

BAM Learning in High Level of Connection Sparseness

Christophe Tremblay

School of Psychology, 136 Jean-Jacques Lussier, Ottawa, ON K1N 6N5
ctrem040@uottawa.ca

Sylvain Chartier

School of Psychology, 136 Jean-Jacques Lussier, Ottawa, ON K1N 6N5
sylvain.chartier@uottawa.ca

Abstract

Bidirectional Associative Memories (BAMs) are artificial neural networks that can learn and recall various types of associations. Although BAMs have shown great promise at modeling human cognitive processes, these models have often been investigated under optimal conditions in which the network is fully connected. Whereas some BAM models have shown to be robust to connection sparseness, those particular models could not handle highly sparse connectivity, unlike the human brain. This paper shows that a particular type of BAM can perform learning and recall under higher levels of sparse connectivity by increasing input dimensionality. This study provides a better understanding of the conditions impacting the convergence of the learning in BAM models and introduces a new avenue of research in learning in biological levels of sparseness, namely network dimensionality.

Introduction

A successful approach to modeling human classification and recall of various associations has consisted of distributing information over parallel networks of processing units. Brain inspired recurrent associative memories offer the ability to develop attractors for each pattern through feedback connections such as the Hopfield model (Hopfield 1982). Kosko (1988) generalized the Hopfield model to a heteroassociation, creating a new class of neural network models, the Bidirectional Associative Memory (BAM). Numerous BAM models were developed following Kosko's (e.g. Xu, Leung and He 1994; Zhuang, Huang and Chen 1993) that showed improvements such as performing multi-step pattern recognition or learning real-valued correlated patterns. In this paper, a BAM model

proposed by Chartier & Boukadoum (2006, 2011), the Bidirectional Heteroassociative Memory (BHM) is used. Although the BHM has shown great promise at modeling human cognitive processes, this model has often been investigated under optimal conditions in which the network is fully connected. In other words, the BHM does not take into account sparseness, unlike what is known of biological neural networks.

Evidence from neuroscience research has shown that the brain is sparsely connected (Brecht, Schneider, Sakmann and Margrie 2004, Fiete, Hahnloser, Fee and Seung 2004). On a global level, neurons are generally connected to only 1×10^2 neurons of the total 1×10^{11} neurons in the brain, therefore extremely sparsely connected. Extreme sparseness also holds on a local level where for example, it was shown that neurons of the hippocampus are connected to not more than 5% of other neurons (Amaral, Ishizuka and Clairborne 1990; D'Este, Towsey and Diederich 1999). Although the human brain only makes up 2% of body mass, it consumes 20% of the body's resources (Kety 1957; Clarke and Sokoloff 1999). Based on this evidence, if the brain was fully connected, as in the case of artificial neural networks, it would need to consume about 400% ($\frac{0.2}{0.05} \times 100$) of the current energy resources needed for body functioning. Currently, assuming full connectivity in artificial neural network models contradicts the empirical evidence that biological neural networks cannot afford to be fully connected.

The implementation of sparseness in a BAM model was studied recently in Tremblay et al. (2014), where it was shown that the BHM is robust to connection sparseness as the performances of the network in recall associations are not reduced in medium sparseness conditions. However, high levels of sparseness (i.e. higher than 70%) drastically reduced the performance of the network. Again, this performance is barely more biologically plausible, as a

connectivity level of 30% (sparseness level of 70%) as opposed to 5% as observed in the hippocampus would result in 6 times the actual energy consumption of the brain. Another sparsely connected BAM model was proposed in Bhatti (2009), which required a minimum interconnection of 15%. Although this model showed improvements in model performance, it does not achieve the plausible minimum level of connectivity. The problem also holds in standard associative memories where it was shown that connection sparseness leads to reduced performances where even minimal connection sparseness diminished the storage capacity of the memory (Bosh and Kurfess 1998; Shirazi, Shirazi and Maekawa 1993) limiting the biological plausibility of this class of models. The issue of a high minimum connectivity in BAM networks is a serious limitation to the biological plausibility of the model and should be solved for the latter to be considered a good model of neurodynamics.

