

# Emergent behaviors based on episodic encoding and familiarity driven retrieval

Emilia I. Barakova

RIKEN BSI, 2-1, Hirosawa, Wako-shi, Saitama, 351-0198 Japan.  
emilia@brain.riken.jp

**Abstract.** In analogy to animal research, where behavioral and internal neural dynamics are simultaneously analysed, this paper suggests a method for emergent behaviors arising in interaction with the underlying neural mechanism. This way an attempt to go beyond the indeterministic nature of the emergent behaviors of robots is made. The neural dynamics is represented as an interaction of memories of experienced episodes, the current environmental input and the feedback of previous motor actions. The emergent properties can be observed in a two staged process: exploratory (latent) learning and goal oriented learning. Correspondingly, the learning is dominated to a different extent by two factors: novelty and reward. While the reward learning is used to show the relevance of the method, the novelty/familiarity is a basis for forming the emergent properties. The method is strongly inspired by the state of the art understanding of the hippocampal functioning and especially its role in novelty detection and episodic memory formation in relation to spatial context.

## 1 Introduction

Emergence of behavior in biological organisms can be understood by simultaneous study of at least two different scales of organization: neural and behavioral. The term emergence may suggest that the intelligent behavior is merely a consequence of unknown and uncontrolled internal dynamics of a system of connected simple elements. In this paper the emergence of behavior is understood as a consequence of assumed neural dynamics, that closely models the functionality of the hippocampal formation. The functionality is assigned to the neural structures as much as supported by behavioral experiments. Yet, this paper aims at a computational model oriented towards a robot implementation.

The choice of the hippocampal formation as the brain structure where the dynamics causing an emergent behavior takes place is made due to several reasons. Among them, the involvement of the hippocampus in episodic memory formation [1][2][3] and its role in familiarity discrimination, are considered, and illustrated through emergent behaviors within a navigation task.

Many biologically inspired robotic systems are built by simulating insect-like behaviors, for a review see [4]. Such a setting supposes a reactive nature of behavior- it is determined mainly by sensing. The simulations following the functionality of the mammalian brain, that include memory or motivational features are seldom implemented into a robot without a severe simplifications. The theoretical models of the hippocampus and basal ganglia, however, suggest neural solutions that incorporate those features, and

produce results on a behavioral scale, often not in the range of the computational expense, affordable for a realistic robotics task [5][6][7].

This paper puts forward memory driven behavior. In particular, memories of experienced episodes are considered as carriers of contextual information, and therefore as a substrate for emergence of complex behaviors: the animat continuously gathers information about the surrounding world through experiencing sequences of events. Higher organisms can encode such subjectively experienced episodes, so their further actions are determined to a big extent of these remembered episodes. Episodic memory formation is generally associated with the encoding in the hippocampus.

While the emergent phenomena are modelled through the hippocampal system, the behavior manifests itself through a motor system. To rightfully describe the organism-environmental interaction the third, sensory system needs to be modelled. The overall model has to satisfy requirements coming from the behavioral setup as well as requirements that concern memory formation and reuse. The specificity of the model is that it attempts navigation, determined mainly by the experienced memories of connected events (episodes). In doing that we expect to observe manifestation of novel behavior that has emerged on the basis of encoding different memories about similar situations.

Therefore, every individual event has to be seen in context. For a realistic task this is possible only if memorizing is flexible in both encoding and retrieval. Flexibility, as opposed to exact storage puts forward the need to selectively store the incoming perceptual information, judging which is new, or very similar to the experienced one. The criteria of familiarity facilitates the ability to infer appropriate behavior in a novel environment or for performing a novel task. The efficient encoding through the familiarity judgement is a basis for life-long learning.

The paper is organized as follows: In Section 2 suggests a hypothesis and the corresponding working model for novelty gated hippocampal encoding; In Section 3 the framework and the scheme from the last two chapters is bridged to the computations, derived by theoretical findings to lay a computational ground for further modelling. The connection between the two levels of the emergency scheme is shown in short in Section 4. The results that show the efficacy of the novelty encoding are not shown in this study. Instead, some results from the functionality of the overall scheme are shown in Section 5. Section 6 offers a discussion of the state and the perspectives of this research.

