Sensing with Artificial Tactile Sensors: An Investigation of Spatio-temporal Inference

Asma Motiwala, Charles W. Fox, Nathan F. Lepora, and Tony J. Prescott

Adaptive Behaviour Research Group, Department of Psychology
University of Sheffield, Sheffield, UK
{a.motiwala,c.fox,n.lepora,t.j.prescott}@sheffield.ac.uk

Abstract. The ease and efficiency with which biological systems deal with several real world problems, that have been persistently challenging to implement in artificial systems, is a key motivation in biomimetic robotics. In interacting with its environment, the first challenge any agent faces is to extract meaningful patterns in the inputs from its sensors. This problem of pattern recognition has been characterized as an inference problem in cortical computation. The work presented here implements the hierarchical temporal memory (HTM) model of cortical computation using inputs from an array of artificial tactile sensors to recognize simple Braille patterns. Although the current work has been implemented using a small array of robot whiskers, the architecture can be extended to larger arrays of sensors of any arbitrary modality.

Keywords: Pattern recognition, Cortical computation, Hierarchical Temporal Memory, Bayesian inference, Tactile perception.

1 Introduction

Interacting with any environment, real or artificial, involves extracting meaningful information about features in the environment, often from noisy and ambiguous sensory inputs, to guide behavior and appropriately respond to and/or control the environment. Implementing these operations in artificial systems has been a persistent challenge. However, biological organisms implement these with relatively much greater ease, in much less controlled natural environments. Given how difficult it is to design artificial systems that can perform with similar adaptability and precision with real world problems as biological systems, there is increasing motivation to address these problems using strategies that are algorithmically similar to those used by biological systems to solve similar problems in their natural environments.

Although, generating appropriate behavior and motor outputs have been argued to be the main function of brains in biological systems, they can only follow from an accurate understanding of the *state* of the agent's internal and external environment. Hence, the first step in appropriately interacting with any environment is to use sensory information available through different sensors to extract properties of how things change in the world.

Several algorithms of pattern recognition and feature extraction have been suggested to extract meaningful information from sensory data in different contexts and for different types of problems. It has been previously suggested that no algorithm is inherently superior to other algorithms at solving all learning problems, and that the superiority of any algorithm depends largely on the degree to which the assumptions in the algorithm match the properties of the problem space [11]. Hence, given the efficiency with which biological organisms interact with their environment, it can be assumed that their brains have evolved to implement algorithms that best exploit the properties of real world objects and events.

2   Theoretical Background: Hierarchical Temporal Memory and Cortical Computation

The key problem an organism faces in learning about its environment is to extract consistent and stable features in the environment rather than features in the sensory data itself, *i.e.* the organism needs to learn an internal model of *hidden causes* in the world based on information available from its noisy percepts.

The hierarchical temporal memory (HTM) model is a theory of cortical computation. It is a hierarchical probabilistic model, that uses assumptions based on the properties of objects and events in the world to suggest how the cortex learns internal representations of causes from its inputs, and how it uses these for inference with ambiguous sensory information [10]. Based on the *common cortical algorithm assumption* [16], the HTM model attempts to characterize a generic algorithm with which any arbitrary region of the cortex learns and infers, irrespective of the modality or the level of processing. In other words, it aims to characterize a cortical algorithm that best exploits the statistical properties of causes in the world, using a tight set of constraints regarding the known anatomy and physiological behavior of the cortex.

The Problem of Invariance in Learning Representations: Why the Hierarchy Needs to be Spatio-temporal

The HTM model argues that representing a spatio-temporal hierarchy of causes is critical for the brain's ability to infer with ambiguous sensory information. It has been previously shown in the visual cortex that detecting spatial coincidences in its inputs allows a cortical region to represent more complex spatial patterns than those in lower levels in the cortical hierarchy [14]. However, the problem in learning consistent and stable representations of causes is that a single cause can give rise to several spatially dissimilar percepts. For example: people are exceptionally good at recognizing faces. Face recognition is invariant to the size, position and view of the face, despite the fact that different combinations of these generate widely different visual patterns.

