

Bidirectional Associative Memory for Short-term Memory Learning

Christophe Tremblay (ctrem040@uottawa.ca)

School of Psychology, 136 Jean-Jacques Lussier
University of Ottawa, ON K1N 6N5 Canada

Nareg Berberian (nberb062@uottawa.ca)

School of Psychology, 136 Jean-Jacques Lussier
University of Ottawa, ON K1N 6N5 Canada

Sylvain Chartier (sylvain.chartier@uottawa.ca)

School of Psychology, 136 Jean-Jacques Lussier
University of Ottawa, ON K1N 6N5 Canada

Abstract

Previous research has shown that Bidirectional Associative Memories (BAM), a special type of artificial neural network, can perform various types of associations that human beings are able to perform with little effort. However, considering a simple association problem, such as associating faces with names, iterative type BAM networks usually take hundreds and sometimes thousands of learning trials to encode such associations correctly, whereas humans in some conditions learn much faster. The present study therefore proposes an adjustment to a particular type of BAM network that increases its performance in a rapid learning condition while processing memory capacity is limited. Results show that the modification to the original learning rule of the BHM leads to improved performance when rapid learning is required. Moreover, the model preserves its high memory load capacity in standard learning. This study could lead to improved cognitive models that can adapt their behavior in function of the contextual conditions.

Keywords: Artificial intelligence; Connectionist models; Bidirectional associative memory; Short-term memory learning.

Introduction

Like human beings, artificial neural networks can discriminate, identify, and categorize perceptual patterns (Faussett, 1994; Haykin, 2009). Bidirectional Associative Memories (BAMs) have been proposed as models of neurodynamics. As such, they are able to develop attractors allowing the network to perform various types of recall under noisy conditions and successfully carry out pattern completion. Many advances were made since Kosko introduced the BAM in 1988; see Acevedo-Mosqueda, Yanez-Marquez & Acevedo-Mosqueda (2013) for a review. Although the majority of improvements made BAM models better suited for learning a wider set of associations and brought greater performance, such improvements have been at the expense of learning times. Nowadays, BAM models often take hundreds and sometimes thousands of learning cycles to correctly associate patterns. Consequently, researchers have attempted to enhance convergence times and performance (Nong & Bui, 2012). Several one-shot learning rules were developed, showing almost perfect

performance in some cases (Ritter, Diaz-Deleon, & Sussner, 1999; Wu & Pados, 2000; Acedo, Yanez, & Lopez, 2006a; 2006b). Although these improvements have successfully overcome the limitations of the model presented previously, these models were most often applied to engineering problems rather than providing legitimate models of cognition. For example, good connectionist models of cognition should adapt their connection weights locally, without homuncular knowledge, backpropagation of the error or complex learning procedure (O'Reilly, 1998). Therefore, models should show the ability to learn quickly with excellent performance, while possessing limited memory capacity when encoding time is short (one shot or few shot learning). On the other hand, they should present good performance and good memory load capacity when the encoding time allowed is longer.

The present study proposes the introduction of a recency parameter in a modified BAM network (Chartier & Boukadoum, 2006b; 2011), which leads to improved performance in rapid learning situations, but with limited storing capacity. The proposed process therefore allows for learning in conditions where a limited number of iterations is required. The remainder of the paper is divided as follows: The background section describes the architecture of the model and the internal properties of the network, while the method section presents the learning and recall procedures of the model. The results section describes the recall performance tasks of the model under various level of noise. The final section discusses the results and provides a conclusion to the study.

Background

The model proposed by Chartier & Boukadoum (2006b; 2011) uses an unique matrix for each layer. This Bidirectional Heteroassociative Memory (BHM) is able to learn correlated patterns for bipolar patterns as well as for real-valued patterns.

Architecture

The network is made of two Hopfield-like neural networks interconnected in a head-to-tail fashion, providing a

recurrent flow of information that is processed bidirectionally. The network's architecture is shown in Figure 1 where $\mathbf{x}(0)$ and $\mathbf{y}(0)$ represent the initial vector-states, \mathbf{W} and \mathbf{V} are the weight matrices and t is the current iteration number.

