Transition cells for navigation and planning in an unknown environment

N. Cuperlier, M. Quoy, C. Giovannangeli, P. Gaussier, P. Laroque

ETIS-UMR 8051
Universite de Cergy-Pontoise - ENSEA
6, Avenue du Ponceau
95014 Cergy-Pontoise France
cuperlier@ensea.fr

Abstract. We present a navigation and planning system using vision for extracting non predefined landmarks, a dead-reckoning system generating the integrated movement and a topological map. Localisation and planning remain possible even if the map is partially unknown. An omnidirectional camera gives a panoramic images from which unpredefined landmarks are extracted. The set of landmarks and their azimuths relative to a fixed orientation defines a particular location without any need of an external environment map. Transitions between two locations recognized at time t and t-1 are explicitly coded, and define spatio-temporal transitions. These transitions are the sensory-motor unit chosen to support planning. During exploration, a topological map (our cognitive map) is learned on-line from these transitions without any cartesian coordinates nor occupancy grids. The edges of this map may be modified in order to take into account dynamical changes of the environment. The transitions are linked with the integrated movement used for moving from one place to the others. When planning is required, the activities of transitions coding for the required goal in the cognitive map are enough to bias predicted transitions and to obtain the required movement.

1 Introduction

Several biomimetic models allow to perform navigation tasks even without relying on localisation neither on maps (see [5] for a review of several insects like strategies). Nevertheless these models are constrained to use different "routes" for each goal to reach and can not exhibit some interesting behaviors like shortcut etc... Hence, in most of bio-inspired models, like in [1, 2], localisation is based on particular neurons found in the rat hippocampus (particularly CA3, CA1 and dentate gyrus (DG), regions and also in the entorhinal cortex (EC)) named "place cells" (PC). A map of the environment may be built by linking these PC. One can refer to [3, 4] for a comparative review of localisation and mapping models. In our "rodent like" model, we also use place cells (layer modelling EC see section 4) that learn patterns specific of a given location (spatial landmarks constellation, see section 3), but we do not directly use them to plan or construct

a map. We rather use neurons ("transition cells") that explicitly code for these spatio-temporal transitions (in the layer modelling CA3/CA1). Details of their creation and arguments in favor of such a coding are given in section 5. During exploration, these transition cells are created and allow to learn a cognitive map whose construction is explained in section 6. When a plan is needed, transitions are predicted and are then biased via top-down information from the cognitive map (section 7).

Hence we propose here a unified neuronal framework based on an hippocampal and prefrontal model where vision, place recognition and dead-reckoning are fully integrated (see Fig. 2 for an overview of the architecture). All neurons activity are analogous. There is no symbolic programming nor predefined object of high cognitive level. No assumption are made about the structure of the environment. We will conclude with improvements that may be proposed in our model.

2 Material and methods

The robot is a koala platform (40*30cm) with six wheels. It has infrared sensors for obstacle detection. A low level obstacle avoidance mechanism is implemented (not described here). Images are taken by a panoramic camera at low resolution. A rectangular image (1500×240 pixels) is obtained from the panoramic image which is originally circular (640×480 pixels).
Since our robotic model is inspired from the animat approach [13], we use three contradictory animal like motivations (eating, drinking, and resting). Each one associated with a satisfaction level that decreases over time and increases when the robot is on the proper source according to coupled differential equations [11]. When a level of satisfaction falls bellow a given threshold, the corresponding motivation is triggered so that the robot has to reach a place allowing to satisfy this need. Hence this place becomes the goal to reach. More sources can be added and one can increase the number of sources associated with a given motivation.

