Neuro-inspired navigation strategies shifting for robots: Integration of a multiple landmark taxon strategy

K Caluwaerts[1,2], A Favre-Félix[1], M Staffa[1,3],
S N'Guyen[1], C Grand[1], B Girard[1] and M Khamassi[1]

[1] Institut des Systèmes Intelligents et de Robotique (ISIR) UMR7222, Université Pierre et Marie Curie, CNRS, 4 place Jussieu, 75005 Paris, France
[2] Reservoir Lab, Electronics and Information Systems (ELIS) department, Ghent University, Sint-Pietersnieuwstraat 41, 9000 Ghent, Belgium
[3] Dipartimento di Informatica e Sistemistica, Università degli Studi di Napoli Federico II, Via Claudio 21, 80125, Naples, Italy
mehdi.khamassi@isir.upmc.fr

Abstract. Rodents have been widely studied for their adaptive navigation capabilities. They are able to exhibit multiple navigation strategies; some based on simple sensory-motor associations, while others rely on the construction of cognitive maps. We previously proposed a computational model of parallel learning processes during navigation which could reproduce in simulation a wide set of rat behavioral data and which could adaptively control a robot in a changing environment. In this previous robotic implementation the visual approach (or taxon) strategy was however paying attention to the intra-maze landmark only and learned to approach it. Here we replaced this mechanism by a more realistic one where the robot autonomously learns to select relevant landmarks. We show experimentally that the new taxon strategy is efficient, and that it combines robustly with the planning strategy, so as to choose the most efficient strategy given the available sensory information.

1 Introduction

Neurorobotic researches provide a multidisciplinary approach that can both (i) contribute to robotics by taking inspiration from the computational principles underlying animals' behavioral flexibility and (ii) contribute to neurobiology by using robots as platforms to test the robustness of current biological hypotheses [1, 2, 3]. Several neurorobotic projects have in particular focused on the study of spatial cognition inspired by rodents' neurophysiological substrates for navigation. Indeed rats are able to show highly adaptive behaviors whose reproduction could help improve current robots' decisional autonomy. Thus several previous robots have been endowed with biomimetic models to enable them to build a cognitive map of the environment and to efficiently plan trajectories in it [4, 5, 6, 7, 8, 9].

However, an ability that has not yet been thoroughly investigated in Neurorobotics is the coordination of multiple navigation strategies. Indeed, rats and more generally mammals are able to learn to select the most appropriate strategy for a navigation problem, to avoid costly computations associated with their cognitive map when a simple sensorimotor strategy is enough, and to shift from one strategy to another in response to environmental changes [10]. Among the numerous possible strategies, experimental neuroscience studies of strategy interactions favored two main families:

- Response strategies, resulting from the learning of direct sensorimotor associations (like moving towards a cue indicating the goal, which is called a *taxon strategy*).
- Place strategies, where the animal builds an internal representation (or cognitive map) of the various locations of the environment, using the configuration of multiple allocentric cues. It can then use this information to choose the direction of the next movement by planning a path in a graph connecting the places with the actions allowing the transitions from one place to another (*topological planning strategy*).

It has been shown that the multiple navigation strategies of rodents are operated by parallel independent memory systems [11, 12], which can result in *cooperative* or *competitive* behaviors, depending on the experimental protocol. Place strategies would rely on the Hippocampus, with its ability to encode places in the so-called *place cells* [13] and to contribute to trajectory planning computations within a cognitive map [14]. Lesions of the hippocampal system impair place strategies while sparing response strategies [15, 16]. In contrast, lesions of the striatum impair the expression of response strategies while sparing place strategies [16, 11].

In previous work, we proposed a modular computational model that can explain a wide range of behavioral data recorded in rodent laboratory maze tasks [17]. We implemented the model in a robot and showed that it enabled the robot to benefit from the particular advantages of each strategy (the planning strategy being efficient far from the goal, the taxon strategy being more precise in adjusting the robot's trajectory near the goal based on vision) by autonomously learning which strategy was the most efficient in each part of the environment [18]. Moreover, the robot could efficiently adapt to changes in the goal location by detecting contextual changes that require new place-strategy associations.

