Chapter 8

Learning and synaptic plasticity

Our brains enable us to produce models of our surrounding environment and to predict the effect of our actions on that environment. In particular, they allow us to learn from experience so, in an ever-changing environment, we can be more successful at our endeavors than any organism whose behaviors and responses are genetically preprogrammed before birth. The initial building of these models and the lifelong process of updating them in response to new information rely on changes in neural circuits within our brains. That is, synaptic connections between neurons change in their pattern and in their strengths through development and through lifelong learning experiences. Such changes in connections are known as synaptic plasticity.

8.1. Hebbian plasticity

In his 1949 book, The Organization of Behavior[1], Donald Hebb wrote the following prescient sentence:

> *When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one of both cells such that A's efficiency, as one of the cells firing B, is increased.*

The sentence, known as Hebb's postulate or Hebb's rule, was suggested as a method for producing the circuitry needed for persistent activity via a reverberatory loop of excitatory feedback (Chapter 6, Figures 6.3-6.5). While Hebb's rule contains a causal direction, so that cell A should fire before cell B for the connection strength to change, in this chapter the term Hebbian plasticity will be used if coactivity of two excitatory neurons leads to an increase in connection strength in either direction.

Box 8.1. Hebbian plasticity: A change in synaptic strength produced by a positive correlation, which should be causal, between presynaptic spikes and postsynaptic spikes.

Experimental support for Hebbian plasticity has accrued since 1966[2] when long-term potentiation between excitatory cells was first discovered by Lømo then later characterized by Bliss and Lømo[3] using hippocampal slices from rabbits. Later, in 1997[4,5], the observation of spike-timing dependent plasticity provided evidence for the causal aspect of Hebb's rule. The biophysical mechanisms underlying Hebbian plasticity, which manifests as an increase in the strength of synaptic connection between the two neurons, have been gradually revealed over time thanks to numerous painstaking studies.

The particular mechanisms for Hebbian plasticity vary across cell-types and brain areas, but all depend on a transient increase in calcium concentration within the postsynaptic cell in the vicinity of the synapse[6-9]. Such an increase can arise via influx through calcium channels[10] (Section 4.5), NMDA receptors[11] (Section 5.1.2), or efflux from internal stores[12].

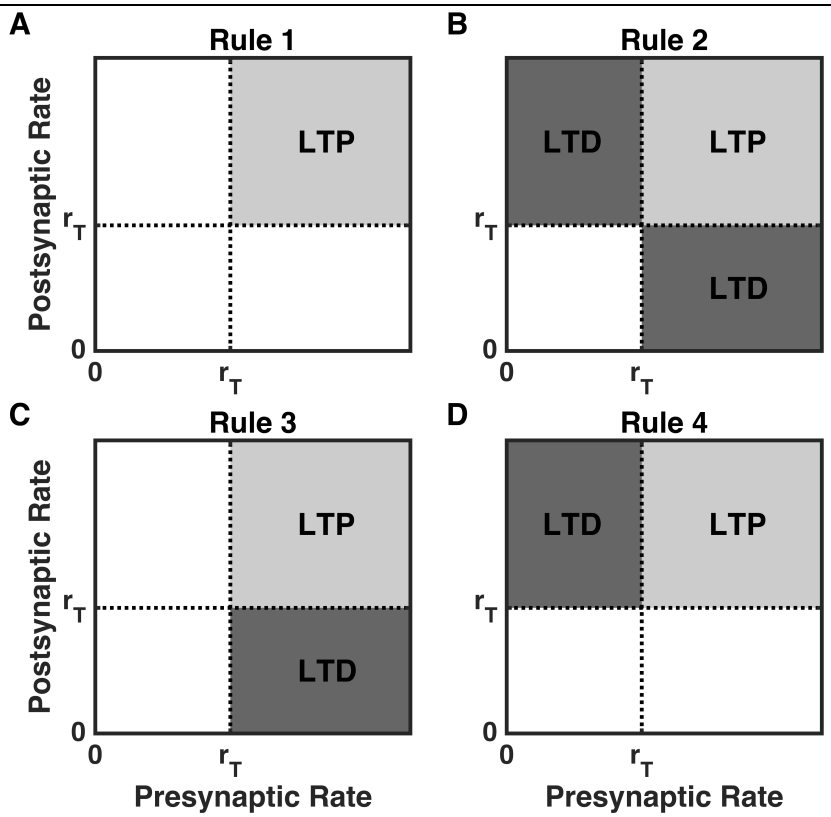

Figure 8.1. Alternative forms of Hebbian plasticity. A. Potentiation only (symmetric). If both presynaptic and postsynaptic cells' activities are above threshold then connections between the two cells strengthen. Otherwise no change in synaptic weight occurs. To prevent runaway excitation a global normalization is needed. (Rule 1, Tutorial 8.1). **B.** Potentiation and depression (symmetric). Potentiation occurs as in A, but if one neuron's activity is above threshold while the other neuron's activity is below threshold then any connection between the two cells weakens. If both neurons have low activity there is no change. (Rule 2, Tutorial 8.1). **C**, **D.** Potentiation and depression (asymmetric). Potentiation occurs as in A. Depression only occurs if the presynaptic cell's activity is above-threshold while the postsynaptic cell's activity is below-threshold in C, while depression only occurs if the postsynaptic cell's activity is above-threshold while the presynaptic cell's activity is below threshold in D. This figure is produced by the online code - `plasticity_four_rules.m`.

The NMDA mechanism is most easily understood[13]: A presynaptic spike causes glutamate to be released so that it binds to the postsynaptic receptors where it can remain for many tens of milliseconds. If, during that period, the postsynaptic membrane potential rises to a high level—as it could if the postsynaptic cell fires a spike or burst of spikes—then the NMDA channel opens and admits calcium ions. Importantly this mechanism is directional—if the postsynaptic cell is active before the presynaptic cell, the high membrane potential—which returns to baseline after a few milliseconds—never coincides with the glutamate binding, so the NMDA channels do not open. We simulate this process in Section 8.3.6.

Box 8.2. Long-term potentiation (LTP): An increase in synaptic strength that persists for longer than a few tens of minutes.

Box 8.3. Long-term depression (LTD): A decrease in synaptic strength that persists for longer than a few tens of minutes.

It is important to note that Hebbian plasticity on its own is destabilizing for two reasons. First, Hebb's postulate only provides a rule for an increase in the strength of synapses (as in Figure 8.1A)—if the only possibility were for synapses to strengthen, they would all strengthen by chance eventually, leading to runaway excitation in the brain. So, an opposing mechanism is required to allow synapses to reduce their strength. Indeed, in parallel with the work on long-term potentiation, rules and mechanisms for long-term depression have also been elucidated[14,15].

Hebb, of course, was aware of such a need. In order for his hypothesized reverberatory loop of excitation to be contained within a specific subset of cells, it was important that those cells inhibit, rather than excite other cells outside of the subset. Therefore, the active cells should weaken their excitatory connections to less active excitatory cells, whose occasional, spontaneous spikes would otherwise cause synaptic strengthening. The observed data on long-term depression (not to be confused with short-term depression of Section 5.3.1) follow this rule qualitatively—particular protocols for inducing long-term depression include weak depolarization of the postsynaptic cell (to mimic its being only weakly excited) while causing the presynaptic cell to fire spikes[16], or low frequency stimulation of presynaptic axons[17].

Even with such an opposing form of synaptic plasticity, the two mechanisms together are still destabilizing because they are each a form of positive feedback: The change produced enhances the likelihood of further change. That is, potentiated synapses are likely to cause even more activity in the postsynaptic cell following presynaptic activity, resulting in further potentiation. Similarly, depression of excitatory synapses increases the likelihood of low activity in the postsynaptic cell at the time of presynaptic activity, resulting in increased likelihood of further depression. The end point of these processes alone, is for a subset of synapses to ever increase in strength up to whatever their biologically constrained maximum might be, while a complementary subset of synapses ever decreases in strength until the synapses are eliminated. Later in this chapter, in Section 8.6, we consider homeostatic processes that counter such synaptic extremism.

8.1.1. Modeling Hebbian plasticity

We will first simulate Hebbian plasticity using a firing rate model—we will explore simulations with spiking neurons in Section 8.3, when we investigate the consequences of spike-timing dependent plasticity. In the simpler method here, we ignore the causal aspect of Hebb's rule (*i. e.*, the order of firing does not matter) and determine the change in connection strength between two units via the covariation of their firing rates. In particular, if the two units have high activity at the same time then we strengthen the connection between them. This can be written using a binary function for the rate of change of connection strength, so that plasticity is either on or off, as:

$$\tau \frac{dW_{ij}}{dt} = \Theta(r_i - r_T) \cdot \Theta\left(r_j - r_T\right) \quad \text{Eq. 8.1}$$

where $\Theta(x)$ is the Heaviside function, defined as $\Theta(x) = 1$ if $x > 0$ and $\Theta(x) = 0$ if $x < 0$; W_{ij} is the strength of connection from unit i to unit j; τ sets the rate of change; and r_T is the threshold rate for inducing Hebbian plasticity (see Figure 8.1A).

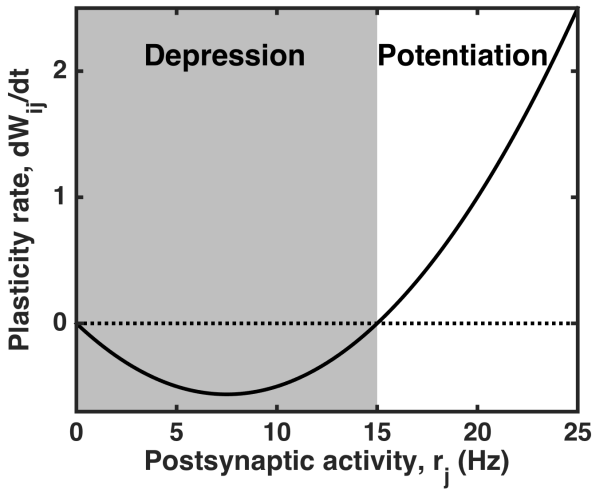

Figure 8.2. Plasticity rule with quadratic dependence on postsynaptic firing rate. The cross-over from synaptic depression to potentiation occurs when the postsynaptic firing rate increases above the threshold rate, which is 15Hz in this example. In some models the threshold rate can vary slowly over time, increasing if the postsynaptic cell has a long-term high firing rate and decreasing if the postsynaptic cell has a long-term low firing rate. Such metaplasticity, as proposed in the BCM model [18], was an early suggestion of homeostasis (see Section 8.6) to counter the instabilities of Hebbian plasticity. This figure is produced by the online code `quadratic_plasticity_figure.m`.

We shall investigate the impact of such a rule, in combination with a rule for reducing synaptic strength in Tutorial 8.1. Other rules are possible, in particular the rate of change of synaptic strength may be a continuous, graded function of firing rates rather than a binary function. A rule that incorporates such rate-dependence and includes depression as well as potentiation is:

$$\tau \frac{dW_{ij}}{dt} = r_i r_j \left(r_j - r_T\right). \quad \text{Eq. 8.2}$$

In Eq. 8.2, r_i, the rate of the presynaptic cell, simply scales the rate of plasticity. However, r_j, the rate of the postsynaptic cell, both scales the rate of plasticity and alters the sign of plasticity, as the rule yields synaptic depression if $r_j < r_T$ and synaptic potentiation if $r_j > r_T$ (Figure 8.2). This formula can also be derived from plasticity rules in spiking neurons (*cf.* Eq. 8.7 in Section 8.3.4, Triplet STDP, and see the derivation in Appendix B of this chapter).

Box 8.4. Pattern completion. A network produces a correct pattern based on prior examples when provided with a partial or corrupted input of that pattern, just as yu cn dcphr thse wrds.

Box 8.5. Pattern separation. A network produces different responses for different patterns, so the particular pattern that is completed depends on which one the input most closely resembles.

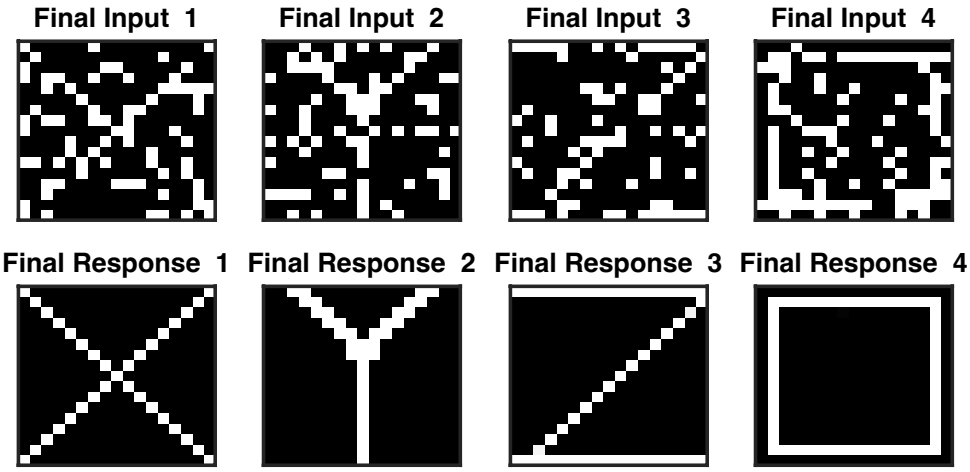

Figure 8.3. Pattern completion and pattern separation of corrupted exemplars following Hebbian learning. Each small square in the 17x17 array corresponds to a unit. The coordinated activity of a set of units can be entirely produced by input (top row) or, after learning, can represent a particular memory stored in the connections between units (bottom row). In this example, inputs generate an activity pattern with approximately 80% overlap with one of the "ideal" templates (20% of units have their input switched sign, top row). After learning, the structure of feedback connections within the network causes the units to retrieve the activity corresponding to the closest "ideal" template (bottom row). That is, the network has learned an "ideal" template, even though the exact template was never presented during Hebbian learning—the corrupted exemplars were sufficient. See Tutorial 8.1. This figure was produced by the online code `Tutorial_8_1.m`.

8.2. Tutorial 8.1 Pattern completion and pattern separation via Hebbian learning

Neuroscience goals: produce a powerful model of associative learning in the brain; gain awareness of how different synaptic plasticity parameters impact such learning. Computational goals: converting matrices to vectors and back; implementing matrix multiplication to calculate the results of multiple integrals.

In this tutorial, you will train a circuit undergoing Hebbian synaptic plasticity with many approximate examples of four distinct patterns, then test whether the circuit reproduces the desired pattern when presented with new approximations to each pattern. The system is a type of autoassociative network[19], the most well-known of which is the Hopfield network[20]. These networks have the property of pattern completion (retrieve a complete pattern given partial input) and pattern separation (produce distinct responses to similar inputs according to the most similar learned pattern of each input). The learned patterns become attractor states of the network (see Figure 8.3).

1a) Define four matrices of size 17x17 (you can call them `pattern1`, `pattern2`, `pattern3`, and `pattern4`) whose entries are 1s and 0s arranged in a distinct pattern that you can recognize (in the lower panel of Figure 8.3, the light entries correspond to active neurons with an entry of 1 in the pattern and the dark entries correspond to inactive neurons with an entry of 0 in the pattern). You may choose your own distinct patterns (with neurons defined on a 17x17 grid) or use the patterns in the lower row of Figure 8.3. Do not use patterns with a lot of overlap between any of them.

b) Set up a loop of 400 trials and on each trial:

(i) Select at random one of your 4 patterns to be the input. (The numpy function `random.randint` is useful here. Otherwise, you can take a random number between 0 and 1, multiply it by 4, add 0.5 to your result and round to the nearest integer).

(ii) Randomly, with probability of 0.1 (*i. e.*, use the `random.rand` function), flip entries from 0 to 1 or 1 to 0, to generate a corrupted pattern, `input_pattern`, with on average 90% overlap with the original one.

(iii) Assign each element of `input_pattern` to produce the applied current to a particular firing-rate unit (so you will have $N_{units} = 17^2 = 289$ units). Simulate for 1s the set of firing-rate units whose activity will be stored in an array of size $N_T \times N_{units}$, where N_T is the number of time-points using:

$$\tau_r \frac{dr_i}{dt} = -r_i + \frac{r^{max}}{1 + \exp[-(I_i - I_{th})/\Delta_I]}$$

$$I_i = \sum_{j=1}^{N_{units}} r_j W_{ji} + I_i^{app}(t)$$

(see Chapter 6, Eqs. 6.1-6.3 for a review of the notation).

Use the parameters $\tau_r = 10$ms, $r^{max} = 50$Hz, $I_{th} = 10$, $\Delta_I = 1$. Initially, the connection matrix between the units, W_{ji}, provides uniform inhibition with $W_{ji} = -0.3/N_{units}$ identical across all connections. All firing rates should be initialized as zero. The applied current, $I_i^{app}(t)$, should only be present for the first 0.5s of the trial. Its value is 50 for those units, i, receiving input (*i. e.*, those units whose entry in `input_pattern` is 1) and its value is 0 otherwise.

