The role of attention in the context of associative memory

Andreas Wichert
Department of Informatics
INESC-ID / IST - Technical University of Lisboa
Portugal
andreas.wichert@ist.utl.pt

Abstract

We model the mechanism of the retrieval of associations from the associative memory during visual scene analysis. During the analysis of the visual scene the retrieval phase of the associative memory is divided into two stages: the attention stage and the binding stage. In the attention stage, an attention window selects patterns representing objects for further access. In the binding stage, the selected patterns form an address vector. The behavior of the model is demonstrated by theoretical analysis and empirical experiment.

Keywords: associative memory, attention, binding, visual scene analysis, searchlight

1 Introduction

Vector representation corresponds to a pattern that mirrors the way the biological sense organs describe the world. A visual scene can be represented by objects and their position in the visual field. Each object is represented by a sub-vector of the vector representing the visual scene. We define a visual category as a set of prototypical visual objects. Our definition is motivated by the verbal category definition. A verbal category is a set of prototypical features [Osherson, 1987], such as red, round and sweet [Tversky, 1977, McClelland and Rumelhart, 1985]. A visual category "tower" is represented in the blockworld as shown in Figure 1. It corresponds to a set of prototypical visual object at certain position. In this paper we model the mechanism of the retrieval of such categories (set of objects) by the associative memory during visual scene analysis.

Suppose 7 objects were recognized in the visual scene. We indicate the 7 visual objects at certain position of the scene by symbols A, B, C, D, E, F, G. The task is to identify a category formed by visual objects represented by the

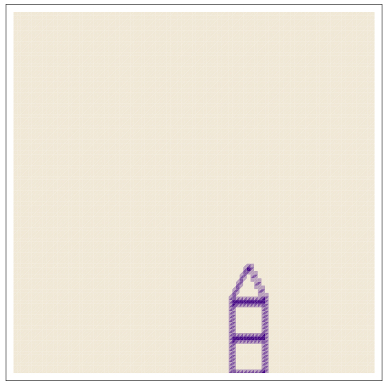

Figure 1: Category “tower” in the blockworld. Blocks can be placed in different positions and picked up and set down. There are two different classes of blocks: cubes and pyramids.

set B, C, G. The task is trivial when working with sets. We check if each of the symbols B, C, G is present in the set that represents the scene. We verify if a set representing a category is a subset of the set representing a scene.

However if the category (set of objects) is stored in an associative memory the task is non trivial. In an associative memory we do not have direct access to the stored information. An associative memory operates on vectors of fixed dimensions. Two pairs of these vectors are always associated and this process of association is called learning. The first of the two vectors is called the address vector and the second, the retrieved vector. After learning, the address vector is presented to the associative memory and the retrieved vector is determined. This process is called association. There is a distinction between heteroassociation and auto-association. An auto-association is present when the retrieved vector represents the reconstruction of the faulty address vector. An heteroassociation is present if both vectors are different. In our model we store auto-associations, means the the address vector and the retrieved vector are the same. Once an retrieved vector is determined, the similarity of the determined retrieved vector and the address vector is calculated. The greater the similarity, the more probable is that the corresponding address-vector was stored in the associative memory.

A set of objects (a category) is represented by a vector by concatenating the sub-vectors which represent the objects. For m sub-vectors there are $m!$ possible orderings of the corresponding sub-vectors. To verify if a set of m sub-vectors representing a category is a subset of the set of n sub-verctors representing a scene there are $\frac{m!}{(n-m)!}$ orderings. We suggest some mechanisms motivated by the searchlight theory to reduce this huge number.

2 Additive associative memory

We divided a vector representing a visual scene into n sub-vectors. A category is represented by m sub-vectors with $m \leq n$. There are L possible m-permutation [Wichert et al., 2008, Wichert, 2009]

$$L = Perm(n, m) = \frac{m!}{(n-m)!} \tag{1}$$

For n=7 and m=3 we would need to pose 210 queries to the associative memory, see Figure 2.