The *temporal slowness principle* argues that causes in the world change slower than noisy percepts [13]. Based on this property of temporal slowness, it has been suggested that the cortex learns to represent invariant and complex representations not just by matching spatial similarity of its inputs but by generalizing over successive spatial patterns in time [8,3]. This *temporal proximity constraint*, that spatial patterns that consistently occur close in time are likely to be associated with the same cause in the world, is a key principle in the HTM model [10].

Spatio-temporal Representations

The cortical hierarchy has been modeled as a hierarchy of HTM nodes where every node implements the same algorithm, and at every level in the hierarchy each node pools information over a sub-set of nodes in lower levels. To maintain a spatio-temporal hierarchy of representations, each HTM node represents recurrent spatial patterns in its inputs, as well as sequences of these spatial patterns. These spatial patterns are patterns of activity in the lower level nodes, that any node pools over (*i.e.* its *children*), that are coincident in time and which recurrently occur in the nodes bottom-up inputs.

According to the temporal proximity constraint, sequences of spatial patterns that consistently occur in sensory inputs are likely to represent a single/coherent cause in the world. Thus by encoding sequences of patterns in its inputs, every node maintains representations of causes that get progressively more complex and invariant as we move up the hierarchy. This is similar to the spatio-temporal properties of cortical representations and mirrors the statistical properties of causes in the world.

Inference

Inference in HTM is probabilistic and implemented using *Bayesian belief propagation* [10]. A key feature of Bayesian probabilistic inference is that it is not only based on estimates of the system's current inputs but uses *prior* information about the statistical properties of patterns in its inputs to compute *posterior probabilities* or *beliefs* over states and causes of the inputs. Probabilistic inference is a powerful way of dealing with the inherent ambiguity in sensory information. Moreover, work on perceptual inference and sensory-motor integration is increasingly suggesting that inference in the cortex is probabilistic [15].

The HTM Model as a Biologically Inspired Technology

Not only is the HTM model a neurobiological theory, that attempts to characterize a cortical algorithm which is hypothesized to best exploit the statistical properties of causes in the world, but it has been developed as a platform to be suitable for several real world applications [18]. The only assumptions in an HTM model are those regarding the nature of the spatio-temporal structure of causes, which are assumed to be invariant features of patterns in the real world, and hence can be applied to a wide range of problems.
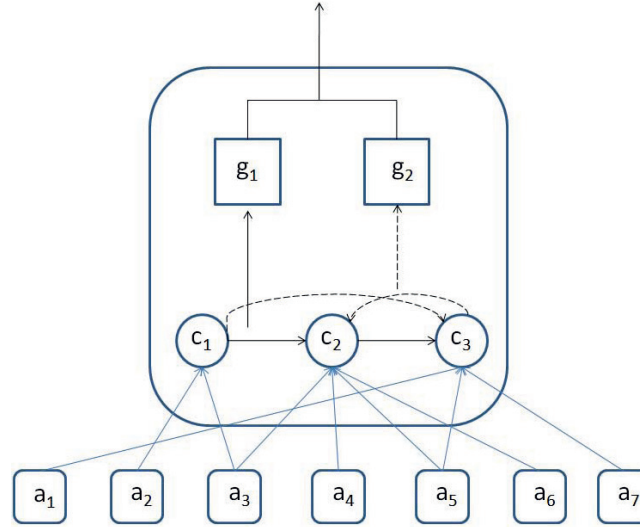
Fig. 1. The figure shows the internal structure of an arbitrary HTM node. The spatial patterns in the node are shown as circles and the sequences as square boxes. Each spatial pattern in the node represents coincident activity in a subset of the node's children and each of the sequences in the node represent sequences of these coincident patterns.

HTM has been previously tested on visual object recognition [10], the model has been extended and tested on sign language recognition [20], as well as for automated design evaluation [12]. HTM has also been tested against standard neural network architectures and has been shown to "produce recognition rates that exceed a standard neural network approach. Such results create a strong case for the further investigation and development of Hawkins' neocortically-inspired approach to building intelligent systems" [21]. In addition, HTM is already being commercially used by Vitamin D, EDSA and Forbes.com, and several other applications are being developed [2].

