SPATIAL LEARNING AND NAVIGATION IN THE RAT: A BIOMIMETIC MODEL

THÈSE N^{O} 3421 (2005)

PRÉSENTÉE À LA FACULTÉ INFORMATIQUE ET COMMUNICATIONS

Institut des systèmes informatiques et multimédias

SECTION D'INFORMATIQUE

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Ricardo CHAVARRIAGA LOZANO

ingénieur électronicien de la Pontificia Universidad Javeriana, Colombie
de nationalité colombienne

acceptée sur proposition du jury:

Prof. W. Gerstner, directeur de thèse
Prof. A. Billard, rapporteur
Prof. F. Schenk, rapporteur
Prof. A. Treves, rapporteur

Lausanne, EPFL
2006

Spatial learning and navigation in the rat: A Biomimetic model

Abstract

Animals behave in different ways depending on the specific task they are required to solve. In certain cases, if a cue marks the goal location, they can rely on simple stimulus-response associations. In contrast, other tasks require the animal to be endowed with a representation of space. Such a representation (i.e. *cognitive map*) allows the animal to locate itself within a known environment and perform complex target-directed behaviour. In order to efficiently perform, the animal not only should be able to exhibit these types of behaviour, but it should be able to select *which* behaviour is the most appropriate at any given task conditions.

Neurophysiological and behavioural experiments provide important information on how such processes may take place in the rodent's brain. Specifically, place- and orientation sensitive cells in the rat *Hippocampus* have been interpreted as a neural substrate for spatial abilities related to the theory of the cognitive map proposed in the late 1940s by Tolman. Moreover, recent dissociation experiments using selectively located lesions, as well as pharmacological studies have shown that different brain regions may be involved in different types of behaviour. Accordingly, one memory system involving the hippocampus and the ventral striatum would be responsible for cognitive navigation, while navigation based on stimulus-response associations would be mediated by the dorsolateral striatum.

Based on these studies, the aim of this work is to develop a neural network model of the spatial abilities of the rat. The model, based on functional properties and anatomical inter-connections of the brain areas involved in spatial learning should be able to establish a distributed representation of space composed of place-sensitive units. Such a representation takes into account both internal and external sensory information, and the model reproduces physiological properties of place cells such as changes in their directional dependence.

Moreover, the spatial representation may be used to perform cognitive navigation. Modelled place cells drive an extra-hippocampal population of action-coding cells, allowing the establishment of place-response associations. These associations encoded in synaptic connections between place- and action-cells are modified by means of reinforcement learning. In a similar way, simple sensory input can be used to establish stimulus-response associations. These associations are encoded in a different set of action cells which corresponds to a different neural substrate encoding for non-cognitive navigation strategies (i.e. *taxon* or *praxic*). Both cognitive and non-cognitive navigation strategies compete for action control to determine the actual behaviour of the agent.

Tests of the performance of the model show that it is able to establish a representation of space, and modelled place cells reproduce some physiological properties of their biological counterparts. Furthermore, the model reproduces goal-based behaviour based on both cognitive and non-cognitive strategies as well as behaviour in conflicting situations reported in experimental studies in animals.

Résumé

Les animaux sont capables d'adapter leur comportement selon la tâche spécifique qu'ils doivent accomplir. Face à des tâches de navigation (comme trouver de la nourriture, retourner au nid, etc.) ils peuvent utiliser différentes stratégies selon les informations disponibles. Lorsque le but est visible, l'animal peut utiliser une stratégie basée sur des associations du type stimulus-réponse (S-R). Dans d'autres situations, s'il n'existe aucune indication sensorielle de la position du but, l'animal doit utiliser une représentation spatiale. Une telle représentation (*carte cognitive*) permet à l'animal de se localiser dans l'environnement et d'accomplir des tâches spatiales plus complexes. L'animal doit non seulement être capable d'utiliser ces deux types de stratégies, mais il doit aussi pouvoir choisir laquelle est la plus appropriée en fonction du contexte environnemental.

Les données issues de l'expérimentation neurobiologique chez le rat fournissent des informations importantes sur les structures du cerveau qui sont impliquées dans ces différentes stratégies. Les *cellules de lieu* dans l'hippocampe du rat semblent être la base neurologique de la carte cognitive comme proposée par Tolman dans les années 1940. De même, les données expérimentales suggèrent que si l'hippocampe est impliqué dans des tâches de navigation cognitive, il ne l'est pas dans les tâches de navigation basées sur des associations S-R. La partie dorsolatérale du striatum semble être impliquée dans ce type de navigation.

L'objectif de cette thèse est de développer un modèle neuronal de l'apprentissage spatial chez le rat. Ce modèle, basé sur des données expérimentales (i.e. propriétés fonctionelles et interconnectivité des aires du cerveau impliquées dans ce type d'apprentissage) doit être capable de construire une représentation spatiale composée de cellules de lieu. L'activation de ces cellules dépendant d'informations allothétiques ainsi que d'informations idiothétiques.

Cette représentation spatiale est la base d'un système de navigation cognitive, construit à partir d'associations entre la position spatiale et les différentes actions. Ces associations sont mis à jour en utilisant l'apprentissage par renforcement. D'une manière similaire, les associations entre stimulus sensoriels et actions sont apprises pour développer des stratégies de navigation non cognitives (S-R). Un processus de sélection entre les différents types de stratégies détermine l'action qui est finalement entreprise.

Le modèle a été testé en simulation en utilisant des tâches semblables à celles utilisées dans l'expérimentation animale. Ces tests démontrent que le modèle est capable de construire une représentation spatiale et que les cellules de lieu générées ont des propriétés physiologiques similaires à celles des cellules de lieu trouvées chez le rat. De plus, le modèle reproduit des comportements issus de stratégies de navigation cognitives et non-cognitives. Enfin, il est capable de choisir quel type de stratégie doit être suivie selon les conditions environnementales.

Acknowledgements

I would like to thank the many people who contributed to the achievement of this thesis[1]. First of all, to express my gratitude to my PhD supervisor Prof. Wulfram Gerstner, for his advise and the opportunity he gave me to get introduced to a subject entirely new to me: the neurobiology of animal behaviour. I would also like to thank the members of the thesis committee: Prof. Roger Hersch, Prof. Alessandro Treves, Prof. Françoise Schenk and Prof. Aude Billard.

I am particularly grateful to Dr. Angelo Arleo whose work constitutes the basis in which this project is based on. In the same way, I want to recognise the valuable contributions made by Dr. Thomas Strösslin and Denis Sheynikhovich with whom I had the pleasure to work in the last years and enjoy fruitful discussions related to the subject of this project.

I want to thank my coworkers at the Laboratory of Computational Neuroscience. Their friendship and support in both academic and non-academic issues made of my experience at EPFL a really pleasant one. I am specially indebted to Magnus Richardson, Silvio Borer, José del R. Millán, Eric Sauser, Renaud Jolivet and Julien Mayor. I also want to thank our secretary Brigitte Ramuz for her excellent work, and equally good cakes.

I would like to thank my family and friends from Colombia, who have always supported me and continue to be right by my side even though we live miles away. I want to thank as well, all the people that became my family in Switzerland and make my life in Lausanne a very pleasant experience. Thanks to Carlos Andrés, for his help; To Carolina, my friend and supporter since ever; and, last but not least, thanks to Iveth for her love and for all the happy moments.

[1]This work was supported by the Swiss National Science Foundation under grants no.20-61622.00 and 200020-100265/1.

Contents

Chapter 1

Introduction

1.1 Motivation

Animals continuously interact with their surrounding environment. Perceptual stimuli (e.g. odour, visual cues, etc.) and previous experience modify their behaviour. To locate a food reservoir, return to its nest or find a shortcut the animal needs to be endowed with the ability to locate itself, and interesting locations in the environment.

In particular cases, animals are able to learn how to reach goals which are not specifically signalled by a cue (or sequence of them). The solution of these tasks seems to rely on the use of an internal representation which allows the animal to locate itself within the environment. Such a representation is referred to as *cognitive map* (Tolman, 1948). Behavioural experiments suggest that animals use such a representation (Tolman, 1948; Morris, 1981).

Moreover, place-sensitive cells found in the rat brain have been interpreted as being responsible for coding a representation of space (O'Keefe and Dostrovsky, 1971; O'Keefe and Nadel, 1978; McNaughton et al., 1983; Quirk et al., 1992; Hafting et al., 2005). The firing of such cells, termed *place cells*, is highly correlated to the animals location within the environment. These cells have been found principally in the *hippocampus*, and have prompted a large number of experimental studies on this structure of the rodent brain. Other cells have been found to be sensitive to the animal's head direction (i.e. *head direction cells*). They constitute an allocentric compass system that together with the spatial sensitivity of place seems to be crucial to solve cognitive navigational tasks.

As described above, the animal may use an internal representation in order to navigate toward non cued goals. Such a navigation strategy is called *locale navi-*

gation (Redish, 1999). However, on other occasions the animal can rely on simpler strategies in order to reach an interesting location. For instance, if the goal location is consistently signalled by a visible cue, landmark-guidance behaviour can be used to satisfactorily solve the task, i.e. *taxon navigation*. This type of strategy corresponds to the learning of a simple stimulus-response association, and it can be understood as a simple reactive behaviour i.e. in response to a given stimulus (i.e. light or tone) the animal learns to elicit a particular response (i.e. to orient toward the stimulus and approach it).

Animals should be able to select among the different navigation strategies in order to efficiently solve the various tasks it is faced with (in both experimental laboratories and wildlife conditions). Such a selection seems to be influenced by the characteristics of the task and the animal's experience among other factors.

Dissociation studies suggest that the two navigation strategies mentioned above are mediated by different brain areas (Packard et al., 1989; Devan and White, 1999; K. E. Ragozzino et al., 2001; Packard and Knowlton, 2002; White and McDonald, 2002; Gold, 2004). Some of these studies aim to dissociate the role of different brain areas by observing how localised lesions reduce the performance in one type of tasks, while leaving other tasks unaffected. In the context of navigation, it has been shown that rats with hippocampal lesions are impaired in tasks requiring locale navigation strategies (Morris, 1981; Redish, 1999; Pearce et al., 1998), whereas lesions in the dorsal striatum (Caudato-putamen, CPu) affect the learning of stimulus-response (S-R) associations required to navigate toward visible cues (Packard and Knowlton, 2002).

1.2 Aims of this Thesis

This work continues previous modelling efforts in our laboratory on the computational principles of spatial learning and navigation (Strösslin, 2004; Arleo, 2000). It aims to study and model the neurophysiological properties of the hippocampal place cells found in the rat hippocampus, as well as to reproduce behavioural patterns observed in navigational tasks. Specifically, this work addresses the following questions:

(i) How can an allocentric representation of the environment be built based on available sensory input.

(ii) What are the computational principles that control the selection among different navigation strategies.

Tests of the developed model in a simulated perception-action closed loop are crucial to assess experience-dependent changes in both the neurophysiological properties of the place coding and the simulated behaviour.

The use of experimental paradigms analogous to those used with real animals will be privileged in order to allow direct comparisons (at least at the qualitatively level)

between reported experimental recordings and the results of the model.

1.3 Road-map of this Dissertation

Apart from methods (Chapter 2) and conclusions (Chapter 8), the thesis is divided in two parts. The first part (Chapters 3 to 5) is devoted to the establishment of a spatial representation, while the second part (Chapters 6 and 7) details the modelling of different navigation strategies. The topics treated in each chapter are:

Chapter 3 reviews experimental data on the neural basis of spatial learning in rodents, specifically, location-sensitive cells found in the rat hippocampus. It also presents relevant attempts to model the firing properties of such cells.

Chapter 4 describes how a spatial representation based on external sensory input is constructed in our model. Local views are coded by a set of features stored in a first population of view-sensitive cells. Several local views are then combined one population downstream, yielding vision-driven place cells.

Chapter 5 details how the allothetic representation is combined with internal sensory information onto a single place representation. It first describes the path integrator module and how the allothetic estimation is used to initialise the internal representation and compensate for the accumulation of errors. Then it details the combined representation and presents results obtained in several experimental paradigms. A discussion of the entire model of spatial learning is given at the end of this chapter.

Chapter 6 discusses navigation strategies in rodents. It reviews behavioural and pharmacological experimental approaches aimed at dissociate the neural substrates responsible for different types of goal-directed behaviour. It also reviews some of the proposed models of animal navigation.

Chapter 7 proposes our navigational model, describing how different navigation strategies are implemented and compete for action-control. This chapter also presents the results obtained in simulations following experimental paradigms analogous to those used with animals.

Chapter 8 summarises the contributions and possible future developments of this work.

Chapter 2

Methods

2.1 Neural Networks

The model proposed in this thesis relies on the artificial neural networks paradigm. It is a distributed parallel system suitable to store and process complex patterns. This approach is also referred to as connectionists networks

Artificial neural networks are inspired by biological neural networks, and their computational power resides in the massive interconnection of (relatively) simple processing units (i.e. neurons). Biological neurons receive a large number of input connections on a spatial structure termed *dendrites*, and their output is transmitted through the *axon* output connection. *Synapses* are the elementary component mediating the interaction between neurons. Each synapse is characterised by its "efficacy" for transmitting information from a pre-synaptic neuron to the post-synaptic one.

Artificial neural networks (Haykin, 1994) are intended to model the adaptive properties of biological networks in order to (i) Acquire the knowledge required to solve a problem, by means of *non symbolic learning*, and (ii) Store this knowledge in the connectivity pattern (and weights) between the neurons forming the network (i.e. synaptic plasticity).

2.1.1 Rate coding model

The firing activity r_i of a neuron i correspond to the temporal average of the number of spikes emitted by that neuron during a time window Δt. A signal I corresponds to the net input to the neuron, and yields an output $r_i = f(I)$, representing the mean firing rate of the neuron i. The function f is termed *transfer function*.

Although the rate code model is unable to capture specific timing properties of the neurons activity, it allows simple calculations that make it suitable for simulations of large populations of neurons.

2.1.2 Reinforcement learning : TD learning

Reinforcement learning (RL) is a computational framework for an agent to learn how to achieve a goal through interactions with the environment. At any given moment, the agent perceives (some) properties of the environment (i.e. the *state*), and selects an *action* in order to maximise the received reward in a long term (Sutton and Barto, 1998; Doya, 2002). The reward signal, provided by the environment, tells the agent whether or not its actions were good. Although there exists a large number of different implementations of the reinforcement learning paradigm, in this section we only review the basic principle of *temporal difference* (TD) learning.

Assuming discrete state, action and time spaces, the RL problem can be stated as follows; given an observed state $s(t) = \{s_1, ..., s_n\}$, the agent takes an action $a(t) = \{a_1, ..., a_m\}$, according to its *policy*. A policy π corresponds to the mapping between states and actions, and it may be deterministic ($a = G(s)$) or stochastic ($P(a|s)$).

As a consequence of the performed action $a(t)$, the perceived environment changes to a new state $s(t+1)$, and a reward signal, $r(t+1)$, is provided. Both the new state and the reward may be given by a deterministic or stochastic process. The goal of reinforcement learning is to find an optimal policy that maximises the expected sum of future rewards.

In general, a *state action function* $V(s)$ can be defined as the cumulative future reward when the agent is in the state s, and follows a given policy,

$$V(s_t) = E^{\pi}[r_{t+1} + \gamma r_{t+2} + \gamma^2 r_{t+3} + ...] \tag{2.1}$$

where $E^{\pi}[\cdot]$ corresponds to the expected value when following the policy π. The factor $\gamma \in [0, 1]$ is the *discount rate*, which gives less importance to rewards expected far in the future.

Analogous to the state value function (Equation (2.1)), an *action function* can be defined corresponding to the expected future reward (Q-value) after taking action a in state s,

$$Q(s, a) = E^{\pi}[r_{t+1} + \gamma r_{t+2} + \gamma^2 r_{t+3} + ... \ |s_t = s, a_t = a] \tag{2.2}$$

neglecting the expectation signs, we have the approximate condition,

$$Q(s_{t-1}, a_{t-1}) = r_t + \gamma Q(s_t, a_t) \tag{2.3}$$

Deviations from the consistency condition in Equation (2.3) can be used as an error signal to learn the action value function.

$$\delta_t = r_t + \gamma Q(s_t, a_t) - Q(s_{t-1}, a_{t-1}) \tag{2.4}$$

The signal δ is called the temporal difference (TD) error. In general, the old estimation of the action value function $Q(s_t, a_t)$ is corrected in proportion to the error signal, $\Delta Q(s_t, a_t) = \eta\delta_t$, where η corresponds to the learning rate.

As mentioned, policies can be either deterministic or stochastic. A policy which always selects the action with the highest expected return is called a *greedy policy*. It tries to maximise the reward based on the current knowledge of the environment (i.e. exploitation). Another type of policy, ϵ-*greedy*, selects the greedy action most of the time, but in some cases (with probability ϵ) selects randomly an arbitrary action among the possible actions. Another way of selecting is the so-called Boltzmann selection,

$$P(a_i|s) = \frac{exp[\beta Q(s, a_i)]}{\sum_{j=1}^{m} exp[\beta Q(s, a_j)]} \tag{2.5}$$

where the parameter β controls the stochasticity of the policy. $\beta = 0$, corresponds to a random selection. As β increases, the probability of selecting the action with the highest Q-value become close to one (when $\beta \to \infty$, it corresponds to a greedy policy).

2.1.3 Simulating experimental paradigms

This thesis follows the animat approach, in which the modelling of cognitive functions takes explicit interest in the interactions between the modelled individual and its environment (Guillot and Meyer, 2001).

The performance of the model is assessed through experimental tests using a simulated robot, i.e. *agent*, based on the physical characteristics of the Kephera robot. The simulated agent is proposed as the embodiment of an artificial rat which is able to gather external sensory information from the environment and use that information, along with internal information (i.e. odometry) to select among future possible actions.

A simulation environment has been developed in our laboratory in order to test the models of spatial learning. It allows to specify different testing environments, as well as the exploratory behaviour of the agent. At each time step, the agent moves accordingly to the selected exploratory behaviour. Once the movement has been done, the simulation environment computes both the external and internal sensory signals corresponding to the new location. This information is then fed into the neural network model and the synaptic connections are modified.

Three types of environments were simulated in this thesis: an open environment, radial arm, mazes and a linear track.

Open environments. Square and rectangular environments are used in this thesis. Two types of exploration behaviours can be used in these environments: (i) when exploring *randomly*, at each time step the agent selects one direction of movement at random and displaces itself a constant distance in that direction. (ii) *Directed* explorations lead the agent to visit a series of significant locations in the environment. At every time-step the agent moves a fixed distance toward the next significant location. In both cases, uniform noise in the orienting direction may be added. Open field arenas have been extensively used in both behavioural and neurophysiological experiments (Muller et al., 1987; O'Keefe and Burgess, 1996; Markus et al., 1995; Fyhn et al., 2004)

Radial mazes. Two types of radial mazes are simulated, a plus maze and a 8-arm maze. The trajectory on the mazes was simulated by defining a law of movement that makes the animal direct itself toward the centre of the maze and the ends of the arms in an interleaved manner. Exploration of mazes in this thesis is equivalent to a directed trajectory in an open environment following the shape of the desired radial maze. No walls were added to signal the arms of the maze. These types of mazes have been used to assess the directionality of place cells (Muller et al., 1994; Markus et al., 1995).

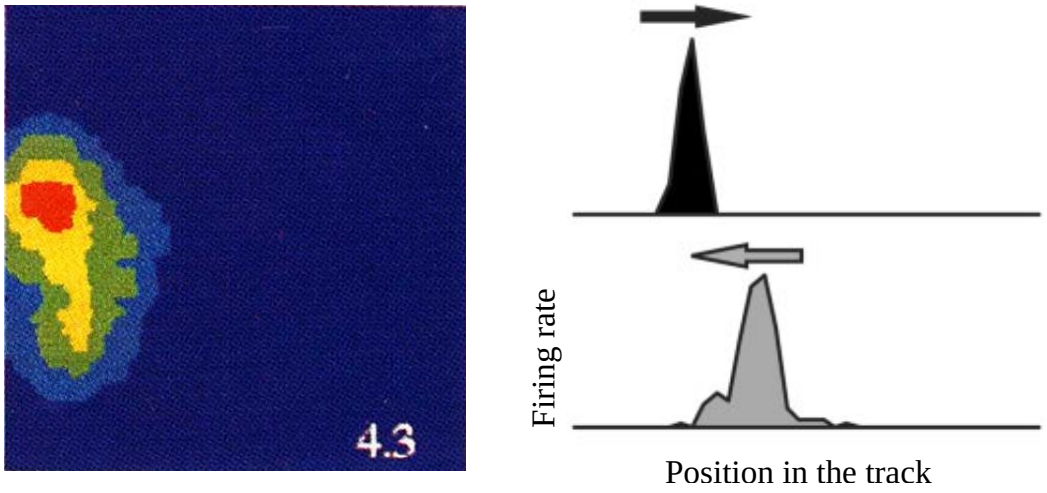
Linear track. A linear track with variable length is used to assess changes in the place code. The track has a maximum length of 188 cm and a box is placed at one of the ends of the track. The box can be displaced among five different positions in order to shorten the length of the track. This type of track is used by Gothard, Skaggs, and McNaughton (1996) to study the interactions between internal and external cues governing the place representation.

2.2 Analysing spatial learning

2.2.1 Place fields

Given a place sensitive cell i, its receptive field, or *place field* $P_i(x^*)$, is computed by estimating its firing rate at each location x^* in the environment. In the general case the environment is segmented in a finite number of bins of equal size and the average firing rate is computed for every segment.

A graphic representation of a place field is shown in Figure 2.1 for both a square arena (two-dimensional environment) and a linear track (one-dimensional environment). Two dimensional place fields are usually represented as a contour plot where the locations in the plot correspond to locations on the environment and the mean activity at each position is coded by the lines or colours of the contour. Place fields in



(a) (b)

Figure 2.1: Graphic representation of a place field. (a) Place field of a CA3 cell in a square environment (122cm x 122cm) surrounded by 61 cm high walls. The 5 colours from dark blue to red, are scaled so that each represents 20% of the peak rate shown in white in the bottom right corner (4.3 Hz). The location of the peak activity of the cell (coded in red) occurs close to the left wall. Redrawn from (O'Keefe and Burgess, 1996). (b) Place fields of a CA1 cell measured in a linear track. The place field is represented in one dimension (horizontal axis) corresponding to the position along the length of the track. The vertical axis shows the activity of the cell. The arrow signals the animal's direction of movement (top, right-bound direction; bottom, left-bound direction). Adapted from (Battaglia et al., 2004).

linear tracks are represented as a two dimensional plot where the horizontal axis corresponds to the position along the length of the track, and the vertical axis corresponds to the cell's firing rate.

Throughout this thesis, when the place field of simulated cells is measured, the parameters of the model are kept constant and the agent is located in different locations uniformly covering the environment. The cell activity is recorded for 8 different heading directions, at each location. Cell activity is colour coded from blue to red corresponding to minimal and maximal activity, respectively.

It should be noticed that there is a significant difference in the measurement of the biological and modelled place cells. In the former case, data is gathered while the animal explores the environment, and the activity is averaged over a significant amount of time. In the case of modelled cells, the activity is measured all over the environment with a constant set of parameters (no learning occurs during the measure

of the receptive fields). Because of this, place fields of the modelled cells may include regions of the environment which haven't been explored before.

2.2.2 Information content of place cell activity

Place cells are supposed to convey information about the location of the animal within the environment. The amount of information and the rate of information conveyed by the activity of a cell can be computed using an approach based on information theory (Skaggs et al., 1992).

In the general case, the information conveyed by a discrete variable X about another discrete variable Y is equal to the mutual information of X and Y, given by,

$$I(Y|X) = \sum_{i,j} p(y_i|x_j) log_2 \frac{p(y_i|x_j)}{p(y_i)} p(x_j) \tag{2.6}$$

where x_j and y_i are the possible values of X and Y, respectively and $p(y|x)$ is the probability of y_i given x_j.

