A Computational Model of the Cerebral Cortex

Thomas Dean
Department of Computer Science
Brown University
Providence, RI 02812
tld@cs.brown.edu

Abstract

Our current understanding of the primate cerebral cortex (neocortex) and in particular the posterior, sensory association cortex has matured to a point where it is possible to develop a family of graphical models that capture the structure, scale and power of the neocortex for purposes of associative recall, sequence prediction and pattern completion among other functions. Implementing such models using readily available computing clusters is now within the grasp of many labs and would provide scientists with the opportunity to experiment with both hard-wired connection schemes and structure-learning algorithms inspired by animal learning and developmental studies. While neural circuits involving structures external to the neocortex such as the thalamic nuclei are less well understood, the availability of a computational model on which to test hypotheses would likely accelerate our understanding of these circuits. Furthermore, the existence of an agreed-upon cortical substrate would not only facilitate our understanding of the brain but enable researchers to combine lessons learned from biology with state-of-the-art graphical-model and machine-learning techniques to design hybrid systems that combine the best of biological and traditional computing approaches.

Introduction

In the last decade, researchers have made significant progress in understanding the structure and function of the cerebral cortex and associated regions of the human/primate brain. Multi-cell recordings and imaging techniques such as fMRI have enabled scientists to build upon the foundational work of Mountcastle, Hubel and Wiesel, and others. Experimental studies have yielded insights not only concerning low-level processing but higher-order cognitive functions as well (see, for example, the review articles in a recent issue of *Science* on encoding and retrieving of episodic memory (Miyashita 2004) and the role of the medial frontal cortex in cognitive control (Ridderinkhof *et al.* 2004)).

We argue that current theoretical models of the neocortex are sufficiently rich in their predictive power and detailed in their specification that they warrant a concerted effort to implement and subject to computational experiment. In particular, we argue that the neocortex makes an ideal candidate for implementation given its relatively homogeneous structure, the state of our knowledge concerning its function and circuitry, and its powerful inferential capability.

While there has been substantial progress refining models of individual neurons since the work of McCulloch and Pitts, it is our contention that a more aggregate model of cortical circuitry is necessary to instantiate a model that approaches the scale and complexity of the human cortex. And, moreover, that only at this scale and complexity will the instantiated model be useful as a basis for experimental study and as a potential component in more general, hybrid computational architectures. To that end we adopt a Bayesian, graphical-models approach (Jordan 1998) for its expressive modeling capability and potential integration with other inferential components based on graphical models.

Several researchers have developed mathematical models of the neocortex that accord well with experimental results and provide directions for implementation, e.g., (Anderson 2003; George & Hawkins 2005; Lee & Mumford 2003; Rao & Ballard 1996; Zemel 2000). In this paper, we generalize Lee and Mumford's (2003) model of hierarchical Bayesian inference in the visual cortex and borrow from Anderson and Sutton's (1997) *Network of Networks* model for ideas about learning and connectivity in large graphical models. The challenge is to design and build a composite generative graphical model that combines top-down and bottom-up hierarchical inference, employs invariant and compositional representations for encoding and retrieving patterns, and supports associative recall, pattern completion and sequence prediction among other functions.

Graphical Models of the Neocortex

In the early 1990s, Mumford (1991; 1992), proposed a computational theory of the neocortex. Building on this work he subsequently described (1994; 2002) how Grenander's pattern theory (Grenander 1993) could potentially model the brain in terms of a generative model in which feedback is used to resolve ambiguities. Dayan *et al*'s *Helmholtz machine* (1995) provided a first approximation to realizing this approach by using feedback-implementing priors. However feedback in the Helmholtz machine is utilized only in support of learning and not as an integral part of inference. Rao

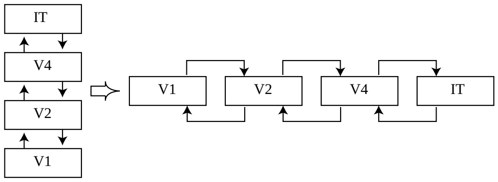

Figure 1: The first four regions of the visual cortex are shown on the left as a stack depicting its hierarchical organization; regions higher in the stack capture general visual features while those lower in the stack capture specific — typically more transient — features. The alternative depiction on the right underscores the fact that the regions are physically arranged as patches on a planar cortical sheet.