Note that in Python, `input_pattern` is a matrix, typing "`input_pattern.flatten()`" produces the required vector of length N_{units}.

(iv) At the end of each trial, alter the connection strengths according to Rule 4 (Figure 8.1D) which can be written as:

$$\Delta W_{ji} = \varepsilon_+ \int_{t=0}^{t=1} \Theta(r_j - r_T) \cdot \Theta(r_i - r_T) dt - \varepsilon_- \int_{t=0}^{t=1} \Theta(r_T - r_j) \cdot \Theta(r_i - r_T)\, dt$$

where the time-integrals are across the trial of 1s, $r_T = 25$Hz, $\varepsilon_+ = 0.1/N_{units}$ and $\varepsilon_- = \varepsilon_+/4$. In this equation, the rate of the presynaptic cell is r_j and the rate of the postsynaptic cell is r_i. The Heaviside functions, for example $\Theta(r_i - r_T)$, give 1 if the

term in parenthesis is greater than zero (*e. g.* if $r_i > r_T$ in the example) and otherwise 0. In Python the numpy function `heaviside(rate-r_thresh,0)` can be used to produce a binary thresholded rate "`binary_rate`". The parameters ε_+ and ε_- determine the rate of plasticity—if too slow, more trials are needed, if too fast, random errors can be imprinted in the circuit.

These integrals can be evaluated rapidly via matrix multiplication. For example, the first of the two terms can be evaluated as:

```
epsilon_plus*(binary_rate.transpose()@binary_rate)*dt
```

if firing rate is stored as a matrix, `rate`, with the first index (row number) corresponding to the time point and the second index (column number) corresponding to the neural unit number. `dt` is the time-step used in the code (so `1/dt` is the number of time-points), `r_thresh` the threshold, r_T, and `epsilon_plus` is the simulation's term for the parameter ε_+.

(v) Ensure no connection strengths pass beyond a maximum value of $W_{max} = 8/N_{units}$, or a minimum value of $W_{min} = -8/N_{units}$, by clamping them to these values if necessary.

c) Once the 400 trials with plasticity are completed, simulate four new trials, one with each input pattern but now with 20% of entries to the input matrix flipped from 0 to 1 or from 1 to 0. Plot, using "`imshow`" the set of firing rates at the end of the simulation (*i. e.*, the last time-point), being careful to assign the rate of the unit labeled i (where i runs from 1 to $N_{units} = 289$) to its correct location on the 17x17 grid. For example, you can use the numpy `reshape` command in Python, to convert a vector to a matrix.

<u>OPTIONAL</u>

2) Repeat the learning paradigm using one or more of the other plasticity rules (Rules 1-3 of Figure 8.1) and/or with altered values for ε_+ and ε_-. Assess whether some rules are better than others at achieving perfect recognition of corrupted patterns by using increasingly corrupted patterns during learning and testing (only allowing 400 trials of learning).

Note: When you use Rule 1, in order to prevent all connections from becoming maximally strong, you should (after updating the connections) renormalize them to maintain a fixed mean incoming connection strength to each unit. For example, a line of code like:

```
W = W - np.mean(W,0)+ W0_mean;
```

where `W0_mean` is the initial fixed connection strength ($-0.3/N_{units}$), ensures the mean incoming synaptic inputs to each unit never change (if rows represent presynaptic cells and columns represent postsynaptic cells, so the mean of each column is subtracted).

8.3. Spike-timing dependent plasticity (STDP)

Box 8.6. Spike-timing dependent plasticity (STDP): Changes in connection strength that depend on the relative timing of presynaptic and postsynaptic spikes, not just their rates.

While Hebb's postulate contains an important causal direction as cell A can "take part in" firing cell B without cell B taking part in firing cell A, the rules for strengthening of connections described so far in this chapter lack such causality. Indeed, they may be summarized by the distorted shortening of Hebb's postulate, "cells that fire together wire together". Yet, shortly before the turn of the millennium, the first experimental data were published demonstrating the importance of the temporal order of neural activation[4,5].

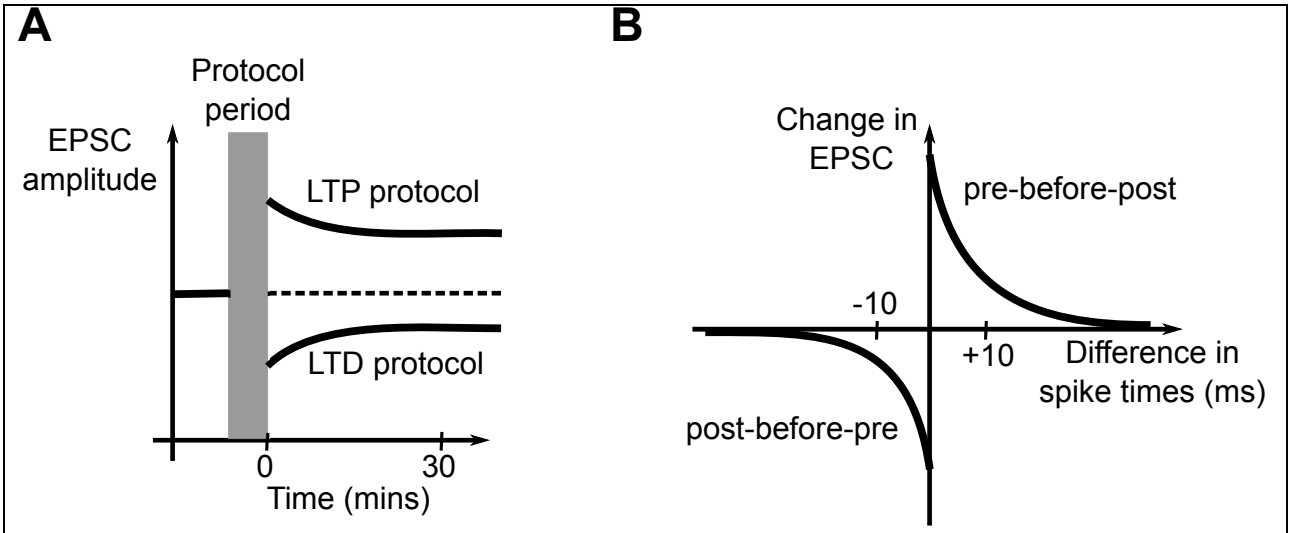

Figure 8.4. Spike-timing dependent plasticity. A. The amplitude of excitatory post-synaptic current (EPSC) can change following a pairing protocol. The change in amplitude can be sustained for tens of minutes or many hours following the protocol (the experiments are *in vitro* so there is a time-limit for sustaining healthy cells). The protocols comprise several minutes of combined presynaptic spikes and postsynaptic depolarization or hyperpolarization or spiking, and last several minutes. Depending on the protocol long-term potentiation (LTP, an increase in evoked EPSC) or long-term depression (LTD, a decrease in evoked EPSC) can be produced. **B.** In a protocol using 10Hz trains of presynaptic spikes (a spike every 100ms) the change in synaptic strength between two excitatory cells is positive if the postsynaptic cell is caused to spike within 20ms following each presynaptic spike (pre-before-post), and is negative if it is caused to spike in a similar time-window preceding each presynaptic spike (post-before-pre). The dependence on the time-difference between the two spikes can be fitted as two exponentials. Such spike-timing dependent plasticity demonstrates a causal requirement for a change in connection, a requirement that was present in Hebb's original postulate[1].

In particular, in protocols with electrodes implanted in both of a pair of connected excitatory neurons, a connection increased in strength when the postsynaptic neuron produced an action potential within a small time-window (on the order of tens of ms) after

the action potential of the presynaptic neuron. Conversely, when the ordering of spikes was reversed the synaptic connection strength decreased. This process means that in a bidirectionally connected pair (a common motif in cortical circuits) whenever one connection strengthens, the reciprocal connection weakens, because the relative order of presynaptic and postsynaptic spikes is reversed for the reciprocal connection.

Some caveats were noted for these data at the time and over later years[21]—a minimum rate of spike-pairs was required[22], combined with a minimum number of pairings to produce significant change; after a few tens of pairings no further change occurred[23]; the variability in the amount of change across different synapses undergoing the same protocol was huge[5]. Yet to produce simple models of spike-timing dependent plasticity (STDP) whose consequences can be easily explored and understood, we ignore these caveats in this book. We therefore define STDP as a change in synaptic strength that is produced by each pair of pre- and post-synaptic spikes, and that depends only on the time-difference between the two spikes (Figure 8.4B).

Even with this simplification of the process, alternative methods still remain for simulating the total change in synaptic strength arising from the trains of action potentials of two connected cells.

First is the choice of whether to include all spikes from one neuron as the partner to form a spike-pair with each and every spike of the other neuron, or to include only "nearest-neighbor" spikes. At low firing rates this choice makes little difference, but once either of the neurons produces a significant number of spikes separated by intervals smaller than the time-window of STDP, then the choice to omit many spike-pairs can even shift the net result from LTD to LTP, or vice versa (as derived in Appendix A, with result shown in Figure 8.13).

Since the original data used to produce figures such as Figure 8.4B were acquired from the sum of all successive synaptic changes at a fixed rate of pairings, the original data do not constrain which rule should be used. That is, so long as two rules can correctly model the original data, that data cannot favor one rule over the other even when they produce opposite results for a different protocol.

Second is the choice of whether to use a batch method (simulate the spike trains for a trial and only update the synaptic strengths after the simulated trial) or a continuous method (update the synaptic strengths at the time of each spike within a simulated trial). The positive feedback that is inherent in standard rules for STDP—a change in connection leads to a change in spike patterns that makes the same change in connection more likely to happen again—means that the continuous method of updating, which more rapidly employs the positive feedback, leads to greater overall changes in synaptic strength than the batch method. Again, since synaptic changes are typically measured after many pairings whose individual effects can accumulate, the original data do not provide a reason to support one method over another.

Box 8.7. Batch updating method: A protocol with which all synaptic strengths are updated at the end of discrete blocks of simulation, or trials, with the set of synaptic strengths kept constant during each simulated block of time.

Box 8.8. Continuous updating method: A protocol for updating synaptic strengths within a simulation, typically at the time of each spike, but possibly on every time-step.

8.3.1. Model of STDP

The standard model of STDP requires an update of the connection strength, W_{ij}, from neuron i to neuron j whenever either neuron i or neuron j produces a spike. At the time, t_i, of a presynaptic spike the update, ΔW_{ij}, is negative, because any postsynaptic spikes used to update the synapse would be at earlier times, t_j, so the spike order is postsynaptic cell' spike before the presynaptic cell's spike (post-before-pre):

$$\Delta W_{ij} = -A_{-}\exp\left[\frac{-(t_i - t_j)}{\tau_{-}}\right] \quad (\text{with } t_i > t_j). \qquad \text{Eq. 8.3}$$

Conversely, at the time, t_j, of every postsynaptic spike the update is positive, because any presynaptic spikes have arrived at earlier times, so the spike order is presynaptic cell's spike before postsynaptic cell's spike (pre-before-post):

$$\Delta W_{ij} = A_{+}\exp\left[\frac{-(t_j - t_i)}{\tau_{+}}\right] \quad (\text{with } t_j > t_i). \qquad \text{Eq. 8.4}$$

By specifying these changes at the time of each spike we are able to implement the continuous update rule. It is computationally straightforward at the time of each spike of each cell to update weights based on only the most recent spike of the connected cell, or to store all prior spike times of connected cells and add together the contributions of all those prior spikes to changes in synaptic strength. In Tutorial 8.2, we will just assume the most recent spike is used in each update.

When implementing STDP via the batch method, it is straightforward to simply sum the contributions of all presynaptic spikes of a given presynaptic cell for a given postsynaptic spike, then sum over all presynaptic cells for each postsynaptic spike, and finally sum over all postsynaptic spikes in a set of nested loops.

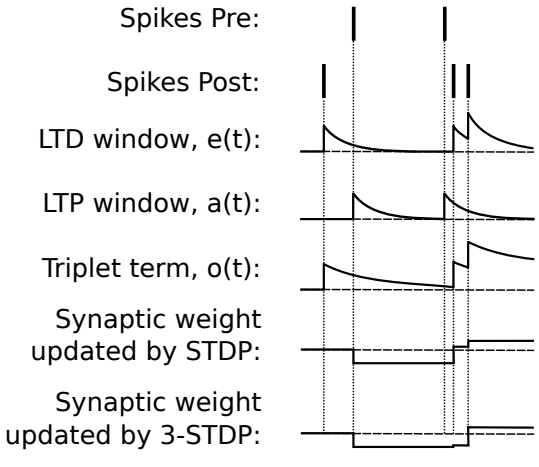

Figure 8.5. Continuous update rule for STDP and triplet-STDP (3-STDP). Synaptic weights can increase (LTP) at the time of a postsynaptic spike (row 2) and decrease (LTD) at the time of a presynaptic spike (row 1). In STDP (row 6) the amount of LTD is given by $e(t)$ (row 3), which depends on the time since the prior postsynaptic spike, whereas the amount of LTP is given by $a(t)$ (row 4), which depends on the time since the prior presynaptic spike. In triplet STDP (bottom row) the amount of LTP is also multiplied by an extra factor, $o(t)$ (row 5), which depends on prior postsynaptic spikes. This leads to minimal LTP at the time of the second postsynaptic spike but enhanced LTP at the time of the third postsynaptic spike that shortly follows.

For the continuous method, it is better to continually update two additional variables for each neuron, i—one of which, $a_i(t)$, represents the impact of all presynaptic (afferent) spikes, the other of which, $e_i(t)$, represents the impact of all postsynaptic (efferent) spikes (Figure 8.5). These variables follow:

$$\frac{da_i}{dt} = -\frac{a_i}{\tau_+} \text{ and } \frac{de_i}{dt} = -\frac{e_i}{\tau_-}. \qquad \text{Eq. 8.5}$$

(If you are modeling neurons that are only either presynaptic or postsynaptic but not both, then only one of these variables per neuron is needed.) At the time of a spike of neuron i, the variables are updated in the same manner, either being incremented by 1 if all spikes are used (*e. g.*, $a_i \rightarrow a_i + 1$) or set to 1 if nearest-neighbor spikes are used (*e. g.*, $a_i = 1$). Then, at the time of each postsynaptic spike of cell j the synaptic strength is incremented by an amount $\Delta W_{ij} = A_+ a_i$ and at the time of each presynaptic spike of cell i the synaptic strength is decremented according to $\Delta W_{ij} = -A_- e_j$.

8.3.2. Synaptic competition via STDP

Box 8.9: Synaptic competition: A situation in which the increase of strength of some synapses causes a decrease in strength of others.

You will notice that each of the update rules has two parameters, an amplitude such as A_+, which determines the change in synaptic strength for near-coincident spikes and a decay time constant such as τ_+, which determines how close together in time the two spikes must be to cause a significant change. The area between each curve and the x-axis is given by the product of amplitude and decay time constant, *i. e.* $A_+\tau_+$ for LTP and $A_-\tau_-$ for LTD. The relative values of these two areas is important, because if two spike trains are uncorrelated such that the differences in their spike times fall with equal likelihood across the region of the STDP window, then the mean amount of potentiation is proportional to the area under the LTP curve while the mean amount of depression is proportional to the area under the LTD curve. (This result assumes all spike pairs are used to update synaptic strength, not just nearest neighbor pairs, but it also holds if only nearest-neighbor pairs are used so long as the mean inter-spike intervals of each cell are significantly longer than the decay time constants of the plasticity rule, *i. e.*, at low firing rates – see Appendix A).

Since we only want net strengthening if there is a causal relationship between presynaptic and postsynaptic spike times—and a causal relationship produces correlations—it is important that the mean amount of LTP is less than the mean amount of LTD for uncorrelated spike trains. That is, in models we require $A_-\tau_- > A_+\tau_+$. Such a requirement allows for competition for the following reasons:

1) If a subset of input synapses is strengthened then postsynaptic activity increases.
2) Increased postsynaptic activity leads to more pairs of pre- and post-synaptic spikes and shorter time intervals between one neuron's spikes and the other's. Both of these effects lead to enhanced synaptic plasticity.
3) If other presynaptic neurons have spikes that are uncorrelated with the additional spikes in the postsynaptic cell then their input synapses to the postsynaptic cell become more rapidly depressed, because the net effect of uncorrelated pre- and post-synaptic spikes is depression.

4) The combined effect is competition—the increase in strength of one subset of input synapses to a neuron causes the decrease in strength of another subset of its inputs.

Therefore, STDP can cause a postsynaptic cell to enhance its response to one set of inputs at the expense of another set of inputs. Such selection is called synaptic competition, which can be useful in two ways that we will investigate in Tutorial 8.2.

First, it can lead to a separation of two distinct sets of mixed inputs[24]. For example, if a mixture of inputs, each originating from activity in the retina of either the left eye or the right eye, impinge in equal amounts on a set of cells, one cell may eventually respond to only the left eye's inputs while another cell responds only to the right eye's inputs. Such selectivity, called ocular dominance, is observed in the primary visual cortex of many species[25,26].