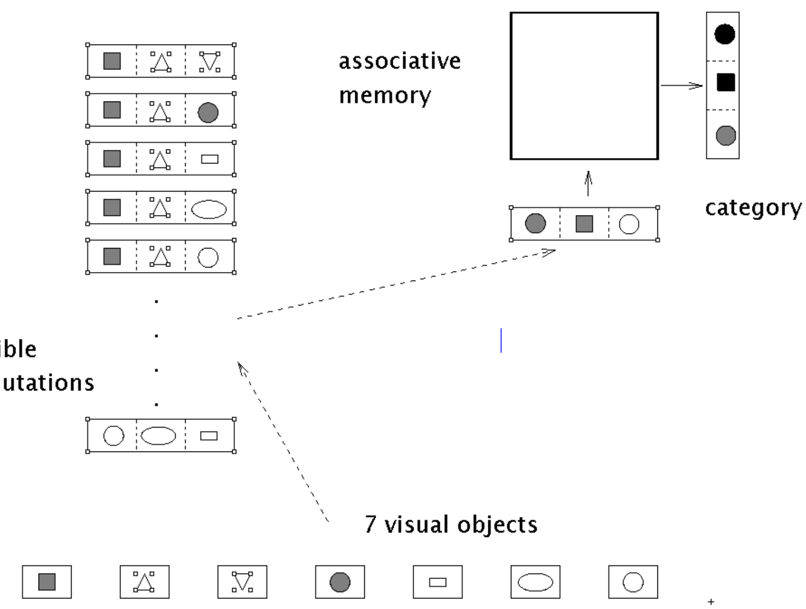

Figure 2: In the retrieval phase L permutations are formed. Each permutation represents a address vector $\vec{x_i}$ $, i \in \{1, \ldots, L\}$. Some of the address vectors represent a category, given that the determined retrieved vector is similar to the address vector.

The number of queries can be reduced, given that the associative memory has the additive property. Additive property of an associative memory indicates that if divided it into parts, each of the parts represents an associative memory, see Figure 3. A part of the associative memory is assigned to the sub-vector representing an object. The whole associative memory stores auto-associations. The parts store hetero associations. The address vector of a part is one sub-vector, the retrieved vector is composed of m concatenated sub-vectors.

An example for an associative memory with the additive property is the formal neural net model integrating the assembly concept [Palm, 1982], also called

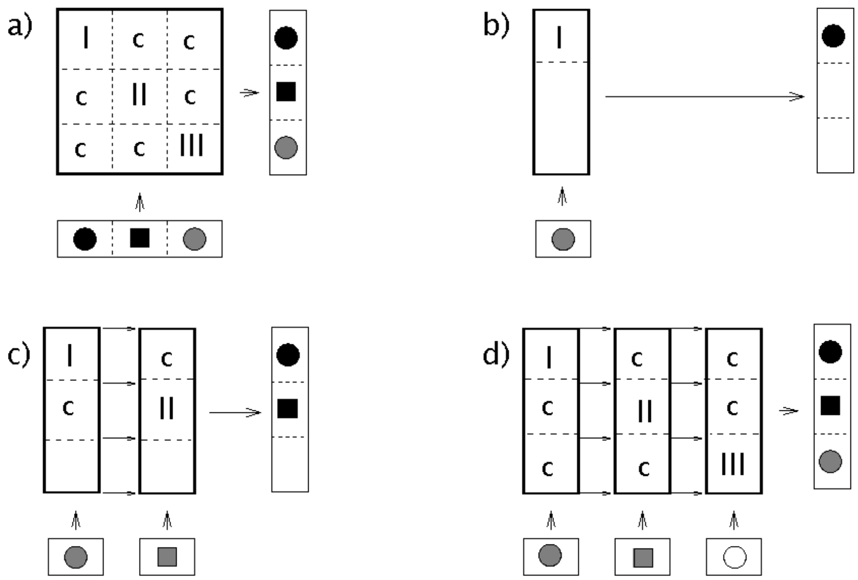

Figure 3: Additive property of associative memory. An associative memory can be divided into parts. Each par itself is an associative memory. A part of the associative memory is assigned to the sub-vector representing an object. The whole associative memory stores auto-associations, the parts which are associative memories store hetero associations. a) An association is stored. I, II, III represent the correlation between the vectors representing the objects, stored in weights (auto-association). C represents correlation between different objects (hetero association). b) Recall of a part of a noisy association by a part of associative memory (different grey value as learned). No correlation between objects is taken into account c) Recall of two parts of an association. by two parts of associative memory. d) Recall of the association by all three parts, the whole associative memory.

Lernmatrix. The biological and mathematical aspects of the Lernmatrix were studied by G. Palm [Palm, 1982, Palm, 1990, Fransén, 1996, Wennekers, 1999]. It was shown that Donald Hebb's hypothesis of cell assemblies as a biological model of internal representation of events and situations in the cerebral cortex corresponds to the formal Lernmatrix model. The Lernmatrix [Steinbuch, 1961, Hecht-Nielsen, 1989] is composed of a cluster of units which represent a simple model of a real biological neuron. The unit is composed of weights which correspond to the synapses and dendrites in the real neuron. They are described by w_{ij} in Figure. 4. T is the threshold of the unit. We call the Lernmatrix simply "associative memory" if no confusion with other models is possible.

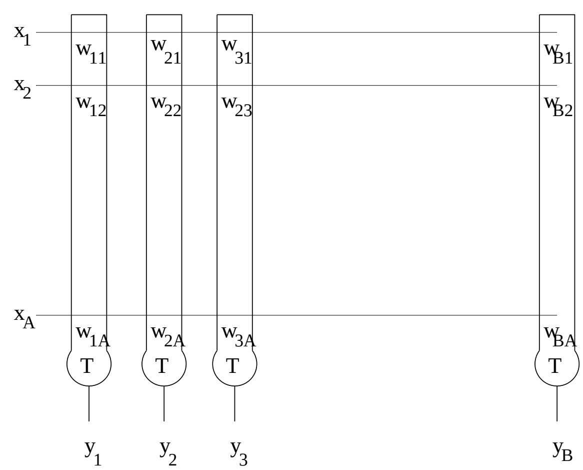

Figure 4: The associative memory is composed of a cluster of units.

