Bidirectional Associative Memory (BAM) Neural Network for Tries Structures Searching

Hind R.Mohamed

Received on: 16/2/2005
Accepted on: 9/3/2006

Abstract

In this work, the relationship between the searching in data structures and neural network is presented.
Bidirectional Associative Memory (BAM)neural network is used for searching any word in trie structures method with Hamming distance by introducing weighting matrix into the connection matrix. This generalization is validated to increase storage capacity, the design algorithm for BAM net is to search word trie structure have $O((P^2(n+m))$ computational complexity, where P is the number of prototype pairs and n,m are the dimensions of input / output bipolar vectors.

الخلاصة

تم في هذا البحث دراسة العلاقة بين الشبكات العصبية والترتيب في هياكل البيانات. استخدمت الذاكرة المقترنة الاتجاهية (BAM) للبحث عن كلمة ممثلة بإحدى طرائق البحث المسماة الهياكل المجربة.
تضـــمن العمـــل استخدام مسافة Hamming لتقديم مصفوفة الأوزان داخل مصفوفة الارتباط.ولزيادة خاصية الخـــزن لBAM تمـــتلك خوارزمـــية (BAM)خاصية البحث عن العنصر $O((P^2(n+m))$ من الحسابات المعقدة حيث ان P هي عدد نماذج الأزواج وان n,m هي للأبعاد لمتجهات الثنائية القطب (١ ، -١) للإدخال والإخراج.

1- Introduction

Neural networks are a large number of simple element called neurons. A neural network is basically a dense interconnection of simple and highly interconnected neurons[5].

Over the past several years the term associative memory
has come to refer collectively to a large class of non-linear artificial neural networks with feedback, with primary applications in the fields of pattern recognition and content addressable memory devices[7]. Arguably, associative memories trace their origin to the correlation matrix memory studies by Steinbuch and Kohonen [2].A few years later ,Hopfield's model propelled associative memories to testate of a main-stream neural-network research area. The original bidirectional associative memory (BAM)was proposed by Kosko[3]extending the Hopfield auto-associative memory to a bidirectional one. It can associat input pattern with a different stored output patten of a stored pattern pair. More layers, or interconnections inside each layer, the symmetrical BAM using the Hamming stability learning algorithm (SBAM) achieves the highest performance [6].To overcome the drawback, an asymmetrical BAM (ABAM) was proposed [3], which requires linear independence of stored patterns limiting its storage

capacity. Inferred from that the capacity of feed for-ward multilayer network and radial basis function network is greater than the number of network neurons, the general BAM (GBAM)[4] used linear reparability condition and increased the capacity slightly greater than the number of neurons in its layer[7].Data structure is a collection of data elements whose organization is characterized by accessing operations that are used to store and retrieve the individual data elements[1].

Searching is the process for finding a certain value among a number of values .There are a number of progressively more complex searching algorithm[4] . There are many fields for the data structure in this work , we take the searching by trie structure using BAM neural network .

2- Bidirectional Associative Memory (BAM) neural network and trie structure

The Bidirectional Associative Memory (BAM)neural network consists of two or more layers. For two layers, an input layer and an output layer. BAM does not stop learning when input values reach the output layer[2].The learning phase stops when the network becomes stable, no change between input and output values during two consecutive cycles. The pattern sets for training and running and output results can only have two values (0,1) or(-1,1)[7].

BAM is useful for pattern recognition or with noisy and corrupted patterns. BAM can also "forget" if there are many pattern in it. We use in many examples of BAM net the Hamming distance which is defined as the number of different bits in two binary or bipolar vectors X1 and X2 and it is denoted by H[X1,X2].The average Hamming distance between the vectors is 1/n H[X1,X2], where n is the number of components in each vector[5]. Figure (1) shows the topology of a BAM, showing the two fields of neurons connected by synapses[2].