Second, if a sensory neuron is receiving a subset of correlated inputs amidst a majority of uncorrelated inputs, it is likely that the correlated subset contains the most useful information from the environment. In this case STDP allows the neuron to respond more strongly to the correlated set and to "filter out" the noise produced by the uncorrelated inputs[24]. Such plasticity produces a more faithful neural response to environmental signals.

8.3.3. Sequence learning via STDP

If a sequence of external stimuli causes neurons to fire in a particular order repeatedly, then STDP strengthens connections from those neurons that fire early in the sequence to those neurons that fire later in the sequence (Figure 8.6A). These strengthened connections can be sufficient to cause the postsynaptic neuron to fire following a presynaptic spike, in the absence of the external stimuli[27]. Thus, STDP provides a mechanism for sequence completion[28]. Relatedly, in the phenomenon of prospective coding, neurons that fire at specific locations on a track (called place cells) shift their receptive fields to earlier locations if an animal passes along the track many times[29]. The explanation[30-32] relies upon a temporally asymmetric plasticity rule (like STDP, but suggested before the discovery of STDP) strengthening connections in a directional manner from neurons that respond at one location to neurons that respond at a later location on the route. Following sufficient synaptic strengthening, those neurons that initially responded later are driven to respond earlier.

Perhaps sequence learning is most important in auditory processing—sound is inherently dynamic and words with meaning are formed by sequences of phonemes, whose order is essential. It takes minimal effort for us to continue phrases or verses of songs we know, when provided with the beginning. A similar type of learning is essential for juvenile songbirds to learn a stereotypical tune from an adult songbird, known as the "tutor"[33]. The songs are sequences of syllables—shorter musical phrases—which can either be learned as a fixed sequence or with some variability based on transition probabilities between syllables[34,35]. Therefore, the songbird system has proven to be of great help in our understanding of the plasticity mechanisms that can underlie the learning and generation of sequences[36].

8.3.4. Triplet STDP

The standard method of implementing STDP, in which all spike-pairs are included in the update of synaptic strength, lacks the qualitative rate-dependence observed in early experiments[37]. These early experiments showed that synaptic stimulation at high frequency produced potentiation, whereas at low frequency it produced depression. Later STDP protocols suggested a similar frequency-dependence of synaptic plasticity, with potentiation favored as the rate of the postsynaptic cell increased (*cf.* Figures 8.1A and 8.2).

The "nearest-neighbor" approach to implementing STDP does include such a qualitative rate-dependence (Figure 8.13), such that the crossover rate for the postsynaptic cell to switch the plasticity from depression to potentiation is. However, in this implementation, the crossover rate increases with the presynaptic firing rate, so that high presynaptic rates favor depression, unlike the observed data.

To better extract and model the combined rate and timing dependence of STDP, in the years following its initial discovery, various groups assessed the impact of more complex or naturalistic protocols including triplets and quadruplets of pre- and postsynaptic spikes on synaptic strength[38,39]. Here we discuss a rule, triplet STDP[40], that fits these later empirical data extremely well and that produces both firing-rate dependence as well as spike-timing dependence of the direction of synaptic plasticity.

As its name suggests, the triplet rule relies on the relative times of a triplet of three spikes. These can be two presynaptic and one postsynaptic, or one presynaptic and two postsynaptic (Figure 8.5). In the formulation used here, we only include the latter pattern and only for the potentiating term of STDP, maintaining the depression term's dependence on only a pair of spikes (one presynaptic and one postsynaptic).

Box 8.10. Triplet STDP: Changes in synaptic strength that depend on the relative times of up to three different spikes from the connected neurons, such as one from the presynaptic cell and two from the postsynaptic cell.

In this formulation, the modification of the standard STDP rule is to include an additional multiplicative factor when potentiating synapses. The factor decays exponentially from the time of the previous postsynaptic spike. Therefore, the amplitude of potentiation depends on two exponential factors. The first factor is the one we have seen already that is based on the time-difference between the presynaptic spike and postsynaptic spike with a time constant on the order of 20ms. The second is an extra term, whose decay time constant is slower, on the order of 100ms. Specifically, we consider all triplets of spike-times t_i, t_j, and $t_{j'}$, where t_i is the time of a presynaptic spike, t_j and $t_{j'}$ are the times of postsynaptic spikes with $t_i < t_j$ and $t_{j'} < t_j$. We then increase the synaptic strength by an amount:

$$\Delta W_{ij} = A_3^+ \exp\left(\frac{t_i - t_j}{\tau_+}\right) \exp\left(\frac{t_{j'} - t_j}{\tau_3}\right). \qquad \text{Eq. 8.6}$$

In this case, the continual update rule can be more efficient as we can represent the contribution to the extra, final term of Eq. 8.6 from all the postsynaptic spikes of a neuron with a single variable, $o(t)$, which increases by one at the time of each postsynaptic spike and, thereafter decays with the appropriate time constant (Figure 8.5).

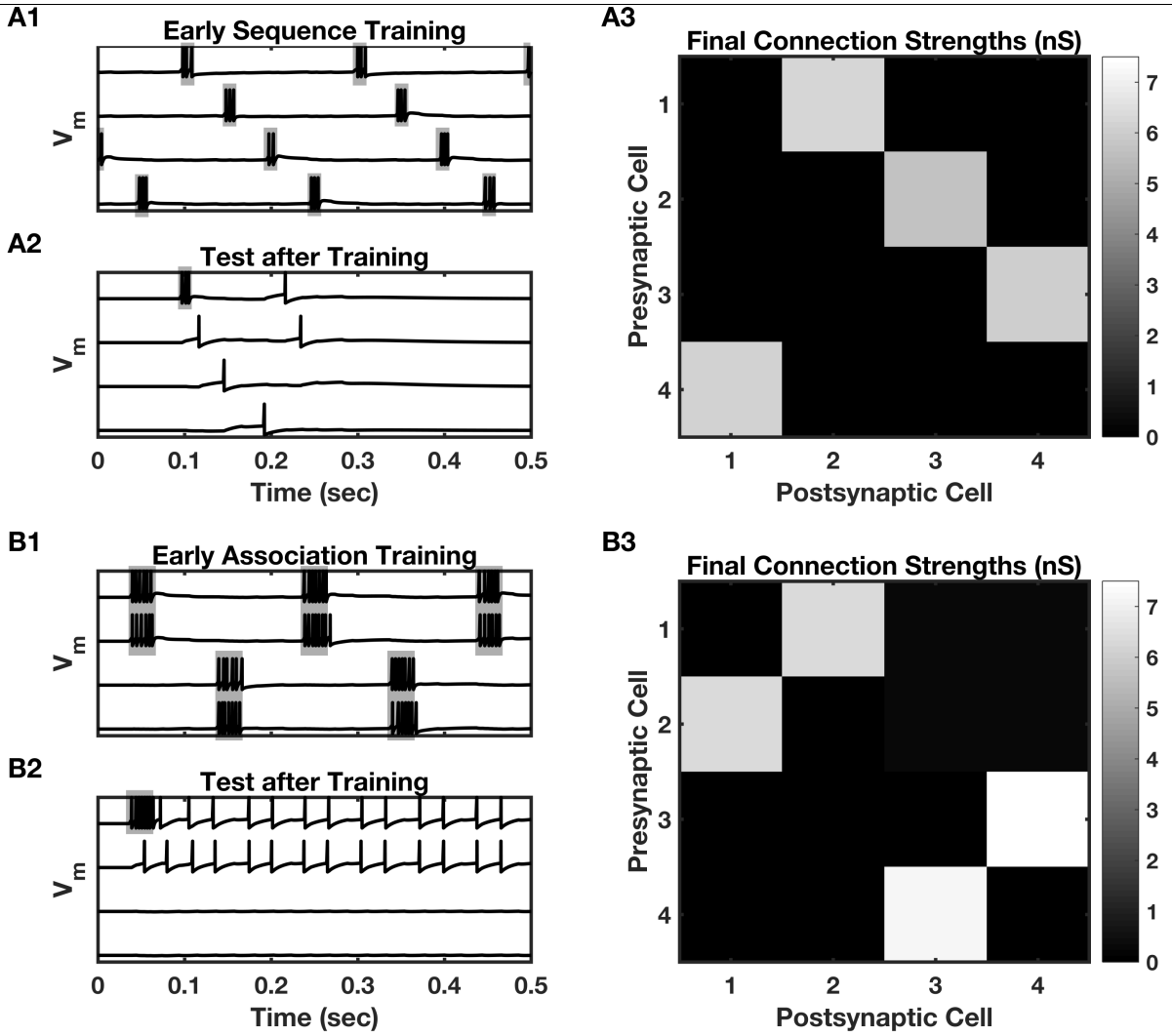

Figure 8.6. Encoding of sequences and paired associations using triplet STDP. Four excitatory neurons are stimulated repeatedly by external input (gray shaded rectangles), either in sequence (**A1**) or simultaneously as pairs of two neurons (**B1**). During training the connections between the four neurons undergo synaptic plasticity using a mechanism of triplet-STDP (Figure 8.5 and Eq. 8.6). After training, stimulation of a single neuron (single gray shaded rectangle) causes a replay of the sequence (**A2**) or of activity in the neuron paired with the stimulated neuron (**B2**) according to the prior training. These responses arise because triplet-STDP can encode asymmetric synaptic weights in a unidirectional pattern following sequence training (**A3**), or autoassociate the paired neurons in a symmetric, bidirectional pattern following paired training (**B3**). This figure is based on the works of Pfister, Clopath, Busing, Vasilaki, and Gerstener[40,41]. The code used to generate this figure is available online as `STDP_recurrent_final.m`. Feedback inhibition is used in the code to enhance stability and it is worth noting that following further training in pattern B, the pairs of neurons would maintain activity in the delay between stimulations. Thereafter the sequence of stimulations lacks a silent period of activity between pairings, so the pairs of neurons connect with each other in sequence (as in A) until eventually all neurons are highly connected. Therefore, the simple model used here is ultimately unstable.

The term producing depression is unchanged from that of standard STDP (Eq. 8.3), so depression is present at low postsynaptic rates. Therefore, when it is rare for two postsynaptic spikes to fall within a window of 100ms of each other, potentiation is suppressed so depression dominates if spike times are uncorrelated. However, at high postsynaptic firing rates, potentiation is possible and—if its peak amplitude, A_3^+, is high enough—potentiation can dominate, even when spike times are uncorrelated (producing a positive feedback that could drive coactive synapses to their maximum strength).

In Appendix B, we solve for the rate dependence of triplet STDP when the presynaptic and postsynaptic spike trains are produced as two uncorrelated Poisson processes at the given rates. The result is:

$$\frac{dW_{ij}}{dt} = r_i r_j \left(A_3^+ \tau_+ \tau_3 r_j - A_- \tau_- \right), \qquad \text{Eq. 8.7}$$

which increases linearly with the presynaptic rate, r_i, but is quadratic in the postsynaptic rate, r_j, switching from net depression to net potentiation when the postsynaptic rate increases past a threshold of $A_- \tau_- / (A_3^+ \tau_+ \tau_3)$. Notice the equivalence of Eq. 8.7, produced from a model of spike timing, with Eq. 8.2 and Figure 8.2, which were based on firing-rate models.

8.3.5. A note on spike-timing dependent plasticity

Any empirical rule for altering the synaptic strength between two neurons on the basis of the times of presynaptic and postsynaptic spikes is an approximation that ignores a great deal of intracellular complexity. The electrophysiological and biomolecular processes set into motion in the postsynaptic cell, particularly those in the neighborhood of one synapse, depend on many additional factors—local densities of ion channels, local concentrations of enzymes, proximity and activity of other synapses, distance of the synapse from the soma, and the size and shape of the cell. Such factors are typically unknown and can be treated as a source of random variability, or simply ignored if the system is already variable enough, or one can attempt to account for as much of the underlying biophysics as possible.

However, if we want to simulate many neurons and their interactions, it can be convenient to use an empirical rule for updating synaptic strengths, rather than attempt to simulate all of the intracellular processes that are rarely measured in a functioning circuit. In many cases the empirical plasticity rule can substitute for a more complex process. For example, in some cases a dendritic calcium spike may be essential to induce synaptic plasticity and without such a spike no plasticity is induced[42]. Since such calcium spikes are highly correlated with a period of high frequency dendritic input and with postsynaptic activity, we can simply produce a rule for synaptic plasticity that requires a high rate of postsynaptic spikes or presynaptic inputs to produce a change in strength. Such a procedure is more efficient than extending the simulations of each neuron to the level where we can more accurately predict the timings of the underlying calcium spikes.

8.3.6. Mechanisms of spike-timing dependent synaptic plasticity

Given the importance of synaptic plasticity for brain development and memory formation, a large wealth of data has been accrued over the last 50 years with the goal of pinning down the signaling pathways involved. While significant differences can arise across brain areas, across species, across developmental stages, and across cell-types, for the majority

of—if not all—mechanisms, the time-course and local calcium concentration is the key initiating factor for synaptic plasticity. While in some cells it is calcium entry through voltage-gated calcium channels that is most important, one of the most understood mechanisms that can relate causality of neural firing to the level of calcium concentration is when considering calcium entry through NMDA receptors.

NMDA receptors require both the binding of glutamate—the neurotransmitter released by an excitatory presynaptic cell in the cortex—and depolarization of the membrane potential in order to open and admit entry to calcium ions. Being situated in the membrane of the postsynaptic cell, they are the ideal mechanism for detecting a coincidence of presynaptic and postsynaptic spikes, so long as the postsynaptic spike causes significant depolarization of the membrane potential in the vicinity of the synapse. For synapses on the dendrites far from the soma, a dendritic calcium spike may be essential for the necessary depolarization to occur.

Moreover, NMDA receptor activation is strongest when the presynaptic spike precedes the postsynaptic spike. This is because the glutamate ligand remains bound to the receptor for many tens of millisecond (50ms is a typical time constant at *in vivo* temperature) whereas the membrane potential decays within a millisecond of a spike. Therefore, NMDA receptors can implement the causal aspect of "pre-before-post" in Hebb's postulate that is measured in STDP protocols (see Figure 8.7).

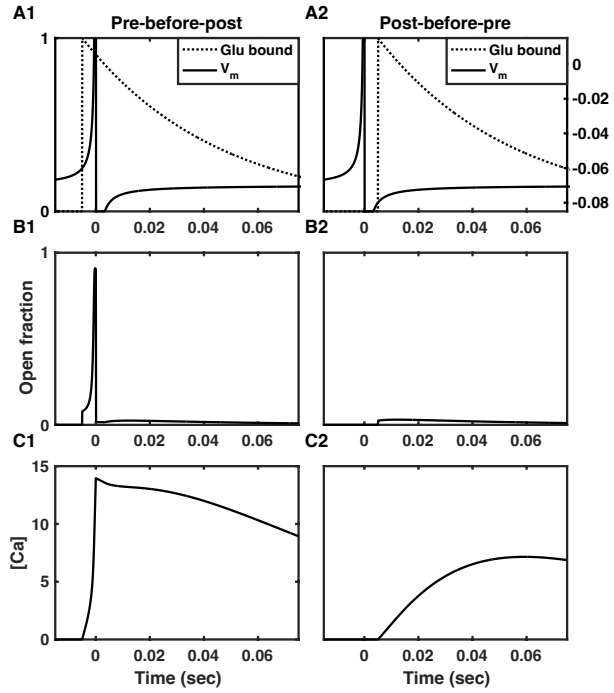

Figure 8.7. NMDA receptor activation as a temporal-order detector. A1. If the presynaptic spike precedes the postsynaptic spike, the transient postsynaptic depolarization (solid line, scale on right axis) overlaps with the period of glutamate binding to the NMDA receptor (dotted line, scale on left axis). **B1.** As a result, NMDA receptors open rapidly to admit calcium. **C1.** Calcium concentration then rises rapidly to a high level in a manner that could lead to LTP. **A2.** If the postsynaptic spike precedes the presynaptic spike, then the rapid membrane depolarization is over before glutamate binds to the NMDA receptor. **B2.** Receptors open only partially. **C2.** Calcium concentration rises slowly to a lower peak value in a manner that could lead to LTD. This figure is produced by the online code `NMDA_activation.m`.

8.4. More detailed empirical models of synaptic plasticity

In nearly all simulations based on spiking neurons, the membrane potential of the neuron is the most significant variable in the model. I say "nearly all" because it is possible to

simulate the spikes of a neuron simply by gathering statistics of its inputs and its history of spike times to determine (usually probabilistically) whether a spike should be generated at any instant. However, when the membrane potential itself is simulated—as in all spiking models considered in this book—then there is minimal simulation cost to basing synaptic plasticity on the membrane potential rather than the spike times alone.

In a successful voltage-dependent rule that is compatible with the empirical data[41], the timing of presynaptic spikes combines with the recent history of postsynaptic membrane potential to determine the direction and magnitude of any synaptic plasticity. Depression arises at the time of a presynaptic spike whenever it follows a period of moderate depolarization. Potentiation arises at the time of high postsynaptic membrane potential (equivalent to the time of a spike) if this follows both a presynaptic spike (as in STDP) and a period of moderate depolarization of the postsynaptic membrane potential. Therefore, the rule is similar to triplet-STDP, the main distinction being that for both depression and potentiation the initial postsynaptic spike is replaced by a period of moderate depolarization that need not include a spike. In particular, LTD can occur following presynaptic spikes in the absence of postsynaptic spiking, as observed empirically.

