Map-Based Spatial Navigation: A Cortical Column Model for Action Planning

Louis-Emmanuel Martinet[1,2,3], Jean-Baptiste Passot[2,3], Benjamin Fouque[1,2,3], Jean-Arcady Meyer[1], and Angelo Arleo[2,3]

[1] UPMC Univ Paris 6, FRE2507, ISIR, F-75016, Paris, France
[2] UPMC Univ Paris 6, UMR 7102, F-75005, Paris, France
[3] CNRS, UMR 7102, F-75005, Paris, France
louis-emmanuel.martinet@upmc.fr

Abstract. We modelled the cortical columnar organisation to design a neuromimetic architecture for topological spatial learning and action planning. Here, we first introduce the biological constraints and the hypotheses upon which our model was based. Then, we describe the learning architecture, and we provide a series of numerical simulation results. The system was validated on a classical spatial learning task, the Tolman & Honzik's *detour* protocol, which enabled us to assess the ability of the model to build topological representations suitable for spatial planning, and to use them to perform flexible goal-directed behaviour (e.g., to predict the outcome of alternative trajectories avoiding dynamically blocked pathways). We show that the model reproduced the navigation performance of rodents in terms of goal-directed path selection. In addition, we present a series of statistical and information theoretic analyses to study the neural coding properties of the learnt space representations.

Keywords: spatial navigation, topological map, trajectory planning, cortical column, hippocampal formation.

1 Introduction

Spatial cognition calls upon the ability to learn neural representations of the spatio-temporal properties of the environment, and to employ them to achieve goal-oriented navigation. Similar to other high-level functions, spatial cognition involves parallel information processing mediated by a network of brain structures that interact to promote effective spatial behaviour [1,2]. An extensive body of experimental work has investigated the neural bases of spatial cognition, and a significant amount of evidence points towards a prominent role of the hippocampal formation (see [1] for recent reviews). This limbic region has been thought to mediate spatial learning functions ever since location-selective neurones — namely hippocampal *place cells* [3], and entorhinal *grid cells* [4] — and orientation-selective neurones — namely *head-direction cells* [5] — were found by means of electrophysiological recordings from freely moving rats.

Hippocampal place cells, grid cells, and head-direction cells are likely to subserve spatial representations in allocentric (i.e., world centred) coordinates, thus

C. Freksa et al. (Eds.): Spatial Cognition VI, LNAI 5248, pp. 39–55, 2008.

providing *cognitive maps* [3] to support spatial behaviour. Yet, to perform flexible navigation (i.e., to plan detours and/or shortcuts) two other components are necessary: goal representation, and target-dependent action sequence planning [6]. The role of the hippocampal formation in these two mechanisms remains unclear. On the one hand, the hippocampus has been proposed to encode topological-like representations suitable for action sequence learning [6]. This hypothesis mainly relies on the recurrent dynamics generated by the CA3 collaterals of the hippocampus [7]. On the other hand, the hippocampal space code is likely to be highly redundant and distributed [8], which does not seem adequate for learning compact topological representations of high-dimensional spatial contexts. Also, the experimental evidence for high-level spatial representations mediated by a network of neocortical areas (e.g., the posterior parietal cortex [9], and the prefrontal cortex [10]) suggests the existence of an extra-hippocampal action planning system shared among multiple brain regions [11]. This hypothesis postulates a distributed spatial cognition system in which *(i)* the hippocampus would take part to the action planning process by conveying redundant (and robust) spatial representations to higher associative areas, *(ii)* a cortical network would elaborate more abstract and compact representations of the spatial context (accounting for motivation-dependent memories, action cost/risk constraints, and temporal sequences of goal-directed behavioural responses). Among the cortical areas involved in map building and action planning, the prefrontal cortex (PFC) may play a central role, as suggested by anatomical PFC lesion studies showing impaired navigation planning in rats [12]. Also, the anatomo-functional properties of the PFC seem appropriate to encode abstract contextual memories not merely based on spatial correlates. The PFC receives direct projections from sub-cortical structures (e.g., the hippocampus [13], the amygdala [14], and the ventral tegmental area [15]), and indirect connections from the basal ganglia through the basal ganglia - thalamocortical loops [16]. These projections provide the PFC with a multidimensional context, including emotional and motivational inputs [17], reward-dependent modulation [18], and action-related signals [16]. The PFC seems then well suited to *(i)* process manifold spatial information [19], *(ii)* encode the motivational values associated to spatio-temporal events [6], and *(iii)* perform supra-modal decisions [20]. Also, the PFC may be involved in integrating events in the temporal domain at multiple time scales [21]. The PFC recurrent dynamics regulated by the modulatory action of dopaminergic afferents [22] may permit to maintain patterns of activity over long time scales. Finally, the PFC is likely to be critical to detecting cross-temporal contingencies, which is relevant to the temporal organisation of behavioural responses, and to the encoding of retrospective and prospective memories [21].

1.1 Cortical Columnar Organisation: A Computational Principle?

The existence of cortical columns was first reported by Mountcastle [23], who observed chains of cortical neurones reacting to the same external stimuli simultaneously. Cortical columns can be divided in six main layers including: layer I, mostly containing axons and dendrites; layer IV, receiving sensory inputs from

sub-cortical structures (mainly the thalamus); and layer VI, sending outputs to sub-cortical brain areas (e.g., to the striatum and the thalamus). Layers II-III and V-VI constitute the so called *supragranular* and *infragranular* layers, respectively. The anatomo-functional properties of cortical columns have been widely investigated [24]. Neuroanatomical findings have indicated that columns can be divided into several *minicolumns*, each of which is composed of a population of interconnected neurones [25]. Thus, a column can be seen as an ensemble of interrelated minicolumns receiving inputs from cortical areas and other structures. It processes these afferent signals and projects the responses both within and outside the cortical network. This twofold columnar organisation has been suggested to subserve efficient computation and information processing [24].

1.2 Related Work

This paper presents a neuromimetic model of action planning inspired by the columnar organisation of the mammalian neocortex. Planning is defined here as the ability, given a state space S and an action space A, to "mentally" explore the $S \times A$ space to infer an appropriate sequence of actions leading to a goal state $s_g \in S$. This definition calls upon the capability of *(i)* predicting the consequences of actions, i.e. the most likely state $s' \in S$ to be reached when an action $a \in A$ is executed from a state $s \in S$, *(ii)* evaluating the effectiveness of the selected plan on-line. The model generates a topological representation of the environment, and it employs an activation-diffusion mechanism [26] to plan goal-directed trajectories. The activation-diffusion process is based on the propagation of a reward-dependent activity signal from the goal state s_g through the entire topological network. This propagation process enables the system to generate action sequences (i.e., trajectories) from the current state s towards s_g.

