

A Computational Model of Neocortical-Hippocampal Cooperation and its Application to Self-Localization

Michail Maniadakis and Panos Trahanias

Institute of Computer Science
Foundation for Research and Technology-Hellas (FORTH)
71110 Heraklion, Crete, Greece

and
Department of Computer Science
University of Crete
71409 Heraklion, Crete, Greece
{`mmaniada,trahanias`}@ics.forth.gr

Abstract. Recently many computational modules of hippocampal system have been proposed, investigating mainly the development of place cells, similar to mammals. In most approaches, however, place cells are not employed by other structures for further use. We propose a biologically plausible computational model of neocortical-hippocampal cooperation, which is based on familiarity recognition by neocortex, followed by a recall process in the hippocampus. Our model is implemented and tested in a simulated robotic platform, which shows that neocortex is able to interact with hippocampus for the development of a self-localization behaviour.

1 Introduction

The hippocampus is one of the most studied areas of the mammalian cortex because of its prominent role in the memorization of spatial information. Different groups of cells have been detected in the rat's hippocampus, which preferably fire when the rat is in a particular portion of its environment, but they are largely independent of its orientation and actual view [1]. These cells are usually termed *place cells*. Similar cell groups have also been detected in the hippocampus of other mammals.

The cortical structures included in hippocampal system are organized in two major components, namely parahippocampal region and hippocampal formation. Following recent trends in the area, we focus our study in the investigation of the entorinal cortex (EC) from parahippocampal region and dentate gyrus (DG) and Amon's horn structures CA3, CA1 from hippocampal formation. Lately, place cells have been detected in all these structures.

A number of hippocampal computational models have been proposed in the literature, which are able to develop place cells based on allothetic sensory stimuli. Early approaches to hippocampal computational models consist of an arrangement of appropriately connected neurons on a planar map. For example

in [2] allothetic visual information is used to perform quantization of the environment, while idiothetic motor information is integrated in a feed-forward neural model to connect locations. A similar idea is used by [3] but with a much more sparse representation of the environment. However, according to [4], [5], the existence of a topographical relation between environmental location and hippocampal cells seems not valid.

Since the anatomical structure of CA3 consists of a high number of synapses within pyramidal cells, it is usually assumed to perform relational computations. This is taken into consideration by recent hippocampal models which use CA3 structures with recurrently connected neurons. For example in [6] a neural model with recurrent connections similar to CA3 is proposed, which operates in two modes (learning-recall) and is able to implement place cells, encoding different environments. Still, this model ignores to a certain extent the functional contributions of other areas. A combination of planar map with recurrent connections which use attractor dynamics is presented in [7].

The projection from CA1 to EC is usually omitted in many proposed models. This is a very critical design decision, since a recurrent cellular structure is computationally represented by a feed forward one. A computational model with re-entrant projections from CA1 to EC is presented in [12], but it is not tested for the development of place cells.

Certain hippocampal models are oriented towards the development of navigation abilities [8], [2], as suggested by the implied assumption that hippocampal results must become useful for the development of various behaviours. These models assume that information about goal location is given in the hippocampus, even though there is no experimental evidence of its existence [9]. Our approach complies with recent findings which support the belief that hippocampal system is not directly involved in navigation, since it has been experimentally proved that rats with hippocampal lesions are able to navigate to visible goals [10],[11]. Thus, in the present study we don't investigate navigational abilities of hippocampus. In contrast, we focus on the cooperation of neocortical and hippocampal structures to infer self-location. Additionally, this information can be further used by the neocortical structures (prefrontal, motor cortex) responsible for action planning and execution.

A similar work investigating neocortical-hippocampal cooperation [12] uses separate input and output EC structures to mediate interaction. However, it has only been tested in the memorization of static relational information and not in the memorization of continuous spatial information and the development of place cells. The interaction of hippocampus with neocortex is also investigated in [9], mainly focusing on the egocentric-allothetic transformation of sensory stimuli.

In summary, existing hippocampal models exhibit shortcomings, with more important ones being the following:

- They assume the afferent input to hippocampal system from cortical association areas without implementing a computational model of neocortex.
- They represent EC with an input layer, which makes impossible the development of place cells in EC.
- They omit the re-entrant projection from CA1 to EC.
- They don't investigate cooperation of hippocampal and neocortical areas.

- They assume global view of the environment, which largely simplifies the self-localization problem.

