

Suzanna Becker

Department of Psychology

McMaster University

1280 Main Street West

Hamilton, ON, L8S 4K1

becker@mcmaster.ca

Computational models are having an increasing impact on neuroscience, by shedding light on the neuronal mechanisms underlying information processing in the brain. In this chapter, we review the contribution of computational models to our understanding of how the brain represents and processes information at three broad levels: 1) sensory coding and perceptual processing, 2) high-level memory systems, 3) representations that guide actions. So far, computational models have had the greatest impact at the earliest stages of information processing, by modeling the brain as a communication channel and applying concepts from information theory. Generally, these models assume that the goal of sensory coding is to map the high-dimensional sensory signal into a (usually lower-dimensional) code that is optimal with respect to some measure of information transmission. Four information-theoretic coding principles will be considered here: 1) Linsker's Infomax principle, 2) Barlow's redundancy reduction principle, 3) Becker and Hinton's Imax principle, and 4) Risannen's Minimum Description Length (MDL) principle. Each of these principles can be used to derive unsupervised learning rules, and has been applied to model multiple levels of cortical organization. Moving beyond perceptual processing to high-level memory processes, the hippocampal system in the medial temporal lobe (MTL) is a key structure for representing complex configurations or episodes in long-term memory. In the hippocampal region, the brain may use very different optimization principles aimed at the memorization of complex events or spatio-temporal episodes, and subsequent reconstruction of details of these episodic memories. Here, rather than recoding the incoming signals in a way that abstracts away unnecessary details, the goal is to memorize the incoming signal as accurately as possible in a single learning trial. Most efforts

at understanding hippocampal function through computational modeling have focused on sub-regions within the hippocampal circuit such as the CA3 or CA1 regions, using “off-the-shelf” learning algorithms such as competitive learning or Hebbian pattern association. More recently, Becker proposed a global optimization principle for learning within this brain region. Based on the goal of accurate input reconstruction, combined with neuroanatomical constraints, this leads to simple, biologically plausible learning rules for all regions within the hippocampal circuit. The model exhibits the key features of an episodic memory system: the capacity to store a large number of distinct, complex episodes, and to recall a complete episode from a minimal cue, and associate items across time, under extremely high plasticity conditions. Finally, moving beyond the static representation of information, we must consider the brain not simply a passive recipient of information, but as a complex, dynamical system, with internal goals, and the ability to select actions based on environmental feedback. Ultimately, models based on the broad goals of prediction and control, using reinforcement-driven learning algorithms, may be the best candidates for characterizing the representations that guide motor actions. Several examples of models are described that begin to address the problem of how we learn representations that can guide our actions in a complex environment.

2.1 Introduction

How does the brain process, represent and act on sensory signals? Through the use of computational models, we are beginning to understand how neural circuits perform these remarkably complex information processing tasks. Psychological and neurobiological studies have identified at least three distinct long-term memory systems in the brain: 1) the perceptual/semantic memory system in the neocortex learns gradually to represent the salient features of the environment; 2) The episodic memory system in the medial temporal lobe learns rapidly to encode complex events, rich in detail, characterizing a particular episode in a particular place and time; 3) the procedural memory system, encompassing numerous cortical and sub-cortical structures, learns sensory-motor mappings. In this chapter, we consider several major developments in computational modeling that shed light on how the brain learns to represent information at three broad levels, reflecting these three forms of memory: 1) sensory coding, 2) episodic memory, and 3) representations that guide actions. Rather than providing a comprehensive review of all models in these areas, our goal is to highlight some of the key developments in the field, and to point to the most promising directions for future work.

2.2 Sensory Coding

At the earliest stages of sensory processing in the cortex, quite a lot is known about the neural coding of information, from Hubel and Wiesel's classic findings of orientation selective neurons in primary visual cortex (Hubel and Wiesel, 1968) to more recent studies of spatio-temporal receptive fields in visual cortex (DeAngelis et al., 1993) and spectro-temporal receptive fields in auditory cortex (Calhoun and Schreiner, 1998; Kowalski et al., 1996). Given the abundance of electrophysiological data to constrain the development of computational models, it is not surprising that most models of learning and memory have focused on the early stages of sensory coding. One approach to modeling sensory coding is to hand-design filters, such as the Gabor or difference-of-Gaussians filter, so as to match experimentally observed receptive fields. However, this approach has limited applicability beyond the very earliest stages of sensory processing for which receptive fields have been reasonably well mapped out. A more promising approach is to try to understand the developmental processes that generated the observed data. Note that these could include both learning and evolutionary factors, but here our focus is restricted to potential learning mechanisms. The goal is then to discover the general underlying principles that cause sensory systems to self-organize their receptive fields. Once these principles have been uncovered, they can be used to derive models of learning. One can then simulate the developmental process by exposing the model to typical sensory input, and comparing the results to experimental observations. More importantly, one can simulate neuronal functions which might not have been conceived by experimentalists, and thereby generate novel experimental predictions.

Several classes of computational models have been influential in guiding current thinking about self-organization in sensory systems. These models share the general feature of modeling the brain as a communication channel and applying concepts from information theory. The underlying assumption of these models is that the goal of sensory coding is to map the high-dimensional sensory signal into another (usually lower-dimensional) code that is somehow optimal with respect to information content. Four information-theoretic coding principles will be considered here: 1) Linsker's Infomax principle, 2) Barlow's redundancy reduction principle, 3) Becker and Hinton's Imax principle, and 4) Risannen's Minimum Description Length (MDL) principle. Each of these principles has been used to derive models of learning, and has inspired further research into related models at multiple stages of information processing.

2.2.1 Linsker's Infomax Principle

How should neurons respond to the sensory signal, given that it is noisy, high-dimensional and highly redundant? Is there a more convenient form in which to encode signals so that we can make more sense of the relevant information and take appropriate actions? In the human visual system, for example, there are hundreds of millions of photoreceptors converging onto about two million optic nerve fibers. By what principle does the brain decide what information to discard and what to preserve?