## **2 The Model**

Approaching the novelty and familiarity discrimination problem from the perspective of an embodied agent has the following meaning: first, novelty has to be considered in relation to behavior; and second, the information, that has to be judged for novel or familiar is derived by the experienced episodes of events. The episode paradigm will be elaborated on first.

Episodic and autobiographical memories have intrinsic relation to sensory and perceptual events (Conway [8]). The term “episodic memory” is used differently by the researchers. In definition, given in [8], which extends the widely accepted definition of Tulving [9], the episodic memory has event specific, sensory-perceptual details of recent experiences that lasted for comparatively short periods of time (minutes to hours).

To make this definitions more transparent for the purposes of our model, lets clarify that perception includes sensing, as well as memory, anticipation behavioral goals, etc. [10][11]. Further in this text sensing and perception will be distinguished, so that sensing implies the external (sensory-bound) patterns, while the perception is the combined representation, formed by the sensory, memory and eventually top-down behavioral or attentional influence.

The encoding of episodes takes place in CA3 area, while the representation, that guides the goal-oriented behavior is formed in CA1 area. The overall processing, naturally includes the contextual temporal information and accounts for novelty, and efficient encoding.

The computation in proposed model accentuates on the representation of CA1 area, as a physical component with comparative function. The relation between the incoming in and outgoing from CA1 area signals is as follows. The information from the learned episode from CA3 and the sensory-bound information coming directly from EC forms the pattern in CA1 that controls the upcoming behavior and signals for novelty.

This functionality has been confirmed by experimental studies to a large extend. Our aim is a model that closely resembles the actual computations in the hippocampus and is applicable for a robotic implementation. Therefore we hypothesize the exact mechanism of novelty detection and context transfer between the two representations. We assume, that the two representations - in CA1 and in CA3 area act together for the novelty detection and future action choice. The CA1 representation determines the action choice and signals for novelty, while CA3 supplies it with contextual information.

The computational scheme, that illustrates the hypothesized functionality is as follows. Two simultaneously active neural networks, corresponding to CA3 and CA1 area perform the major computations. The neurons in CA3 area account for temporal aspect and the formation of episodes, the representation in CA1 area is prone to detect novelty.

The third structure, EC provides the input patterns to both areas. The same patterns, as formed in EC area are projected simultaneously to CA1 area trough the direct pathway, and to CA3 area (trough DG, which is not modelled here) to further reach CA1 area. The projection to CA3 area is mainly topological. In the itinerary of the indirect pathway, in the CA3 area the broad lateral connectivity promotes formation of episodes, where the temporal characteristics of the formed episodes are taken into consideration.

The representation, projected to CA1 area is not assumed topological, i.e. the connectivity between EC and CA1 is full. The pattern that reaches CA1 area via the direct pathway is organized on pattern similarity. Since the same projected pattern from EC area reaches within a small time interval the areas in CA1 and CA3, the connection between the currently most active neurons in this two areas is also strenghtened.

Let the representation, that is formed in CA1 area is denoted by  $E$ . It evolves under the action of the following competing influences: sensory  $S$ , memory  $M$  and behavioral  $B$ .

$$\frac{dE}{dt} = f(S + M + B) - \alpha E \quad (1)$$

where  $f$  denotes a functional dependence, and  $\alpha E$  is a self inhibitory term.

