

# A theoretical model of the hippocampal-cortical memory system motivated by physiological functions in the hippocampus

Minoru Tsukada

*Department of Information & Communication Engineering  
Faculty of Engineering  
Tamagawa University*

*6-1-1, Tamagawagakuen, Machida, Tokyo 194, JAPAN*

## ABSTRACT

Based on the physiological evidence, we propose a hypothesis on how association and dissociation of event memories are made in the hippocampal-cortical memory system. It is postulated that the association / dissociation of memory is carried out by indexing the representations of events (memory contents) with temporal codes. The memory contents are supplied from the sensory association cortices, while the temporal codes are supplied from decision making / motivation area. The two inputs are mixed (indexing) in the ento-perirhinal area. Indexed signals are fed to hippocampus where connection / disconnection of memory contents occur depending on the kind of index. Finally, association / dissociation of event memories is made in the association cortex according to a covariance rule: two events memories are associated when direct cortico-cortical inputs and indirect inputs from the hippocampus are positively correlated through the consolidation made in hippocampus, and they are dissociated when two inputs are negatively correlated in consequence of the disconnection made in the hippocampus.

*Keywords: A hippocampal-cortical memory system, index theory, temporal pattern codes.*

### 1. The hippocampal-cortical memory system corresponds to a neural network with three layers

The hippocampal-cortical memory system consists of the following components; the neocortex, the hippocampus and the ento-perirhinal (ento-perirhinal, and parahippocampal cortices) area. The entorhinal cortex (EC) is the major source of projections to the hippocampus. It receives projections from the unimodal and polymodal areas of the frontal, temporal and parietal lobes of the neocortex, via the adjacent perirhinal (PR) and parahippocampal (TH/TH) cortical area. The entorhinal cortex also receive other direct input from the orbital frontal cortex, the cingulate cortex, the insular cortex and the superior temporal gyrus. Thus, the ento-perirhinal cortical area makes up a part of the interface between the neocortex and the hippocampus. The hippocampal networks consist of the dentate gyrus (DG), and the CA3 and CA1 regions where cells are interconnected by widespread and overlapping projections [1]. The widespread efferent projections leave the hippocampus by way of the subiculum (S) flow to the EC area and continue to the neocortex [2]. Physiologically Squire and Zola-Morgan [3] posited that the neocortex represents visual object quality (in area TE) and object location (in area PG) as long-term memory, while the hippocampus plays an important role at the time of learning in establishing long-term memory which is then stored in the neocortex. The hippocampal process, which works only temporary as short-term memory, serves as a device for forming conjunctions between ordinarily unrelated events or stimulus features which are processed and represented by distinct cortical sites. After learning, memory stored in the neocortex becomes independent of the intervening temporal lobe structure. Thus, we assume that the memory system is composed of neural networks with three layers; the input layer, the hidden layer and the output layer. The neocortex may be thought of as the input-output layer, where information is represented in terms of the activities of assemblies of neurons and it serves as long-term memory. The hippocampus corresponds to the hidden layer where short-term memory is developed. The ento-perirhinal cortical area is an interface between the neocortex and the hippocampus. Schematically this memory system is shown in Fig. 1.

### 2. Neural coding in the memory system

Almost all of the models of the associative memory studied previously are based upon the algorithm of taking auto- and cross-correlation of different inputs [4-7]. This process is a kind of computation of similarity of input patterns. To ensure an accurate recall, the above models store different patterns by transforming them into mutually orthogonal signals. However, it would be generally accepted a simple psychological inspection that the event