3 Autonomous landmark extraction and recognition based on characteristic points

In order to reduce problems induced by luminance variability, we only use the gradient image as input of the system. Next, curvature points (corresponding to robust focal points) are detected by filtering this gradient image with a Difference Of Gaussian. Two processes then occur in parallel: first a log-polar transform of the local area extracted around each focal point is computed. Connection's weights of neurons are then modified to learn these small images. This allows to improve the pattern recognition when small rotations and/or scale variations on these small images occur [6, 7]. These images are landmarks, and by extension, we also name the coding neurons landmarks. Second, for each landmark, an angular position relative to the north given by a compass is computed [8, 9]. Thus, this

visual system provides both a *what* and a *where* information: the recognition of a 32×32 pixels small images in log-polar coordinates, and the azimuth of the corresponding focal point. *What* and *where* informations are then merged in a product space leading to a spatial landmark constellation. The number of
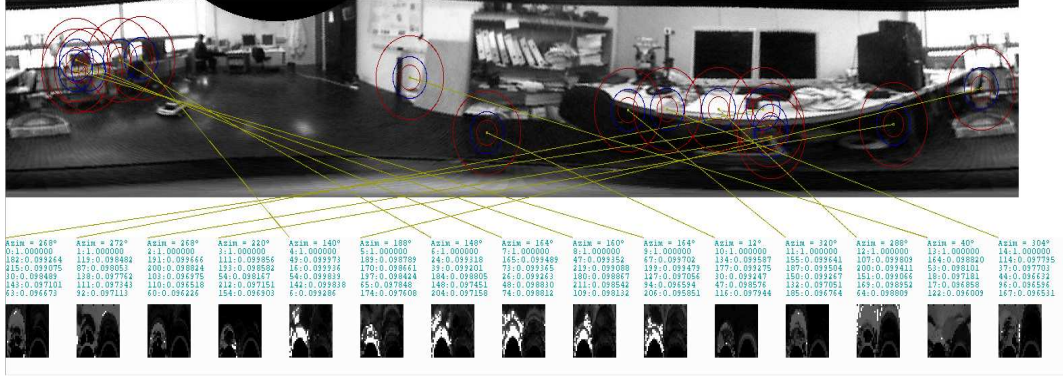

Fig. 1. Image taken from a panoramic camera. Below are 15 examples of 32×32 log-polar transforms taken as landmarks and their corresponding position in the image.

landmarks needed is a tradeoff between the robustness of the algorithm and the speed of the process. If all landmarks were fully recognized, only three of them would be needed. But as some of them may not be recognized in case of changing conditions like luminance or occlusion, taking a greater number is enough to guarantee the robustness.

4 Autonomous place building

The spatial landmarks constellation resulting from the visual input treatement characterizes one location. This constellation can thus be learned on a neuron of EC (place recognition at time t see fig. 2). The neuron coding for this location is called a "place cell" as the one found in the rat's hippocampus [9] since these cells fire when a rat is at a particular location in its environment. The activity of a PC results from the computation of the distance between the learned and the current local view. Thus, the activity of the kth PC can be expressed as follows:

$$P_k = \frac{1}{l_k} \left(\sum_{i=1}^{N_L} \omega_{ik}.f_s(L_i).g_d(\theta_{ik}^L - \theta_i) \right) \qquad (1)$$

with $l_k = \sum_{i=1}^{N_L} \omega_{ik}$ the number of landmarks used for the kth PC, where $\omega_{ik} = \{0, 1\}$ expresses the fact that landmark i has been used to encode PC k, with N_L the number of learned landmarks, L_i the activity of the landmark i, $f_s(x)$ the activation function of the neurons in the landmark recognition group,

θ_{ik}^L the learnt azimuth of the ith landmark for the kth PC, θ_i the azimuth of the current local view interpreted as the landmark i. d is the angular diffusion parameter which defines the shape of the function $g_d(x)$. The purpose of $f_s(x)$ and $g_d(x)$ is to adapt respectively the dynamics of *what* and *where* groups of neurons. They are defined as follow :

$$g_d(x) = \left[1 - \frac{|x|}{d.\pi}\right]^+$$
$$f_s(x) = \frac{1}{1-s}\left[x - s\right]^+$$

where $[x]^+ = x \ if \ x > 0$, $and \ 0 \ otherwise$.

The s parameter rescales the activity of the landmark neuron over s between 0 and 1. The d parameter modulates the weight of the angular displacement.