3   Methods

The current work aimed to investigate how inference using the HTM architecture could be implemented to recognize simple braille patterns using sensory inputs from an array of artificial whiskers.

Probabilistic inference and hierarchical representations are not new concepts, in either biology or pattern recognition. Cortical regions have been shown to be connected in a hierarchical manner [7] and several neurobiological models as well as statistical models of pattern recognition use hierarchical organization [17]. In addition, the idea that inference in the cortex is probabilistic has been popularized as the *Bayesian brain* hypothesis and belief propagation has generated

great interest as a biologically plausible inference strategy [5]. However, HTM is unique in that it uses spatio-temporal hierarchical representations for probabilistic inference, in contrast to most neurobiological and statistical models which either have spatial or temporal structure [17].

This spatio-temporal architecture was particularly well suited for the investigation of tactile perception since neither the spatial nor the temporal components in the sensory data can be ignored. The amount of information available from any single tactile sensor at any given instant is extremely sparse. Hence it is critical to pool over multiple sensors as well as consider how the inputs evolve over time to learn and infer about the properties of different stimuli.

Data Acquisition

The sensory input to the model was obtained from an array of three artificial whisker sensors that have been developed on the BIOTACT project [1]. The stimuli used were braille-like patterns (shown in Figure 2) and the size of the braille patterns was chosen to match the spatial resolution of the sensor array. The stimulus was swept over the whisker sensors using the XY positioning robot in a plane perpendicular to the shaft of the whiskers. The arrangement of the whiskers and the stimulus being swept over it are shown in Figure 2.
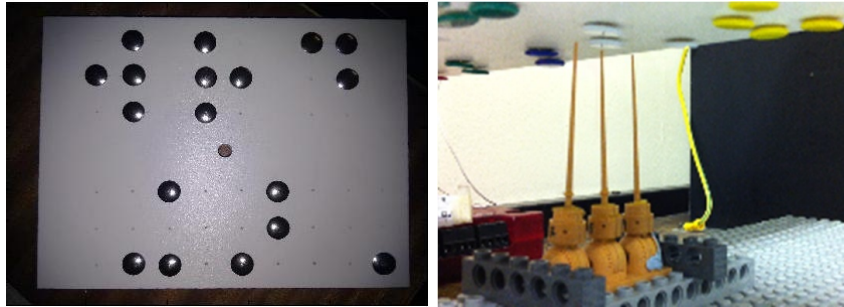

Fig. 2. The left panel shows an example of a braille pattern used as a stimulus. The right panel shows how the robot whiskers were set up and how the braille stimulus was swept over them for the current implementation.

The sensor, which has been previously described in [6], detects deflection in the whisker from its resting position on two orthogonal axes x and y and converts it into two voltage signals which are proportional to the whisker displacement on the two axes. For the current implementation, the displacement amplitude in the whiskers, when the stimulus was moved over them, was used. The signal from the sensors was segmented and an *activation* $a_m(t)$, for each sensor m at timestep t, was defined as a function of the displacement amplitude $d_m(t)$ over the tth segment,

$$a_m(t) = \frac{1}{1 + e^{-d_m(t)}} \tag{1}$$

where $d_m(t)$ is the deflection in the whisker sensor m at time t and $a_m(t)$ is the corresponding activation of the sensor. The whisker deflection signal was segmented and converted into an activation for each sensor to simplify the implementation of the model. The duration of the time window for each segment was defined to match the relative frequency with which non-zero patterns tend to occur in the inputs. Defining the activation function is not critical to the working of HTM, but since braille patterns are essentially binary, this transformation was used to make the *activation patterns* more representative of the characteristics of the stimuli (as shown in the lower panel in Figure 3).