Infomax principle Linsker proposed a model of self-organization in sensory systems based on the *Infomax principle*: Each neuron adjusts its connection strengths or weights so as to maximize the amount of Shannon information in the neural code that is conveyed about the sensory input (Linsker, d). In other words, the Infomax principle dictates that neurons should maximize the amount of mutual information between their input \mathbf{x} and output y :

$$I_{\mathbf{x};y} = \langle \ln [p(\mathbf{x}|y)/p(\mathbf{x})] \rangle$$

Assuming that the input consists of a multi-dimensional Gaussian signal with additive, independent Gaussian noise with variance $V(n)$, for a single neuron whose output y is a linear function of its inputs and connection weights \mathbf{w} , the mutual information is the log of the signal-to-noise ratio:

$$I_{\mathbf{x};y} = \frac{1}{2} \ln \frac{V(y)}{V(n)}$$

Linsker showed that a simple, Hebb-like weight update rule approximately maximizes this information measure. The center-surround receptive field (with either an on-center and off-surround or off-center and on-surround spatial pattern of connection strengths) is characteristic of neurons in the earliest stages of the visual pathways including the retina and lateral geniculate nucleus (LGN) of the thalamus. Surprisingly, Linsker's simulations using purely uncorrelated random inputs, and a multi-layer circuit as shown in Figure 2.1, showed that neurons in successive layers developed progressively more "Mexican-hat" shaped receptive fields (Linsker, c,b,a), reminiscent of the center-surround receptive fields seen in the visual system. In further developments of the model, using a two-dimensional sheet of neurons with local-neighbor lateral connections, Linsker (e) showed that the model self-organized topographic maps with oriented receptive fields, such that nearby units on the map developed similarly oriented receptive fields. This organization is a good first approximation to that of the primary visual cortex.

Independent
Components
Analysis

The Infomax principle has been highly influential in the study of neural coding, going well beyond Linsker's pioneering work in the linear case. One of the major developments in this field is Bell and Sejnowski's Infomax-based Independent Components Analysis (ICA) algorithm, which applies to nonlinear mappings with equal numbers of inputs and outputs (Bell and Sejnowski, 1995). Bell and Sejnowski showed that when the mapping from inputs to outputs is continuous, nonlinear and invertible, maximizing the mutual information between inputs and outputs is equivalent to simply maximizing the entropy of the output signal. The algorithm therefore performs a form of ICA.

Infomax-based ICA has also been used to model receptive fields in visual cortex. When applied to natural images, in contrast to PCA, Infomax-based ICA develops oriented receptive fields at a variety of spatial scales that are sparse, spatially localized, and reminiscent of oriented receptive fields in primary visual cortex (Bell and Sejnowski, 1997). Another variant of nonlinear Infomax developed by Okajima and colleagues (Okajima, 2004) has also been applied to modeling higher levels

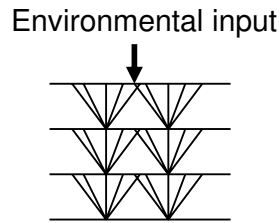


Figure 2.1 Linsker's multilayer architecture for learning center-surround and oriented receptive fields. Higher layers learned progressively more Mexican-hat-like receptive fields. The inputs consisted of uncorrelated noise, and in each layer, center-surround receptive fields evolved with progressively greater contrast between center and surround.

of visual processing, including combined binocular disparity and spatial frequency analysis.

2.2.2 Barlow's Redundancy Reduction Principle

The principle of preserving information may be a good description of the very earliest stages of sensory coding, but it is unlikely that this one principle will capture all levels of processing in the brain. Clearly, one can trivially preserve all the information in the input simply by copying the input to the next level up. Thus, the idea only makes sense in the context of additional processing constraints. Implicit in Linsker's work was the constraint of dimension reduction. However, in the neocortex, there is no evidence of a progressive reduction in the number of neurons at successively higher levels of processing.

Barlow proposed a slightly different principle of self-organization based on the idea of producing a *minimally redundant* code. The information about an underlying signal of interest (such as the visual form or the sound of a predator) may be distributed across many input channels. This makes it difficult to associate particular stimulus values with distinct responses. Moreover, there is a high degree of redundancy across different channels. Thus, a neural code having minimal redundancy should make it easier to associate different stimulus values with different responses.

The formal, information-theoretic definition of redundancy is the information content of the stimulus, less the capacity of the channel used to convey the

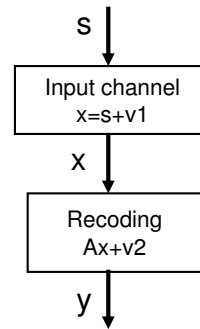


Figure 2.2 Atick and Redlich's learning principle was to minimize redundancy in the output, y , while preserving information about the input, x .

information. Unfortunately, quantities dependent upon calculation of entropy are difficult to compute. Thus, several different formulations of Barlow's principle have been proposed, under varying assumptions and approximations. One simple way for a learning algorithm to lower redundancy is reduce correlations amongst the outputs (Barlow and Földiák, 1989). This can remove second-order but not higher-order dependencies.

Atick and Redlich proposed minimizing the following measure of redundancy (Atick and Redlich, 1990):

$$R = 1 - \frac{I_{y;s}}{C_{out}(y)}$$

channel capacity

subject to the constraint of zero information loss (fixed $I_{y;s}$). $C_{out}(y)$, the output channel capacity, is defined to be the maximum of $I_{y;s}$. The channel capacity is at a maximum when the covariance matrix of the output elements is diagonal, hence Atick and Redlich used:

$$C_{out}(y) = \frac{1}{2} \prod_i \left[\frac{R_{yy}}{N_v 2^2} \right]_{ii}$$

Thus, under this formulation, minimizing redundancy amounts to minimizing the *the channel capacity*. This model is depicted in Figure 2.2. This model was used to simulate retinal receptive fields. Under conditions of high noise (low redundancy), the receptive fields which emerged were Gaussian-shaped spatial smoothing filters, while at low noise levels (high redundancy) on-center off-surround receptive fields resembling second spatial derivative filters emerged. In fact, cells in the mammalian retina and lateral geniculate nucleus of the thalamus dynamically adjust their filtering characteristics as light levels fluctuate between these two extremes under conditions of low versus high contrast (Virsu et al., 1977; Shapley and Victor, 1979a). Moreover, this strategy of adaptive rescaling of neural responses has been shown to be optimal with respect to information transmission (Brenner et al., 2000).

