Towards binary encoding in Bidirectional Associative Memories

Thaddé Rolon-Merette[1a], Damiem Rolon-Merette[1b] & Sylvain Chartier[1c]

[1] School of Psychology at the University of Ottawa, ON, Canada
[a]Trolo068@uottawa.ca, [b]Drolo083@uottawa.ca, [c]Sylvain.chartier@uottawa.ca

Abstract

Bidirectional Associative Memories (BAMs) are Artificial Neural Networks frequently utilized in cognitive modeling. While bipolar encoding is commonly used in BAMs for optimal performance, binary encoding presents interesting properties. As such, this study introduces a novel transmission function for binary encoding and compares its performance to the conventional bipolar transmission function. To evaluate, an auto-association learning task and a noisy recall task were implemented. Results revealed that despite longer learning times, binary encoding preserves or enhances the properties observed in binary encoding. Findings are promising from a cognitive perspective, as they open the possibility of building intricate models of human cognition.

Introduction

Artificial Neural Networks (ANNs) are computational models inspired by the structure and function of biological neural networks. They can be an interesting approach towards explaining cognitive processes [Hasson et al, 2020]. One notable group of ANNs used in cognitive modeling is Bidirectional Associative Memories (BAMs), which operate based on a neurodynamic perspective. BAMs use feedback weights to learn pairs of stimuli and are noise tolerant, capable of recalling inputs when only given partial information [Acevedo-Mosqueda et al., 2013]. BAMs usually use bipolar encoding, where input vectors are composed of values of -1 and 1, since it increases learning performance over binary encoding, where input vectors are composed of 0 and 1 [Kosko, 2021]. However, when using ANNs for cognitive modelling, they must be built on principles grounded in processes occurring in the brain while avoiding methods that merely enhance computational efficiency [O'Reilly, 1998]. Binary encoding is believed to be more biologically plausible since it is a closer representation of the presence and absence of spikes. Furthermore, it provides the absorbent property of 0, which could allow the implementation of more cognitive processes, like true sparseness, gating, filtering and more. Therefore, this paper introduces a binary transmission function for the BAM, derived from its bipolar implementation, and compares their learning speed and recall performance under noise.

Methodology

This study used a BAM with two layers where the weights, **W** and **V**, are feedback connections between each layer, as shown in Figure 1 and in [Chartier & Boukadoum, 2006].

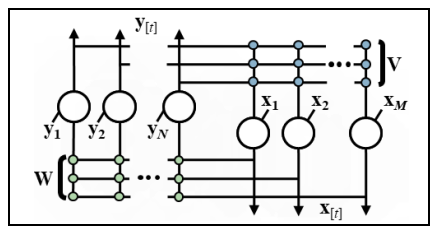

Figure 1. The BAM's architecture.

Each layer receives a set of external inputs, $\mathbf{x}[0]$ and $\mathbf{y}[0]$, which are given to the network and cycled through the weights matrices t times to generate $\mathbf{x}[t]$ and $\mathbf{y}[t]$. Weights are updated using a Hebbian/anti-Hebbian learning rule defined by equations 1a and 1b:

$$\text{1a)}\quad \mathbf{W}_{[k+1]} = \mathbf{W}_{[k]} + \eta\left(\mathbf{y}_{[0]} - \mathbf{y}_{[t]}\right)\left(\mathbf{x}_{[0]} + \mathbf{x}_{[t]}\right)^{\mathrm{T}}$$

Here, η is the learning parameter, and k is a given learning trial. **V** is updated using an equivalent process. The newly introduced binary transmission function is a variation of the cubic bipolar function outlined in [Rolon-Merette et al., 2018]. It maintains the cubic form but with saturating limits at 1 and 0, as expressed by equation 2:

$$\text{(2)}\ \forall i, \dots, M, \mathbf{x}_{i[t+1]} = \begin{cases} 1, & if\ \mathbf{W}\mathbf{x}_{i[t]} > 1 \\ 0, & if\ \mathbf{W}\mathbf{x}_{i[t]} < 0 \\ 3(\mathbf{W}\mathbf{x}_{i[t]})^2 - 2(\mathbf{W}\mathbf{x}_{i[t]})^3, & Else \end{cases}$$

Where **M** is the number of units in each layer, i is the index unit, and t is the time index. A similar process is used to obtain $\mathbf{y}_{i[t+1]}$ by replacing $\mathbf{W}\mathbf{x}_{i[t]}$ with $\mathbf{V}\mathbf{y}_{i[t]}$.

To compare the performance of both transmission functions, an auto-association task was conducted where the network had to learn the 26 letters of the alphabet. Each letter was represented by a 49-dimensional input vector, where black pixels indicated 1s and white pixels were either 0 for binary encoding or -1 for bipolar encoding. An example of these input vectors is provided in Figure 2.

Figure 2. Inputs of the auto-association task

The task was learned under three different conditions, which were established by varying the learning parameter for slow (η = 0.001), medium (η = 0.005), and fast (η = 0.01) learning. Learning was stopped when a minimum mean square error (MSE) was achieved between the stored and actual input vectors of 10e-5. Following training, the network's recall performance was evaluated on a random pixel flip noise task, with the number of pixels flipped ranging from 0 (0%) to 24 (50%). For each learning condition and noise percentage, learning and recall were repeated for 1000 trials with different randomly generated noise. Finally, an additional recall task was used on each learning condition to determine the proportion of spurious attractors. This was done by giving 1000 random input vectors to the network to find the number of vectors that stabilize in a spurious state divided by the total number of vectors. To ensure the network would self-stabilize, the number of cycles *t* during recall was set to 100.

