

Capacidad y Recuperación en Memorias Asociativas Bidireccionales (BAM) Discretas

J.A. Horas* y E.A. Bea

Universidad Nacional de San Luis - Facultad de Ciencias Físico-Matemáticas y Naturales
Departamento de Física - Instituto de Matemática Aplicada (IMASL)
Ejército de los Andes 950 - (5700) San Luis - Argentina
e-mail: jhoras@unsl.edu.ar

Realizamos un estudio comparativo de la performance de tres tipos de memorias asociativas bidireccionales discretas. En el primero y segundo de ellos la forma de codificar los patrones a grabar es Hebbiana mientras que en el último se usa otra codificación. Con el objetivo de aumentar la extensión de las cuencas de atracción, se aplica un procedimiento de desaprendizaje también Hebbiano que modifica la intensidad de las conexiones entre neurodos, obteniéndose interesantes resultados.

La performance es estudiada para diversos tipos de patrones (biased y no biased). Mostramos el comportamiento para varios tamaños de la red utilizando también la influencia de los parámetros relevantes.

We perform a comparative study on the performance of three coding schemes of discrete bidirectional associative memory (BAM). In the first and second we use Hebbian coding while in the last another coding is used. We apply an unlearning procedure which is also Hebbian, with the goal of enlarging the basins of attraction. This procedure also modifies the synapsis intensities.

We study the performance for biased and not biased patterns. We also show the behavior for several network sizes, and the parameter influence is also analyzed.

I. Introducción

La capacidad de actuar como memorias asociativas es una de las importantes características de las redes neuronales artificiales[1,2]. Entre los muchos modelos de memorias asociativas se encuentra la Memoria Asociativa Bidireccional (BAM)[3], cuyo mayor interés es su capacidad de lograr heteroasociación con una pequeña matriz de correlación. Puede decirse que la BAM es similar a la usual memoria autoasociativa de Hopfield pero con dos capas en vez de una. En la fig. 1 esquematizamos su topología.

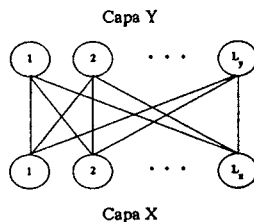


Figura 1: arquitectura de la BAM.

La performance de una memoria asociativa es determinada por dos aspectos centrales que no son independientes, por un lado la capacidad para grabar el mayor número posible de patrones y por el otro la habilidad que debe tener una red para recuperar los patrones grabados si se la inicia con una versión

incompleta o ruidosa de los mismos, una medida de esto último esta dada por el tamaño de las cuencas de atracción.

En otro tipo de memorias asociativas, la incorporación de un proceso de desaprendizaje[2] mejora la performance. Mostramos aquí, por primera vez a nuestro conocimiento, que lo mismo ocurre para una memoria heteroasociativa como la BAM.

En resumen, en este trabajo:

- 1) Se mantiene la codificación Hebbiana en la regla de aprendizaje de una BAM.
 - 2) Se adiciona un proceso, también Hebbiano, llamado de desaprendizaje (uBAM).
 - 3) Se usa otro esquema de codificación (eBAM) [4].
- analizando la performance en cada caso para diferentes condiciones.

II. El modelo

La idea básica de la BAM (ver figura 1) es:

Grabación

Como hay dos capas en la BAM una de ellas se denota como campo f_x y tiene L_x neurodos correspondientes al patrón X , y la otra se denota como campo f_y y tiene L_y neurodos correspondientes al patrón Y . La matriz de conexiones o de acoplamientos sinápticos está dada por

$$W = \sum_{i=1}^M Y^{(i)} X^{(i)T} \quad (1)$$

* Investigador del CONICET.

donde los M pares de patrones a grabar $(X^{(1)}, Y^{(1)}), \dots, (X^{(M)}, Y^{(M)})$ son bipolares, de L_x (o L_y) componentes $+1$ o -1 .

Recuperación

Luego de la presentación de un patrón inicial o de prueba X (o Y) la BAM comienza a evolucionar desde ese estado inicial y recupera un cierto par de patrones de acuerdo a la siguiente dinámica

$$Y'_{BAM} = \text{sgn}(WX) = \text{sgn}\left(\sum_{i=1}^M (Y^{(i)} X^{(i)T}) X\right) \quad (2)$$

$$X'_{BAM} = \text{sgn}(W^T Y) = \text{sgn}\left(\sum_{i=1}^M (X^{(i)} Y^{(i)T}) Y\right) \quad (3)$$

La convergencia de este proceso está asegurada puesto que la red siempre llega, después de un tiempo finito, a un estado estable o punto fijo que es uno de los mínimos del Hamiltoniano o energía de la red, dada por

$$H_{BAM}(X, Y) = -Y^T W X = -X^T W^T Y \quad (4)$$

Desafortunadamente la codificación dada por (1) no puede garantizar el ideal de que cada par nominal sea un punto fijo, aún cuando el número de pares a grabar sea pequeño. A fin de moderar este problema, manteniendo el mismo esquema de codificación, el modelo se completa con un proceso de desaprendizaje sobre la matriz de conexiones sinápticas, de acuerdo al siguiente proceso:

- 1) Inicialización al azar (Random Shooting).
- 2) Relajación: el sistema debe relajar a una configuración estacionaria debido a la existencia de una función energía.
- 3) Desaprendizaje: la matriz de conexiones es actualizada según

$$W^{(d)} = W^{(d-1)} - \varepsilon Y^{(d)} X^{(d)T} \quad (5)$$

El procedimiento es local, iterativo y puramente Hebbiano. Puede obtenerse[2]

$$D_{\text{optimo}} = \frac{M}{\varepsilon} \left[c_0 + \frac{c_1}{\frac{1}{2}(L_x + L_y)} \right] \quad (6)$$

con $c_0=0.28$ y $c_1=10$, para el número óptimo de iteraciones del procedimiento.

Por último, para la BAM con codificación exponencial (eBAM) [4] se tiene

$$Y'_{eBAM} = \text{sgn}\left(\sum_{i=1}^M Y^{(i)} b^{X^{(i)} X}\right) \quad (7)$$

$$X'_{eBAM} = \text{sgn}\left(\sum_{i=1}^M X^{(i)} b^{Y^{(i)} Y}\right) \quad (8)$$

donde $b=2$.

Al igual que en la BAM, la convergencia de este proceso está asegurada pues existe el Hamiltoniano

$$H_{eBAM}(X, Y) = -\sum_{i=1}^M b^{X^{(i)} X} - \sum_{i=1}^M b^{Y^{(i)} Y} \quad (9)$$

III. Resultados numéricos

A fin de analizar la influencia del proceso de desaprendizaje sobre la performance de la BAM, se ha estudiado el mismo en dos casos de heteroasociación. En todas las simulaciones, se han tomado redes con igual número de neuronas en ambas capas, sin pérdida de generalidad, y la actualización es sincrónica.

Caso 1:

Los patrones son bipolares generados al azar y con la misma probabilidad de ser $+1$ o -1 en cada bit. Se generan el número conveniente de patrones para cada punto, según se indica en cada gráfica.

Capacidad

En las figuras 2 y 3 se grafican el número de eventos exitosos versus el número de patrones grabados para una BAM con y sin desaprendizaje. Definimos un evento exitoso cuando la BAM recupera completamente un par de patrones grabados. Las condiciones de operación y otros parámetros son descriptos en cada figura. En todas ellas se promedia sobre 5 redes y se calcula la desviación standard que es denotada por barras.

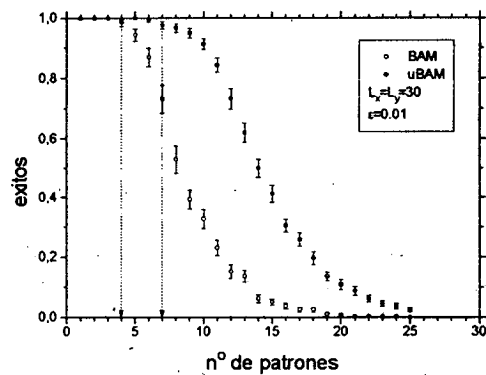


Figura 2: nro. de eventos exitosos versus el nro. de patrones grabados, para redes con 30 neuronas en cada capa. Las flechas indican la capacidad de almacenamiento máximo de cada red que se inicializa presentando los patrones nominales.

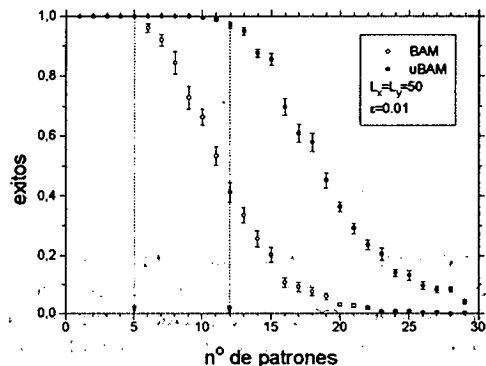


Figura 3: idem anterior para redes con 50 neuronas en cada capa.

Recuperación

En las figuras 4 y 5 se muestran gráficas que dan el comportamiento de ambas BAM, cuando las redes son inicializadas con versiones ruidosas que se diferencian de los patrones nominales por el porcentaje de inversión de bits (ruido) que se especifica en cada caso.