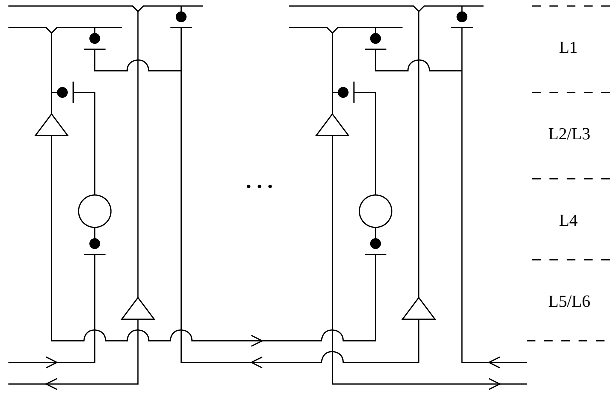

Figure 2: A cross section of the cortical sheet depicting the columnar structure and the six layers of the cerebral cortex. The triangles in Layers 2 and 3 represent supragranular pyramidal cells, the circles in Layer 4 spiny stellate cells, and the triangles in Layers 5 and 6 infra-granular pyramidal cells. Black dots represent dendrites. The cells on the left schematically represent a region consisting of many columns at one level in a hierarchy of features communicating with a second region — shown on the right — at a higher level in the hierarchy (after (Braitenberg & Schuz 1991)).

and Ballard's (1996) predictive coding Kalman filter model employs feedback during inference, but their approach is limited due to its reliance on linear models.

Lee and Mumford's (2003) model for hierarchical inference in the visual cortex provides a significant step forward by taking advantage of the confluence of ideas from statistics, neuroscience and cognitive and computational science. Specifically, they propose a generative model of the visual cortex based on hierarchical Bayesian inference; their model emphasizes the role of feedback and meets head on the challenge of coping with the combinatorial explosion of conflicting interpretations. In the remainder of this section, we present Lee and Mumford's model emphasizing its hierarchical structure and pattern of feedback required to facilitate simultaneous top-down and bottom-up inference.

Figure 1 shows the first four regions of the visual cortex arranged on the left as a hierarchy of increasingly abstract visual features and their associated processing units, and on the right schematically as they appear distributed across the cortical sheet. The figure also depicts the postulated interaction between regions implementing both top-down and bottom-up communication. The bottom-up communications are used to combine more primitive features into more abstract ones and the top-down communications are used to exploit expectations generated from prior experience.

The basic feed-forward and feedback circuitry is depicted in a highly schematic form in Figure 2. Feed-forward projections originate in Layer 3 and terminate in Layer 4 while feedback projections originate in Layers 5 and 6 and terminate in Layer 1. The exact circuitry is more complicated but it does appear that the cortex is fully capable of implementing information flows of the sort shown in Figure 1.

Lee and Mumford provide a Bayesian account of the hierarchical inference implied in Figure 1. When applied to vision, Bayes' rule enables us to combine previous experience in the form of a *prior* probability $P(x_1|x_B)$ on x_1 — the hidden variable — and x_B — the context or background knowledge — with an imaging model $P(x_\mathrm{O}|x_1, x_B)$ relating the observations x_O to the other variables:

$$P(x_\mathrm{O}, x_1|x_B) = P(x_\mathrm{O}|x_1, x_B)P(x_1|x_B) \qquad (1)$$

Typically we assume that the imaging model does not depend on the background knowledge, thereby allowing us to rewrite Equation 1 as:

$$P(x_1|x_\mathrm{O}, x_B) = \frac{P(x_\mathrm{O}|x_1, x_B)P(x_1|x_B)}{P(x_\mathrm{O}|x_B)}$$

where the $P(x_\mathrm{O}|x_B)$ is the normalizing factor, Z_1, required so that $P(x_1|x_\mathrm{O}, x_B)$ is a proper probability distribution.