Similar learning principles have been applied by Atick and colleagues to model higher stages of visual processing. Dong and Atick modelled redundancy reduction across time, in a model of visual neurons in the lateral geniculate nucleus of the thalamus (Dong and Atick, 1995). In their model, neurons with both lagged and non-lagged spatio-temporal smoothing filters emerged. These receptive fields would be useful for conveying information about stimulus onsets and offsets. Li and Atick (1994) modeled redundancy reduction across binocular visual inputs. Their model generated binocular, oriented receptive fields at a variety of spatial scales, similar to those seen in primary visual cortex. Bell and Sejnowski's Infomax-based ICA algorithm (Bell and Sejnowski, 1995) is also closely related to Barlow's minimal redundancy principal, since the ICA model is restricted to invertible mappings; in this case, the maximization of mutual information amounts to reducing statistical dependencies amongst the outputs.

2.2.3 Becker and Hinton's Imax Principle

The goal of retaining as much information as possible may be a good description of early sensory coding. However, the brain seems to do much more than simply preserve information and recode it into a more convenient form. Our perceptual systems are exquisitely tuned to certain regularities in the world, and consequently, to irregularities which violate our expectations. The things which capture our attention and thus motivate us to learn and act are those which violate our expectations about the coherence of the world — the sudden onset of a sound, the appearance of a looming object or a predator.

In order to be sensitive to *changes* in our environment, we require internal representations which first capture the regularities in our environment. Even relatively low-order regularities, such as the spatial and temporal coherence of sensory signals, convey important cues for extracting very high-level properties about objects. For example, the coherence of the visual signal across time and space allows us to segregate the parts of a moving object from its surrounding background, while the coherence of auditory events across frequency and time permits the segregation of the auditory input into its multiple distinct sources.

Becker and Hinton (1992) proposed the Imax principle for unsupervised learning, which dictates that signals of interest should have high mutual information across different sensory channels. In the simplest case, illustrated in Figure 2.3, there are two input sources, x_1 and x_2 conveying information about a common underlying Gaussian signal of interest, s , and each channel is corrupted by independent, additive Gaussian noise:

$$x_1 = s + n_1 \quad x_2 = s + n_2$$

However, the input may be high-dimensional and may require a nonlinear transformation in order to extract the signal. Thus the goal of the learning is to transform the two input signals into outputs, y_1 and y_2 , having maximal mutual information.

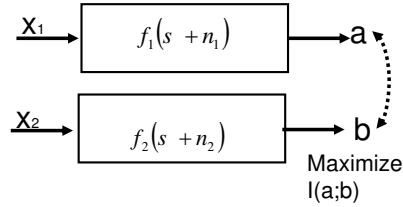


Figure 2.3 Becker and Hinton’s Imax learning principle maximizes the mutual information between features a and b extracted from different input channels.

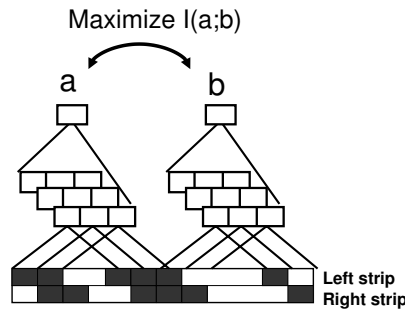


Figure 2.4 Imax architecture used to learn stereo features from binary images.

Because the signal is Gaussian and the noise terms are assumed to be identically distributed, the information in common to the two outputs can be maximized by maximizing the following log signal-to-noise ratio (SNR):

$$I_{y_1;y_2} \approx \log \left(\frac{V(y_1 + y_2)}{V(y_1 - y_2)} \right)$$

This could be accomplished by multiple stages of processing in a nonlinear neural circuit like the one shown in Figure 2.4. Becker and Hinton (1992) showed that this model could extract binocular disparity from random dot stereograms, using the architecture shown in Figure 2.4. Note that this function requires multiple stages of processing through a network of nonlinear neurons with sigmoidal activation functions.

The Imax algorithm has been used to learn temporally coherent features (Becker, 1996; Stone, 1996), and extended to learn multidimensional features (Zemel and

Hinton, 1991). A very similar algorithm for binary units was developed by Kay and colleagues (Kay, 1992; Phillips et al., 1998). de Sa's minimizing disagreement algorithm (de Sa, 1994) is a probabilistic learning procedure based on principles of Bayesian classification, but is nonetheless very similar to Imax in its objective to extract classes that are coherent across multiple sensory input channels.

2.2.4 Risannen's Minimum Description Length Principle

Minimum Description Length

The overall goal of every unsupervised learning algorithm is to discover the important underlying structure in the data. Learning algorithms based on Shannon information have the drawback of requiring knowledge of the probability distribution of the data, and/or of the extracted features, and hence tend to be either very computationally expensive or to make highly simplifying assumptions about the distributions (e.g. binary or Gaussian variables). An alternative approach is to develop a model of the data that is somehow optimal with respect to coding efficiency. The Minimum Description Length (MDL) principle, first introduced by Risannen (1978), favors models that provide accurate encoding of the data using as simple a model as possible. The rationale behind the MDL principle is that the criterion of discovering statistical regularities in data can be quantified by the length of the code generated to describe the data.