For a robotic framework it is feasible to consider discrete processing. Therefore formally an  $E$  is a set of  $n$  discrete events  $e$ , defined by a considerable difference in the event representations. A single event  $e$ , that is expressed by a specific firing pattern in CA1 area is defined by  $s$ ,  $m$  and  $b$ ,

$$e = \{s, m, b\} \quad (2)$$

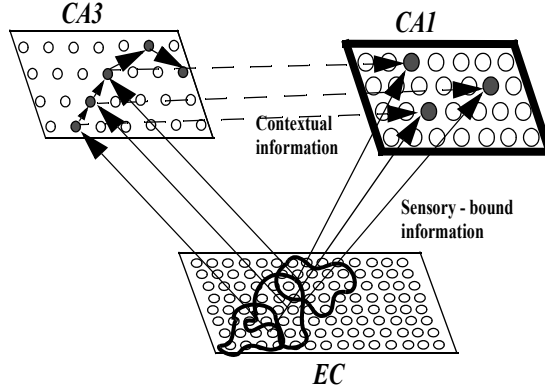
where the sensory component  $s$  introduces the current influence from the external world and constitutes by feedforward connections; memory component  $m$  represents the temporal impact of the sensory patterns, and is constituted by the lateral connections; the behavioral component  $b$  represents the influence, that the previous action has brought on the current even. All three components are multidimensional vectors. The change from one to another event requires a change in at least one component, that is bigger than an internal threshold, indicating the detection of a novel event.

The memory component consists of encoded episodes of events (memory with a temporal reference)  $M$  as formed in CA3 area.  $M$  is a set of  $n$  discrete events occurring in a temporal order  $m_t, (t \in [1, \dots, n])$  defined by a considerable difference in the event representations.

$$M = \{m_t, t \in [1, \dots, n]\} \quad (3)$$

Learning of an episode means that associations can rapidly be formed among items presented in temporal proximity. This is especially valid for events, or routes, where the temporal order is of importance. Therefore by episode learning the order or the temporal association has appeared important rather than or along with another common feature, and this has influenced the encoding. So, the events that have been learned as an episode will tend to be recalled together and after each other, even if presentation order is changed. In addition, the following two dependencies between the temporally related connections are important: contiguity and asymmetry. Contiguity means that stronger associations are formed between stimuli that occur near each other in time than between those that are separated by a larger interval. Asymmetry determines that the forward associations are stronger than backward associations. To summarize, the proposed model, as shown in Fig. 1., is as follows. Three structures, resembling EC, CA1 and CA3 areas form the representation, that is further used for navigation. The computations are performed in the superficial CA1 and CA3 areas. The representation, which is denoted by  $E$  describes the activation pattern in CA1 area. It contains set of elements, which are not necessarily in the order of their temporal appearance. In contrast,  $M$  is a set of temporally related patterns. It refers to the activation pattern of CA3 area. The two representations are physically and computational connected.

The pattern that reaches CA1 area via the direct pathway is organized on pattern similarity, not on topological principle. Since the same projected pattern from EC area reaches within a small time interval the areas in CA1 and CA3, the connection between the currently most active neurons in this two areas is also strengthened.



**Fig. 1.** Schematic representation of the proposed model. The CA3 and CA1 respond differently to the same patterns. The CA3 area provides temporal context to the CA1 representation.

This automatically activates the complete episode to which the pattern in CA3 area belongs, and therefore the contextual information from this episode is transferred to CA1 area.

### 3 Biological Learning

This section outlines the biologically plausible representational and learning mechanisms within and between the different neural structures. Some of the formulas are adopted from the related work (like place cells definition, the specific Hebbian and Temporal Hebbian learning schemes). These are the preliminary elements that will support the choice of the model and justify the simplifications that are made for the robotics implementation.

It is widely known that the cells in the rat hippocampus fire when the rat is at particular location of the environment [12]. Because of that property, these cells are called also place cells. If the rat moves through the environment, at every particular position a number of place cells fire. Cells that code for places in nearest vicinity fire most strongly, while the cells that code for more distant locations fire less. The activity of the place cell firing can be modelled by a Gaussian for the open environments, where place cells show non-directional firing. Therefore, the movement of a simulated rat at every position of the environment is characterized by a particular pattern of firing, containing of the active place cells in its vicinity. The activity of each cell can be represented in the following way:

$$s_{i,t}(r) = -\left(\frac{\|r - c_i\|}{\exp(2\sigma^2)}\right) \quad (4)$$

where  $c_i$  is the location in the space of the center of the cell  $i$ 's place field,  $r$  is the position of the simulated rat, and  $\sigma$  represents the width of the place field.  $s$  corresponds to the sensory representation within a single event.