However, in our previous work the taxon strategy was simplified by having visually access only to the single intramaze cue that indicates the goal location, as in rat experiments, and had to learn how to orientate itself towards this landmark. In contrast, the planning strategy had full access to all landmarks in order to learn a cognitive map of the environment. This was biologically acceptable since biologists assume that taxon strategies rely only on intramaze cues while planning strategy rely on extramaze cues, and since a single intramaze cue is used in most protocols [15, 16]. However in a situation with multiple intramaze cues, the real issue is to learn which cue leads to reward. Moreover, the capability to orient toward a chosen cue is probably hardwired in the Superior

Colliculus [19]. From a robotic point of view, this is also not satisfying because one would want the robot to autonomously learn to select relevant landmarks within its visual field.

In this work, we extended the taxon strategy of our model by having it learn by reinforcement which visual landmarks it should orient toward. We first made an experiment where the taxon strategy competed with a random exploration strategy to show that the taxon could succesfully learn and progressively win the competition against the exploration strategy. Then we combined this taxon strategy with the whole model, thus competing with the planning and exploration strategies. We found that the robot successfully learned to rely less and less on the exploration strategy through learning. Moreover, at the end of learning the robot learned to prefer the taxon strategy in subparts of the environment where the robot could robustly perceive salient landmarks near the goal, while it learned to prefer the planning strategy in parts of the environment where individual landmarks were less reliably associated with goal reaching.

2 Computational model

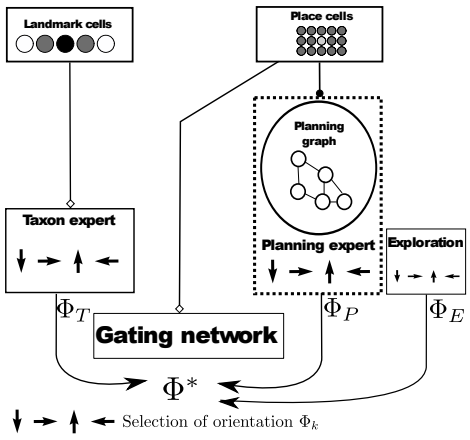

Fig. 1. Overview of the computational model. Different strategies (taxon/planning/exploration) are connected to the gating network. Each strategy has a dedicated expert which proposes actions (Φ_T for the taxon, Φ_P for the planning, Φ_E for the exploration). The gating network decides which of the experts is the winner in the current situation and then the action Φ^* from this strategy is performed.

The model (Fig. 1) is composed of three navigation strategy experts (the taxon, the planning and the exploration), which propose directions for the next movement Φ_k, and a gating network, which learns to choose the most efficient

strategy depending on the current position. As stated in the introduction, it is identical to the one described in detail in [18], except for the taxon expert.

The exploration expert just provides a randomly chosen direction of movement, redrawn every three timesteps, so as to ensure that the selection of this strategy on successive timesteps results in coherent movements. It does not learn at all.

The planning expert builds a graph of states and transitions, based on the place-cells activity provided by a simplified model of the hippocampus (states), and the experience of which action allows to go from one place to another. Places where rewards are encountered are learned, so that path planning can then be computed.

The taxon expert implements a standard neural implementation of a Q-learning algorithm. It takes as input state space a representation of the perceived landmarks configuration (namely, which landmark is visible and at which distance) and learns in each state to choose a landmark towards which the robot should orient. The following transformation of this choice into an egocentric direction of movement using the camera information is hardwired (see 2.1 below for more details).

Finally, the gating network also implements a Q-learning algorithm, which takes as input the activity of the place cells (hence the current estimation of position), and learns to choose the strategy which is the most efficient in each position for maximizing future reward.