A large number of learning algorithms have been developed based on the MDL principle, but only a few of these have attempted to provide plausible accounts of neural processing. One such example was developed by Zemel and Hinton (1995), who cast the auto-encoder problem within an MDL framework. They proposed that the goal of learning should be to encode the total cost of communicating the input data, which depends on three terms, the length of the code, c , the cost of communicating the model M — which depends on the coding cost of communicating how to reconstruct the data, M^{-1} , and the reconstruction error:

$$Cost = Length(c) + Length(M^{-1}) + Length(|d - d'|)$$

as illustrated in Figure 2.5. They instantiated these ideas using an autoencoder architecture, with hidden units whose activations were Gaussian functions of the inputs. Under a Gaussian model of the input activations, it was assumed that the hidden unit activations, as a population, encode a point in a lower-dimensional implicit representational space. For example, a population of place cells in the hippocampus might receive very high-dimensional multi-sensory input, and map this input onto a population of neural activations which codes implicitly the animal's spatial location — a point in a two-dimensional Cartesian space. The population response could be decoded by averaging together the implicit coordinates of the hidden units, weighted by their activations. Zemel and Hinton's cost function incorporated a reconstruction term and a coding cost term that measured the fit of the hidden unit activations to a Gaussian model of implicit coordinates. The weights of the hidden units and the coordinates in implicit space were jointly optimized with

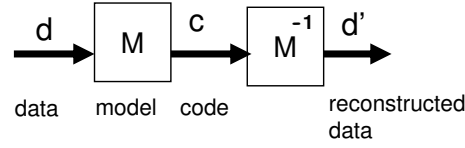


Figure 2.5 Minimum Description Length (MDL) principle.

respect to this MDL cost.

Algorithms which perform clustering, when cast within a statistical framework, can also be viewed as a form of MDL learning. Nowlan derived such an algorithm, called maximum likelihood competitive learning (MLCL), for training neural networks using the EM algorithm (Nowlan, 1990; Jacobs et al., 1991). In this framework, the network is viewed as a probabilistic, generative model of the data. The learning serves to adjust the weights so as to maximize the log likelihood of the model having generated the data:

$$L = \log P(\text{data} \mid \text{model}).$$

If the training patterns, $I^{(\alpha)}$, are independent,

$$\begin{aligned} L &= \log \prod_{\alpha=1}^n P(I^{(\alpha)} \mid \text{model}) \\ &= \sum_{\alpha=1}^n \log P(I^{(\alpha)} \mid \text{model}). \end{aligned}$$

The MLCL algorithm applies this objective function to the case where the units have Gaussian activations and form a mixture model of the data:

$$\begin{aligned} L &= \sum_{\alpha=1}^n \log \left[\sum_{i=1}^m P(I^{(\alpha)} \mid \text{submodel}_i) P(\text{submodel}_i) \right] \\ &= \sum_{\alpha=1}^n \log \left[\sum_{i=1}^m y_i^{(\alpha)} \pi_i \right] \end{aligned}$$

where the π_i 's are positive mixing coefficients that sum to one, and the y_i 's are the unit activations:

$$y_i^{(\alpha)} = \mathcal{N}(\vec{I}^{(\alpha)}, \vec{w}_i, \Sigma_i)$$

where $\mathcal{N}(\cdot)$ is the Gaussian density function, with mean \vec{w}_i and covariance matrix Σ_i .

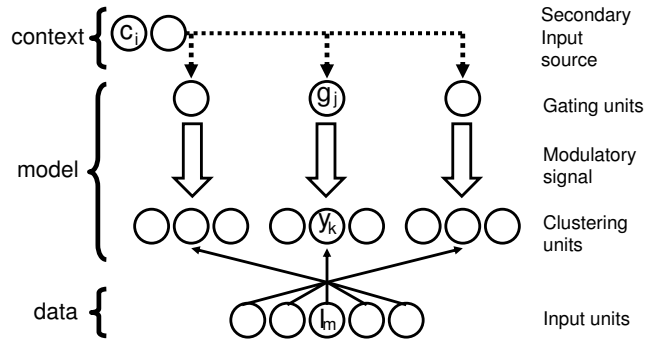


Figure 2.6 Contextually modulated competitive learning.

The MLCL model makes the assumption that every pattern is independent of every other pattern. However, this assumption of independence is not valid under natural viewing conditions. If one view of an object is encountered, a similar view of the same object is likely to be encountered next. Hence, one powerful cue for real vision systems is the temporal continuity of objects. Novel objects typically are encountered from a variety of angles, as the position and orientation of the observer, or objects, or both, vary smoothly over time. Given the importance of temporal context as a cue for feature grouping and invariant object recognition, it is very likely that the brain makes use of this property of the world in perceptual learning. Becker (1999) proposed an extension to MLCL that incorporates context into the learning. Relaxing the assumption that the patterns are independent, allowing for temporal dependencies amongst the input patterns, the log likelihood function becomes:

$$L = \log P(\text{data} \mid \text{model}) \\ = \sum_{\alpha} \log P(I^{(\alpha)} \mid I^{(1)}, \dots, I^{(\alpha-1)}, \text{model})$$

To incorporate a contextual information source into the learning equation, a contextual input stream was introduced into the likelihood function:

$$L = \log P(\text{data} \mid \text{model}, \text{context}) \\ = \sum_{\alpha} \log P(I^{(\alpha)} \mid I^{(1)}, \dots, I^{(\alpha-1)}, \text{model}, \text{context})$$

as depicted in Figure 2.6. This model was trained on a series of continuously rotating images of faces, and learned a representation that categorized people's faces according to identity, independent of viewpoint, by taking advantage of the temporal continuity in the image sequences.

Many models of population encoding apply to relatively simple, one-layer feed-forward architectures. However, the structure of neocortex is much more complex. There are multiple cortical regions, and extensive feedback connections both within and between regions. Taking these features of neocortex into account, Hinton has developed a series of models based on the Boltzmann machine (Ackley et al., 1985), and the more recent Helmholtz machine (Dayan et al., 1995) and Product of Experts (PoE) (Hinton, 2000; Hinton and Brown, 2000). The common idea underlying these models is to try to find a population code that forms a causal model of the underlying data. The Boltzmann machine was unacceptably slow at sampling the “unclamped” probability distribution of the unit states. The Helmholtz machine and PoE model overcome this limitation by using more restricted architectures and/or approximate methods for sampling the probability distributions over units’ states, see Figure 2.7A). In both cases, the bottom-up weights embody a “recognition model”; that is, they are used to produce the most probable set of hidden states given the data. At the same time, the top-down weights constitute a “generative model”; that is, they produce a set of hidden states most likely to have generated the data. The “wake-sleep algorithm” maximizes the log likelihood of the data under this model and results in a simple equation for updating either set of weights:

$$\Delta w_{kj} = \varepsilon s_k^\alpha (s_j^\alpha - p_j^\alpha)$$

where p_j^α is the target state for unit j on pattern α , and s_j^α is the corresponding network state, a stochastic sample based on the logistic function of the unit’s net input. Target states for the generative weight updates are derived from top-down expectations based on samples using the recognition model, whereas for the recognition weights, the targets are derived by making bottom-up predictions based on samples from the generative model. The Products of Experts model advances on this learning procedure by providing a very efficient procedure called ‘brief Gibbs sampling’ for estimating the most probable states to have generated the data, as illustrated in Figure 2.7B).