The sensory as well as behavioral signals are encoded into a constellations of the active place cells. The unique pattern of activity corresponds to a certain position  $r$  in the environment. The level of activity of every place cell depends on the distance between the rat position and the place fields centers. Fig. 2.a) shows two single activation patterns from the rat route, which represent the pattern of activation of the simulated EC area. These patterns are dependant on the external-world and are further transmitted through the direct pathway.

The learning mechanism of the feedforward connections from a EC area, to both CA1 and CA3 area is the modified hebbian rule as shown in Eq. (5).

$$\Delta w_{ij}^{CA-EC} = \alpha_1 g( EC_i CA_j - w_{ij}^{CA-EC} CA_j^2 ) \quad (5)$$

where  $\alpha_1$  is learning rate, notation  $CA-EC$  shows the starting and destination layer of the connection (coming from EC, reaching CA layer) the indices  $i$  and  $j$  denote neurons on the input and output layer, correspondingly. The CA layer is not denoted as CA1 or CA3, because the learning rule is used for EC-CA1 as well as EC-CA3 learning. The term  $w_{ij}^{CA-EC} CA_j^2$  of Eq. (5) is needed due to internal instability of the Hebbian rule. The difference in the feedforward learning in EC-CA1 and EC-CA3 stems from the different connectivity. In EC-CA3 area the predominant are the topological connections - the simulation is done as the neurons from the first (EC) layer project to a topologically correspondent area, in a way that every input neuron is connected to 20% of the output neurons. Differently, the learning between EC and CA1 area is done on the self-organizing principle, since the connectivity between those two layers is full. The lateral inhibition connections, denoted as LI have a sharpening effect on the transmitted to the CA areas activations.

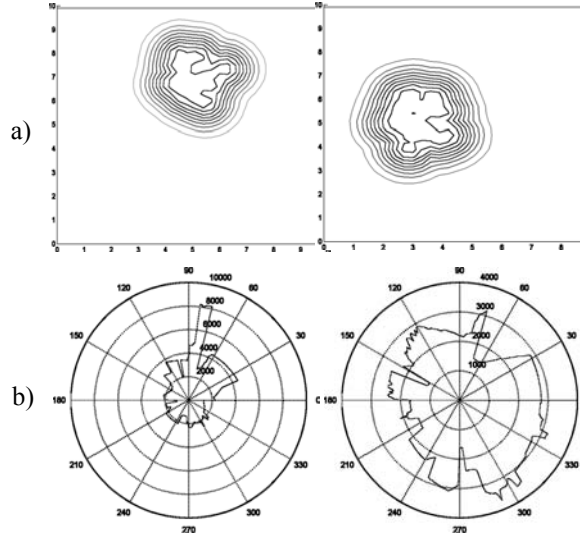
$$\Delta w_j^{LI-CA} = \alpha_2 g( LI \cdot CA_j - \alpha_3 w_j^{LI-CA} CA_j^2 ) \quad (6)$$

where  $\alpha_2, \alpha_3$  are learning rates,  $g$  is a gating factor,

The biological and computational plausibility of the learning process as described in Eq. (5),(6) have been shown previously in [13] where the learning parameters choices is explained in detail.