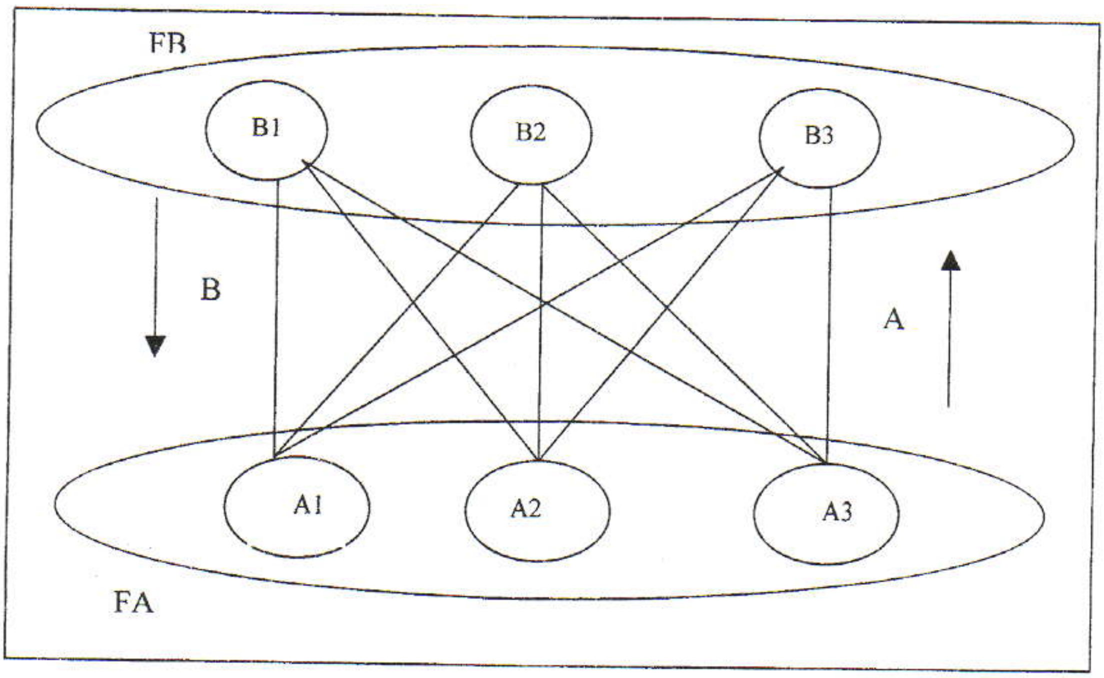


Figure (1): BAM with two fields

The trie structures is a method for searching an element(digital or word) in a list.We examine briefly the feasibility of using m-arry tree(m>=2).The branching criterion at a particular level in such a tree will be based on a portion of the key value rather than its entire key value.

A trie structure is a complete m-arry tree in which each node consists of m components .These components correspond to letters and digits. A trie structure occur frequently in the area of information organization and retrieval. The method of searching in trie is analogous to the notion of digital sorting. In particular the branching at each node of level k depends on the kth character of a key.

3- BAM algorithm

1- [Initialize]

Let (x^i, y^i); $i = 1,2,3, \ldots\ldots\ldots, p$;

be the desired bipolar pattern pairs. Dimensions of $x\ i$ and $y\ i$ are n and m respectively.

2-[BAM learns the pairs using transition matrix]

$i=1,2,3,\ldots\ldots\ldots n$

$$W = \sum (y^i x)^{iT} / (x^{iT} x^i) \qquad \ldots\ldots\ldots\ldots(1)$$

3-If a pattern x is applied to input layer, each y^i is summed in output layer with a weighting factor proportional to

$$(x^{iT} x^i) / (x^{iT} x^i) = 1 - 2dh(x^i, x) \qquad \ldots\ldots\ldots\ldots(2)$$

where $2dh(x^i,x)$ is $(\sum x\ k \neq x\ ^i k)/n$, The Hamming distance divided by n.. W^T is used in backward association.

4- Considering the different usefulness of each pixel information, diagonal matrix is:

$$\Lambda^i = \mathrm{diag}(\lambda^i 1, \ldots\ldots, \lambda^i n) \qquad \ldots\ldots\ldots\ldots(3)$$

can be multiplied before the correlation matrix[2]. We suggest a new forward and backward transition matrices Wf and Wb as follows:

$$Wf = \sum y^i x^{iT} \Lambda^i f / x^{iT} \Lambda^i f\ x^i \qquad \ldots\ldots\ldots\ldots(4)$$

$$Wb = = \sum y^i x^{iT} \Lambda^i b / x^{iT} \Lambda^i b\ x^i \qquad \ldots\ldots\ldots\ldots(5)$$