2.3 Models of Episodic Memory

Moving beyond sensory coding to high-level memory systems in the medial temporal lobe (MTL), the brain may use very different optimization principles aimed at the memorization of complex events or spatio-temporal episodes, and subsequent reconstruction of details of these episodic memories. Here, rather than recoding the incoming signals in a way that abstracts away unnecessary details, the goal is to memorize the incoming signal as accurately as possible in a single learning trial. The hippocampus is a key structure in the MTL that appears to be crucial for episodic memory. It receives input from most cortical regions, and is at the point of convergence between the ventral and dorsal visual pathways, as illustrated in

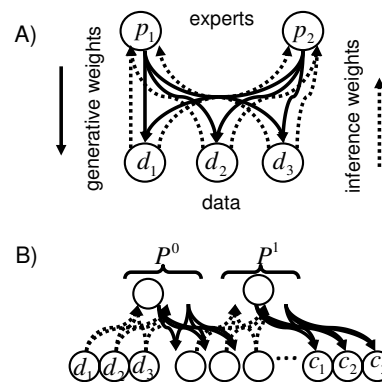


Figure 2.7 Hinton’s Product of Experts model, showing A) the basic architecture, and B) Brief Gibbs sampling, which involves several alternating iterations of clamping the input units to sample from the hidden unit states, and then clamping the hidden units to sample from the input unit states. This procedure samples the “unclamped” distribution of states in a local region around each data vector and tries to minimize the difference between the clamped and unclamped distributions.

Figure 2.8 (adapted from Mishkin et al. (1997)). Some of the unique anatomical and physiological characteristics of the hippocampus include the following: 1) the very large expansion of dimensionality from the entorhinal cortex (EC) to the dentate gyrus (DG) (the principal cells in the dentate gyrus outnumber those of the EC by about a factor of 5 in the rat (Amaral et al., 1990)); 2) the large and potent mossy fiber synapses projecting from CA3 to CA1, which are the largest synapses in the brain and have been referred to as ‘detonator synapses’ (McNaughton and Morris, 1987); and 3) the extensive set of recurrent collateral connections within the CA3 region. In addition, the hippocampus exhibits unique physiological properties including: 1) extremely sparse activations (low levels of activity), particularly in the dentate gyrus where firing rates of granule cells are about 0.5Hz (Barnes et al., 1990; Jung and McNaughton, 1993)), and 2) the constant replacement of neurons (neurogenesis) in the dentate gyrus: about about 1% of the neurons in the dentate gyrus are replaced each day in young adult rates (Martin Wojtowicz, University of Toronto, unpublished data).

In 1971 Marr (1971) put forward a highly influential theory of hippocampal coding. Central to Marr’s theory were the notions of a rapid, temporary memory store mediated by sparse activations and Hebbian learning, an associative retrieval system mediated by recurrent connections, as well as a gradual consolidation process by which new memories would be transferred into a long-term neocortical store. In the decades since the publication of Marr’s computational theory, many researchers have built on these ideas and simulated memory formation and retrieval in Marr-like models of the hippocampus. For the most part, modelers have focused on either

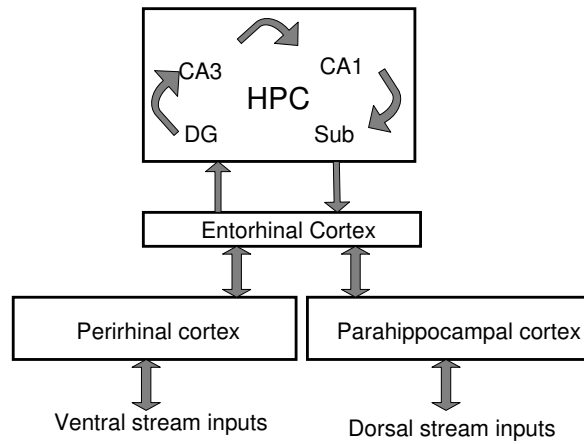


Figure 2.8 Some of the main anatomical connections of the hippocampus. The hippocampus is a major convergence zone. It receives input via the entorhinal cortex from most regions of the brain including the ventral and dorsal visual pathways. It also sends reciprocal projections back to most regions of the brain. Within the hippocampus, the major regions are the dentate gyrus (DG), CA3 and CA1. The CA1 region projects back to the entorhinal cortex, thus completing the loop. Note that the subiculum, not shown here, is another major output target of the hippocampus.

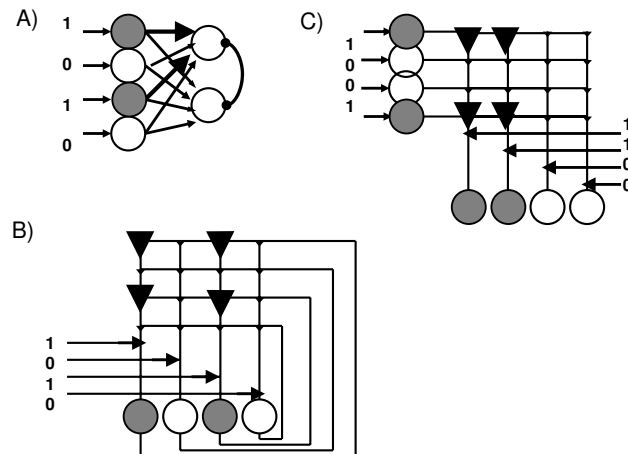


Figure 2.9 Various models have been proposed for specific regions of the hippocampus, for example, A) models based on variants of competitive learning have been proposed for the dentate gyrus; B) many models of the CA3 region have been based upon the recurrent auto-associator, and C) several models of CA1 have been based on the heteroassociative network, where the input from the entorhinal cortex to CA1 acts as a teaching signal, to be associated with the (non-driving) input from the CA3 region.