This paper proposes a solution to this limitation by increasing input dimensionalities. It is showed that such an increase leads to better performances for a same memory load ratio. In other words, having a higher input dimensionality is not a burden but rather a blessing in natural system. This study provides a better understanding of the conditions that impact learning and recall of BAM networks and suggests that the BAM is a versatile neural network model that can take into account many cognitive and biological constraints simultaneously.

The remainder of the paper is divided as follows: Section II describes the network's architecture. Section III shows simulation results regarding the network's performance in learning and recalling autoassociations. Section IV discusses the results and provides a conclusion of our work.

Model

The model proposed by Chartier & Boukadoum (2006, 2011) is made of two Hopfield-like neural networks interconnected in head-to-tail fashion, providing a recurrent flow of information that is processed in a

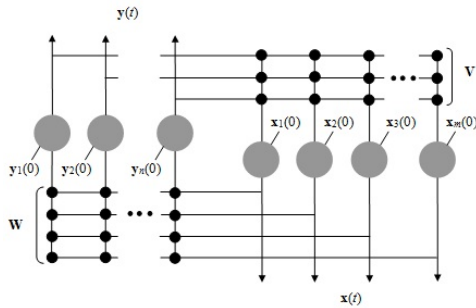


Figure 1: Architecture of the BAM.

bidirectional fashion. The network's architecture is presented in fig. 1 where $x(0)$ and $y(0)$ represent the initial vectors-states, \mathbf{W} and \mathbf{V} are the weight matrices and t is the current iteration number.

Transmission Function

The transmission function is defined by the following equations:

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

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

and

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

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

where N and M are the number of units in each layer, i is the unit index, δ is a general transmission parameter and a and b are the activation. These activations are obtained the usual way: $a(t) = \mathbf{W}x(t)$ and $b(t) = \mathbf{V}y(t)$. In short, the equation is made of a cubic function with hard limits added at 1 and -1. This function has the advantage of exhibiting grey-level attractor behaviour (Chartier and Boukadoum 2006).

Learning Rule

As for the learning rule, the weight connection's modification is done following a Hebbian/anti-Hebbian approach (Storkey and Valabregue 1999)/(Begin and Proulx 1996):

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

where $x(0)$ and $y(0)$ are the initial inputs to be associated, and k is the learning trial number. However, the learning differs from the original BHM by replacing the learning parameter η by a matrix of learning parameters (Tremblay et al. 2014). This matrix of learning parameters allows the introduction of sparseness. These matrices are given by:

$$\mathbf{A} = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1m} \\ \alpha_{21} & \alpha_{22} & & \alpha_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nm} \end{bmatrix} \quad (3a)$$

for linking the x-layer to y-layer and

$$\mathbf{B} = \begin{bmatrix} \beta_{11} & \beta_{12} & \dots & \beta_{1n} \\ \beta_{21} & \beta_{22} & \dots & \beta_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{m1} & \beta_{m2} & \dots & \beta_{mn} \end{bmatrix} \quad (3b)$$

for linking the y-layer to the x-layer. In a standard BHM, the network convergence is guaranteed if the learning parameter $(\alpha_{ij} \boxtimes \beta_{ij})$ is smaller than the threshold found with:

$$(\alpha_{ij} \boxtimes \beta_{ij}) = \frac{1}{2(1-2\delta)\text{Max}[M,N]}, \quad \delta \neq \frac{1}{2} \quad (4)$$

where M and N are respectively the dimensionality of the input and its association (Chartier & Boukadoum 2006).

Simulation

This simulation assesses the performances of a sparse BHM on a random auto associative task using large input dimensionalities.

Methodology

The BHM used was described in the model section. Sparseness was applied by setting a number of learning parameters to 0. It is noted from Equation (2) that if $(\alpha_{ij} \boxtimes \beta_{ij}) = 0$ for a given connection, then $w_{ij}(k+1) = w_{ij}(k)$. The task performed is an autoassociation of random bipolar patterns. The input dimensions varied from 500 to 2000, while the memory load (number of input patterns compared to their dimensionality) was kept constant to a value of 10% for every condition tested. In addition to this, the degree of sparseness was tested for 80% and 90% throughout the simulations. The patterns were generated randomly without repetition within the list of associations. The transmission function parameter (δ) was set to 0.2 throughout the simulations and the number of iterations through the network before the weight matrices are updated was set to $t = 1$.