Topological map learning and path planning have been extensively studied in biomimetic robotics (see [27] for a review). Here we focus on model architectures that take inspiration from the anatomical organisation of the cortex, and implement an activation-diffusion planning principle. Burnod [28] proposed one of the first models of the cortical column architecture, called "cortical automaton". He also described a "call tree" process that can be seen as a neuromimetic implementation of the activation-diffusion principle. Several action selection models were inspired by Burnod's hypothesis. Some of these works employed the cortical automaton concept explicitly [29,30,31]. Others used either connectionist architectures [32,33,34] or Markov decision processes [35]. Yet, none of these works took into account the multilevel coding property offered by the possibility to refine the cortical organisation by adding a sublevel to the column, i.e. the minicolumn. The topological representation presented here exploits this idea by associating the columnar level to a compact representation of the environment, and by employing the minicolumn level to characterise the agent behaviour. In order to validate the model, we have implemented it on a simulated robotic platform, and tested it on the classical Tolman & Honzik's navigation task [36]. This protocol allowed us to assess the ability of the system to learn topological

representations, and to exploit them to perform flexible goal-directed behaviour (e.g., planning *detours*).

2 Methods

2.1 Single Neurone Model

The elementary computational units of the model are artificial firing-rate neurones i, whose mean discharge $r_i \in [0, 1]$ is given by

$$r_i(t) = f\Big(V_i(t) \cdot (1 \pm \eta)\Big) \ . \tag{1}$$

where $V_i(t)$ is the membrane potential at time t, f is the transfer function, and η is a random noise uniformly drawn from $[0, 0.01]$. V_i varies according to

$$\tau_i \cdot \frac{dV_i(t)}{dt} = -V_i(t) + I_i(t) \ . \tag{2}$$

where $\tau_i = 10$ ms is the membrane time constant, and $I_i(t)$ is the synaptic drive generated by all the inputs. Eq. 2 is integrated by using a time step $\Delta_t = 1$ ms. Both the synaptic drive $I_i(t)$ and the transfer function f are characteristic of the different types of model units, and they will be defined thereafter.

2.2 Encoding Space and Actions: Minicolumn and Column Model

The main inputs to the cortical model are the location- and orientation-selective activities of hippocampal place and head-direction cells, respectively [3,5]. The hippocampal place field representation is built incrementally as the simulated animal (i.e., the *animat*) explores the environment, and it provides the system with a continuous distributed and redundant state representation S [37,38]. A major objective of the cortical model was to build a compact state-action representation $S \times A$ suitable for topological map learning and action planning.

In the model, the basic component of the columnar organisation is the minicolumn (vertical grey regions in Fig. 1). An unsupervised learning scheme (Sec. 2.3) makes the activity of each minicolumn selective to a specific state-action pair $(s, a) \in S \times A$. Notice that a given action $a \in A$ represents the allocentric motion direction of the animat when it performs the transition between two locations $s, s' \in S$. According to the learning algorithm, all the minicolumns selective for the same spatial location $s \in S$ are grouped to form a higher-level computational unit, i.e. the column (see c and c' in Fig. 1A). This architecture is inspired by biological data showing that minicolumns inside a column have similar selectivity properties [39]. Thus, columns consist of a set of minicolumns that are incrementally recruited to encode all the state-action pairs $(s, a_{1\cdots N}) \in S \times A$ experienced by the animat at a location s. During planning (Sec. 2.4), all the minicolumns of a column compete with each other to locally infer the most appropriate goal-directed action.

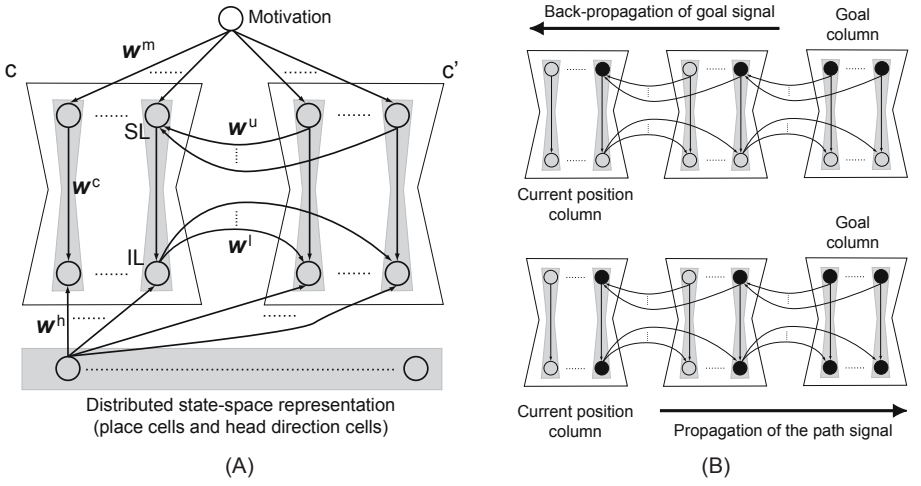

Fig. 1. The cortical model and the implementation of the activation-diffusion process. **(A)** Columns (c and c') consist of sets of minicolumns (vertical grey regions), each of which contains a supragranular (SL) and an infragranular (IL) layer unit. **(B)** Top: back-propagation of the motivational signal through the network of SL neurones. Bottom: forward-propagation of the goal-directed action signal through the IL neurones.

Every minicolumn of the model consists of two computational units, representing supragranular layer (SL) and infragranular layer (IL) neurones (Fig. 1A). The discharge of SL and IL units simulates the mean firing activity of a population of cortical neurones in layers II-III, and V-VI, respectively. Each minicolumn receives three different sets of afferent projections (Fig. 1A): *(i)* Hippocampal inputs conveying allocentric space coding signals converge onto IL neurones; these connections are plastic, and their synaptic efficacy is determined by the weight distribution $\boldsymbol{w}^h$ (all the synaptic weights of the model are within the maximum range of $[0, 1]$). *(ii)* Collateral afferents from adjacent cortical columns converge onto SL and IL neurones via the projections $\boldsymbol{w}^u$ and $\boldsymbol{w}^l$, respectively. These lateral connections are learnt incrementally (Sec. 2.3), and play a prominent role in both encoding the environment topology and implementing the activation-diffusion planning mechanism. *(iii)* SL neurones receive projections $\boldsymbol{w}^m$ conveying motivation-dependent signals. As shown in Sec. 2.4, this input is employed to relate the activity of a minicolumn to goal locations.