In this work, we present a detailed hippocampal model with separate, biologically plausible, computational modules for each hippocampal area (EC, DG, CA3, CA1) and we test its interaction with neocortex. Similar to other models we use a large number of recurrent connections in CA3 to perform relational computations. Instead of representing EC by a poor input layer of sensory allothetic information, we provide a model of neocortical sensory association area (AC) with afferent and efferent EC connectivity. Our model is able to develop place cells in all hippocampal areas, similar to mammalian hippocampus. This is done by employing familiar features detected in AC, to recall the exact location in hippocampus. Our approach is based on the assumption [12] that neocortex supports familiarity recognition, while hippocampus supports conflict resolution and recall. In contrast to [12], we don't use a separate output area to accept hippocampal results, but the same AC structure, since it is bidirectionally connected to hippocampus. Even though our approach is currently tested in a spatial problem, we believe that it is also able to encode individual events within episodes, whether spatial or not [5], by encoding the relation within event features. The existence of a detailed computational model for AC, the single model for EC, together with the recurrent connectivity within AC-EC and EC-CA1 modules, constitute the main contributions of our approach.

In the following section we present the details of our approach for the hippocampal processing of sensory stimuli and the cooperation with neocortex. The results from the application of our model in a spatial learning task of a simulated robot are presented in section 3. Finally, conclusions and suggestions for further work are drawn in the last section.

2 Methodology

Our approach is in accordance to the flow of information in the mammalian central nervous system. The general layout is shown in Fig 1. Environmental information reaches the somatotopically organized sensory cortex which is located in the medial parietal area. There are separate areas to receive signals for different senses. Posterior Parietal Cortex receives afferent projection from sensory cortices to perform high order processing. This part of the cortex undertakes egocentric to allocentric transformation of sensory information, and also relates different stimuli in the Association Cortex (AC). Experimental evidence for ego-allo transformation in Posterior Parietal Cortex is given in [9]. AC projects to EC which is the gate of hippocampus. We have used the three standard areas to represent hippocampal formation, which all receive afferent connections from EC. DG examines current information to emphasize novel features and projects to CA3 which perform temporal and spatial relational processing. The results of CA3 together with the detailed input from EC are passed to CA1 to perform fine tuning. The integration of the results with new sensory information takes place in EC which receives afferent projection from CA1. Efferent projections from EC to AC are used to store the statistical regularities of environmental stimuli in neocortex for further use (e.g. action planning). In the following, we give a

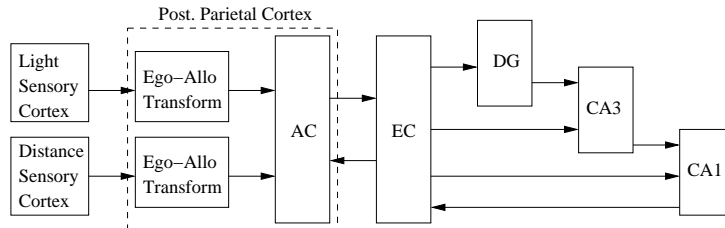


Fig. 1. The flow of information in the proposed neocortical-hippocampal model

simple, yet effective, computational model for egocentric - allocentric transformation of sensory stimuli, and the description of the computational module used to represent all cortical areas.

2.1 Egocentric-Allocentric Transform of Sensory Information

Animals receive egocentric information from their sensors which is modulated by their orientation in the environment. However, it has been experimentally shown that hippocampal system process allocentric (orientation invariant) information. In order to perform the egocentric to allocentric transformation, current orientation is needed. Cells encoding head direction have been observed in presubiculum and anterior thalamus [13]. Presubiculum projects to posterior parietal area, which is assumed to perform egocentric to allocentric mapping. Thus, allocentric information is fed in the hippocampus. This is a common hypothesis for all hippocampal computational models.

The existing computational models assume a hard-wired transformation from egocentric to allocentric information. Instead, we have implemented a simple computational module to perform this transformation, given the current orientation ϕ of the animal. For the sake of simplicity we assume that the number of head-direction (HD) neurons is equal to the number of light or distance sensors; let this number be M (it is also possible to formulate it in the case where the mentioned numbers are different; this formulation is omitted here due to space limitations). Each HD neuron has a preferred direction θ of maximal activation and follows the gaussian model (this is similar to the activation model of real HD cells [13]). Let us assume that the information of the i -th egocentric distance sensor is given by h_i . The allocentric distance measure is achieved by the following summation over all HD neurons

$$f_i = \frac{\sum_{j=0 \dots M-1} h_{(i+j)\%M} e^{-(\phi - \theta_{(M-j)\%M})^2}}{\sum_{j=0 \dots M-1} e^{-(\phi - \theta_{(M-j)\%M})^2}} \quad (1)$$

where f_i is the new orientation invariant measure. This formula has a slight smoothing effect in sensory stimuli, which is due to the averaging performed. Intuitively, it considers stimuli from all sensors, rotated by certain angles, and weighted each time by a factor that is proportional to the matching of rotation and head direction. It is interesting to observe that this formula can be directly used to combine our approach with other computational models that develop HD cells (e.g. [14]).

