A Solid Bridge Between the Biology and Cognition of Learning

Randall C. O'Reilly[1]

[1] Dept of Psychology,
Center for Neuroscience,
and Institute of Cognitive Science,
University of Colorado Boulder
E-mail: randy.oreilly@colorado.edu

A significant gap has persisted between the known biological mechanisms of long-term potentiation and depression (LTP/D) and the learning mechanisms that can solve difficult cognitive problems. In broad brushstrokes, the biology appears to support various forms of unsupervised or self-organizing learning mechanisms (e.g., Hebbian learning), whereas more powerful error-driven learning mechanisms appear necessary to solve challenging cognitive problems like learning to recognize many different types of visual objects, or learning to pronounce English words. We report a novel learning mechanism that was derived in part from a highly-detailed biological model (developed by other researchers) of known LTP/D mechanisms, and naturally exhibits a powerful combination of error-driven and self-organizing learning. This learning mechanism solves a range of challenging cognitive problems, while exhibiting superior generalization and other desirable properties relative to existing error-driven learning mechanisms. As such, it appears to represent the strongest bridge yet between the biology and cognition of learning.

Since Francis Crick's sharp critique (*1*) in 1989 on the biological implausibility of the error-driven backpropagation algorithm (*2*) (which has been used to simulate a wide range of cognitive phenomena and played a critical role in the resurgence of connectionist models (*3*)), the divide between the biology and cognition of learning has remained, and perhaps grown wider. At the biological level, a huge amount has been learned since 1989 about the detailed neural mechanisms and chemical signalling pathways involved in long-term potentiation and depression (LTP/D). The current consensus is that the biology supports some form of spike timing dependent plasticity (STDP) (**?**), which is a form of Hebbian self-organizing learning that is sensitive to the precise timing of sending and receiving spikes. A number of computational models have attempted to show how STDP arises as a result of the known biological mechanisms (**?**). The current

"last word" in this literature is a highly-detailed model by Urakubo and colleagues (**?**), which replicates several STDP experiments, including those with spike triplets and quadruplets that provide strong constraints on the temporal and other dynamics of the underlying mechanisms. The range of data captured by this model at different levels of biological and chemical analysis is extremely impressive, and it represents a modern high point in the bottom-up approach to neural modeling.

From the cognitive side of things, error-driven learning mechanisms (like backpropagation) remain the dominant approach for solving hard cognitive problems. Although there are many examples where Hebbian or STDP-like learning mechanisms have exhibited some useful forms of learning, these are inevitably with small numbers of highly simplified stimuli (**?**, **?**, **?**). Any attempt to capture people's ability to recognize thousands of different visual objects, or correctly pronounce thousands of different written words, invariably relies on the power of error-driven learning. Mathematically, this is not surprising: biologically-realistic Hebbian-style self-organizing learning mechanisms are good for capturing various low-order statistics about the input stimuli, but have no guarantee of solving complex input-output mapping problems. In contrast, error-driven learning mechanisms like backpropagation are mathematically guaranteed to converge on a minimum-error solution (modulo the problem of local minima in the error-surface). Recent developments in learning with deep networks (having many internal representation layers) rely on backpropagation and another variant of error-driven learning to produce impressive results (**?**).

Thus, there is a significant and troubling gap between the biological mechanisms that we know underlie learning, and the evident capabilities of the learning mechanism at a cognitive level. A large number of potential solutions have been offered to bridge this gap, including various proposals for how error-backpropagation can be made considerably more compatible with the biology (**?**, *4, 5*), and various ways of leveraging global neuromodulatory signals driven by reward prediction error (e.g., dopamine) to drive local synaptic learning in a partially-error-driven manner (**?**). However, the gap remains, in the former case because critical holes remain in the linkage between the biology and the necessary learning mechanisms, and in the latter case because these reinforcement-driven learning mechanisms are very slow to learn, if they converge at all, and do not appear to scale well to large, hard problems.

We report here some significant progress in bridging the gap, by deriving a learning rule directly from the biologically-detailed Urakubo et al model, and integrating it with some ideas from both the BCM (Bienenstock-Cooper-Munro; (**?**)) and CHL (Contrastive Hebbian Learning (**?**)) algorithms. Specifically, we subjected the Urakubo et al. (*?*) model to extensive testing with systematically varying sequences of sender and receiver firing patterns, and measured the resulting synaptic efficacy (*weight*) change produced by the model. A highly regular pattern emerged, which can be captured with a linear correlation value of

$r = .894$ using the learning function shown in Figure 1. In retrospect, this function is perhaps not particularly surprising: it is essentially a linearized version of the long-known LTP/D functional dependence on levels of intracellular Calcium: low levels produce LTD (weight decrease), while high levels above a given threshold produce LTP (weight increase). In addition, a more rounded, polynomial version of this same function lies at the heart of the BCM learning rule. Interestingly, the extracted curve is a function of the *time integral* of neural activity at the synapse over a roughly 1 second period (i.e., sum over 1s of number of sending spikes times sum over 1s of number of receiving spikes). The effects of detailed variations in spike timing get largely washed out when integrating over a realistically-long spike train, as has been noted both in the empirical literature on STDP with spike triples and quadruplets (**?**), and the STDP computational modeling literature (**?**).