Experimental place cell formation has also been tested in outdoor environments [10]. The result confirmed the mathematical model which predicts that the size of the place field grows proportionaly with the landmark distance.
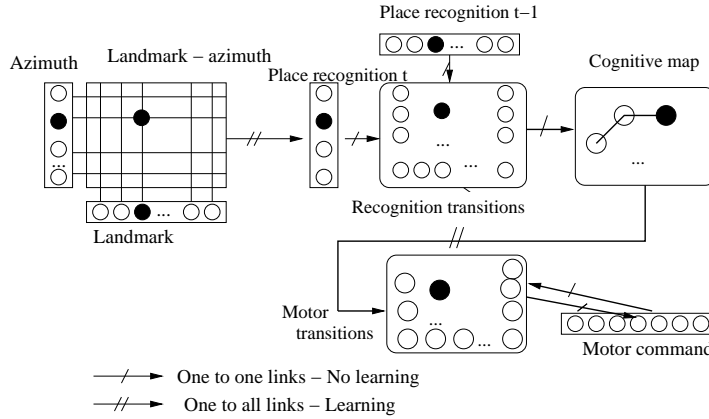

Fig. 2. Sketch of the model. From left to the right: merging landmarks and their azimuth, then learning of the corresponding set on a place cell. Two successive place cells define a transition cell. They are used to build up the cognitive map and are also linked with the integrated movement performed.

If the robot is at the exact position where the PC has learned, its activity is maximal (equal to one). When the robot moves from this position, the activity of this PC decreases. Hence the PC keeps a certain amount of activity around the learned position that is named the *place field* of a PC. Consequently, we have to use a rule that controls the recruitment of a new neuron to encode a new location. This mechanism is performed autonomously, without any external signal, relying only on the PC's population activity. If the activities of all previously learned place cells are below a *given* recognition threshold (R.T), then a new neuron is recruited. At a given place, every existing place cell responds with an

analog recognition value that may be seen as the robot position probability. If at a given place several PC respond with activities greater than the R.T, a competition takes place so that the most activated one wins and codes the current location. The density of locations learned depends on the level of this threshold, but also on the robot position in the environment. Namely, more locations are learned near walls or doors due to the fast changes in the angular position that can occur near landmarks, or in the (dis)appearance of landmarks caused by these obstacles. In other locations, small changes produce a small variation in the place cell activity. When the environment has been entirely explored, and thus fully covered by place cells, a PC responds specifically for each location (see Fig. 6). Consequently the PC neural layer gives our robot a way to localize itself inside the environment it has explored.

5   Autonomous building of transition cells

A natural question is "why using transitions instead of places"? In order to briefly answer this question, we have to focus first on how to plan using place cells. Several bio-inspired approaches rely on place cells, but to better illustrate our approach we will only describe briefly our past-model which allows to easily underline the problem. First a place cell may be linked with the movement needed to reach a goal without any map. This sensory-motor association may be generalized to the whole environment [11]. However, this simple reactive mechanism is not enough in environments composed of several rooms, or when there are contradictory motivations. A cognitive map will solve these drawbacks (see section 6). Two different approaches of this cognitive map exploitation have been proposed. First, the selection of the action in a place cell based model can be realized by an external mechanism applied to the cognitive map: the gradient algorithm. But, if this solution is enough for a navigation task, it might be more difficult to find an external mechanism for more complex tasks like robot arm control. Moreover from a biological point of view, using an external algorithm "looking for" the gradient of activity leads to the famous problem of the homonculus: "who is looking ?" Second, as a consequence, the action selection mechanism has to be integrated. This can be performed by associating an action with a place, thus defining a sensori-motor unit. But then, the choice of the direction to follow may be ambiguous. Indeed, in some place, several actions can be associated with the same place (see fig. 3) like in the T-maze example. In this case, which movement should select the robot if it must go to C?

In order to solve this problem, we do not directly use PC for planning in our model, but rather transitions between the two PC winning the recognition competition: respectively at time t (in EC) and time $t-1$ (in DG). Such spatio-temporal transitions are explicitly coded on neurons called *transition cells*. The idea of this coding has been inspired by a neurobiological model of timing and temporal sequences learning in the hippocampus [12]. Motivation from such a
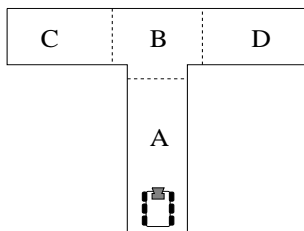
Fig. 3. In this example, from place B the robot had learned during exploration that it can go either to C by turning right or D by turning left. Both movements are thus linked with place B.

coding comes from the fact that transitions are better suited for sensori-motor association than places since only one direction can be linked with a transition: the movement used to go from A to B with the $t$ransition cell AB (see fig. 2).