In the case of early vision, x_O denotes the output of the lateral geniculate nucleus (LGN), x_1 denotes the features computed by V1, and x_B denotes all the other information — contextual and background — available at the time of inference. Thus V1 computes the most likely values of x_1 by finding the *a posteriori* estimate of x_1 that maximizes $P(x_1|x_\mathrm{O}, x_B)$.

Lee and Mumford extend this simple model to account for processing at multiple regions in the temporal cortex including V1, V2, V4 and the inferotemporal cortex (IT). We assume that each of these regions is responsible for computing features at different levels of abstraction. Using the chain rule and the simplifying assumption that in the sequence $(x_\mathrm{O}, x_\mathrm{V1}, x_\mathrm{V2}, x_\mathrm{V4}, x_\mathrm{IT})$ each variable is independent of the other variables given its immediate neighbors in the sequence, we write the equation relating the four regions as

$$P(x_\mathrm{O}, x_\mathrm{V1}, x_\mathrm{V2}, x_\mathrm{V4}, x_\mathrm{IT}) = \\ P(x_\mathrm{O}, x_\mathrm{V1})P(x_\mathrm{V1}, x_\mathrm{V2})P(x_\mathrm{V2}, x_\mathrm{V4})P(x_\mathrm{V4}, x_\mathrm{IT})P(x_\mathrm{IT})$$

resulting in an undirected graphical model or Markov random field (MRF) based on the chain of variables:

$$x_\mathrm{O} \leftrightarrow x_\mathrm{V1} \leftrightarrow x_\mathrm{V2} \leftrightarrow x_\mathrm{V4} \leftrightarrow x_\mathrm{IT}$$

From this it follows that

$$P(x_\mathrm{V1}|x_\mathrm{O}, x_\mathrm{V2}, x_\mathrm{V4}, x_\mathrm{IT}) = P(x_\mathrm{O}|x_\mathrm{V1})P(x_\mathrm{V1}|x_\mathrm{V2})/Z_1$$
$$P(x_\mathrm{V2}|x_\mathrm{O}, x_\mathrm{V1}, x_\mathrm{V4}, x_\mathrm{IT}) = P(x_\mathrm{V1}|x_\mathrm{V2})P(x_\mathrm{V2}|x_\mathrm{V4})/Z_2$$
$$P(x_\mathrm{V4}|x_\mathrm{O}, x_\mathrm{V1}, x_\mathrm{V2}, x_\mathrm{IT}) = P(x_\mathrm{V2}|x_\mathrm{V4})P(x_\mathrm{V4}|x_\mathrm{IT})/Z_4$$

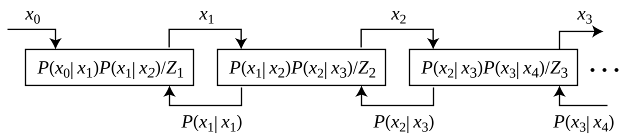

Figure 3: A schematic of the hierarchical Bayesian framework proposed by Lee and Mumford (2003). The regions of the visual cortex are linked together in a Markov chain. The activity in the ith region is influenced by bottom-up feed-forward data x_{i-1} and top-down probabilistic priors $P(x_i|x_{i+1})$ representing feedback from region $i+1$. The Markov property plays an important computational role by allowing units to depend only on their immediate neighbors in the Markov chain.

and, given that in graphical models you need only potentials $\phi(x_1, x_2)$ indicating the preferred pairs of values of the directly linked variables x_1 and x_2, we have