Let X, be the location of the agent; and Y, the activity of a place sensitive cell. The rate (bits/s) at which the cell transfers information about the animal's or agent's location can be computed as (Skaggs et al., 1992),

$$I = \int_x r(x)\ log_2 \frac{r(x)}{\bar{r}} p(x) dx \tag{2.7}$$

where x is the spatial location, $r(x)$ is the mean firing rate at location x, $p(x)$ is the probability density for the rat being at location x, and $\bar{r} = \int_x r(x)p(x)dx$ is the overall mean firing rate.

Equation (2.7) gives a measure of information rate in bits per second. Additionally, by dividing by the overall firing rate of the cell, a measure of the information per unit of activity is obtained. This measure tells us how much information is obtained when the cell is active. It can also be interpreted as a measure of the specificity of the cell. For instance, a cell with a broad receptive field will give little information about location, since it fires in a large area of the environment.

The measure I can be interpreted as how many bits of information a given cell provides in a brief interval of time, whereas $(I/\bar{r})$ corresponds to the amount of information conveyed by the cell every time it is active (i.e. every time it fires a spike). Lets consider the simulated cells shown in Figure 2.2a and Figure 2.2b. The place fields are approximated as a two-dimensional Gaussian with standard deviation σ. Intuitively, cells with narrow place fields (small σ) will provide more information about the spatial location than cells with broader place fields. In concordance to this, the information conveyed by the cell when it fires $(I/\bar{r})$ decreases when the width (σ)

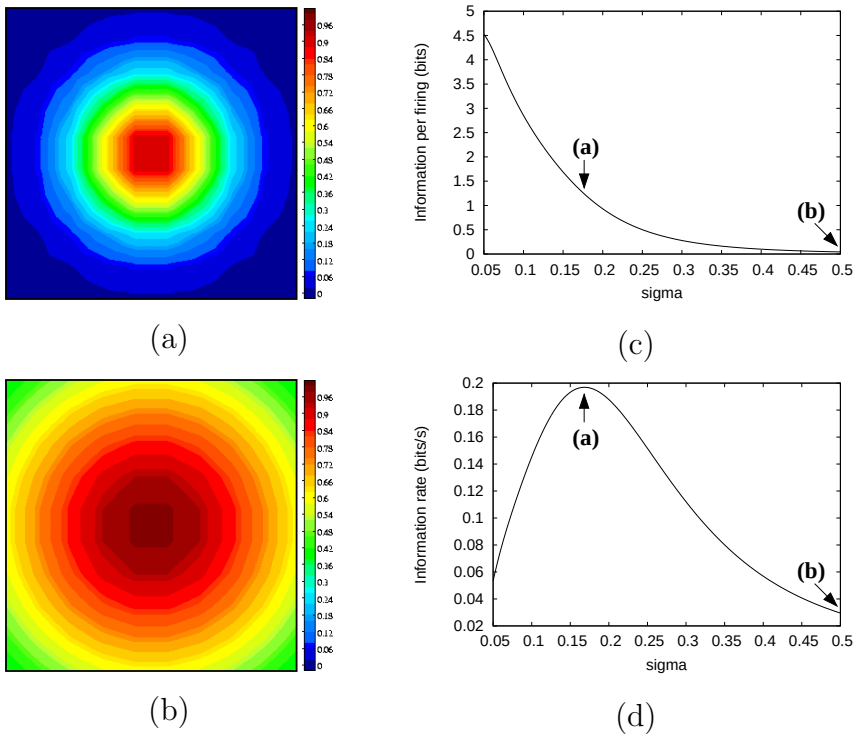


Figure 2.2: Information conveyed by a simulated place cell. The place field of the cell corresponds to a two-dimensional Gaussian, with standard deviation (a) $\sigma = 0.16846$ (b) $\sigma = 0.5$. (c) Information (bits) conveyed by the cell by units of activity. (d) Information rate (bits/s) for different values of sigma. See text for discussion.

of the place field increases, as shown in Figure 2.2c[1]. However, if the cell is only active in a reduced region of the environment, the probability for that cell to be active will be small. As a consequence of this, the cell will rarely provide any information about the location of the animal (even though, in the rare cases it fires, the cell activity conveys a lot of information). Figure 2.2d shows the information transfer rate (bits/sec) for different widths of the receptive fields. As described above, cells with both very narrow and very broad place fields transfer information at a very low rate.

In the general case, Equation (2.7) can be used to compute information rate about variables other than spatial location, such as head direction and running speed (Skaggs et al., 1992).

[1]In this case, it is assumed that all locations within the environment have the same probability of being occupied ($p(x) = cte,\ \forall x$).

2.2.3 Assessing place cell directionality

Even though the term place cell suggests that location is the unique determinant of firing of hippocampal cells, there exist several other factors which influence hippocampal activity (See Section 3.2). In particular, the head direction (or the direction of movement) of the animal seems to modulate the firing of place cells in certain experimental conditions. Place cells seem to be directionally selective when the animal is running in a linear track, whereas they show little dependence on the head direction when the animal freely explores a bi-dimensional environment.

The information theoretic approach described above can be used to evaluate the directional selectivity of place cell firing. The less directional the cell is the less information its activity conveys about the orientation of the animal.

A measure of how much the cell firing depends on the heading direction of the animal can be obtained by computing the ratio of the firing rates at the same location when the animal is facing two opposite directions (Markus et al., 1995). A *directionality index* can be computed as,

$$D_i = \max_{\theta} \frac{|r_i(\theta) - r_i(\theta + 180^\circ)|}{r_i(\theta) + r_i(\theta + 180^\circ)} \tag{2.8}$$

Where $r_i(\theta)$ is the activity of cell i when the animal is heading in direction θ. An index $D_i = 0$ corresponds to a cell which fires equally for all heading directions (i.e. an omni-directional cell). In contrast, an index value $D = 1$ corresponds to a cell which fires maximally when the animal is heading in one direction while being silent in the opposite direction.

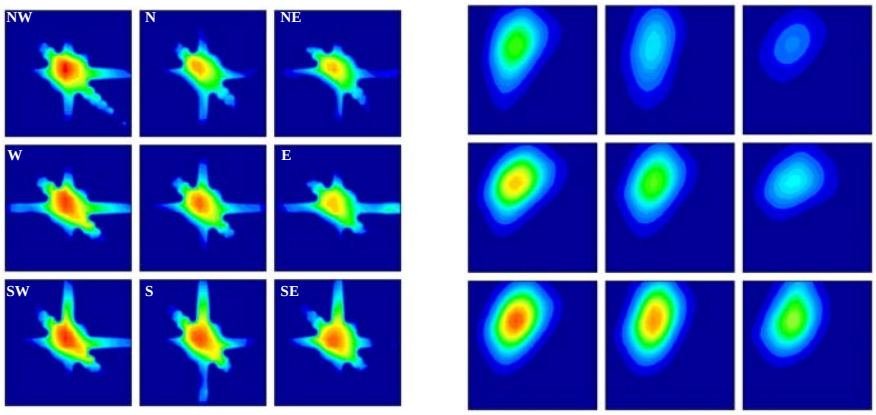
Markus et al. (1995) limit this analysis to the inner circle which fell within the boundaries of the region where the cell was maximally active [2].

An example of two simulated place cells with different directionality properties is shown in Figure 2.3. Nine place fields are shown for each cell. The plot in the centre corresponds to the average activity over all directions. The remaining eight plots show the activity of the cell for specific heading directions (N, for north; E, for east; S, for south; W, for west, and so on).

2.2.4 Population vector coding

The place fields of the place-sensitive cells can be interpreted as an ensemble of radial basis functions coding for the animal's location. Assuming a highly distributed populations of cells, an estimation of the animal's location can be obtained by mean of *population vector decoding* (Georgopoulos et al., 1986). Being p_i the peak firing location of cell i and $r_i(t)$ the activity of that cell at time t, the agent's estimated

[2] Defined as the area of adjacent bins with a firing rate greater than two standard deviations above the mean firing rate in the entire apparatus.



(a) (b)

Figure 2.3: Place field of two simulated cells with different directionality properties. The centre plot shows the average firing overall heading directions. Surrounding plots show the firing rate when the agent is heading in one specific direction. (a) Cell with low dependency of heading direction (D=0.31). This cell corresponds to a simulated cell located in the centre of a 8-arm maze. (b) Strongly directional cell (D=0.74).

location at that moment $\hat{p}(t)$ can be approximated as the centre of mass of the population activity,

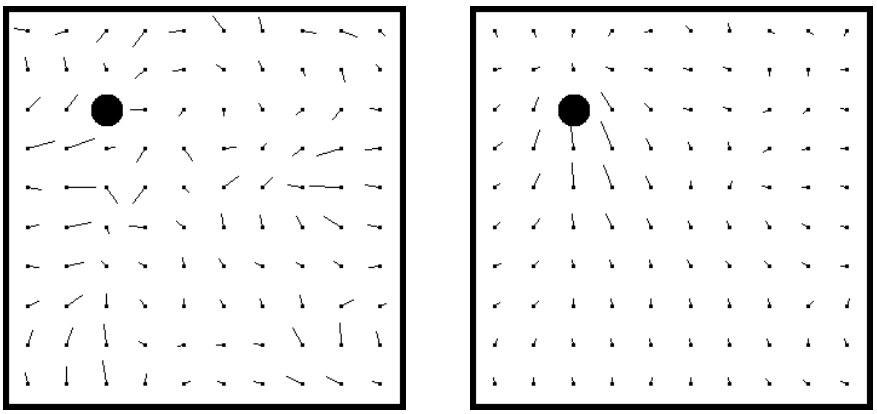
$$\hat{p}(t) = \frac{\sum_i r_i(t) p_i}{\sum_i r_i(t)} \tag{2.9}$$

The decoded position $\hat{p}(t)$ corresponds to an approximate estimation of the real location. It constitutes a good approximation for the case of large neural populations covering uniformly the environment (Salinas and Abbott, 1994). Similar decoding techniques have been used to interpret signals in other brain areas (Georgopoulos et al., 1986; Salinas and Abbott, 1994).

2.2.5 Navigation Maps

The spatial representation can be used to achieve goal-directed navigation. The model presented in this work, performs navigation by learning an association between location and action. In other words, for every position in the environment, the navigation system provides the direction in which the agent should move in order to reach the goal location.

The associations between location and actions are shown in a *navigation map* which shows a set of sampling points in the environment, and a line pointing in



(a) (b)

Figure 2.4: Navigation maps for a goal-searching task in a simulated square environment. The black dots correspond to the sampling points, and the location of the goal is marked by the green filled circle. (a) Navigation map before learning with random initialisation. (b) After learning, the parameters of the model have been updated such that it provides the correct association between location and action (i.e. direction of movement) has been acquired. This is reflected in the navigation map as most of the lines point to the goal location.

the direction provided by the model. In the case when the model has successfully learnt the navigation task, the lines will point toward the location of the goal (usually denoted by a filled circle).

Chapter 3

Spatial representation in rodents

In recent years a wealth of studies have focused on the role of the Hippocampus in spatial learning and navigation, triggered by the finding of place sensitive cells in this area (O'Keefe and Dostrovsky, 1971). The activity of these cells, termed *place cells*, is highly correlated to the location of the animal within the environment. It is believed that, since the activity of a place cell codes for a particular location, the ensemble of these cells may encode a representation of space in an external (i.e. allocentric) frame of reference.

The existence of place sensitive cells and experimental evidence of the role of hippocampus in navigation, is often linked to the theory of the *cognitive map* proposed by Tolman (1948). This map is a representation of the environment in external coordinates which allows the animal to perform goal directed navigation even when the goal location is not signalled by a unique local cue (i.e. *locale navigation*). It also explains the ability of taking shortcuts, and latent learning.

3.1 The rat hippocampus and place cells

Place cells have been measured primarily in the Hippocampus, as well as the medial Entorhinal cortex (EC), the Subiculum (Sb) and the Parasubiculum (PaSb).

3.1.1 Anatomy

The hippocampus is composed of the Dentate Gyrus and the Hippocampus proper (or *Cornu ammonis*, CA). The latter is further subdivided in four subregions denoted

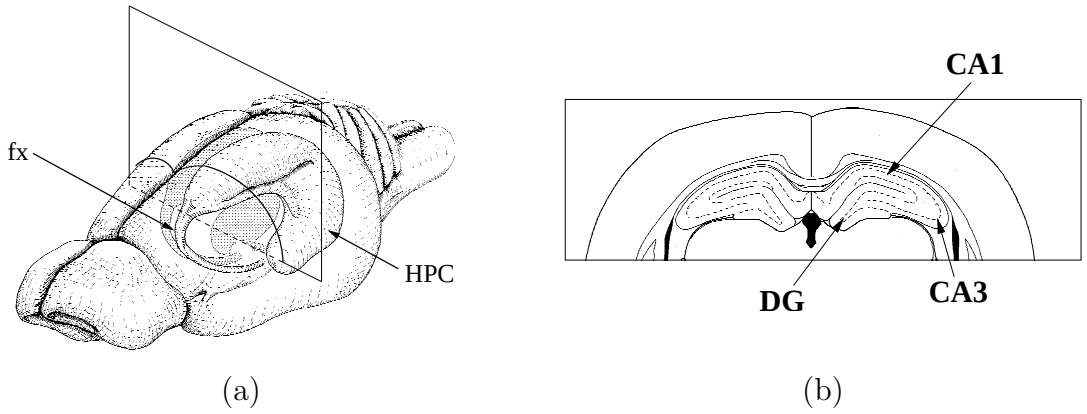


Figure 3.1: Anatomy of the hippocampal formation. (a) Schematic rat brain. Hippocampal formation highlighted. Adapted from (Amaral and Witter, 1995). (b) Coronal section (Bregma −3.80mm). Adapted from (Praxinos and Watson, 1998). HPC, Hippocampal formation; fx, Fimbria fornix; DG, Dentate gyrus; CA1/CA3, Cornu ammonis subregions.

CA1 to CA4, but CA1 and CA3 are the most distinguishable areas.

Main projections from/to the Hippocampal formation

The hippocampus receives projections from *(i)* the EC via the *perforant path*. These projections carry information from most of the multimodal and unimodal associative areas. *(ii)* Sub-cortical areas via the *fornix fibre bundle*. These include projections from the thalamus, the hypothalamus, the brainstem, and the amygdala. *(iii)* Subcortical cholinergic and GABA-ergic inputs from the septal region modulate the hippocampal activity, and are believed to be responsible for generating the theta rhythm (Buzsáki, 1984).

The hippocampal formation has two main output pathways. One pathway projects to subcortical areas via the subiculum. It innervates thalamic nuclei, amygdala and the nucleus accumbens. The second pathway projects to several cortical areas through the Entorhinal cortex.

Figure 3.2 shows a schematic representation of the principal connectivity between the different structures composing the hippocampal formation.

Entorhinal cortex

The Entorhinal Cortex is the cortical gateway to the hippocampus. EC follows the neocortical six-layered organisation. This structure is usually subdivided in medial

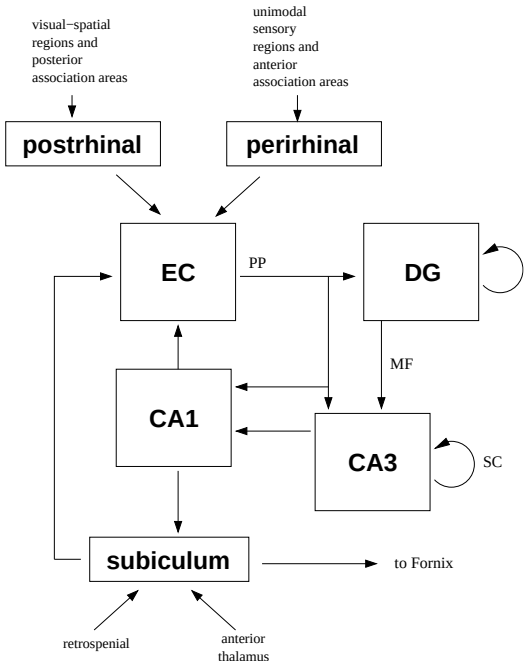


Figure 3.2: Connectivity of the main areas of the Hippocampal formation. EC, Entorhinal cortex; DG, Dentate gyrus; CA1,CA3, Cornu ammonis subregions. PP, Perforant path; SC, Schaffer collaterals; MF, Mossy fibres.

(mEC) and lateral (lEC) areas.

Both mEC and lEC receive projections from sensory associative areas (visual, auditory and somatosensory) as well as from the parietal, temporal and frontal areas via perirhinal and postrhinal cortices (Witter, 1993; Suzuki and Amaral, 1994; Insausti et al., 1997; Liu and Bilkey, 1997). Olfactory bulb and piriform cortex project to the EC via perirhinal cortex (Kosel et al., 1981; Witter et al., 1989; Schwerdtfeger et al., 1990; Liu and Bilkey, 1997; Biella and De Curtis, 2000; Witter et al., 2000; Insausti et al., 2002). Outputs from subiculum and CA1 are projected mainly to mEC, but also to lEC. There are no connections from DG and CA3 (Amaral and Witter, 1989; Witter, 1993).

EC outputs target mainly perirhinal, orbitofrontal and piriform cortices. Other targeted areas are the temporal, frontal and occipital areas (Silva et al., 1990; Witter, 1993; Insausti et al., 1997). Projections from EC to the hippocampus (DG, CA1, CA3 and SB) are conveyed via the *perforant path* (Amaral and Witter, 1989; Witter, 1993).

Dentate gyrus

DG is a major recipient of EC projections via the perforant path. These projections target granule cells in DG, whose axons collateralise to mossy cells and then project onto the CA3 where they synapse on the proximal dendrites of pyramidal cells (Claiborne et al., 1986; Amaral and Witter, 1995; Hastings et al., 2002). A particular characteristic of this structure is that in DG neurogenesis takes place throughout the entire life (Bayer, 1982; Kuhn et al., 1996; Ciaroni et al., 1999; Hastings et al., 2002).

Hippocampus proper (Cornu ammonis)

The hippocampus proper is subdivided in four subregions CA1 to CA4, although only CA1 and CA3 are often distinguished. CA1 was the first structure where place-sensitive cells were recorded by O'Keefe and Dostrovsky (1971).

CA3 and CA1 receive projections from the EC via the perforant path. CA3 is also innervated by DG via the mossy fibres (Amaral and Witter, 1989; Amaral, 1993; Witter, 1993; Amaral and Witter, 1995; Hastings et al., 2002).

CA3 and CA1 project via the Schaffer fibre bundle to subiculum. CA3 projects onto CA1 via the Schaffer fibres. Additionally, CA3 cells have recurrent connections via the Schaffer collaterals (Amaral and Witter, 1989; Amaral, 1993; Amaral and Witter, 1995).

Subiculum

The Subicular complex consists of subiculum (Sb), the pre-subiculum (prSb), whose dorsal part forms the post-subiculum (poSb) and the para-subiculum (paSb).

Sb receives input from EC and CA1 (Amaral et al., 1991). Retrospenial cortex projects onto paSb, and prSb is innervated by temporal and parietal lobes. Thalamic nuclei projects onto prSb and poSb (Burgess et al., 1999).

Outputs from Sb project to the nucleus accumbens (NA) and the septal complex via the fornix fibre bundle. Prefrontal cortex, amygdala and thalamus are also targeted by the Sb. EC receives projections from prSb and paSb (Amaral and Witter, 1989; Witter, 1993).

3.1.2 Hippocampal place cells

Place sensitive cells have been found in the different regions of the hippocampus, the CA1 being the first region where place cells were recorded (O'Keefe and Dostrovsky, 1971).

Pyramidal cells in the CA1 and CA3, and granule cells in the DG show clear location dependent activity. Both types of cell exhibit directional selectivity when the animal explores radial arm mazes and linear tracks.

CA3-CA1 place fields tend to cover uniformly the environment, and at any moment only a small proportion of these cells is active. The set of overlapping place fields is thought to encode a distributed representation of the environment. In general, any given cell may have a place field in several but not all the environments (Kubie and Ranck, 1983), and there is no relationship between the spatial topology of the coding and the anatomical structure of the CA3-CA1 populations (O'Keefe and Conway, 1978; Kubie and Ranck, 1983; Muller and Kubie, 1987; Thompson and Best, 1989).

When the animal follows repeatedly the same route, CA1 place fields, which are initially symmetrical, expand backward with respect to the animal's direction of movement. Such a change only take place during the same block of trials, and its effect disappear the next day (Mehta et al., 2000).

CA1 place fields have been recorded even after selective lesions of the DG, (McNaughton, Barnes, et al., 1989) and partial and complete lesions of the CA3 (Mizumori et al., 1989; Brun et al., 2002, respectively). Furthermore, directional selectivity in linear tracks is also observed in rats with complete disruption of CA3 inputs to this area. Place fields remain stable across sessions and the firing activity was theta-modulated. These studies suggest that an intact CA3 is not required to maintain the place and directional properties of CA1 neurons.

Subicular place fields are broader than those in the hippocampus proper. Furthermore, subiculum exhibits a similar place field topology across different environments (Sharp and Green, 1994; Sharp, 1997, 1999), and seems to be largely influenced by self-motion information (McNaughton et al., 1996), suggesting that this area may be a neural substrate for spatial coding based on path integration.

Initial recordings of place cells in the EC, reported broad single peaked firing fields (Quirk et al., 1992). More recent data have shown a gradual change of information related activity in the EC along the dorsolateral to ventromedial axis of the medial entorhinal cortex (mEC) (Fyhn et al., 2004). Localised, multi-peaked place fields where found in the dorsolateral band of the mEC. These place fields remain stable even after disruption of hippocampal projections to the EC. Cell activity is as spatially selective as hippocampal place cells, and less modulated by the animal's direction of movement.

Changes in the shape of the environment (i.e. from circular to square arenas) do not affect place fields of mEC cells. In a general sense, mEC largely preserves the firing topology across environments. When changing from one environment to another, mEC place fields are highly correlated, and cells active in one environment are likely to be also active in the second environment (Quirk et al., 1992).

Recently Hafting et al. (2005) reported a directionally oriented, topographically organised structure in the dorsal medial EC. Cells in this area (termed, *grid cells*) fire at the vertex of a regular grid of equilateral triangles. Neighbouring cells have a common orientation and spacing, but their vertex locations differ. The grid organisation is anchored on external landmarks, but persists if they are absent. This suggests

that these cells constitute the basis of a spatial representation based on idiothetic cues (i.e. path integration).

3.1.3 Theta rhythm and sharp waves

Hippocampal EEG is characterised by two different patterns according to the animal's ongoing behaviour. When the animal moves (actively or passively), hippocampal activity shows a 6-12Hz oscillation called *theta rhythm*. This oscillations are also observed during REM sleep. Additionally, firing is synchronised to a *gamma oscillation* (40-100 HZ) throughout the whole hippocampal formation. In contrast, during slow-wave sleep or when the animal is in awake-immobility the hippocampal EEG shows large field high irregular amplitude signature, termed *sharp waves*. During each sharp wave event, a high frequency ripple volley of 140-200Hz occurs.

Theta rhythm has been proposed to originate from cholinergic and GABA-ergic projections from the septal region (Buzsáki, 1984; Hasselmo and Bower, 1993; Miller, 1991; Winson, 1978). Inactivation of septal inputs affects CA3 place fields, while leaving CA1 fields unaffected. Additionally, it impairs acquisition of place learning tasks (Brandner and Schenk, 1998), and working memory tasks (Mizumori et al., 1989).

Theta rhythm has been proposed to modulate the input to the hippocampus (Chrobak and Buzsáki, 1996; Chrobak et al., 2000); as well as to mediate the balance between encoding and retrieval by controlling the relative strength of feed-forward and recurrent connections in the hippocampal formation (Hasselmo, Bodelon, and Wyble, 2002; Hasselmo, Hay, et al., 2002; Rogers and Kesner, 2003).

It exists a significant relationship between the firing of place cells and the theta rhythm. It has been observed that the phase when the place cell fires a spike, precesses to lower values as the animal moves further in the cell's place field (O'Keefe and Recce, 1993). This phenomenon, known as *phase precession*, has been proposed to play a key role in encoding of sequences by the hippocampus (Skaggs et al., 1996), as it reproduces the temporal order of activation of cells with overlapping place cells on a short time scale ($< 10ms$) (Mehta et al., 2002).