The arrangement of the whisker sensors used in the current implementation is similar to the way in which the microvibrissae of the rat are used to explore patterns for shape recognition (see Figure 4). It has been shown in behavioral studies that the microvibrissae are critically involved in object/pattern recognition, as
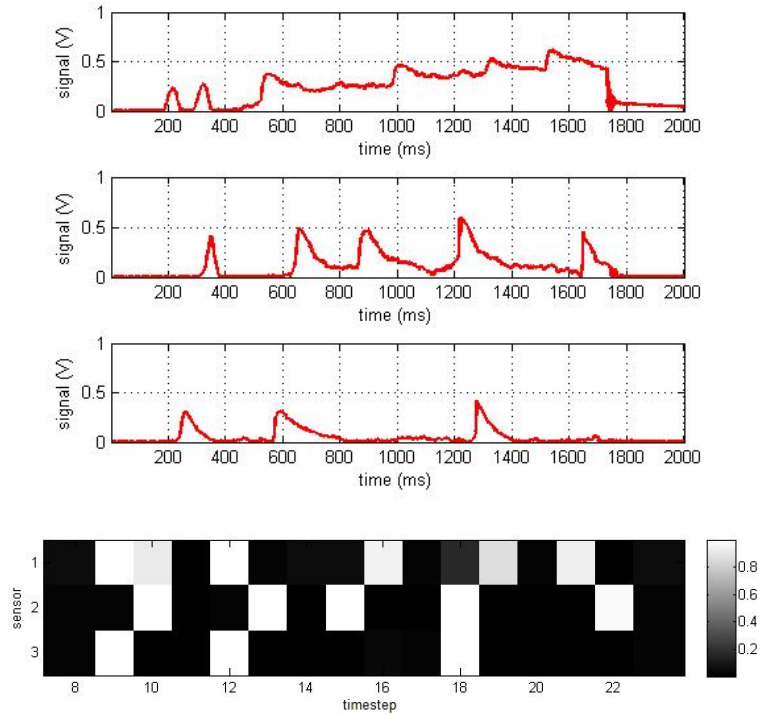

Fig. 3. The upper panels in the figure shows an example of a set of unprocessed signals from the whisker sensors. The top panel is an example of noisy sensory data, corrupt with a significantly large low frequency component and the second panel is an example of data which clearly reflects the features in the stimulus. These whisker deflection timeseries were filtered to remove low frequency components, segmented and the corresponding activation patterns for each sensor at timesteps corresponding to the segments of the timeseries are shown in the lower panel.
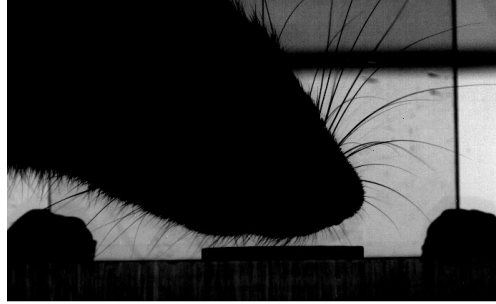
Fig. 4. The figure shows an image of a rat using its microvibrissae to explore a coin

opposed to the role of the macro-vibrissae in object localization [4]. And hence, the current implementation of HTM is also being suggested as a model of processing in the primary sensory cortex associated with the microvibrissae.

HTM Implementation

Inference in HTM nodes is with probabilistic message passing based on the Bayesian Belief Propagation algorithm [19] and the complete mathematical description of inference in HTM has been previously presented [10].

A key feature in Bayesian probabilistic inference is that it is based not only on current evidence, but also past evidence to resolve ambiguity. To this end, current evidence, of the form of a *likelihood* distribution, is combined, using the
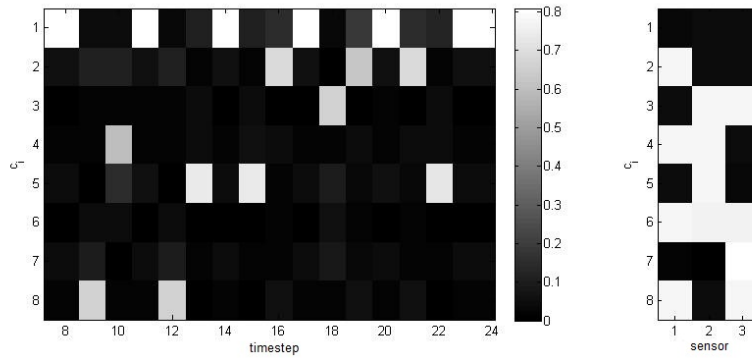