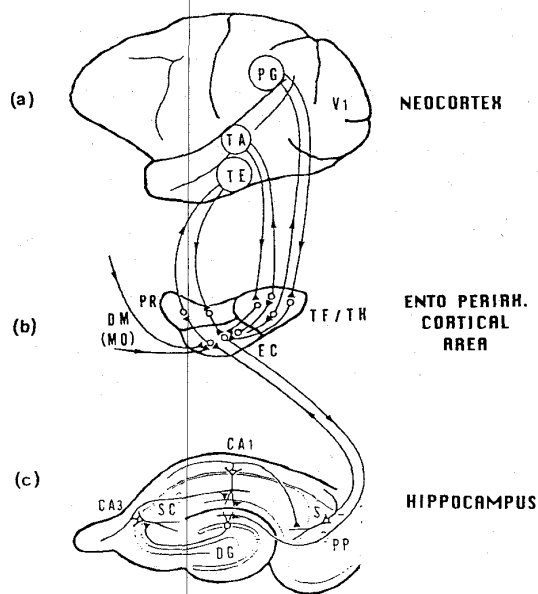


Fig.1. Schematic drawing of the hippocampal-cortical memory system.

(a) The networks in the neocortex show representations of visual object quality (in area TE), object location (in area PG) and information for auditory recognition (in area TA) as long-term memory.

(b) The entorhinal cortex (EC), perirhinal cortex (PR) and parahippocampal gyrus (TF/TH) are an interface between the neocortex and the hippocampus, and receive projections from the neocortex and are the major source of projections to the hippocampus. The ento-perirhinal cortices also receives projections from the decision making area (DM) or the motivational area (MO). All these projections are reciprocal.

(c) The hippocampus consists of the following stages: the dentate gyrus (DG), the CA3 and CA1 regions and plays an important role at the time of learning in establishing long-term memory.

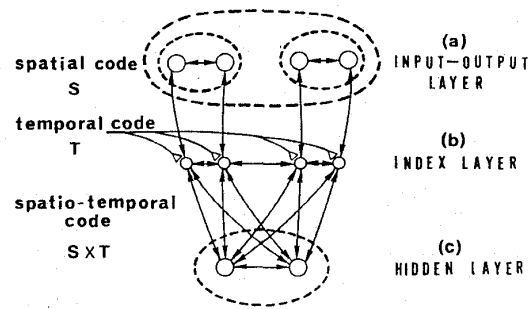


Fig.2. A proposed model of the hippocampal-cortical memory system.

The input-output layer of the network corresponds to the neocortex, the index layer corresponds to the ento-perirhinal cortical area and the hidden layer correspond to the hippocampus.

(a) Activation of an assembly of neurons in the neocortex represents a perceptual feature as a spatial code; visual object quality in area TE, visual object location in area PG, information for auditory recognition in area TA, etc..

(b) In the index layer, the spatial vector is indexed by a temporal code sent from the decision making area i.e., it is transformed into a spatio-temporal vector of which the temporal part represents the index and the spatial part represents the information content.

(c) The hippocampus works as a device for connecting/disconnecting information contents depending on the index. The role of this system is only temporary, and this hippocampal control becomes consolidated as permanent memory in the neocortex as time passes.

memories could be associated or dissociated independently of the similarity of the contents.

Here we propose a coding method in the memory system in which memories are associated or dissociated independently of the similarity of contents. In our hypothesis a memory is temporarily represented by a spatio-temporal pattern (a short-term memory) of which the spatial part represents the content of memory and the temporal part represents the index of memory. Memories are associated or dissociated by indexes. In this coding scheme, similar spatial vectors with near distance can be separated by assigning some particular index, and different spatial vectors with far distance can be connected by assigning some other index. Schematically this system is shown in Fig. 2.

## 2. 1. A sparse coding of assemblies of neurons in the neocortex

Several neurophysiologists suggest that the memory storehouse for final representations of events is the area of association cortices [8], in which events are represented by activities of groups or assemblies of neurons [9] rather than those of single neurons. Representation of the memory of some event seems to be distributed among

assemblies of neurons, as the distribution of synaptic strengths of neural networks. Pulm [10] and Amari [11] theoretically demonstrated that a sparse coding by assemblies of neuron is the most desirable method from the view point of information capacity. Physiological evidence supporting a sparse coding has recently been found in Inferior Temporal [8]. We postulate that the spatial pattern of the sparse activities of assemblies of neurons in the neocortex represents events. Memories of different modalities and submodalities are stored in different cortical areas: visual object quality in area TE, visual object location in area PG, information for auditory recognition in area TA, etc.