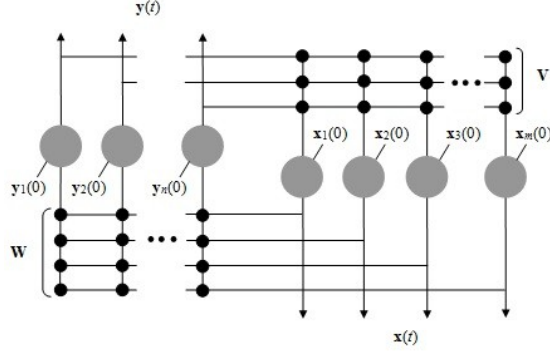


Figure 1: Architecture of the BAM.

Transmission Function

The transmission function is based on the classic Verhulst equation extended to a cubic form with a saturating limit at ± 1 (Chartier & Boukadoum, 2006b):

$$\forall i, \dots, N, y_{i[t+1]} = f(\mathbf{a}_{i(t)}) \quad (1a)$$

$$= \begin{cases} 1, & \text{if } \mathbf{a}_{i(t)} > 1 \\ -1, & \text{if } \mathbf{a}_{i(t)} < -1 \\ (\delta + 1)\mathbf{a}_{i(t)} - \delta \mathbf{a}_{i(t)}^3, & \text{else} \end{cases}$$

and

$$\forall i, \dots, M, x_{i[t+1]} = f(\mathbf{b}_{i(t)}) \quad (1b)$$

$$= \begin{cases} 1, & \text{if } \mathbf{b}_{i(t)} > 1 \\ -1, & \text{if } \mathbf{b}_{i(t)} < -1 \\ (\delta + 1)\mathbf{b}_{i(t)} - \delta \mathbf{b}_{i(t)}^3, & \text{else} \end{cases}$$

where N and M are the number of units in each layer, i is the unit index, δ is a general transmission parameter and \mathbf{a} and \mathbf{b} are the activations. These activations are obtained the usual way: $\mathbf{a}(t) = \mathbf{W}\mathbf{x}(t)$ and $\mathbf{b}(t) = \mathbf{V}\mathbf{y}(t)$. Figure 2 illustrates the shape of the transmission function for $\delta = 0.2$.

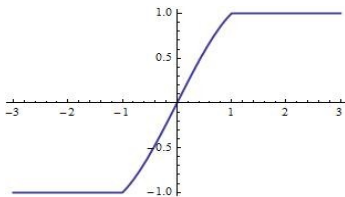


Figure 2: Transmission function for $\delta = 0.2$.

Contrary to sigmoid-type function, there are no asymptotic behaviors in the transmission. This function has the advantage of exhibiting grey-level attractor behaviour

which contrasts with other BAMs that can only develop bipolar attractors (Chartier & Boukadoum, 2006b).

Learning Rule

The connection weights are modified following a Hebbian/anti-Hebbian approach (Storkey & Valabreque, 1999; Bégin & Proulx, 1996):

$$\begin{aligned} \mathbf{W}(k+1) &= \mathbf{W}(k) + \eta(\mathbf{y}(0) - \mathbf{y}(t))(\mathbf{x}(0) + \mathbf{x}(t))^T \\ \mathbf{V}(k+1) &= \mathbf{V}(k) + \eta(\mathbf{x}(0) - \mathbf{x}(t))(\mathbf{y}(0) + \mathbf{y}(t))^T \end{aligned} \quad (2)$$

where η is the learning parameter controlling for the speed of convergence and k is the learning trial number. Connection weights are initiated at 0 and $\mathbf{x}(0)$ and $\mathbf{y}(0)$ are the initial inputs to be associated. The network has converged when $\mathbf{x}(0) = \mathbf{x}(t)$ or $\mathbf{y}(0) = \mathbf{y}(t)$. Thus, each weight matrix converges when the feedback is equal to the initial inputs. In the BHM, the network convergence is guaranteed if the learning parameter η is set according the following condition (Chartier & Boukadoum, 2006a):

$$\eta < \frac{1}{2(1-2\delta)\max[M,N]}, \quad \delta \neq \frac{1}{2} \quad (3)$$

where M and N are respectively the dimensionality of the input and its association. The η parameter was set to a lower value than the threshold found in (3) for every simulation performed. The learning rule (2) acts much like a long-term memory where the learning convergence is longer, but exhibits an increased storage capacity and has a better-defined attractor.