5-Applying a pattern x, each $y\ i$ is summed with a weighting factor:

$$(x^{iT} \Lambda^i x) / (x^{iT} \Lambda^i x^i) = 1 - 2\, d\Lambda^i(x^i, x) \qquad \ldots\ldots\ldots(6)$$

if new distance measure, $d\Lambda^i(x^i, x)$, is defined as

$$(\sum x\ k \neq x^i_k\ \lambda^i_k) / (\sum_k \lambda^i_k) \qquad \ldots\ldots\ldots\ldots(7)$$

where $\lambda^i k$ reflects the usefulness of k-th input element. Wb is used in backward association.[3]

6-The generalized radial basis function (GRBF) network can be expressed as

$$y = \sum w^i f((x - c^i)^T \Lambda^i ((x - c^i)) \qquad \ldots\ldots\ldots\ldots(8)$$

where activation function $f(u) = \exp(-u)$. Although GRBF does not include convergence procedure, GABAM can be translated as GRBF with $w^i = y^i$, $c^i = x^i$, and $f(u) = 1 - u/2$ from the static point. [5]

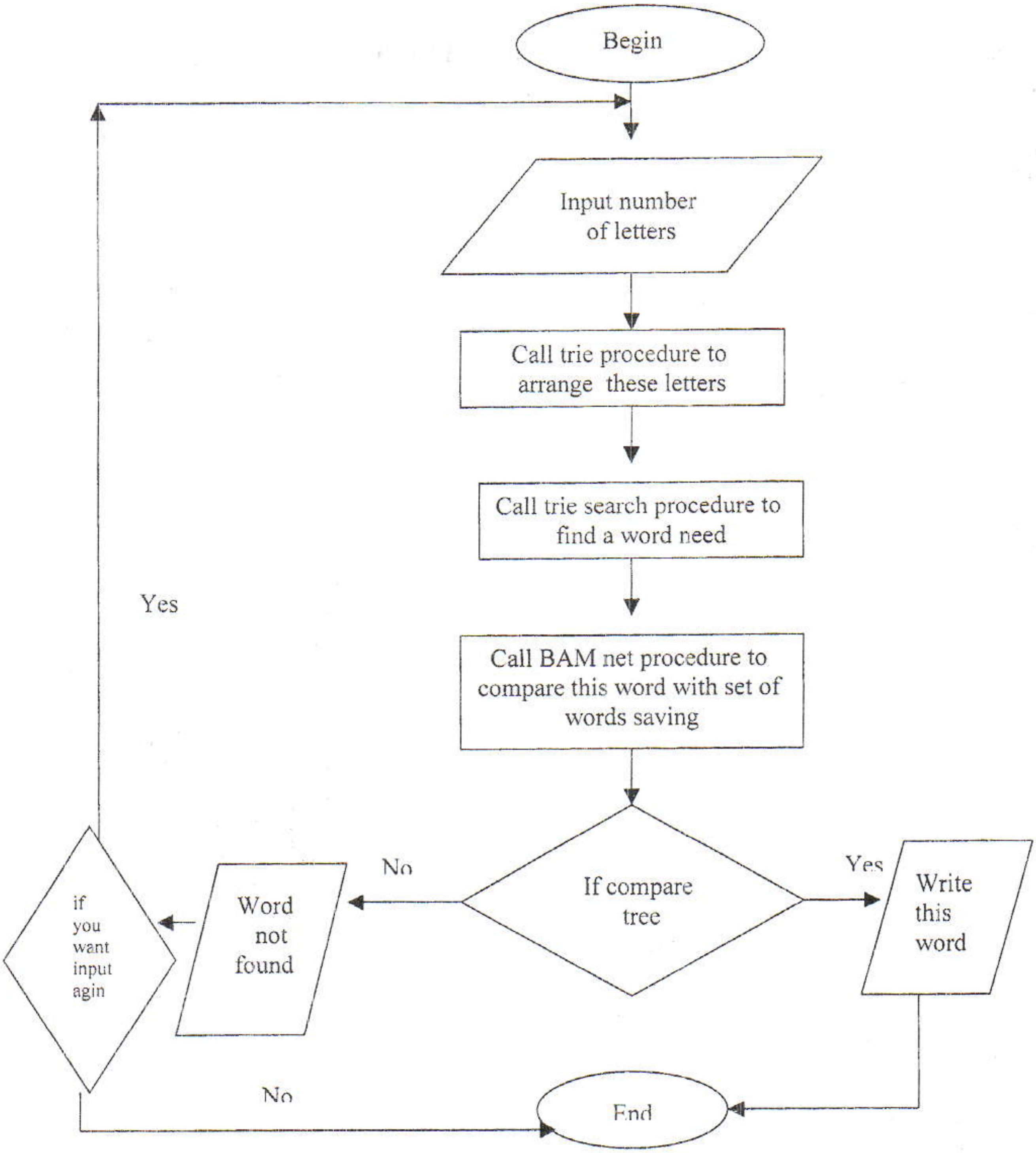
4- Design of system

Our system contain the trie structure for searching a set of 29 English words shown in table(1) consists of 12 nodes ,each of which is a vector of 27 elements. There are subnet for each word call by procedure of learning and each element contains either a dash , or the desired word , or a node number.

A blank symbol space(b) used to denote the end of a word during the scan of the key. Node 1 is the root of the trie. The architecture for BAM net contain single- layer nonlinear feedback has n units(12 node)in its X-layer and m units (27 node) in its Y-layer. The connections between the layers are bidirectional. If the weight matrix for signals sent from the X-layer to the Y-layer is W, the weight matrix for signals sent from the Y-layer to the X-layer is W^T.

Table(1): A trie structure for a list of letters

	1	2	3	4	5	6	7	8	9	10	11	12