A specificity of the [17] model is that the adaptive navigation experts (here taxon and planning) receive reward signals even when they did not generate the movement, so that they can learn from each other: for the taxon, the Q-value of the landmark whose direction is the closest to the direction of the real movement (no matter if it was generated by the taxon itself, by the planning or by the exploration expert) is updated; for the planning the position of the reward is recorded whichever expert led the robot there.

The planning expert (place cells + planning graph) has been extensively described in [17, 18]. In following sections, we provide the equations for the newly implemented taxon expert and for its coordination with other experts by the gating network.

2.1 Taxon strategy

The taxon learns to choose the best landmark to guide the orientation behavior, based on the current distance configuration of the landmarks, using a Q-learning algorithm [20]. This means that the taxon selects a landmark (action) based on the landmark configuration (state) that the robot sees and then proposes the direction towards this landmark to the gating network (section 2.2).

Indeed, the visual system automatically detects the N_L visual landmarks, based on the color of contrasted patches and on their shape, and evaluates their distance based on the binocular disparity. The distance (in meters) of each landmark is discretized in five possible ranges $([0, 0.5],]0.5, 1],]1, 2],]2, 3],]3, +\infty[)$,

and for each landmark l a corresponding 5-component vector I_l with a one on the component corresponding to the detected distance is produced, and zeros elsewhere. The input I of the taxon is the concatenation of all these landmark distance vectors. The output $O = W^{taxon} I$ is a N_L-long vector, attributing a value to each of the possible landmark choice. The final selection of the landmarks is simply greedy, the chosen landmark L is:

$$L = \underset{i \in [0, N_L]}{\arg\max}\, O_i \tag{1}$$

as the exploration necessary for the convergence of such an algorithm is handled by the exploration expert, and the regulation of the amount of exploration results directly from the gating network choices (see below).

A standard neural implementation of a Q-learning algorithm [20] is then used to learn the $[N_L \times 5] \times [N_L]$ weight matrix W^{taxon} associating landmark distance configurations to landmarks. The prediction error δ is computed and the matrix W^{taxon} updated accordingly:

$$\delta(t+1) = r_{t+1} + \gamma \max_{i \in [0, N_L]} (W^{taxon} I(t+1))_i - W^{taxon} I(t) \tag{2}$$

$$W^{taxon}_{i,L} \leftarrow W^{taxon}_{i,L} + \alpha \delta(t+1) \tag{3}$$

When the taxon is not chosen by the gating network, and thus has not chosen the last direction of movement, the landmark $\hat{L}$ whose direction $d_{\hat{L}}$ is the closest to the chosen direction d is updated as follows:

$$\hat{L} = \underset{i \in [0, N_L]}{\arg\max}\, d_i \cdot d \tag{4}$$

$$W^{taxon}_{i,\hat{L}} \leftarrow W^{taxon}_{i,\hat{L}} + \alpha \delta(t+1) \tag{5}$$

2.2 Gating network

The gating network learns the Q-value $g^k(t)$ of each expert k (called gating-values), based on a matrix of weights $z^k_j(t)$:

$$g^k(t) = \sum_j^{N_{PC}} z^k_j(t) n^{PC}_j(t) \tag{6}$$

The selection probability of an expert is then computed as follows:

$$\mathrm{P}(\Phi^*(t) = \Phi^k(t)) = \frac{g^k(t)}{\sum_i g^i(t)} \tag{7}$$

Here $\Phi^k(t)$ is the action proposed by expert k at time t. $\Phi^*(t)$ is the final action proposed by the gating network. The gating network is a strategy selection mechanism instead of an action selection mechanism. It selects a winning

strategy ($*$) at each action step and the action (a new heading direction) proposed by this strategy will be executed unless a higher priority mechanism (e.g. hardwired wall avoidance) is activated. This is an imporant part of the system, as the gating network is independent of the actions proposed by the strategies when it selects a strategy.. If the executed action was not produced by any of the strategies feeding into the gating network (e.g. wall avoidance), the gating network and the strategies themselves can still learn as the global reward and executed action are shared between all strategies and the gating network.