$$P(x_{\mathrm{V1}}|x_{\mathrm{O}}, x_{\mathrm{V2}}, x_{\mathrm{V4}}, x_{\mathrm{IT}}) = \phi(x_{\mathrm{O}}|x_{\mathrm{V1}})\phi(x_{\mathrm{V1}}|x_{\mathrm{V2}})/Z_{\mathrm{O,V2}}$$
$$P(x_{\mathrm{V2}}|x_{\mathrm{O}}, x_{\mathrm{V1}}, x_{\mathrm{V4}}, x_{\mathrm{IT}}) = \phi(x_{\mathrm{V1}}|x_{\mathrm{V2}})\phi(x_{\mathrm{V2}}|x_{\mathrm{V4}})/Z_{\mathrm{V1,V4}}$$
$$P(x_{\mathrm{V4}}|x_{\mathrm{O}}, x_{\mathrm{V1}}, x_{\mathrm{V2}}, x_{\mathrm{IT}}) = \phi(x_{\mathrm{V2}}|x_{\mathrm{V4}})\phi(x_{\mathrm{V4}}|x_{\mathrm{IT}})/Z_{\mathrm{V2,IT}}$$

where $Z_{i,j} = Z(x_i, x_j)$ is required for normalization.

The potentials must be learned from experience and are essential to the model. Roth and Black (2005) have taken ideas from the sparse coding of image patches and applied them to homogeneous Markov random fields to obtain translation-invariant models of local image statistics. Exploiting the Products-of-t-distributions model (Welling, Hinton, & Osindero 2003) and contrastive divergence (Hinton 2002), they learn a Gibbs distribution with a rich set of potentials defined using learned filters. In contrast to previous approaches that use a pre-determined set of filters, the translation-invariant learning method produces filters, as well as other distribution parameters, that properly account for spatial correlations in the data.

The resulting prior is trained using a generic database of natural images and can be exploited in any Bayesian inference method that requires a spatial prior. They demonstrate a denoising algorithm that is remarkably simple (about 20 lines of Matlab code), yet achieves performance close to the best special-purpose wavelet-based denoising algorithms. The advantage over the wavelet-based methods lies in the generality of the prior and its applicability across different vision problems.

The Roth and Black learning algorithm is a batch method and their Markov random field model is not hierarchical. By contrast, George and Hawkins (2005) present a simple on-line algorithm that learns the parameters (conditional probability density functions) of a hierarchical (three-level, tree-structured) model level-by-level starting from the lowest level. They demonstrate the translation invariant recognition capabilities of their model on simple 32×32 line drawings. The results of Roth and Black and George and Hawkins provide evidence that it is possible to learn the parameters for the Lee and Mumford model.

Figure 3 derives from Figure 1 modeling cortical regions as local experts, each encoding knowledge about the probabilistic relationships among features in a hierarchy of such features. Each expert seeks to maximize the probability of its computed features (often referred to as *beliefs*) by combining bottom-up, feed-forward feature selections with top-down, feedback expectations (priors). As information propagates up and down the hierarchy, the top-down and bottom-up *messages* change to reflect the combined expertise within the hierarchy. *Loopy belief propagation* (Murphy, Weiss, & Jordan 2000) — related to *turbo decoding* (Berrou, Glavieux, & Thitimajshima 1993) and Pearl's belief-propagation algorithm (Pearl 1986a; 1986b) — seems particularly appropriate for inference in networks of the sort illustrated in Figure 3. In loopy belief propagation, local messages propagate up and down the hierarchy as the system moves toward a global equilibrium state.

Particle filtering is another candidate for inference in Lee and Mumford's hierarchical Bayesian model; particle filtering generalizes the Kalman filter and dispenses with the restrictive requirements of Gaussian noise and linear dynamics. Particle filtering operates much like Pearl's belief-propagation algorithm in which regions corresponding to random variables in the graphical model communicate by message passing. Communication can be asynchronous providing a computationally simple framework not unlike that hypothesized for communication among regions in the visual cortex. By combining belief propagation with particle filtering, Lee and Mumford anticipated algorithms for approximate inference via non-parametric belief propagation (Isard 2003; Sudderth *et al.* 2003).

Instead of inferring a single preferred feature value x_i for the ith region in the cortical hierarchy, particle filtering generates a set of n values $\{x_i^{(1)}, x_i^{(2)}, ..., x_i^{(n)}\}$, one for each region. The n values constitute a sample of the distribution associated with ith region and serve as a proxy for the full distribution thereby avoiding the problem of keeping track of the potentially exponential number of possible states of the system. Lee and Mumford relate the algorithmic behavior of particle filtering to the ideas of Zemel and others on population coding (Zemel 2000).