b	-	-	-	-	me	-	-	-	-	-	-	GO
A	-	-	AC	-	-	-	-	-	-	-	-	-
B	2	-	-	-	-	-	-	-	-	-	-	-
C	3	-	-	DEL	-	-	-	-	-	-	-	-
D	4	-	-	-	-	-	-	-	-	-	END	-
E	5	BEGIN	-	-	-	-	GET	-	-	-	-	-
F	6	-	-	-	-	-	-	-	-	-	-	-
G	7	-	-	-	-	-	-	-	-	-	-	-
H	-	-	THEN	-	-	-	-	-	-	-	-	-
I	-	IF	-	-	ton	-	-	-	-	-	-	-
J	-	-	-	-	-	-	-	-	-	-	-	-
K	-	-	-	-	-	-	-	-	-	-	-	-
L	-	-	CLOSE	ELSE	-	-	-	-	-	-	-	-
M	-	-	-	-	-	-	-	-	-	-	-	-
N	NO	-	-	so	-	-	-	-	-	-	-	-
O	OPEN	-	TO	DO	-	-	-	-	-	-	-	-
P	8	-	-	-	-	-	-	-	-	-	-	-
Q	-	-	-	-	-	-	-	-	-	-	-	-
R	REPEAT	-	FREE	PROC	-	-	-	-	-	-	-	-
S	STOP	-	-	-	ke	-	-	-	-	-	-	-
T	9	-	-	PUT	-	-	-	-	-	-	ENTEY	GOTO
U	-	-	-	-	-	-	-	-	-	-	-	-
V	-	-	-	-	-	-	-	-	-	-	-	-
W	10	-	-	-	-	-	-	-	-	-	-	-
X	-	-	-	-	pe	she	-	-	-	-	-	-
Y	-	BY	-	EXIT	-	-	-	-	-	-	-	-
Z	-	-	-	-	-	-	-	-	-	-	-	-



Figure(2): Flow chart of the structure of design system

5-BAM Learning

A neural network is learnt by modifying the synapses between its neurons. All synaptic information is contained in 12 - 27 connection matrix M(n-by-p).Every matrix M between FA and FB produces a stable BAM .All inputs quickly map to a pattern of stable reverberation, but different connection matrices encode different(A,B)associations as stable reverberations set of associations{(A1,B1),(A2,B2),.......,(Am,Bm)} by summing bipolar correlation matrices. A BAM stores and recalls associations(Ai,Bi) that are learned by summing correlation on matrices. The matrix element Mij indicates the symmetric(distance dependent)synapse between neurons Ai and Bj. The synapse is excitatory if Mij>0, inhibitory if Mij<0.

Procedure for trie search performs a search of a letter .if the search fails Row and Col are assigned a value of zero , we again assume that name contains only alphabetic characters.