Learning is sped up by using action generalization and eligibility traces, the detailed equations for these techniques are to be found in [18]. To modify the Q-values, a modified Q-learning algorithm [20] is applied:

$$\Delta z_j^k(t) = \xi \delta(t) e_j^k(t+1) \tag{8}$$

$$\delta(t) = R(t+1) + \gamma \max_k(g^k(t+1)) - g^{k^*}(t) \tag{9}$$

where ξ is the learning rate of the algorithm and δ the reward prediction error, and $e_j^k(t+1)$ the eligibility trace.

The reward prediction error δ is based on the observed reward when performing action $\Phi*$ and the future expected reward (g^{k^*} is the activation of the winning output neuron and γ the discount factor). The eligibility trace e_j^k allows the reinforcement of previously selected strategies and the strategies proposing a direction close to the one proposed by the winning strategy [20]:

$$e_j^k(t+1) = \nu(t)\Psi(\Phi^*(t) - \Phi^k(t))r_j^{PC}(t) + \lambda e_j^k(t). \tag{10}$$

The eligibility traces depend on a time-varying value $\nu(t)$ which is a measure of the quality of the sensory input as estimated by the robot, $r_j^{PC}(t)$, the place cells' activations and a decay factor λ [18].

The gating network is a simple but effective way to combine competition and cooperation between strategies, as experimentally observed in rat behavior [15]. While the gating network itself only directly provides competition, the strategies cooperate by sharing rewards and their actions (e.g. the taxon uses the executed action for learning, instead of its proposed action (Eq. 4)). Hence the gating network is advantageous for strategies as they can learn from each other, while at the global level the performance can also increase because the best performing strategy can be used in each situation.

3 Experimental setup

The robot is a mobile platform equipped with two motorized wheels and a rotating head on which sensors are fixed (see Fig. 2-left). We use here the two frontal cameras, and given their limited aperture (60 degrees), the robot regularly makes head rotation movements so as to acquire panoramas. The combination of these panoramas with a memory of the previously seen landmarks allows to generate

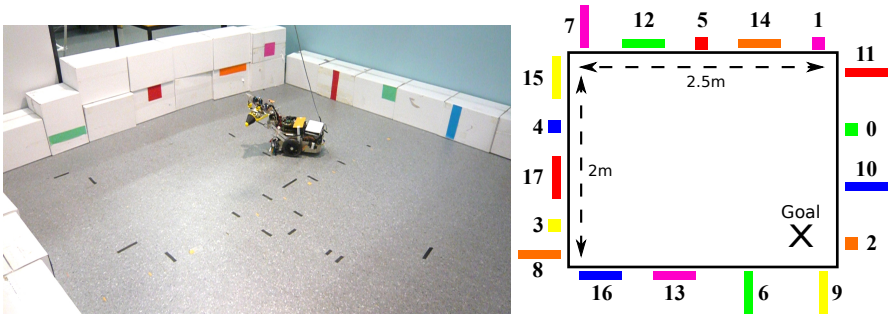

Fig. 2. Open field arena used for the experiments with the Psikharpax rat robot. The arena is surrounded by 18 landmarks differentiated by their color and shape. The goal is located in the south-east part of the arena.

estimations of the landmark configuration over 360 degrees (see [18] for a detailed description of the different layers of the visual system of the model). The landmarks are identified based on their color and shape (only unique combinations were used so as to avoid aliasing, see Fig. 2-right), and their distance was roughly estimated using binocular discrepancy.

The robot makes discrete movements, moving 10 cm at each timestep in an open 2 m by 2.5 m environment (Fig. 2). 18 different landmark cues are distributed around the arena. An invisible 20 cm diameter zone (314 cm^2, or 1/160th of the environment) is defined as the goal location that the robot has to learn to efficiently reach by trial-and-error. When the robot reaches this zone, the reward is set to one, then the robot is moved to a different starting location to begin a new trial. Anywhere else the reward is 0. The egocentric reference frame has the neck of the robot as its origin and the orientation is defined by the direction of the head.