Two pairs of binary vectors are associated, this process of association is called learning. The first of the two vectors is called the address vector and the second, the retrieved vector. After learning, the address vector is presented to the Lernmatrix and the retrieved vector is determined.

Learning In the initialization phase of the associative memory, no information is stored. Because the information is represented in weights, they are all initially set to zero. In the learning phase, pairs of binary vector are associated. Let $\vec{x}$ be the address vector and $\vec{y}$ the retrieved vector, the learning rule is:

$$w_{ij}^{new} = \begin{cases} 1 & if \quad y_i \cdot x_j = 1 \\ w_{ij}^{old} & \text{otherwise} \end{cases} \tag{2}$$

The whole associative memory stores auto-associations, the parts which are associative memories store hetero associations. For auto-associations $\vec{x} = \vec{y}$.

This rule is called the binary Hebbian rule [Palm, 1982]. Every time a pair of binary vectors is stored, this rule is used.

Retrieval In the *one-step* retrieval phase of the associative memory, a fault tolerant answering mechanism recalls the appropriate answer vector for a address vector $\vec{x}$. For the presented address vector $\vec{x}$, the most similar learned $\vec{x^l}$ address vector regarding the Hamming distance is determined and the appropriate retrieved vector $\vec{y}$ is identified. For the retrieval rule, the knowledge about the correlation of the components is sufficient. The retrieval rule for the determination of the retrieved vector $\vec{y}$ is:

$$y_i = \begin{cases} 1 & \sum_{j=1}^{A} w_{ij} x_j \geq T \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

where T is the threshold of the unit. The threshold is set as proposed by [Palm et al., 1997] to the maximum of the sums $\sum_{j=1}^{A} w_{ij} x_j$:

$$T := \max_{1 \leq i \leq B} \Big\{ \sum_{j=1}^{A} w_{ij} x_j \Big\}. \tag{4}$$

Only the units which are maximal correlated with the address vector are set to one.

Backward projection For the computation of the reliability of a part of the associative memory the answer of a backward projection is required. A part is a associative memory that stores hetero associations. The retrieved patterns of parts are sub-vectors of the address vector.

The backward projection corresponds to a bidirectional associative memory (BAM) [Kosko, 1992]. Anatomical studies suggest the presence of cell groups that project reciprocally onto each other [Braitenberg and Schüz, 1991]. This time the learned matrix is cued with the retrieved vector and the best address vector is retrieved. Formally, $\vec{y}$ is the address vector, and the retrieved vector which should be determined is $\vec{x^l}$. The categorization rule for the determination of the retrieved vector $\vec{x^l}$ is:

$$x_j^l = \begin{cases} 1 & \sum_{i=1}^{B} w_{ij} y_i \geq T^* \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

This means that the synaptic matrix used is a transposition of the matrix which is used for the forward projection. T^* is the threshold of the unit. The threshold is set to the maximum sum $\sum_{j=1}^{B} w_{ij} y_j$:

$$T^* := \max_{1 \leq j \leq A} \Big\{ \sum_{i=1}^{B} w_{ji} y_i \Big\}. \tag{6}$$

Reliability of the answer Let $\vec{x}$ be the question vector and $\vec{y}$ the retrieved vector that was determined by the associative memory for example by a part of the associative memory. First, the vector $\vec{x^l}$ which belongs to the vector $\vec{y}$ is determined. These two vectors form together a vector pair $\vec{x^l}$ $\vec{y}$ which is stored in the associative memory. It was either created by learning, $\vec{x^l}$ and $\vec{y}$ were learned together, or created through overlap with other already learned vector pairs. The vector $\vec{x^l}$ is determined by a backward projection of the vector $\vec{y}$. In the second step, the similarity of the stored address vector $\vec{x^l}$ to the actually presented vector $\vec{x}$ is determined. The greater the similarity of the vector $\vec{x^l}$ to the vector $\vec{x}$, the more reliable the retrieved vector $\vec{y}$.

3 Permutations

If the associative memory has the additive property then the number of L possible m-permutation $L = Perm(n, m) = \frac{m!}{(n-m)!}$ can be reduced. We divide the associative memory into m parts. A vector representing a visual scene is divided into n sub-vectors. The parts of an additive associative memory can be ordered so that l calls of the associative memory (or $l \cdot m$ calls of the associative memory parts) are performed.

$$l = \frac{1}{m} \cdot \sum_{i=1}^{n} \frac{n!}{(n-i)!} < L. \tag{7}$$

Proof

Each part is identified by index $i \in \{1, .., m\}$. The index $j \in \{1, .., n\}$ indicates the a sub-vector representing an object. Notation p(i,j): i shows the part number, j shows which sub-vector that is the address vector of this part. $part(2, 3)$ second part of the associative memory with address sub-vector 3.