If one additional variable should be simulated to make better contact with realistic plasticity mechanisms, then the calcium concentration in the postsynaptic cell is the clear choice. However, care is needed because the calcium concentration varies tremendously across different domains and organelles of any cell. The calcium concentration that drives independent plasticity at each synapse is necessarily different from synapse to synapse. So, if calcium concentration is to be used then a separate variable is needed for each synapse. In Section 8.6 we will see that somatic calcium concentration, which depends on the cell's firing rate, can be used for coordinated, cell-wide homeostatic regulation of all synapses in the cell.

Box 8.11. Homeostatic regulation: Control of a system via feedback from the measurement of at least one variable, with the goal of ensuring the system remains in a stable and useful operating regime.

8.5. Tutorial 8.2. Competition via STDP

In this tutorial, you will see how STDP can produce a selective response in a neuron by preferentially strengthening one subset of its incoming synapses while weakening all other inputs. This tutorial is based on a paper by Sen Song and Larry Abbott (2001)[43].

1) Generation of input spike trains and synaptic conductance.

a) Define 50 input spike trains, grouped into two subsets (a, b) of 25. Each subset follows a single time-varying rate (r_a or r_b) given by:

$$r_{a,b}(t) = \frac{r_{max}}{2}\left[1 + \sin\left(2\pi\nu t + \phi_{a,b}\right)\right],$$

with maximum rate of $r_{max} = 60$Hz, frequency of oscillation of $\nu = 20$Hz, and phase offsets of $\phi_a = 0$ and $\phi_b = \pi$. The two phase-offsets can represent stimuli received via the two eyes.

b) Produce an array of input spikes, with 50 rows and a number of columns equal to the number of time-steps per trial (see below). Produce spikes according to an

inhomogeneous (time-varying) Poisson process with rate $r_a(t)$ for rows 1-25 (the first set of 25 inputs) and with rate $r_b(t)$ for rows 26-50 (the second set of 25 inputs). Each row represents a single trial of inputs, for which you can use a time-step of $\delta t = 0.1$ms and a duration of 0.5s.

c) Produce a vector of initial values of the synaptic strength, G_i, of each input synapse, labelled i (one per distinct incoming spike train) with all entries sampled from a Normal distribution with mean 500pS and standard deviation 25pS.

d) Produce a single vector representing the total (summed) synaptic excitatory conductance, $G_{syn}(t)$. Increment $G_{syn}(t)$ by an amount, G_i, at the time of each incoming spike at each input synapse, i. Between spikes $G_{syn}(t)$ decays exponentially to zero, with time constant $\tau_{syn} = 2$ms.

2) Simulate the spike-train of a leaky-integrate and fire (LIF) neuron receiving these inputs.

a) Simulate the membrane potential of an LIF neuron (see Chapter 2) that follows:

$$C_m \frac{dV}{dt} = G_L(E_L - V) + G_{syn}(E_{syn} - V); \text{ where if } V > V_{th} \text{ then } V \to V_{reset}$$

with parameters $C_m = 100$pF, $G_L = 5$nS, $E_L = -70$mV, $E_{syn} = 0$mV, $V_{th} = -50$mV, and $V_{reset} = -80$mV, initialized with $V(0) = E_L$.

b) Record all the spike times of the LIF neuron produced during the trial.

c) Plot the total synaptic conductance, $G_{syn}(t)$, and membrane potential, $V(t)$, as a function of time.

3) Update the input synaptic strengths using the STDP rule after the trial (Batch method).

a) For each input neuron and each input spike compare the time, t_{pre}, of the input (presynaptic) spike to the time, t_{post}, of each spike of the LIF neuron (the postsynaptic spike). Change the synaptic strength, G_i, of that input connection according to the rule:

$$G_i \to G_i + \Delta G_{LTP} \exp\left(\frac{t_{pre} - t_{post}}{\tau_{LTP}}\right) \text{ if } t_{pre} < t_{post},$$

$$\text{or } G_i \to G_i - \Delta G_{LTD} \exp\left(\frac{t_{post} - t_{pre}}{\tau_{LTD}}\right) \text{ if } t_{pre} > t_{post},$$

with $\Delta G_{LTP} = 20$pS, $\Delta G_{LTD} = 25$pS, and $\tau_{LTP} = \tau_{LTD} = 20$ms.

b) Ensure synaptic strengths are neither less than zero nor greater than a maximum bound, $G^{max} = 2nS$.

c) Record the mean synaptic strength for each set of 25 inputs and plot the set of synaptic strengths as a function of input number.

d) Explain how the synaptic strengths change across trials and how the changing synaptic strengths impact the time-variation of the input conductance and the pattern of the LIF neuron's spike train. Repeat the simulation one or two times and explain any differences you see.

4) "Train" the system.

a) Repeat 1-3 for 200 trials, but only producing plots every 20 trials.

b) Plot the mean synaptic strengths of the two sets of 25 inputs as a function of trial number.

5) Alter the phase offsets of two stimuli.

Repeat 1) - 4) using a phase-offset that alternates across trials, such that $\phi_a = 0$ on all trials, while $\phi_b = \pi/2$ on odd trials and$\phi_b = -\pi/2$ on even trials. The sign

reversals represent stimuli moving or eyes moving in either direction—input from one eye does not consistently precede the other.

6) Correlated versus uncorrelated stimuli.

a) In this question, the phase-offset of each oscillation producing the 50 inputs will vary, so you should produce a set of 50 rate vectors,

$$r_i(t) = \frac{r_{max}}{2}[1 + \sin(2\pi\nu t \pm \phi_i)],$$

where $\phi_i = 0$ for $1 \leq i \leq 10$ and ϕ_i is selected as a random number from 0 to 2π for $i > 10$. On odd trials set the phase offset as $+\phi_i$, and on even trials set it as $-\phi_i$. Use a frequency of $\nu = 10$Hz and otherwise repeat questions 1) - 4) with no further alterations.

b) Comment on any differences you see and plot the mean synaptic strength of the correlated group separately from the mean synaptic strength of uncorrelated inputs, as a function of trial number.

c) Plot the final synaptic strength of each neuron as a function of its phase offset.

8.6. Homeostasis

Homeostasis (literally "similar standing" or "equilibrium") is the process by which all living things maintain a stable internal environment that is conducive to their wellbeing. Perhaps the most well-known example is temperature regulation—we sweat to cool down if we are too hot and shiver to warm up if we are too cold. Since biophysical processes depend on chemical reactions, whose rates depend on temperature, salt concentrations, and pH, regulation of all these parameters is important. Also, at the cellular and organismal level properties such as size and energy use/supply must be tightly regulated.

Regulation is a term for negative feedback. "Feedback" means that the outputs of a system are monitored and used to alter the system's inner workings, or its inputs. The "negative" part means that the effect of feedback is in the direction that counteracts or ameliorates any change, whereas its opposite, positive feedback, would enhance any change. It is a curiosity of language that positive feedback, which sounds desirable in human interactions, is very bad for a complex system as it pushes processes toward instability and extremes. On the other hand, in biology and engineering, negative feedback is valuable and indeed essential, to ensure those processes remain in their desired operating regime.

Neural circuits have their own specific needs of such regulation or control. Space is at a premium in the brain, so neurons need the right level of connectivity and, therefore, appropriate dendritic ramification (branching). Neural activity is energetically costly, so neurons should not fire excessively, but nor should they be always unresponsive since they take up space and resources. Finally, the role of a neuron is to process information—minimally to pass on information that it receives, but usually to process the information with a useful filter or operation. Therefore, neurons should respond with more spikes at some times and fewer spikes at other times—*i. e.,* they should possess an input-dependent range of firing rates. Both their connectivity and their internal excitability contribute to their ability to produce the necessarily variable, information-rich responses.

When simulating neurons or neural circuits, you will have noticed that their behavior depends quite sensitively on parameters such as the maximum calcium conductance, or the strength of synaptic connections within a circuit. If these parameters

were not regulated *in vivo*, then neural circuits would quickly produce excessive activity (a cause of epilepsy) or become quiescent. These model parameters represent the combined effect of numerous proteins, which must be expressed, translocated, and inserted to the appropriate part of the cell membrane. Since the proteins are not static, but are replaced on a timescale of hours to days in a process called turnover, any maintenance of conductance or synaptic strength is based on a dynamic equilibrium as proteins are manufactured and inserted at the rate at which they are lost. The maintenance of these levels in the range that allows neurons to operate is, therefore, an example of homeostasis.

8.6.1. Firing rate homeostasis

For a neuron to convey or process information it should be able to modulate its firing rate in response to changes in its inputs. In particular, any neuron that is so unexcitable it remains quiescent, or that is so excitable it fires maximally at all times, would convey no information.

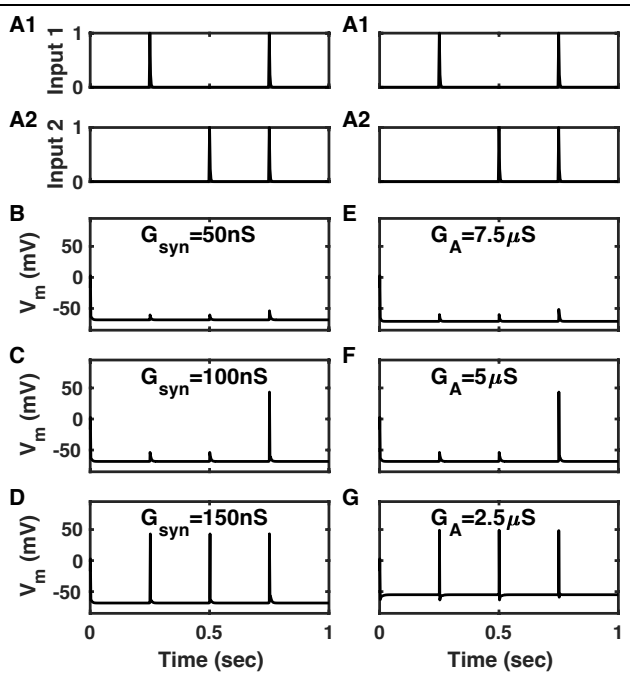

Figure 8.8. Example of the need for homeostasis in a neuron that responds to coincident inputs. A1-2) Synaptic inputs 1 and 2 to the cell. These should not produce a response individually (first spike in each panel) but only when they arrive together (second spike in each panel). **B, E)** If synaptic conductance is too low (B) or A-type potassium conductance too high (E), then the neuron never responds. **C, F)** At the standard level of synaptic and A-type potassium conductance, the neuron responds with a spike only to coincident inputs. **D, G)** If synaptic conductance is too high (D) or A-type potassium conductance is too low (G) then the neuron responds to every input, even single ones. Hebbian plasticity can rescue neither the unresponsive case (B, E) nor the overly responsive case (D, G). Figure produced by the online code `CS_3spikesin.m`.

Figures 8.8C & 8.8F show the typical response needed of neurons, albeit in the most simplified case of two inputs. In the example, the neuron is required for coincidence detection, so should produce a spike whenever its two inputs arrive at the same time, but not when they arrive alone. If the neuron is not excitable enough (Figure 8.8E) or its incoming synapses are too weak (Figure 8.8B) it does not respond to the coincident inputs. None of the types of synaptic plasticity we have looked at would rescue this situation—in fact, they may decrease synaptic strength making it even harder for the neuron to respond.

Conversely, if the neuron is too excitable (Figure 8.8G) or its incoming synapses are too strong (Figure 8.8D) then it responds to every input—a situation that would worsen when any of the Hebbian-like plasticity mechanisms we have considered cause strengthening of the incoming synapses.

You will notice that when the neuron fails, in one situation it does not respond to any input (Figures 8.8B & 8.8E) so it would be quiescent, whereas in the other situation (Figures 8.8D & 8.8G) it responds to every input so it would have excessively high activity. Therefore, if the neuron could monitor its firing rate and adjust the strengths of its afferent (incoming) synapses and/or its intrinsic excitability via a homeostatic firing-rate dependent mechanism, it could arrive at the situation shown in Figures 8.8C & 8.8F, where it performs a computation.

It is worth noting that the desired behavior (Figures 8.8C & 8.8F) could be achieved in a number of ways—for example, the neuron still responds as a coincident detector even with a 5-fold increase in synaptic strengths, so long as the A_type conductance is simultaneously increased 3-fold. This is a common feature of neural systems—there is a large range of "sweet-spots" producing useful function that may be maintained by coordinated or correlated changes in different parameters[44]. However, the systems have an even larger space of possible parameter combinations that lead to loss of function.

Box 8.12. Tetrodotoxin (TTX): A sodium channel blocker.

Box 8.13. Bicuculline: An antagonist of (meaning a chemical that binds without activating) $GABA_A$ chloride channels.

Both intrinsic excitability and excitatory synaptic strengths have been shown to change in a homeostatic manner following strong perturbations from equilibrium[45,46]. For example, if neurons are prevented from producing spikes (by application of tetrodotoxin, TTX, a venom found in species such as pufferfish) for a day or more, they become more excitable and their excitatory inputs become stronger. Conversely, application of bicuculline, which leads to more rapid firing by preventing inhibitory inputs from affecting a neuron, causes a reduction of excitatory synaptic strengths and a reduction of intrinsic excitability[47].

While the underlying molecular mechanisms are still being elucidated, overwhelming evidence favors a calcium-dependent negative feedback pathway[48]. The negative feedback is such that an increase in a neuron's firing rate causes an increase of somatic calcium, which ultimately causes a reduction in AMPA receptors or sodium channels (or increase in potassium channels) and a decrease in the neuron's firing rate. Here we gain some idea of the versatility of calcium-dependent signaling pathways. While a transient, high-amplitude pulse of calcium in the vicinity of a synapse can lead to local insertion of extra AMPA receptors to strengthen synapses in an LTP protocol, a long-lasting small increase in calcium in the soma can have the opposite effect on all of a neuron's incoming excitatory synapses.

To understand the impact of such homeostasis on circuit activity, we can bypass the internal mechanisms and let either a neuron's excitatory inputs or its intrinsic

excitability—for example, via the spiking threshold of a LIF model—depend negatively on its average firing rate:

$$\tau_g \frac{dG_E}{dt} = r_{goal} - r(t), \qquad \text{Eq. 8.8}$$

or

$$\tau_{th} \frac{dV_{th}}{dt} = r_{goal} - r(t). \qquad \text{Eq. 8.9}$$

In Eqs. 8.8 and 8.9, the time constants, τ_g and τ_{th}, should be long, on the order of hours, to mimic the observed slowness of homeostasis. However, when simulating the effect of homeostasis on a network it is sometimes necessary to reduce the time constants, in order to simulate hours rather than days of activity—the latter can take inordinately long when a time-step at the sub-millisecond level is needed to capture spikes, especially if large networks are being modeled.

8.6.2. Homeostasis of synaptic inputs

A particular need for homeostasis arises at the level of synaptic inputs to a cell, since those inputs undergo changes throughout the life of an animal as it develops and learns. Most Hebbian mechanisms—even STDP—lead to instability via positive feedback: stronger synapses are more likely to produce a postsynaptic response, which for Hebbian mechanisms leads to increased strengthening of those synapses and a greater response in the postsynaptic cell. In a recurrent circuit this effect is exacerbated, as the increased postsynaptic response produces greater input to other cells, leading to the potential for runaway excitation (as in Figure 6.4C).

One of the first suggestions for homeostasis of synaptic inputs, the 'BCM rule'[18] was via 'metaplasticity', meaning that the parameters of the synaptic plasticity mechanism should change slowly over time. In this case, the boundary between potentiation and depression of synapses that depends on postsynaptic firing rate (Figure 8.2) would slowly increase if mean postsynaptic rate were too high (making potentiation harder to achieve) or slowly decrease if the rate were too low (making potentiation easier to achieve).

However, the results of TTX treatment suggest that the synapses strengthen or weaken in a homeostatic manner without a requirement for synaptic input. Moreover, the distribution of synaptic strengths scales multiplicatively, a result that is consistent with the change in synaptic strength being proportional to the initial synaptic strength. These results suggest an actual rule of the form of Eq. 8.8, but with an extra term on the right-hand-side:

$$\tau_g \frac{dG_E}{dt} = G_E\left[r_{goal} - r(t)\right]. \qquad \text{Eq. 8.10}$$

A rule like that of Eq. 8.10 keeps the ratios of all synaptic strengths onto a neuron the same, as they all change in proportion. In this manner, any information stored in the synaptic strengths via prior Hebbian learning is maintained, while the overall input scales up or down to ensure the mean postsynaptic firing rate approaches r_{goal}.