The same basic organization of linked regions illustrated in Figure 1 and the same forward-backward, belief-propagation algorithm that serves for inference involving hierarchically arranged visual features serves equally well in processing time-series data of the sort arising in language and sequences of data from other sensory modalities. Variants of the cortical model outlined above can be used to represent hierarchical hidden Markov models (Fine, Singer, & Tishby 1998), factored Markov decision processes (Boutilier, Dean, & Hanks 1999) and hierarchical reinforcement learning models (Barto & Mahadevan 2003).

Architectural Issues

One of the biggest challenges in implementing and applying the above model arises in *wiring* the cortical regions corresponding to variables (or, more likely, sets of variables) in the graphical model. If we imagine starting with a large, un-

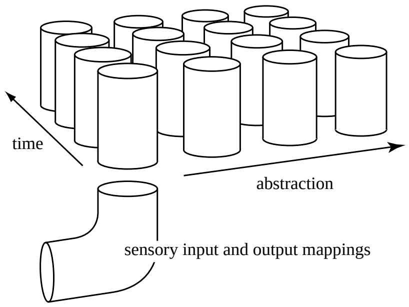

Figure 4: Architecture illustrating the spatial layout of the columns responsible for sequence prediction and those responsible for inferring high-level (compositional) features from lower-level features.

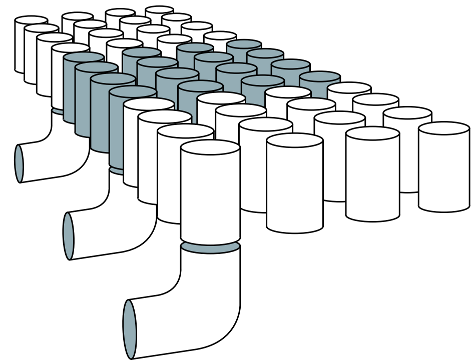

Figure 5: Representing multiple modalities by simply replicating the architecture shown in Figure 4.

differentiated Markov random field, the task seems nigh on hopeless. If, however, we start with a roughly planar topology and regular neighborhood structure, there are a variety of strategies we can use to institute non-local connections.

As mentioned earlier, the cortex consists of a layered sheet with a more-or-less uniform cellular structure. Neuroanatomists have identified what are called *columns* corresponding to groups of local cells running perpendicular to the cortical surface (see Figure 2). In a special issue of the journal *Cerebral Cortex* devoted to cortical columns, Mountcastle (2003) writes "The basic unit of cortical operation is the *minicolumn* [...] [containing] on the order of 80-100 neurons [...] The minicolumn measures of the order of 40-50 μm in transverse diameter, separated from adjacent minicolumns by vertical cell-sparse zones which vary in size in different cortical areas." These minicolumns are then grouped into cortical columns which "are formed by the binding together of many minicolumns by common input and short-range horizontal connections."

If we take the cortical column (not the minicolumn) as our basic computational module as in (Anderson & Sutton 1997), then the gross structure of the neocortex consists of a dense mat of inter-columnar connections in the outermost layer of the cortex and another web of connections at the base of the columns. The inter-columnar connectivity is relatively sparse (something on the order of 10^{15} connections spanning approximately 10^{11} neurons) and there is evidence (Sporns & Zwi 2004) to suggest that the induced inter-columnar connection graph exhibits the properties of a *small-world graph* (Newman, Watts, & Strogatz 2002); in particular, that it has low diameter (the length of the longest shortest path separating a pair of vertices in the graph) thereby enabling low-latency communication between any two cortical columns. Even simple local connectivity of the sort suggested in Anderson and Sutton's (2003) *Network of Networks* model provides a straightforward architecture that allows us to simultaneously handle temporal sequences of sensor input and support a hierarchy of increasingly abstract (compositional) features as suggested in the Lee and Mumford visual cortex model.