The first $part(1, j)$ of the associative memory computes n times the retrieved vectors for the n objects so that the reliability of the answer can be determined.

$$\sum_{j=1}^{n} part(1, j) = n \tag{8}$$

For each n first $parts(1, j)$ with $j \in \{1, .., n\}$ of the associative memory there are $n - 1$ different combinations for the second parts of the associative memory, namely the remaining $n - 1$ objects, excluding the one which was already processed by the first part of the associative memory.

In total $n \cdot (n - 1)$ copies of the second part of the associative memory.

The number of copies of a part of an associative memory can be defined recursively, $i,\ i \in \{1, \ldots, m\}$

$$\#part(1) = n$$

number of copies of the first part of the associative memory,

$$\#part(i + 1) = \#part(i) \cdot (n - i).$$

number of copies of the $i + 1$ part of the associative memory.

Summed there are:

$$\sum_{i=1}^{m} \#part(i) = \sum_{i=1}^{m} \frac{n!}{(n-i)!} \tag{9}$$

The number of associative memories corresponds to l and the number of the parts is m times more, so the number of pars is $l \cdot m$,

$$l \cdot m = \sum_{i=1}^{m} \frac{n!}{(n-i)!} \leq \frac{n!}{(n-m)!} \cdot m = L \cdot m \tag{10}$$

Example: All possible 3-permutation of 7 objects are composed to form address vectors, $n = 7$, $m = 3$ see Figure 2. In this case $L = 210 = Perm(10, 3) = \frac{7!}{(7-3)!}$.

- $\#part(1) = 7$ copies of the first part.
- $\#part(2) = 7 \cdot (7 - 1) = 42$ copies of the second part.
- $\#part(3) = 42 \cdot (7 - 2) = 210$ copies of the third part.

There are 259=7+42+210 parts of the associative memory, the parts would correspond to $l = 86.3 = 259/3$ associative memories.

This number l can be reduced considerably by the fact that the parts are associative memories [Wichert et al., 2008]. The entire associative memory can be composed only from those parts which recognized some objects. By this constraint, the number of possible combinations is reduced. However this doesn't mean, that all recognized combinations of m objects represent categories. In Figure 5 we see the arrangement of the parts in m layers repeated over n objects. This arrangement is used for the retrieval phase. The address vectors of the parts are represented by the connection between them and the corresponding objects. In the attention stage the retrieved patterns of $m \cdot n$ parts are determined, and the corresponding parts whose reliability of the answer are above a certain threshold are marked. In the binding stage the associative memories are formed successively from marked parts over different objects. The formed associative memories determine the retrieved vectors.

1. For all parts $\forall i,\ i \in \{1, \ldots, m\}$ and for all sub-vector $\forall j,\ j \in \{1, \ldots, n\}$ $part(i, j)$ the retrieved vectors and the reliability of the answer are determined $(n \cdot m)$.
2. If the reliability of the answer of the $part(i, j)$ is above a certain threshold, the part is marked.
3. From different marked parts whose retrieved patterns were determined by different sub-vectors associative memories are formed.
4. The corresponding retrieved patterns of whole associative memories are determined.

R calls of the associative memory are performed. If (i) is the number of marked parts $p(i,j)$ then

$$R \leq n + \frac{1}{m} \cdot \sum_{i=1}^{m} \prod_{j=1}^{i} v(j) \leq l < L \tag{11}$$

Proof

For all m parts we determine over n objects whether the objects are recognized or not, $m \cdot n$ times. The parts which recognize objects are marked, $v(i)$ is the number of marked parts of number i. The number of copies of a part of an associative memory can be defined recursively, $i,\ i \in \{1, \ldots, n\}$

$$\#\mathit{marked\ part}(1) = v(1)$$

number of copies of the first part of the associative memory,

$$\#\mathit{marked\ part}(i+1) = \#\mathit{marked\ part}(i) \cdot v(i+1)$$

number of copies of the $i+1$ part of the associative memory. Summed there are:

$$\sum_{i=1}^{m} \mathit{marked\ part}(i) \leq \sum_{i=1}^{n} \prod_{j=1}^{i} v(j) \tag{12}$$

marked parts of the associative memory, less than or equal is used in the equation because marked parts over the same object should be not considered.

The permutational grow is avoided by the proposed search mechanism resulting in Equation 11. In the next section we describe a coding mechanism that avoids combinatorial explosion by possibly representing the same object many times at different locations.