All experiments presented here follow an exploration phase where the robot moves randomly without getting reward and builds place cells and a cognitive map (topological links between places) of the environment based on vision and odometry. This process and its robustness to the number of landmarks, to noise and to the model's parameters have been extensively described in [18]. Here we focus on the learning processes that both allow the taxon strategy to progressively select appropriate landmarks to orient toward, and simultaneously allow the gating network to progressively select the most appropriate strategy (among taxon, planning and exploration strategies) in each part of the arena.

4 Results

In all experiments, the robot is initially positioned in one of the three corners, far from the goal. If after 5 meters of movement the robot has still not found the goal, it is guided directly towards the goal as in rat experiments and as in our previous robotic work [18]. During guiding, the strategies (taxon/planning) can

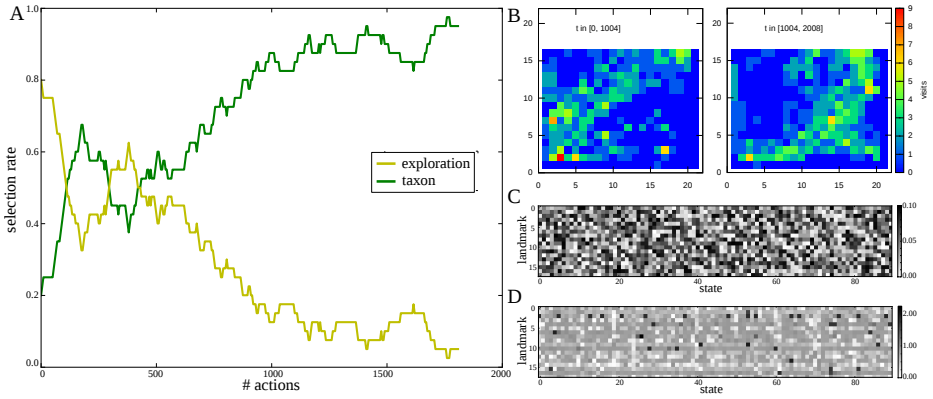

Fig. 3. Behavioral adaptation in the taxon experiment. A: the selection rate (averaged over 200 actions) illustrates that the model has successfully learned to prefer the taxon strategy over time. B: histogram of the number of times the taxon strategy is selected by the gating network at each position in the environment, during the first (left) and the second (right) half of the experiment. C and D: the taxon's weight matrices (associating the landmark distance configuration (state) with the landmark to aim for (action)) before and after learning (note the change of scale).

still learn, hence guiding can speed up learning. Guiding is also used to lead the robot back to one of the corners after receiving a reward to start a new trial.

4.1 Evaluation of the taxon

As we modified the taxon from our previous work, we first test the efficiency of this new expert, by running the model without planning. An experiment is run for 2008 timesteps, corresponding to 32 goal-reaching sequences (*i.e.* 32 trials).

The results confirm that the taxon has learned and has become more and more efficient, as it is predominantly selected by the gating network (Fig. 3, A). The histogram locations occupied by the robot when the taxon strategy is selected (Fig. 3, B) shows a pattern that switches from random locations in the first half of the experiment (left), to locations on trajectories leading to the goal for the second half (right). Finally, the matrices of weights at the beginning of the experiment (Fig. 3, C) and at the end (Fig. 3, D) have evolved from a random pattern to the selection of a limited number of landmarks, most of them close to the goal (mainly landmarks #2, #6, #9 and #10).