The algorithm for this search is:

```
1-[Initialize]
  Col=1
2-[Perform the search]
  Repeat for k=1,2,3,....length(name)
Row=index('bABC,......X,Y,Z',sub(name,k,1))
    Repeat While trie[Row,Col] ≠ '-'
     If trie[Row,Col]=name then
       Return
     If index('0 1 2 3 4 5 6 7 8 9 ',sub(trie[Row,Col],1,1))=0 then
       Write('unexpected', name, 'found')
        Row=Col=0
        Return
    Else Col=trie[Row,Col]
3-[missing name]
  write('name not found')
   Row=col=0
   Return
```

6-Discussion

The representation for search character by BAM neural network given in table(2)

Table(2):branches for all characters have branches

index	Character name	Number of branches	Name of branches	notes
1	B	2	E,Y	-
2	C	3	A,H,L	-
3	D	2	C,D	-
4	E	3	L,N,X	N have 2 branches
5	F	3	L,O,R	-
6	G	2	E,O	O have 2 branches
7	P	2	R,V	-
8	T	2	H,O	-
9	W	2	H,R	-
10	A,I,N,O,R,S	-	-	Contain one word

As an example of a search for award in trie structure shown in table(1)we trace through search for the word(END) by numbers of steps:

1-The letter E tells us that we should go from node 1 to node 5.

2-The second letter(N) is then used to select the appropriate element in node 5.

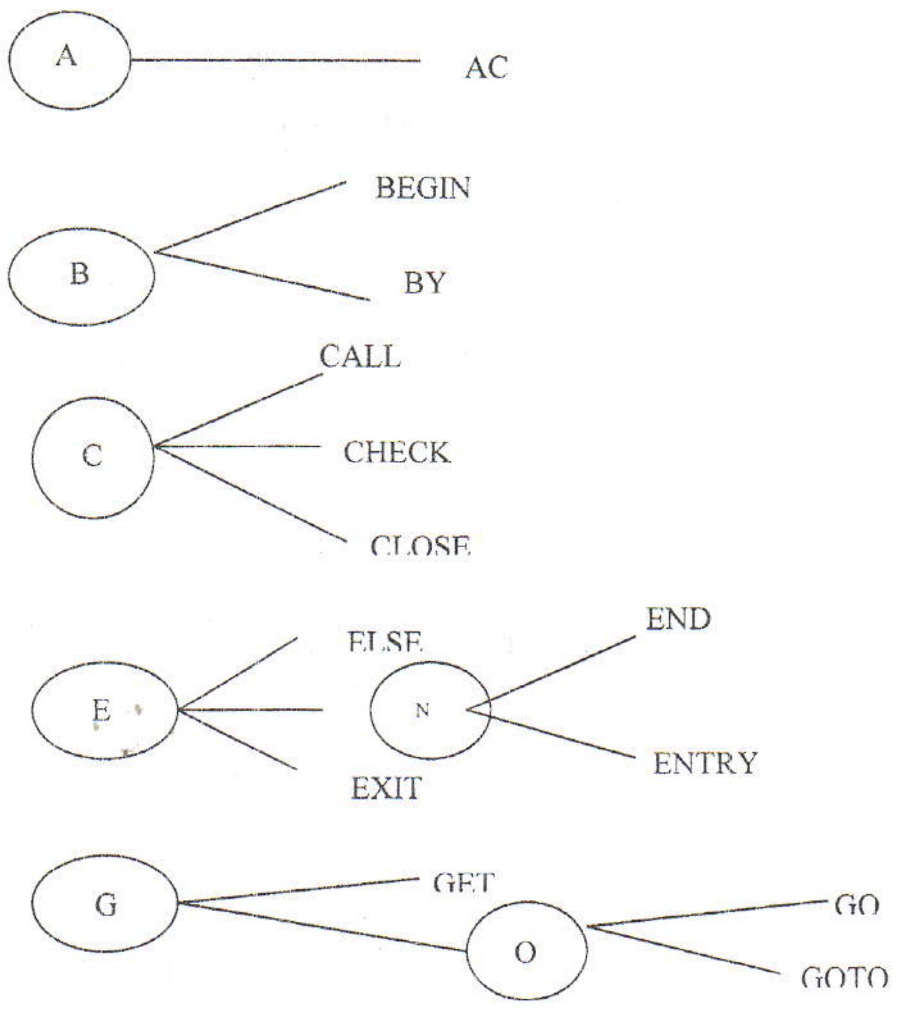
3-The entry corresponding to label n transfers us to node 11.

4-At node 11, the letter (d) is finally used to find the desired word.