4 Coding

The visual system recognizes objects in an image. It was suggested [Gross and Mishkin, 1977] that the brain includes two mechanisms for visual categorization [Posner and Raichle, 1994]: one for the representation of the object and the other for the representation of the localization [Kosslyn, 1994]. The first mechanism is called the *what* pathway and is located in the temporal lobe. The second mechanism is called the *where* pathway and is located in the parietal lobe. According to this division, the identity of a visual object can be coded apart from its location. A visual scene can be either represented by an image or by objects and their position in the visual field. Objects are represented by pictograms together with their corresponding position in the image. This is a simple form of structured and compressed representation of a mental image. The visual buffer consists of the categorical and coordinate spatial encoding in the parietal and the temporal lobes and of an image representation in the occiptal lobe. The visual buffer

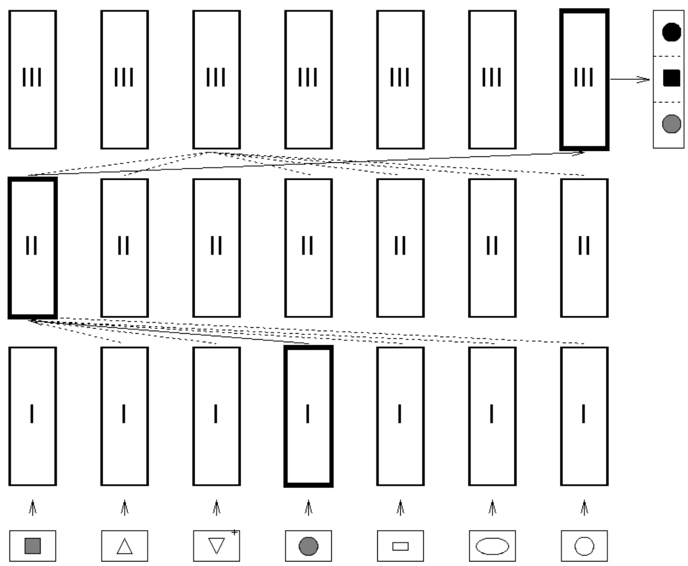

Figure 5: The retrieval phase is subdivided into two stages: the attention stage and the binding stage. The figure shows the arrangement of the parts in three (=m) layers repeated over seven (=n) objects. Because each part is an associative memory, we can arrange the parts hierarchically over the sub-vectors representing the objects of visual the scene. An object is a address vector to a part (indicated in the figure by a part over the object). We have a hierarchy of three parts. However, in our figure, the connections are not shown. This arrangement is used for the retrieval phase. The question vectors of the parts are represented by the connection between them and the corresponding objects. In the attention stage, parts are marked. In the binding stage, the associative memories are formed. From the thick boxes, the association of the example of Figure 1 and 2 is formed.

in the parietal and the temporal lobes is formed by a fixed number of the so called *cognitive entities* [Anderson, 1995]. Cognitive entities represent objects and their position in the image. Each cognitive entity represents the identity of the object and its position is given by Cartesian coordinates.

Cognitive entities are represented in the temporal and the parietal lobes by associative fields. A cognitive entity corresponds to a convergence zone [Damasio and Damasio, 1994] with many feedforward/feedback loops between the neurons of the corresponding associative fields.

4.1 Modeling cognitive entities

The identity of an object is represented by a binary pattern which is normalized for size and orientation. Its location in the x-axis is represented by a binary vector of the size of the abscissa of the pictogram representing the object. The location in the y-axis is likewise represented by a binary vector of the size of the coordinate of the pictogram representing the object. A binary bar of the size and position of the object in the pictogram of the state represents the location and size (see Figure 6) in each of those vectors. The three vectors that compose the cognitive entity are called associative fields. Each associative field is represented by a binary vector of a fixed dimension; each cognitive entity is formed by the concatenation of the associative fields.

Through the binarization, we achieve a higher level of simplification. This model can be easily extended by additional associative fields to represent additional attributes of the objects, beside the position in an image. For example, its color, or its position in a three dimensional mental image.

By this form of coding is possible to represent an object many times at different locations avoiding a combinatorial explosion. Each position of the same object is stored in an associative memory represented by the weight matrix. Each different position corresponds to a different vector part (associative field), represented by particular zones of the weight matrix.

4.2 Associations

A cognitive entity is represented by a binary vector formed by the concatenation of binary vectors which represent the three associative fields. For the category tower (see Figure 1) the address and retrieved vectors are represented by a binary vector formed by the concatenation of three binary sub-vectors which represent the cognitive entities. Both the question and the answer vectors have dimension 900 because each cognitive entity is described by a binary vector of dimension 300. The representation of the category "tower":

```
0000110000
0000110000
0001001000
0001001000
0010000100
0010000100
0100000010
0100000010
1000000001
1111111111
```