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 (2).
- 4) Repetition of steps 1) to 3) until all pairs have been presented.
- 5) Repetition of steps 1) to 4) until the mean square error is lower than 10^{-4} .

The network was then tested on a recall task with pixel flipped noise (flipping a proportion of pixels, varying from

0% to 50%). Each recall task was repeated 200 times for a given pattern. Mean performances are reported. Performances are given by the number associations correctly reconstructed ($\text{MSE} < 1$, where MSE is mean squared error between the reconstructed and target vector) over the total number of associations.

Results

Figure 2 (top) presents the simulation results for the 80% sparseness condition while Figure 2 (bottom) presents the same results with 90% sparseness. Both show a clear increase in the performances of the network as the dimensionalities are increased, even with memory load kept constant. Of course, if the noise level is too extreme (higher than 42% pixel flipped) then performances are poor in every conditions tested. Results also show a bigger impact of dimensionality on performances when sparseness levels are higher (i.e. 90%).

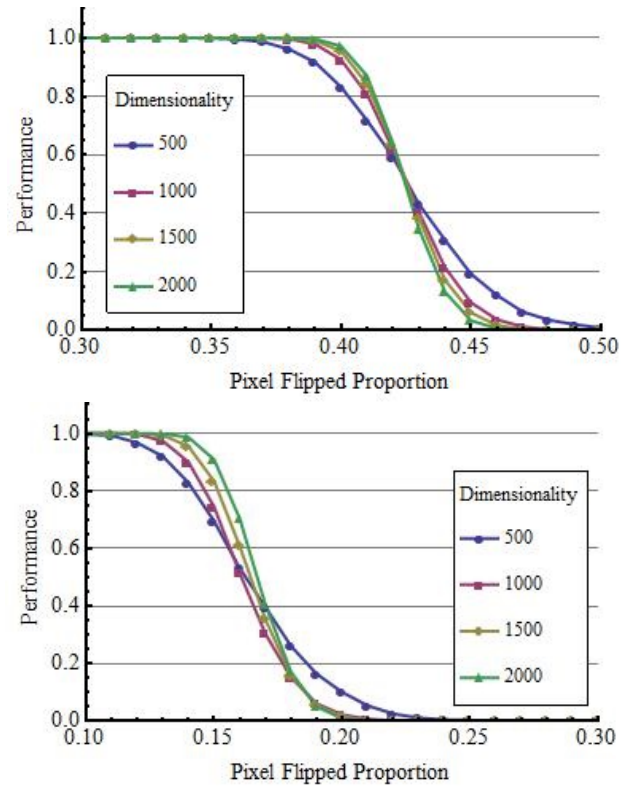


Figure 2: Performances of the BHM on a recall task with pixel flipped under varying levels of input dimensionality for (top) 80% sparseness and (bottom) 90% sparseness.

Discussion

As the results showed, the BHM is indeed capable of handling high levels of sparseness (i.e. 80%). The results are consistent with the ones reported in other types of

connectionist models where sparse connectivity was shown to still lead to stable solutions (Shirazi, Shirazi and Maekawa 1993; Bosch and Kurfess 1998, Hoyer and Hyvarinen 2002; Bhatti 2009). However, the influence of input dimensionality on network performances was never clearly explored in connectionist models. Also, our results show that unlike what was shown in Tremblay et al. (2014), sparseness levels higher than 70% do not lead to reduced performances as long as the input dimensions are high enough. These also add to the results found in Bhatti (2009), where increased dimensionalities could fix for the decreased maximum memory load in sparse network. In short, these results show that the network size should be taken into account in future research on sparseness tolerance where high level of dimensionality, as is seen biological neural networks (Curcio et al. 1990), could lead to a higher tolerance to connection sparseness. The results also show that network sparseness, at a natural level, does not systematically lead to reduced performances and that sparseness therefore represents a clear advantage.