Figure 4 illustrates how sensor sequences and hierarchies of increasingly abstract features might be handled using just local connections. This architecture would allow us to implement hidden Markov models rather easily. Of course, the "elbow pipe" shown in Figure 4 labeled "sensory input and output mappings" masks additional dimensions that give rise to the need for longer connections. Retinotopically mapped visual input has its own spatial relationships to preserve while other sensor modalities have similar representational requirements. One might imagine a separate patch of cortex with local connections as in Figure 4 for each sensor modality and for each effector modality.

Assuming the obvious replication of structure, the architecture shown in Figure 5 illustrates the need for more than simply local connections. Different modalities would be linked by longer-range connections thereby allowing us to notice correlations across neural activity corresponding to sensors and effectors occurring over time. In implementing a multi-modal architecture, it may be useful to take a hint from nature and selectively emphasize different regions in a computational analog of blood flow and cellular ATP depletion. There's a danger in thinking that if the body had unlimited energy reserves it would perform global rather than energy-mediated local optimizations; the economics of blood flow and energy conversion may have an important role in both expediting learning and avoiding over-fitting and diffuse, hard-to-handle models.

The primate brain consumes more energy than any other organ, but even so we can only afford to activate a small fraction of our total complement of neurons, and for any given cognitive task the active cortical structures tend to be clustered rather than uniformly distributed over the entire cortex. Moreover, when faced with a new task to learn — one for which no existing, healthy structure can easily be adapted and generalized to cover — the brain tends to co-opt underutilized cortical structures. These two principles — any task can require only a fraction of the total processing units and new tasks are assigned underutilized and spatially separate units — might serve to direct wiring over time and impose a constraint to avoid over fitting, analogous to the use of minimum-description-length (MDL) priors in statistical machine learning.

Representational Issues

The Lee and Mumford model addresses the basic structure of hierarchical inference in the visual cortex but provides little detail concerning how information is represented within and between the different levels in the hierarchy. We continue to underestimate the importance of designing — or, preferably, learning — hierarchical representations for solving problems involving perceptual inference and image, speech and text processing in particular. Cognitive scientists, linguists, neural modeling experts and machine vision researchers generally agree on the need to find a way to exploit the compositional structure of language and natural scenes in order to avoid searching in the combinatorial space of possible interpretations (Chomsky 1986; Grenander 1993). It may be somewhat optimistic, but there appears to be a consensus opinion emerging at least as regards the compositional structure of a representational solution (Geman, Potter, & Chi 2002; Jackendoff 2002).

While implementing a cortical model along the lines described above does not directly address the representation problem, an agreed-upon substrate and the existence of fast, scalable implementations will serve science in three ways. First and most obviously, building upon a standard foundation promotes sharing. Some of the most interesting experiments are likely to involve combining theories regarding different modalities, and hence will require combining representations from experts in different areas; these experiments will be facilitated by having initially developed these representations separately in a shared framework. Second, there is every reason to believe that size does matter in gleaning secrets from the brain; whether the representational components consist of new primitive features or composition rules in a suitable pattern language, it is likely to require a large number of such components to span the gap from raw input to abstract concepts, and the computation required to exhibit even the most basic competence will easily swamp the capabilities of standard workstations. Third, while the architecture is sure to evolve, the existence of an agreed-upon standard will focus effort on synthetic and biologically-inspired representations, contributions that are often discouraged as being ad hoc unless embedded in a shared framework.

There is evidence to suggest that the brain makes extensive use of invariant representations particularly in the visual system where retinal circuits are largely contrast invariant, cells in V1 and V2 exhibit invariance with respect to position and size, and receptive fields in IT are sensitive across a wide range of poses. Invariants are important in enabling the visual cortex to capture essential characteristics while ignoring irrelevant details (Fukushima 1980; Riesenhuber & Poggio 1999) and it is likely that invariants play a role in processing other sensory modalities. Unfortunately, invariant representations can mask differences reducing selectivity (Geman 2004). The whole point of designing compositional hierarchical representations is to increase selectivity and reduce search by combining top-down expectations and bottom-up inference. Managing the tradeoff between invariance and selectivity is just one of the representational challenges faced in building an artificial cortex.