Learning Rule Modification

In order to lower the time to learn associations, the memory capacity has to be decreased. One way to accomplish this is by introducing a recency parameter ($0 \leq \beta \leq 1$). This parameter removes from the memory associations that are not reinforced enough. The resulting learning rule after modification is given by:

$$\begin{aligned} \mathbf{W}(k+1) &= \beta \mathbf{W}(k) + \eta(\mathbf{y}(0) - \mathbf{y}(t))(\mathbf{x}(0) + \mathbf{x}(t))^T \\ \mathbf{V}(k+1) &= \beta \mathbf{V}(k) + \eta(\mathbf{x}(0) - \mathbf{x}(t))(\mathbf{y}(0) + \mathbf{y}(t))^T \end{aligned} \quad (4)$$

If $\beta=1$ then the learning is accomplished in the same fashion as in Equation 2. This learning rule can be simplified to the following hebbian/anti-hebbian equation in the case of auto association where $\mathbf{y}(0) = \mathbf{x}(0)$:

$$\begin{aligned} \mathbf{W}(k+1) &= \beta \mathbf{W}(k) + \eta(\mathbf{x}(0)\mathbf{x}(0)^T - \mathbf{x}(t)\mathbf{x}(t)^T) \\ \mathbf{V}(k+1) &= \beta \mathbf{V}(k) + \eta(\mathbf{y}(0)\mathbf{y}(0)^T - \mathbf{y}(t)\mathbf{y}(t)^T) \end{aligned} \quad (5)$$

Simulations

The simulations are to assess the performance of the Short-term BHM on a recall task in comparison to the standard Hopfield, Kosko and BHM networks.

Methodology

Learning was carried out according to the following procedure:

- 1) Random selection of a pair of patterns ($\mathbf{x}(0)$ and $\mathbf{y}(0)$).
- 2) Computation of $\mathbf{x}(t)$ and $\mathbf{y}(t)$ according to the transmission function (1).
- 3) Computation of the weight matrices update according to (4).
- 4) Repetition of steps 1) to 3) until all of the pairs have been presented.
- 5) Repetition of steps 1) to 4) for an *a-priori* set number of epochs.

The transmission parameter (δ) was set to 0.2 throughout the simulations and the number of iterations to perform by the network before the weight matrices were updated was set to $t = 1$. The network was tested on an auto-association and hetero-association task that consisted of 26 stimuli placed on 7x7 grids (Figure 3). The auto-association task was an association of uppercase stimuli only, whereas the hetero-association consisted of the association between uppercase and lower case stimuli. The recency parameter (β) was set to 0.99 and 0.995 for the rapid setting and at 1.0 for the standard long-term setting. In the rapid setting, instead of presenting all the patterns at once, the network was limited to only one subset at a given time. In other words, rather than learning all stimuli in one epoch, the network limited itself to grouped associations of a maximum of 5 associations.

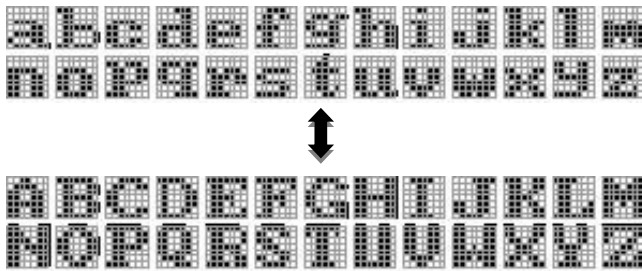


Figure 3: Patterns used for the simulation

Following the learning phase, the network was tested on a recall task according to the following procedure:

- 1) Selection of an input pattern $\mathbf{x}(0)$.
- 2) Computation of $\mathbf{y}(1)$ according to the transmission function (1).
- 3) Comparison with the target value $\mathbf{y}(0)$
- 4) Repetition of steps 1) to 3) until all of the patterns have been presented.

In this situation a given pattern iterated until a steady state. Recall performance was recorded for the level of flipped pixels varying from 0 to 24 (0 to $\approx 50\%$). The network was tested on grouped associations of 2, 3, 4 and 5 patterns. The network was tested 200 times for every condition and the average performance was computed.