The recall that the encoding procedure places the associations (Ai,Bi) at or near system energy minima.

The synaptic matrix encodes my particular computational problem, storing (A1,B1) and (A2,B2) in a parallel distributed network. The matrix element Mij indicates the symmetric (distance dependent) synapse between neurons Ai and Bi.

They are some statuses for representation of characters figure (3) shows:



figure(3): some statuses for representation of characters

We use character pairs of pixels 5×3 .For network processing , each one of the prototype is victimized row by row and black pixels are coded by +1,while white pixels are coded by -1.

Figure(4) shows this representation for character(A) and character (C).

The procedure which application a BAM net to associate letters with simple bipolar codes in my work is given by 5*3 pattern for each character. To explain the implication for Row 2 in tabale(1) we know that this row has 2 words,the first is (AC) and the second is (call). The word (AC) representation in BAM net is as in the following bipolar codes:

A(1-,1)

$$\begin{pmatrix} 1 & -1 \\ -1 & 1 \\ 1 & -1 \\ -1 & 1 \\ 1 & -1 \\ -1 & 1 \\ -1 & 1 \\ -1 & 1 \\ -1 & 1 \\ -1 & 1 \\ 1 & -1 \\ -1 & 1 \\ -1 & 1 \\ 1 & -1 \\ -1 & 1 \end{pmatrix}$$

C(1,1)

$$\begin{pmatrix} -1 & -1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ -1 & -1 \\ -1 & -1 \\ 1 & 1 \\ -1 & -1 \\ -1 & -1 \\ 1 & 1 \\ -1 & 1 \\ -1 & -1 \\ -1 & -1 \\ 1 & 1 \\ 1 & 1 \end{pmatrix}$$

weight matrix to store both

$$\begin{pmatrix} 0 & -2 \\ 0 & 2 \\ 2 & 0 \\ 0 & 2 \\ 0 & -2 \\ -2 & 0 \\ 0 & 2 \\ -2 & 0 \\ -2 & 0 \\ 0 & 2 \\ 0 & -2 \\ -2 & 0 \\ -2 & 0 \\ 2 & 0 \\ 0 & 2 \end{pmatrix}$$

To see the bidirectional nature of the net, observe that Y vector be used as input .For signals from Y- layer to X-layer, the weight is the transpose of the vector W.

$$W^T= \begin{bmatrix} 0 & 0 & 2 & 0 & 0 & -2 & 0 & -2 & -2 & 0 & 0 & -2 & -2 & 2 \\ -2 & 2 & 0 & 2 & -2 & 0 & 2 & 0 & 0 & 2 & -2 & 0 & 0 & 0 \end{bmatrix}$$

For the input vector associated with pattern A namely(-1,1), we have

$$(-1,1)W^T=(-1,1) \begin{bmatrix} 0 & 0 & 2 & 0 & 0 & -2 & 0 & -2 & -2 & 0 & 0 & -2 & -2 & 2 \\ -2 & 2 & 0 & 2 & -2 & 0 & 2 & 0 & 0 & 2 & -2 & 0 & 0 & 0 \end{bmatrix}$$

$$= (-2 \ \ 2 \ \ -2 \ \ 2 \ \ -2 \ \ 2 \ \ 2 \ \ 2 \ \ 2 \ \ 2 \ \ -2 \ \ 2 \ \ 2 \ \ -2 \ \ 2)$$

$$= (-1 \ \ 1 \ \ -1 \ \ 1 \ \ -1 \ \ 1 \ \ 1 \ \ 1 \ \ 1 \ \ 1 \ \ -1 \ \ 1 \ \ 1 \ \ -1 \ \ 1)$$

Similarly, if we input the vector associated with pattern C, we must obtain $(1,1)W^T$.