Implementation Issues

In terms of actually implementing a cortex-scale architecture, immediate progress can be made using medium-sized (100-200 processors running on 50-100 nodes with Myrinet/Gigabit interconnect) computing clusters and existing interconnect technologies to support 10^9 intermediate-scale functional units, each unit accounting for several thousand neurons in aggregate and arranged hierarchically in columns with virtual connections allowing upwards of 10^{14} unit-to-unit communications per millisecond. MPI codes accelerating local message passing should provide reasonable performance for a wide range of experiments (see Johansson and Lansner (2003) for a discussion of implementing cortex-sized artificial neural networks on clustered computers). Somewhat smaller, but still large-enough-to-be-interesting networks should fit on smaller clusters. While not discounting the technical challenges involved in implementing high performance cortex-sized models, these challenges are surmountable now, and will become increasingly simple to manage with continuing improvements in the performance of affordable hardware.

Conclusions

The neocortex represents a set of capabilities and a level of robustness that artificial intelligence researchers have long aspired to. Our understanding of the neural basis for inference in the cortex has advanced to a point where computation and representation are taking center stage and computer scientists can contribute to and draw inspiration from the work on neural modeling. Our choice of mathematical tools exploits the fact that Bayesian principles and graphical models have become common within the neural modeling community due to their providing clear semantics and an expressive medium for capturing neural function at multiple levels of detail. The time is opportune for creating a computational framework based on statistical techniques for modeling the neocortex and tools for implementing cortex-scale models on available hardware. This framework could serve much the same role in developing biologically-inspired computational architectures as the Bayesian reformulation of the Quick Medical Reference model (Shwe *et al.* 1991) did in pushing computer-aided medical diagnosis, while simultaneously advancing research on learning and inference algorithms for very-large graphical models.

References

Anderson, J., and Sutton, J. 1997. If we compute faster, do we understand better? *Behavior Ressearch Methods, Instruments and Computers* 29:67–77.

Anderson, J. 2003. Arithmetic on a parallel computer: Perception versus logic. *Brain and Mind* 4:169–188.

Barto, A., and Mahadevan, S. 2003. Recent advances in hierarchical reinforcement learning. *Discrete Event Dynamic Systems: Theory and Applications* 13:41–77.

Berrou, G.; Glavieux, A.; and Thitimajshima, P. 1993. Near shannon limit error-correcting coding: Turbo codes. In *Proceedings International Conference on Communications*, 1064–1070.

Boutilier, C.; Dean, T.; and Hanks, S. 1999. Decision theoretic planning: Structural assumptions and computational leverage. *J. of Artificial Intelligence Research* 11:1–94.

Braitenberg, V., and Schuz, A. 1991. *Anatomy of the Cortex*. Berlin: Springer-Verlag.

Chomsky, N. 1986. *Knowledge of Language: Its Nature, Origin and Use*. New York, NY: Praeger.

Dayan, P.; Hinton, G.; Neal, R.; and Zemel, R. 1995. The Helmholtz machine. *Neural Computation* 7(8):889–904.

Fine, S.; Singer, Y.; and Tishby, N. 1998. The hierarchical hidden Markov model: Analysis and applications. *Machine Learning* 32(1):41–62.

Fukushima, K. 1980. Neocognitron: a self organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological Cybernnetics* 36(4):93–202.

Geman, S.; Potter, D.; and Chi, Z. 2002. Composition systems. *Quarterly of Applied Mathematics* LX:707–736.

Geman, S. 2004. Invited talk at CVPR 2004: Invariance and Selectivity in the Ventral Visual Pathway.

George, D., and Hawkins, J. 2005. A hierarchical Bayesian model of invariant pattern recognition in the visual cortex. In *Proceedings of the International Joint Conference on Neural Networks*. Montreal, CA: IEEE.

Grenander, U. 1993. *General Pattern Theory: A Mathematical Study of Regular Structures*. New York, NY: Oxford University Press.

Hinton, G. 2002. Training products of experts by minimizing contrastive divergence. *Neural Computation* 14:1771–1800.