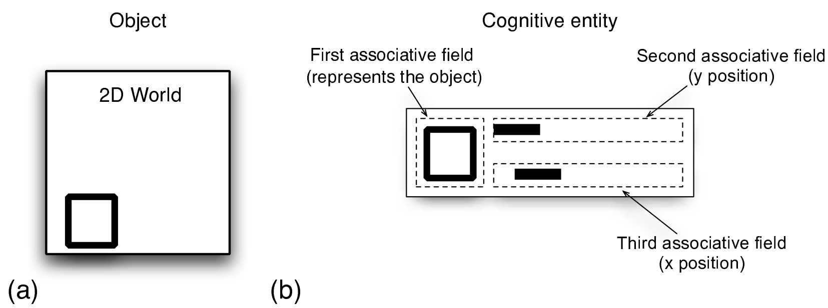

Figure 6: Representation of an object in a 2D world (a) by a cognitive entity (b). The identity of an object is represented in the first associative field by a binary pattern which is normalized for size and orientation. Its location corresponding to the abscissa is represented by a binary vector in the second associative field. The location corresponding to the ordinate is likewise represented by a binary vector in the third associative field of the size of the ordinate of the pictogram representing the state. A binary bar of the size and position of the object in the pictogram of the state represents the location.

```
0000000000000000000011111111110000000000000000000000000000000000000000000000000000000000000000000000

0000000000000000000000000000001111111111000000000000000000000000000000000000000000000000000000000000

1111111111
1000000001
1000000001
1000000001
1000000001
1000000001
1000000001
1000000001
1000000001
1111111111

0000000000111111111000000000000000000000000000000000000000000000000000000000000000000000000000000000

0000000000000000000000000000001111111111000000000000000000000000000000000000000000000000000000000000

1111111111
1000000001
1000000001
1000000001
1000000001
1000000001
1000000001
1000000001
1000000001
1111111111

1111111111000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000

0000000000000000000000000000001111111111000000000000000000000000000000000000000000000000000000000000
```

For the category "flower" (see Figure 7) the address and retrieved vectors are represented by a binary vector formed by the concatenation of two binary sub-vectors which represent the cognitive entities.

Both the question and the answer vectors have dimension 600 because each cognitive entity is described by a binary vector of dimension 300. Note that the

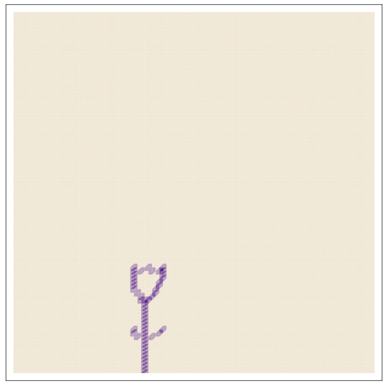

Figure 7: Category "flower" is composed of a stalk and a florescence.

binary bar representing the size and position of the y-axis of the stalk has a double size compared with the florescence. The representation of the category "flower":

```
01000100010
01010010110
01000000010
01000000010
01000000100
01000000100
01000001000
00100001000
00010010000
00011100000

0000000000000000000011111111100000000000000000000000000000000000000000000000000000000000000000000000

0000000000111111111100000000000000000000000000000000000000000000000000000000000000000000000000000000

00001000000
00001000000
00001000000
01001001000
00101100000
00001000000
00001000000
00001000000
00001000000
00001000000

1111111111111111111100000000000000000000000000000000000000000000000000000000000000000000000000000000

0000000000111111111100000000000000000000000000000000000000000000000000000000000000000000000000000000
```

Ten associations representing ten different positions of the category "tower" are learned by the associative memory which is composed of three parts. The same first two parts of the associative memory are used to learn ten associations representing ten different positions of the category "flower". Twenty associations are learned corresponding to categories tower and flower in different possible positions. After learning, a weight matrix of dimension $x = y = 900$ emerges, each of the three parts with the size $x = 300$ and $y = 900$

4.3 Retrieval

In the corresponding example we indicate the recognition of two categories, flower and tower of the pictogram of the Figure 8 represented by nine different objects.

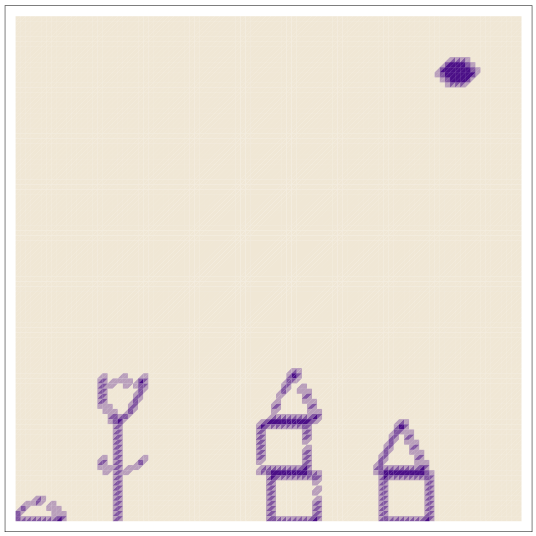

Figure 8: Visual representation of the world with nine different objects and with categories "flower" and "tower" with noise. The state of the world is described by pictograms of 100×100 pixels. The pictograms are represented by nine cognitive entities ($= n$).

The corresponding retrieved vectors values of those associative memories are determined. If the If the reliability of the answer of the $part(i,j)$ is above a certain threshold, the part is marked. $3+2+1$ parts of the associative memory are marked, see Table 1. Three possible categories are recognized corresponding to the flower, sub-category tower and the tower, see Table 2 and 3.