Sharp waves, in turn, have been proposed to modulate the output from the hippocampus to the cortex (Chrobak et al., 2000). During both slow-wave sleep and quiet wakefulness reactivation of the temporal pattern of activation of hippocampal place cells have been observed (Wilson and McNaughton, 1994; Skaggs and McNaughton, 1996; Kudrimoti et al., 1999). The repetition of firing sequences has been interpreted as a reactivation of memory traces and supports the hypothesis that there is a "replay" of hippocampal activity to the neocortex, which plays an important role for memory consolidation (McClelland et al., 1995). Recent data suggest that the number of replays during inter-trials non-exploring states within a session increases with the experience suggesting that the replay may be the effect of experience-dependent

changes in the network excitability (Jackson et al., 2005).

3.2 Firing determinants of place cells

As the term suggests, the discharge of place cells is highly correlated with the location of the animal. However, place cells are also sensitive to other factors like speed, orientation, reward, odour or tones, among others (McNaughton et al., 1983; Markus et al., 1994, 1995). It is believed that the hippocampus codes for a combination of multimodal stimuli defining a particular episode (*context*).

In any environment, only a subset (20-25%) of the pyramidal cells are active, while the rest of the population remains silent (Thompson and Best, 1989). When returned to a familiar environment the firing pattern of place cells is consistent with previous experiences in the same environment (Muller et al., 1987; Thompson and Best, 1989). One cell may be active in several environments, even though the active subsets in different environments are highly uncorrelated (Kubie and Ranck, 1983; Thompson and Best, 1989).

Furthermore, hippocampal place cells do not show a topological organisation, anatomically neighbouring cells do not fire in adjacent locations in the environment (O'Keefe and Conway, 1978; Kubie and Ranck, 1983; Muller and Kubie, 1987; Thompson and Best, 1989).

3.2.1 Sensory information

Place cell activity is influenced by both idiothetic and allothetic inputs, and there exists a wealth of experimental data focused on the main determinants of place cell firing, as well as the interactions between different sources of information.

Distal vs local visual cues. Several studies suggest that self-localisation in rats relies strongly on visual cues (Etienne et al., 1996; Maaswinkel and Whishaw, 1999). Furthermore place cell firing mainly depends on distal visual cues. Rotation of distal landmarks produces a corresponding rotation on the hippocampal place code, whereas rotation of proximal cues does not. However, if the location of a reward is tied to the location of a local cue, place cells firing may be determined by that local cue neglecting distal cues (Breese et al., 1989; Biegler and Morris, 1993, 1996; Gothard, Skaggs, Moore, and McNaughton, 1996).

Non-visual external sensory input. When visual cues are not available, rats can rely on other sensory modalities to drive the place cells firing. Olfactory cues have been observed to contribute to the stability of place fields (Lavenex and Schenk, 1996; Save et al., 2000).

Furthermore, when exploring an environment containing three-dimensional objects as local cues, blind rats display an exploratory pattern that led them to make

contact with the objects more often than sighted rats (Save et al., 1998). Presumably, this behaviour allows the animal to calibrate its internal estimations (i.e. path integrator) using tactile information.

Idiothetic cues. Hippocampal principal cells show strong place dependency in absence of visual cues. In this case, they may be driven by movement related internal signals. For instance, removing visual cues does not alter the activity of hippocampal place code (O'Keefe and Conway, 1978; Hill and Best, 1981; Pico et al., 1985; Muller and Kubie, 1987; O'Keefe and Speakman, 1987; McNaughton, Leonard, and Chen, 1989; Quirk et al., 1992); and place cells in visually symmetric environments have asymmetric visual fields (Sharp et al., 1990). Moreover, when the place cells are established in darkness, they persist in subsequent light conditions (Quirk et al., 1990; Markus et al., 1994).

Save et al. (1998) recorded place cell activity in early blind rats, and found place field similar to those recorded in sighted rats. It suggests that motion-related information (and possibly non visual sensory information) may be sufficient to form a stable place code in the hippocampal formation.

Interaction between internal and external information. Despite the fact that internal motion-related information may be used to update place cell firing (by path integration) (M. L. Mittelstaedt and Mittelstaedt, 1980; Etienne and Jeffery, 2004), this information is prone to accumulate errors over time. This requires a process of re-calibration based on external sensory input to take place in order to maintain an accurate estimation of position (H. Mittelstaedt, 1983; Gothard, Skaggs, and McNaughton, 1996; Etienne et al., 1996; Save et al., 2000; Redish et al., 2000).

Gothard, Skaggs, and McNaughton (1996) study the interactions between internal and external information. They observe that, in case of mismatch between the information provided by external and internal cues, the population activity changes in order to align itself with one of the two. Gothard, Skaggs, and McNaughton (1996) conclude that the dynamics of these changes were determined by the degree of the perceived mismatch. Using a similar protocol Redish et al. (2000), conclude that in addition to the mismatch, a temporal delay (different from an animal to another, but consistent within animal) is required to make the transition.

This data suggest that both types of information compete in a dynamical process to control place cells firing.

3.2.2 Geometrical properties of the environment

O'Keefe and Burgess (1996) reported that environmental features control the location and shape of the place fields of hippocampal cells. They measured CA1-CA3 pyramidal cells of rats exploring four different boxes (a small square; a horizontal rectangle; a vertical rectangle; and a large square) located within a rectangular room. The location of peak firing in different boxes was found to depend on environmental

features such as the distance to one or several walls, or the ratio of the distances to two opposite walls. Furthermore, the place fields change according to manipulations of the environment. In some cases, in boxes differing along one dimension the place fields either stretches or reveal a second peak along the changing direction.

Based on this data, they suggest that these place fields may result of the conjunction of several independent subcomponents tuned to respond maximally at a given distance from a specific wall (O'Keefe and Burgess, 1996) or landmark (Burgess et al., 1994).

3.2.3 Direction of movement

Experimental data have shown that the firing of place cells doesn't depend on the heading direction when the animal is foraging freely in an open environment. In contrast, on radial mazes or linear tracks the firing of these cells is becomes strongly directionally selective (Muller et al., 1994; Gothard, Skaggs, and McNaughton, 1996). Furthermore, place directionality also appears in open environments when the animal is constrained to follow the same path between points of special significance (Markus et al., 1995).

In the experiment described in the previous section, O'Keefe and Burgess (1996) observed that stretched or double peaked receptive fields are also modulated by the direction of movement, such that the peak of the field shifts toward the wall away from which the animal is moving. This may suggest that omni-directional place fields observed in open environments are the result of several superimposed directional subcomponents.

Recently, Battaglia et al. (2004) showed that the directionality of CA1 cells of rats running in linear tracks can be changed by manipulation of multi-modal, local sensory cues. They compared the directionality properties of place cells in rats running along linear tracks in either cue-poor or cue-rich environments. The cell's activity when the rat explores a linear track which is surrounded with salient, multimodal cues is less directional than in cue-poor environments. Additionally, the place coding of cue-rich environment seem to be sparser and to change more quickly in space (i.e. the place fields are more specific).

Most of the bidirectional cells found in this experiment don't fire at the same location in the track. Place fields seem to be shifted backward relatively to the opposite direction. These may suggest that the place cells code for locations ahead of the animal, which can be interpreted as *prospective coding*. Similar effects were also reported by O'Keefe and Burgess (1996). This phenomenon have been addressed by models which takes into account the effects of experience and plasticity in place coding populations (Mehta et al., 2000). Some models (Tsodyks et al., 1996) have also proposed the prospective coding as the source of the phase precession phenomenon (O'Keefe and Recce, 1993).

The fact that the availability of local cues affects the directionality of place cells support the idea that external stimuli might have a view-independent representation which projects onto the hippocampus. In cue-rich environments the views taken at opposite directions may have a greater correlation than in the cue-poor environment thus yielding a less directional place code.

3.3 Modelling hippocampal place cells

A significant amount of effort has been devoted in the last 20 years to modelling the role of the hippocampus in spatial learning and navigation. This section reviews some of the models which have been proposed. Some of these models rely on the idea that location-specific activity can be driven by convergent input from sensory input responding to environmental stimuli available at each place cell's firing field (Zipser, 1985). The external sensory input available at each location, termed *local view*[1], can be used for later recognition of the same location. Other models combine this information with internal estimations of location based on the accumulation of proprioceptive signals

This section reviews some of the models which have been proposed. All the models are based upon the recognition of a local view (at different levels of abstraction) which, in some cases may be combined with a location estimation based on path integration.

3.3.1 Sharp (1991)

Sharp (1991) proposes a model of place cells based on local-view recognition by means of competitive learning. The system, represented in Figure 3.3 consists of a three-layer feed-forward neural network. Every layer is fully connected to the following layer and the connections are updated using a Hebbian learning rule.

The first layer is composed of two types of sensory cells which respond to specific stimuli. Type 1 cells are active whenever the distance d from the agent location to a specific external cue is equal or lower to a fixed threshold θ_d. The activity of Type 2 cells depends not only on the distance ($d < \theta_d$), but also in the bearing angle (ϕ) between the agent's heading and the external cue.The parameters for both types of cells (distances and angle range) are assigned stochastically at the beginning of the simulation and remain fixed from then on.

The middle layer is divided in three winner take-all (WTA) clusters of 20 cells each. The input of a cell j in this layer is equal to $h_j = \sum_i (w_{ji} r_i)$ where $r_i \in [0, 1]$ is the activity of a sensory cell i. The weights w_{ji} are initialised randomly, and their sum

[1]The term local view corresponds to the set of multimodal sensory input available at one specific location, and is not necessarily limited to visual input.

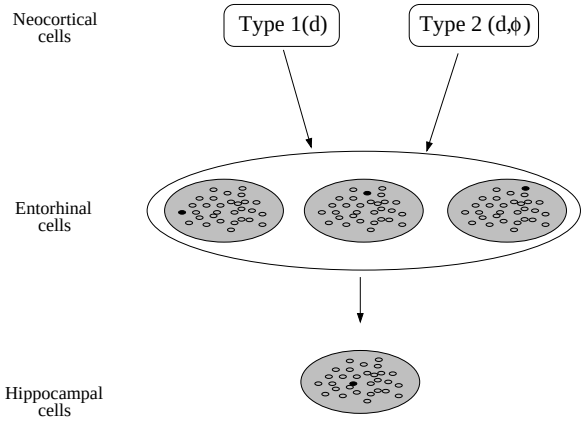


Figure 3.3: Model proposed by (Sharp, 1991).

is normalised. In every cluster, only the cell with the highest input $h^* = max_j(h_j)$ is active, and the synaptic connections to that cell are updated by Hebbian learning.

The activity of the second layer (i.e. three active cells.) is propagated to the third layer. This layer is composed of one cluster of 20 cells, and the same WTA mechanism as described above is applied to compute the firing of the units and to update the weights. Computer simulations of this model yield omni-directional place sensitive cells in this layer, when the agent explores a circular environment. Furthermore, if the movement of the agent is restricted to follow paths similar to the eight-arm maze, the firing of the simulated place cells also depends on the direction of movement (i.e. head direction) in agreement with reported recordings in rats, as described in Section 3.2.

Sharp associates every layer of this model to putative brain structures. Namely the first layer, in charge of sensory process, would correspond to neocortical areas. The second layer, which integrates information from several sensory cells, corresponds to the entorhinal cortex, while the third layer, composed of place sensitive cells, is associated to the hippocampus.

This model relies strongly on the availability of distinguishable visual cues, and therefore it doesn't reproduce experimental findings showing that the hippocampal place fields are maintained in absence of external visual cues (Quirk et al., 1990) or even in blind rats (Save et al., 1998). In addition, the winner-take-all competition that takes place in the model contrasts with neurophysiological data suggesting a distributed, redundant place code in the hippocampal formation.

3.3.2 Burgess *et al.*(1994)

Burgess et al. (1994) propose a model which, like the one previously described, propagates information from landmark sensitive cells to a multilayer neural network.

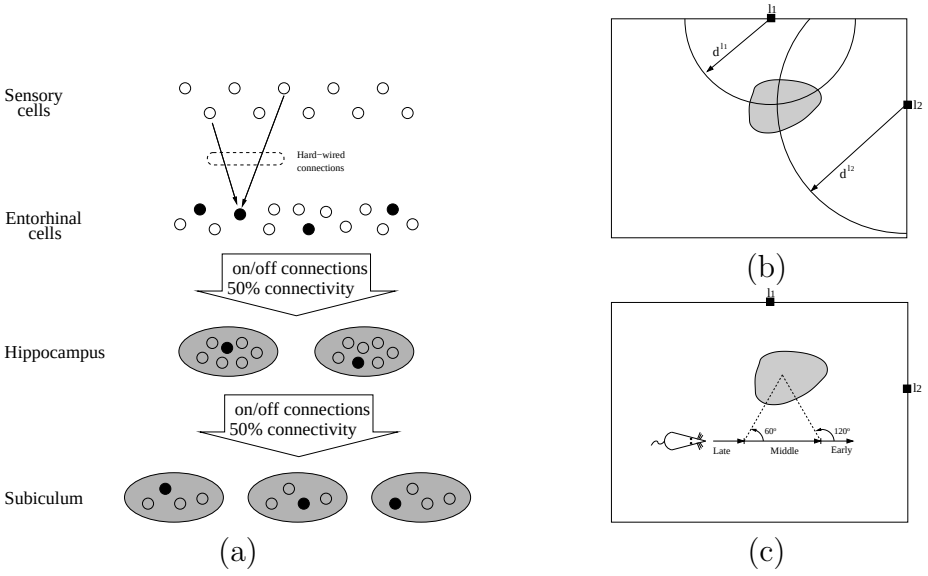


Figure 3.4: Model proposed by Burgess et al. (1994). (a) Schematic representation of the model. Shaded areas correspond to clusters where winner take all competition takes place. Adapted from (Strösslin, 2004). (b) Tuning curves of two sensory cells sensitive to landmarks l_1 and l_2, and preferred distances d^{l_1} and d^{l_2}, respectively. The activity of an entorhinal cell, driven by these cells will depend on the product of its afferents' firing rate. (c) Theta phase precession. Depending on the angle α between the agent's head direction and the location of the receptive field, the entorhinal cell firing will take place in the early, middle or late phase of the theta cycle, as the animal crosses the place field.

As in the model by Sharp (1991), a first layer of sensory cells projects onto a second layer associated to the EC. Activity is then propagated downstream to two more layers of place sensitive cells, which the authors relate to the CA1-CA3, and the subiculum (Burgess and O'Keefe, 1996; Burgess et al., 1994). A schematic representation of the model is shown in the Figure 3.4a.

Cells in the sensory layer have a broad tuning to the distance to external cues. As a function of distance each sensory cell i will have a triangular tuning curve, peaked at a preferred distance d_i^l from a landmark l. For each landmark l, there are 15 cells associated to it, with different preferred distances d_i^l. These cells project onto the entorhinal layer. Every entorhinal cell e receives projections from two sensory cells i and j, and its firing rate $r_e = \lfloor r_i r_j / 2 \rfloor$, approximate a radial basis function around the intersection of the tuning curves of its afferents, as shown in Figure 3.4b.

Entorhinal cells are then connected to the Hippocampal layer through binary connections. The connection from a cell i to a cell j is "turned on" ($w_{ji} = 1$)

whenever both cells are active in the same time step (i.e. Hebbian learning). The Hippocampal layer is subdivided in 5 cluster of 50 cells each, and a winner take-all mechanism takes place within every cluster.

Hippocampal cells, in turn, project onto the subicular layer following the same Hebbian learning as before. A competitive learning also takes place in this layer, but in this case the population is subdivided in 10 clusters of 25 cells each. Given that subicular cells have to compete with less cells than hippocampal cells, place field in this layer will be larger than hippocampal place fields.

In addition, as the sensory cells are driven by metric information (i.e. distance to landmarks), the model reproduces changes in the place field characteristics following changes in the geometry of the environment, as reported by experimental data (O'Keefe and Burgess, 1996) .

The model also includes a relationship between place coding and the theta rhythm. Specifically, each time step is assumed to last one third of the theta cycle and, correspondingly each theta is divided in early, middle and late phase. The phenomenon of phase precession is modelled to be originated in the firing of the Entorhinal cells (EC). For cell $e \in EC$, sensitive to two landmarks l_1 and l_2, the agent will compute the angle α between its heading direction and the egocentric direction to the centroid of the cell's place field. For $|\alpha| < 60°$ (i.e. the place field is 'ahead' of the agent) the cell e will fire at the early phase of the theta cycle, if $|\alpha| > 120°$ (i.e. the place field is 'behind'), the cell fires at a late phase; otherwise, the cell fires in the middle phase of the theta cycle (Figure 3.4c). This mechanism reproduces the effect of phase precession described in Section 3.1.3.

This model assumes that landmarks are perfectly distinguishable, and an accurate estimation of the distance to those landmarks is available to the agent. In addition, the theta phase dependence of the entorhinal activity requires the computation of the bearing angle to the cell's receptive field. Finally, this model relies on external sensory input and does not take into account idiothetic signals, being incapable of reproduce location dependent activity in absence of external cues.

3.3.3 Wan, Redish and Touretzky (1994–1997)

Wan et al. (1994) proposed a model composed of separate populations coding for local view, head direction, path integrator and place code. These populations interact with each other to form a stable representation of space.

Local view is defined in simulation as an arrangement of landmarks. Each landmark i is defined by a triplet $[T_i, r_i, \theta_i]$ corresponding to the type, the distance and the bearing angle. The head direction population provides the current heading direction Φ which is used with the egocentric angle θ_i to compute the allocentric landmark angle ϕ_i.

The head direction estimation Φ is updated by integration of angular velocity

signals derived from vestibular and efferent motor signals. In case of disorientation, the local view system can reset the current heading estimation by comparing the egocentric and allocentric bearing of landmarks.

Like the head direction population, the path integrator system adds up motor efference copies to update an estimation of the agent's position (x, y).

Local view and path integration information are combined in the place code population. Units are recruited and tuned to respond to: (i) two randomly chosen landmarks (each one defined by the triplet $[T_i, r_i, \theta_i]$); (ii) retinal angle difference $(\alpha = \theta_i - \theta_j)$ between two different landmarks; (iii) position information (x, y) given by the path integrator. Place units compute a fuzzy combination of these parameters, which drops out missing or unreliable inputs.

Computations in the model are performed in an algorithmic manner, and the local view system relies on an abstract visual system able to identify the landmark type, and measure its distance and bearing angle.

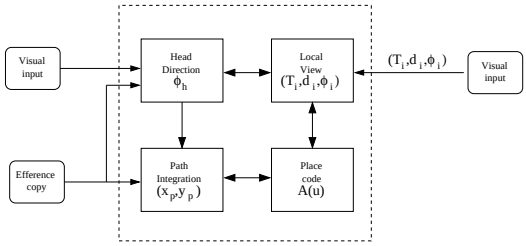


Figure 3.5: Schematic diagram of the model proposed by Wan et al. (1994).

3.3.4 Gaussier *et al.*(2002–2005)

Unlike other models of hippocampal spatial learning, Gaussier and colleagues propose that hippocampal cells code for *transition* between places (rather than place itself), driven by place sensitive cells in the entorhinal cortex (Gaussier et al., 2002; Banquet et al., 2005).

The input of the model is a set of landmarks detected in a real panoramic image. Each landmark is defined by its *type* and its *bearing angle*, this information is combined in a "what" and "where" matrix. Neurons coding for the type of landmark form a winner-take-all network, whereas the population coding the bearing angle spreads its activity to neighbouring cells to allow "generalisation".

Interesting locations (i.e. near to a goal) are encoded by recruiting a location sensitive unit. This unit is connected to active cells in the what-where matrix. In subsequent time steps, the activity of the place cell is computed in two phases, first,

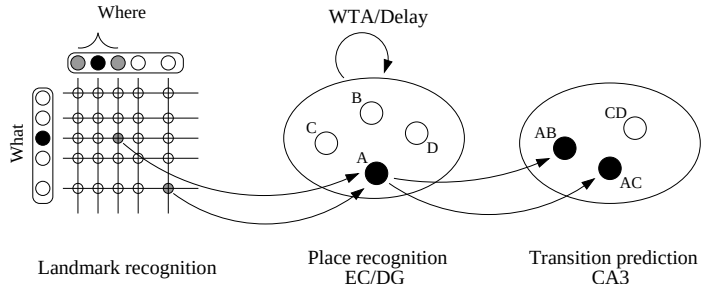


Figure 3.6: Place recognition in the model proposed by (Gaussier et al., 2002). Local views are processed in a "what"-"where" matrix driven by cells tuned to the type and bearing of landmarks in the visual field. A winner-take-all competition takes place among place cells to achieve place recognition. Transitions between sequentially active place cells are learnt by means of a delay loop which allows to associate the previously active cell with the currently active cell. Learnt transitions can later be used to predict possible future locations.

the initial activation is computed by,

$$act_i = 1 - \frac{1}{\pi N_i} \sum_{k=1}^{N} V_{ik}[v_k|\Theta_{ik} - \theta_{ik}|_\pi + \pi(1 - v_k)] \tag{3.1}$$

Where N is the number of perceived landmarks, N_i the number of landmarks perceived when the cell was created, Θ_{ik} is the bearing of landmark k from location i, θ_{ik} is the bearing of the same landmark from the current position. V_{ik} is set to one if landmark k was visible when the cell i was recruited and zero otherwise; $v_k$is equal to one if landmark k is currently visible. $|.|_\pi$ is the absolute value modulo π. Second a winner-take-all mechanism resets the activity of all cells but the winning cells to zero. This module is proposed to reside in the dentate gyrus and entorhinal cortex.

A delay in this layer allows the next layer to encode place transitions. Previously active cells connect to transition cells using a Hebbian-type rule.

3.3.5 Arleo *et al.*(2000–2004) and Strösslin *et al.*(2004)

The model proposed initially by Arleo (Arleo et al., 2004; Arleo and Gerstner, 2000), and developed further by Strösslin (2004) is composed by two pathways processing idiothetic and allothetic information, respectively.

The idiothetic information is processed in an algorithmic manner in which the position estimation is updated by integrating signals coding the velocity and the direction of movement of the agent. This position estimation is neurally represented in a population of cells (denoted PI cells, for path integration) whose receptive fields

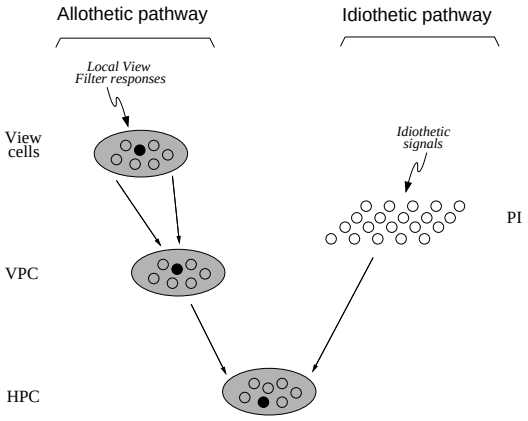


Figure 3.7: Model proposed by Arleo and colleagues. See text for explanation.

are modelled as a two-dimensional Gaussian around a preferred location x_i. The firing rate r_i of a PI cell i is,

$$r_i = exp\left(- \frac{(\vec{p} - \vec{p_i})^2}{2\sigma^2} \right) \tag{3.2}$$

where p is the estimated location, and σ is the width of the receptive field.

The allothetic pathway uses high-dimensional visual input. Two-dimensional images are transformed into a vector of features corresponding to the response to a set of Gabor filters applied on specific locations of the image (Arleo et al. (2004) use a log-polar grid, whereas Strösslin (2004) uses an uniform square grid). In both cases more than 1000 features were used to code for each local view.

At each location, the agent takes a snapshot of the local view and filters the image. The response of the filters is stored in visual cells (VC), whose activity depends on the difference between the current local view, and the filters responses it has stored. Activity of VCs is then propagated onto another population where information from several local views is integrated to form a vision-based place representation. In earlier versions of the model (Arleo et al., 2004) local views were taken in the four cardinal directions and successful recognition relies on the accurate estimation of these directions. Strösslin (2004) in contrast, uses a rather artificial visual processing which requires no alignment and yields directional place fields in the vision based place representation.