Isard, M. 2003. PAMPAS: real-valued graphical models for computer vision. In *Proceedings of the 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition: Volume I*, 613–620.

Jackendoff, R. 2002. *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford, UK: Oxford University Press.

Johansson, C., and Lansner, A. 2003. Mapping of the BCPNN onto cluster computers. Technical Report TRITA-NA-P0305, Department of Numerical Analysis and Computing Science Royal Institute of Technology, Stockholm, Sweden.

Jordan, M., ed. 1998. *Learning in Graphical Models*. Cambridge, MA: MIT Press.

Lee, T. S., and Mumford, D. 2003. Hierarchical Bayesian inference in the visual cortex. *Journal of the Optical Society of America* 2(7):1434–1448.

Miyashita, Y. 2004. Cognitive memory: Cellular and network machineries and their top-down control. *Science* 306(5695):435–440.

Mountcastle, V. B. 2003. Introduction to the special issue on computation in cortical columns. *Cerebral Cortex* 13(1):2–4.

Mumford, D. 1991. On the computational architecture of the neocortex I: The role of the thalamo-cortical loop. *Biological Cybernetics* 65:135–145.

Mumford, D. 1992. On the computational architecture of the neocortex II: The role of cortico-cortical loops. *Biological Cybernetics* 66:241–251.

Mumford, D. 1994. Neuronal architectures for pattern-theoretic problems. In *Large Scale Neuronal Theories of the Brain*. MIT Press. 125–152.

Mumford, D. 2002. Pattern theory: the mathematics of perception. In *Proceedings of the International Congress of Mathematicians: Volume III*.

Murphy, K.; Weiss, Y.; and Jordan, M. 2000. Loopy-belief propagation for approximate inference: An empirical study. In *Proceedings of the 16th Conference on Uncertainty in Artificial Intelligence*. Morgan Kaufmann.

Newman, M.; Watts, D.; and Strogatz, S. 2002. Random graph models of social networks. *Proceedings of the National Academy of Science* 99:2566–2572.

Pearl, J. 1986a. Fusion, propagation, and structuring in belief networks. *Artificial Intelligence Journal* 29:241–288.

Pearl, J. 1986b. On evidential reasoning in a hierarchy of hypotheses. *Artificial Intelligence Journal* 28:9–16.

Rao, R. P. N., and Ballard, D. H. 1996. Dynamic model of visual recognition predicts neural response properties in the visual cortex. *Neural Computation* 9:721–763.

Ridderinkhof, K.; Ullsperger, M.; Crone, E. A.; and Nieuwenhuis, S. 2004. The role of medial frontal cortex in cognitive control. *Science* 306(5695):443–446.

Riesenhuber, M., and Poggio, T. 1999. Hierarchical models of object recognition in cortex. *Nature Neuroscience* 2(11):1019–1025.

Roth, S., and Black, M. J. 2005. Fields of experts: A framework for learning image priors with applications. In *Proceedings IEEE Conference on Computer Vision and Pattern Recognition*.

Shwe, M.; Middleton, B.; Heckerman, D.; Henrion, M.; Horvitz, E.; Lehmann, H.; and Cooper, G. 1991. Probabilistic diagnosis using a reformulation of the INTERNIST-1/QMR knowledge base I: The probabilistic model and inference algorithms. *Methods of Information in Medicine* 30(4):241–255.

Sporns, O., and Zwi, J. D. 2004. The small world of the cerebral cortex. *Neuroinformatics* 2(2):145–162.

Sudderth, E.; Ihler, A.; Freeman, W.; and Willsky, A. 2003. Nonparametric belief propagation. In *Proceedings of the 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition: Volume I*, 605–612.

Welling, M.; Hinton, G.; and Osindero, S. 2003. Learning sparse topographic representations with products of Student-t distributions. In S. Becker, S. T., and Obermayer, K., eds., *Advances in Neural Information Processing Systems 15*. Cambridge, MA: MIT Press. 1359–1366.

Zemel, R. 2000. Cortical belief networks. In Hecht-Neilsen, R., ed., *Theories of the Cerebral Cortex*. New York, NY: Springer-Verlag.