SL neurones discharge as a function of the motivational signals mediated by the $\boldsymbol{w}^u$ and $\boldsymbol{w}^m$ projections. The synaptic drive $I_i(t)$ depolarising a SL neurone i that belongs to a column c is given by:

$$I_i(t) = \max_{i' \in c' \neq c} \left\{ w^u_{ii'} \cdot r_{i'}(t) \right\} + w^m_i \cdot r_m \ . \tag{3}$$

where i' indexes other SL neurones of the cortical network; w^m_i and r_m are the weight and the intensity of the motivational signal, respectively. In the current

version of the model the motivational input is generated algorithmically, i.e. $w_i^m = 1$ if column c is associated to the goal location, $w_i^m = 0$ otherwise, and the motivational signal $r_m = 1$. The membrane potential of unit i is then computed according to Eq. 2, and its firing rate $r_i(t)$ is obtained by means of an identity transfer function f.

Within each minicolumn, SL neurones project onto IL units by means of non-plastic projections $\boldsymbol{w}^c$ (Fig. 1A). Thus, IL neurones are driven by hippocampal place (HP) cells h (via the projections $\boldsymbol{w}^h$), IL neurones belonging to adjacent columns (via the collaterals $\boldsymbol{w}^l$), and SL units i (via $\boldsymbol{w}^c$). The synaptic drive of a IL neurone $j \in c$ is:

$$I_j(t) = \max\Big\{ \sum_{h \in HP} w_{jh}^h \cdot r_h(t) \ , \ \max_{j' \in c' \neq c} \big\{ w_{jj'}^l \cdot r_{j'}(t) \big\} \Big\} + w_{ji}^c \cdot r_i(t) \ . \quad (4)$$

where j' indicates other IL neurones of the network; $w_{ji}^c = 1$ if the SL neurone i and the IL neurone j belong to the same minicolumn, $w_{ji}^c = 0$ otherwise. The membrane potential $V_j(t)$ is computed by Eq. 2, and a sigmoidal transfer function f is employed to calculate $r_j(t)$. The parameters of the transfer function change online to adapt the electroresponsiveness properties of IL neurones j to the strength of their inputs [40].

2.3 Unsupervised Growing Network Scheme for Topological Map Learning

The topological representation is built incrementally as the animat explores the environment. At each location visited by the agent at time t the cortical network is updated if-and-only-if the infragranular layers of all existing minicolumns remain silent, i.e. $\sum_j \mathcal{H}(r_j(t) - \rho) = 0$, where j indexes all the IL neurones, $\mathcal{H}$ is the Heaviside function (i.e., $\mathcal{H}(x) = 1$ if $x \geq 0$, $\mathcal{H}(x) = 0$ otherwise), and $\rho = 0.1$ (see [38] for a similar algorithmic implementation of novelty detection in the hippocampal activity space). If at time t the novelty condition holds, a new group of minicolumns (i.e., a new column c) is recruited to become selective to the new place. Then, all the simultaneously active place cells $h \in HP$ are connected to the new IL units $j \in c$. Weights w_{jh}^h are initialised according to

$$w_{jh}^h = \mathcal{H}(r_h - \rho) \cdot r_h \ . \quad (5)$$

For $t' > t$, the synaptic strength of these connections is changed by unsupervised Hebbian learning combined to a winner-take-all scheme. Let c be the column selective for the position visited by the animat at time t', i.e. let all the $j \in c$ be the most active IL units of the network at time t'. Then:

$$\Delta w_{jh}^h = \alpha \cdot r_h \cdot (r_j - w_{jh}^h) \ . \quad (6)$$

with $\alpha = 0.005$. Whenever a state transition occurs, the collateral projections $\boldsymbol{w}^l$ and $\boldsymbol{w}^u$ are updated to relate the minicolumn activity to the state-action space $S \times A$. For instance, let columns c and c' denote the animat position

before and after a state transition, respectively (Fig. 1A). A minicolumn $\theta \in c$ becomes selective for the locomotion orientation taken by the animat to perform the transition. A new set of projections $w^l_{j'j}$ is then established from the IL unit $j \in \theta$ of column c to all the IL units j' of the column c'. In addition, at the supragranular level, a new set of connections $w^u_{ii'}$ is learnt to connect all the SL units of column c', i.e. $i' \in c'$, to the SL unit i of the minicolumn $\theta \in c$. The strengths of the lateral projections are initialised as:

$$w^l_{j'j} = w^u_{ii'} = \beta_{LTP} \quad \forall i', j' \in c' \; . \tag{7}$$

with $\beta_{LTP} = 0.9$. Finally, in order to adapt the topological representation online, a synaptic potentiation-depression mechanism can modify the lateral projections $\boldsymbol{w}^l$ and $\boldsymbol{w}^u$. For example, if a new obstacle prevents the animat from achieving a previously learnt transition from column c to c' (i.e., if the activation of the IL unit $j \in \theta \in c$ is not followed in the time by the activation of all IL units $j' \in c'$), then a depression of the $w^l_{j'j}$ synaptic efficacy occurs:

$$\Delta w^l_{j'j} = -\beta_{LTD} \cdot w^l_{j'j} \quad \forall j' \in c' \; . \tag{8}$$

where $\beta_{LTD} = 0.5$. The projections $w^u_{ii'}$ are updated in a similar manner. A compensatory potentiation mechanism reinforces both $\boldsymbol{w}^l$ and $\boldsymbol{w}^u$ connections whenever a previously experienced transition is performed successfully:

$$\Delta w^l_{j'j} = \beta_{LTP} - w^l_{j'j} \quad \forall j' \in c' \; . \tag{9}$$

$w^u_{ii'}$ are updated similarly. Notice that $\boldsymbol{w}^l, \boldsymbol{w}^u \in [0, \beta_{LTP}]$.