2.2 The Cortical Computational Module

We have implemented a general computational module to represent cortical areas. This model bears similarities with other works in the definition of synapses [14], and the computational model of the neurons [6]. In contrast to the majority of hippocampal models, a separate module is employed to represent each area of hippocampus and neocortex.

Synapse Definition Each cortical computational module consists of a population of excitatory and inhibitory neurons. A rectangular plane with both sets of neurons uniformly distributed simulates the cortical area. Thus, an excitatory and inhibitory grid is defined on the cortical plane with each neuron occupying a predefined position. In order to achieve common spatial properties for neurons in the middle and neurons in the borders of the plane, we assume that opposite planar sides are met and the neurons near by can be connected. Interconnectivity of neurons follows the general rule of locality [14] for short and long range interactions. Short-range interaction concerns the intracortical connectivity while long-range the cortico-cortical connectivity.

The local intracortical connectivity is expressed by the rule that close neurons are more likely to be connected. It is simulated with a stochastic gaussian linking. If d is the distance of neurons a, b then $t = e^{-\frac{d^2}{\sigma}}$ defines the probability of synaptic connection for those neurons. Variance σ determines the sharpness of synaptic connection. Using a random number $r \in [0, 1]$ a synapse from a to b is defined if $r < t$. The synapse is assigned with a random weight $w_{ab} \in [0, 1]$ to represent its strength. The same process is repeated for all pairs of neurons (excitatory to excitatory, excitatory to inhibitory, inhibitory to excitatory) to define the local connectivity within a cortical module. We assume that inhibitory neurons are used to enforce separability within local planar areas, thus inhibitory-inhibitory synapses are not used, and only three different sets of synapses are defined W^{EE} , W^{EI} , W^{IE} . In that way, bi-directionally neural pairs can be defined, with the flexibility of different weight in each direction.

Long range cortical connectivity follows the general rule that neighbouring cells project to neighbouring areas. This is simulated in our model by a gaussian stochastic projection. Thus, different cortical modules are connected using their spatial properties inherited by the planar model. Interconnectivity of modules concerns linking of the respective excitatory neurons and defines one-way synapses. For each cortical module B, which receives afferent projection from a module A, a set of input neurons Inp is defined, equal to the number of excitatory neurons in module A. They are located in the same positions as their respective excitatory neurons in module A. The connectivity from input neuron a to excitatory neuron b is defined by employing the same gaussian stochastic linking described above, with σ_{InpE} connectivity variance. Thus a set of input synapses V with random synaptic weights is defined in module B.

Neuron Model All neurons of the cortical module, follow a modified version of the Wilson-Cowan neuronal model, similar to [6]. Let p represent the potential, and q the activation of a neuron. The potential of each neuron is updated based

on the afferent input information, and the excitatory and inhibitory signals accepted by neighbouring neurons. This is expressed mathematically, in a single form for both excitatory and inhibitory neurons, by:

$$\frac{1}{\mu}\Delta p_b = -p_b + \sum_{v_{ab} \in V} v_{ab} \text{inp}_a + \sum_{w_{ab} \in W^{EE}} w_{ab} q_a + \sum_{w_{ab} \in W^{EI}} w_{ab} q_a - \sum_{w_{ab} \in W^{IE}} w_{ab} q_a \quad (2)$$

where μ presents the membrane time constant. Then, the activation of the neuron is defined using the non-linear sigmoid function

$$q_b = \frac{1}{1 + e^{-\alpha(p_b - \beta)}} \quad (3)$$

where β stands for the threshold, and α is the slope of the activation function.

Learning Rules A learning process adjusts the initially random strength of synapses to encode the spatial properties of the environment. We have used four biologically plausible Hebbian-like rules to train cortical modules. Principal Component Analysis rule is used to maximize information flow within neurons, while Anti-Hebbian learning gives us the possibility to develop a novelty detection mechanism. We have also used Postsynaptic and Presynaptic rules, which simulate the biological heterosynaptic and homosynaptic learning. The pure Hebbian rule is not used here, to avoid monotonic changes in synaptic efficacy. Thus synaptic weights can be freely adjusted in the range [0,1]. Assuming that there is a synapse with strength z_{ab} from neuron a with activation q_a to neuron b with activation q_b , then learning rules are described below.