Fig. 5. The left panel in the figure shows the likelihood $y(i)$ which the node receives over its coincidence patterns. Each column in the image represents a probability distribution over spatial patterns c_i for each time step. The right panel shows the activation expected in each sensor corresponding to each of the spatial patterns c_i represented in the node.

Bayes Theorem, with a *prior* distribution, to obtain a *belief* distribution over states or events.

The feedforward input to the HTM node at each time step is the likelihood of the spatial pattern of activation in the whisker sensors that the node pools over. The activation patterns can be represented in m-dimensional space, where m is the number of sensors/children the spatial pattern is detected over. The likelihood y can then be defined by soft classification of data points in this space into clusters associated with the node's spatial patterns using any measure of degree belonging of data points to each of the clusters,

$$y_t(i) = P(A(t)|c_i) \tag{2}$$

where c_i represents coincident pattern i and $A(t) = a_1(t), a_2(t)...a_m(t)$. The coincident patterns represented in the node correspond to sub-characters in the braille pattern.

The likelihood y is combined with a prior distribution $\pi(c)$ to obtain a belief $Bel(c)$ over the node's spatial patterns.

$$Bel_t(c_i) = \frac{1}{Z} y_t(i)\pi_t(c_i) \tag{3}$$

where Z is the normalization constant.

Sequences represented in the node are assumed to be Markov chains, *i.e.* they are sequences of states in which the state of the system at any given time t is dependent on the previous states of the system, hence, the prior $\pi(c)$ and the likelihood $\lambda(g)$ are computed from the prior $\pi(g)$ and the likelihood y, respectively, using transition probabilities between the node's spatial patterns.

$$\pi_t(c_i) = \sum_{c_j} \sum_{g_r} P(c_i(t)|c_j(t-1), g_r)\pi(g_r), \tag{4}$$

and

$$\lambda_t(g_r) = P(A(0 \rightarrow t)|g_r) = \sum_{c_i} y(i) \sum_{c_j} P(c_i(t)|c_j(t-1), g_r) \tag{5}$$

where $P(c_i(t)|c_j(t-1), g_r)$ is the transition probability represented within the node. All the internal representations of the node were predefined in the current implementation.

This likelihood $\lambda_t(g)$ is the feedforward message that the node sends to the node above it in a hierarchy (*i.e* its *parents*) and the prior, which is sent to the node's children, is a function of the belief $Bel_t(c)$.

4   Results

For the current implementation the working of a single HTM node was instantiated. The coincidence patterns in the node represent single columns of Braille alphabets (see Figure 2), and the sequences represent entire Braille alphabets.
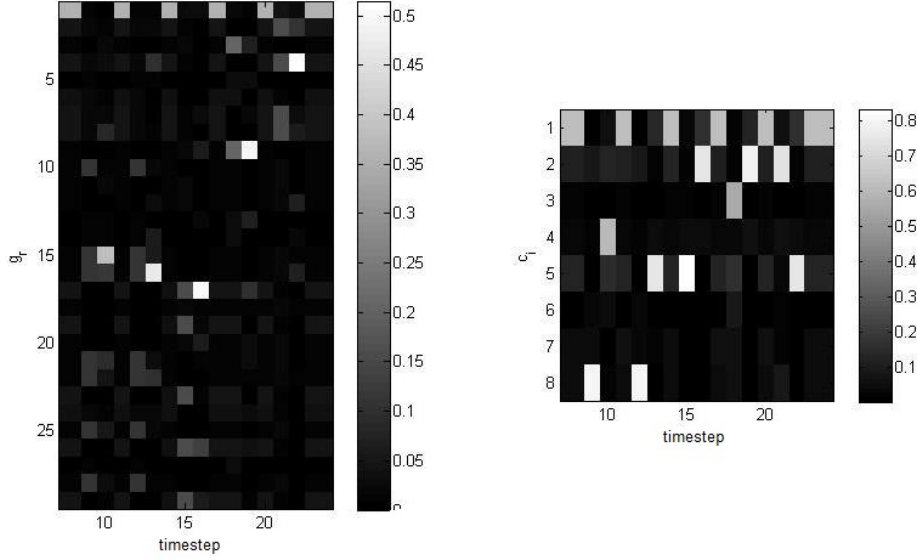
Fig. 6. The right panel in figure shows the node's belief over its coincidence patterns $Bel(c)$ and the left panel shows the feedforward likelihood over its sequences $\lambda(g)$. Each column in both images represent a probability distribution at that time step.