4.2 Full model (taxon+planning+exploration)

We then test the full model, in an experiment lasting 2664 timesteps, corresponding to 43 goal-reaching sequences. Here, the strategy selection evolution (Fig. 4-A) shows that exploration is rapidly replaced by the taxon and planning strategies as they become more and more efficient. Unfortunately, this experiment is not long enough for convergence of the gating network's learning

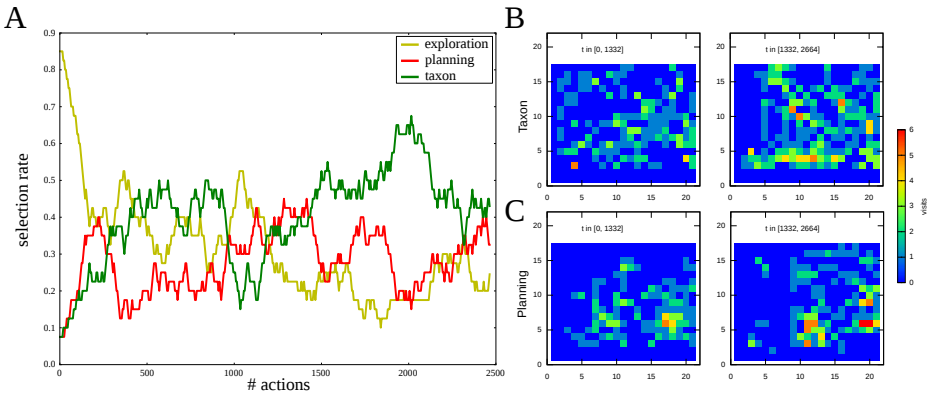

Fig. 4. Behavioral adaptation in the full model experiment. A: evolution over time of the selection rate (averaged over 200 actions) of the three stategies, illustrating the progressive learning of the taxon and planning strategies. B and C: histograms in the first (left) and the second (right) half of the experiment for the taxon (B) and planning (C) strategies indicating how many times the strategy was selected by the gating network. Similarly to the previous experiment, the taxon learns direct trajectories to the goal, while the planning is used in areas where the taxon is less efficient.

mechanism, so that exploration is still selected 20% of the time. Nevertheless, the histograms of positions where the taxon (Fig. 4-B) is selected show that the taxon has learned very similarly to the previous experiment: during the second part of the experiment ($t \in [1332 : 2664]$) the taxon is mainly recruted along direct paths to the goal (*i.e.* along the southern wall and along the diagonal from the northwestern corner until the goal). Moreover, the planning is selected in places where the taxon is less efficient (Fig. 4-C), thus showing a complementary recruitment of the two strategies by the gating network. Quantitatively, excepting moments when the robot is exploring, the taxon and planning strategies are both selected during a substantial proportion of time during the second part of the experiment (respectively 62% and 38% of the time). More precisely, there are particular regions within the arena where the taxon is not sufficiently efficient and where the planning is thus preferred (Fig. 5). This is especially clear along the eastern wall, where landmark #2 is often not seen. This is also the case in the south central part of the arena, which falls outside the paths followed by the taxon. At the end of the experiment, exploration is mainly used far away from the goal, where neither taxon nor planning are yet fully efficient.

5 Discussion

We have presented the integration of a multiple landmark taxon in our strategy selection model allowing an autonomous robot to navigate in an initially unknown environment. The model selects among two parallelly learned navigation strategies: a response strategy learning to orient towards relevant cues in the vi-

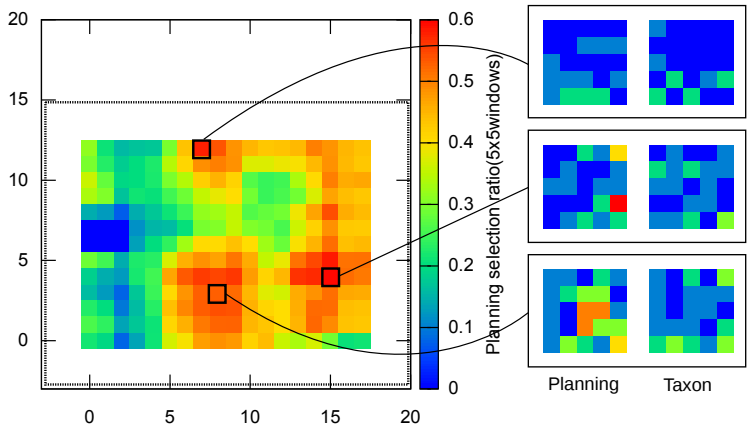

Fig. 5. Relative proportion of selection of the planning strategy over the taxon strategy during the second part of the full model experiment, averaged over $50cm \times 50cm$ sliding windows. Note the 0 to 60% scale. The planning strategy is overtly preferred in three different regions. The right insets show, for three points (black squares) in these regions, the corresponding windows of Planning and Taxon selection extracted from figure 4. The upper region seems to be an artefact, given the low number of measurements; however in the two others, the taxon is clearly less chosen, being locally less efficient or reliable.