Both populations, the one based on path integration (PI) and the vision based place code (VPC) project onto a population where a combined place code is formed. This population is associated to the hippocampus, and Hebbian learning is used to integrate information from external and internal sensory signals.

Apart from the updating rule for the PI population described above, activity of these cells is also continuously modified such as it tends to be synchronised with

the allothetic estimation. This allows the system to reset the PI when the agent is first located into the environment, and also compensate for the drifting caused by the accumulation of errors in the estimation of self-movements (Chavarriaga and Gerstner, 2004; Strösslin, 2004).

3.3.6 Modelling place cells directionality properties : Brunel and Trullier (1998)

Brunel and Trullier (1998) proposed that CA3 place cells are intrinsically directional, as they are driven by direction dependent entorhinal signals. Non directionality in the case of exploration of random environments is the result of modification of recurrent connections in the hippocampal CA3.

They propose a model of CA3 composed of two interconnected population of excitatory and inhibitory cells. Cells were modelled as integrate and fire neurons. Excitatory cells have pre-determined place and heading direction selectivity, and are supposed to model for population of pyramidal cells with similar place tuning. The inhibitory population represents CA3 inter-neurons and has no spatial selectivity. Cells are interconnected by binary synapses that can be in two possible states (J_{low} and J_{high}). Transition between these states depends on the average activity of pre-synaptic and post-synaptic cells. Transitions are modelled after experimental observation of long term potentiation and depression (LTP and LTD, respectively) of hippocampal neurons (Bliss and Collingridge, 1993).

Consistently with experimental data, random exploration of open environments *decreases* the directionality dependence of hippocampal cells. Such a reduction in the directionality of the place code is due to the uniform exploration of the same location at several heading directions. This allows synapses connecting cells with similar spatial selectivity (i.e same location but different head directions) to be strengthen. The model predicts that there are not changes in the sensory input to the hippocampus (via the entorhinal cortex), and even weak directional modulations of the input signals get amplified by the recurrent dynamics of the CA3 network.

Two recent experimental studies on the firing properties of mEC cells seem to contradict such hypothesis. First, entorhinal cell have been reported to have a small directionality dependence after exploration of open environments (they are indeed less directional than CA1 cells). Such a directionality persists even after removal of hippocampal input to EC. Second, Brun et al. (2002) have reported that after disruption of CA3 $\rightarrow$ CA1 projections, CA1 cells show strong directional modulation after post-surgical training on a linear track. These results suggest that the directionality properties of CA1 cells do not necessarily require a functional CA3 to be observed. However, they do not discard that recurrent connections in CA3 to be one of several mechanism contributing to define the directional properties of the hippocampal place

code.

Chapter 4

Allothetic spatial representation

A location may be recognised by identifying the sensory information available at that location. According to this, place-sensitive activity can be observed in cells tuned to respond to specific environmental stimuli available at each location (Zipser, 1985). In our model, an allothetic place representation is built by storing and comparing local views. During exploration, the agent samples the available external stimuli (local view) and stores its features in local-view tuned cells (VC, for View cells). From then on, the activity of a VC will depend on the grade of similarity between the local view it stores, and the currently available stimuli.

Taking into account that several local views may correspond to the same location (for example, views taken at the same place with different gaze orientations), VC population projects onto a second population where several local views may be combined. Cells in this population are referred to as Allothetic place cells, APCs.

This chapter describes how the local views are encoded in the VCs of our model (Section 4.1) and combined in the APC population Section 4.2. Place-sensitivity properties of the APC activity are shown in Section 4.3.1, dependence on head direction and geometric properties of the environment in Section 4.3.2 to 4.3.4.

4.1 Encoding local views

External sensory cues, and specially visual cues seem to play an important role in the establishment of the hippocampal place code (see Section 3.2).

Similar to the model of Hartley et al. (2000) a spatial representation based in metric information to the walls is proposed. At each location, the local view is

coded as a vector $\vec{d}$ of N_{dir}. Each component d_ϕ corresponds to the distance to the closest wall in the allocentric head direction ϕ. Let Ψ be the visual field of the agent, and ψ the current head direction, the feature vector correspond to $\vec{d} = [d_{\psi-(\Psi/2)}, .., d_\psi, .., d_{\psi+(\Psi/2)}]$, coding for N_{dir} directions.

A population of cells encoding the encountered local views is built incrementally; at each location the agent takes a view of its environment and recruits a new cell which stores the vector $\vec{d^i}$, and the agent's head direction ψ^i. This population is referred to as *view cells* (VC).

The activity of a given VC i depends on the difference of the current local view ($\vec{d}$) and the stored features ($\vec{d^m}$). When computing this difference, local views are aligned according to the direction in which they were perceived, and only the $N_\Omega \leq N_{dir}$ features in the overlapping region are taken into account for the comparison. Furthermore, the difference in the gaze direction ($\psi^i - \psi$) also modulates the response of a VC such that it will be maximally active only if the difference between the views is small, and they were taken at the same head direction. The VC activity is then computed as follows,

$$r_i^{VC} = exp\left[-\frac{1}{2\sigma^2_{\text{VC}}}\left(\frac{1}{N_\Omega}\sum_{\phi\in\Omega}(|d^i_\phi - d_\phi|)\right)^2\right] exp\left[-\frac{(\psi^i - \psi)^2}{2\sigma^2_{\text{VCANG}}}\right] \tag{4.1}$$

where the first Gaussian depends on the difference between the features of the local views, and the second depends on the difference in the head direction. ψ^i is the heading direction at the moment when the VC i was recruited, and ψ is the current heading direction.

4.2 Allothetic place cells

The VC population projects downstream onto another population where information from several local views is integrated to build a vision based place representation. This population is termed APC, for*allothetic place cells*.

Each time a new VC is recruited, a new APC is selected and its activity set to maximal rate $r_i^{APC} = 1$. Connections from all active VCs j with activity $r_j^{VC} > \theta_{VC}$ to APC i are updated to a value $w_{ij} = r_j^{VC} r_i^{APC}$. In subsequent time steps, initialised connections w_{ij} are updated using a Hebbian rule (Equation (4.2)) to allow the integration of information from several local views into a single APC.

$$\Delta w_{ij} = \eta_{APC} r_i^{APC}(r_j^{VC} - w_{ij}) \tag{4.2}$$

In consequence, an APC may receive pre-synaptic input from VCs coding for the same location but with different heading directions. In this way, the directionally-dependent VC coding may be used to code for location in an omni-directional manner.

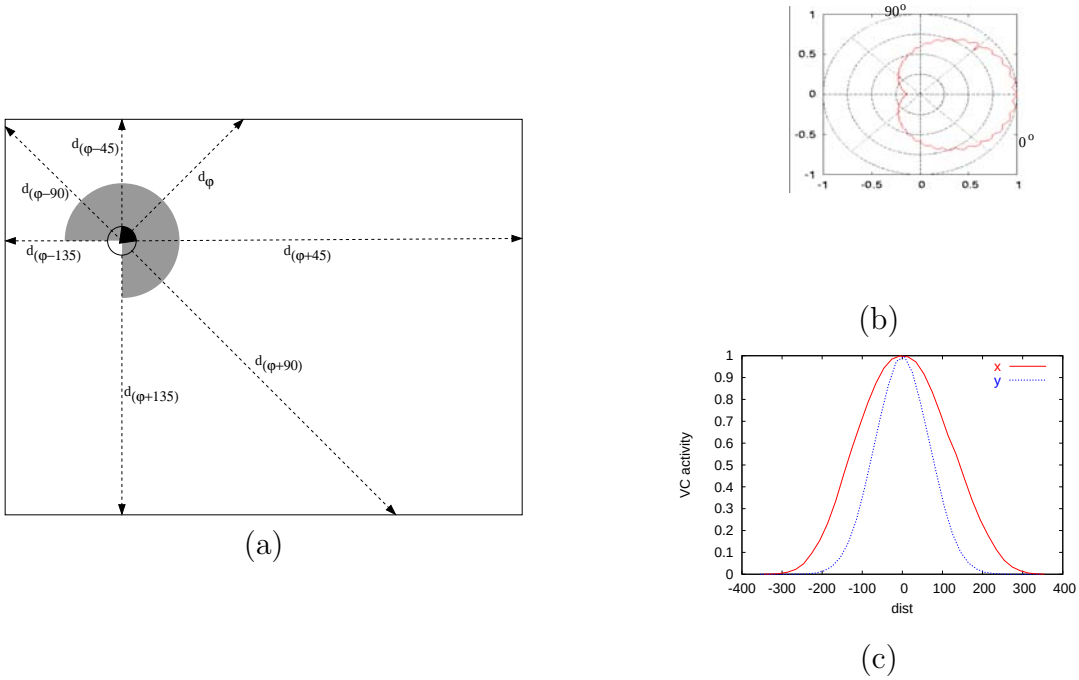


Figure 4.1: (a) Encoding of local view taken with head direction $\psi = 45°$. The cell stores the distances d_ϕ to the closest walls in different directions (ϕ) in the visual field (shaded area). d_ψ, corresponds to the feature (distance to wall) at the centre of the visual field. (b) Response of a VC ($\sigma_{\text{VC}} = 0.05$, $\sigma_{\text{VCANG}} = 90°$) with respect to changes in the heading direction. (b) Response of the same cell to movements in the vertical (y) and horizontal (x) directions (with constant heading, $\psi = 0°$).

Ideally, an omni-directional APC will receive inputs from VCs coding for all possible heading directions. However, given that the visual field Ψ of the agent is smaller than 360°, the cell may be able to fire in response to a subset of its inputs (i.e. VCs coding for stimuli inside the view field). Assuming that the APC receives input from n VCs coding for uniformly distributed heading directions, a subset of $(\frac{360°}{\Psi})n$ pre-synaptic cells may be sufficient to activate the post-synaptic APC cell.

The activity of an APC i is computed as,

$$r_i^{APC} = \left\lfloor \tau \frac{\sum_j w_{ij} r_j^{VC}}{\sum_j w_{ij}} \right\rceil \tag{4.3}$$

where $\lfloor x \rceil = x$ if $0 < x < 1$, $\lfloor x \rceil = 1$ for $x > 1$, and zero otherwise. The scale factor $\tau = \frac{360}{\Psi}$ accounts for the limited view field.

Connectivity between VC and APC depends on the simultaneous activation of cells in both populations. As mentioned above, local views corresponding to the

same location and taken at different orientation may converge onto a single APC, with uniformly distributed weights yielding an omni directional place cell firing. For this to happen, the agent should be able to explore uniformly the same location with different head directions - as is the case when the animal freely explores an environment. In contrast, if the agent only experiences local views corresponding to a small subset of orientations (as in linear tracks and directed search tasks), the APC will receive projections from VCs coding for these orientations. Therefore, the activity of this APC will be directional-dependent. This is illustrated in Figure 4.2.

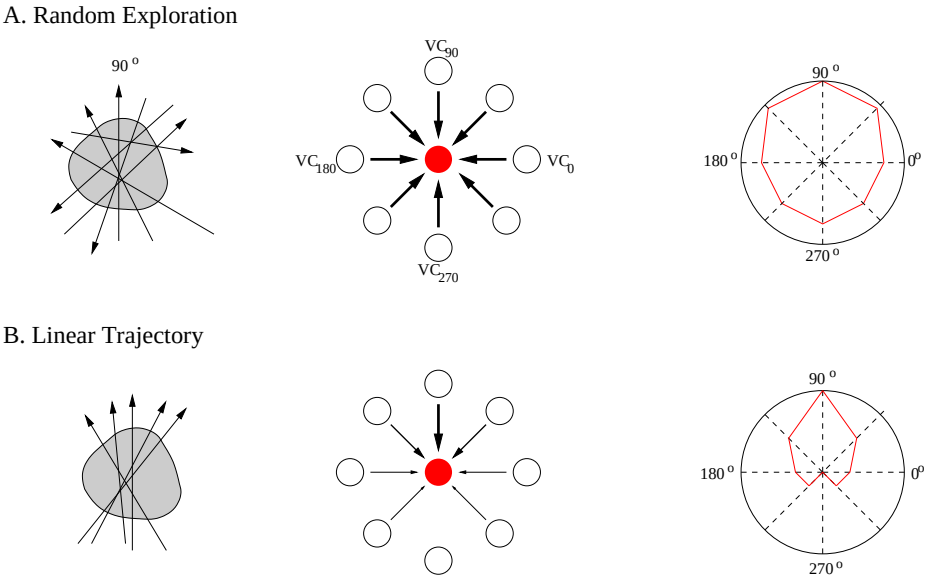


Figure 4.2: Schematic representation of VC to APC connectivity. (a) Random exploration of an open environment. *Left.* Crossing directions of an APC place field (grey region). As the agent can move freely, it is able to cross the field in any possible direction. *Centre.* Synaptic connections from VCs (open circle) to an APC (red circle). As the agent crosses the region in different directions ϕ, it can sample local views at several bearings. So that, VCs coding for such directions (VC_ϕ) get connected to the APC. *Right.* Weights of connections from VCs to APCs. Given an uniform sampling, connections strength will tend to be equal for all possible head directions. Therefore the APC response will not be directionally dependent. (b) Linear track. As the animal always crosses the place field in similar directions, the APC will only receive pre-synaptic input for VCs in a subset of possible directions, and the weight profile will lead to a strong directionality dependence.

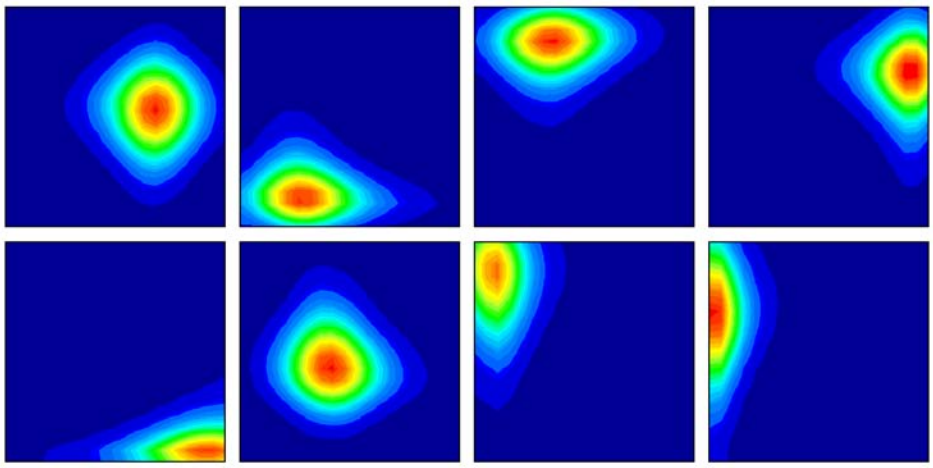


Figure 4.3: Place fields of eight APCs. Each plot shows the mean activity of an APC at different locations in the environment.

4.3 Results

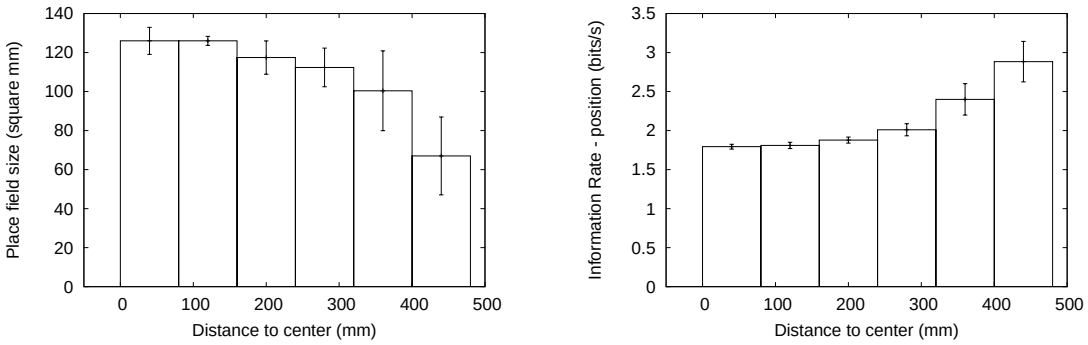
4.3.1 Spatial selectivity of APCs

To test the spatial selectivity of APCs a random exploration of a square environment (80cm x 80cm) was simulated. At each time step the animal moves a constant distance (60 mm) in a random direction and takes a local view from that location. VCs and APCs are recruited and synaptic connections are created and updated as described in the previous section.

After exploration, the place fields of a sample of APCs are measured by recording their activity at locations covering uniformly all the environment. Figure 4.3 shows the receptive fields of 8 APCs. Each square shows the colour coded mean activity of the cell for every position in the environment, where red corresponds to the maximal activity.

Place fields recorded in rats have been reported to be of elliptical or circular shape, and have been largely modelled as two-dimensional Gaussian. However, place fields located close to the walls seem to reflect the structure of the environment (as if the field *hugs* the surrounding walls (Muller et al., 1987)). Modelled place fields shown in Figure 4.3 reproduces this feature as fields located in the periphery of the environment are elongated along the neighbouring wall. Furthermore, the model predicts that place fields located near to the walls will be narrower than those located in the centre of the environment, as perceived local views will change faster if the agent is moving close to the external stimuli.

The size of the place fields was defined as the region where the cell's activity



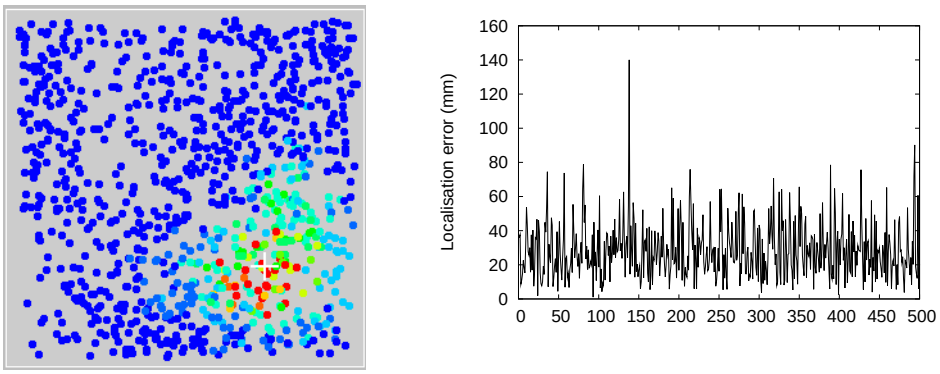
(a) (b)

Figure 4.4: Changes of place fields with respect to the peak firing location. Random exploration of an open environment. (a) The mean size of the receptive field decreases as the field location is closer to the walls. (b) Information transfer rate about position. Consistent with changes in the field size, the firing of cells coding for locations in the periphery conveys more information about location than cells coding for central locations. This suggests that the place coding is sparser in the centre of the environment (far from the stimuli).

is higher than half of the peak activity. Figure 4.4a shows the average place field size with respect to the distance of the peak firing location to the centre of the environment. It shows that cells firing at the centre of the environment (therefore far from surrounding walls) have larger place fields than those firing in the periphery. Alternatively, we measure the position information transfer rate of each cell depending on its peak firing location (Section 2.2.2). Given the larger fields at the centre of the environment, individual cell firing in that region conveys less information about position than cells firing far from the walls (Figure 4.4b).

These results suggests that the place coding in the centre of the environment (far from external cues) is less sparse than in the periphery. Accordingly, Battaglia et al. (2004) measured the population activity of CA1 cells of rats exploring a circular track in cue-poor and cue-rich environments. They report that the sparseness of the population place code (measured as the correlation of the population vector) is higher in cue-rich environments.

An estimation of the agent's location can be obtained by means of *population vector decoding*, as described in Section 2.2.4. Being p_i the location where an APC i is most active, the estimated agent location based on visual input $\hat{p}^{APC}$ corresponds



(a) (b)

Figure 4.5: (a) APC population activity. Each dot corresponds to an APC and is located at the position where the cell was recruited. The colour indicates the level of activity. The white cross marks the estimated location using population vector (Equation (4.4)). (b) APC position error after exploration. The localisation error (euclidean distance between the real location and the population estimate) was computed for 500 randomly chosen locations. The mean error for these locations is 28.5 mm (size of the environment 80cm x 80cm).

to,

$$\hat{p}^{APC} = \frac{\sum_i r_i^{APC} p_i^{APC}}{\sum_i r_i^{APC}} \tag{4.4}$$

The activity of the APC population is illustrated in Figure 4.5a. Active APCs are concentrated in a particular region of the environment (where the agent is located in), whereas cells coding for other locations are silent. The location estimation error computed as the Euclidean distance between the real location and the population estimate (computed according to Equation (4.4)) is shown in Figure 4.5b.

4.3.2 Directionality properties of APCs

During exploration APCs are connected to simultaneously active VCs, integrating information from several local views. Provided that the agent explores one location from different heading directions, APCs coding for such location may be active independently of the orientation of the agent.

As an example, assume the agent takes a local view with heading direction $\psi_i = 0°$. A new VC i is recruited to code for that view, as well as a new APC j. A synaptic connection from $w_{ji} = 1$ is also created as described above. If the agent rotates in its

place and takes a new local view at $\psi = \psi_k$ a new VC k will be recruited. At this point the activity of the previously created APC j depends on the pre-synaptic VC i (see Figure 4.1), if $r_j^{APC} > \theta_{APC}$ a connection w_{ik} will be established. The APC then, will be responsive for both heading directions ψ_i and ψ_k. If the agent keeps rotating, the APC will be connected to VPCs coding for all possible orientations ($\psi \in [0°, 360°]$), and therefore it will become omni-directional.

Figure 4.6 illustrates how the weight distribution changes depending on the exploratory behaviour of the agent. Two cases are shown: In the first case, the agent first turns in place (at each time step the agents orientation changes by $\Delta\psi = 4°$) to uniformly sample local views at all possible directions. Then, the agent's movement rule is changed such that the agent only perceives local views in head directions close to 0° (at each time step the agents orientation is randomly chosen to a value $\psi \in [-5°, 5°]$). The second case corresponds to the opposite situation, the agent first only explores a subset of head directions and then turns in place to explore views all around itself. In the first case, projections to an APC are initially uniformly distributed and, when the agent movement is circumscribed to a subset of directions, the cell becomes strongly connected to VCs coding to such directions (Figure 4.6a). This corresponds to a cell whose response becomes strongly directionally dependent. In the opposite case, a cell can reduce its directionality when the agent freely explores stimuli at all possible head directions (Figure 4.6b). The weight distribution at two different moments of the exploration (red and blue arrows in Figure 4.6b) is shown in Figure 4.6c. Initially the APC is strongly connected to VCs coding for stimuli at 0° (red line) and, after the exploratory behaviour changes it receives about equal contributions from VCs at every possible head directions (blue line).

The changes in the weight distributions determines the directionality dependence of the APC activity. As shown in Figure 4.6, the directionality of a cell depends strongly on the exploratory behaviour and can can change from non-directional to directional and vice-versa, according to these. Such changes have also been observed in rats switching between directed and non-directed tasks (Markus et al., 1995).

The model assumes that APCs are intrinsically directional at the moment they are created, and as the agent explores the environment and is able to integrate local views corresponding to different heading directions, the cell's response becomes less dependent on the heading direction. According to this assumption, the directional properties of place cells are influenced by the exploratory behaviour of the agent. As confirmed by experimental data showing that place cells recorded in open environments are less directional than those recorded in linear tracks and mazes.

The directionality dependence of place cells created in the random exploration described above was assessed using the same method proposed by (Markus et al., 1995). See Section 2.2.3.