2.4 Action Planning

The model presented here aims at developing a high-level controller determining the spatial behaviour based on action planning. Yet, a low-level reactive module subserves the obstacle-avoidance behaviour. Whenever the proximity sensors detect an obstacle, the reactive module takes control and prevents collisions. Also, the simulated animal behaves in order to either follow planned pathways (i.e., exploitation) or improve the topological map (i.e., exploration). This exploitation-exploration tradeoff is governed by an ϵ-greedy selection mechanism, with $\epsilon \in [0, 1]$ decreasing exponentially over time [38].

Fig. 1B shows an example of activation-diffusion process mediated by the columnar network. During trajectory planning, the SL neurones of the column corresponding to the goal location s_g are activated via a motivational signal r_m (Eq. 3). Then, the SL activity is back-propagated through the network by means of the lateral projections $\boldsymbol{w}^u$ (Fig. 1B, top). During planning, the responsiveness of IL neurones (Eq. 4) is decreased to detect coincident inputs. In particular, the occurrence of the SL input r_i is a necessary condition for a IL neurone j to fire. In the presence of the SL input r_i, either the hippocampal signal r_h or the inter-column signal r'_j is sufficient to activate the IL unit j. When the back-propagated

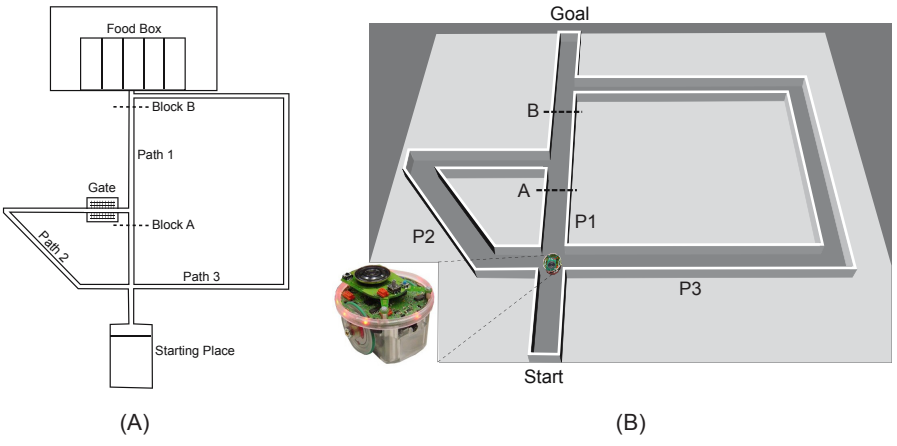

Fig. 2. (A) Tolman & Honzik's maze (adapted from [36]). The gate near the second intersection prevented rats from going from right to left. **(B)** The simulated maze and robot. The dimensions of the simulated maze were taken so as to maintain the proportions of the Tolman & Honzik's setup. Bottom-left inset: the real *e-puck* mobile robot has a diameter of 70 mm and is 55 mm tall.

goal signal reaches the minicolumns selective for the current position s this coincidence event occurs, which triggers the forward propagation of a goal-directed path signal through the projections $\boldsymbol{w}^l$ (Fig. 1B, bottom). Goal-directed trajectories are generated by reading out the successive activations of IL neurones. Action selection calls upon a competition between the minicolumns encoding the $(s, a_{1\ldots N}) \in S \times A$ pairs, where s is the current location, and $a_{1\ldots N}$ are the transitions from s to adjacent positions s'. For sake of robustness, competition occurs over a 10-timestep cycle. Notice that each SL synaptic relay attenuates the goal signal by a factor $w^u_{ii'}$ (Eq. 3). Thus, the smaller the number of synaptic relays, the stronger the goal signal received by the SL neurone corresponding to the current location s. As a consequence, because the model column receptive fields are distributed rather uniformly over the environment, the intensity of the goal signal at a given location s is correlated to the distance between s and the target position s_g.

2.5 Behavioural Task and Simulated Agent

In order to validate our navigation planning system, we chose the classical experimental task proposed by Tolman & Honzik [36]. The main objective of this behavioural protocol was to demonstrate that rodents undergoing a navigation test were able to show some "insights", e.g. to predict the outcome of alternative trajectories leading to a goal location in the presence of blocked pathways. The original Tolman & Honzik's maze is shown in Fig. 2A. It consisted of three narrow alleys of different lengths (Paths 1, 2, and 3) guiding the animals from a starting position (bottom) to a feeder location (top).

Fig. 2B shows a simulated version of the Tolman & Honzik's apparatus, and the simulated robot[1]. We emulated the experimental protocol designed by Tolman & Honzik to assess the animats' navigation performance. The overall protocol consisted of a training period followed by a probe test. Both training and probe trials were stopped when the animat had found the goal.

Training period: it lasted 14 days with 12 trials per day. The animats could explore the maze and learn their navigation policy.

- During Day 1, a series of 3 *forced runs* was carried out, in which additional doors were used to force the animats to go successively through P1, P2, and P3. Then, during the remaining 9 runs, all additional doors were removed, and the subjects could explore the maze freely. At the end of the first training day, a preference for P1 was expected to be already developed [36].
- From Day 2 to 14, a block was introduced at place A (Fig. 2B) to require a choice between P2 and P3. In fact, additional doors were used to close the entrances to P2 and P3 to force the animats to go first to the Block A. Then, doors were removed, and the subjects were forced to decide between P2 and P3 on their way back to the first intersection. Each day, there were 10 "Block at A" runs that were mixed with 2 non-successive free runs to maintain the preference for P1.

Probe test period: it lasted 1 day (Day 15), and it involved 7 runs with a block at position B to interrupt the common section (Fig. 2B). The animats were forced to decide between P2 and P3 when returning to the first intersection point.

For these experiments, Tolman & Honzik used 10 rats with no previous training. In our simulations, we used a population of 100 animats, and we assessed the statistical significance of the results by means of an ANOVA analysis (the significant threshold was set at 10^{-2}, i.e. $p < 0.01$ was considered significant).

2.6 Theoretical Analysis

A series of analyses was done to characterise the neural activities subserving the behavioural responses of the system. We recall that one of the aims of the cortical column model was to build a spatial code less redundant than the hippocampal place (HP) field representation. Yet, it is relevant to show that the spatial properties (e.g., spatial information content) of the neural responses were preserved in the cortical network.