By far, the representation made within the layer denoted as CA3 has not the intrinsic capacity for temporal association. This quality is obtained by applying a hebbian rule with asymmetric time window over the lateral connections only. The asymmetric time window has been simulated to correspond to the experimental measurements as found by [14], see also [15]. Due to computational difficulties it has been scaled symmetrically in the interval [-1 1]. The lateral learning rule is adapted from [16], so that it overcomes the computational constrains of the asymmetric time window function - Eq.(7).

$$\Delta w_t = \sum_{\tau=0}^{bound} H(\tau)v(t)u(t-\tau) + H(-\tau)v(t-\tau)u(t) \quad (7)$$



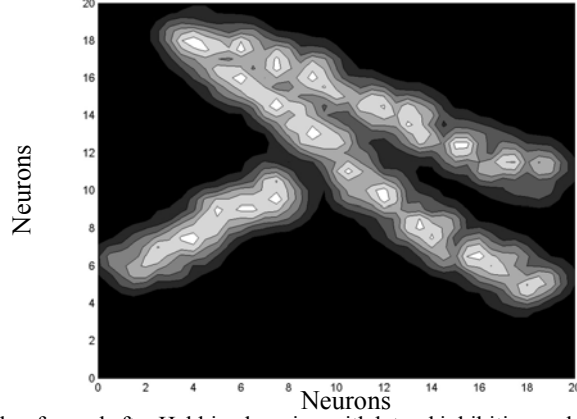
**Fig. 2.** Samples of sensory patterns provided by a) the simulation of the place cells formation process, and b) the omnidirectional range sensor of a robot. The patterns in Fig. 2.b) will be discussed in the following section.

where  $H$  stays for the asymmetrical time window function. Due to the so introduced temporal aspect, the learning in CA3 results in episode learning. A plot of a sample learned episode performed by the set of equations (5-7) is shown in Fig. 3.

#### 4 Connection between the Neural and Behavioral Representations

The actor-critic model [18] most closely resembles the behavioral part of the organism-environmental interaction, in the following sense: At any moment  $t$ , the embodied animat is able to choose an action on the environment, as well as environment provides it with a feedback. In the case considered here the simulated animat can choose from 8 possible actions - or more exactly directions of movement.

By a physical robot they are to be restricted to 3. The critic is the feedback influence, that reaches CA1 area in this model. The actor-critic mechanism regards the actions of the animat that are directed in finding a particular goal. The familiarity gating is entirely performed in the feed-forward track. The goal oriented behavior will be used only for illustration of the animat behavior, and not to optimize its performance. The mechanism of actor-critic optimization, based on place cells representation is suggested in [17]. However, the two studies have completely different objective - in [17] the reinforcement algorithm causes improved behavior, while in our case the learned episodes are a substrate of the emergent behavior, i.e. the first paper models dopamine like effects, while we work on episodic memory based behavior.



**Fig. 3.** Episodes, formed after Hebbian learning with lateral inhibition and novelty gating. Every input pattern activates multiple neurons. The actual path corresponds to the neurons, activated at highest.

The actor-critic mechanism used for the simulations is modified from the original [18][17] in a way, that it accounts for a more realistic movement of both rat and robot. To be consistent with the text above and with the previous notation, the critic is denoted with  $b$ . The value of  $b$  when the animal is at a position  $r$  as represented in CA1 area, has the following form:

$$b(r^{CA1}) = \sum_i w_i f_i(r^{CA1}) \quad (8)$$

where  $w_i$  is the weight between the output cell and the  $i$ -th place cell as formed in CA1 area. At the made experiments, the actor makes use of 8 action cells  $a_j, j = 1 \dots 8$ . At position  $r$ , the activity of the each action cell is:

$$a_j(r) = \sum_i v_{ij} f_i(r^{CA1}) \quad (9)$$

where  $a_j$  stays for the  $j$ -th action cell, and  $v$  is the adaptive weight between the action cell and the  $i$ -th place cell. The first step for the movement direction is taken randomly with a probability  $P_j$ . However, the next movement direction is chosen in a random way, but the possibilities are restricted according to the choices made in the previous movements  $P_j(t-1), P_j(t-2)$ , so there is not a random walk like trajectory, but a smoother orbits with eventual sudden turns.

The actor weights are adapted according to:

$$\Delta v_{ij} \propto \delta_t f_i(r_t) g_j(t) \quad (10)$$

where  $g_j(t) = 1$  if action  $j$  was chosen at the moment  $t$ ,  $g_j(t) = 0$  otherwise.