Before going further about transitions, note that as transitions link two successively recognized PC, transitions like AA are also coded. These kind of transitions are the equivalent of PC in transition coding. No movements are linked with these transition cells. We only associate a movement to a transition linking two different PC. An internal signal is computed from the automatic detection of a new wining PC at time t by temporal differences on EC. This signal is used to trigger the sensori-motor association.

A relevant question is about the growth of the number of transition cells created while exploring the environment. This number is intimately linked with the number of place cells. This number of place cells created for a fixed R.T value depends on the complexity of the environment. The degree of complexity of an environment relies mainly on two factors: the number and the location of its landmarks and the number of obstacles.

Thus we have performed several tests setting one of these parameters to underline the impact of the second upon the ratio between created transition cells over created place cells. Each simulation lasts 50000 cycles. This number has been chosen high enough to ensure that the robot has learned a complete cognitive map of the environment [1]. The results of both tests shown here are the average on ten simulation results. We have first studied the impact of obstacles configuration in three environments of increasing complexity. Tests have been performed for a single, a two and a four rooms environment. The number of landmarks have been fixed at a high value. The ratio remains stable around the mean value 5.45 for all environments once the cognitive map of the environment is complete (see table 1). The second study shows how this ratio evolves for an environment with the same complexity of obstacles but with simple to half landmark number. These tests have been done on the two and four rooms environment with the

---

[1] We consider the cognitive map is complete when the robot becomes unable to detect new places or new transitions.

same experimental set-up than the previous study. The number of landmarks increases the number of particular cases in which a landmark previously visible becomes invisible (or the reverse) and consequently decreases the activity of all previously known place cells. This finally results in the creation of a higher number of place cells. The results (see Table 2) show a stable ratio of a mean value still around 5.35 for simple environments. This ratio does not depend on the value of R.T (but the number of place cells increases with increasing R.T). The stability of this ratio can be explained as follows: since the number of a place cell's neighbours is necessary limited and that a transition is a link between "*adjacent*" place cells, only a few transitions can be created from a given place cell. To conclude, there is no combinatorial explosion of the number of created transitions. Thus, they can be memorized and used for planning purpose.

Param / Env	One room	Two rooms	Four rooms
nbPC	133.8(2.85)	606.2(6.89)	643.7(9,88)
nbT	735.8(19.80)	3389.2(56.38)	3281,2(48,80)
ratio	5.49(0.06)	5.59(0.08)	5.09(0,04)

Table 1. Results of the experiments on the ratio of the number of place cells (nbPC) created over the number of transitions created (nbT) according to the number of rooms in the environment. Standard deviation is given into brackets. This ratio remains stable. There are at most six times more transition cells than place cells. R.T is set at 0.97.

Param / Env	Two, many land.	Two, few land.	Four, many land.	Four, few land.
nbPC	606.2(6.89)	364.3(5.75)	643.7(9,88)	295.5(4.94)
nbT	3389.2(56.38)	1951.2(35.30)	3281,2(48,80)	1591.8(26.64)
ratio	5.59(0.08)	5.35(0.03)	5.09(0,04)	5.38(0,05)

Table 2. Results of the experiments on the ratio of the number of place cells (nbPC) created over the number of transitions created (nbT) according to the number and configuration of landmarks in the environment: with two rooms (first column for many landmarks and second column for few landmarks) and with four rooms (third column for many landmarks and fourth column for few landmarks). Standard deviation is given into brackets. This ratio remains stable. There are at most six times more transition cells than place cells. R.T is set to 0.97.