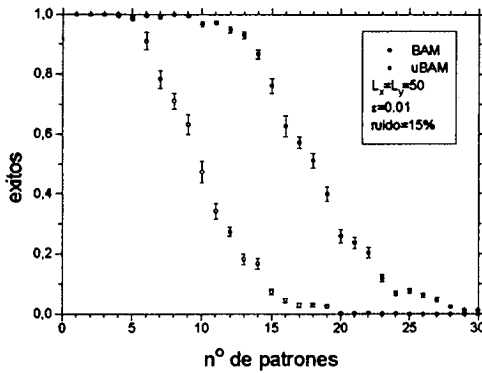


Figura 4: nro. de eventos exitosos versus el nro. de patrones grabados, para redes con 50 neuronas en cada capa. Las redes son inicializadas con versiones ruidosas de los patrones nominales.

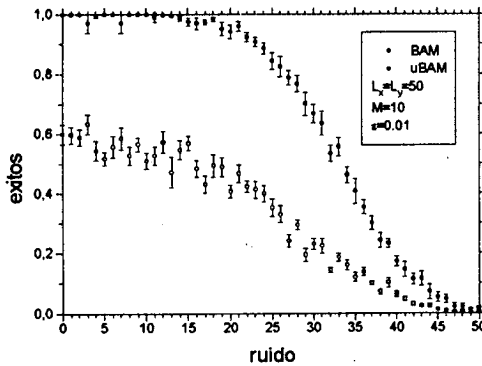


Figura 5: nro. de eventos exitosos versus porcentaje de inversión de bits sobre 10 patrones nominales, para redes con 50 neuronas en cada capa.

Caso 2:

Para completar este estudio preliminar se muestra en la fig. 6, resultados de capacidad y recuperación para patrones en que la probabilidad de adquirir los valores +1 y -1 para cada bit no es igual. En la recuperación se inicializa con patrones en que se ha invertido un cierto número de bits elegidos al azar. En la fig. 7 se muestran resultados para el caso de la eBAM.

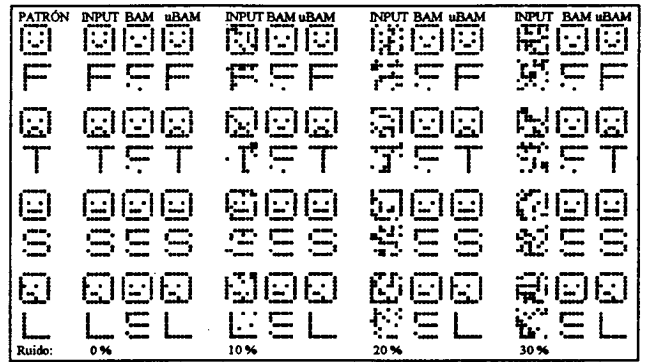


Figura 6: resultados de recuperación para las BAM y uBAM, inicializadas con versiones ruidosas de 4 patrones biased. En cada capa hay 64 neuronas.

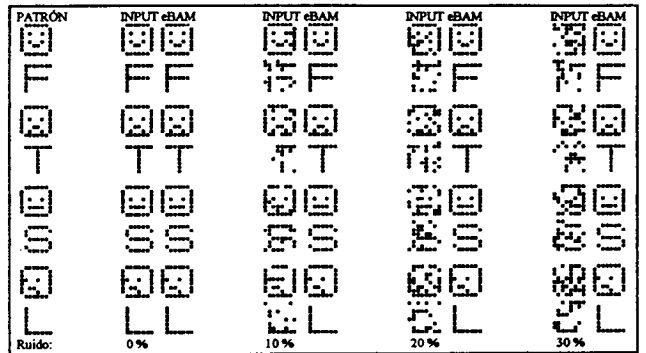


Figura 7: resultados de recuperación para la eBAM, inicializada con versiones ruidosas de 4 patrones biased. En cada capa hay 64 neuronas.

IV. Conclusiones

Se ha mostrado aquí la influencia del proceso de desaprendizaje aplicado a una BAM. Las simulaciones muestran interesantes resultados en la performance. Estos resultados preliminares impulsan a continuar este trabajo en varias direcciones, como por ejemplo

- Formular otra dependencia paramétrica para el número óptimo de pasos de desaprendizaje en que se permita variar la profundidad ϵ de estos.

- Aplicar el proceso de desaprendizaje, adecuado convenientemente, a otras formas de codificación.

Referencias

- 1- J.A. Freeman and D.M. Skapura, *Neural Networks: Algorithms, Applications, and Programming Techniques*. Addison-Wesley, 4 127-141 (1991).
- 2- E. Domany, J.L. van Hemmen and K. Schulten, *Models of Neural Networks*. Berlin: Springer-Verlag, 1 55-85 (1991).
- 3- B. Kosko, *IEEE Trans. Syst. Man, Cybern.*, 18 49-60 (1988). B. Kosko, *App. Opt.*, 26 4947-4960 (1987).
- 4- T. Wang, X. Zhuang and X. Xing, *IEEE Trans. Neural Networks*, 3 1010-1018 (1992).