The set of stimuli S consisted of the places visited by the animat. For the analyses, the continuous two-dimensional input space was discretized, with each location $s \in S$ defined as a 5 x 5 cm square region of the environment. The size of the receptive field of a neurone j was taken as $2{\cdot}\sigma_S(j)$, with $\sigma_S(j)$ denoting the standard deviation around the mean of the response tuning curve.

A *spatial density* measure was used to assess the level of redundancy of a neural spatial code, i.e. the average number of units necessary to encode a place:

$$D = \langle \sum_{j \in J} \mathcal{H}(r_j(s) - \sigma_J(s)) \rangle_{s \in S} \, . \tag{10}$$

[1] The model was implemented by means of the Webots© robotics simulation software.

with $r_j(s)$ being the response of the neurone j when the animat is visiting the location $s \in S$, and $\sigma_J(s)$ representing the standard deviation of the population activity distribution for a given stimulus s.

Another measure was used to analyse the neural responses, the *kurtosis* function. This measure is defined as the normalised fourth central moment of a probability distribution, and estimates its degree of peakedness. If applied to a neural response distribution, the kurtosis can be used to measure its degree of sparseness across both population and time [41]. We employed an average population kurtosis measure $\bar{k}_1 = \langle k_1(s) \rangle_{s \in S}$ to estimate how many neurones j of a population J were, on average, responding to a given stimulus s simultaneously. The kurtosis $k_1(s)$ was taken as:

$$k_1(s) = \langle \, [r_j(s) - \bar{r}_J(s))/\sigma_J(s)]^4 \, \rangle_{j \in J} \, . \tag{11}$$

with $\bar{r}_J(s) = \langle r_j(s) \rangle_{j \in J}$. Similarly, an average *lifetime* kurtosis $\bar{k}_2 = \langle k_2(j) \rangle_{j \in J}$ was employed to assess how rarely a neurone j responded across time. The $k_2(j)$ function was given by:

$$k_2(j) = \langle \, [r_j(s) - \bar{r}_j)/\sigma_j]^4 \, \rangle_{s \in S} \tag{12}$$

with $\bar{r}_j = \langle r_j(s) \rangle_{s \in S}$, and σ_j being the standard deviation of the cell activity r_j.

Finally, we used an information theoretic analysis [42] to characterise the neural codes of our cortical and hippocampal populations. The mutual information $MI(S;R)$ between neural responses R and spatial locations S was computed:

$$MI(S;R) = \sum_{s \in S} \sum_{r \in R} P(r,s) \log_2 \left(\frac{P(r,s)}{P(r)P(s)} \right) \tag{13}$$

where $r \in R$ indicated firing rates, $P(r,s)$ the joint probability of having the animat visiting a region $s \in S$ while recording a response r, $P(s)$ the a priori probability computed as the ratio between time spent at place s and the total time, and $P(r) = \sum_{s \in S} P(r,s)$ the probability of observing a neural response r. The continuous output space of a neurone, i.e. $R = [0,1]$, was discretized via a binning procedure (bin-width equal to 0.1). The $MI(S;R)$ measure allowed us to quantify the spatial information content of a neural code, i.e. how much could be learnt about the animat's position s by observing the neural responses r.

3 Results

3.1 Spatial Behaviour

Day 1. During the first 12 training trials, the animats learnt the topology of the maze and planned their navigation trajectory in the absence of both block A and B. Similar to Tolman & Honzik's findings, our results show that the model learnt to select the shortest goal-directed pathway P1 significantly more frequently than the alternative trajectories P2, P3 (ANOVA, $F_{2,297} = 168.249$,

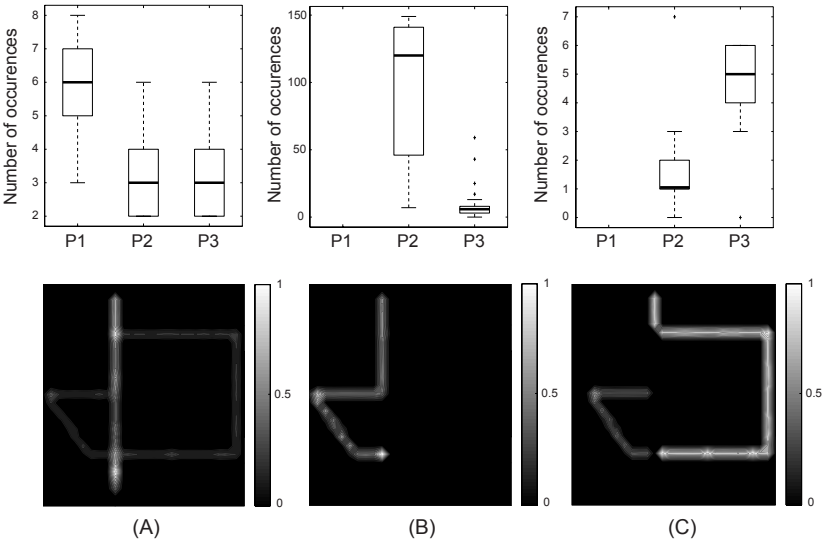

Fig. 3. Behavioural results. **Top row:** mean number of transits through P1, P2, and P3 (averaged over 100 animats). **Bottom row:** occupancy grid maps. **(A)** During the first 12 training trials (day 1) the simulated animals developed a significant preference for P1 (no significant difference was observed between P2 and P3). **(B)** During the following 156 training trials (days 2-14, in the presence of block A, Fig. 2B) P2 was selected significantly more frequently than P3. **(C)** During the last 7 trials (day 15, test phase), the block A was removed whereas the block B was introduced. The animats exhibited a significant preference for P3 compared to P2.

$p < 0.0001$). The quantitative and qualitative analyses reported on Fig. 3A describe the path selection performance averaged over 100 animats.

Days 2-14. During this training phase (consisting of 156 trials), a block was introduced at location A, which forced the animats to update their topological maps dynamically, and to plan a *detour* to the goal. The results reported by Tolman & Honzik provided strong evidence for a preference for the shortest *detour* path P2. Consistently, in our simulations (Fig. 3B) we observed a significantly larger number of transits through P2 compared to P3 (ANOVA, $F_{1,198} = 383.068$ $p < 0.0001$), P1 being ignored in this analysis (similar to Tolman & Honzik's analysis) because blocked.