8.6.3. Homeostasis of intrinsic properties

We saw in Section 2.3.1, that neurons of different sizes can function alike, to a first approximation, if the specific conductance—that is, conductance per unit area—of each ion channel is the same. This implies that, for example, if a neuron's activity should not change

as it grows, then the density of ion channels should remain constant. If the proteins comprising ion channels have a fixed lifetime, the total rate of manufacture and insertion into the cell's membrane of the proteins should be proportional to the cell's total membrane area. Yet it is unlikely that the cell's combined machinery of transcription and translation would, in the absence of modification, lead to protein expression at a rate that is proportional to the surface area of its membrane. Therefore, a homeostatic feedback process is likely to be necessary to ensure neurons retain their functional role as they change in size[49].

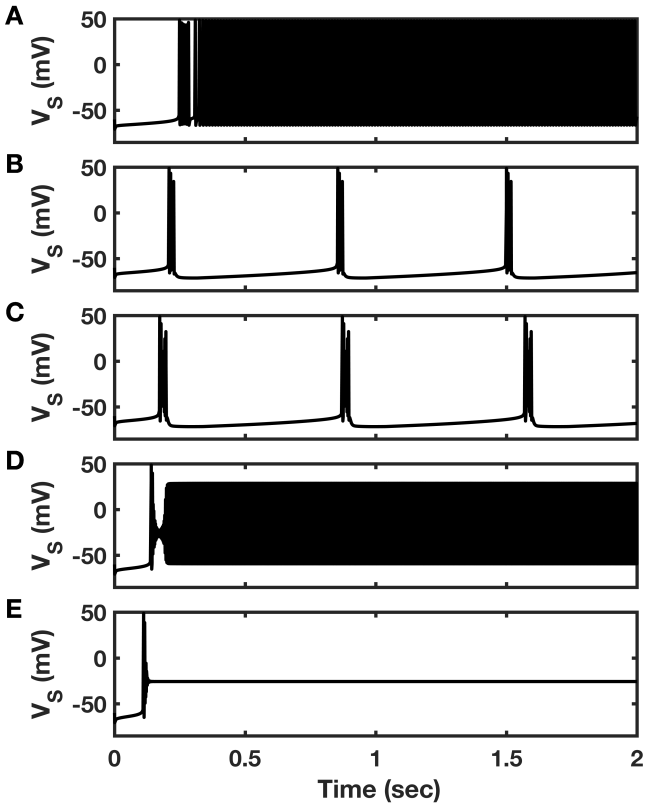

Figure 8.9. Need for homeostasis to produce intrinsic bursts. The somatic membrane potential V_S, indicates the pattern of neural firing in the Pinsky-Rinzel model (Section 4.7) when the maximum calcium conductance in the dendrite is: **A)** $A_D \times 1\mu$S, or **B)** $A_D \times 1.5\mu$S, or **C)** $A_D \times 2\mu$S, or **D)** $A_D \times 2.5\mu$S, or. **E)** $A_D \times 3\mu$S (where A_D is relative dendritic area). Here, mean membrane potential is non-monotonic in the maximum calcium conductance (with lowest mean membrane potential in panel C) so a feedback mechanism based on the membrane potential or firing rate would not be able to return dendritic calcium channel density to its ideal level (panel C) following all deviations. For example, panels A and D are similar, but require opposite homeostatic responses. This figure was produced by the online code `PR_euler_5cellversions.m`.

If a neuron is a pacemaker, its ability to produce bursts at the requisite frequency depends on parameters such as the density of calcium channels (Figure 8.9), the relative sizes of regions of the cell with distinct profiles of channel densities, and how well these different regions are electrically connected (Tutorial 4.3). It is not always clear how these parameters can be controlled. For example, from Figure 8.9, it is not apparent from the membrane potential whether a neuron spiking tonically at a high rate, as in panels A or D, should respond homeostatically by producing more calcium channels (the correct response in A) or fewer calcium channels (the correct response in D). It turns out that monitoring of dendritic calcium would answer this question, so could serve a role in regulation of bursting, as that quantity (not unexpectedly) increases monotonically with the maximum calcium conductance.

On the other hand, the frequency of bursts—an important quantity in the pacemaker of any motor system, in particular for heart rate or respiration rate—depends

in this model on the after-hyperpolarization-activated potassium conductance, G_{KAHP} (Figure 8.10). An **increase** in this conductance slows the rhythm of the pacemaker cell, which in turn causes a reduction of net calcium and a reduction in the calcium-activated potassium current, I_{KCa}, yielding an overall **decrease** in flow of potassium ions out of the cell (recall, potassium channels provide an outward current). Regulation of rhythmic frequency by G_{KAHP} would be difficult, therefore, as the system is operating in a regime such that the net effect on potassium concentration is opposite to the direct effect produced by the channel. Moreover, a feedback mechanism based on potassium concentration would be confounded by a change in the other dendritic potassium conductance G_{KCa}.

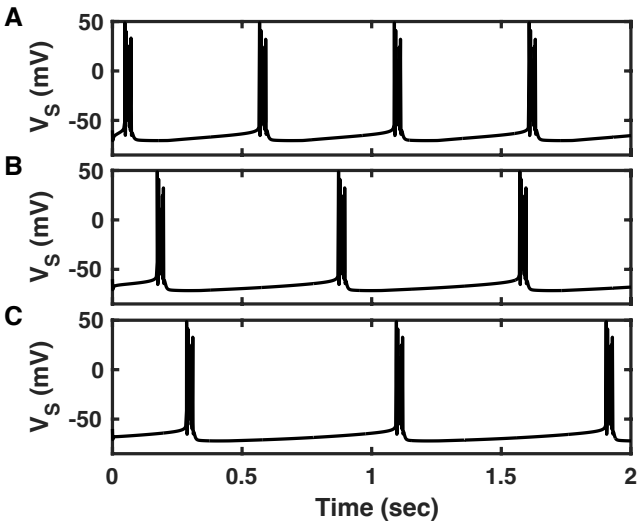

Figure 8.10. Need for homeostasis to control frequency of an intrinsically bursting neuron. The frequency of bursting depends on the maximum after-hyperpolarization conductance, a potassium conductance in the dendrites, with values of: **A)** $A_D \times$ 2nS; **B)** $A_D \times$ 4nS; and **C)** $A_D \times$ 6nS, where A_D is the relative dendritic area. In these examples the mean potassium current across the dendritic membrane decreases (because bursts are less frequent) even as this particular potassium conductance increases. The figure is produced by the online code `PR_euler_3cellversions_freq.m`.

8.7. Supervised learning

Examples of learning can be grouped into two broad categories, supervised and unsupervised. In the first part of this chapter we focused on Hebbian plasticity, which can underlie much unsupervised learning. The generation of preferential neural responses to the most common patterns of sensory input allows our brains to extract prevalent features in the environment (as in Tutorial 8.1 and Figure 8.3). This process is an example of unsupervised learning, which enables us to learn to recognize objects we have seen before, but cannot teach us what to do with those objects nor how to decide what to look at. The latter requires us to ascribe values and likelihoods to the outcomes that can be achieved if we take a particular action. The values and likelihoods of outcomes depend on our current state and on whatever environmental inputs we receive. The process of learning how to improve our actions in any situation in order to achieve a better outcome requires supervised learning.

Box 8.14. Supervised learning: Learning that depends on feedback based on the outcome of a behavior.

Box 8.15. Reinforcement learning: Supervised learning, in which the feedback signal is a scalar quantity, either positive or negative depending on the outcome of the behavior.

Supervised learning involves feedback, so that we can adapt our actions based on their consequences, not simply on the correlation of events. If unsupervised learning were the only type present, all we would have would be random, ill-formed habits—the reason for an action would simply be the history of such actions in the same circumstances, much as persistent repetitions of a sequence result in its spontaneous occurrence under STDP (Figure 8.6A). Such repetition, like a song replaying in our head, however annoying it is to us, is independent of whether the action provides a beneficial outcome for the animal. The addition of a feedback signal provides the necessary information for us to alter our behavior in a beneficial direction. The feedback signal can be binary ("good" or "bad"), or scalar (graded from "extremely good" through "OK" to "extremely bad"), or multidimensional ("I can feel myself tipping a bit to the left so should shift my weight a bit to the right"). The feedback itself can be mediated through self-sensation (*e. g.*, balance or proprioception), through direct sensory input (*e. g.*, pain from heat, or pleasure from food), or more acquired measures (*e. g.*, advice, social approval, or the promise of future reward).

The form of supervised learning that we will consider here is reinforcement learning[50]. In reinforcement learning the feedback provided after an action is a scalar signal, meaning it can be ascribed a numerical quantity. A better than expected outcome produces a positive reinforcement signal, while a worse than expected outcome produces a negative reinforcement signal. Unlike more sophisticated forms of supervised learning, reinforcement learning does not contain within the feedback any information as to how the action should be altered in order to improve the outcome. It is therefore ideal for use in situations with discrete alternatives. For example, if there are only two alternatives, positive reinforcement (commonly called "reward") should lead to a greater likelihood of a repetition of a choice, while negative reinforcement (commonly called "punishment") should lead to a greater likelihood of switching to the other alternative.

Since supervised learning is intimately connected with changes in behavior, it can be thought of as shaping our decisions. Indeed, in the task we consider later, (the weather prediction task, Section 8.7.3) the supervised learning is framed in terms of its impact on a decision-making circuit. The quantification and evaluation of human decisions is at the heart of the field of economics, so the study of how neural circuits respond to feedback in order to shape future behavior has been termed neuroeconomics.

Box 8.16. Neuroeconomics: The study of the neural basis of decision making in situations where the outcomes of those decisions are quantifiable.

8.7.1. Conditioning

Box 8.17. Unconditioned stimulus: A stimulus that causes a response or behavior automatically, or innately, without any need for training to produce the response.

Box 8.18. Conditioned stimulus: A stimulus which causes a particular behavioral response only after training (conditioning) in which it is associated with an unconditioned stimulus.

Box 8.19. Classical conditioning: Also called Pavlovian conditioning, the altering of an animal's behavior independent the animal's behavior, by associating a conditioned stimulus with an unconditioned stimulus.

Box 8.20. Operant conditioning: Also called instrumental conditioning, the altering of an animal's behavior, often through reinforcement learning, such that some responses to a set of stimuli are rewarded and/or others are punished.

The study and shaping of an animal's responses through presentation of stimuli coupled with reinforcement is the subject of conditioning. Two categories of conditioning should be clearly distinguished, classical conditioning and operant conditioning.
In "classical", also called "Pavlovian", conditioning, an initial stimulus (the conditioned stimulus, CS) is paired with a reinforcement signal (the unconditioned stimulus, US) on all trials, whatever action is taken by the animal. Such conditioning causes any innate, reflexive, or previously learned response to the unconditioned stimulus to become a response to the conditioned stimulus.

In "operant", also called "instrumental", conditioning, the reinforcement signal depends upon the animal's action. Therefore, the study of operant conditioning becomes important if we want to understand decision-making in terms of why we choose some actions or outcomes over others. The key to operant conditioning is learning to associate the value of a later outcome with the decision(s) that led to that outcome.

Box 8.21 Reward prediction error. The difference between the actual value of an outcome, positive if good, negative if bad, following an action compared to the expected value beforehand obtained from the weighted average of all possible outcomes.

8.7.2. Reward prediction errors and reinforcement learning
A neurotransmitter that plays an important role in conditioning is dopamine. While the complete set of circumstances that lead to dopamine release—including novelty and, in certain contexts, aversion—is still being explored, it has been long established that the activity of many dopamine-releasing neurons signals a reward prediction error [51-53].

A reward prediction error is positive if a rewarding stimulus, such as food, is received at a time when it was not entirely expected. Conversely, if an expected reward is omitted, the reward prediction error is negative. The greater the expectation of reward at a given time, as based on past experience, the smaller the reward prediction error when the reward is received, and the more negative the reward prediction error if the reward is not received. Importantly, the reward prediction error transfers from the primary reward, such as juice delivery, to the prior cues or stimuli that reliably predict such rewards. The phasic

(*i. e.*, transient, time-locked) activity of a large subset of dopaminergic neurons appears to follow these patterns typifying a reward-prediction error signal.

Such dopamine signals can be used as the scalar feedback signal underlying reinforcement learning. In general, reinforcement learning enables an animal to progress from a period of exploration, during which different alternative responses or behaviors are tried out, to a period of exploitation, once the consequences—usually evaluated in terms of the "value" of the state reached following a given action in a given situation—have been learned. Reinforcement learning can also help the animal learn to change established behaviors, so as to overcome intrinsic biases or undo previous learning when it is no longer applicable. For reinforcement learning to be effective, the likelihood of repeating an action should increase with the value of the state reached by that action.

	Postsynaptic rate	Reward prediction error	Change in synaptic strength
High	High	+	+
High	Low	+	−
Low	High	+	0
Low	Low	+	0
High	High	−	−
High	Low	−	+
Low	High	−	0
Low	Low	−	0

Table 8.1. A three-component plasticity rule that can be used for supervised learning in problems such as the weather-prediction task. The reward prediction error is positive (+) when reward is delivered during training (before an animal is certain the response is correct and the reward will arrive) and is negative (−) when the potential reward does not arrive. When reward arrives, the synaptic plasticity is Hebbian (as in Figure 8.1C) whereas in the absence of reward the response is anti-Hebbian. Other combinations of rule are possible and different versions of Hebbian plasticity are possible (*cf*. Figures 8.1A-D). The impact of these different variants will be tested in Tutorial 8.3. A general requirement for all useful rules is that the change in synaptic strength is either zero or equal to the parity of the three prior columns. That is, if we convert the rates of the presynaptic cell and of the postsynaptic cell to values of "+1" for high rate and "-1" for low rate, then multiply together those two values with the sign of the reward prediction error, the product indicates the required direction of any change in synaptic strength (*i. e.*, a result of "+1" means increase strength and "-1" means decrease strength).

In the following section we will study a simple model of reinforcement learning (Figure 8.11) that can be simulated using circuits of model neurons[54]. We will implement

reward as a binary signal that is either present or absent on any trial, with a probability that depends on the stimuli that are present and the action that is chosen. The underlying plasticity rule will depend on three terms: presynaptic activity, representing the stimulus; postsynaptic activity representing the response; and dopamine release, indicating the reinforcement signal (see Table 8.1).

8.7.3. The weather prediction task

The weather prediction task (Figure 8.11) is one that requires the subject to learn how well different stimuli predict the desired response. The task is difficult because stimuli appear together, so that good performance requires the subject to weight the more predictive stimuli more strongly than the less predictive ones, but on any one trial it is not clear which of the presented stimuli were more associated with the reward and so should be reinforced.

The name of the task arises from the need to choose appropriate clothing, or whether to bring an umbrella when heading out, based on a number of pieces of information such as cloud cover, weather forecast, time of year, current temperature, and so on. Ideally, we would always bring an umbrella or raincoat on a day-trip when it rains and avoid the extra baggage on a day-trip without rain. While in the real-world the costs of the two errors are not equal, in laboratory tasks set to mimic this type of decision, a correct choice in either direction is equally rewarded.

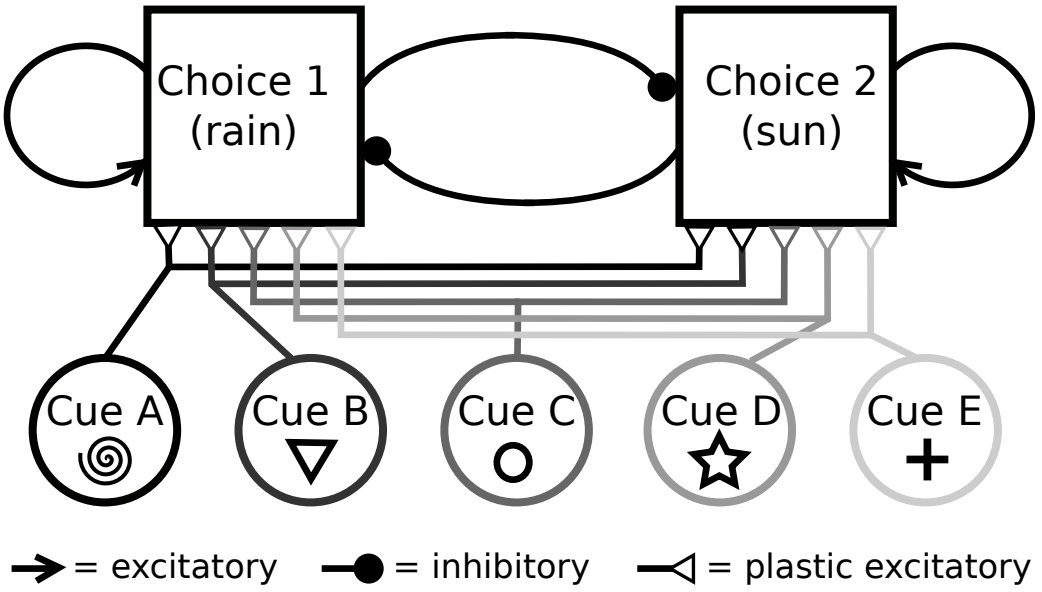

Figure 8.11. The weather-prediction task: learning to respond in an uncertain environment. The task is a two-alternative forced choice task that could be considered as similar to the choice of clothing to wear based on a prediction of the day's weather—often an uncertain affair. In each trial of the task, one of the responses is considered correct and rewarded. Depending on the correct response, each cue appears with a given probability. For example, Cue A appears with 95% probability if choice 1 is correct but only with 5% probability if choice 2 is correct, while Cue E appears with 5% probability if choice 1 is correct but 95% probability if choice 2 is correct. Cue C is uninformative,

appearing with 50% probability, whichever choice is correct. Cues B and D are intermediate, with probabilities of 75% versus 25% and 25% versus 75% respectively if choice 1 versus choice 2 is correct. In the standard version of the task each choice is correct with 50% probability on any trial. The model circuit used to solve the task—*i. e.*, to produce correct choices given the uncertain cues—relies on reward-dependent plasticity at the excitatory synapses between the cue-responsive units and the decision-making units. Initially the noise in the decision-making circuit generates random choices (the exploration stage) but the synaptic plasticity causes biases so that eventually the response is dominated by activity of the most predictive cues.