- *PCA Rule* [15]: This well known rule has the effect of adjusting the weights to pass the principal component of presynaptic activation to postsynaptic unit. The mathematical expression of the rule is $\Delta z_{ab} = q_b(q_a - q_b z_{ab})$.
- *AntiHebbian Rule* [16]: It adjusts synaptic strength to decorrelate neurons a, b . The mathematical expression of the rule is $\Delta z_{ab} = k + \frac{-2q_a q_b}{q_b^2 + 1}$, where $k > 0$ is a small forgetting factor, to avoid vanishing.
- *PostSynaptic Rule* [17]: Generally it increases the synaptic weight, but it can decrease it when postsynaptic unit is active and the presynaptic unit is not. Its mathematical expression is $\Delta z_{ab} = z_{ab}(q_a - 1.0)q_b + (1.0 - z_{ab})q_a q_b$.
- *PreSynaptic Rule* [17]: Generally it increases the synaptic weight, but it can decrease it when presynaptic unit is active and the postsynaptic unit is not. Its mathematical expression is $\Delta z_{ab} = z_{ab}(q_b - 1.0)q_a + (1.0 - z_{ab})q_a q_b$.

3 Experimental Results

In order to evaluate our neocortical-hippocampal model we have tested it on a simulated Khepera robot. We have used a modified version of the khepera simulator, with uniformly distributed sensors and similar range of view for both light and distance sensors. The goal of our experiments was to implement a self-localization process of the robot in the environment, based on the familiarities detected in the neocortical structure, and the recall of location by hippocampus.

We have performed numerous tests of our model, with a large range of parameter values. In all cases, the results obtained have shown the validity of our model. In the following we present a detailed sample experiment. We specify the actual experimental setup and then report on the obtained results.

3.1 Experimental Setup

Our model (Fig. 1) receives information from 8 distance and 8 light sensors, which are uniformly distributed in a circle around the robot. This somatotopic relation is preserved in the Sensory Cortex, similar to mammals. Each sensor is assigned a position in a circle of the Distance or Light Sensory cortex. Both circles have diameter of 60, and a random centre. Then, environmental information is transformed from egocentric to allocentric measures using the process described in section 2.1. Both of the allocentric Distance and Light Sensory Cortex are projected to AC and then to hippocampus via EC. After hippocampal processing, recalled memory is projected back to AC, again via EC. The variances σ_{InpE} employed for the cortico-cortical projections are shown in Table 1.

We use planes of common size 100×100 distance units for all cortical modules. The number of excitatory and inhibitory neurons (ExcN, InhN) used in each cortical module is shown in Table 2. Similar to actual biological models, the number of excitatory neurons is much larger than the inhibitory. To avoid equality of hippocampal cortical modules, neuron locations are perturbed by a rectangular noise of length 6 in each side. The same table shows the rest of the parameters used for cortical module construction. All modules have a small number of recurrent interconnections within excitatory neurons depending on σ_{EE} . These neural connections simulate synapses within cortical layers. However, in CA3 a large value of σ_{EE} is used to simulate high recurrent connectivity within CA3 pyramidal cells. The connectivity within excitatory and inhibitory neurons is defined by σ_{EI}, σ_{IE} . It is clear that excitatory and inhibitory neurons have a local and global effect, respectively. This model of local excitation (through excitatory neurons) and global inhibition (through inhibitory neurons), imposes competition within different areas of the cortical plane.

The neurons in most modules have relatively high values of potential change μ to capture the fast environmental changes, together with a high slope α for sharp representation. Only CA3 neurons have a large range of variance in potential change to explore temporal relations (some neurons are updated rapidly and some others very slowly). This fact, in conjunction with the large number of synapses in CA3, allows a spatio-temporal relational processing in this module. The thresholds in each module have been set experimentally and do not have a special biological meaning.