Results

Figure 4 presents an example of the first 10 patterns recalled in a noiseless (0 flipped pixel) situation for both auto-association and hetero-association tasks. The orange dashed lines represent the demarcation between previously learned associations and the associations that have just been learned. The model was compared to the results of Hopfield's model (1982) as well as Kosko's (1988). For both networks, contrary to the BHM, there are no memory traces between the past and current association. In other words, the connection weights are reset to zero between the learning of a given group. The connection weights had to be set to zero since both Hopfield and Kosko's model cannot perform the task otherwise as they suffer from memory overload. It is as if we are comparing the performance of a single BHM with several independent Hopfield or Kosko models. Although this situation is different, it was included for comparison purposes using optimal conditions for Hopfield and Kosko.

The results (Figure 4) for the auto-association of the short-term memory show that previously learned associations tend to be erased as new associations are made, particularly when the correlation is very high between two patterns (for example, the stimulus E and F). When patterns are presented in groups of two, the short-term network makes no mistakes in associating the patterns presented within the step; this also holds for conditions where patterns are presented in groups of 5. The Hopfield network shows perfect performance when the input patterns are learned in groups of two. However, when presented in groups of 5, the network makes several mistakes even in the absence of noise. These results are even more disastrous for hetero-associations, where the network can barely recall any associations. Hence, Kosko's network is not able to learn any of the associations grouped in pairs, whereas the short-term BAM is able to learn all associations whether they are presented in groups of 2 or 5. Results for the standard BHM were not shown because it could learn and recall perfectly in all of the previous situations.

Figure 5 depicts Monte Carlo simulations of the network performance on recall tasks with flipped pixels. As can be seen in figure 5a and 5b, the short-term memory outperforms the standard BHM. However, the results for the independent Hopfield networks with sequences of 2 and 4 items are very similar to the results of the short-term version of the BHM. Results for sequences of 5 inputs are not reported since it was shown in Figure 4 that the Hopfield networks could not perfectly recall more than 4 associations even in a noiseless condition. In addition, more epochs (15 rather than 5) systematically lead to increased performance in the BHM. In other words, using the batch of training sets for a greater number of cycles before updating the weight matrices leads the short-term BHM to better recall performance. For example, the performance for input sequences of 5 improved if the number of epoch is increase from 5 to 15 epochs. However, for the remaining conditions, the number of epochs has little or no impact on the performance. Evidently, because