It turns out that a three-component plasticity rule is sufficient to produce the correct behavior, at least qualitatively[54]. Moreover, the required rule has some evidence in favor of it from the impact of dopamine on synaptic plasticity[55,56]. The basis of the rule is one of reward-dependent plasticity: If a decision is correct it leads to reward and this reward causes (or enhances) Hebbian plasticity between the neurons that led to the reward. In particular, in the circuit of Figure 8.11, if specific stimuli are present and a specific choice is made, then if that choice results in reward the connections from the neurons responsive to those stimuli to the neurons generating that choice should be strengthened. We discuss a few alternative three-component plasticity rules that could lead to successful in the following paragraphs.

Rule 1: In order to prevent chance coincidences from strengthening all connections, a method for reducing synaptic strengths is also necessary. For example, the standard rules for synaptic depression with high presynaptic activity and low postsynaptic activity (Rule 3 of Figure 8.1 and/or Figure 8.2) can be used, if a requirement of a positive reward signal is added. In this case, when the correct choice is made, the synapses being depressed correspond to those promoting the unchosen, incorrect alternative. While such a method has the benefit of counterbalancing potentiation with depression, it has the disadvantage of only producing synaptic change when a correct response is made. The problem with such a plasticity rule kicking in only when a pulse of dopamine arrives following a correct choice, is that the circuit could get 'stuck', always producing incorrect responses and without the ability to alter its behavior. That is, one could not unlearn behaviors that initially, by chance, seemed advantageous but later produced negative outcomes.

Rule 2: Therefore, alternatively, it can be better to ensure that incorrect choices cause a decrease in the likelihood of that stimulus-response combination occurring again. To achieve this, the same synapses that are potentiated when a choice is correct should be depressed in an anti-Hebbian manner, if the same choice made with the same stimuli proves to be incorrect. Table 8.1 indicates such a situation.

Rule 3: Finally, it is worth thinking about the desired plasticity mechanisms from a Bayesian perspective. If a particular stimulus is very rare, but when it appears a particular choice is always correct, then that stimulus is highly predictive of the required choice. Therefore, the synapses connecting neurons responsive to the stimulus with neurons generating the correct choice should remain strong and not be diminished on the many trials when the presynaptic neurons are inactive because the stimulus did not appear. That is, the probability of the response given the stimulus is much higher than the probability of the stimulus given the response when the stimulus is rare. If the strength of the synapse

were to alter when a response occurred (postsynaptic activity) without the stimulus (presynaptic activity) then the final state of the synapse would convey, primarily, information about the probability of the stimulus given the response. However, optimal decision-making requires the probability of the response given the stimulus. This argument suggests Rule 3 of Figure 8.1 is more appropriate than Rules 2 or 4, at least in cases when stimuli are not equally likely—a hypothesis we can test in Tutorial 8.3.

In laboratory tasks the reward arrives immediately after the choice, so that one can assume stimulus-related activity and choice-related activity coincide with reward delivery. This is the assumption behind the model of Tutorial 8.3. However, in reality a reward may arrive after all the predictive activity has dissipated. This leads to a problem of eligibility: Of all the neurons that have been active in the brain prior to the reward, which ones should have their synapses potentiated? It turns out that with a slowly decaying 'eligibility trace' [57], it is possible to solve this problem over time and even for the reward signal to propagate back to arise at the time of predictive stimuli, as seen *in vivo*[58].

8.7.4. Calculations required in the weather prediction task.

In the weather prediction task, a subject must predict which of two responses is the rewarded one, based on which symbols appear beforehand. Each symbol appears on a given trial with a probability that depends on which of the two responses is correct (rewarded) on that trial. Across many trials, the subject can learn the degree to which a symbol is predictive of one response or the other and choose the response that is the most likely to be correct given the observed set of symbols.

The optimal method—in terms of maximizing the number of rewarded responses made— for combining together the information from the presented symbols is for the subject to multiply together the likelihood ratios (or add the log-likelihood ratios) produced by each symbol's presence or absence.

Box 8.22. Likelihood ratio: the ratio of the probabilities of two different outcomes (in this case sunny versus rainy) given the observed data.

Example.

Suppose Choice 1 and Choice 2 are rewarded equally often, and three symbols can appear. Symbol A appears on ¾ of the trials when Choice 1 is rewarded and ¼ of the trials when Choice 2 is rewarded. Symbol B appears on 2/3 of the trials when Choice 1 is rewarded and 1/3 of the trials when Choice 2 is rewarded. Symbol C appears on 1/3 of the trials when Choice 1 is rewarded and 2/3 of the trials when Choice 2 is rewarded. If symbols A and C appear, what is the preferred choice?

We use the notation for probability (see Section 1.4.3) where $\tilde{B}$ means 'NOT B', so that $P(\tilde{B}) = 1 - P(B)$. We want to compare the probability that Choice 1 is correct:

$$P(1|A,\tilde{B},C) = \frac{P(1)P(A,\tilde{B},C|1)}{P(A)P(\tilde{B})P(C)} = \frac{P(1)P(A|1)P(\tilde{B}|1)P(C|1)}{P(A)P(\tilde{B})P(C)} \qquad \text{Eq. 8.11}$$

with the probability that Choice 2 is correct:

$$P(2|A,\tilde{B},C) = \frac{P(2)P(A,\tilde{B},C|2)}{P(A)P(\tilde{B})P(C)} = \frac{P(2)P(A|2)P(\tilde{B}|2)P(C|2)}{P(A)P(\tilde{B})P(C)}. \qquad \text{Eq. 8.12}$$

If $P\left(1\middle|A,\tilde{B},C\right) > P\left(2\middle|A,\tilde{B},C\right)$ the subject should make Choice 1, otherwise the subject should make Choice 2. We make the comparison by testing if the likelihood ratio is greater than 1, *i. e.*, we make Choice 1 if $\frac{P\left(1\middle|A,\tilde{B},C\right)}{P\left(2\middle|A,\tilde{B},C\right)} > 1$. This likelihood ratio is given by the ratio of the priors, $\frac{P(1)}{P(2)} = 1$ multiplied by the individual, stimulus-specific likelihood ratios:

$$\begin{aligned}\frac{P\left(1\middle|A,\tilde{B},C\right)}{P\left(2\middle|A,\tilde{B},C\right)} &= \frac{P(1)}{P(2)} \cdot \frac{P(A|1)}{P(A|2)} \cdot \frac{P\left(\tilde{B}\middle|1\right)}{P\left(\tilde{B}\middle|2\right)} \cdot \frac{P(C|1)}{P(C|2)} \\ &= (1)\left(\frac{3}{1}\right)\left(\frac{1}{2}\right)\left(\frac{1}{2}\right) \\ &= \frac{3}{4}.\end{aligned} \qquad \text{Eq. 8.13}$$

So, since $\frac{P\left(1\middle|A,\tilde{B},C\right)}{P\left(2\middle|A,\tilde{B},C\right)} < 1$ the subject should make Choice 2. Notice that the absence of Symbol B provided information that swung the balance in favor of Choice 2 from Choice 1, so the absence of a common stimulus also provides information that should impact our decisions.

In Tutorial 8.3 we will assess how well a neural circuit can achieve a similar result through reward-dependent plasticity.

8.8. Tutorial 8.3. Learning the weather-prediction task in a neural circuit.

Neuroscience goals: learn how a single reinforcement signal can lead to changes in specific connections in a manner conducive to the learning of tasks.
Computational goals: keep track of connections in different types of connectivity matrix.

In this tutorial, you will produce a firing-rate model network comprising 5 input units (labeled A-E) that provide feedforward input to 2 units in a decision-making circuit (labeled Choice 1 and 2, Figure 8.11). You will simulate 800 trials. On each trial the weather is sunny with probability $P(sun)$, in which case Choice 2 is correct (so $P(2) = P(sun)$) otherwise it is rainy, in which case Choice 1 is correct (so $P(1) = 1 - P(sun)$). During a trial the input units receive applied current if the corresponding cue is present, otherwise they receive no applied current. If it is sunny the cue corresponding to input units A, B, C, D, and E, is present with respective probabilities, $P(A|2) = 0.95$, $P(B|2) = 0.75$, $P(C|2) = 0.5$, $P(D|2) = 0.25$, and $P(E|2) = 0.05$. Whereas if it is rainy the cue corresponding to input units A, B, C, D, and E, is present with respective probabilities, $P(A|1) = 0.05$, $P(B|1) = 0.25$, $P(C|1) = 0.5$, $P(D|1) = 0.75$, and $P(E|1) = 0.95$.

Simulate the firing rate of each unit, i, as a sigmoid function of its input current, $I_i(t)$, without noise, as:

$$\tau \frac{dr_i}{dt} = -r_i + \frac{r_{max}}{1 + \exp\left[\frac{I_{th} - I_i(t)}{I_\sigma}\right]},$$

with parameters $\tau = 20$ms, $r_{max} = 100$Hz, $I_{th} = 50$, and $I_\sigma = 5$. Each trial should last 500ms, with a time-step of $\Delta t = 0.001$ sec.

The input current, $I_i(t)$, for the input units whose corresponding cue is present on a trial should be set as $I_i(t) = 50$ for $t > 0.1$ sec (otherwise the input current, $I_i(t)$, for the input units is zero.

The input current to the decision-making units depends on the rates of the input units, through a 5×2 matrix of connections, $W_{ji}^{(In)}$, and the rates of the decision-making units through a 2×2 matrix of recurrent connections, $W_{ji}^{(Rec)}$, plus an additional noise term, $I^{(rnd)}(t)$:

$$I_i(t) = W_{Ai}^{(In)} r_A + W_{Bi}^{(In)} r_B + W_{Ci}^{(In)} r_C + W_{Di}^{(In)} r_D + W_{Ei}^{(In)} r_E + W_{1i}^{(Rec)} r_1 + W_{2i}^{(Rec)} r_2 + I^{(rnd)}(t)$$

Initialize all of the connections from input units to decision-making units, $W_{ji}^{(In)}$, to a value of $W_0 = 0.2$. These will be updated by the plasticity rule. The decision-making circuit has fixed connections with self-excitation $W_{11}^{(Rec)} = W_{22}^{(Rec)} = 0.5$ and cross-inhibition, $W_{12}^{(Rec)} = W_{21}^{(Rec)} = -0.5$. The noise current is generated independently for each decision-making unit and independently on each time-step, by repeatedly sampling from a Gaussian distribution of unit standard deviation (`randn()` in Matlab) and multiplying by a scaling factor of $\sigma_I / \sqrt{\Delta t}$ where $\sigma_I = 1$. Each separate trial should have independent noise (*i. e.*, regenerate the noise on every trial).

Use a threshold of 40Hz and determine which of the two decision-making units reaches that rate first (typically only one of the units achieves a high rate)—that unit represents the choice made, (unit 1 = rain, unit 2 = sun), so you should add code to determine if the choice is correct, in which case reward is received.

The synapses between input units and decision units will be updated at the end of each trial according to a three-component rule (Table 8.1): a) low/high rate of input unit; b) low/high rate of decision-making unit; and c) correct/incorrect response.

One goal of this tutorial is to compare and contrast the success (in terms of achieving a high rate of correct responses) and stability (in terms of connection strengths remaining within a limited range) of different rules in different conditions. To achieve this, for each simulation calculate the cumulative number of correct responses and the number correct in the final 100 trials. Plot the synaptic strengths as a function of trial number. Compare the behavior of each rule when connection strengths are unbounded versus bounded such that $0 \leq W_{ij} \leq 2W_0$.

For each rule plot the difference in synaptic strength to the two pools as a function of the log likelihood ratio due to a given input unit's activity. For example, for input unit A, you would plot $W_{A2}^{(In)} - W_{A1}^{(In)}$ against $\log\left[\frac{P(A|2)}{P(A|1)}\right]$, where $P(A|2)$ and $P(A|1)$ are the probabilities given in the question, while $W_{A2}^{(In)}$ and $W_{A1}^{(In)}$ are the connection strengths produced by learning. In this manner, for each rule you will obtain 5 data points, one for each input unit. Which rule produces the most linear result? A linear result would indicate that the total difference in input currents to the two decision-making units is proportional to the log-likelihood ratio based on the probability of each unit producing the correct response.

<u>Rule A.</u>

Assume the reward prediction error, E, is $+0.5$ on correct trials and -0.5 on incorrect trials. Update synapses only from input units that were active in the trial (*i. e.*, the corresponding symbol was present).

Update connections from active input units to the active decision-making unit by an amount, which is positive on correct trials and negative on incorrect trials:

$$\Delta W_{ij} = +\varepsilon E$$

and from active input units to the inactive decision-making unit by an amount, which is negative on correct trials and positive on incorrect trials:

$$\Delta W_{ij} = -\varepsilon E,$$

where $\varepsilon = 0.04$ sets the plasticity rate.

Treat a trial without a decision as incorrect, so that ΔW_{ij} is positive from all active input units to both decision-making units.

Rule B.

Follow Rule A, but calculate the reward prediction error as $E = R - \langle R \rangle$, where $R = 1$ if the trial was rewarded and $R = 0$ otherwise, while $\langle R \rangle$ is the mean reward across the prior 10 trials (assume $\langle R \rangle = 0.5$ over the first 10 trials).

Rule C.

Calculate the reward prediction error as in Rule A, but update synapses using soft bounds, so that if the calculated ΔW_{ij} is negative, multiply it by the current value of W_{ij} before reducing the connection strength and if ΔW_{ij} is positive, multiply it by a factor of $W_{max} - W_{ij}$ before increasing the connection strength. Use $W_{max} = 2W_0$. That is, update connections from active input units to the active decision-making unit by an amount:

$$\Delta W_{ij} = +\varepsilon E\left(W_{max} - W_{ij}\right) \quad \text{if } E > 0$$

or

$$\Delta W_{ij} = +\varepsilon E W_{ij} \quad \text{if } E < 0.$$

Similarly, update connections from active input units to the inactive decision-making unit by an amount:

$$\Delta W_{ij} = -\varepsilon E\left(W_{max} - W_{ij}\right) \quad \text{if } E < 0$$

or

$$\Delta W_{ij} = -\varepsilon E W_{ij} \quad \text{if } E > 0.$$

Rule D.

Follow Rule B, but only update synapses to the active decision-making unit from active input pools (as in Figure 8.1A).

Rule E.

Follow Rule D, but also update synapses to the active decision-making unit from inactive input units (as in Figure 8.1D) by an amount:

$$\Delta W_{ij} = -\varepsilon E.$$

PART B. OPTIONAL

Alternative training protocols

In the standard training protocol, $P(sun) = 0.5$ and 800 trials are used.

1) Try $P(sun) = 0.2$. How do the final connection strengths change and does performance change? Which rules cope better with this change? If performance deteriorates try to explain the reason.

2) Increase the number of trials to 4000. For which plasticity rules do the connection strengths reach a steady state? Explain your findings.

CHALLENGE

Generate a rule such that connection strengths reach a stable steady state and the network performs well with both alternative training protocols.