5 Visual attention and the human brain

We try to describe the process of forming the permutations through a biological plausible model. Instead of forming permutations of all possible cognitive entities, we will only form permutations of the interesting ones.

The retrieval phase is subdivided into two stages: the *attention stage* [Allport, 1989, Posner and Raichle, 1994] and the *binding stage*. The process of marking during the attention stage corresponds to the mechanism of attention window that selects a pattern in the visual buffer for further access to the visual system in the human brain [Kosslyn, 1994]. The successive examination of the cognitive

$part(i,j)$	j=1	j=2	j=3	j=4	j=5	j=6	j=7	j=8	j=9
part(1,j)		M		M			M		
part(2,j)			M		M				
part(3,j)						M			

Table 1: Attention stage: If the reliability of the answer of the $part(i,j)$ is above a threshold, the part is marked, indicated by M. $7+3+1$ parts of the associative memory are marked. Rows correspond to part numbers and columns correspond to cognitive entities. In this example noise is present. $3+2+1$ parts of the associative memory are marked. Three parts are marked for the first part of the associative memory; two for the second part of the associative memory and one for the third part of the associative memory.

$part(1,j)$	$part(2,j)$	marked
2	3	M
2	5	
4	3	
4	5	M
7	3	
7	5	

Table 2: Binding stage of two parts: The corresponding retrieved patterns of whole associative memories determined. If the reliability of the answer is above a threshold, a category is recognized, indicated by M. Two categories are recognized, flower and the sub-category tower, one cube on another (incorporated in the category tower, the pyramid is missing).

entities corresponds to the spotlight theory [Downing and Oinker, 1985].

Attention is linked to a spotlight that is focused on the cued location and shifted as necessary [Kosslyn, 1994, Posner and Raichle, 1994]. The brain is supposed to have an internal attentional searchlight that moves around from one visual object to the next with steps as big as 70ms.

The searchlight should fulfill following criteria:

- It should be able to sample the activity in the visual cortex and decide where the action is;
- It must be able to turn off its beam;
- It should move to the next place demanding attention and repeat the process.

Crick proposed that the searchlight is controlled by the thalamus [Crick, 2003]. In his view, the expression of the searchlight is the production of rapid firing in a subset of active thalamus neurons and the corresponding fields in the visual cortex.

$part(1,j)$	$part(2,j)$	$part(3,j)$	marked
2	3	6	
2	5	6	
4	3	6	M
4	5	6	
7	3	6	
7	5	6	

Table 3: Binding stage of three parts: The corresponding retrieved patterns of whole associative memories determined. If the reliability of the answer is above a threshold, a category is recognized, indicated by M. One category is recognized, the category tower (despite noise).

All visual input to the cortex passes through the thalamus, and there is a rapid movement of the searchlight from place to place. However the brain must know what it is searching for. That aspect involves the parietal and the temporal lobes. As stated, we suppose that the visual buffer consists of the object and coordinate spatial encoding in parietal and temporal lobe and of an image representation in the occiptal lobe. The searchlight mechanism is dependent on the objects represented by cognitive entities and on the associative memory.

During the examination, the interesting cognitive entities are determined by the associative memory.

Each part is tested if it is interesting in succession; if it recognizes an object represented by the cognitive entity then it is marked. How could the marking be realized in human brain? The marking of the corresponding parts could be represented by a synchronous oscillation of the cognitive entities [Gary and Singer, 1989], [Singer and Gary, 1995]. Only when the attention is focused on certain parts are they bound to a whole object [Triesmann and Gormican, 1988], [Damasio and Damasio, 1994], [Posner and Raichle, 1994].

In the following binding stage only the interesting cognitive entities are examined by the searchlight and bound together. It is proposed that there are two searchlight mechanisms, one that determines the interesting objects, and a second one that bounds the marked cognitive entities. The bounded cognitive entities serve as input to the associative memory, which determines if they correspond to any stored association. The described process of attention and binding doesnt include goal driven behavior [Corbetta et al., 2008]. During goal driven behavior top down signals dynamic interaction with sensory or bottom-up information by activating and deactivating certain associative memory parts. The proposed process of attention can be extend beyond visual scene analysis. We can form a global associative memory from different other associative memories by arranging them on the diagonal of the global associative memory and storing the correlation between different objects, see Figure 3. (visual objects, sount, tatile information, etc..)

6 Conclusion

The searchlight determines all possible permutations of objects representing a visual scene. This is necessary, because in an associative memory we do not have direct access to the stored information. Associative memory with additive property can be represented by the sum of its parts. Because each part is an associative memory, we can arrange the parts hierarchically over the subvectors representing the objects of visual the scene. This architecture reduces the computational complexity and highlights the role of attention during visual scene analysis.

7 Acknowledgments

This work was supported by Fundao para a Cencia e Tecnologia (FCT) (INESC-ID multiannual funding) through the PIDDAC Program funds.