the CA3 or CA1 fields, using variants of Hebbian learning, for example, competitive learning in the dentate gyrus and CA3 (Rolls, 1989; McClelland et al., 1995; Hasselmo et al., 1996), Hebbian auto-associative learning (Marr, 1971; McNaughton and Morris, 1987; Rolls, 1989; Treves and Rolls, 1992; Kali and Dayan, 2000; O'Reilly and Rudy, 2001), temporal associative learning (Levy, 1996; Wallenstein and Hasselmo, 1997; Gerstner and Abbott, 1997; Stringer et al., 2002) in the CA3 recurrent collaterals, and Hebbian hetero-associative learning between EC-driven CA1 activity and CA3 input (Hasselmo and Schnell, 1994) or between EC-driven and CA3-driven CA1 activity at successive points in time (Levy et al., 1990). The key ideas behind these models are summarized in Figure 2.9.

In modeling the MTL's hippocampal memory system, Becker (2005) has shown that a global optimization principle based on the goal of accurate input reconstruction, combined with neuroanatomical constraints, leads to simple, biologically plausible learning rules for all regions within the hippocampal circuit. The model exhibits the key features of an episodic memory system: high storage capacity, accurate cued recall, and association of items across time, under extremely high plasticity conditions.

The key assumptions in Becker's model are as follows:

- During encoding, dentate granule cells are active whereas during retrieval they are relatively silent.
- During encoding, activation of CA3 pyramidal cells is dominated by the very strong mossy fiber inputs from dentate granule cells.
- During retrieval, activation of CA3 pyramidal cells is driven by direct perforant path inputs from the entorhinal cortex combined with time-delayed input from CA3 via recurrent collaterals.
- During encoding, activation of CA1 pyramidal cells is dominated by direct perforant path inputs from the entorhinal cortex.
- During retrieval, CA1 activations are driven by a combination of perforant path inputs from the entorhinal cortex and Shaffer collateral inputs from CA3.

Becker proposed that each hippocampal layer should form a neural representation that could be transformed in a simple manner — i.e. linearly — to reconstruct the original activation pattern in the entorhinal cortex. With the addition of biologically plausible processing constraints regarding connectivity, sparse activations, and two modes of neuronal dynamics during encoding versus retrieval, this results in very simple Hebbian learning rules.

It is important to note, however, that the model itself is highly nonlinear, due to the sparse coding in each region and the multiple stages of processing in the circuit as a whole; the notion of linearity only comes in at the point of *reconstructing* the EC activation pattern from any one region's activities. The objective function made use of the idea of an implicit set of reconstruction weights from each hippocampal region, by assuming that the perforant path connection weights could be used in reverse to reconstruct the EC input pattern. Taking the CA3 layer as an example, the CA3 neurons receive perforant path input from the entorhinal cortex, $EC^{(in)}$ associated with a matrix of weights $W^{(EC,CA3)}$. The CA3 region also receives input connections from the dentate gyrus, DG with associated weights $W^{(DG,CA3)}$ as well as recurrent collateral input from within the CA3 region with connection weights $W^{(CA3,CA3)}$. Using the transpose of the perforant path weights, $(W^{(EC,CA3)})^T$, to calculate the CA3 region's reconstruction of the entorhinal input vector:

$$EC^{(reconstructed)} = W^{(EC,CA3)^T} CA3 \quad (2.1)$$

the goal of the learning is to make this reconstruction as accurate as possible. To quantify this goal, the objective function Becker proposed to be maximized here is the cosine angle between the original and reconstructed activations:

$$\begin{aligned} Perf^{(CA3)} &= \cos(EC^{(in)}, W^{(EC,CA3)^T} CA3) \\ &= \frac{(EC^{(in)})^T (W^{(EC,CA3)^T} CA3)}{\|EC^{(in)}\| \|W^{(EC,CA3)^T} CA3\|} \end{aligned} \quad (2.2)$$

By rearranging the numerator, and appropriately constraining the activation levels and the weights so that the denominator becomes a constant, it is equivalent to

maximize the following simpler expression:

$$Perf^{(CA3)} = (W^{(EC,CA3)} EC^{(in)})^T CA3 \quad (2.3)$$

which makes use of the locally available information arriving at the CA3 neurons' incoming synapses: the incoming weights and activations. This says that the incoming weighted input from the perforant path should be as similar as possible to the activation in the CA3 layer. Note that the CA3 activation, in turn, is a function of both perforant path and DG input as well as CA3 recurrent input. The objective functions for the dentate and CA1 regions have exactly the same form as equation 2.3, using the DG and CA1 activations and perforant path connection weights respectively. Thus, the computational goal for the learning in each region is to maximize the overlap between the perforant path input and that region's reconstruction of the input. This objective function can be maximized with respect to the connection weights on each set of input connections for a given layer, to derive a set of learning equations.

By combining the learning principle with the above constraints, Hebbian learning rules are derived for the direct (monosynaptic) pathways from the entorhinal cortex to each hippocampal region, a temporal Hebbian associative learning rule is derived for the CA3 recurrent collateral connections, and a form of hetero-associative learning is derived for the Shaffer collaterals (the projection from CA3 to CA1).

Of fundamental importance for computational theories of hippocampal coding is the striking finding of neurogenesis in the adult hippocampus. Although there is now a large literature on neurogenesis in the dentate gyrus, and it has been shown to be important for at least one form of hippocampal-dependent learning, surprisingly few attempts have been made to reconcile this phenomenon with theories of hippocampal memory formation. Becker (2005) suggested that the function of new neurons in the dentate gyrus is in the generation of novel codes. Gradual changes in the internal code of the dentate layer were predicted to facilitate the formation of distinct representations for highly similar memory episodes.

Why doesn't the constant turnover of neurons in the dentate gyrus, and hence the constant rewiring of the hippocampal memory circuit, interfere with the retrieval of old memories? The answer to this question comes naturally from the above assumptions about neuronal dynamics during encoding versus retrieval. New neurons are added only to the dentate gyrus, and the dentate gyrus drives activation in the hippocampal circuit only during encoding, not during retrieval. Thus, the new neurons contribute to the formation of distinctive codes for novel events, but not to the associative retrieval of older memories.