8.9. Eyeblink conditioning

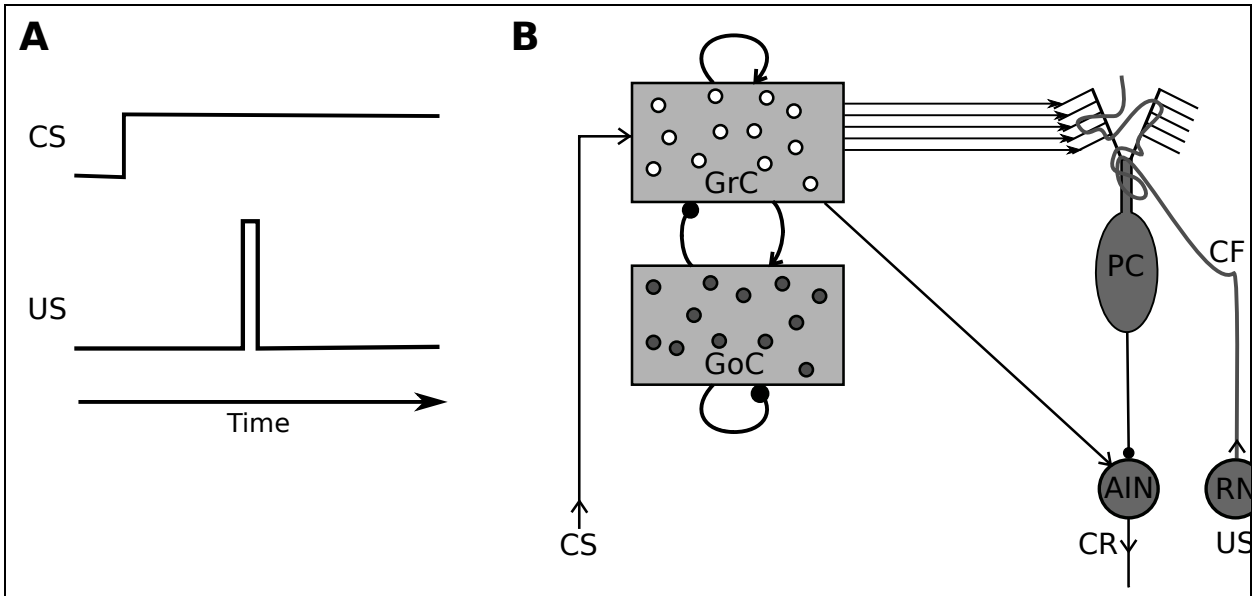

Figure 8.12. Simplified model of cerebellar eyeblink conditioning. A. The conditioned stimulus (CS) is typically a tone, which stays on during a trial. The unconditioned stimulus (US), typically a puff of air into the eye, arrives a fixed time interval later. **B.** In the simplified model circuit, the CS provides input to excitatory Granule Cells (GrC), which are in a recurrent circuit with inhibitory Golgi Cells (GoC), and which provide input to the inhibitory Purkinje Cell (PC) via parallel fibers. The US, as well as producing the automatic response of eyelid closure, also initiates the reinforcement signal for learning. To this end, it generates transient activity in the excitatory cells of the Red Nucleus (RN), whose activity is transmitted via the Climbing Fiber (CF), which winds around the dendrites of the Purkinje Cell. The CF provides a reinforcement signal to the synapses from the Granule Cells to the Purkinje Cell. The output of the Anterior Interpositus Nucleus (AIN) causes the conditioned response, when a temporary pause in the Purkinje cell's activity causes disinhibition, while it is receiving activity from the Granule Cells (model connections from GrC to AIN are not present empirically, but in this simplified model, they are necessary to ensure the AIN can only be active when the CS is on). The model is based on the work of Medina et al [59].

Eyeblink conditioning is a form of classical conditioning in which an animal (often a rabbit) learns to associate a tone (the conditioned stimulus, CS) with an aggravating puff of air aimed at its eyes (the unconditioned stimulus, US). The puff of air reliably follows the tone after a predetermined time interval that can range from 150ms to 1.5s. The puff of air produces an unconditioned response, the eyeblink, which, after a sufficient number of trials (more than 100 for a rabbit!), is produced just before the air-puff arrives. While this task is known as classical conditioning because the final response, the eyeblink, is the same as the innate response to the US, it could arguably be thought of as instrumental conditioning: the animal must learn to make the eyeblink at the correct time, and when it does so it avoids the 'punishment' of an aggravating puff of air in the eye.

The protocol can be one of delay conditioning, in which the tone remains on, or trace conditioning in which the tone switches off to leave a silent pause before the air-puff. The distinction between trace or delay may not seem important, but trace conditioning requires processing in the hippocampus, whereas delay conditioning may be entirely a function of the cerebellum.

In Tutorial 8.4 we will simulate a simplified cerebellar model (Figure 8.12) for delay conditioning, based on models by Buonomano, Medina, and Mauk[59,60], which provide an example of the Marr-Albus-Ito theory of cerebellar function[61-63]. To solve the delay-conditioning task, two requirements must be satisfied. First, some function of the neural activity following the CS must reliably predict the US at the correct time. Therefore, the neural activity must contain information about time. Second, a plasticity rule must allow the circuit to extract that time information in order to produce the response at the correct time. The plasticity rule—at the synapses onto the cerebellar Purkinje cell—achieves the latter by association of activity in a particular subset of granule cells with a reinforcement signal arriving via the climbing fiber (see Figure 8.12).

While the subset of active units in our circuit is strongly dependent on time, so can be used as an indication of time since CS onset, we use a chaotic circuit to produce such a time signal—and as seen in Section 7.8, activity in a chaotic circuit is not robust to changes in initial conditions or to noise fluctuations. Therefore, in our simulations the initial conditions must be identical on each trial and the neural responses should be noise-free, to ensure identical dynamics in every trial. Such a circuit is valuable here as a demonstration of cerebellar learning arising from highly irregular granule cell activity as occurs in the cerebellum. However, we still do not understand the means by which neural circuits robustly represent time, so this important topic is a matter of ongoing theoretical and experimental investigation[64,65].

8.10. Tutorial 8.4. A model of eyeblink conditioning.

In this tutorial, we will use a model network found online as "`Figure_7.13.py`", comprised of a large number of strongly coupled excitatory and inhibitory units that produced the "high-dimensional chaos" shown in Figure 7.13. We will make the false assumption that the complex trajectory of neural activity is reproducible from a given starting point, in order to train the activity pattern of a particular point in time to produce an eye-blink. A model of the circuit we will produce—highly simplified from the actual circuitry in the cerebellum—and a schematic of the task is shown in Figure 8.12.

All neurons will be treated as firing-rate model units with sigmoidal f-I curves following (see Chapter 6, Eqs. 6.1-6.3 for notation):

$$\tau_i \frac{dr_i}{dt} = -r_i + f(I_i)$$

where

$$f(I_i) = \frac{r_i^{max}}{1 + \exp\left[-\left(I_i - I_i^{th}\right)/I_i^{\sigma}\right]},$$

and I_i is the input current to the cell-group. Other parameters of the f-I curve depend on the cell-type in the circuit as described in Table 8.2.

Cell	Symbol	τ_i (sec)	r_i^{max} (Hz)	I_i^{th}	I_i^{σ}
Granule Cell (E)	$r_i = r_{GrC}$	0.010	100	10	1
Golgi Cell (I)	$r_i = r_{GoC}$	0.005	100	10	1
Purkinje Cell	$r_i = r_{PC}$	0.010	30	40	10
Climbing Fiber	$r_i = r_{CF}$	0.002	30	15	1
Anterior Interpositus Nucleus	$r_i = r_{AIN}$	0.010	100	25	4

Table 8.2. Properties of different cell types in the cerebellar eye-conditioning model of Tutorial 8.4.

The granule and Golgi cells will be simulated in the chaotic network available on line (see next paragraph). For other cells in this model, their input currents are:

$$I_{PC}(t) = \sum_n W_n^{GP} r_n^{GrC}$$
$$I_{CF}(t) = I_{US}(t)$$
$$I_{AIN}(t) = W^{GN} \sum_n r_n^{GrC} - W^{PN} r_{PC},$$

where r_n^{GrC} is the firing rate of the granule cell labeled n, which has a unique connection strength, W_n^{GP}, to the Purkinje cell. Initially set $W_n^{GP} = \frac{10}{N_{GrC}} = 0.05$ for all such connections, though these strengths will change during learning. Also set $W^{GN} = \frac{5}{N_{GrC}} = 0.025$ and $W^{PN} = 30$. $N_{GrC} = 200$ is the number of granule cells.

You can use the code `Figure_7.13.py` to implement the balanced network of 200 excitatory granule cells coupled to 200 inhibitory Golgi cells. In this combined inhibitory-excitatory network, each connection is present with a fixed probability, $p^{(conn)} = 0.05$. Connection strengths are drawn randomly from a uniform distribution then inputs are scaled, so the sum of excitatory connection strengths to each granule cell is 90, as is the sum of inhibitory connection strengths to each granule cell, whereas the sum of excitatory connection strengths to each Golgi cells is 80, as is the sum of inhibitory connection strengths to each Golgi cell. In addition to the input currents to these cells already generated in the network, you should include an additional input of $I_{CS}(t)$ to each granule cell, which will initiate activity in the network at the time of the CS.

In each trial, of duration 2s, you should initialize the firing rates of all cells at zero. At a time of 200ms, the conditioned stimulus will switch on, to a value of $I_{CS}(t) = 2$ for $t >$

200ms. At a later time, t_{US}, the unconditioned stimulus will be switched on for 20ms, to a value of $I_{US}(t) = 100$, for $t_{US} \leq t \leq t_{US} + 25$ms. At all other times these inputs are fixed at zero.

Learning will occur via plasticity of the connection strengths, W_n^{GP}, between granule cells and Purkinje cell. The plasticity depends on the overlap between the synaptic gating variable of the synapse and the activity of the climbing fiber. Each synaptic gating variable follows the firing rate of the corresponding granule cell, according to:

$$\tau_s \frac{dS_n}{dt} = -S_n + \frac{r_n^{GrC}}{r_{GrC}^{max}},$$

where $\tau_s = 50$ms (and the factor $r_{GrC}^{max} = 100$Hz ensures $0 \leq S_n \leq 1$).

The learning rule is then

$$W_n^{GP} \rightarrow W_n^{GP}\left\{1 - \Delta\left[\int S_n(t) r_{CF}(t) dt - \Theta\right]\right\},$$

where you can evaluate the integral by summing the product of $S_n(t)$ and $r_{CF}(t)$ across all time bins then multiplying by the time-step in the simulation. If this integral is larger than the threshold, $\Theta = 0.1$, it indicates the climbing fiber fires when there is synaptic input from that particular granule cell and the corresponding synapse depresses. Otherwise the synapse increases in strength (as is necessary to prevent all synapses from eventually depressing to zero). The factor $\Delta = 0.1$ sets the rate of change of synaptic strength.

Update the synaptic strengths at the end of each 2-second trial using the above method, then constrain the strengths to be no less than zero and no greater than a maximum value of $\frac{50}{N_{GrC}} = 0.25$.

1) Train the system for 200 trials with the unconditioned stimulus arriving 800ms after the conditioned stimulus (*i. e.*, $t_{US} = 1$ sec, when $t_{CS} = 200$ ms). Plot the total input current to the Purkinje cell, the firing rate of the Purkinje cell, and the firing rate of the interpositus nucleus before and after training. (It can be useful to plot these variables every 20 trials, overwriting the figure each time, to assess progress through the simulation).
Note: since there is no added noise in this simulation, you can speed it up by increasing the value of Δ and correspondingly reducing the number of trials if the time taken is too long. A time-step of 0.2ms should be sufficiently fine for this simulation of firing-rate model neurons.

2) Repeat, but with $t_{US} = 500$ ms.

3) Repeat, but with $t_{US} = 1500$ ms. Is there a limit to the range of values of t_{US} that will produce a correctly timed conditioned response?

4) How do you expect your results to change (if at all) if the synaptic time constant for granule cells to the Purkinje cell is reduced from 50ms to 2ms? Explain any changes both in terms of the function of the circuit and how it would impact the animal's behavior.

5) The granule cells of the cerebellum are the most numerous cells of a given class in the brain. Do you expect the results to be more reliable or less reliable if the circuit contained more granule cells and Golgi cells? Test your expectation if time allows. Care is needed when reducing the number of cells to ensure each cell receives some inhibitory input and

some excitatory input—either increase connection probability, or sample networks until all cells receive connections, or use another method for ensuring each neuron receives both excitatory and inhibitory connections.

Final notes. In this tutorial, we have simulated a circuit that can achieve a conditioned response. The circuit has many similarities with the circuitry of the cerebellum that is responsible for eyeblink conditioning, but it is important to note some differences.

1) In reality, the Purkinje cell is active (due to other inputs) prior to the conditioned stimulus.

2) There are no direct connections from granule cells to the AIN, though AIN does receive input when the conditioned stimulus is present.

3) We have omitted a parallel path, which may be dominant, in which other Purkinje cells increase their firing rate at the time of the unconditioned stimulus over the course of learning. The timed output of these cells inhibits a parallel pathway that, when active, prevents eyelid closure.

4) The tiniest bit of noise or change in initial conditions would destroy the reliable alignment of activity in a specific subset of granule cells with the unconditioned stimulus. Such easy disruption of information means the network we have produced would be ineffective in practice. How networks can have chaotic properties but respond reliably and robustly to certain stimuli (such as the conditioned stimulus) is a question of ongoing, active research[64].

Questions for Chapter 8.

1) How does standard STDP (Eqs 8.3-8.4) fully implement Hebb's postulate and what does STDP include that is not mentioned in Hebb's postulate?

2) A type of synapse is found which increases in strength when the postsynaptic cell fires before the presynaptic cell, and decreases in strength when the cell's fire in reverse order. Exponential fits of $-0.02e^{-\Delta t/10}$ and $+0.015e^{-\Delta t/15}$ are made respectively for the fractional amount of depression and potentiation when any pair of presynaptic and postsynaptic spikes are separated by an absolute time difference of Δt measured in ms.

a) Will the synapse strengthen or depress when the two neurons spike in an uncorrelated manner? (Assume all possible spike-pairs contribute to plasticity.)

b) How will the change in synaptic strength impact the correlation between spikes from the two neurons?

c) Do you expect the rule to lead to a stable non-zero solution of synaptic strength?

3) Two homeostatic processes are operating on a neuron simultaneously, one changing its synaptic strength, the other its threshold. The threshold rule responds linearly to firing rate with a set point of 5Hz.

a) If the synaptic rule responds linearly to the neuron's firing rate, what can its set point or range of set points be for a stable solution to exist?

b) If the synaptic rule responds quadratically to the neuron's firing rate, what can its set point or range of set points be for a stable solution to exist?

4) Explain the benefit of a signal for positive or negative reinforcement in supervised learning being proportional to the reward prediction error, rather than being just proportional to the reward itself.

8.11. Appendix A: Rate-dependent plasticity via STDP between uncorrelated Poisson spike trains.

Here we will consider successively the two distinct types of standard rule for STDP, first using all spike pairs and second only using the most recent spike of the other neuron during each update (*i. e.*, nearest-neighbor pairs).

We assume both the presynaptic and the postsynaptic neuron produce spikes as a Poisson process with fixed rates (r_i for the presynaptic cell and r_j for the postsynaptic cell). All spike times are uncorrelated with each other. The probability of a neuron producing a spike in a time interval δt is constant for each neuron and proportional to its firing rate ($r_i \delta t$ for the presynaptic cell and $r_j \delta t$ for the postsynaptic cell). We will use the result (Section 3.7) that for a neuron emitting spikes as a Poisson process, the probability, $P(T_D)\delta T_D$ that the time since the most recent spike is in a given time interval, T_D to $T_D + \delta T_D$ is exponential: $PT_D \delta T_D = r\exp(-rT_D)\delta T_D$.

1) Plasticity due to all pairs of spikes.

If all pairs of spikes contribute to plasticity, then for each postsynaptic spike the average amount of LTP is calculated as the sum over all time-differences, T_D, of the probability of a presynaptic spike occurring that amount earlier ($r_i \delta T_D$) multiplied by the amount of LTP generated by a presynaptic spike at that amount earlier ($A_+ \exp[-T_D/\tau_+]$). Mathematically the sum over all time-differences becomes an integral over Δt, so the expected (mean) amount of LTP per postsynaptic spike is:

$$\int_0^\infty A_+ \exp[-T_D/\tau_+]\, r_i dT_D = A_+ \tau_+ r_i. \qquad \text{Eq. 8.14}$$

Since the rate of postsynaptic spikes is r_j the mean rate of change of synaptic strength due to the LTP window is $A_+ \tau_+ r_i r_j$.

A similar argument based on the amount of LTD per presynaptic spike leads to the mean rate of synaptic change due to the LTD window as $-A_- \tau_- r_i r_j$.

Thus, the net rate of plasticity is proportional to the rate of production of spike-pairs, which is the product of the two firing rates, $r_i r_j$, and the mean level of potentiation or depression per spike pair depends on the total area of the STDP curve, $A_+ \tau_+ - A_- \tau_-$.

2) Plasticity due to only the most recent spike.

In the second formulation, whenever the postsynaptic neuron emits a spike, the amount of LTP is $A_+ \exp[-T_D/\tau_+]$ where T_D is the time since the most recent presynaptic spike. Now, to obtain the expected (mean) amount of LTP per postsynaptic spike we multiply the amount of LTP for each possible time-interval by the probability of that time-interval—which is $r_i \exp(-r_i T_D) dT_D$ for a Poisson process—then sum over all possible intervals. Or, using calculus to convert the summation to an integral we have:

$$\int_0^\infty A_+ \exp[-T_D/\tau_+]\, r_i \exp(-rT_D) dT_D = \frac{A_+ \tau_+ r_i}{1 + \tau_+ r_i}. \qquad \text{Eq. 8.15}$$

Since the rate of postsynaptic spikes is r_j the mean rate of change of synaptic strength due to the LTP window is $A_+\tau_+ r_i r_j/(1+\tau_+ r_i)$. Notice that this is like the result when all spikes are included, but includes the denominator of $(1+\tau_+ r_i)$, which becomes important if the presynaptic rate is not much less than $1/\tau_+$. Similarly, the mean rate of change of synaptic strength due to the LTD window is $-A_-\tau_- r_i r_j/(1+\tau_- r_j)$ so that the net expected rate of change of synaptic strength is:

$$\frac{dW_{ij}}{dt} = r_i r_j \left(\frac{A_+\tau_+}{1+\tau_+ r_i} - \frac{A_-\tau_-}{1+\tau_- r_j} \right). \qquad \text{Eq. 8.16}$$

Interestingly, if we keep the presynaptic rate fixed and vary the postsynaptic rate, r_j, the resulting dependence is biphasic, with increasing net depression at low postsynaptic rates switching to potentiation at high postsynaptic rates (Figure 8.13). However, unlike the observed data, depression is favored at high presynaptic rates and potentiation is favored at low presynaptic rates.