References

[Allport, 1989] Allport, A. (1989). Visual attention. In Posner, M., editor, *Foundations of cognitive science*, pages 631–682. The MIT Press.

[Anderson, 1995] Anderson, J. A. (1995). *An Introduction to Neural Networks.* The MIT Press.

[Braitenberg and Schüz, 1991] Braitenberg, V. and Schüz, A. (1991). *The anatomy of the cortex. Statistics and Geometry.* Springer-Verlag.

[Corbetta et al., 2008] Corbetta, M., Patel, G., and Shulman, G. L. (2008). The reorienting system of the human brain: From environment to theory of mind. *Neuron*, 58:306–324.

[Crick, 2003] Crick, F. (2003). Function of the thalamic reticular complex: The searchlight hypothesis. In Baars, B. J., Banks, W. P., and Newman, J., editors, *Essential Sources in the Scientific Study of Consciousness.* MIT Press.

[Damasio and Damasio, 1994] Damasio, A. R. and Damasio, H. (1994). Cortical systems for retrieval of concrete knowledge: The convergence zone farmework. In Koch, C. and Davis, J. L., editors, *Large-Scale Neural Theories of the Brain*, chapter 3, pages 61–74. The MIT Press.

[Downing and Oinker, 1985] Downing, C. and Oinker, S. (1985). The spatial structure of visual attention. In Posner, M. and Marin, O., editors, *Attention and performance XI*, pages 171–187. Erlbaum, Hillsdale, N.J.

[Fransén, 1996] Fransén, E. (1996). *Biophysical Simulation of Cortical Associative Memory.* PhD thesis, Stockhokms Universitet, Dept. of Numerical Analysis and Computing Science Royal Institute of Technology, S-100 44 Stockholm, Sweden.

[Gary and Singer, 1989] Gary, C. and Singer, W. (1989). Stimulus specific neural oscillations in orientation columns of cat visual cortex. *PNAS*, 86:1698–1702.

[Gross and Mishkin, 1977] Gross, C. and Mishkin (1977). The neural basis of stimulus equivalence across retinal translation. In Harnad, S., Dorty, R., Jaynes, J., Goldstein, L., and Krauthamer, editors, *Lateralization in the nervous system*. Academic Press, New York.

[Hecht-Nielsen, 1989] Hecht-Nielsen, R. (1989). *Neurocomputing*. Addison-Wesley.

[Kosko, 1992] Kosko, B. (1992). *Neural Networks and Fuzzy Systems: A Dynamical Systems Approach to Machine Intelligence*. Prentice-Hall.

[Kosslyn, 1994] Kosslyn, S. M. (1994). *Image and Brain, The Resolution of the Imagery Debate*. The MIT Press.

[McClelland and Rumelhart, 1985] McClelland, J. and Rumelhart, D. (1985). Distributed memory and the representation of general and specific memory. *Journal of Experimental Psychology: General*, 114:159–188.

[Osherson, 1987] Osherson, D. N. (1987). New axioms for the contrast model of similarity. *Journal of Mathematical Psychology*, 31:93–103.

[Palm, 1982] Palm, G. (1982). *Neural Assemblies, an Alternative Approach to Artificial Intelligence*. Springer-Verlag.

[Palm, 1990] Palm, G. (1990). Assoziatives Gedächtnis und Gehirntheorie. In *Gehirn und Kognition*, pages 164–174. Spektrum der Wissenschaft.

[Palm et al., 1997] Palm, G., Schwenker, F., Sommer, F., and Strey, A. (1997). Neural associative memories. In Krikelis, A. and Weems, C., editors, *Associative Processing and Processors*, pages 307–325. IEEE Press.

[Posner and Raichle, 1994] Posner, M. I. and Raichle, M. E. (1994). *Images of Mind*. Scientific American Library, New York.

[Singer and Gary, 1995] Singer, W. and Gary, C. (1995). Visual feature integration and the temporal correlation hypothesis. *Ann.Rev.Neuroscience*, 18:555–586.

[Steinbuch, 1961] Steinbuch, K. (1961). Die Lernmatrix. *Kybernetik*, 1:36–45.

[Triesmann and Gormican, 1988] Triesmann, A. and Gormican, S. (1988). Feature analysis in early vision. *Psyhology Review*, 95:15–30.

[Tversky, 1977] Tversky, A. (1977). Feature of similarity. *Psychological Review*, 84:327–352.

[Wennekers, 1999] Wennekers, T. (1999). *Synchronisation und Assoziation in Neuronalen Netzen*. Shaker Verlag, Aachen.

[Wichert, 2009] Wichert, A. (2009). Sub-symbols and icons. *Cognitive Computation*, 1(4):342–347.

[Wichert et al., 2008] Wichert, A., Pereira, J. D., and Carreira, P. (2008). Visual search light model for mental problem solving. *Neurocomputing*, doi:10.1016/j.neucom.2007.08.019.