sual field; a place strategy building a map of place cells and planning trajectories between different locations in the arena. This model constitutes an extension to a previously published model of multiple navigation strategies [17] which was tested in simulation to replicate a series of rat behavioral experimental results [15, 16], and which was previously applied to robotics in tasks involving a single intramaze cue [18].

Here we first show that the new taxon can succesfully learn to orient towards relevant landmarks to reach the goal, and thus that the gating network can parallely learn to prefer the taxon strategy over the exploration strategy. Then we combine this taxon strategy with the whole model, thus competing with the planning and exploration strategies. We find that the robot successfully learns to rely less and less on the exploration strategy through learning. Moreover, at the end of learning the robot has learned to prefer the taxon strategy in subparts of the environment where the robot can robustly perceive salient landmarks near the goal, while it has learned to prefer the planning strategy in parts of the environment where perceived individual landmarks are less reliably associated with goal reaching.

These results show that the model generalizes well with a taxon that takes into account all visual landmarks in the environment. They also validate in a new experiment the ability of the model to use the specific advantages of each strategy in each subpart of the environment (*e.g.* a local taxon strategy combined with a global but coarse path planning strategy). The complete validation of the

model will require to test it in non-stationary environments (*e.g.* changes in goal location), as shown with the previous simplified taxon expert [18]. Future work will also test the ability of the model to achieve more complex robotic tasks involving a larger environment and the apparition/vanishing of new objects that can constitute obstacles for the robot.

Besides the interest of such modelling approach to contribute to a better formalization of rodent navigation behavior, this work has also the potential of contributing to mobile robotics. Indeed, the bio-inspired ability to rapidly switch between several behavioral strategies, and to memorize which strategy is the most efficient and appropriate in each subzone of the environment could help improve current control architectures for robots. Multi-layered control architectures with different levels of decisions have become more and more popular in robotics and are now widely used [21, 22]. Such architectures raise issues such as managing the interactions between submodules, coordinating multiple competing learning processes and providing alternative solutions to motion planning in situations where such strategy is limited [23]. Indeed the planning strategy can be approximate when coping with uncertainties, *e.g.* when there is perceptual aliasing as we illustrated in [18], and can also require high computational costs and long times to propagate possible trajectories through mental maps [21]. In contrast, in situations where animals have developed habits under the form of cue-guided taxon or response strategies to solve a particular task, they can perform quick and accurate decision-making. Taking inspiration from computational models of how mammals progressively shift from costly decision-making to habits as a function of a speed-accuracy trade-off may constitute the basis of great future advances in robotics [24].

Acknowledgments

This research was funded by the EC FP6 IST 027819 ICEA Project and a Ph.D. fellowship of the Research Foundation - Flanders (FWO).

References

[1] Pfeifer, R., Lungarella, M., Iida, F.: Self-organization, embodiment, and biologically inspired robotics. Science **318** (2007) 1088–1093
[2] Arbib, M., Metta, G., van der Smagt, P.: Neurorobotics: From vision to action. In: Handbook of robotics. Springer-Verlag, Berlin (2008) 1453–1480
[3] Meyer, J.A., Guillot, A.: Biologically-inspired robots. In: Handbook of robotics. Springer-Verlag, Berlin (2008) 1395–1422
[4] Arleo, A., Gerstner, W.: Spatial cognition and neuro-mimetic navigation: a model of hippocampal place cell activity. Biological Cybernetics **83**(3) (2000) 287–299
[5] Krichmar, J., Seth, A., Nitz, D., Fleischer, J., Edelman, G.: Spatial navigation and causal analysis in a brain-based device modeling corticalhippocampal interactions. Neuroinformatics **3**(3) (2005) 147–169
[6] Meyer, J.A., Guillot, A., Girard, B., Khamassi, M., Pirim, P., Berthoz, A.: The Psikharpax project: towards building an artificial rat. Robotics and Autonomous Systems **50**(4) (2005) 211–223