Results

The average learning times are shown in Figure 3. Results indicate that learning time is, on average, three times longer than bipolar regardless of learning condition.

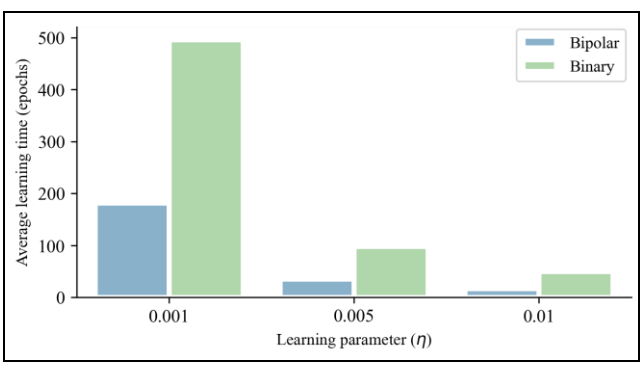

Figure 3 – Average learning time

The average recall performances under the noisy pixel flip task are shown in Figure 4. Results show the traditional graceful degradation found in BAMs. However, binary encoding showed slightly superior recall performances (%) on average) for each learning condition compared to bipolar. No significant effect was found by manipulating the learning parameter.

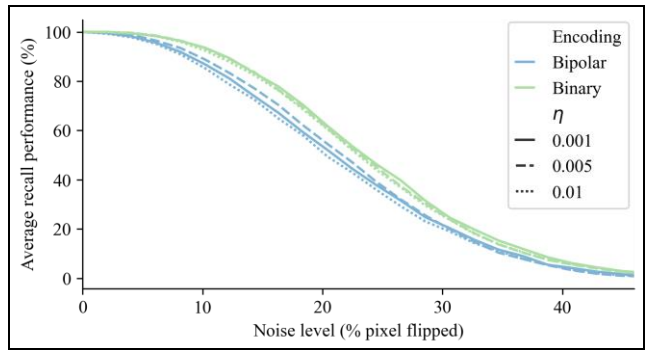

Figure 4 – Average recall performance under noise.

The proportion of spurious attractors present in the network was, on average, at least 20% lower with binary encoding, as shown in Figure 5. This could explain why the recall performance under noise was better with binary.

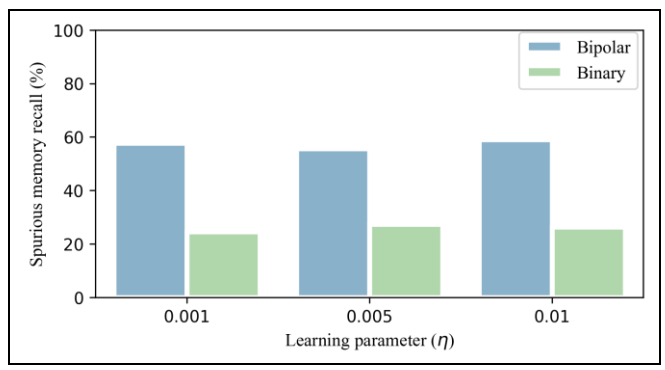

Figure 5 – Average proportion of spurious attractors.

Conclusion

This study aimed to introduce a new transmission function for binary encoding in a BAM that maintained the learning properties and performances obtained by its bipolar counterpart. Results showed that the binary's ability to learn the task remained unaffected despite expected longer learning times. Furthermore, the model showed better recall under noise and fewer spurious attractors with binary encoding. These findings suggest that binary encoding can be used without sacrificing performance or losing the properties previously observed in bipolar encoding. Further research should evaluate binary encoding when using different architectures and cognitive tasks (Rolon-Mérette, et al., 2019) and explore the additional properties gained.

References

Acevedo-Mosqueda, M. E., Yanez-Marquez, C., & Acevedo-Mosqueda, M. A. (2013). Bidirectional associative memories: Different approaches. *ACM Computing Surveys (CSUR*), *45*(2), 1-30.

Chartier, S., & Boukadoum, M. (2006). A bidirectional heteroassociative memory for binary and grey-level patterns. *IEEE Transactions on Neural Networks*, *17*(2), 385-396.

Hasson, U., Nastase, S. A., & Goldstein, A. (2020). Direct fit to nature: an evolutionary perspective on biological and artificial neural networks. *Neuron*, *105*(3), 416-434.

Kosko, B. (2021). Bidirectional associative memories: unsupervised Hebbian learning to bidirectional backpropagation. IEEE *Transactions on Systems, Man, and Cybernetics: Systems*, *51*(1), 103-115.

O'Reilly, R. C. (1998). Six principles for biologically based computational models of cortical cognition. *Trends in cognitive sciences*, *2*(11), 455-462.

Rolon-Merette, T., Rolon-Merette, D., & Chartier, S. (2018). Generating Cognitive Context with Feature-Extracting Bidirectional Associative Memory. *Procedia computer science*, *145*, 428-436.

Rolon-Mérette, T., Rolon-Mérette, D., Calderini, M., & Chartier, S (2019). Different Brain, Same Prototype? Cognitive Variability within a Recurrent Associative Memory. *Proceedings of the 17th International Conference on Cognitive Modeling (pp. 192–197). Montreal, Canada.*