The matrias of (A) and (C) are :

A

-1	1	-1
1	-1	1
1	1	1
1	-1	1
1	-1	1

C

-1	1	1
1	-1	-1
1	-1	-1
1	1	-1
-1	1	1

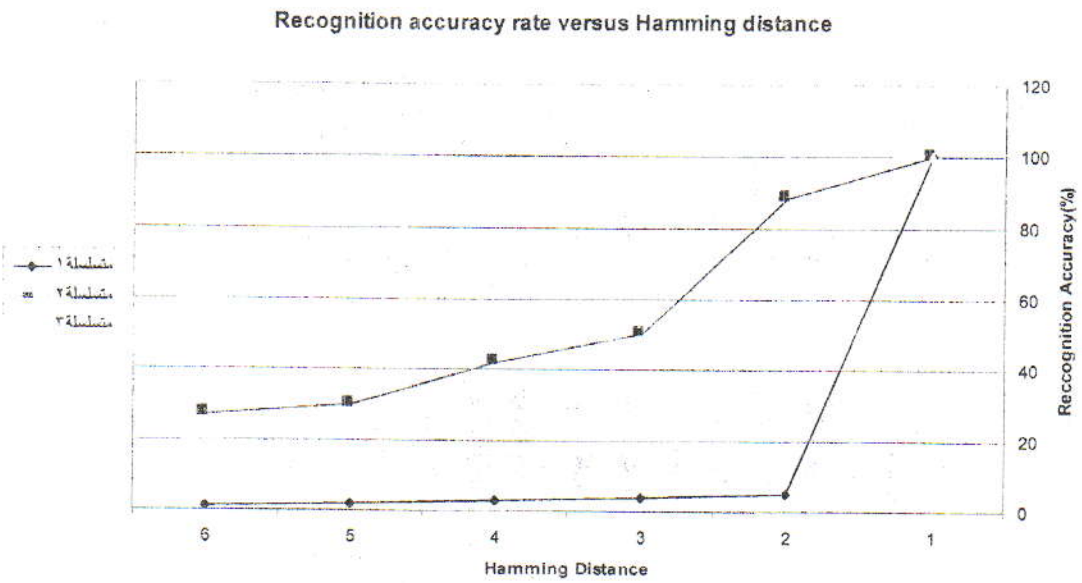
Figure(4): The matrixes of (A) and (C)

The upper bound on the memory capacity of the BAM is corresponding on n and m where n is the number of X-layer unites and m is the number of Y-layer unites.

We show that it can be extended to$(2^n,2^m)$if an appropriate nonzero threshold value is chosen for each unit,Choice was based on a combination of heuristics and an exhaustive search.

The hamming distance of the word (AC) differ in the 3rd ,6^{th},8^{th},9^{th},12^{th},13^{th} and 14^{th} positions. This gives an averages hamming distance between these vectors of 7/15 .The average hamming distance between corresponding Y vector is 12.

Figure(5) shows the Recognition accuracy rate versus Hamming distance where the first series is BAM . The second is trie structure and the third is (BAM and trie structure)



Figure(5): The Recognition accuracy rate versus Hamming distance

7-Conclusions

1-BAM become saturated when the number of patterns stored is greater than the minimum of the input layer count and the output layer node count.

2-each element in M-ary node be either a pointer to another trie node, pointer to a binary trie node, or string descriptor giving the address and length of key being stored in the string space.

Similarly, no leaf node contains a left pointer, a right pointer ,and a string descriptor. A leaf node contains only string descriptor.

3-Input patterns that are separated by a small Hamming distance are mapped to output vectors that are also so separated , while input vectors that are separated by large Hamming distance go to correspondingly distant(dissimilar) output patterns.

This is analogous to the behavior of a continuous function.

4-The best compromise situation in terms of space and running time occurs when only a few levels of a trie are used for the first few characters of the key and then some other structure.

References

[1] B. David , T. John ," Understanding Algorithms and data structure" , McGraw-Hill companies,1997.

[2] B. Kosko, "Bidirectional associative memories," IEEE Trans. Syst.,Man,Cybern., vol. 18, pp. 49–60, Jan. 1988.

[3] B. Kosko, "constructing An associative memories,", Byte , p 137-145 , September , 1987.

[4] G. Paul , "An introduction to data structure with applications " , McGraw-Hill companies ,NewYork,1992.

[5] H. Kang, "Multilayer associative neural network (MANN): Storage ca-paucity versus perfect recall," IEEE Trans. Neural Networks, vol. 5, pp.812–822, Sept. 1994.

[6] P. K. Simpson, "Higher-ordered and interconnected bidirectional associative memories," IEEE Trans. Syst., Man, Cybern., vol. 20, pp.637–653, May/June 1990.

[7] W.Yingquan and A. dimitris," A feed forward Bidirectional Associative Memory" , IEEE Transactions on neural networks, Vol.11, No.4, July, 2000.