2.4 Representations That Guide Action Selection

Moving beyond the question of how information is represented, we must consider the brain not simply a passive storage device, but as a part of a dynamical computational system that acts and reacts to changes within its environment, as

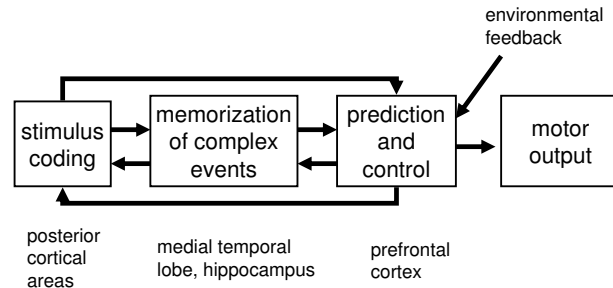


Figure 2.10 Architecture of a learning system that incorporates perceptual learning, episodic memory and motor control.

illustrated in Figure 2.10. Ultimately, models based on the broad goals of prediction and control may be our best hope for characterizing complex dynamical systems which form representations in the service of guiding motor actions.

Reinforcement learning algorithms can be applied to control problems, and have been linked closely to specific neural mechanisms. These algorithms are built upon the concept of a value function, $V(s_t)$, which defines the value of being in the current state s_t at time t to be equal to the expected sum of future rewards:

$$V(t) = r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \dots + \gamma^n r_{t+n} + \dots$$

The parameter γ , chosen to be in the range $0 \leq \gamma \leq 1$, is a temporal discount factor which permits one to heuristically weight future rewards more or less heavily according to the task demands. Within this framework, the goal for the agent is to choose actions that will maximize the value function. In order for the agent to solve the control problem – how to select optimal actions, it must first solve the prediction problem – how to estimate the value function. The TD-learning algorithm (Sutton and Barto, 1981; Sutton, 1988) provides a rule for incrementally updating an estimate \hat{V}_t of the true value function at time t by an amount called the TD-error: TD-error = $r_{t+1} + \gamma \hat{V}_{t+1} - \hat{V}_t$, which makes use of r_t , the amount of reward received at time t , and the value estimates at the current and the next time step. It has been proposed that the TD-learning algorithm may be used by neurobiological systems, based on evidence that firing of midbrain dopamine neurons correlates well with TD-error (Montague et al., 1996).

TD-learning

Q-learning

The Q-learning algorithm (Watkins, 1989) extends the idea of TD-learning to the problem of learning an optimal control policy for action selection. The goal for the agent is to maximize the total future expected reward. The agent learns incrementally by trial and error, evaluating the consequences of taking each action

in each situation. Rather than using a value function, Q-learning employs an action-value function, $Q(s_t, a_t)$, which represents the value in taking an action a_t when the state of the environment is s_t . The learning algorithm for incrementally updating estimates of Q-values is directly analogous to TD-learning, except that the TD-error is replaced by a temporal difference between Q-values at successive points in time.

Becker and Lim (2003) proposed a model of controlled memory retrieval based upon Q-learning. People have a remarkable ability to encode and retrieve information in a flexible manner. Understanding the neuronal mechanisms underlying strategic memory use remains a true challenge. Neural network models of memory have typically dealt with only the most basic operations involved in storage and recall. Evidence from patients with frontal lobe damage indicates a crucial role for the prefrontal cortex in the control of memory. Becker and Lim's model was developed to shed light on the neural mechanisms underlying strategic memory use in individuals with intact and lesioned frontal lobes. The model was trained to simulate human performance on free recall tasks involving lists of words drawn from a small set of categories. Normally when people are asked repeatedly to study and recall the same list of words, their recall patterns demonstrate progressively more categorical clustering over trials. This strategy thus appears to be learned, and correlates with overall recall scores. On the other hand, when patients with frontal lobe damage perform such tests, while they do benefit somewhat from the categorical structure of word lists, they tend to recall fewer categories in total, and tend to show lower semantic clustering scores. Becker and Lim postulated a role for the prefrontal cortex (PFC) in self-organizing novel mnemonic codes that could subsequently be used as retrieval cues to improve retrieval from long-term memory. Their model is outlined in Figure 2.11.

The "actions" or responses in this model are actually the activations generated by model neurons in the PFC module. Thus, the activation of each response unit is proportional to the network's current estimate of the Q-value associated with that response, and response probabilities are calculated directly from these Q-values. Learning the memory retrieval strategy involved adapting the weights for the response units so as to maximize their associated Q-values. Reinforcement obtained on a given trial was self-generated by an internal evaluation module, so that the PFC module received a reward whenever a non-repeated study list item was retrieved, and a punishment signal (negative reinforcement) when a non-list or repeated item was retrieved. The model thereby learned to develop retrieval strategies dynamically in the course of both study and free recall of words. The model was able to capture the performance of human subjects with both intact and lesioned frontal lobes on a variety of types of word lists, in terms of both recall accuracy and patterns of errors.

The model just described addresses a rather high level of complex action selection, namely, the selection of memory retrieval strategies. Most work on modelling action selection has dealt with more concrete and observable actions such as the choice of lever-presses in a response box or choice of body-turn directions in a