The results from the current implementation show how the probability distributions over the node's spatial patterns and sequences, *i.e.* the distributions over the nodes internal representations, change over time in response to the inputs from the whisker sensors (as shown in Figure 6).

Since the spatial patterns and sequences were predefined in the node, the accuracy with which the node inferred could be easily verified by matching the patterns in the node, over which the probability distributions peaked, with the patterns in the stimulus. For example, the first spatial pattern as well as the first sequence represented in the node correspond to blank patterns in the stimulus. The time steps at which the distributions peak on the first pattern indicate that the the node has inferred blank patterns in the stimulus at those time steps (see Figure 5). It is evident from number of time steps between blank patterns that sequences of non-zero patterns in the data are only two patterns long. This is a true of all braille characters which are formed by precisely arranged raised dots which occur only in six positions (two columns of three rows each) for each character (as shown in Figure 2). In other words, the patterns associated with the peaks in the belief distribution over the node's spatial patterns and the likelihood over its sequences match those in the stimulus, *i.e.* the patterns over which the node infers the highest degree of certainty match the actual patterns in the stimuli.

The results also show that the likelihood distribution over the sequences has a relatively high level of ambiguity, *i.e.* it is relatively broad, on the first time step when any non-zero sequence begins in the data. However, the distribution over the sequences sharpens over the next time step. Even in the general case, with sequences of any arbitrary length, it would be expected that the distribution over the node's sequences would become progressively sharper as more evidence is accumulated. In other words, since it is assumed that causes can share common sub-components, the degree of certainty in the state of a cause increases as more evidence becomes available.

5   Discussion

The sensor array, stimulus as well as the model architecture were all kept minimal in the current implementation to investigate thoroughly the inner workings of the HTM model. Although the data set of stimulus patterns used here was very simple, the spatio-temporal structure of braille patterns as a tactile stimuli is sufficiently rich to use them to test pattern recognition in HTM. The results show that the discussed architecture can successfully use spatio-temporal patterns in inputs to infer complex causes. This architecture can further be extended to receive sensory inputs from a larger array of sensors and by implementing a more elaborate hierarchy of nodes, to represent and infer with more complex spatio-temporal patterns. It is anticipated that insights from the current work are likely to be instrumental, within the general framework of developing robotic multiwhisker touch, for recognizing complex behaviorally relevant objects and patterns.

Key Challenges in Using HTM

For the current implementation, since the structure of the stimulus used was extremely simple, it was relatively straightforward to predefine the internal representations in the HTM node with which inference was implemented. However, a key challenge with more extended HTM hierarchies is that the nature of the node's internal representations become less intuitive as we go further up in the hierarchy. Hence, it becomes crucial that adequate learning mechanisms are defined for the model to extract these representations directly from its inputs.

The temporal proximity constraint is a powerful idea in learning invariant representations of stable causes in the world based on rapidly changing, ambiguous sensory inputs. Pooling over local spatial coincidences in a hierarchical arrangement allows representing progressively more complex spatial patterns from simpler components/features. By analogy, patterns of transitions between spatial patterns can be used to learn hierarchical representations of larger temporal patterns in inputs to represent stable/invariant features of causes in the world. However, learning complex sequences from transition probabilities between spatial patterns is not the same as learning complex spatial patterns from local coincidences. Information about inputs at previous time steps is not available to

the system from successions in the same sort of way as is the information about larger spatial patterns from local spatial coincidences. Hence, learning associations between complex sequences of patterns and higher level causes is still a key challenge.