As expected, the directionality of APCs changes depending on the experience of the cell (i.e. how well the head direction space has been explored). To measure these

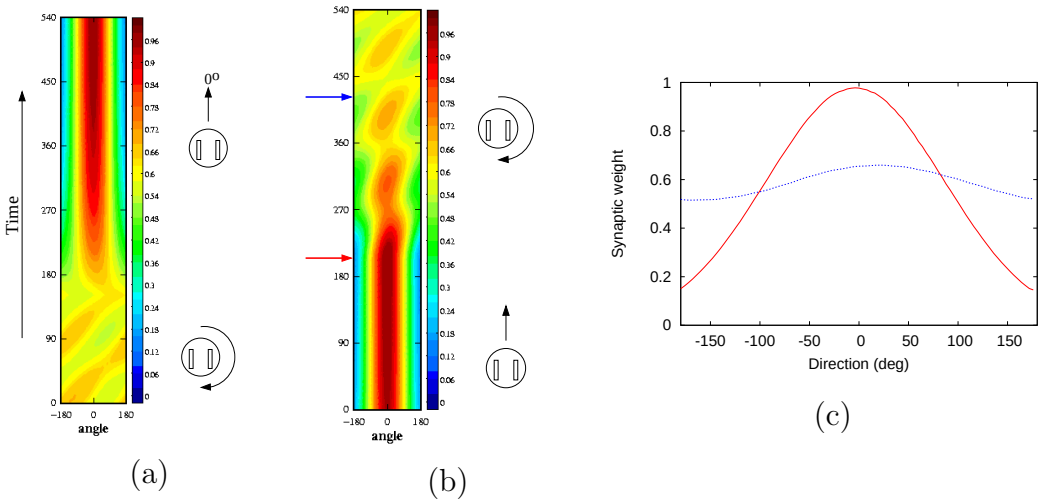
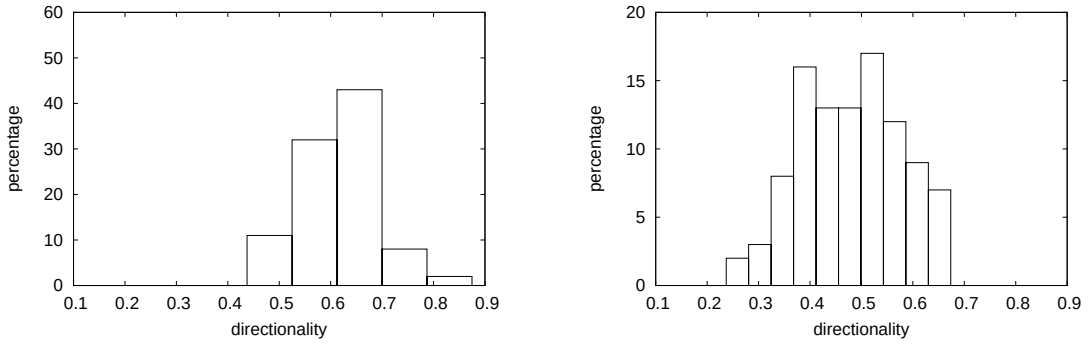


Figure 4.6: Changes in the connectivity pattern from VCs to an APC according to the exploratory behaviour. (a) The agent first samples uniformly all possible heading directions, by rotating in the same place, and (after 150 time steps) the agent orientation changes only in the range of $[-5°, 5°]$. Synaptic weights form VCs to an APC change accordingly from being uniformly distributed to a localised distribution around 0°. (b) In the opposite case, if the exploration changes from a localised to an uniform sampling, the weight distribution can become uniformly distributed again. (c) Weight distribution at two stages of the exploration shown in (b), the red and blue plot correspond to the weights at the time steps marked by the red and blue arrows in (b), respectively. *Red line*, after only sampling a subset of the head directions, the APC is strongly connected to VCs coding for local views close to 0°. *Blue line*, the distribution becomes more uniform if the agent perceives local views at all possible head directions.

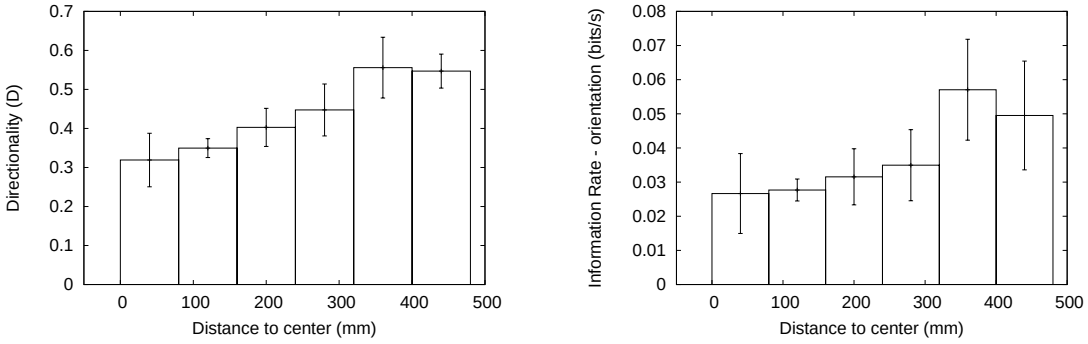
changes Figure 4.7 shows the directionality histogram of 100 APCs created before the first 500 time steps. Two histograms are shown, Figure 4.7a shows the directionality of these cells after 500 time steps, whereas Figure 4.7b shows the same cells after 10000 time steps. Clearly the activity of these cells reduce its dependence of heading direction as the exploration goes on.

Both measures of directionality and orientation-information transfer rate show that orientation dependency of place cells after exploration is modulated by the cell's preferred location (Figure 4.8). Place fields at the centre of the environment are less directional than in the periphery. This can be explained by the exploratory path of the animal, as at locations close to walls the rat's head cannot point toward the centre of the arena (as the rat's body cannot penetrate the wall). These "kinematic" constraint leads to a more directional exploration of the periphery of the environment (Taube et al., 1990; Muller et al., 1994).



(a) (b)

Figure 4.7: Directionality histogram at two moments of a random exploration. A sample of 100 APCs recruited early on during training (before of 500 time steps) is shown. (a) Directionality at time step $t = 500$ (mean = 0.59, SD = 0.14). (b) Directionality at time step $t = 10000$ (mean = 0.48, SD = 0.1).



(a) (b)

Figure 4.8: Random exploration of an open environment. (a) Place field directionality after exploration vs distance to the centre of the environment. (b) Orientation-related information conveyed by the cells.

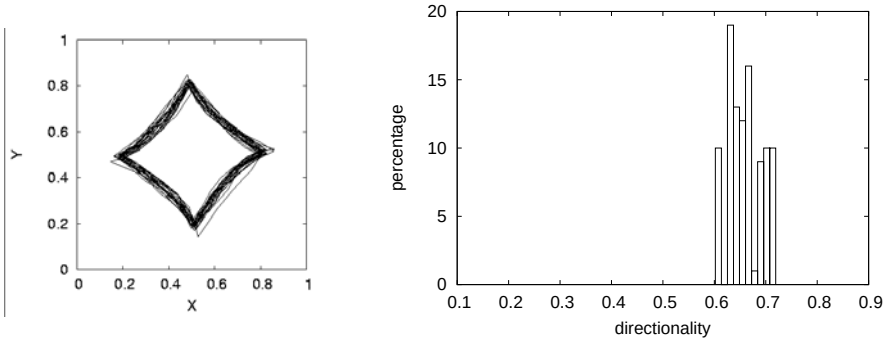
As shown before, the directionality dependence of APCs decrease when the agent explores freely an open environment. Conversely, when the agent moves in a linear trajectory (ex. arms of a maze, or a directed search in an open environment) its movements are restricted to two directions of movement along the track. As the agent only samples a subset of (opposite) heading directions, it may not be able to integrate information for several local views. In consequence, APCs should remain directional if the agent follows a linear trajectory.

A simulation of a linear trajectory (*directed search*) was performed in the same environment as above. The agent visits sequentially four points located near the middle of the walls. The movements of the agent are governed by the following expression,

$$\phi(t+1) = \phi^{goal}(t) + N \tag{4.5}$$

where $\phi(t)$ is the direction of movement at time t, $\phi^{goal}(t)$ is the bearing of the next goal, and N is a noise factor (uniform distribution $[-15°, 15°]$). Given this exploratory behaviour, at each location the agent will tend to perceive stimuli located at gaze directions close the bearing of the point is heading to. An example of the trajectory followed by the agent is shown in Figure 4.9a, a random trajectory is also shown for illustration. The directionality histogram, after exploration (10000 time steps) of 100 APCs created in the first 500 time steps of exploration is shown in Figure 4.9b.

The temporal evolution of the average directionality of place cells for both types of exploration (random and linear trajectories) is shown in Figure 4.10.



(a) (b)

Figure 4.9: Exploration following a linear trajectory in an open environment. (a) Agent's trajectory. The agent visits sequentially (counter clockwise) four locations of the environment. (b) Directionality histogram of cells recruited in the first 500 time steps (mean = 0.66, SD = 0.03).

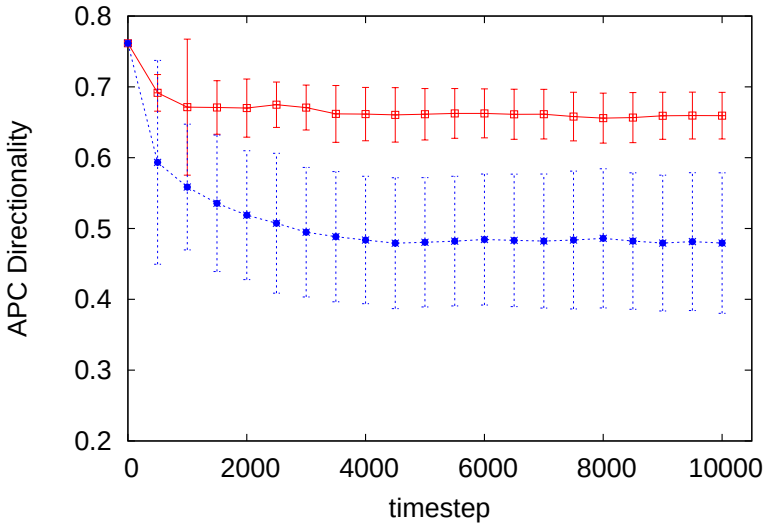


Figure 4.10: Average directionality at different stages of the exploration. *Blue*, random exploration. *Red*, directed search. The directionality value (D = 0.761) at time step zero, corresponds to the directionality index of the response of a single VC (see Figure 4.1).

The directionality of place cells can also been assessed by measuring how much information they convey about location and orientation (See Section 2.2.2). Figure 4.11 shows the average rate at which the sampled cells transfer information about position (a) and orientation (b). This figure shows that the main determinant of cell activity (i.e. the variable they convey the most information about) is the location of the agent. Furthermore, place cells created during a random exploration convey less information about the agent's orientation than cells recruited during a directed trajectory, confirming the increased directionality dependence of the later.

Figures 4.12 and 4.13 show the receptive fields of place cells after random and directed exploration respectively. External plots show the place field corresponding to eight different head directions. The centre plot shows the average activity over all directions.

4.3.3 Exploring radial mazes

Radial mazes are one of the most common paradigms in animal experimentation. Like in other tracks cells tend to be directional, but non-directional cells have been reported in the centre of both 8-arm (Muller et al., 1994; Markus et al., 1995) and plus mazes (Markus et al., 1995). Opposite to the arms of the maze, when the animal

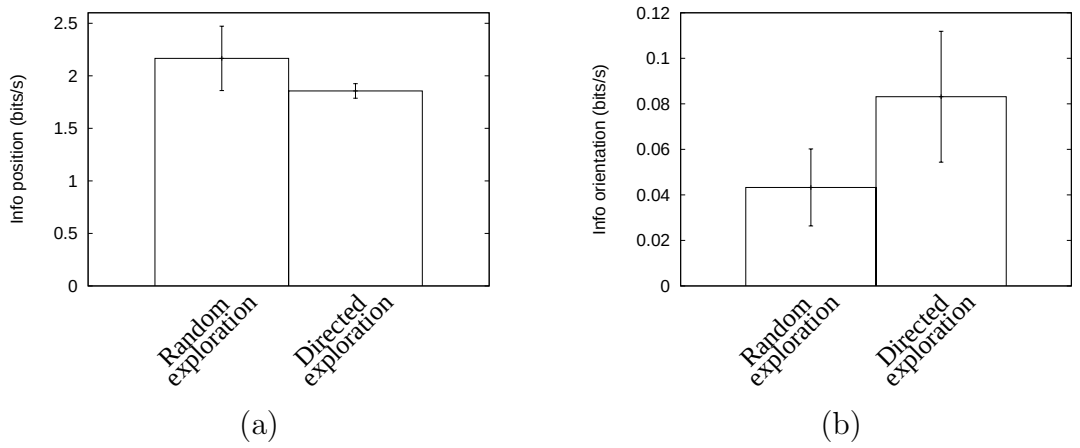


Figure 4.11: Information rate transfer of APC firing after random and directed explorations. (a) Information about position. (b) Information about orientation.

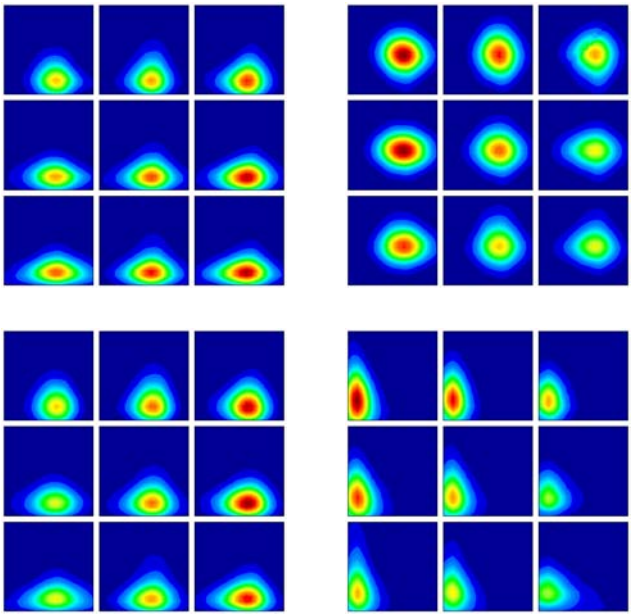


Figure 4.12: Receptive fields of APC cells after a random exploration.

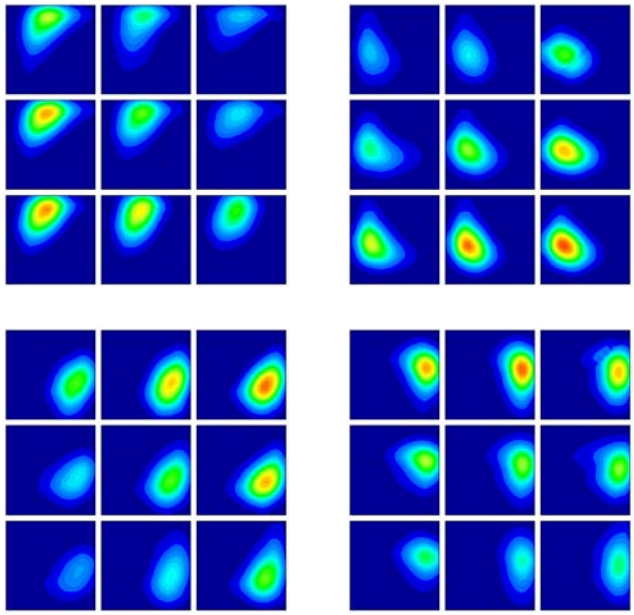


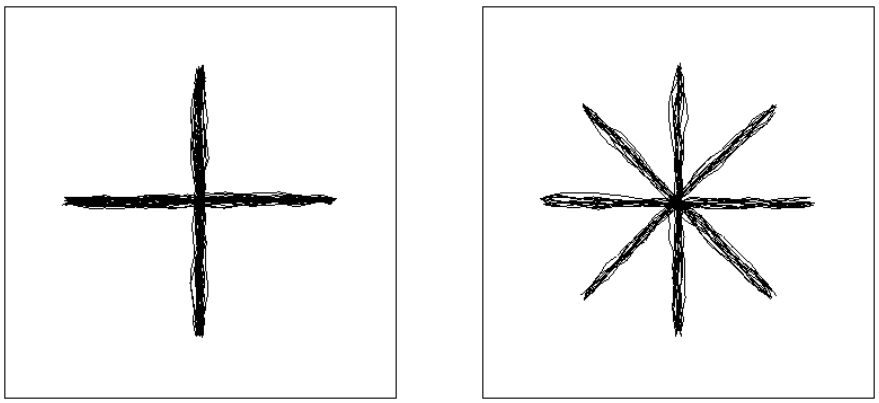
Figure 4.13: Receptive fields of APC cells after a linear exploration.

crosses the centre it is able to do it in several directions. This allows the integration of different local views, reducing the directionality dependence of the cell.

Simulations of the trajectory in a maze were performed in the same way as the linear trajectory in the previous section. A series of points corresponding to the ends of each arm and the centre of the maze was defined, and at each time step the agent moves toward one of these points. The maze has a diameter of 122 cm. Trajectories in both mazes are shown in Figure 4.14.

A sample of 250 cells is selected such that their peak firing locations are uniformly distributed across the environment. The directionality histogram of the sampled cells is shown in Figure 4.15. As observed in linear trajectories, the population is highly directional in both the plus and the 8-arm maze. The Figure also shows the directionality histograms reported by Markus et al. (1995).

As a result of the model, cells coding for locations in the arms of the maze are highly directional, whereas both directional and non-directional cells were found in the centre of the maze. Figure 4.16 shows the colour coded directionality of sampled cells depending on their peak firing location. Cells activity changes from being highly directional in the ends of the maze to being non-directional close to the centre of the maze (average directionality at the centre of the mazes is $D_8 = 0.368$ in the 8-arm maze and $D_4 = 0.409$ in the plus maze).



(a) (b)

Figure 4.14: Exploration of radial mazes. (a) Plus maze. (b) 8-arm maze.

4.3.4 Geometric determinants of APCs

As described in Section 3.2 place field changes accordingly to geometric manipulations of the environment. Based on their experimental results, O'Keefe and Burgess (1996) proposed a model in which place cell firing depends mainly on the geometrical properties of the environment (i.e. distance to walls or landmarks). See Section 3.3.2.

As in those models, the sensory system proposed in this chapter also takes into account the geometrical properties of the environment and therefore, is able to reproduce the changes in receptive fields as a consequence of manipulations of the environment. We test the same conditions described in (O'Keefe and Burgess, 1996): a small square (SS, 61x61cm.); a vertical rectangle (VR, 61x122cm.); a horizontal rectangle (HR, 122,61cm); and a large square (LS, 122x122cm.). In order to asses the changes of the cells activity the agent first explores the VR environment during 4500 time steps, and then the receptive fields are measured in all four environments.

Figure 4.19 shows the place field of four simulated cells in the different boxes described above. Firing rates for each cell were rescaled such as the maximal firing rate in the training environment (VR) is equal to one. The peak activity is shown in the bottom-left corner of each plot. Cell activity is colour coded such that red correspond to the maximal firing rate. The location of the peak firing of cells shown in Figure 4.19(a) and (b) doesn't change, but the rate of the later decreases significantly in the SS and HR environments. The distance to the south wall seem to be one of the determinants of the place field of that cell. In contrast, (c) and (d) show cells whose place fields exhibit a double peak in the HR and LS environments. The shape of the receptive field suggest that the distance to both the east and west walls contribute to the place-response of these cells.

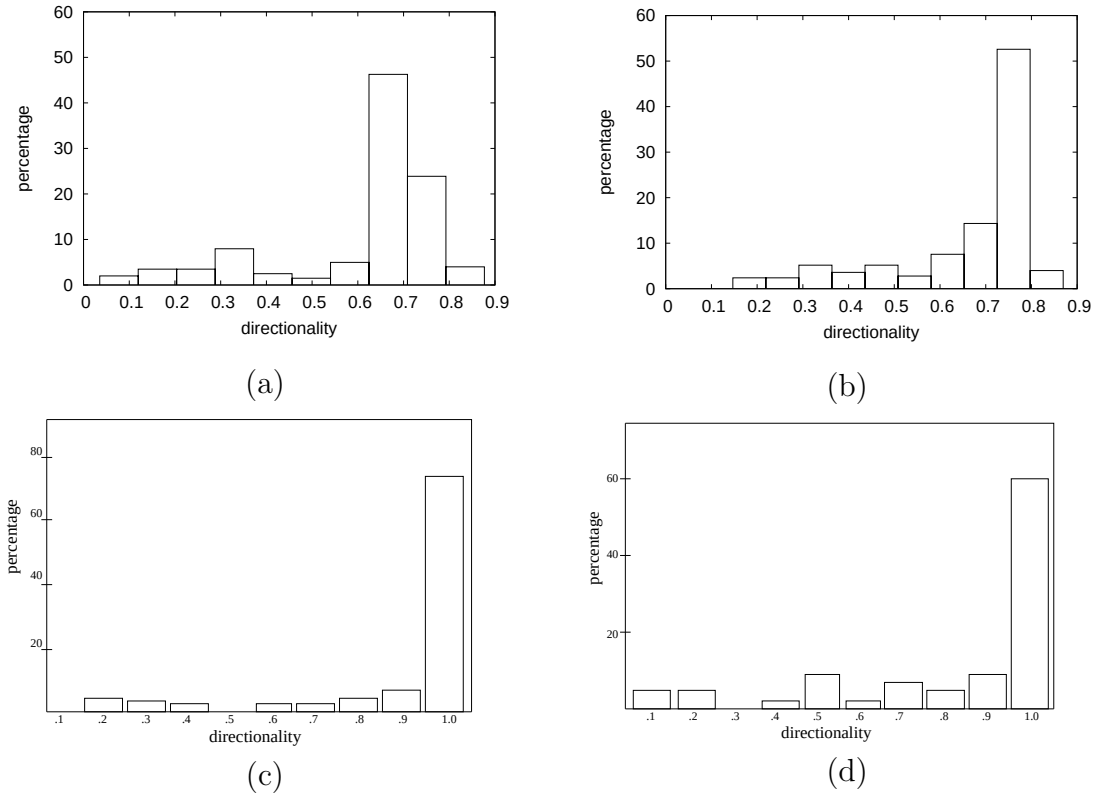
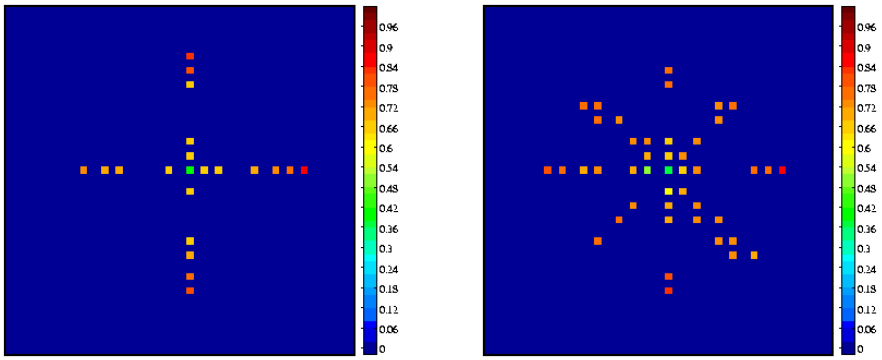


Figure 4.15: Histogram of the magnitude of directionality of place fields in our model (*top*), and cells reported by Markus et al. (1995) (*bottom*). (percentage of cells). (a)-(c) plus-maze. (b)-(d) 8-arm maze.



(a) (b)

Figure 4.16: Directionality of place fields depending on the peak firing location. Each point codes for the average directionality of cells coding for that location. In both (a) plus and (b) 8-arm mazes cells firing in the centre of the maze are less directional than those firing in the arms.