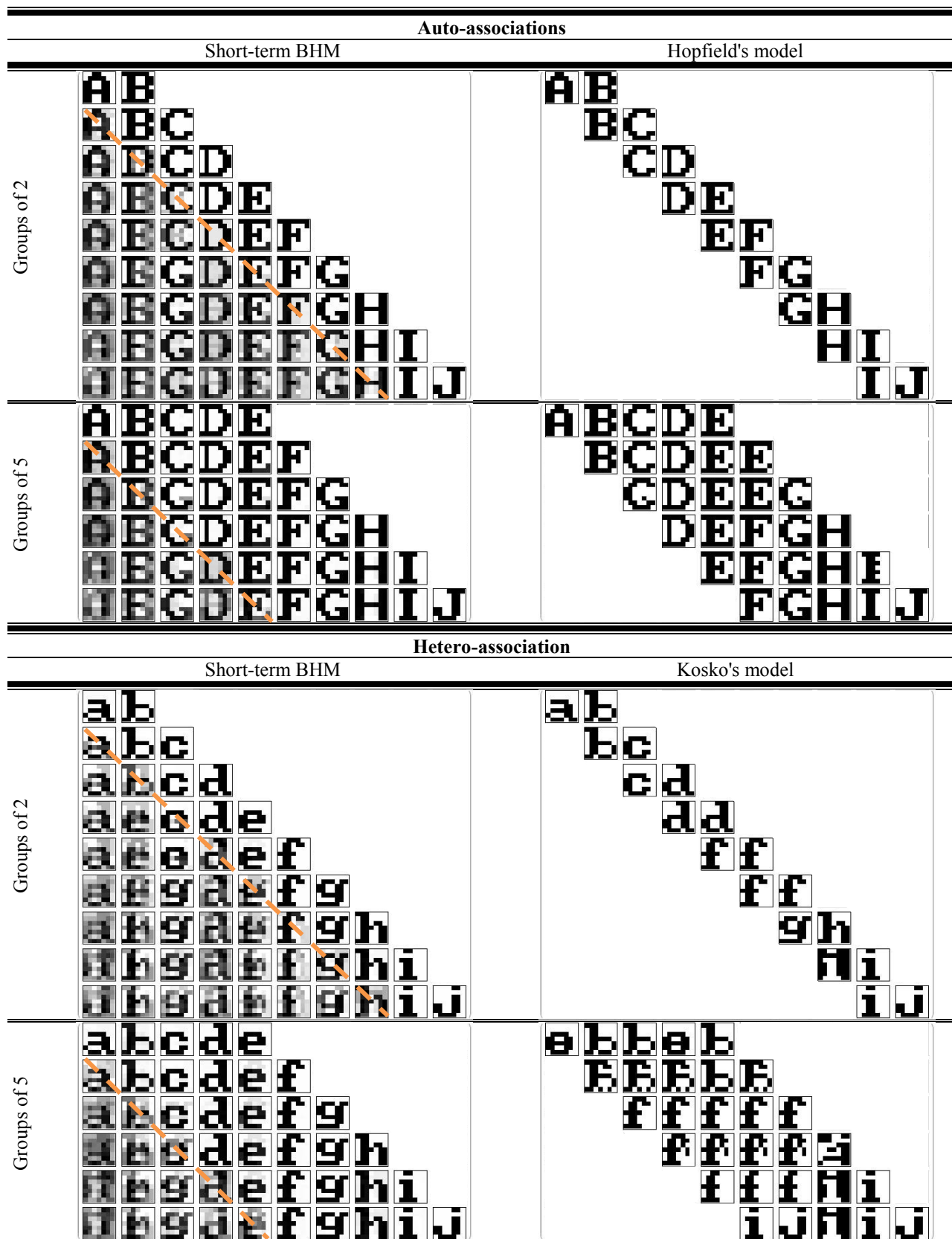


Figure 4: Association recall for auto and hetero-association learning

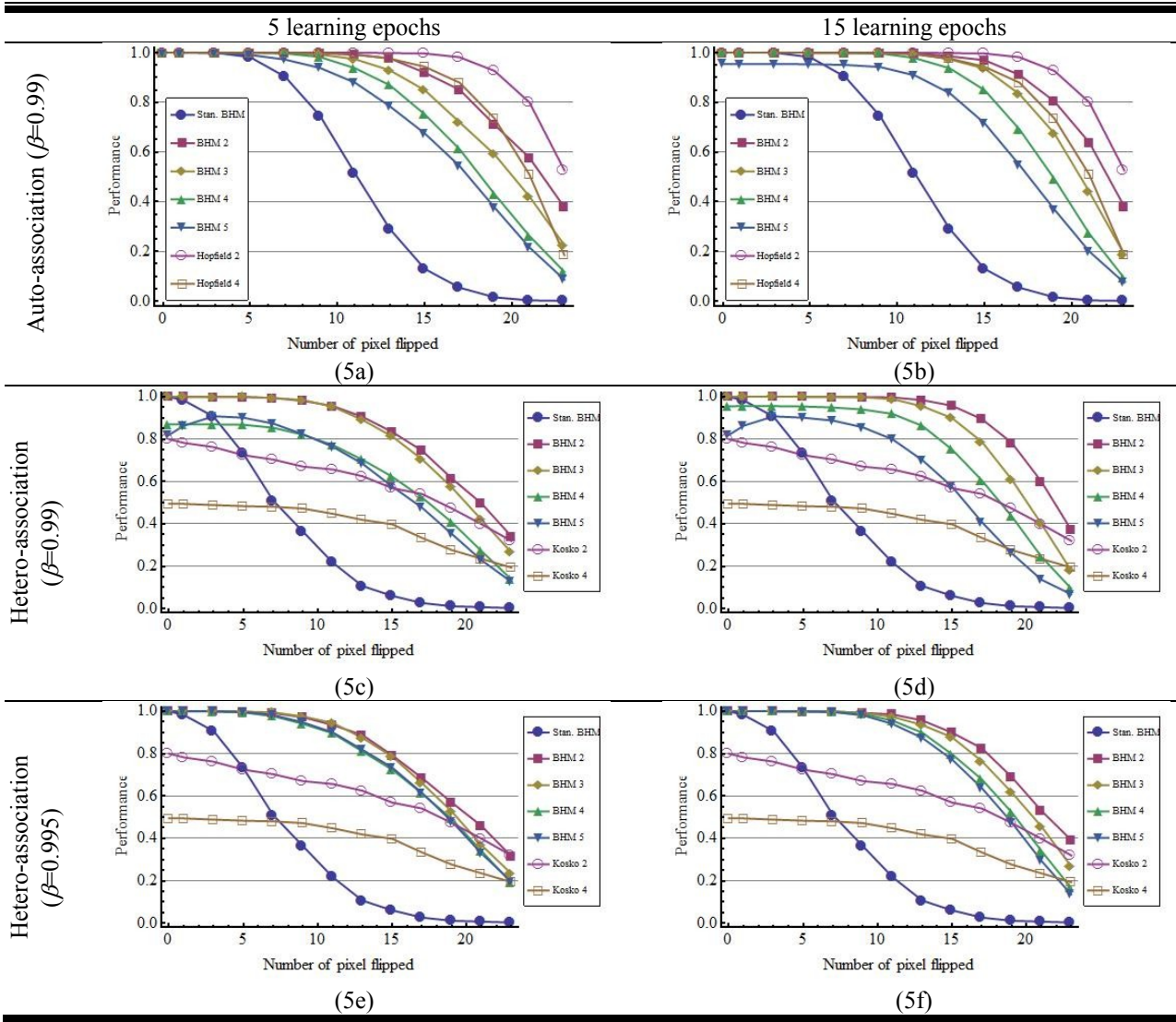


Figure 5: Performances of the network on a recall task with pixels flipped. The legend is laid out as follow: Stan. BHM are the results for a standard BHM. BHM 2, BHM 3, BHM 4 and BHM 5 are the results for the short-term version of the BHM for input sequences of respectively 2, 3, 4 and 5 patterns. Finally, Hopfield and Kosko are the results for the Hopfield and Kosko networks for input sequences of respectively 2 and 4 patterns

Hopfield network uses a strict Hebbian learning rule, anything other than one epoch will lead to worse results (Bégin & Proulx, 1996). Finally, for both the Hopfield and BHM network, the performance is reduced as the lengths of the sequences are increased.

Similar results hold for the more difficult task of hetero-associations (5c and 5d) where a trend similar to what was presented for auto-associations is presented. However, in some conditions, such as simulations performed on sequences of 4 and 5 inputs for the short-term BHM, the network could not systematically recall the task 100% of the time (5c and 5d). We therefore tested the network using a value of $\beta = 0.995$, which led to perfect performance on recalls with little or no noise (5e and 5f). Moreover, the results for Kosko's BAM are drastically reduced in comparison to the previous results on auto-

associations. In fact, the independent networks cannot learn perfectly sequences of 2 nor 4 associations. In no condition could Kosko's BAM perform as well as the standard or the short-term BHM.

Discussion and Conclusion

The results show that the modification to the learning rule of the standard BHM leads to increased performance in conditions where rapid learning is necessary. Those results are similar to other research done in the field of bidirectional associative memory, which led to perfect performance on recalls with little or no noise (5e and 5f). Moreover, the results for Kosko's BAM are drastically reduced in comparison to the previous results on auto-associations. In fact, the associative memories, such as Nong & Bui (2012), in which a BAM model was proposed

with both faster convergence and improved performance. However, the proposed network remains a compelling model of human cognition as it retains the original properties of the standard BHM network (Chartier & Boukadoum, 2006), which was not the case in Nong & Bui (2012). In other words, the same BHM model can be used for both short-term and long-term memory encoding. The results show that the use of short-term memory is advantageous for short sequences of inputs for both auto- and hetero-associations.