Finally, it has been pointed out that, "if it were not for the fact that we can see, we might reasonably think it is impossible" [9]. This is true not only for vision but for several problems that nature seems to have found elegant solutions for and this becomes strikingly evident in trying to replicate these functions in artificial systems. Biomimetic robotics is a powerful tool to reverse engineer solutions available in biological systems, that have evolved to efficiently handle several hard problems, and to optimize these strategies for specific problems of interest.

References

1. Biotact consortium, `http://www.biotact.org`
2. Numenta customers, `http://www.numenta.com/about-numenta/customers.php`
3. Berkes, P., Wiskott, L.: Slow feature analysis yields a rich repertoire of complex cell properties. Journal of Vision 5(6) (2005)
4. Brecht, M., Preilowski, B., Merzenich, M.: Functional architecture of the mystacial vibrissae. Behavioural Brain Research 84(1-2), 81–97 (1997)
5. Doya, K.: Bayesian brain: Probabilistic approaches to neural coding. The MIT Press, Cambridge (2007)
6. Evans, M., Fox, C.W., Pearson, M.J., Lepora, N.F., Prescott, T.J.: Whisker-object contact speedaects radial distance estimation. In: IEEE International Conference on Robotics and Biomimetics (2010)
7. Felleman, D.J., Van Essen, D.C.: Distributed hierarchical processing in the primate cerebral cortex. Cerebral Cortex 1(1), 1 (1991)
8. Földiák, P.: Learning invariance from transformation sequences. Neural Computation 3(2), 194–200 (1991)
9. Frisby, J., Stone, J.: Seeing: the computational approach to biological vision. The MIT Press, Cambridge (2009)
10. George, D., Hawkins, J.: Towards a mathematical theory of cortical micro-circuits. PLoS Comput. Biol. 5(10), e1000532 (2009)
11. George, D.: How to make computers that work like the brain. In: Proceedings of the 46th Annual Design Automation Conference, DAC 2009, pp. 420–423. ACM, New York (2009)

12. Hartung, J., McCormack, J., Jacobus, F.: Support for the use of hierarchical temporal memory systems in automated design evaluation: A first experiment. In: Proceedings of the ASME 2009 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference, San Diego, CA, USA (2009)
13. Hinton, G.: Connectionist learning procedures. Artificial Intelligence 40(1-3), 185–234 (1989)
14. Hubel, D., Wiesel, T.: Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. The Journal of Physiology 160(1), 106 (1962)
15. Knill, D., Pouget, A.: The Bayesian brain: the role of uncertainty in neural coding and computation. Trends in Neurosciences 27(12), 712–719 (2004)
16. Mountcastle, V.: An organizing principle for cerebral function: The unit model and the distributed system. In: Edelman, G., Mountcastle, V. (eds.) The Mindful Brain. MIT Press, Cambridge (1978)
17. Numenta: Hierarchical temporal memory: Comparison with existing models. Tech. rep., Numenta (2007)
18. Numenta: Problems that fit htm. Tech. rep., Numenta (2007)
19. Pearl, J.: Probabilistic reasoning in intelligent systems: networks of plausible inference. Morgan Kaufmann, San Francisco (1988)
20. Rozado, D., Rodriguez, F., Varona, P.: Optimizing hierarchical temporal memory for multivariable time series. In: Diamantaras, K., Duch, W., Iliadis, L.S. (eds.) ICANN 2010. LNCS, vol. 6353, pp. 506–518. Springer, Heidelberg (2010)
21. Thornton, J., Gustafsson, T., Blumenstein, M., Hine, T.: Robust character recognition using a hierarchical Bayesian network. In: Sattar, A., Kang, B.-h. (eds.) AI 2006. LNCS (LNAI), vol. 4304, pp. 1259–1264. Springer, Heidelberg (2006)