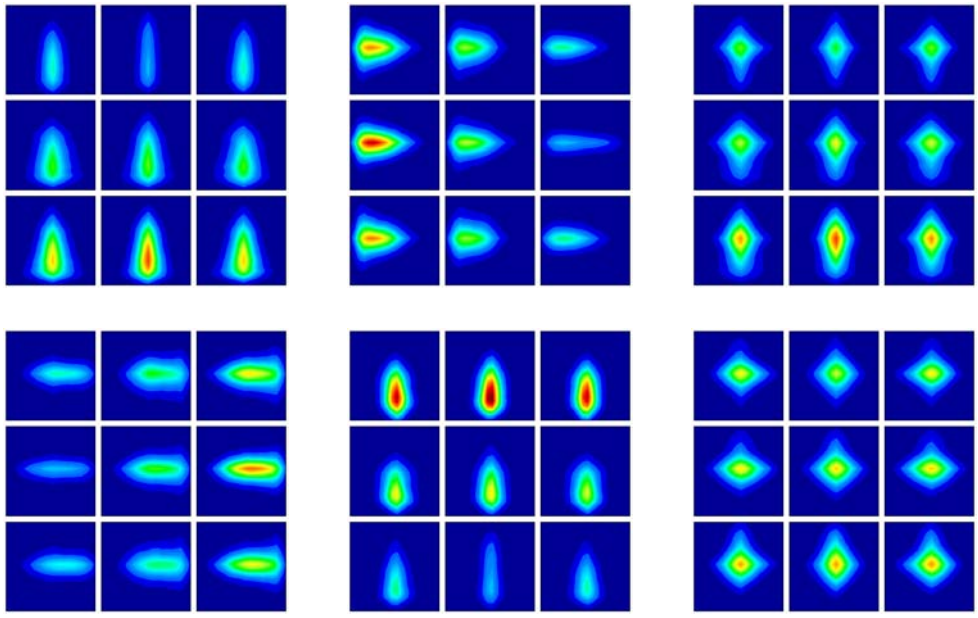


Figure 4.17: Example of 6 receptive fields of APC cells after exploration of a plus maze.

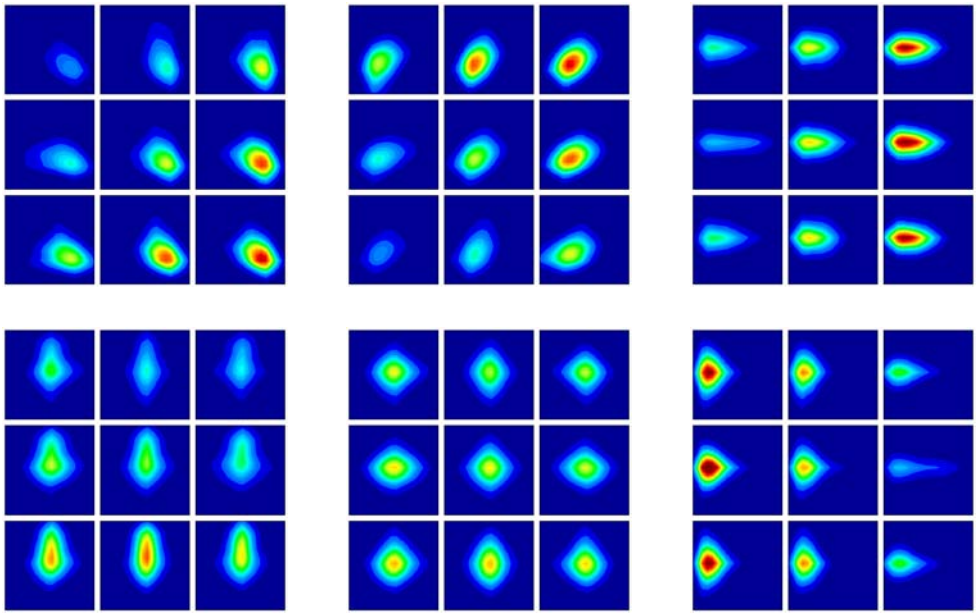


Figure 4.18: Example of 6 receptive fields of APC cells after exploration of a 8-arm maze.

Double peaked activity of place cells after changes in the geometric properties of the environment suggest that the place fields are determined by the combination of several directional subcomponents. As an example, the place field in Figure 4.19d seems to be determined by both the eastern and western wall, and in some extent the south wall. This supports the idea that omni-directional response is indeed the conjunction of directionally dependent subcomponents as previously described.

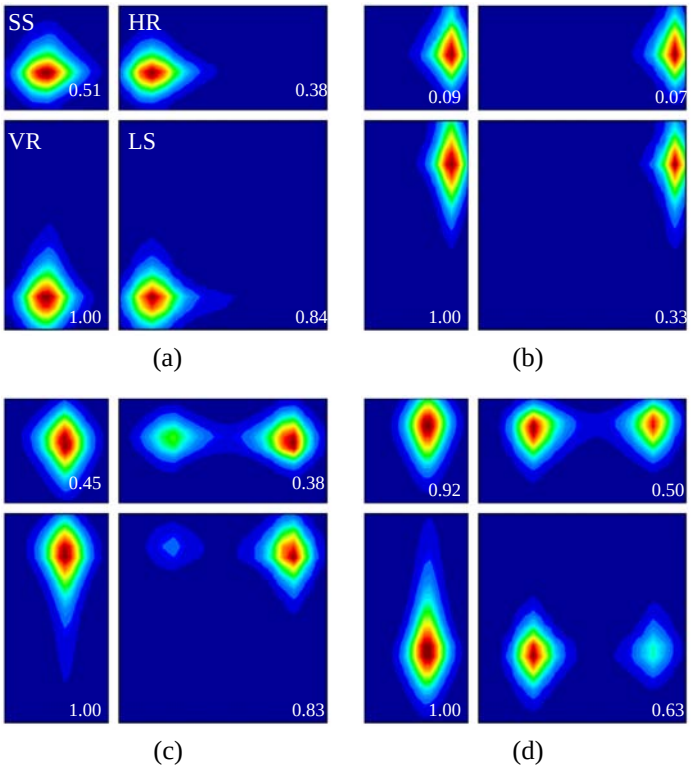


Figure 4.19: Changes of place cell activity as a consequence of geometrical manipulations of the environment. The four receptive fields correspond to the activity of the same cell in the different environments. The shape of the receptive field changes accordingly to manipulations of the environment geometry.

Chapter 5

Combined place representation

In experiments, place cells can maintain stable fields in darkness, or absence of salient visual cues. In these cases, idiothetic signals such as proprioception or vestibular information may contribute to the place coding in the hippocampus. Behavioural experiments have studied the ability of animals to navigate in absence of cues, typically they observe how the animal is capable of returning back to the nest location after exploring the environment. This ability is referred to as *path integration* or *dead reckoning*, as the name suggests it is believed that the animal integrates information about its movements in order to compute the orientation, and distance that leads to the nest location.

In our model, an idiothetic space representation composed by a population of cells driven by path integration is included. These cells, arranged according to predefined spatial preferences are algorithmically updated based on the integration of odometry signals. This representation is then combined with the allothetic representation described in the previous chapter to establish and maintain a robust, stable representation of space. A block diagram of the architecture of the complete spatial learning model is shown in Figure 5.1.

In this chapter, we first describe how the idiothetic space representation is updated (Section 5.1) and calibrated (Section 5.2). Moreover, the combined representation is detailed in Section 5.3. Finally, the results obtained in several experimental paradigms are presented (Sections 5.4.1 -5.4.3).

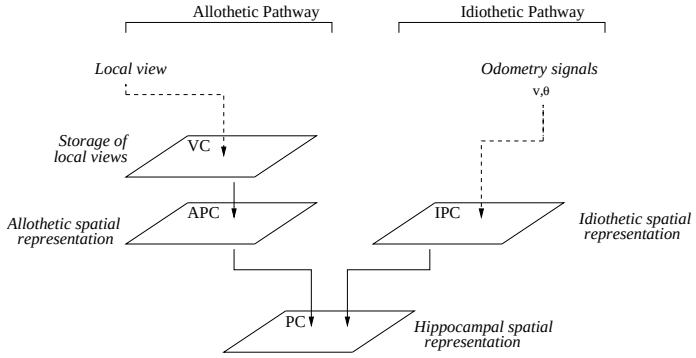


Figure 5.1: Architecture of the model of spatial learning. Dashed arrows correspond to algorithmical transformation of sensory inputs. Full arrows correspond to synaptic projections between populations of cells. These projections are updated using a Hebbian learning rule (see Equation (5.7)). VC, View cells; APC, Allothetic place cells; IPC, Idiothetic place cells; PC, Place cells.

5.1 Idiothetic place cells

As in previous models developed in our laboratory (Arleo et al., 2004; Strösslin, Sheynikhovich, et al., 2005), a location estimation based on idiothetic signals is encoded in a population of rate-coded cells (IPC, for *Idiothetic place code*). Each one of these cells has a preferred location (i.e. where its activity is maximal), and they are distributed uniformly across the whole environment. The population activity is algorithmically updated based on self-motion information.

At each time step, proprioceptive signals provide the direction of movement ϕ and the length of the movement ΔP, so that an estimation of position based on path integration can be computed as follows,

$$x^{\mathrm{IPC}}(t) = x^{\mathrm{IPC}}(t-1) + \Delta P \cos(\phi) \tag{5.1}$$

$$y^{\mathrm{IPC}}(t) = y^{\mathrm{IPC}}(t-1) + \Delta P \sin(\phi) \tag{5.2}$$

where $\mathbf{P}^{\mathrm{IPC}}(t) = [x^{\mathrm{IPC}}(t), y^{\mathrm{IPC}}(t)]$ is the estimated position at time t. Each cell $i \in$ IPC has a preferred position $\vec{p_i}$ for which it fires maximally and its place field correspond to a two dimensional Gaussian with width σ_{IPC} centred around $\mathbf{p_i} = (x_i, y_i)$. According to this and, based on the estimated position $\mathbf{P}^{\mathrm{IPC}}(t)$, the activity of the IPC population is updated as follows,

$$r_i(t) = exp\left(- \frac{|\mathbf{P}^{\mathrm{IPC}}(t) - \mathbf{p_i}|^2}{2\sigma^2_{\mathrm{IPC}}} \right) \tag{5.3}$$

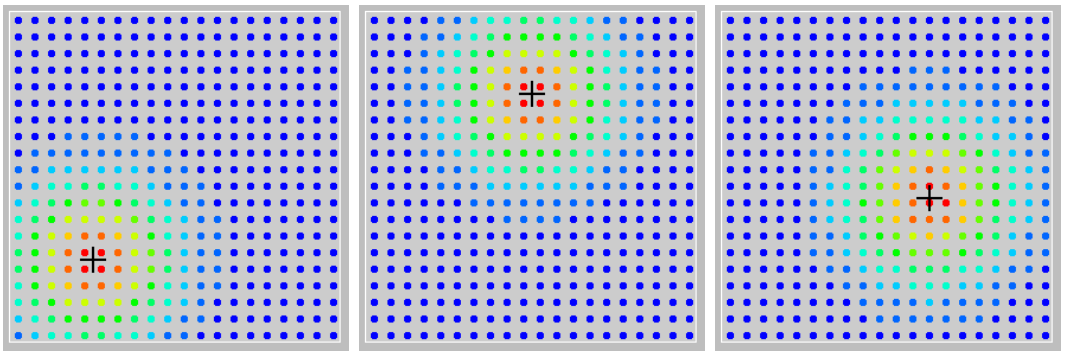


Figure 5.2: Activity of the PI population. From left to right, the plots show the population activity when the agent is located at coordinates (x=200, y=200), (x=400, y=600), and (x=500, y=350), respectively. Cells are arranged uniformly to cover the whole environment. PI activity is modelled as a two dimensional Gaussian around the cell's preferred location. Activity is colour coded from blue (silent) to red (maximal activity). The black cross shows the location estimated by computing the centre of mass of the population activity.

Figure 5.2 shows the IPC population activity when the agent is located at three different positions in the environment. The environment is a square box (80x80cm), and the cells (coloured dots) are arranged in a rectangular grid covering uniformly all the possible positions within the box. Cell activity is colour coded from blue to red, corresponding to the range from minimal to maximal firing. An estimation of the agent's position can be obtained by computing the centre of gravity of the ensemble activity (i.e. population vector decoding), analogous to Equation (4.4). In the figures, the estimated location is marked by the black cross.

5.2 Path integrator calibration

Location estimation based on path integration has an inherent problem as error signals are integrated and accumulated over time. Because of this, the estimation becomes inaccurate with time if no calibration process is performed (Save et al., 2000). External sensory input can be used to perform this calibration, as experimental findings suggest an interaction between internal and external sensory signals to control place cell firing (Gothard, Skaggs, and McNaughton, 1996; Redish et al., 2000; Save et al., 2000).

An iterative, algorithmic interaction is proposed to synchronise the location estimations made by the allothetic and idiothetic populations. Let $\hat{\mathbf{P}}^{\mathrm{IPC}}$ be the location estimation given by Equation (5.1), and $\hat{\mathbf{P}}^{\mathrm{APC}}$ be the location estimated by the APC

population (Equation (4.4)). A new estimated location is computed according to,

$$\mathbf{P}^{\mathrm{IPC}}(t) = (1-\beta)\hat{\mathbf{P}}^{\mathrm{IPC}}(t) + \beta\hat{\mathbf{P}}^{\mathrm{APC}}(t) \tag{5.4}$$

IPC activity is then updated using this new estimation ($\mathbf{P}^{\mathrm{IPC}}$) according to Equation (5.3). The parameter β corresponds to the relative significance of the external input in controlling the place code (Strösslin, Sheynikhovich, et al., 2005; Chavarriaga and Gerstner, 2004). In the simulations presented in this chapter $\beta = 0.1$.

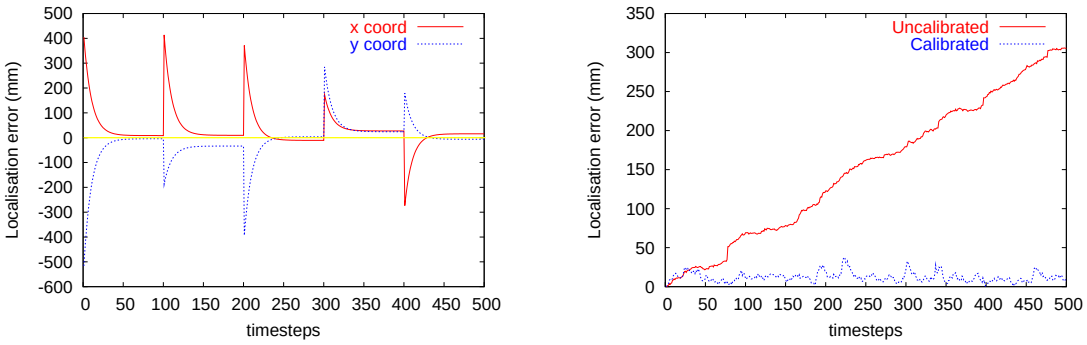
This mechanism allows the model to reset the activity of the IPC when the agent first enters the environment. At this point, internal information provides no useful information about location, and the animal has to rely on external input. A simulation of this phenomenon is illustrated in Figure 5.3a. Every 100 time-steps, the agent is located at a random location in the environment and the PI population is set to a different location. While the agent remains at the same location, the internal representation is updated according to the external input until both representations coincide. The figure shows the difference between the real and estimated location for both the x and y coordinates. It can be seen that this distance decreases with time, allowing the path integrator to become aligned with the allothetic representation.

Notice that, when entering into the environment, the head direction estimation should also be calibrated. A mechanism similar to the one explained above can be used to perform this calibration as described by Strösslin (2004). In the scope of this thesis we simply assume that the calibration of the head direction estimation takes place without modelling it explicitly.

The interactions defined by Equation (5.3) continuously update the path integrator so as to make it coincide with the estimation of the allothetic pathway. This correction compensates for the inherent drifting of the PI representation due to the accumulation of errors. In order to test this, a simulation of a random walk was performed in an environment already explored (Figure 4.5). Gaussian noise was added to the estimation of the distance travelled at each time step (mean equal to 0.5% of the estimated distance, variance equal to 1mm). Figure 5.3 shows the location error for both the uncalibrated (red line) and calibrated (blue) path integrators. The localisation error was measured at each time step as the Euclidean distance between the actual location of the agent $\mathbf{P}(t) = [x(t), y(t)]$ and the IPC population estimate $\mathbf{P}^{\mathrm{IPC}}(t) = [x^{\mathrm{IPC}}, y^{\mathrm{IPC}}]$,

$$e(t) = ((x^{\mathrm{IPC}} - x)^2 + (y^{\mathrm{IPC}} - y)^2)^{\frac{1}{2}} \tag{5.5}$$

The figure clearly illustrates that the use of external information yields a more accurate, bounded estimation of location than the uncalibrated integrator.



(a) (b)

Figure 5.3: Path integrator calibration using external input. (a) Resetting of the PI population. Every 100 time-steps the agent is located at a random location in the environment and remains there for the subsequent steps. After the movement, the idiothetic estimation is not accurate but it is gradually updated using external input to reduce the estimation error. (b) Continuous correction of the PI estimation. Localisation error of an uncalibrated and calibrated path integrator (red and blue lines, respectively). When not calibrated, the PI estimation is quickly degraded due to the accumulation of errors over time. In contrast, the proposed calibration mechanism compensate for the drifting and bounds the estimation error (localisation error $<$ 37.85mm, mean error = 12.199 mm).

5.3 Place cells : Combining allothetic and idiothetic representations

APC and IPC populations code a spatial representation based on external and internal sensory information, respectively. These representations are combined and projected onto a single population of place cells (PCs).

Similar to APCs, the PC population grows incrementally during exploration. In order to control the growth of the population, PCs are added to the model only if there are not enough cells coding for the current location. That is, at every time step t, a new PC is recruited only if,

$$\sum_i \mathcal{H}(r_i(t) - \epsilon) < A \tag{5.6}$$

where $r_i(t)$ is the activity of PC i at time step t, and $\mathcal{H}$ is the Heaviside function. Such a limitation prevents an unbounded growth of the populations of the model, while allowing a certain level of redundancy in it.

During exploration, new PCs are recruited and connected to simultaneously active AP and IP cells, as described above for the $VC \rightarrow APC$ connections. APC $\rightarrow$ PC connections are updated according to a Hebbian rule,

$$\Delta w_{ij} = \eta_{PC} r_i^{PC} (r_j^{APC} - w_{ij}) \tag{5.7}$$

where r_i^{PC} is the activity of the post-synaptic PC cell i, and r_j^{APC} is the activity of the pre-synaptic APC cell j.

The activity of a place cell is computed as the weighted average of its pre-synaptic inputs (in both the APC and IPC population),

$$r_i^{PC} = \left\lfloor \frac{\sum_j w_{ij} r_j}{\sum_j w_{ij}} \right\rceil \tag{5.8}$$

PCs combine both the idiothetic and allothetic information into a single spatial representation.

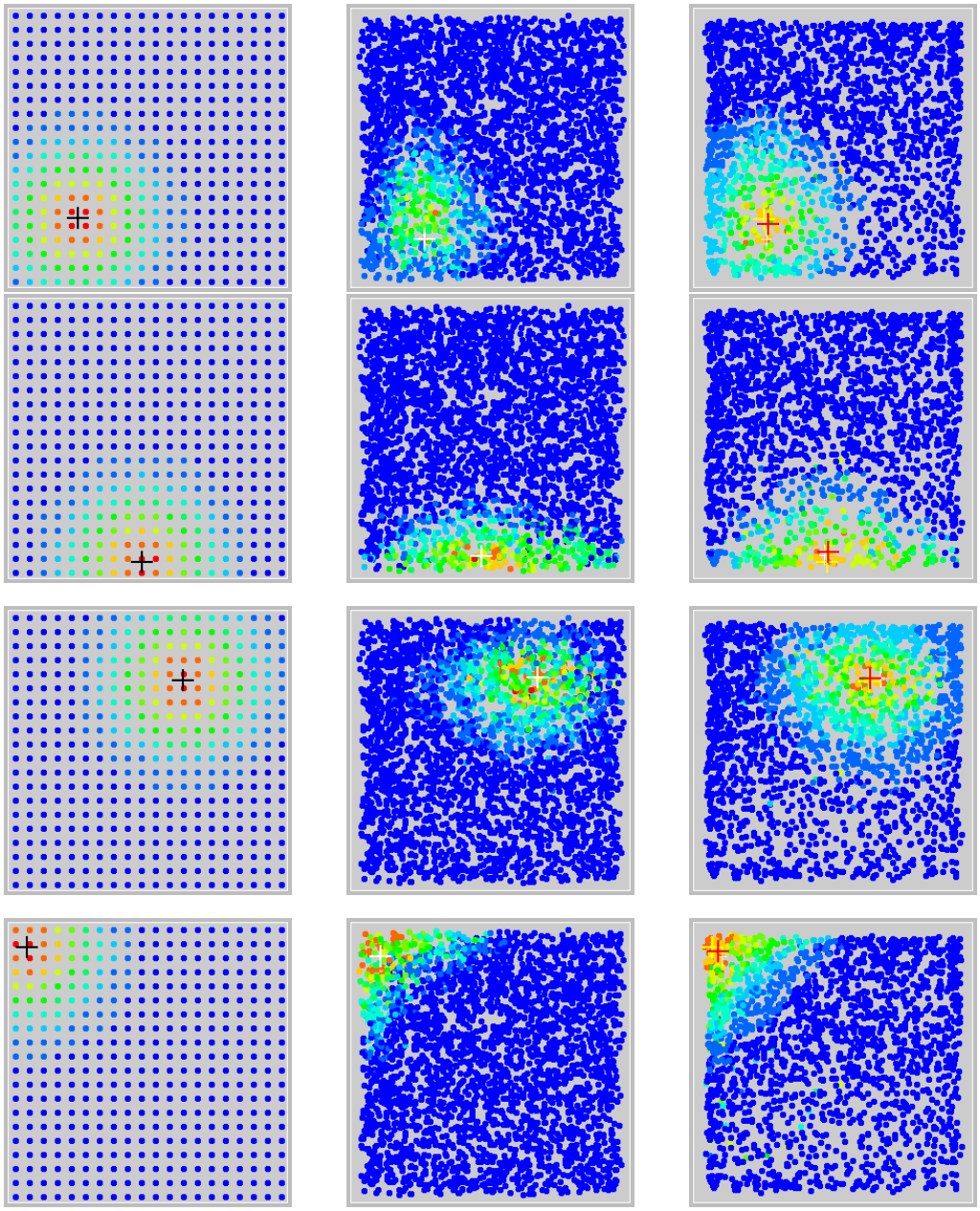
5.4 Results

5.4.1 Spatial selectivity of PCs

A random exploration of a square environment (80cm x 80cm) was performed. The population activity of the PI, APC and PC populations is shown in Figure 5.4, for four different locations of the environment (coordinate (0,0) corresponds to the bottom-left corner of each plot). In each plot, a coloured dot represents one cell and is located at the centre of the cell's place field. Activity is colour coded from blue to red corresponding to range from minimal to maximal activity. The figure illustrates that (i) only a localised sub-set of the population is active at any moment, and (ii) the ensemble population activity provides an accurate estimation of the agent's location.

Place fields of simulated cells after a random exploration are shown in Figure 5.5. Place fields show circular or elliptical shapes and, as observed with APC place fields, when the field is located close to the walls, it elongates along the wall following the structure of the environment.

Place cells in this module combine the non-directional representation from the idiothetic pathway with the representation from the allothetic stream. Correspondingly, changes in the directionality properties of the allothetic place cells (Section 4.3.2) will be reflected at this stage. That is, random exploration of open environments will result in a population of place cells which are non-directional, whereas a linear path yields directionally dependent cells. The histogram of directionality indexes of the combined place cells for both a random exploration and a directed trajectory is shown in Figure 5.6.



(a) (b) (c)

Figure 5.4: Population activity when the agent is at three different locations in the environment [coordinates from top to bottom (200,200); (380,50); (500,600) (20,740)]. (a) PI cells. (b) Allothetic place cells (3000 cells). (c) Place cells (2000 cells). Each dot represents one cell and is located at the centre of the cell's place field. Activity is colour coded from blue (silent cells) to red (maximal activity).