Different learning rules were used to adjust the various types of synapses in each module. Similar to the learning functionality of neocortex and hippocampus, we assume that hippocampus is specialized for rapid memorization, while neocortex is specialized for slowly learning about statistical regularities in the environment. Thus, the learning rate for AC was set to 0.01, while hippocampal structures learned with a three times faster rate. Learning rules applied in each module are shown in Table 3. Since PCA rule adjust synaptic weight to obtain large information flow, it is used in the input synapses of most modules. Post-synaptic rule is used in CA3 input to allow EC and DG structures to modulate

Table 1. Variances σ_{InpE} used for the definition of synaptic projections within layers. DSC and LSC stands for Distance and Light Sensory Cortex, respectively

	DSC	LSC	AC	EC	DG	CA3	CA1
AC	15	15	-	10	-	-	-
EC	-	-	10	-	-	-	10
DG	-	-	-	10	-	-	-
CA3	-	-	-	12	10	-	-
CA1	-	-	-	12	-	10	-

Table 2. The parameters used for the construction of cortical modules. Each neuron of a module gets a random value of potential change μ in the defined range

	ExcN	InhN	σ_{EE}	σ_{EI}	σ_{IE}	μ	α	β
AC	64	36	10	20	30	[0.8-0.9]	2.5	1.8
EC	49	25	12	15	20	[0.7-0.8]	3.0	1.2
DG	49	25	12	15	20	[0.7-0.9]	3.0	0.5
CA3	49	25	20	25	30	[0.2-0.9]	2.0	0.8
CA1	49	25	10	15	25	[0.8-0.9]	3.0	0.4

its activity based on their activation. Postsynaptic rule has a disjunctive effect when it is used in EE synapses, while Presynaptic rule has a conjunctive functionality. Antihebbian rule enforces decorrelation within neurons, and it is used here similar to [16] as a novelty detector. The forgetting factor k was set to 0.05. PCA rule is used in EI synapses with a sharp inhibitory effect while Postsynaptic rule with a more smooth one. Similarly, PCA enforces sharp inhibition when it is used in IE synapses, while Presynaptic rule performs smooth inhibition.

Table 3. Each cortical computational module uses a different set of learning rules to adjust synaptic weight.

	Input-Exc.	Exc.-Exc.	Exc.-Inh.	Inh.-Exc.
AC	PCA	PreSyn.	PostSyn.	PreSyn.
EC	PCA	PostSyn.	PCA	PreSyn.
DG	PCA	AntiHeb.	PCA	PCA
CA3	PostSyn.	PostSyn.	PostSyn.	PreSyn.
CA1	PCA	PostSyn.	PCA	PCA

It should be noted that the experimental setup described here is by no means bound to the specific parameter values that are detailed. The later are enumerated for the completeness of the presentation of experimental results. Indeed, we have verified experimentally that these parameters can vary with practically no effect on the performance of the proposed model.

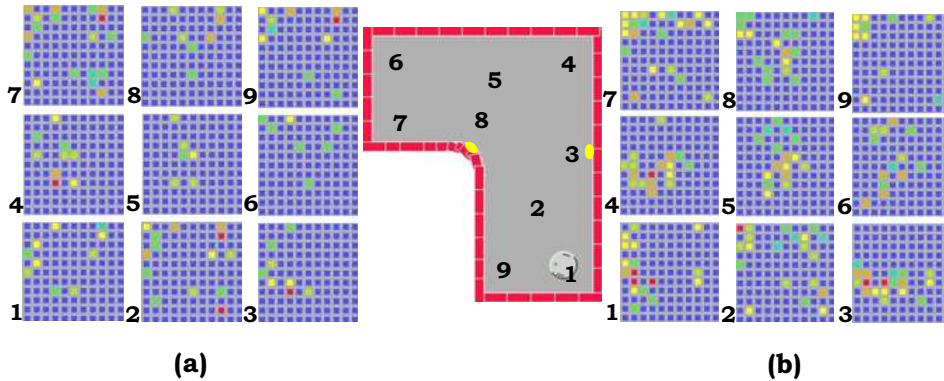


Fig. 2. Results of Self-localization via place-cell development in (a) CA1 and, (b) AC

3.2 Findings

In our experiment, we set the simulated robot to move around in the environment in a wonder mode for 10000 steps. The range of view of the robot was limited to a maximum of two sides at a time (those which are vertically connected). We have used an environmental shape with many angles to enforce the locality of robot view. Also, two light sources were used, to mix stimuli from different sensors. The simulated workspace employed is depicted in the middle of Fig 2; each light source is represented on it by a yellow oval.

The results of place cell development in CA1 for different locations (those marked with the numbers 1 through 9) are shown in Fig 2(a). Evidently, these results indicate the effective development of place cells in CA1 and compare promisingly to other results from the literature [2],[6]. Furthermore, similar neural firing is present in the other areas of hippocampal system (EC,DG,CA3).