Day 15. Seven probe trials were performed during the 15th day of the simulated protocol, by removing the block A and adding a new block at location B. This manipulation aimed at testing the "insight" working hypothesis: after a first run through the shortest path P1 and after having encountered the unexpected block B, will animats try P2 (wrong behaviour) or will they go directly through P3 (correct behaviour)? According to Tolman & Honzik's results, rats behaved as predicted by the insight hypothesis, i.e. they tended to select the longer but

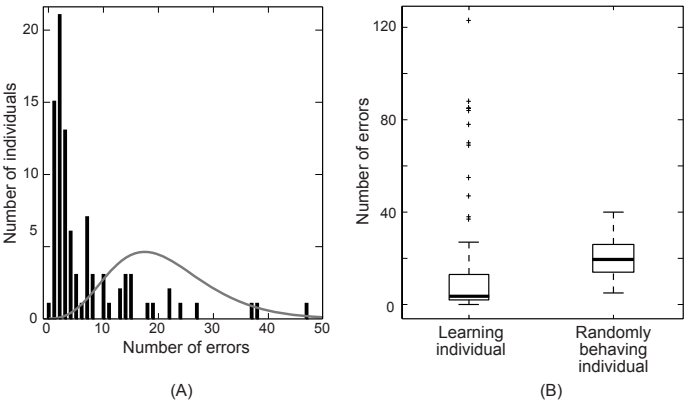

Fig. 4. Comparison between a learning and a randomly behaving agent. **(A)** Error distribution of learning (black histogram) versus random (grey line) animats. **(B)** Mean number of errors made by the model and by a randomly behaving agent.

effective P3. The authors concluded that rats were able to inhibit the previously learnt policy (i.e., the "habit behaviour" consisting of selecting P2 after a failure of P1 during the 156 previous training trials). Our probe test simulation results are shown in Fig. 3C. Similar to rats, the animats exhibited a significant preference for P3 compared to P2 (ANOVA, $F_{1,198} = 130.15$, $p < 0.0001$). Finally, in order to further assess the mean performance of the system during the probe trials, we compared the action selection policy of learning animats with that of randomly behaving (theoretical) animats. Fig. 4A provides the results of this comparison by showing the error distribution over the population of learning agents (black histogram) and randomly behaving agents (grey curve). The number of errors per individual are displayed in the boxplot of Fig. 4B. These findings indicate a significantly better performance of learning animats compared to random agents (ANOVA, $F_{1,196} = 7.4432$, $p < 0.01$).

3.2 Analysis of Neural Activities

Fig. 5A contrasts the mean spatial density (Eq. 10) of the HP receptive fields with that of cortical column receptive fields. It is shown that, compared to the upstream hippocampal space code, the cortical column model reduced the redundancy of the learnt spatial code significantly (ANOVA, $F_{1,316} = 739.2$, $p < 0.0001$). Fig. 5B shows the probability distribution representing the number of active column units (solid curve) and active HP cells (dashed line) per spatial location $s \in S$. As shown by the inset boxplots, the distribution kurtosis was significantly higher for column units than for HP cells (ANOVA, $F_{1,198} = 6057$, $p < 0.0001$). To further investigate this property, we assessed the average population kurtosis $\bar{k}_1$ (Eq. 11) of both columnar and HP cell activities (Fig. 5C). Again, the columnar population activity exhibited a significantly higher kurtosis

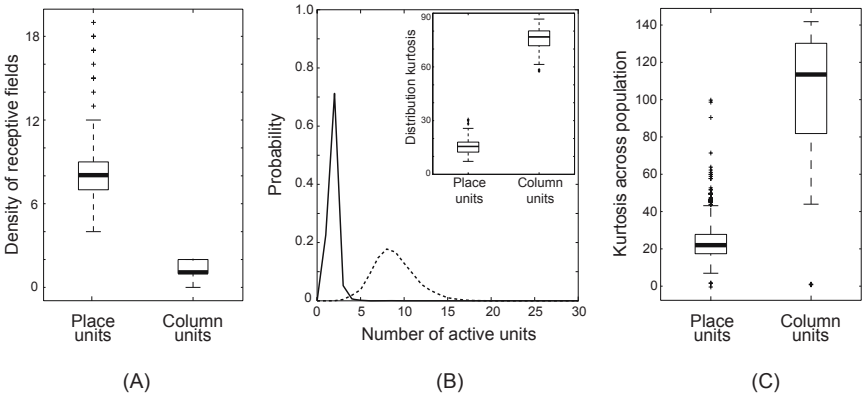

Fig. 5. (A) Spatial density of the receptive fields of HP cells and cortical column units. **(B)** Probability distribution of the number of active column units (solid line) and active HP cells (dashed line) per spatial location $s \in S$. Inset boxplots: kurtosis measures for the two distributions. **(C)** Population kurtosis of columnar and hippocampal assemblies.

than the HP cell activity (ANOVA, $F_{1,3128} = 14901$, $p < 0.0001$). These results suggest that, in the model, the cortical column network was able to provide a sparser state-space population coding than HP units.

In a second series of analyses, we focused on the activity of single cells, and we compared the average lifetime kurtosis $\bar{k}_2$ (Eq. 12) of cortical and HP units. As reported on Fig. 6A, we found that the kurtosis across time did not differ significantly between cortical and HP units (ANOVA, $F_{1,2356} = 2.2699$, $p < 0.13$). This result suggests that, on average, single cortical and HP units tended to respond to a comparable number of stimuli (i.e., spatial locations) over their lifetimes. Along the same line, we recorded the receptive fields of the two types of units. Figs. 6B,C display some samples of place fields of cortical and HP cells, respectively. As expected, we found a statistical anticorrelation between the lifetime kurtosis and the size of the receptive fields. The example of Fig. 6D shows that, for a randomly chosen animat performing the whole experimental protocol (15 days), the size of hippocampal place fields was highly anticorrelated to the HP cells' lifetime kurtosis (correlation coefficient $= -0.94$). These results add to those depicted in Fig. 5 in that the increase of sparseness at the level of the cortical population (compared to HP cells) was not merely due to an enlargement of the receptive fields (or, equivalently, to a decrease of the lifetime stimulus-dependent activity).