Although the objective of the study was not to model human data, the fact that traces of older memories remain encoded in the connection weights is an interesting component to the model since real-time problem solving, as constantly done by human beings, requires memory traces to be kept. This could lead to a priming effects, where learning performance could increase, resulting in a faster learning during long-term memory encoding. On the other hand, erasing such information, as is done in Hopfield's and Kosko's models, would force for constant relearning of the same associations. The BHM network proposed also exhibits some advantages in comparison to the Hopfield and Kosko networks as the BAM network is much better a hetero-association and can tackle longer input sequences.

In conclusion, this study introduces a modification to the original learning rule of the BHM, which leads to improved performance when rapid learning is required. It showed that BAM models are well suited to perform associations using a faster learning phase. Hence, they appropriately encompass biological and environmental limitations. Future research should evaluate the performances when the sequence is set randomly between epochs (i.e. abc, bac, acb). In addition, future research should explore the dynamics obtained as β is varied systematically under various dimensions of stimuli. The parameter could allow for more freedom on the number of items that can be kept in memory, which could vary in relation to contextual settings. These results are a step towards the construction of a larger model of working memory, since it is shown that the encoding of shorter sequences of input patterns leads to a different dynamic. Therefore, artificial neural networks used for real time problem solving (such as in robotics for example) should have a short-term memory component as well as a long-term memory.

Acknowledgments

This research was partly supported by the Natural Sciences and Engineering Research Council of Canada. The authors would like to thank Kaia Myers-Stewart for her comments on an earlier version of this text.

References

Acevedo-Mosqueda, M. E., Yanez-Marquez, C., & Acevedo-Mosqueda, M. A. (2013). Bidirectional associative memories: Different approaches. *ACM Computing surveys*, 45(2), 18:1-18:30.
 Acevedo, M. E., Yanez, C., & Lopez, I. (2006a). A new

model of BAM alpha-beta bidirectional associative memories. *Lecture Notes in Computer Science*, 4263, 286-295.
 Acevedo, M. E., Yanez, C., & Lopez, I. (2006b). Alpha-beta bidirectional associative memories based translator. *International Journal on Computational Science and Network Security*, 6, 190-194.
 Bégin, J., & Proulx, R. (1996). Categorization in unsupervised neural networks: The eidos model. *IEEE Transactions on Neural Networks*, 7(1), 147-154.
 Chartier, S., & Boukadoum, M. (2006a). A Sequential Dynamic Heteroassociative Memory for MultiStep Pattern Recognition and One-to-Many Association. *IEEE Transactions on Neural Networks*, vol. 17, no. 1, pp. 59-68.
 Chartier, S., Boukadoum, M. (2006b). A bidirectional heteroassociative memory for binary and grey-level patterns. *IEEE Transactions on Neural Networks*, 17(2), 385-396.
 Chartier, S., Boukadoum, M. (2011). Encoding static and temporal patterns with a bidirectional heteroassociative memory, *Journal of Applied Mathematics*, 2011, 1-34.
 Fausett, L. (1994). *Fundamentals of neural networks: Architectures, algorithms, and applications*. Upper Saddle River, NJ: Prentice Hall, inc.
 Haykin, S. (2009). *Neural networks and learning machines*. (3rd ed.). Upper Saddle River, NJ: Pearson Education, inc.
 Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Science*, 79(8), 2554-2558.
 Kosko, B. (1988). Bidirectional associative memories. *IEEE Transactions on Systems, Man and Cybernetics*, 18(1), 49-60.
 Nong, H. T., & Bui, T. D. (2012). A fast iterative learning strategy for Bi-directional associative memory. *Proceedings of the 2012 Conference on Information Science and Technology*, 2, 182-184.
 O'Reilly, R. C. (1998). Six principles for biologically-based computational models of cortical cognition. *Trends in Cognitive Sciences*, 2(11), 455-462.
 Ritter, G. X., Diaz-Deleon, J., L. & Sussner. (1999). Morphological bidirectional associative memories. *Neural Networks*, 12, 851-867.
 Storkey, J., & Valabregue, R. (1999). The basins of attraction of a new Hopfield learning rule. *Neural Networks*, 12(6), 869-876.
 Wu, Y. & Pados, D. A. (2000). A feedforward bidirectional associative memory. *IEEE Transactions on Neural Networks*, 11(4), 859-866.