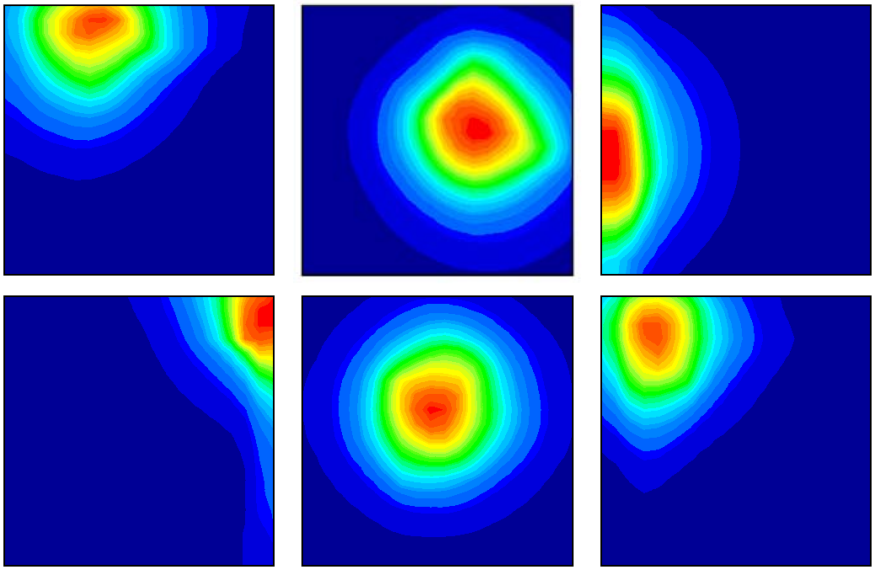
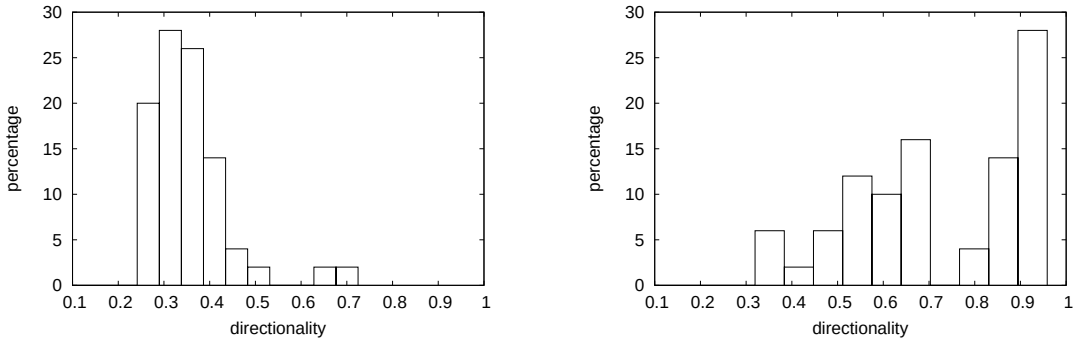


Figure 5.5: Place fields of eight place cells after a random exploration of an open environment.



(a) (b)

Figure 5.6: Directionality of place cells. (a) Random exploration (D, mean = 0.35 SD = 0.102). (b) Directed exploration (D, mean = 0.70 SD = 0.212).

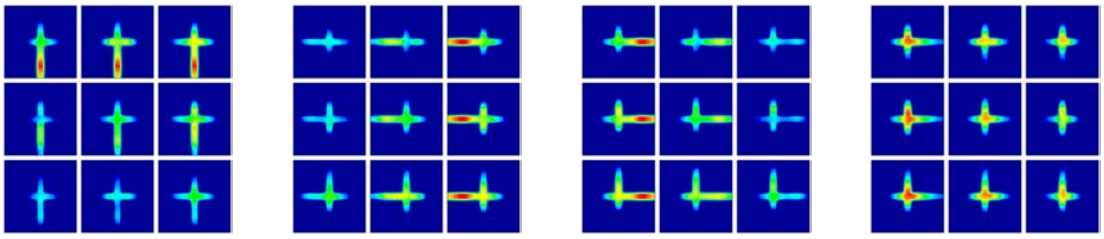


Figure 5.7: Directional contours of four place cells exploring a plus maze.

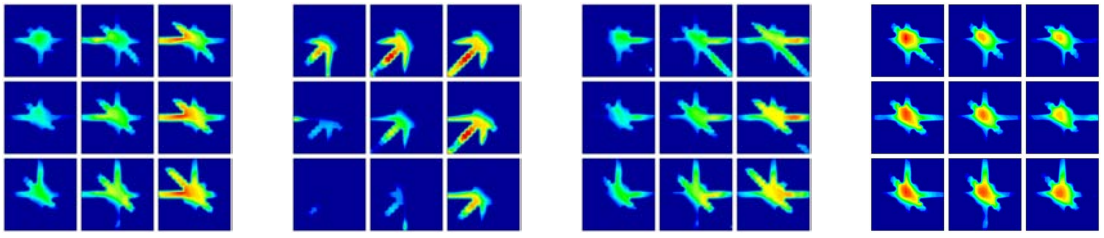


Figure 5.8: Directional contours of four place field after exploring a 8-arm maze.

5.4.2 Exploring radial mazes

As described in Section 4.3.3, exploration in radial mazes was simulated. Examples of resulting place fields are shown in Figure 5.7 and Figure 5.8 for plus- and 8arm-mazes, respectively. Figure 5.9 shows the directionality histograms after exploration of both types of maze.

As observed in the APC population, cells in the centre of the maze are less directional than those in the arms. Both the directionality index proposed by Markus et al. (1995), and the information content about directionality reflect such changes (Figure 5.10a-b). In contrast, the information content about position is about equal for cells at any location in the maze (Figure 5.10c).

Comparing the directionality histogram for the modelled cells with the results reported by Markus et al. (1995), we observe that the modelled population tend to be less directional than experimentally recorded cells. A significant part of the less-directional cells fire at location close to the centre of the maze.Markus et al. (1995) provide no detailed information about the place field location of the recorded cells which does not allow us to directly compare the results shown in Figure 5.10. The contribution of the non-directional idiothetic pathway may also reduce the directionality of PCs. To our knowledge no systematic study has been performed to study the changes in the directionality of place cells in absence of visual cues.

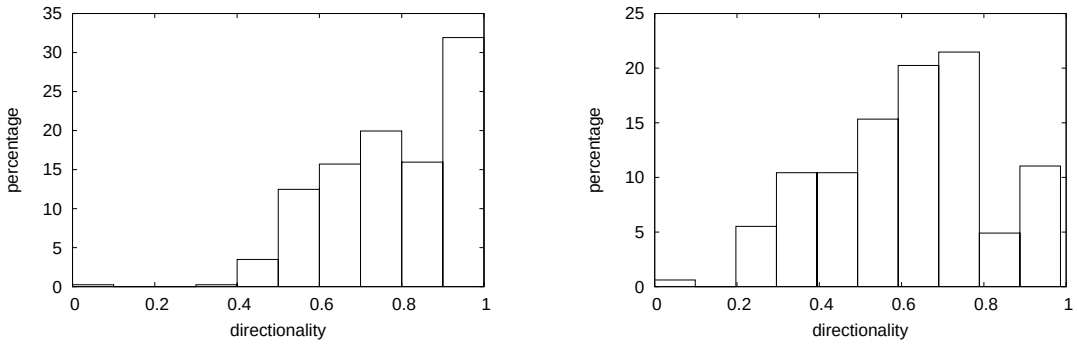


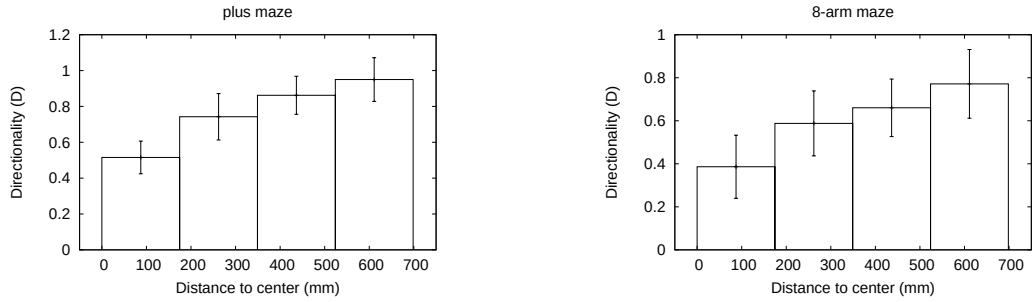
Figure 5.9: Directionality histogram (a) plus maze (D, mean = 0.72 SD = 0.197). (b) 8-arm maze (D, mean = 0.61 SD = 0.195).

5.4.3 Interaction between internal and external cues

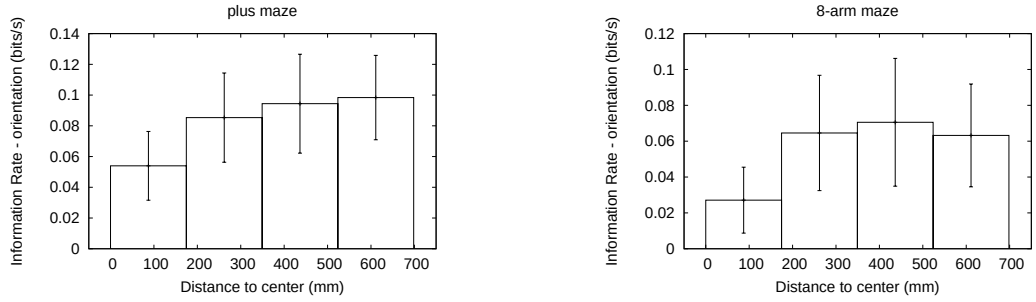
In the proposed model two types of interactions between internal and external cues are proposed. (i) idiothetic and allothetic place coding populations project onto the place cell population. This allows the disambiguation of visually similar locations by use of internal cues; as well as reduce the effect of drifting of the path integration in the combined place code. (ii) idiothetic coding is aligned with the allothetic estimation (see Section 5.2). Using this calibration process the IPC can be reset when the agent enters into the environment.

The interactions between internal and external cues were studied by Gothard, Skaggs, and McNaughton (1996). They trained rats to shuttle back and forth in a linear track with a movable box on one end and a fixed reward site on the other end. The box remains at the same location throughout the entire training period (box1). During testing, the size of the track was changed by moving the box during the outbound journey of the rat (when the box is behind the rat, and presumably outside of the rat's visual field). The box was randomly moved between five locations (box1 to box5).

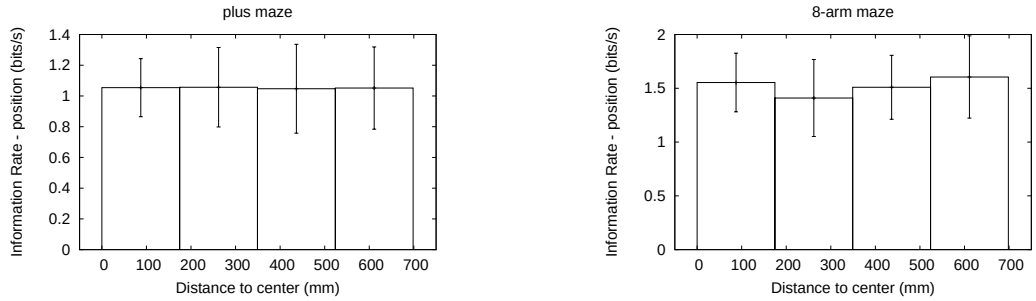
Cell recordings show that in the initial part of the journey cells fire at fixed distances to the point of departure, whereas toward the end of the journey, cells were aligned to the destination. This may suggest that initially, cells firing is mostly driven by path integration and, as the animal runs along the track the animal perceives a mismatch between external and internal cues. This mismatch prompts a correction mechanism, which aligns the representation to the external cues (perceived destination end). Similar experimental results were also reported by Redish et al. (2000).



(a)



(b)



(c)

Figure 5.10: Directionality firing properties of PCs depending on the peak firing location after exploration of a plus (*left*) and 8-arm maze (*right*). (a) Directionality index. (b) Information about orientation. (c) Information about position.

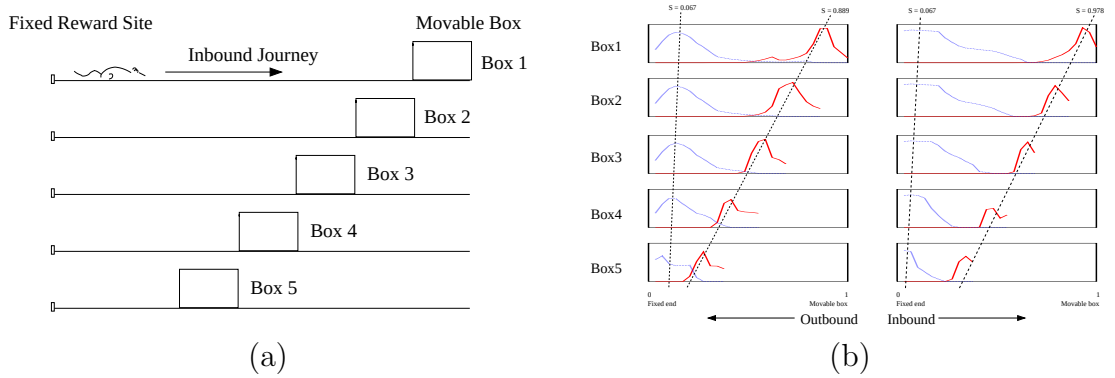


Figure 5.11: (a) Experimental setup proposed by Gothard, Skaggs, and McNaughton (1996) The rat is trained in a linear track to alternate between a fixed site and a movable box (box1). The rat is later tested in the same track, but the box is moved in every lap between five positions (box 1 to box5). (b) Place fields of four modelled place cells for each track configuration.

The receptive fields of four PCs is shown in Figure 5.11, after a simulation using an experimental protocol analogous to the one used by Gothard, Skaggs, and McNaughton (1996). Two of the cells fire during the inbound journey and the others fire during the outbound journey. The location of its receptive fields changes accordingly to those reported by Gothard, Skaggs, and McNaughton (1996). That is, at the beginning of the journey, cells tend to fire at a fixed distance from the starting point, while cells at the end of the journey seem to be fire at fixed distances from the destination point (signalled by allothetic signals).

Changes in the place cell firing are quantified by computing the *displacement slope* (S) of each cell with respect to the different configurations (box1 to box5) (Gothard, Skaggs, and McNaughton, 1996). The slope corresponds to a linear fit of place field shifts in the box2 to box5 conditions with respect to box1. Place field shifts are determined by the maximum cross-correlation of the place field with respect to the box1 condition, and the magnitude is normalised such as a displacement of 1 corresponds to a displacement equal to the displacement of the box. A slope of 0 corresponds to a place field which remains at the same location with respect to the fixed end of the maze independent of the box location.

The slope of place cells with respect to the their peak firing location in the box1 condition is shown in Figure 5.12. In the x axis, 0 corresponds to the fixed end, and 1 to the box location. It shows that changes in the displacement slope of the cells changes consistently with respect to their peak firing location. Cells firing close to the fixed site on the track ($x \rightarrow 0$) tend to have a slope close to 0, and the slope

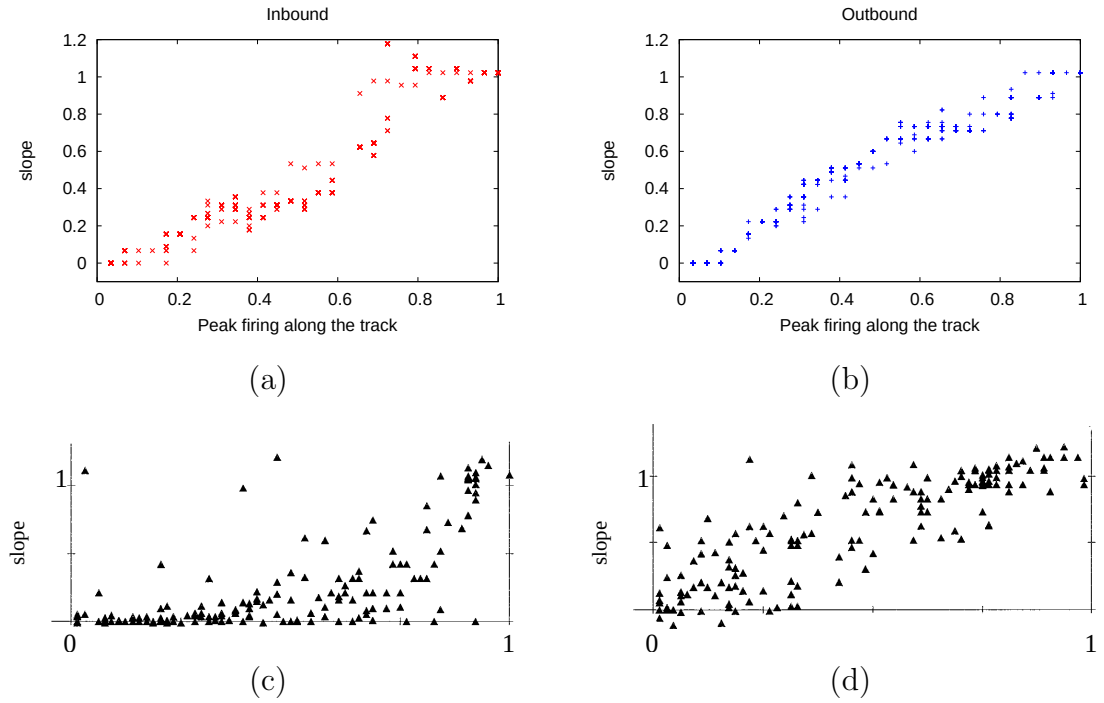


Figure 5.12: Displacement slope of place cells versus the peak firing location in the box1 condition. Slopes are normalised to 0 when the place fields do not change, and 1 when the place field shifts the same distance as the box. Cells firing close to the box tend to fire at a fixed distance to the box. *Top.* Simulated cells obtained with our model. (a) Inbound journey. (b) Outbound journey. *Bottom.* Experimental results reported by Gothard, Skaggs, and McNaughton (1996). (c) Inbound journey. (d) Outbound journey.

increases as the cell's maximal firing approaches the box location ($x \rightarrow 1$).

The population activity of place cells is assumed to code for a representation of the location of the animal. The similarity of the representation for each box condition can be observed by computing the correlation of the population vectors for each condition with the population activity in the full length track (box 1 condition). This correlation is shown in Figure 5.13. In all conditions, population activity at the beginning and end of the track shows a high correlation (coded in red). During the inbound journey, a discontinuity can be observed in the box4 and box5 conditions. This can be interpreted as an abrupt change in the frame of reference controlling the place representation. In contrast, a more gradual change is observed during the outbound journey. In the former case, the visible presence of the box in the

middle of the journey increased the mismatch between internal- and external-based representations prompting a strong change in the combined place code.

The precise transition point between the box-aligned and the room-aligned frames of reference depends in our model on the mismatch of the two representations. In our simulations, slopes in both the inbound and outbound journey change linearly in contrast to the results of Gothard, Skaggs, and McNaughton (1996) (Figure 5.12c-d). In their case, they report that the pattern of the slope curves resemble a *hysteresis* loop, suggesting that place cell activity is primarily driven by previous activity (i.e. path integration), than incoming sensory input. However, other factors like the saliency of external stimuli may affect the process re-alignment of the spatial representation. In their experiment, Gothard, Skaggs, and McNaughton (1996) made no manipulation of the external stimuli to assess if the same pattern of hysteresis was observed in different environment conditions.

Furthermore, Redish et al. (2000) suggest that such a persistence may be generated by the extensive training in the box 1 condition. They use a slightly different experimental protocol in which the box changes positions during training. They report that the displacement slopes change linearly for both inbound and outbound and propose that the realignment process occurs consistently after a temporal delay, as opposed to taking place at some specific point during the journey. This delay changes from one animal to the other, but is consistent for every individual. The existence of such a delay, they claim, supports the idea that a dynamic alignment process takes place among the internal and external-based representations in order to define the spatial coding. The calibration process depicted in Section 5.2 qualitatively reproduces the realignment of the place code in case of mismatch between the internal and external cues.

5.5 Discussion

The model developed here combines two separate representations of space driven by internal and external sensory information onto a single robust representation of space. A direct interaction between the allothetic and the idiothetic representation compensate for accumulating errors in the latter and allows the resetting of the path integrator upon re-entry in known environments. As in previous models developed in our laboratory, place cells populations grow incrementally during exploration.

Simulations of exploration of open environments, linear tracks and radial mazes were performed, in order to assess the spatial selectivity and directionality dependence of place cells (in both APC and PC populations). The interaction between internal and external cues was also observed following an experimental paradigm analogous to those used in animals (Gothard, Skaggs, and McNaughton, 1996). The parameters used in the simulations are shown in Table 5.1. Let us summarise our results of

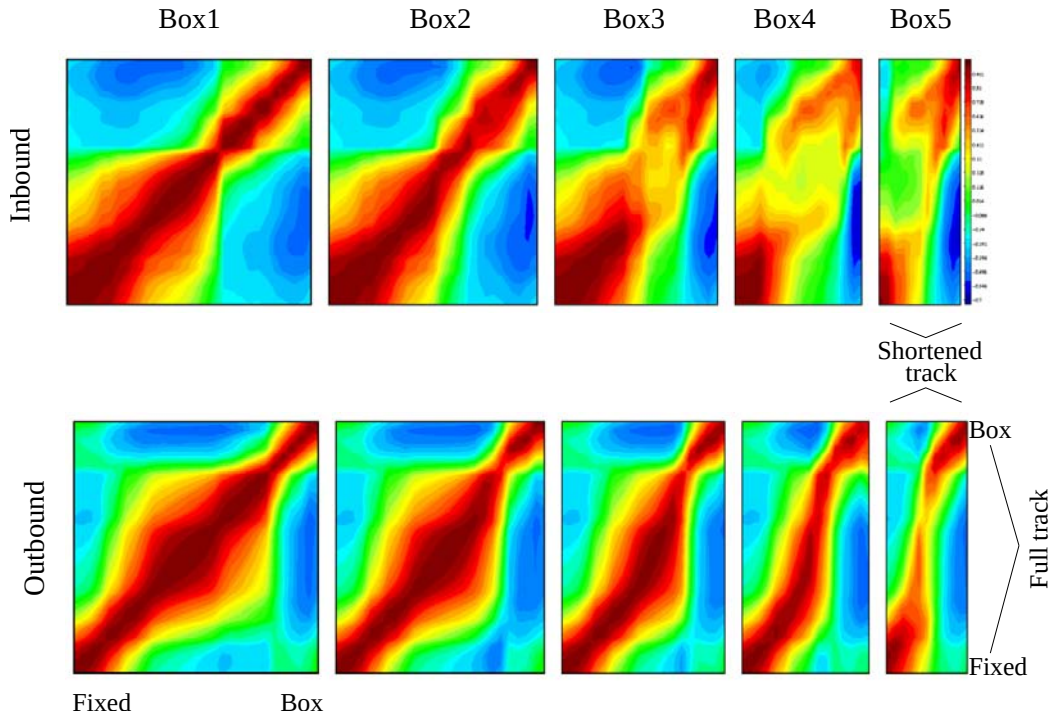


Figure 5.13: Population vector correlation between the firing pattern in the full and the shortened track for both the inbound (*top*) and outbound (*bottom*) journeys. Vertical axis corresponds to locations in the full length track (box1 condition). Horizontal axis represents location in the different testing situations (box1 to box5). Highly correlated areas are shown in red, and denote a correspondence between the firing pattern in the full and shortened tracks. The first plot on the left of each row (box1) is an autocorrelation of the population activity in the box 1 condition.

Chapters 4 and 5:

Spatial selectivity. Cells in the model show localised, overlapping place fields and the activity of the population assembly forms a distributed, redundant representation of space. Such a representation can be interpreted by means of population vector decoding.