Three key insights transform this simple learning function into a powerful learning mechanism that naturally integrates features from both error-driven and self-organizing BCM learning mechanisms. First, the key idea behind BCM is that the potentiation threshold (θ_p, where LTD transitions into LTP) should be a function of the overall activity of the receiving neuron: neurons that are highly active should have higher thresholds, while less active ones should have lower thresholds. This results in a homeostatic dynamic that causes neurons to carve the overall representational space into roughly equal-activity chunks, and typically results in individual neurons representing statistically reliable and separable elements of the environment. There is ample evidence for this BCM-like homeostatic mechanism in the brain (**?**), so we start by adopting that into our model.

However, BCM is purely self-organizing, and does not offer any guarantee of solving challenging cognitive mapping problems (and nor has there been any demonstration thereof). The second insight shows how error-driven learning can arise from the time-integral nature of the learning rule (which has not been incorporated into any BCM-like rules that we are aware of). Specifically, transient synaptic activations (less than roughly 300msec in duration) result in LTD, while activity that is more persistent results in LTP. This suggests a form of error-driven learning where errors correspond to transient activations, which is a very natural dynamic for networks that generate expectations about sequelae for a given state, and then experience the actual outcomes: incorrect expectations are squelched by the correct outcomes (resulting in LTD), while correct expectations experience a long period of activation (resulting in LTP) (*4–6*). However, a critical feature of error-driven learning is that correct expectations correspond to zero error, and should not experience any weight change at all, instead of the strong LTP that this mechanism would suggest.

The third insight addresses this issue, by noting that the most recent period of neural activity is most strongly weighted in an exponential running average computation of the sort typically used to adapt the

potentiation threshold θ_P in the BCM algorithm. Thus, if a neuron is stably active in both the expectation and outcome phases, one might expect the potentiation threshold to be higher than for a neuron that only came on in the outcome phase, but was not active in the expectation phase. An extreme version of this dynamic can result in pure error-driven learning, which is essentially equivalent to a temporally-integrated version of the CHL learning rule (see supplemental info). To integrate the BCM and error-driven learning elements in a clean and effective way, we set the potentiation threshold to be the MAX of the longer-term BCM-style running average, and the medium-term average over the current trial of processing, reflecting the error-driven component:

$$\theta_P = MAX(\overline{x_m y_m}, \gamma_P \overline{x_l}) \tag{1}$$

This MAX formulation ensures that which-ever factor is stronger between the long time-average activation or current-trial activation will dominate, producing a stronger enforcement of the relevant constraints than an alternative mixture of these two factors. Note that the Urakubo model itself does not exhibit a dynamic potentiation threshold, and nor is there substantial empirical data to help constrain the design of such a mechanism. Therefore, we have been guided here primarily by functional performance, which is quite good with the above formulation.

Finally, we note that the same argument about recent short-term factors dominating a running average computation also apply to the synaptic activation term that drives learning, which means that the outcome phase is going to have a stronger influence on learning than the expectation phase. This helps to ensure that outcome states get sufficiently active to drive LTP, and further helps to drive the error-driven learning dynamic described above (see supplemental for more details).

To summarize, the resulting learning rule states that learning is based on the time integrated synaptic (sender * receiver) activation (with stronger influence from the most recent firing), relative to a dynamic threshold based on medium and longer term running averages, with the shape of the learning function derived directly from the very detailed Urakubo et al model. This rule is purely local to the synapse, and can operate in continuous time with no other phase-based synchronization or similarly "artificial" dynamics or mechanisms that were potentially problematic in earlier formulations (**?**). Both error-driven and BCM-style learning dynamics emerge naturally from the network, and appear to work synergistically together to produce very good learning results across a wide range of different cognitively challenging learning tasks (Figures 2-4, and supplemental info).

Although there are still several features of this learning rule that remain to be empirically tested, the learning rule overall seems to avoid many of the most obvious problems from prior attempts to bridge the gap between the biology and cognition of learning.

References

1. F. H. C. Crick, *Nature* **337**, 129 (1989).

2. D. E. Rumelhart, G. E. Hinton, R. J. Williams, *Nature* **323**, 533 (1986).

3. D. E. Rumelhart, J. L. McClelland, *Parallel distributed processing: Explorations in the microstructure of cognition*, vol. 1 and 2 (MIT Press, Cambridge, MA, 1986).

4. R. C. O'Reilly, *Neural Computation* **8**, 895 (1996).

5. R. C. O'Reilly, Y. Munakata, *Computational Explorations in Cognitive Neuroscience: Understanding the Mind by Simulating the Brain* (The MIT Press, Cambridge, MA, 2000).

6. J. L. McClelland, *International Perspectives on Psychological Science, Volume I: Leading themes*, P. Bertelson, P. Eelen, G. D'Ydewalle, eds. (Erlbaum, Hillsdale, NJ, 1994), pp. 57–88.

7. Supported by ONR grants N00014-00-1-0246 and N00014-03-1-0428, and NIH grant MH069597-01. We thank the staff of the Colorado Neurological Institute for their assistance, and Karen Richardson for help in administering cognitive tasks to patients and senior participants. We also thank Tim Curran, Gary McClelland, and Yuko Munakata for helpful comments.