The recalled locations are also transferred to AC (Fig 2(b)) and are readily available for further use (e.g. motor control). The activation in this module is generally higher since it is directly connected to environmental stimuli. However, its activation is modulated by the knowledge in hippocampal system. It is clear, that even if the somatotopic relation of sensory stimuli is preserved, learning results do not consist a pure circular projection of environmental information within cortical layers. Learning rules together with inter- and intra-cortical synapses perform the appropriate transformations to store feature relations in synaptic weights. We believe that this novel result, i.e. transfer of recalled location to AC, confirms the validity and suitability of the proposed model. Additionally, in order to test the stability of our model we let the robot run for 50000 steps and we got very similar results.

4 Conclusions

We have presented a model of neocortical-hippocampal cooperation and its application to self-localization. It is based on a general computational module to represent all cortical areas. Our approach complies with the main structural,

functioning and learning properties of the mammalian cortex, constituting it a biologically plausible model. The introduced model has been successfully tested in the development of place cells for localization, similarly to the mammalian hippocampus. Moreover, this knowledge is also projected to neocortex for further use. Since AC in mammals is one of the areas which relate sensory and motor processing, our future work aims at the use of spatial knowledge for action planning in the motor cortex. Furthermore we believe that our model is also able to encode general events within episodes (not only spatial) since it stores the relations between features.

References

1. O'Keefe J.: Place units in the hippocampus of the freely moving rat. *Exp. Neurol.*, 51, pp. 78-109, 1976
2. Arleo A., Gerstner W.: Spatial cognition and neuro-mimetic navigation: A model of hippocampal place cell activity. *Biological Cybernetics*, 83, pp.287-299, 2000.
3. Hafner V.V.: Learning places in Newly Explored Environments. *Proc. SAB2000*.
4. O'Keefe J., Burgess N., Donnett J., Jeffery K., Maguire E.: Place cells, navigational accuracy, and the human hippocampus. *Ph. Tran. R. Soc.* 353, pp. 1333-1340, 1998.
5. Eichenbaum H., Dudchenko P., Wood E., Shapiro M., Tanila H.: The hippocampus, memory and place cells: Is it a spatial memory or a memory space? *Neuron*, vol. 23, pp.209-226, 1999.
6. S. Kali, P. Dayan: The involvement of recurrent connections in area CA3 in establishing the properties of place fields: a model. *J. Neurosc.* 20(19) pp.7463-77 2000.
7. Samsonovich A. McNaughton B.L.: Path integration and Cognitive Mapping in a Continuous Attractor Neural Network Model. *J. Neurosc.*, 17(15), pp.5900-20, 1997.
8. Hasselmo M. E., Hay J., Ilyn M., Gorchetchnikov A.: Neuromodulation, theta rhythm and rat spatial navigation. *Neural Networks*, 15, pp.689-707, 2002.
9. Burgess N., Becker S., King J., O'Keefe J.: Memory for events and their spatial context: models and experiments. *Ph. Trans. R. Soc.*, 356, pp. 1493-1503, 2001.
10. Jarrard L.E.: On the role of hippocampus in learning and memory in the rat. *Behav. Neural Biology*, 60, pp. 9-26, 1993.
11. Morris R.G.M., Garrud P., Rawlins J.N.P., O'Keefe J.: Place navigation impaired in rats with hippocampal lesions. *Nature*, 297, pp. 681-683, 1982.
12. Norman K.A., O'Reilly R.C.: Modelling Hippocampal and Neocortical Contributions to Recognition Memory: A complementary Learning Systems Approach University of Colorado, Boulder, ICS, Technical Report 01-02, 2001.
13. Taube J.S.: Head direction cells and the neuropsychological basis for a sense of direction. *Progress in Neurobiology*, 55, pp.225-256, 1998.
14. Redish A.D., Elga A.N., Touretzky D.S.: A Coupled attractor model of the rodent head direction system. *Comp. in Neural Systems*, 7, pp. 671-685, 1996.
15. Oja E.: A simplified neuron model as a principal component analyzer. *Journal of Mathematical Biology*, 15, pp.267-273, 1982.
16. Schraudolph N.N., Sejnowski T.J.: Competitive Anti-Hebbian Learning of Invariants *Advances in Neural Information Processing Systems*, 4, pp. 1017-1024, 1992.
17. Urzelai J., Floreano D.: Evolutionary Robotics: Coping with Environmental change. *Proc. GECCO2000*.