In conclusion, the present paper introduces a way to correct for reduced performances in highly sparse BHM conditions by increasing the input dimensionality. By showing that the BHM can handle highly sparse connectivity, this article provides a better understanding of the conditions that impact the convergence of the BHM. This research should have implications for VLSI where extreme sparse connectivity could be used in order to save processing time with no effect on network performances. Finally, this research should also have implications in physical neural networks and in robotics where it is shown that large networks can tolerate extreme sparseness, again saving time and resources.

Acknowledgements

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

References

- Amaral, D. G., Ishizuka, N., and Clairborne, B. 1990. Neurons, numbers and the hippocampal network. *Progress in Brain Research* 83(1):1-11.
- Begin, J. and Proulx, R. 1996. Categorization in unsupervised neural networks: the eidon model. *IEEE Transactions on Neural Networks* 7(1):147-154.
- Bhatti, A. A. 2009. A minimum interconnection direct storage model of a neural bidirectional memory. *Proceedings of the International Conference on Neural Networks IJCNN2009*: 308-315.
- Bosch, H. and Kurfess, F. J. 1998. Information storage capacity of incompletely connected associative memories. *Neural Networks* 11(5): 869-876.
- Brecht, M., Schneider, M., Sakmann, B., and Margrie, T. W. 2004. Whisker movements evoked by stimulation of single pyramidal cells in rat motor cortex. *Nature* 427(6976): 704-710.
- Chartier, S., and Boukadoum, M. 2006. A bidirectional heteroassociative memory for binary and grey-level patterns. *IEEE Transactions on Neural Networks* 17(2): 385-396.
- Chartier, S., and Boukadoum, M. 2011. Encoding static and temporal patterns with a bidirectional heteroassociative memory. *Journal of Applied Mathematics* 2011: 1-34.
- Clarke, D. D. and Sokoloff, L. 1999. Circulation and energy metabolism of the brain. In Siegel G. J., Agranoff, B. W., Albers, R. W., Fisher, S. K., Uhler, M. D. (Eds.) *Basic neurochemistry: Molecular, cellular and medical aspects*. Philadelphia: Lippincott-Raven.
- Curcio, C. A., Sloan, K. R., Kalina, R. E., and Hendrickson, A. E. 1990. Human photoreceptor topography. *The Journal of Comparative Psychology* 292(4): 497-523.
- D'Este, C., Towsey, M., and Diederich, J. 1999. Sparsely-connected recurrent neural networks for natural language learning. *First Workshop on Natural Language Processing and Neural Networks* 1999: 64-69.
- Hopfield, J. J. 1982. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences* 79(8): 2554-2558.
- Hoyer, P. O., and Hyvarinen, A. 2002. A multi-layer sparse connectivity network learns contour coding from natural images. *Vision Research* 42(12): 1593-1605.
- Kety, S. S. 1957. The general metabolism of the brain in vivo. In Richter D. (Eds.) *Metabolism of the nervous system*. London: Pergamon.
- Kosko, B. 1988. Bidirectional associative memories. *IEEE Transactions on Man, Systems and Cybernetics* 18(1): 49-60.
- Shirazi, M. N., Shirazi, M. N., and Maekawa, S. 1993. The capacity of associative memories with malfunctioning neurons. *IEEE Transactions on Neural Networks* 4(4): 628-635.
- Storkey, J., and Valabregue, R. 1999. The basins of attraction of a new Hopfield learning rule. *Neural Networks* 12(6): 869-876.
- Tremblay, C., Dorville, M., Myers-Stewart, K., and Chartier, S. 2014. Spreading-Activation and Sparseness in a Bidirectional Associative Memory. *Proceedings of the 2013 International Joint Conference on Neural Networks IJCNN2013*: 1917-1923.
- Xu, Z. B., Leung, Y., and He, X. W. 1994. Asymmetric bidirectional associative memories. *IEEE Transactions on Man, Systems and Cybernetics* 24(10): 1558-1564.
- Zhuang, Y., Huang, Y., and Chen, S. 1993. Better learning for bidirectional associative memory. *Neural Networks* 6(8): 1131-1146.