Now that we know the number of possible transitions starting from a given place cell, we can use this information for modelling the transition layer. Transition cells building does not rely on a full "matrix" coding the relationships between successively reached places. This would be too memory consuming. Instead, we exploit the fact that a place cell has around 5 neighbours on average to

compress the structure merging these informations (see fig. 4). In order to cope with extreme cases, we allow for a maximal number of 10 neighbours. Consequently the number of neurons of this structure has decreased since we only take into account real possible transitions and not all the combination of place cells. Each neuron of a given line receives projections both from population coding for place cell at time t named PC_t and from population coding place cell at time $t-1$ named PC_{t-1}. Each transition neuron belongs to a particular neighbourhood supervised by a single PC_t neuron (a line in the figure 4). No learning is allowed on those links and their weights are not sufficient to trigger any activity on the associated transition neurons. Conversely, each transition neuron is connected to all the PC_{t-1} neurons through conditional links. The activation of PC_t neurons increases the weights coming from the activated neuron in PC_{t-1}, when no transition neuron already corresponds to this conjunction. Once those weigths are learned, in a prediction mode, the single activity of the corresponding PC_{t-1} neuron allows the activity of the transition neuron even if no signal comes from PC_t.
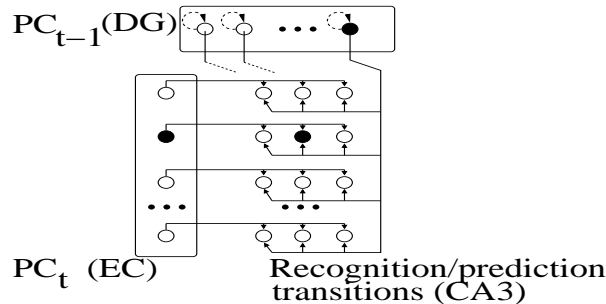

Fig. 4. Transition cells population inputs from population of place cells at time t and at time $t-1$. In order to have a clear figure, only 3 possible transitions are shown on the merging and compression group. For the same reason, connections from only one neuron of PC_{t-1} are drawn.

6   Autonomous cognitive map building

Experiments carried out on rats have led to the definition of cognitive maps used for path planning [14]. Most of cognitive maps models are based on graphs showing how to go from one place to an other [15–21]. They mainly differ in the way they use the map in order to find the shortest path, in the way they react to dynamical environment changes, and in the way they achieve contradictory goal satisfactions. Other works use ruled-based algorithms, a classical functional approach, that can exhibit the desired behaviors, we will not discuss them in this paper, but one can refer to [22].

In our model, learning the cognitive map is performed continuously during the exploration of the unknown environment (latent learning) by linking transition cells successively reached if no link was yet created between these two transitions. Equation 2 shows the learning rule applied to the value of edge $W_{i,j}$ linking vertice j to i. $G(j)$ is the activity of transition j. $\overline{G(i)}$ is the memory term of G(i) that decreases with time. λ is a decay term that allows to forget erroneous transition due to an uncomplete exploration. $\frac{dR}{dt}$ is the variation of the reinforcement. The edge value is increased if the edge is used, and decreased if it is not. After some time, some edges are reinforced. These edges correspond to paths that are often used. In particular, this is the case when some particular locations have to be reached more often than others (see section 7) [11].

$$\frac{dW_{i,j}}{dt} = -\lambda.W_{i,j} + (1 + \frac{dR}{dt}).(1 - W_{i,j}).\overline{G(i)}.G(j) \qquad (2)$$

In the same time, if a source is present at the destination place the corresponding transition is associated with a motivation neuron. After some time, exploring the environment leads to the creation of the cognitive map. The prefrontal cortex is the place in our model where this cognitive map is coded. This seems to be coherent with neurobiological data [23]. This topological map may be seen as a graph where each vertices is a transition and where the edges code for a path between two transitions. No position in a fixed reference is assigned to the vertices of the graph and edges code for adjacence relation only.

7   Autonomous planning using the cognitive map

Some places are more important because they are goals that have to be reached when necessary. When a goal has to be reached, the transitions leading to it are activated. This activation is then diffused on the cognitive map graph, each node taking the maximal incoming value which is the product between the weight on the link and the activity of the node sending the link. After stabilization, this diffusion process gives the shortest path between all nodes and the goal node. This is a neural version of the Bellman-Ford algorithm [2][24, 25] (see fig. 5).