## 5 Emergent Behaviors and Robotic Implications

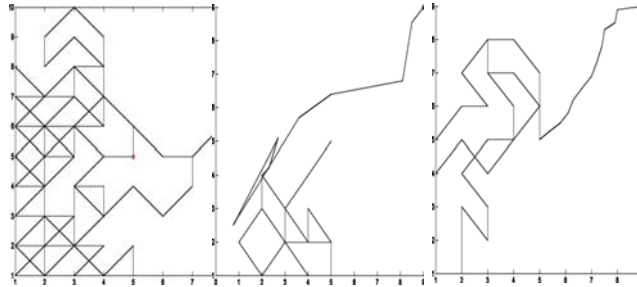
Emergent properties arise at a particular level of system description by interaction of relatively simple lower level components - between themselves and with the environment. In this paper as the lower level of organization is considered the functional simulation of the hippocampal formation, and as the higher level the resulting behavior.

At the lower level, the sensory bound representation of the direct pathway, and the episodic representation of the indirect pathway, come together in CA1 area, where the comparison (novelty/familiarity judgements) takes place. Note, that at the same time, the CA1 area gets input from the current pattern of EC area and a pattern from CA3 area, that does not have incorporated the pattern that EC area currently projects.

In Figure 3b) are shown the omnidirectional views taken by range sensors, when a robot is at a specific position of the environment. The similarity between the nature and the shapes of both types of signals, the one obtained by distance dependant superposition of place cell firing with respect to the rat positioning (Fig. 2.a, the inner contour) and the range sensor readings of a robot with respect to the robots current position, Fig. 2.b), is the basis of our motivation to use the patterns as illustrated in Fig. 2.b) for a robot simulation. For the novelty detection, the network build on the basis of the proposed computational scheme is simplified. The hebbian learning followed by a lateral inhibition is replaced by a modification of the competitive hebbian learning algorithm [19][20], which makes a single neuron to represent an input pattern. It is necessary to organize the map which has fine representation for frequently visited places but coarse representation for others. This problem is elegantly solved by the proposed algorithm. The connections between events within an episode are formed as an internal property of the learning process, that is a principle difference from existing models with temporal history encoding [21]. An on-line comparison of the incoming sensory-bound patterns with those encoded into episodes is made. Once a recognition of the similarity between an experienced episode and the current sequence of events is encountered, the behavior is entirely guided by the encoded memories. When memory of a similar situation is not available, the behavior is random goal-searching. In Fig. 4. are illustrated two groups of behaviors: merely random, when no related memory episode have been encountered, and two memory based behaviors, that are initiated when the similarity between the upcoming environmental sequence appears familiar with already experienced episodes of memories.

## 6 Discussion

This paper features episodic memory encoding as a substrate for emergence of behaviors and goal oriented navigation, performed solely on the recollected episodes. The episode formation and behavioral choices are based on novelty/familiarity discrimination. The impact of novelty is two-fold: it allows an efficient encoding (exploration) phase and it is a basis for flexible reuse of memories in the recall (exploitation) phase. The same computational paradigm is used in both cases, which makes an on-line implementation feasible.



**Fig. 4.** Behavior of the simulated rat. a) random goal-searching behavior; b) and c) familiarity driven behaviors, emerging after a period of exploration.

Mimicking an animal by a robot simulation has its roots in the many parallels, that can be found between them. The most obvious parallels stem from the embodied nature of an animal and a robot: both have a physical body, gather external information and behave accordingly. Going beyond this coarse comparison, we search for a simulation, that aims exactness at the level of internal representation, the hippocampal representation in particular. The functional efficiency of the hippocampal encoding, while performing both tasks - episodic encoding and novelty detection, suggests an optimal computational scheme, which may be beneficial for a robotics task. A detailed hippocampal simulation on its turn is a substrate for realistic emergent behaviors and can benefit the experimental sciences.