Despite their less redundant code, were cortical columns able to provide a representation comparable to that of HP cells in terms of spatial information content? The results of our information theoretic analysis (Eq. 13) suggest that this was indeed the case. Fig. 6E shows that, for a randomly chosen animat,

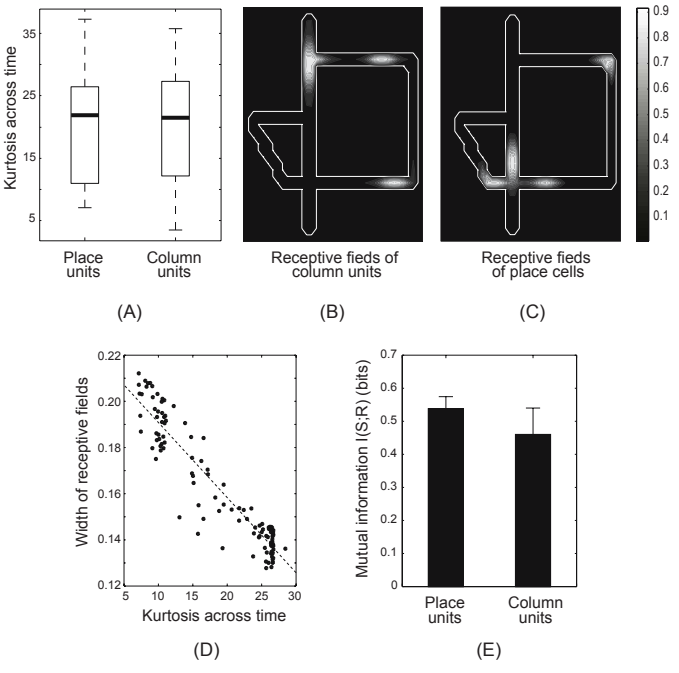

Fig. 6. (A) Lifetime kurtosis for column and HP units. **(B, C)** Samples of receptive fields of three column units and four HP cells. **(D)** Correlation between the lifetime kurtosis and the size of receptive fields. **(E)** Mutual information $MI(S;R)$ between the set of spatial locations S and the activity R for both cortical and HP units.

the average amount of spatial information conveyed by cortical units was not significantly lower than that of HP cells (ANOVA, $F_{1,140} = 0.8034$, $P < 0.3716$).

4 Discussion

We presented a navigation model that builds a topological map of the environment incrementally, and uses it to plan a course of actions leading to a goal location. The model was employed to solve the classical Tolman & Honzik's task [36]. As aforementioned, other models have been proposed to solve goal-directed navigation tasks. They are mainly based on the properties of hippocampal (e.g., [43]), and prefrontal cortex (e.g., [31]) neural assemblies. However, most of these models do not perform action *planning* as defined in this paper (Sec. 1). Samsonovich and Ascoli [43] implement a local path finding mechanism to select the most suitable orientation leading to the goal. Similarly, Hasselmo's model [31] does not plan a sequence of actions from the current location to the goal but it rather infers the first local action to be taken, based upon a back-propagated

goal signal. Yet, these two models rely on discretized state spaces (with predefined grid units coding for places), whereas our model uses a place field population providing a continuous representation of the environment [38]. Also, our model learns topological maps coding for the state-action space simultaneously. In the model by Samsonovich and Ascoli [43] no topological information is represented, but only a distance measure between each visited place and a set of potential goals. Likewise, in Hasselmo's model states and actions are not jointly represented, which generates a route-based rather than a map-based navigation system [1].

We adopted a three-fold working hypothesis according to which *(i)* the hippocampus would play a prominent role in encoding spatial information; *(ii)* higher-level cortical areas, particularly the PFC, would mediate multidimensional contextual representations (e.g., coding for motivation-dependent memories and action cost/risk constraints) grounded on the hippocampal spatial code; *(iii)* neocortical representations would facilitate the temporal linking of multiple contexts, and the sequential organisation (e.g., planning) of behavioural responses. The preliminary version of the model presented here enabled us to focus on some basic computational properties, such as the ability of the columnar organisation to learn a compact topological representation, and the efficiency of the activation-diffusion planning mechanism. Further efforts will be put to integrate multiple sources of information. For example, the animat should be able to learn maps that encode reward (subjective) values, and action-cost constraints. Also, these maps should be suitable to represent multiple spatio-temporal scales to overcome the intrinsic limitation of the activation-diffusion mechanism in large scale environments. Additionally, these multiscale maps should allow the model to infer high-level shortcuts to bypass low-level environmental constraints.

The neurocomputational approach presented here aims at generating cross-disciplinary insights that may help to systematically explore potential connections between findings on the neuronal level (e.g., single-cell discharge patterns), and observations on the behavioural level (e.g., spatial navigation). Mathematical representations permit to describe both the space and time components characterising the couplings between neurobiological processes. Models can help to scale up from single cell properties to the dynamics of neural populations, and generate novel hypotheses about their interactions to produce complex behaviour.

Acknowledgments. Granted by the EC Project ICEA (Integrating Cognition, Emotion and Autonomy), IST-027819-IP.

References

1. Arleo, A., Rondi-Reig, L.: Multimodal sensory integration and concurrent navigation strategies for spatial cognition in real and artificial organisms. J. Integr. Neurosci. 6(3), 327–366 (2007)
2. Dollé, L., Khamassi, M., Girard, B., Guillot, A., Chavarriaga, R.: Analyzing interactions between navigation strategies using a computational model of action selection. In: Freksa, C., et al. (eds.) SC 2008. LNCS (LNAI), vol. 5248, pp. 71–86. Springer, Heidelberg (2008)

