, even occurring simultaneously.

Download English Version:

<https://daneshyari.com/en/article/4946707>

Download Persian Version:

<https://daneshyari.com/article/4946707>

[Daneshyari.com](https://daneshyari.com)