A new concept for hippocampal model underlying behavior simulation that has been presented incorporates the novelty/familiarity gated learning and builds a functional connection between novelty detection and episode formation. A different approach has been taken to the sensory input. The functional and structural resemblance between sensory patterns of a robotic range sensors and place cells representation was observed, so the robotics simulation is based on so derived patterns.

## 7 References

- [1] Bunsey, M. and Eichenbaum, H. B. (1996). Conservation of hippocampal memory function in rats and humans. *Nature (London)*, 379 (6562), 255-257.
- [2] Eichenbaum, H., Dudchenko, P., Wood, E., Shapiro, M., and Tanila, H. "The Hippocampus, Memory, and Place Cells: Is It Spatial Memory or a Memory Space?" *Neuron*. Cell Press, 1999, Vol. 23, pp. 209-226.
- [3] Stark, C.E. and L.R. Squire, "Simple and associative recognition memory in the hippocampal region". *Learn. Mem.* **8**: 190-197
- [4] Franz, M.O. and H.A. Mallot, Biomimetic robot navigation. *Robotics and Autonomous Systems* **30**, 133-153 2000.
- [5] Arleo, A. & Gerstner, W. (2000). Spatial cognition and neuro-mimetic navigation: A model of hippocampal place cell activity. *Biological Cybernetics*, 83, 287-299.
- [6] Burgess, N., Recce, M., & O'Keefe, J. (1994). A model of hippocampal function. *Neural Networks*, 7, 1065-1081

- [7] Touretzky, D. S. & Redish, A. D. (1996). Theory of rodent navigation based on interacting representations of space. *Hippocampus*, 6, 247-270
- [8] Conway, M. A. 2001 "Sensory-perceptual episodic memory and its context: autobiographical memory." *Phil. Trans. R. Soc. Lond. B* **356**, 1375-1384.
- [9] Tulving, E. & Donaldson, W., *Organization of memory*. New York: Academic Press, 1972.
- [10] Barakova, E. and Lourens, T., "Prediction of rapidly changing environmental dynamics for real time behavior adaptation ." *Dynamic Perception*, pp 147-152, 2002. IOS press.
- [11] R. Pfeifer, and C. Scheier, "*Understanding Intelligence*" The MIT Press, 1999.
- [12] O'Keefe, J. and Nadel, L., 1978, *The Hippocampus as a Cognitive Map*, Oxford: Clarendon Press
- [13] Schweighofer N., Doya K. and Lay F., "Unsupervised learning of granule cell sparse codes enhances cerebellar adaptive control" *Neuroscience* 103:35-50
- [14] Zhang, L., Tao, H., Holt, C., Harris, W. and Poo, M., A critical window for cooperation and competition among developing retinotectal synapses. *Nature*, 395:37-44, 1998.
- [15] Markram, H., Lubke, H., Frotscher, M., and Sakmann, B., Regulation of synaptic efficacy by coincidence of postsynaptic aps and epsps. *Science*, 275:213-215, 1997.
- [16] Dayan, P. and Abbott, L.F. (2001) *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*, MIT Press, Cambridge MA.
- [17] Foster DJ, Morris RGM, Dayan P., A model of hippocampally dependent navigation, using the temporal difference learning rule. *Hippocampus* 10:1-16. 2000.
- [18] Sutton, R.S. and Barto, A.G., *Reinforcement Learning: An Introduction* MIT Press, Cambridge, MA, 1998.
- [19] Fritzke B., "Growing cell structures- a self organizing network for unsupervised and supervised learning", *Neural Networks*, vol.7 pp. 1441-1460, 1995.
- [20] Martinetz T., and K. Schulten, Topology representing networks, *Neural Networks*, v.7 n.3, p.507-522, 1994
- [21] Barakova, E. and Zimmer, U., "Dynamic Situation and Trajectory Discrimination by Means of Clustering and Summation of Raw Range Data", AISTA, Australia, 2000.