3. O'Keefe, J., Nadel, L.: The Hippocampus as a Cognitive Map. Oxford University Press, Oxford (1978)
4. Hafting, T., Fyhn, M., Molden, S., Moser, M.B., Moser, E.I.: Microstructure of a spatial map in the entorhinal cortex. Nature 436(7052), 801–806 (2005)
5. Wiener, S.I., Taube, J.S.: Head Direction Cells and the Neural Mechansims of Spatial Orientation. MIT Press, Cambridge (2005)
6. Poucet, B., Lenck-Santini, P.P., Hok, V., Save, E., Banquet, J.P., Gaussier, P., Muller, R.U.: Spatial navigation and hippocampal place cell firing: the problem of goal encoding. Rev. Neurosci. 15(2), 89–107 (2004)
7. Amaral, D.G., Witter, M.P.: The three-dimensional organization of the hippocampal formation: a review of anatomical data. Neurosci. 31(3), 571–591 (1989)
8. Wilson, M.A., McNaughton, B.L.: Dynamics of the hippocampal ensemble code for space. Science 261, 1055–1058 (1993)
9. Nitz, D.A.: Tracking route progression in the posterior parietal cortex. Neuron. 49(5), 747–756 (2006)
10. Hok, V., Save, E., Lenck-Santini, P.P., Poucet, B.: Coding for spatial goals in the prelimbic/infralimbic area of the rat frontal cortex. Proc. Natl. Acad. Sci. USA. 102(12), 4602–4607 (2005)
11. Knierim, J.J.: Neural representations of location outside the hippocampus. Learn. Mem. 13(4), 405–415 (2006)
12. Granon, S., Poucet, B.: Medial prefrontal lesions in the rat and spatial navigation: evidence for impaired planning. Behav. Neurosci. 109(3), 474–484 (1995)
13. Jay, T.M., Witter, M.P.: Distribution of hippocampal ca1 and subicular efferents in the prefrontal cortex of the rat studied by means of anterograde transport of phaseolus vulgaris-leucoagglutinin. J. Comp. Neurol. 313(4), 574–586 (1991)
14. Kita, H., Kitai, S.T.: Amygdaloid projections to the frontal cortex and the striatum in the rat. J. Comp. Neurol. 298(1), 40–49 (1990)
15. Thierry, A.M., Blanc, G., Sobel, A., Stinus, L., Golwinski, J.: Dopaminergic terminals in the rat cortex. Science 182(4111), 499–501 (1973)
16. Uylings, H.B.M., Groenewegen, H.J., Kolb, B.: Do rats have a prefrontal cortex? Behav. Brain. Res. 146(1-2), 3–17 (2003)
17. Aggleton, J.: The amygdala: neurobiological aspects of emotion, memory, and mental dysfunction. Wiley-Liss, New York (1992)
18. Schultz, W.: Predictive reward signal of dopamine neurons. J. Neurophysiol. 80(1), 1–27 (1998)
19. Jung, M.W., Qin, Y., McNaughton, B.L., Barnes, C.A.: Firing characteristics of deep layer neurons in prefrontal cortex in rats performing spatial working memory tasks. Cereb. Cortex 8(5), 437–450 (1998)
20. Otani, S.: Prefrontal cortex function, quasi-physiological stimuli, and synaptic plasticity. J. Physiol. Paris 97(4-6), 423–430 (2003)
21. Fuster, J.M.: The prefrontal cortex–an update: time is of the essence. Neuron. 30(2), 319–333 (2001)
22. Cohen, J.D., Braver, T.S., Brown, J.W.: Computational perspectives on dopamine function in prefrontal cortex. Curr. Opin. Neurobiol. 12(2), 223–229 (2002)
23. Mountcastle, V.B.: Modality and topographic properties of single neurons of cat's somatic sensory cortex. J. Neurophysiol. 20(4), 408–434 (1957)
24. Mountcastle, V.B.: The columnar organization of the neocortex. Brain 120, 701–722 (1997)
25. Buxhoeveden, D.P., Casanova, M.F.: The minicolumn hypothesis in neuroscience. Brain 125(5), 935–951 (2002)

26. Hampson, S.: Connectionist problem solving. In: The Handbook of Brain Theory and Neural Networks, pp. 756–760. The MIT Press, Cambridge (1998)
27. Meyer, J.A., Filliat, D.: Map-based navigation in mobile robots - ii. a review of map-learning and path-planing strategies. J. Cogn. Syst. Res. 4(4), 283–317 (2003)
28. Burnod, Y.: An adaptative neural network: the cerebral cortex. Masson (1989)
29. Bieszczad, A.: Neurosolver: a step toward a neuromorphic general problem solver. Proc. World. Congr. Comput. Intell. WCCI94 3, 1313–1318 (1994)
30. Frezza-Buet, H., Alexandre, F.: Modeling prefrontal functions for robot navigation. IEEE Int. Jt. Conf. Neural. Netw. 1, 252–257 (1999)
31. Hasselmo, M.E.: A model of prefrontal cortical mechanisms for goal-directed behavior. J. Cogn. Neurosci. 17(7), 1115–1129 (2005)
32. Schmajuk, N.A., Thieme, A.D.: Purposive behavior and cognitive mapping: a neural network model. Biol. Cybern. 67(2), 165–174 (1992)
33. Dehaene, S., Changeux, J.P.: A hierarchical neuronal network for planning behavior. Proc. Natl. Acad. Sci. USA. 94(24), 13293–13298 (1997)
34. Banquet, J.P., Gaussier, P., Quoy, M., Revel, A., Burnod, Y.: A hierarchy of associations in hippocampo-cortical systems: cognitive maps and navigation strategies. Neural Comput. 17, 1339–1384 (2005)
35. Fleuret, F., Brunet, E.: Dea: an architecture for goal planning and classification. Neural Comput 12(9), 1987–2008 (2000)
36. Tolman, E.C., Honzik, C.H.: "Insight" in rats. Univ. Calif. Publ. Psychol. 4(14), 215–232 (1930)
37. Arleo, A., Gerstner, W.: Spatial orientation in navigating agents: modeling head-direction cells. Neurocomput. 38(40), 1059–1065 (2001)
38. Arleo, A., Smeraldi, F., Gerstner, W.: Cognitive navigation based on nonuniform gabor space sampling, unsupervised growing networks, and reinforcement learning. IEEE Trans. Neural. Netw. 15(3), 639–651 (2004)
39. Rao, S.G., Williams, G.V., Goldman-Rakic, P.S.: Isodirectional tuning of adjacent interneurons and pyramidal cells during working memory: evidence for microcolumnar organization in pfc. J. Neurophysiol. 81(4), 1903–1916 (1999)
40. Triesch, J.: Synergies between intrinsic and synaptic plasticity mechanisms. Neural Comput. 19(4), 885–909 (2007)
41. Willmore, B., Tolhurst, D.J.: Characterizing the sparseness of neural codes. Netw. Comput. Neural Syst. 12(3), 255–270 (2001)
42. Bialek, W., Rieke, F., de Ruyter van Steveninck, R., Warland, D.: Reading a neural code. Science 252(5014), 1854–1857 (1991)
43. Samsonovich, A., Ascoli, G.: A simple neural network model of the hippocampus suggesting its pathfinding role in episodic memory retrieval. Learn. Mem. 12, 193–208 (2005)