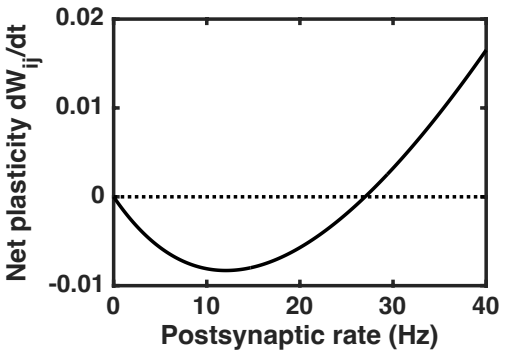

Figure 8.13. Rate-dependence of STDP with nearest-neighbor rule. A plot of $\frac{dW_{ij}}{dt}$ using Eq. 8.16, using parameters $A_+ = 0.01$, $A_- = 0.011$, $\tau_+ = \tau_- = 20$ms, $r_i = 20$Hz, and variable postsynaptic rate, r_j. Notice the similarity in shape with Figure 8.2. This figure was produced by the online code `STDP_Poisson_nn.m`.

8.12. Appendix B: Rate-dependence of triplet STDP between uncorrelated Poisson spike trains

Here we assume plasticity is due to all pairs of spikes. Since the model of depression is unchanged, the rate of change of synaptic strength due to depression is identical to that of Appendix A, Part 1, and is $-A_-\tau_- r_i r_j$ (for symbol meanings see Appendix A).

For the potentiation term, which we add whenever there is a postsynaptic spike (at time, t_j), we must include all possible times for a preceding postsynaptic spike (at time, $t_{j'}$) as well as all possible times for a preceding presynaptic spike (at time, t_i). The summation over all possible times of preceding pre- and post-synaptic spikes requires evaluation of a double integral of the update rule for each triplet of spikes (Eq. 8.6):

$$\Delta W_{ij} = A_3^+ \int_{-\infty}^{t_j} \exp\left(\frac{t_i - t_j}{\tau_+}\right) r_i dt_i \int_{-\infty}^{t_j} \exp\left(\frac{t_{j'} - t_j}{\tau_3}\right) r_j dt_{j'}. \qquad \text{Eq. 8.17}$$

In Eq. 8.17, we have substituted $r_i dt_i$ for the fixed probability of a presynaptic spike in the time interval t_i to $t_i + dt_i$, and $r_j dt_{j'}$ for the fixed probability of a postsynaptic spike in the time interval $t_{j'}$ to $t_{j'} + dt_{j'}$. The integrals can be evaluated to give $\Delta W_{ij} = A_3^+ r_i r_j \tau_+ \tau_3$. Since this is the average increment in synaptic strength per postsynaptic spike, the rate of change of synaptic strength due to potentiation gains an extra factor of the postsynaptic firing rate, r_j, so is $A_3^+ r_i r_j^2 \tau_+ \tau_3$.

Combining the rates of potentiation and depression we see that the net rate of change of synaptic strength is (on average):

$$\frac{dW_{ij}}{dt} = r_i r_j \left(A_3^+ \tau_+ \tau_3 r_j - A_- \tau_- \right). \qquad \text{Eq. 8.18}$$

It can be seen that this term is quadratic in the postsynaptic rate, r_j, and switches from net depression at low postsynaptic rates ($r_j < \frac{A_- \tau_-}{A_3^+ \tau_+ \tau_3}$) to potentiation at high rates ($r_j > \frac{A_- \tau_-}{A_3^+ \tau_+ \tau_3}$). Therefore, Eq. 8.18 produces a curve that is identical to the one proposed for firing-rate models, shown in Figure 8.2.

References for Chapter 8

1. Hebb DO. *Organization of behavior*. New York: Wiley; 1949.
2. Lømo T. Frequency potentiation of excitatory synaptic activity in the dentate area of the hippocampal formation. *Acta Physiol Scand*. 1966;68(suppl. 277):128.
3. Bliss TV, Lømo T. Plasticity in a monosynaptic cortical pathway. *J Physiol*. Apr 1970;207(2):61P.
4. Markram H, Lubke J, Frotscher M, Sakmann B. Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. *Science*. Jan 10 1997;275(5297):213-5.
5. Bi GQ, Poo MM. Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type. *J Neurosci*. Dec 15 1998;18(24):10464-72.
6. Teyler TJ, Cavus I, Coussens C, et al. Multideterminant role of calcium in hippocampal synaptic plasticity. *Hippocampus*. Dec 1994;4(6):623-34. doi:10.1002/hipo.450040602
7. Kirkwood A, Bear MF. Elementary forms of synaptic plasticity in the visual cortex. *Biol Res*. 1995;28(1):73-80.
8. Finch EA, Tanaka K, Augustine GJ. Calcium as a trigger for cerebellar long-term synaptic depression. *Cerebellum*. Sep 2012;11(3):706-17. doi:10.1007/s12311-011-0314-x
9. Cavazzini M, Bliss T, Emptage N. Ca2+ and synaptic plasticity. *Cell Calcium*. Sep-Oct 2005;38(3-4):355-67. doi:10.1016/j.ceca.2005.06.013
10. Blair HT, Schafe GE, Bauer EP, Rodrigues SM, LeDoux JE. Synaptic plasticity in the lateral amygdala: a cellular hypothesis of fear conditioning. *Learn Mem*. Sep-Oct 2001;8(5):229-42. doi:10.1101/lm.30901
11. Debanne D. Associative synaptic plasticity in hippocampus and visual cortex: cellular mechanisms and functional implications. *Rev Neurosci*. Jan-Mar 1996;7(1):29-46.
12. Baker KD, Edwards TM, Rickard NS. The role of intracellular calcium stores in synaptic plasticity and memory consolidation. *Neurosci Biobehav Rev*. Aug 2013;37(7):1211-39. doi:10.1016/j.neubiorev.2013.04.011
13. Brown TH, Chapman PF, Kairiss EW, Keenan CL. Long-term synaptic potentiation. *Science*. Nov 04 1988;242(4879):724-8.
14. Artola A, Brocher S, Singer W. Different voltage-dependent thresholds for inducing long-term depression and long-term potentiation in slices of rat visual cortex. *Nature*. 1990;347:69-72.
15. Collingridge GL, Peineau S, Howland JG, Wang YT. Long-term depression in the CNS. *Nat Rev Neurosci*. Jul 2010;11(7):459-73. doi:10.1038/nrn2867
16. Artola A, Singer W. Long-term depression of excitatory synaptic transmission and its relationship to long-term potentiation. *TINS*. 1993;16:480.
17. Kirkwood A, Bear MF. Homosynaptic long-term depression in the visual cortex. *J Neurosci*. May 1994;14(5 Pt 2):3404-12.
18. Bienenstock EL, Cooper LN, Munro PW. Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex. *J Neurosci*. Jan 1982;2(1):32-48.
19. Willshaw D, Buneman OP, Longuet-Higgins H. Non-holographic associative memory. *Nature*. 1969;222:960-962.
20. Hopfield JJ. Neural networks and physical systems with emergent collective computational abilities. *Proc Natl Acad Sci USA*. 1982;79:2554-2558.

21. Feldman DE. The spike-timing dependence of plasticity. *Neuron*. Aug 23 2012;75(4):556-71. doi:10.1016/j.neuron.2012.08.001
22. Nelson SB, Sjostrom PJ, Turrigiano GG. Rate and timing in cortical synaptic plasticity. *Philos Trans R Soc Lond B Biol Sci*. Dec 29 2002;357(1428):1851-7. doi:10.1098/rstb.2002.1162
23. Tsodyks M. Spike-timing-dependent synaptic plasticity - the long road towards understanding neuronal mechanisms of learning and memory. *Trends Neurosci*. Dec 2002;25(12):599-600.
24. Song S, Miller KD, Abbott LF. Competitive Hebbian learning through spike-time-dependent synaptic plasticity. *Nat Neurosci*. 2000;3:919-926.
25. Goodhill GJ, Lowel S. Theory meets experiment: correlated neural activity helps determine ocular dominance column periodicity. *Trends Neurosci*. Oct 1995;18(10):437-9.
26. Huberman AD. Mechanisms of eye-specific visual circuit development. *Curr Opin Neurobiol*. Feb 2007;17(1):73-80. doi:10.1016/j.conb.2007.01.005
27. Drew PJ, Abbott LF. Extending the effects of spike-timing-dependent plasticity to behavioral timescales. *Proc Natl Acad Sci U S A*. Jun 6 2006;103(23):8876-81. doi:0600676103 [pii]
10.1073/pnas.0600676103
28. Nowotny T, Rabinovich MI, Abarbanel HD. Spatial representation of temporal information through spike-timing-dependent plasticity. Research Support, U.S. Gov't, Non-P.H.S.
Research Support, U.S. Gov't, P.H.S. *Physical review E, Statistical, nonlinear, and soft matter physics*. Jul 2003;68(1 Pt 1):011908.
29. Mehta MR, Barnes CA, McNaughton BL. Experience-dependent, asymmetric expansion of hippocampal place fields. *Proc Natl Acad Sci U S A*. Aug 5 1997;94(16):8918-21.
30. Blum KI, Abbott LF. A model of spatial map formation in the hippocampus of the rat. *Neural Comput*. Jan 1996;8(1):85-93.
31. Gerstner W, Abbott LF. Learning navigational maps through potentiation and modulation of hippocampal place cells. *J Comput Neurosci*. Jan 1997;4(1):79-94.
32. Levy WB. A sequence predicting CA3 is a flexible associator that learns and uses context to solve hippocampal-like tasks. *Hippocampus*. 1996;6:579-590.
33. Doupe AJ, Solis MM, Kimpo R, Boettiger CA. Cellular, circuit, and synaptic mechanisms in song learning. *Ann N Y Acad Sci*. Jun 2004;1016:495-523. doi:10.1196/annals.1298.035
34. Troyer TW, Doupe AJ. An associational model of birdsong sensorimotor learning II. Temporal hierarchies and the learning of song sequence. *J Neurophysiol*. Sep 2000;84(3):1224-39.
35. Warren TL, Charlesworth JD, Tumer EC, Brainard MS. Variable sequencing is actively maintained in a well learned motor skill. *J Neurosci*. Oct 31 2012;32(44):15414-25. doi:10.1523/JNEUROSCI.1254-12.2012
36. Brainard MS, Doupe AJ. What songbirds teach us about learning. *Nature*. May 16 2002;417(6886):351-8. doi:10.1038/417351a
37. Sjöström PJ, Nelson SB. Spike timing, calcium signals and synaptic plasticity. *Curr Opin Neurobiol*. Jun 2002;12(3):305-14. doi:S0959438802003252 [pii]

38. Froemke RC, Dan Y. Spike-timing-dependent synaptic modification induced by natural spike trains. *Nature*. Mar 28 2002;416(6879):433-8. doi:10.1038/416433a
39. Sjöström PJ, Turrigiano GG, Nelson SB. Rate, timing, and cooperativity jointly determine cortical synaptic plasticity. *Neuron*. Dec 20 2001;32(6):1149-64. doi:S0896-6273(01)00542-6 [pii]
40. Pfister JP, Gerstner W. Triplets of spikes in a model of spike timing-dependent plasticity. *J Neurosci*. Sep 20 2006;26(38):9673-82. doi:26/38/9673 [pii] 10.1523/JNEUROSCI.1425-06.2006
41. Clopath C, Busing L, Vasilaki E, Gerstner W. Connectivity reflects coding: a model of voltage-based STDP with homeostasis. *Nat Neurosci*. Mar 2010;13(3):344-52. doi:nn.2479 [pii] 10.1038/nn.2479
42. Brandalise F, Carta S, Helmchen F, Lisman J, Gerber U. Dendritic NMDA spikes are necessary for timing-dependent associative LTP in CA3 pyramidal cells. *Nat Commun*. Nov 16 2016;7:13480. doi:10.1038/ncomms13480
43. Song S, Abbott LF. Cortical development and remapping through spike timing-dependent plasticity. *Neuron*. Oct 25 2001;32(2):339-50. doi:S0896-6273(01)00451-2 [pii]
44. O'Leary T, Williams AH, Caplan JS, Marder E. Correlations in ion channel expression emerge from homeostatic tuning rules. *Proc Natl Acad Sci U S A*. Jul 9 2013;110(28):E2645-54. doi:10.1073/pnas.1309966110
45. Turrigiano GG, Nelson SB. Hebb and homeostasis in neuronal plasticity. *Curr Opin Neurobiol*. Jun 2000;10(3):358-64.
46. Turrigiano GG, Nelson SB. Homeostatic plasticity in the developing nervous system. *Nat Rev Neurosci*. Feb 2004;5(2):97-107. doi:10.1038/nrn1327
47. Turrigiano GG. Homeostatic plasticity in neuronal networks: the more things change, the more they stay the same. *Trends Neurosci*. May 1999;22(5):221-7.
48. Turrigiano G. Homeostatic synaptic plasticity: local and global mechanisms for stabilizing neuronal function. *Cold Spring Harb Perspect Biol*. Jan 2012;4(1):a005736. doi:10.1101/cshperspect.a005736
49. Bucher D, Prinz AA, Marder E. Animal-to-animal variability in motor pattern production in adults and during growth. *J Neurosci*. Feb 16 2005;25(7):1611-9.
50. Barto AG. Reinforcement learning control. Review. *Current Opinion in Neurobiology*. Dec 1994;4(6):888-93.
51. Schultz W. Behavioral dopamine signals. *Trends Neurosci*. May 2007;30(5):203-10. doi:S0166-2236(07)00068-9 [pii] 10.1016/j.tins.2007.03.007
52. Schultz W. Dopamine reward prediction-error signalling: a two-component response. *Nat Rev Neurosci*. Mar 2016;17(3):183-95. doi:10.1038/nrn.2015.26
53. Watabe-Uchida M, Eshel N, Uchida N. Neural Circuitry of Reward Prediction Error. *Annu Rev Neurosci*. Apr 24 2017;doi:10.1146/annurev-neuro-072116-031109
54. Soltani A, Wang XJ. Synaptic computation underlying probabilistic inference. *Nat Neurosci*. Jan 2010;13(1):112-9. doi:nn.2450 [pii] 10.1038/nn.2450
55. Gurney KN, Humphries MD, Redgrave P. A new framework for cortico-striatal plasticity: behavioural theory meets in vitro data at the reinforcement-action interface. *PLoS Biol*. Jan 2015;13(1):e1002034. doi:10.1371/journal.pbio.1002034

56. Morris G, Schmidt R, Bergman H. Striatal action-learning based on dopamine concentration. *Exp Brain Res*. Jan 2010;200(3-4):307-17. doi:10.1007/s00221-009-2060-6
57. Izhikevich EM. Solving the distal reward problem through linkage of STDP and dopamine signaling. *Cereb Cortex*. Oct 2007;17(10):2443-52. doi:bhl152 [pii] 10.1093/cercor/bhl152
58. Schultz W. Dopamine neurons and their role in reward mechanisms. *Curr Opin Neurobiol*. Apr 1997;7(2):191-7. doi:S0959-4388(97)80007-4 [pii]
59. Medina JF, Garcia KS, Nores WL, Taylor NM, Mauk MD. Timing mechanisms in the cerebellum: testing predictions of a large-scale computer simulation. *J Neurosci*. 2000;20(14):5516-25.
60. Buonomano DV, Mauk MD. Neural network model of the cerebellum: Temporal discrimination and the timing of motor responses. *Neural Computation*. 1994;6:38-55.
61. Marr D. A theory of cerebellar cortex. *J Physiol*. Jun 1969;202(2):437-70.
62. Albus JS. A theory of cerebellar function. *Mathematical Bioscience*. 1971;10:25-61.
63. Ito M. Long-term depression. *Annu Rev Neurosci*. 1989;12:85-102.
64. Laje R, Buonomano DV. Robust timing and motor patterns by taming chaos in recurrent neural networks. *Nat Neurosci*. Jul 2013;16(7):925-33. doi:10.1038/nn.3405
65. Merchant H, Harrington DL, Meck WH. Neural basis of the perception and estimation of time. *Annu Rev Neurosci*. Jul 08 2013;36:313-36. doi:10.1146/annurev-neuro-062012-170349