When the robot is in a particular location A, all possible transitions beginning with A are predicted and filtered from the n most activated place cells (similar to the multiple hypothesis position tracking, described in [4], where several position hypothesis can be used in constrast with a 'single' position following). The top-down effect of the cognitive map is to bias these predicted transitions such that the ones chosen by the cognitive map have a higher value. This small bias is enough to select/filter the appropriate transitions via a competition mechanism. This results in a unique movement vector to apply to the robot motor command. See fig. 6 for an illustration of a path followed.

---

[2] The Bellman-Ford algorithm allows to find the shortest path between any node and a goal node of a weighted graph.
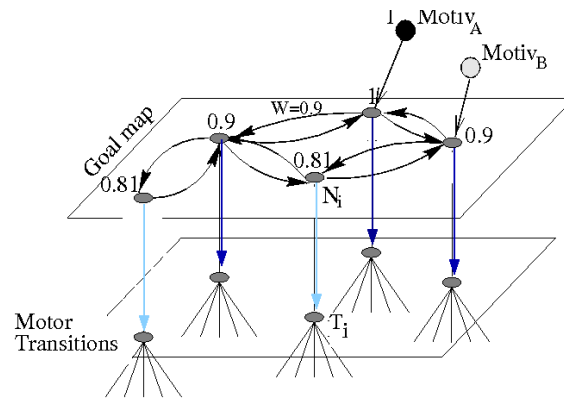
Fig. 5. Diffusion of the activity on the graph corresponding to the cognitive map. Diffusion is starting from the goal. Each vertice keeps the maximal activity coming from its neighbors. Corresponding motor transitions (integrated movement) are then biased by this activity.

Fig. 6. A simulated environment fully explored. Each region represents the place field of a particular place cell. After a full exploration, the entire environment is covered by the place cell population. The curve is an example of a planned path to reach a goal place (presence of a source).

8   Discussion

Exploration periods may be alternated with planning periods. The choice of the behavior is obtained through the self-regulation of two control variables: first the motivational information which allows to trigger a planning behavior, and second, a detection signal triggering a period of exploration. This signal is generated while a new transition is learned meaning that the planning behavior leads the robot in a place still unknown (case of an incomplete map). Planning then restarts as soon as the robot is able to predict transitions from the current place.

Our model currently running on robots (Koala robots and Labo3 robots) has interesting properties in terms of autonomous behavior. However, this autonomy has some drawbacks:

- we are not able to build a cartesian map of the environment because all locations learned are robot centered. However, the places in the cognitive map and the direction used give a skeleton of the environment.
- we have no information about the *exact* size of the rooms or corridors. Again, the cognitive map only gives a sketch of the environment.
- some parameters have to be set, in particular the recognition threshold (section 4). The higher the threshold, the more places are created.

The transitions used in this model may also be the elementary block of a sequence learning process. Thus, we are able to propose a unified vision of the spatial (navigation) and temporal (memory) functions of the hippocampus [26]. However how to go from a graph of transitions to a sequence of transitions of any length is still an open question. This will be part of the next step of the work. The same scaling problem appears when one wants to code several different maps. Each map should be linked with a kind of context signal (which floor or which room) that should be able to "reload" the previous learned map (or a part of it) into the different neural structures used here. Again, models are available and should be tested in simulation and on a robot.

References

1. Hafner, V.V.: Cognitive maps in rats and robots. Adaptive Behavior **13** (2005) 87–96
2. Arleo, A., Gerstner, W.: Spatial cognition and neuro-mimetic navigation: A model of hippocampal place cell activity. Biol. Cybern. **83**(3) (2000) 287–299
3. Filliat, D., Meyer, J.A.: Map-based navigation in mobile robots - I. a review of localisation strategies. Journal of Cognitive Systems Research **4**(4) (2003) 243–282
4. Meyer, J.A., Filliat, D.: Map-based navigation in mobile robots - II. a review of map-learning and path-planing strategies. Journal of Cognitive Systems Research **4**(4) (2003) 283–317
5. Franz, M.O., Mallot, H.A.: Biomimetic robot navigation. Robotics and Autonomous Systems **30** (2000) 133–153
6. Schwartz, L.: Computational anatomy and functional architecture of striate cortex: a spatial mapping approach to perceptual coding. Vision Res. **20** (1980) 645–669
7. Joulain, C., Gaussier, P., Revel, A.: Learning to build categories from perception-action associations. In: International Conference on Intelligent Robots and Systems - IROS'97, Grenoble, France, IEEE/RSJ (1997) 857–864