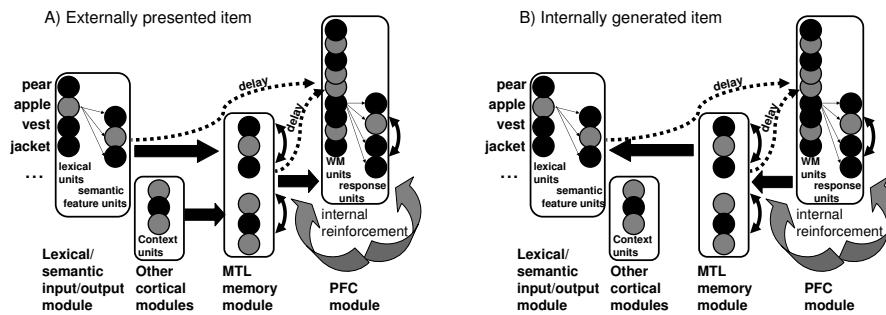


Figure 2.11 Becker and Lim's architecture for modelling the frontal control of memory retrieval. The model operated in two different modes: A) During perception of an external stimulus (during a study phase) there was bottom-up flow of activation. B) During free recall, when a response was generated internally, there was a top-down flow of activation in the model. After an item was retrieved, but before a response was generated, the item was used to probe the MTL memory system, and its recency was evaluated. If the recency (based on a match of the item to the memory weight matrix) was too high, the item was considered to be a repetition error, and if too low, it was considered to be an extra-list intrusion error. Errors detected by the model were not generated as responses, but were used to generate internal reinforcement signals for learning the PFC module weights. Occasionally, a repetition or intrusion error might go undetected by the model, resulting in a recall error.

maze. The advantage of this level of modelling is that it can make contact with a large body of experimental literature on animal behaviour, pharmacology and physiology. Many such models have employed TD-learning or Q-learning, under the assumption that animals form internal representations of value functions, which guide action selection. As mentioned above, phasic firing of dopamine neurons has been postulated to convey the TD-error signal critical for this type of learning. However, in addition to its importance in modulating learning, dopamine plays an important role in modulating action choice. It has been hypothesized that tonic levels of dopamine have more to do with motivational value, whereas the phasic firing of dopamine neurons conveys a learning-related signal (Smith et al., 2005).

Rather than assuming that actions are solely guided by value functions, Smith et al. (2005) hypothesized that animals form detailed internal models of the world. Value functions condense the reward value of a series of actions into that of a single state, and are therefore insensitive to the motivational state of the animal (e.g. whether it is hungry or not). Internal models, on the other hand, allow a mental simulation of alternative action choices, which may result in qualitatively different rewards. For example, an animal might perform one set of actions leading to water only if it is thirsty, and another set of actions leading to food only if it is hungry. The internal model can be described by a Markov decision process (MDP) over a set of internal states, with associated transition function and reward function, as in Figure 2.12 A). The transition function and (immediate) reward value of each state are learned through trial and error. Once the model is fully trained, action selection involves simulating a look-ahead process in the internal model for one or more steps in order to evaluate the consequences of an action. Finally, at the end of the simulation sequence, the animal's internal model reveals whether the outcome is favorable (leads to reward) or not. An illustrative example is shown in Figure 2.12 B). The choice faced by the animal is either to take the right arm of the T-maze to receive a small reward, or to take the left arm and then jump over a barrier to receive a larger reward. The role of tonic dopamine in this model is to modulate the efficacy of the connections in the internal model. Thus, when dopamine is depleted, the model's ability to simulate the look-ahead process to assess expected future reward will be biased toward rewards available immediately rather than more distal rewards. This implements an online version of temporal discounting.

Cousins et al. (1996) found that normal rats trained in the T-maze task in Figure 2.12 B) are willing to jump the barrier to receive a larger food reward nearly 100% of the time. Interestingly, however, when rats were administered a substance that destroys dopaminergic projections to the nucleus accumbens (DA-lesion), they chose the smaller reward. In another version of the task, rats were trained on the same maze except that there was no food in the right arm, and then when given DA-lesions, they nearly always chose the left arm and jumped the barrier to receive a reward. Thus, the DA-lesion was not merely disrupting motor behaviour, it was interacting with the motivational value of the behavioural choices. Note that the TD-error account of dopamine only provides for a role in learning, and would have

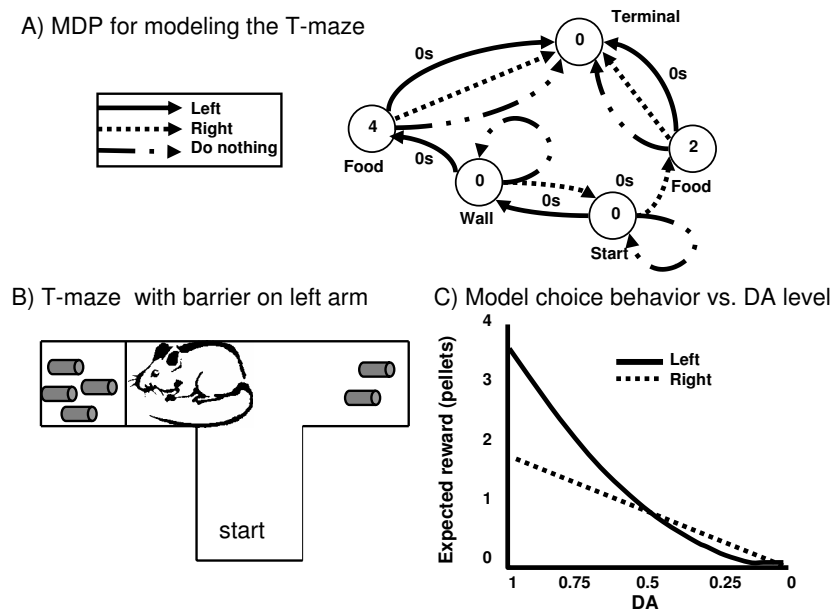


Figure 2.12 The T-maze with

nothing to say about effects of dopamine on behaviour subsequent to learning. Smith et al argued, based on these and other data, that dopamine serves to modulate the motivational choice of the animals, with high levels of dopamine favoring the selection of action sequences with more distal but larger rewards. In simulations of the model, depletion of dopamine therefore biases the choice in favor of the right arm in this task, as shown in Figure 2.12 C).

2.5 New Directions: Integrating Multiple Memory Systems

In this chapter, we have reviewed several approaches to modelling the mind, from low-level sensory coding, to high-level memory systems, to action selection. Somehow, the brain accomplishes all of these functions, and it is highly unlikely that they are carried out in isolation of one another. For example, we now know that striatal dopaminergic pathways, presumed to carry a reinforcement learning signal, affect sensory coding even in early sensory areas such as primary auditory cortex (Bao et al., 2001). Future work must address the integration of these various levels of modeling.