[7] Barrera, A., Weitzenfeld, A.: Biologically-inspired robot spatial cognition based on rat neurophysiological studies. Autonomous Robots **25** (2008) 147–169
[8] Giovannangeli, C., Gaussier, P.: Autonomous vision-based navigation: Goal-oriented action planning by transient states prediction, cognitive map building, and sensory-motor learning. In: Proceedings of the International Conference on Intelligent Robots and Systems. Volume 1., University of California Press (2008) 281–297
[9] Milford, M., Wyeth, G.: Persistent navigation and mapping using a biologically inspired slam system. The International Journal of Robotics Research **29**(9) (2010) 1131–1153
[10] Arleo, A., Rondi-Reig, L.: Multimodal sensory integration and concurrent navigation strategies for spatial cognition in real and artificial organisms. Journal of Integrative Neuroscience **6**(3) (2007) 327–366
[11] Packard, M.G., Knowlton, B.J.: Learning and memory functions of the basal ganglia. Annual Review of Neuroscience **25** (2002) 563–593
[12] Burgess, N.: Spatial cognition and the brain. Year In Cognitive Neuroscience 2008 **1124** (2008) 77–97
[13] O'Keefe, J., Nadel, L.: The Hippocampus as a Cognitive Map. Clarendon Press, Oxford (1978)
[14] Johnson, A., Redish, A.: Neural ensembles in CA3 transiently encode paths forward of the animal at a decision point. Journal of Neuroscience **27**(45) (2007) 12176–12189
[15] Pearce, J., Roberts, A., Good, M.: Hippocampal lesions disrupt navigation based on cognitive maps but not heading vectors. Nature **396**(6706) (1998) 75–77
[16] Devan, B., White, N.: Parallel information processing in the dorsal striatum: Relation to hippocampal function. Journal of Neuroscience **19**(7) (1999) 2789–2798
[17] Dollé, L., Sheynikhovich, D., Girard, B., Chavarriaga, R., Guillot, A.: Path planning versus cue responding: a bioinspired model of switching between navigation strategies. Biological Cybernetics **103**(4) (2010) 299–317
[18] Caluwaerts, K., Staffa, M., N'Guyen, S., Grand, C., Dollé, L., Favre-Félix, A., Girard, B., Khamassi, M.: A biologically inspired meta-control navigation system for the psikharpax rat robot. Bioinspiration and Biomimetics (2012) to appear.
[19] Stein, B., Meredith, M.: The merging of the senses. The MIT Press, Cambridge, MA (1993)
[20] Sutton, R., Barto, A.: Reinforcement Learning: An Introduction. MIT Press (1998)
[21] Gat, E.: On three-layer architectures. In Kortenkamp, D., Bonnasso, R., Murphy, R., eds.: Artificial Intelligence and Mobile Robots: Case Studies of Successful Robot Systems. AAAI Press (1998) 195–210
[22] Kortenkamp, D., Simmons, R.: Robotic systems architectures and programming. In Siciliano, B., Khatib, O., eds.: Handbook of Robotics. Springer-Verlag (2008) 187–206
[23] Minguez, J., Lamiraux, F., Laumond, J.: Motion planning and obstacle avoidance. In Siciliano, B., Khatib, O., eds.: Handbook of Robotics. Springer-Verlag (2008) 827–852
[24] Keramati, M., Dezfouli, A., Piray, P.: Speed/accuracy trade-off between the habitual and goal-directed processes. PLoS Computational Biology **7**(5) (2011) 1–25