8. Tinbergen, N.: The study of instinct. Oxford University Press, London (1951)
9. O'Keefe, J., Nadel, N.: The hyppocampus as a cognitive map. Clarenton Press, Oxford (1978)
10. Giovannangeli, C., Gaussier, P., Banquet, J.P.: Robot as a tool to study the robustness of visual place cells. In: I3M'2005: International Conference on Conceptual Modeling and Simulation (CMS 2005), Marseille (2005) 97–104
11. Gaussier, P., Leprêtre, S., Quoy, M., Revel, A., Joulain, C., Banquet, J.: Experiments and models about cognitive map learning for motivated navigation. In: Interdisciplinary approaches to robot learning. Volume 24. Robotics and Intelligent Systems Series, World Scientific, ISBN 981-02-4320-0 (2000) 53–94
12. Banquet, J., Gaussier, P., Dreher, J., Joulain, C., Revel, A.: Space-Time, Order and Hierarchy in Fronto-Hippocampal System: A Neural Basis of Personality. In: Cognitive Science Perpectives on Personality and Emotion. Volume 124. Elsevier Science BV Amsterdam (1997)
13. Meyer, J.A., Wilson, S.W.: From animals to animats. In Books2-4, B., ed.: First International Conference on Simulation of Adaptive Behavior, MIT Press (1991)
14. Tolman, E.: Cognitive maps in rats and men. The Psychological Review **55**(4) (1948)
15. Arbib, M., Lieblich, I.: Motivational learning of spatial behavior. In Metzler, J., ed.: Systems Neuroscience, Academic Press (1977) 221–239
16. Schmajuk, N., Blair, H.: Place learning and the dynamics of spatial navigation: a neural network approach. Adaptive Behavior **1** (1992) 353–385
17. Franz, M.O., Schölkopf, B., Mallot, H.A., Bülthoff, H.H.: Learning view graphs for robot navigation. Autonomous Robots **5** (1998) 111–125
18. Bachelder, I.A., Waxman, A.M.: Mobile robot visual mapping and localization: A view-based neurocomputationnal architecture that emulates hippocampal place learning. Neural Networks **7** (1994) 1083–1099
19. Trullier, O., Wiener, S.I., Berthoz, A., Meyer, J.A.: Biologically based artificial navigation systems: review and prospects. Progress in Neurobiology **51** (1997) 483–544
20. Schölkopf, B., Mallot, H.A.: View-based cognitive mapping and path-finding. Adaptive Behavior **3** (1995) 311–348
21. Bugmann, G., Taylor, J., Denham, M.: Route finding by neural nets. In Taylor, J., ed.: Neural Networks, Henley-on-Thames, Alfred Waller Ltd. (1995) 217–230
22. Donnart, J., Meyer, J.: Learning reactive and planning rules in a motivationnally autonomous animat. IEEE Transactions on Systems, Man and Cybernetics-Part B **26**(3) (1996) 381–395
23. V. Hok, E. Save, P.L.S., Poucet, B.: Coding for spatial goals in prelimbic-infralimbic area of the rat frontal cortex. Proceedings of the National Academy of Sciences ((to appear in 2005))
24. Bellman, R.E.: On a routing problem. In: Quaterly of Applied Mathematics. Volume 16. (1958) 87–90
25. Revel, A., Gaussier, P., Leprêtre, S., Banquet, J.: Planification versus sensory-motor conditioning: what are the issues ? In: From Animals to Animats : Simulation of Adaptive Behavior SAB'98. (1998) 129–138
26. Banquet, J., Gaussier, P., Quoy, M., Revel, A., Burnod, Y.: A hierarchy of associations in hippocampo-cortical systems: cognitive maps and navigation strategies. Neural Computation **17**(6) (2005)