## 2. 2. A temporal coding in the ento-perirhinal cortical area as an indexer of memory

The ento-perirhinal cortical area may be functioning as an indexer to associate / dissociate memories by using temporally modulated patterns. In the ento-perirhinal cortical area, inputs from the neocortex represented by spatial patterns (events) are modulated temporally by inputs (indexes) from the decision making area. One of important decision making areas may be the frontal cortex. The interaction between both inputs generates a spatio-temporal pattern. One of effective temporal modulation (temporal code) would be rhythms with different second order statistics (serial correlation coefficients) of the impulse sequence. There is at present no direct evidence that the ento-perirhinal cortices use temporal code as an index. However, the possibility of using temporal code as an index is strongly supported by the fact that rhythmic activities are very common in the entorhinal cortex [12] and related areas [13-16].

## 2. 3. Learning rule at the early stage to associate / dissociate events in the hippocampus based on the index

Hippocampus works as a hidden network to construct a new memory structure in the association cortex through the interface of ento-perirhinal area. In this theory, it is assumed that hippocampal network associate / dissociate events according to the kind of indexes joined with them. We propose that hippocampus uses a temporal code as an index. We stress this because homosynaptic and associative LTP of CA1 neurons were shown to be highly sensitive to a temporal modulation of Schaffer collateral stimulation. Our observations showed that the stimulus with positive correlation of successive inter-spike intervals (ISIs) generated a large homo- and associative synaptic LTP, that of negative correlation was ineffective, and the stimulus with no correlation induced a moderate LTP [17,18,19]. Thus we hypothesize that the positive correlation works as the joining index code ( $T^1$ ) which associate different events, negative correlation works as a disjoining index code ( $T^0$ ) which dissociates two events. No correlation works as a neutral code which maintains the present state. One more assumption is introduced. That is, to associate / dissociate indexed events, rough synchronization (spatial correlation) of input events is required. Our proposal is that when two spatial vectors (the representation of events) are in synchrony with the index  $T^1$ , they are associated, and when they are in synchrony with the index  $T^0$ , they are dissociated. This is schematically shown in Fig. 3.

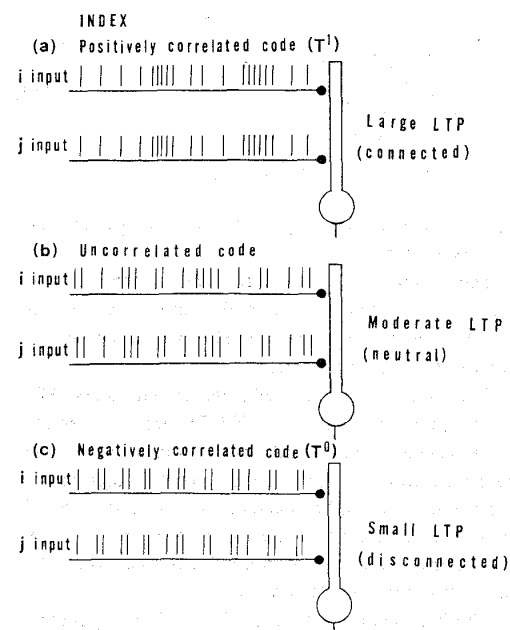


Fig.3. Schematic representation of connected/neutral/disconnected information in the hippocampal network based on the "index theory".

- (a) When both i and j inputs synchronously include the positively correlated temporal code as an index, they produce a large LTP and are connected.
- (b) When both i and j inputs are the synchronous and uncorrelated temporal code or asynchronous time code, they produce a moderate LTP which provides a neutral state.
- (c) When both i and j inputs synchronously include the negative correlated temporal code as an index, they produce a small LTP and are disconnected.

#### **2. 4. Learning rule at final stage to fix the memory in the association cortex as long term memory**

The association cortex functions as the memory store house for the final representations of events. This area receives informations about association / neutral / dissociation through back projection from hippocampus in the form of difference in the activity of neurons. It is hypothesized that association / dissociation of events A and B in the association cortex is determined according to the spatial covariance rule. In the spatial covariance rule, the connection from neuron a to neuron b is strengthened or weakened according to whether the firings of the two neurons are positively correlated or negatively correlated in the spatial domain. The connection is unchanged if there is no correlation between their firings.

Adopting this rule to the hippocampal-association cortex memory system, the mechanisms of the association / dissociation of memories of two events are described as follows: The memory representation of event A in the association cortex is associated with those of event B when the cell assembly of A receive hippocampal input of event B in spatially positive correlation with a input of event A representation. This positive correlation in the activity of neurons is obtained by associative LTP in hippocampus using index  $T^1$ . The two memory representations are dissociated when the inputs from association cortex (representation of event A) and hippocampus (event B representation) are negatively correlated. This negative correlation is achieved by dissociative LTP in hippocampus using index  $T^0$ .

#### **3. Conclusion**

In this paper, based on the physiological evidence, we proposed a theoretical model of a hippocampal-cortical memory system which performs operations to associate and to dissociate events in the memory space. The neocortex works as the storehouse for the final representations of events (long-term memory). The hippocampus works as a molding house which temporarily associate or dissociate different events (thought to be a short-term memory). The ento-perirhinal cortices work as an interface between the neocortex and the hippocampus, and give indexes to the representations of events.

We propose that temporal patterns given by second-order statistics are used as indexing codes. If the event representation in the association cortex is made by the distribution of synaptic weight among assembly of neurons (spatial vector), the indexing code is essentially orthogonal to it. The function of the ento-perirhinal cortices is to take cartesian product of spatial vector (event representation) and temporal vector (index). This operation (indexing) plays two important roles, one to give "orthogonality" to event representations and the other to expand "dimension" of them. These two aspects are indispensable to increase both the capacity and separability of event memory.

#### **References**

- [1] N. Ishizuka, J. Weber and D. G. Amaral, J. Comp. Neurol. 295, 580 (1990).
- [2] L. R. Squire, in the Biology of Memory, ed. by L. R. Squire and E. Lindenlaub (Schattauer Verlag, Stuttgart, 1990) p. 643.
- [3] L. R. Squire and S. Zola-Morgan, Science 253, 1380 (1991).
- [4] T. Kohonen, IEEE Trans. C-21, 353 (1972).
- [5] K. Nakano, IEEE Trans. SMC-12, 380 (1972).
- [6] S. Amari, IEEE Trans. Computers, C-21(11), 1197 (1972).
- [7] J. A. Anderson, Math. Biosci. 14, 197 (1972).
- [8] Y. Miyashita, Nature 1988, 335, 817 (1988).
- [9] K. Tanaka, H. Saito, Y. Fukuda and M. Moriya, J. Neurophysiol. 66, 170 (1991).
- [10] G. Palm, Bio. Cybern., 39, 125 (1981).
- [11] S. Amari, Neural Networks, 2, 451 (1989).
- [12] J. S. Mitchell, JB, Jr. Ranck, Brain Res. 189, 49 (1980).
- [13] J. Holsheimer, Exp Brain Res. 47, 309 (1982).
- [14] C. H. Vanderwolf, Electroenceph. Clin. Neurophysiol. 21 407 (1969).
- [15] S.E. Fox, S. Wolfson, and J. B. Ranck, Jr, Exp. Brain Res. 62, 495 (1986).
- [16] Y.J., Greenstein, C. Pavlides and J. Winso, Brain Res. 438, 331 (1988)
- [17] M. Tsukada et al., in Vision, Memory and the Temporal Lobe, ed. by E. Iwai and M. Mishikin.
- [18] A. Inoue et al., Annual Conference of JNNS 3, 133 (1992).
- [19] M. Tsukada, Concepts in Neuroscience, 3 213-224 (1992).