Sparseness of the place code. The information content about position of the APC activity is modulated by the presence of available external cues. Figure 4.4 shows that after exploration of open environments, place fields located near to the walls are smaller and convey more information about the agent's location than those located in the centre of the arena. Such a change can be explained in terms of changes in the perceived features of the environment. i.e. when the animal moves close to a wall features in the local view change more rapidly than when it is located far from

Allothetic place coding

Parameter	Value	Description
Ψ	275°	Visual field
N_{dir}	21	Number of directions encoded in the VCs
σ_{VC}	0.05	Sensitivity of the VCs
σ_{VCANG}	90°	Angular sensitivity of the VCs
θ_{VC}	0.8	Connectivity threshold of VCs
θ_{APC}	0.7	Connectivity threshold of APCs
η_{APC}	0.01	Learning rate of the APCs

Idiothetic place coding

Parameter	Value	Description
σ_{IPC}	128 mm	Place field width of the IPCs
β	0.1	PI calibration parameter

Combined place coding

Parameter	Value	Description
θ_{PC}	0.8	Connectivity threshold of PCs
η_{PC}	0.1	Learning rate of the PCs

Table 5.1: Main parameters of the model

the periphery. Our model agrees with experimental results which suggest that the sparseness of the hippocampal place code is influenced by the availability of external sensory cues (Battaglia et al., 2004).

It should be noted that this modulation appears in our model as a results of the processing in the allothetic pathway; opposite to other models in which this feature is hard-coded in the definition of the sensory cells (e.g. O'Keefe and Burgess (1996); Kali and Dayan (2000)).

Place cell directionality. The model reproduces the directional dependence of place cells depending on the exploratory behaviour of the animal (or agent). Random exploration of open environment yields omni-directional place cells, while linear trajectories, or exploration of mazes results in directional dependent activity. Omni-directionality, in the model, results from the combination of several directional sub-components (sensitive to specific local views).

The idea that combination of several perceived local views may result on directionally independent place cell activity in rodents has been linked to reported *spatial view fields* in primates. These cells are active when the animal looks at particular places in the environment and their firing is relatively independent of the animal's location (Rolls, 1999). He claim that, in both rats and monkeys hippocampal cells respond to the combination of external stimuli (local views) and, the wide visual field of rats allows them to integrate stimuli in several head directions resulting in place sensitive cells. In contrast, the reduced visual field of primates yields a pattern of activity which is mainly dependent on specific visual stimuli.

Our model contrasts with previous models addressing the issue of place cells directionality (Brunel and Trullier, 1998; Kali and Dayan, 2000), in which this property is a product of experience-dependent changes in recurrent connections among place cells. Such models have proposed the recurrent network in CA3 as the putative locus of directional changes in the hippocampal place code.

Several experimental data support the claim that the representation of external sensory input influences the directionality of place cells. Battaglia et al. (2004) report that the availability of multimodal local cues may contribute to an decreased directionality of place cells recorded in a annular linear track. Furthermore, the changes in the place fields as a result of geometrical manipulations of the environment described by O'Keefe and Burgess (1996) often result in stretching or division of receptive fields, as if sub-components of the place field remained attached to one of several external determinants (i.e. walls surrounding the environment).

Experimental results show that CA3 $\rightarrow$ CA1 projections are not required to develop directional dependency of CA1 cells in linear tracks (Brun et al., 2002). Moreover, mEC cells show a reduced directionality (D=0.45) in open environments after disruption of feedback projections from the hippocampus to the entorhinal cortex (Fyhn et al., 2004). Suggesting that a functional CA3 recurrent network is not required to produce changes in the directionality of both entorhinal and CA1 cells.

The mechanism described here does not discard the possible role of recurrent connections in CA3 in refining the place code and enhance the directional independence of place cells. In fact, recent data reports that the activity of place cells in the medial Entorhinal cortex become more dependent in the direction of movement after bilateral hippocampal lesions (average directionality index for 30 mEC cells change from 0.49 to 0.21, after the lesions) (Fyhn et al., 2004).

Interaction of external and internal sensory input. The model combines the information from both the allothetic and idiothetic pathways in order to maintain a robust, stable space representation. This combination is performed in two ways: (i) Both APCs and IPCs project onto a common population of place cells (PCs). (ii) Path integration is calibrated using external information. The later interaction allows the model to set the IPC activity when the agent first enter the environment (Figure 5.3), as well as reduce the effect of drifting inherent to path integration estimations.

The model also reproduces qualitatively changes in the place code as a result of mismatch between internal and external cues. As reported by Gothard, Skaggs, and McNaughton (1996) the place representation changes the frame of reference it is based on depending on the perceived differences between the allothetic and idiothetic-based representations.

O'Keefe and Burgess (1996) report that double peaked and elongated receptive fields resulting from changes in the geometry of the environment are modulated by the direction of movement. The peak of the field shifts backward with respect to the animal's movement. For example, if the place field stretches in the east-west direction, when the animal moves west the firing rate of the eastern peak is greater than when the animal moves in the opposite direction, and vice-versa. Such a change may be due to a change in the alignment of the place code. Similar to the results reported in Gothard, Skaggs, and McNaughton (1996), when the animal starts to move away from a wall its place representation may remain attached to said wall, and place cells will tend to fire at fixed distances from it.

5.5.1 Putative neural substrates

The model is composed primarily of two pathways processing external and internal sensory input. These pathways project in a feed-forward manner to the combined place representation.

Allothetic place representation. External sensory input drives this population. Perirhinal and postrhinal cortices convey information from unimodal and multimodal sensory areas to the EC. Such areas may be responsible of carrying information of stored and recalled local views alike to the VCs activity in the model.

Idiothetic place representation. This representation is composed by cells having a predefined, environment-independent topological organisation. They are algorithmically updated depending on the proprioceptive signals. The recently found *grid cells*

(Hafting et al., 2005) in the dorsal medial EC show such a topological structure, and constitute a good candidate for implementing the path integrator.

mEC receives projections from the parasubiculum. This area receives head direction and self-movement information through interactions with the retrospenial cortex and anterior thalamus, which maybe the basis for the idiothetic representation (Hargreaves et al., 2005).

It should be noted, however, that IPCs described in this work do not reproduce the firing patterns recently reported for the grid cells, as these cells exhibit multi-peaked regular place fields. A single peaked representation can be obtained by decoding the collective firing of several mEC (Fyhn et al., 2004). IPC population can be interpreted as a read-out of the combined activity of several mEC grid cells.

Although the population activity persists in absence of visual cues, the ensemble activity changes following external cues. Such an interaction may be the basis for the path integrator resetting required when the animal first enter the environment. That interaction in our model is the product of a direct projection from the local-view sensitive cells to the path integrator and does not require an intact hippocampus to be performed (Section 5.2). Correspondingly, experimental results suggest that hippocampal lesions do not disrupt path integration (Alyan et al., 1997).

Combined place representation. Both idiothetic and allothetic representation project onto a single population of place cells. These population does not have recurrent connectivity and there is no topological correspondence between anatomically closed cells and the locations they code. Projections from the APC and IPC population may correspond to the entorhinal afferents to the DG, CA3 and CA1 structures. Taking into account that the PC population does not have recurrent connections it better reflects the sort of computation that may take place in the CA1 structure.

The entorhinal-CA1 pathway have been reported to be sufficient to maintain a stable place representation (Brun et al., 2002), and that feedback connectivity from the HPC to the EC is not required for place-sensitive activity in entorhinal cells (Fyhn et al., 2004).

5.5.2 Limitations

Since the model uses rate coding neurons, it is unable to reproduce specific timing properties of place cells, mainly the modulation of place cell firing by the theta rhythm.

The calibration of the path integrator by the allothetic pathway in the model relies on a parameter (β) which is maintained constant throughout the simulation. Although the nature and strength of such interactions in rats is yet to be clarified; some degree of adaptability may be incorporated to this part of the model.

For instance, the parameter β can be adjusted according to the reliability of both representations. The variance (σ) of the APC activity around its centre of mass

(population vector) can be used as a measure of the reliability of the allothetic code (broad blobs of activity will be considered less reliable than sharp ones). Accordingly, large values of σ will lead to smaller values of β.

The model only takes into account feed-forward connections from the sensory process structures (related to the EC) to the combined place coding population (related to CA1). As a consequence, it does not address the possible computational role of structures like the DG and CA3. Even though this structures seem not to be necessary for the features reproduced by the model (Brun et al., 2002; Fyhn et al., 2004; Hafting et al., 2005), they can contribute to enhance the spatial selectivity and directional properties of place fields in both the CA1 and EC (Fyhn et al., 2004). And be the source of other features not addressed in this work (i.e. replay of sequences of place cell's firing during sharp waves).

Chapter 6

Navigation strategies in rats

6.1 Navigation strategies

Navigation is defined as the capability of planning and performing a path from a current position toward a desired location (Etienne, 1998). This requires the use of available information about both the current and desired location, as well as the animal's ability to adapt its own behaviour to the characteristics of the task to be solved.

In order to solve a navigational task, an animal can select different strategies. These strategies can be categorised as follows (Redish, 1999),

- *Taxon navigation.* These strategies can be applied when the desired target location is uniquely signalled by a salient cue, in these cases the animal can rely on a simple landmark-based behaviour. No knowledge about the agent location is required as the goal can be located in an egocentric frame of reference. These strategies are also referred to as *orienting* or *beaconing* behaviour, as they require the animal to orient to a landmark, or use it as a beacon.

- *Praxic navigation.* This category includes behaviour consisting on complex responses (for instance, a sequence of movements) triggered by simple stimuli. For instance, if the animal has to run from one arm in a radial maze, and perform always the same body turn (i.e. 90° left) at the centre of the maze. Praxic strategies generally consists of *stereotyped movements* independent of the starting point.

- *Locale navigation.* In cases when the goal's location is not signalled by any cue (or sequence of cues), the animal has to rely on an internal representation of space. This representation corresponds to the *cognitive map* as proposed by Tolman (1948).

Both praxic and taxon navigation rely on the association of a stimulus to a specific response (Stimulus-response behaviour, or S-R). In contrast, locale navigation requires the use of complex (internal and external) stimuli to build the representation of space in which the navigation strategy relies on. In the following we will also refer to locale strategies as *place-response*, as opposed to *stimulus-* or *cue-response* corresponding to both praxic and taxon strategies.

As described in previous chapters, the hippocampus seems to encode a representation of space which may be required in order to be able to express place-responses. Besides, different nuclei in the Basal Ganglia (BG) have been linked to the expression of different types of response. Specifically, the Nucleus Accumbens (NAc) in the ventral pathway of the BG has been proposed to encode and select actions based on *place-coding* information from the hippocampus. Furthermore, the dorsal pathway of the BG (including the dorsal striatum or Caudato-Putamen) is believed to be responsible for the selection of *stereotypic stimulus-responses* motor control by a competition among different preset behaviours.

In this chapter, we first describe the anatomical properties of the Basal Ganglia (Section 6.2); then, we review experimental paradigms aimed to test and dissociate the navigation systems mediating each type of navigation strategy (Sections 6.3–6.4). Some theoretical hypothesis on the role of neuromodulators in learning and memory are presented in Section 6.5. Finally, a review of several models of rodent navigation is presented in Section 6.6.

6.2 The basal ganglia

The basal ganglia refer to a brain area comprised of several nuclei in the fore-and midbrain. It is known to be involved in motor disorders such as Parkinson disease or Tourette syndrome. Functional models have proposed that BG are in charge of action selection by means of a competition among different sensory-motor programs, where the selected program gains access to motor control trough a process of selected des-inhibition (Mink, 1996).

The main input structure to the BG is the striatum. It is innervated by almost all cortical areas and the hippocampus. The striatum is further subdivided along the dorso-ventral axis into two structures. First, the dorsal striatum, or *Caudato-Putamen* (CPu) whose main afferents come from cortical areas; second, the *nucleus accumbens* (NAc), in the ventral striatum is innervated by both the hippocampus and cortical areas. Another input structure to the BG is the sub-thalamic nucleus (STN).

The dorsal striatum is anatomically divided in two regions: patch and matrix. The matrix is structurally segregated in isolated neural compartments and projections from this area to output structures of the BG preserve the compartmental segregation, except for projections to the STN. Such a segregated structure has been proposed to implement independent sensory-motor associations (Mink, 1996).

The main outputs of the BG are the *substantia nigra pars reticulata* (SNr) and the *entopeduncular nucleus* (EP)[1]. They project onto the thalamus, which in turn projects back to the cortex, and to the superior colliculus (Amaral and Witter, 1995; Heimer et al., 1995; Mink, 1996).

Other nuclei in the BG are the *substantia nigra pars compacta* (SNc) and the *ventral tegmental area* (VTA). They receive projections from the patch region of the dorsal striatum, and from the ventral striatum.

Dopaminergic neurons in this area project back onto the striatum and tend to synapse on the same spines as cortical and hippocampal afferents. Dopaminergic activity is believed to regulate the plasticity of cortico-striatal synapses (Schultz, 2002). Such regulation seems to be based on the processing of reward signals, analogous to the reward prediction error used in the reinforcement learning paradigm (Schultz et al., 1997; Sutton and Barto, 1998) as described in Section 6.5.

A closed circuit is formed by the cortical projections to the striatum, which then project onto the Thalamus through the output nuclei in BG, and from there back to cortical areas. This cortico-striatal circuit is subdivided in parallel independent loops originated in different regions of cortex, passing through specific sub-regions of the BG nuclei (Alexander and Crutcher, 1990; Alexander et al., 1986), and projecting back from there to the same cortical areas. In the rat, three independent loops may be distinguished along the dorso-ventral axis:

1. Motor loop, consisting on the dorsal pathway of the BG, including the Caudate nucleus (CPu), the Globus pallidus and the lateral part of the SNr. It also involves motor regions of the cortex. It has been suggested to be involved in instrumental memory, and developing of stimulus-response associations (Graybiel, 1998).

2. Associative loop, including the ventral part of the dorsal pathway. It is believed to be involved in working memory and sequence learning.

3. Limbic loop, corresponding to the ventral pathway of the BG, including the ventral striatum (particularly the NAc). It receives projections from the hippocampus and the amygdala. In particular, models of hippocampal-based navigation have proposed the Nucleus accumbens as the putative locus for selecting actions based on place coding projections from the hippocampus (Burgess et al., 1994; Foster et al., 2000; Arleo and Gerstner, 2000).

[1]Analogous to the internal segment of the *globus pallidus* (GPi) in the primate.

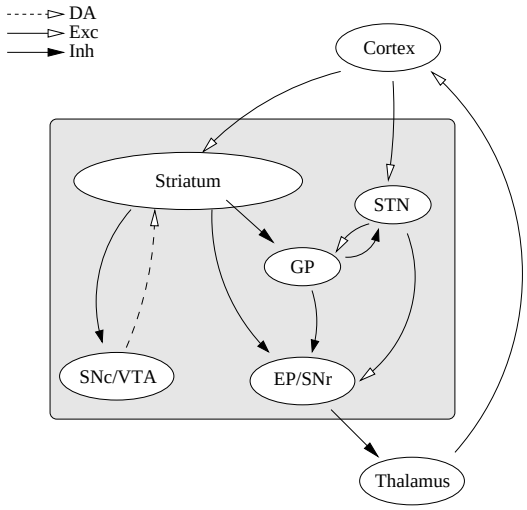


Figure 6.1: Schematic representation of the interconnectivity in the BG. Projections from the hippocampus and amygdala to the ventral part of the striatum are not represented. Dopaminergic projections are represented by dashed arrows. Open arrowhead denotes excitatory connections and filled arrowheads denote inhibitory connections. See the text for further description.

Each one of these loops may be further segregated in independent parallel subcircuits (as in the case of the compartmental organisation of the dorsal striatum-matrix).

6.3 Testing navigation strategies

The Morris water maze is one of the canonical experimental paradigms used to test the navigational capabilities of rodents. It consists of a (usually circular) swimming pool filled with coloured water. At some location of the pool there is an escape platform on which the animal can stand out of the water. Two different versions exist: in the *hidden* version, the platform is just below the surface of the water; In the *visible* water maze, the platform stands above the water, and in some cases a co-localised cue indicates its location. The animal is placed in the pool by the experimenter and is required to swim toward the platform. The performance in the water maze is often assessed by measuring the animal's *escape latency*, i.e. the time required to reach the platform. As the animal improves its performance, it takes less time to escape and its latency decreases.

The availability of visual stimuli (i.e. the co-localised cue or the platform itself) in the visible version of the water maze makes it possible to use taxon strategies to

solve the task. In the hidden version, by contrast, as the platform is not visible, locale strategies are required. Extra-maze cues and internal sensory signals provide complex sensory input which is used to update the spatial representation (the *cognitive map*) required to solve this task.

Radial-arm mazes are also used to assess the navigational skills of animals. In some cases a reward can be consistently located at the same position with respect to the starting arm, thus requiring the animal to perform the same body turn at the centre of the maze (i.e. entering the arm at the left). Taxon strategies can be tested by adding a sensory cue (i.e. a light) signalling the arm in which the animal has to enter (Packard et al., 1989).

These paradigms have been used to validate the theory of multiple memory systems by observing the effect of localised lesions (e.g. in hippocampus or dorsal striatum) in the preference for locale or S-R strategies (Whishaw et al., 1987; Packard et al., 1989; Pearce et al., 1998; Da Cunha et al., 2003). As mentioned, substantial evidence supports the idea that the hippocampus mediates spatial learning[2], whereas the dorsal striatum is involved in S-R learning (See Redish (1999); Packard and Knowlton (2002); White and McDonald (2002) for reviews).

6.4 Dissociating navigation systems

Some studies use a single experimental paradigm to test simultaneously the development of both cue and place-response (Packard and McGaugh, 1996; Oliveira et al., 1997; Devan and White, 1999; Chang and Gold, 2003b). Additionally, the interactions between the different systems have been studied by testing the facilitating effect of lesioning a system which is not related to the task to be solved (Packard et al., 1989; Schroeder et al., 2002; Chang and Gold, 2003a). Another, interesting line of research studies the correlation of neurochemical factors with the expression of different kinds of response (M. E. Ragozzino and Gold, 1995; Colombo and Gallagher, 1998; Janis et al., 1998; Bizon et al., 2003; Chang and Gold, 2003b; McIntyre et al., 2003; Marriott and Korol, 2003). This section reviews some of these studies.

6.4.1 Simultaneous Development of Place and Cue-response

Devan and White (1999) applied an experimental procedure in which cue and place responses are developed simultaneously in the water maze (Figure 6.2). Rats were trained during 9 days, interleaving trials in the visible (days 1,2,4,5,7 and 8) and

[2]Studies exist however, which show that animals with hippocampal impairments can solve the hidden water maze task under certain training conditions. Some of these cases can be explained by a form of learning, different from spatial learning, as responsible for solving the task (Eichenbaum et al., 1990; Whishaw et al., 1995; Pearce et al., 1998; Redish, 2001; Pouzet et al., 2002).

hidden (days 3,6, and 9) version of the maze. The escape platform remained in the same location during the entire training phase. On the 10th day, a test (competition) trial was performed with a visible escape platform in a location different from the one used during training. This testing paradigm allows the dissociation of both types of response, as animals exhibiting cue-response will swim directly toward the visible platform, whereas a place-response behaviour will lead the animal toward the location where the platform was during the training phase. The experimental procedure was applied to rats which were either intact or had lesions in one of the following areas: (i) Hippocampal formation (fornix/fimbria); (ii) dorsolateral striatum; and (iii) dorsomedial striatum. During training, intact animals decrease their escape latency in both the visible and hidden trials. Lesioned animals also improved during training except for animals with hippocampal lesions which learn the visible version of the task but were unable to reach the hidden platform in blocks 3, 6, and 9. In the competition test, consistent with other studies, hippocampal and dorsomedial striatal lesions produced a preference for cue-response, as opposed to lesions in the dorsolateral striatum which biased the animal to exhibit place-responses. Among the control animals, nearly half of them swam first to the location of the platform used during training (place-response) and then toward the visible platform (cue-response). This supports the idea of both types of response developing simultaneously in intact animals, and shows the ability of switching from a first behaviour to a second one if the first fails.

Other experiments have shown that, in some experimental conditions, rats change gradually from using a place strategy in early phases of training to a S-R behaviour, after longer training. Packard and McGaugh (1996) trained rats in a T-maze with a fixed starting position to go toward a baited arm (e.g. arm pointing to the west). Both the starting position and the bait location remain fixed during the training phase. A test trial was performed after rotation of the maze by 180°, to observe whether the rat follows a place-response behaviour, going toward the same arm location (west), or if it repeats the same response as in training, performing the same body turn at the junction of the maze (entering the eastern arm). After 8 days of training, most of the intact animals exhibited place-response, but after 16 days most of them performed the same body turn as during training. Functional inactivation of the dorsolateral striatum at day 16 led the animals to exhibit place-response, showing that this type of memory is preserved even if it is not behaviourally manifested.

6.4.2 Studying Interactions Between Systems

The precise mechanism of the interactions between the different neural navigation systems is not clear yet. Studies have shown opposite effects of lesions in the hippocampus and the dorsolateral striatum when solving a place or cue-response task respectively. The idea of a competition among those systems is supported by studies

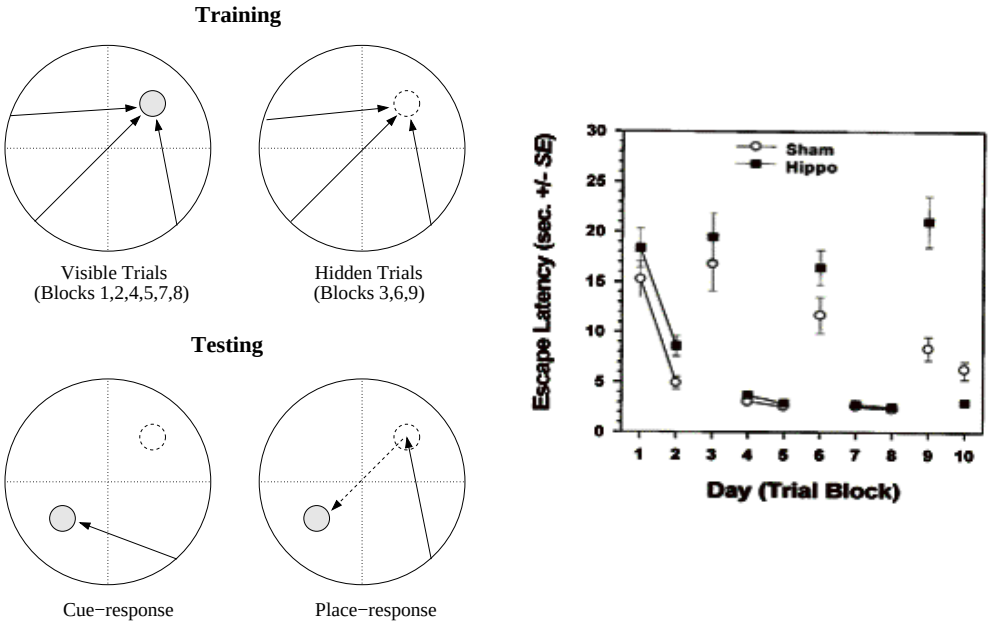


Figure 6.2: (a) Water maze used by Devan and White (1999). *Top*. Training phase. The rats were trained to reach an escape platform at a fixed location. During days 1,2,4,5,7, and 8 the platform was visible (left). In the remaining days (3,6, and 9) the platform was below the surface of the water (right). The interleaved training in both the visible and hidden version of the water maze allows the simultaneous development of place and cue-responses. *Bottom*. On day 10, a competition trial is performed with a visible platform at a different location from the one used in training. The filled circle marks the new position of the platform. The dotted circle shows the location of the platform during training. Cue-response will guide the animal toward the visible platform (left). In contrast, place-response will lead the animal toward the location used during training (right). Animals choosing place-response later changed to cue-response and swam toward the visible platform (dotted-arrow). (b) Escape latency of the control animals (Sham) and animals with hippocampal lesions (Hippo). Redrawn from Devan and White (1999).

showing that lesions in one system may improve the performance in tasks associated to the unlesioned brain area (i.e. animals with hippocampal lesions outperforms control animals in cue-response tasks) (Packard et al., 1989; Devan and White, 1999; Schroeder et al., 2002; Chang and Gold, 2003a). This suggests that the intact area (eliciting appropriate behaviour) no longer has to compete with others for action selection, leading to a better performance (Gold, 2004).

For instance, Chang and Gold (2003a) trained rats in a plus-maze with a varying starting point to go into a baited arm which can be either always at the same room location (place version), or defined by the same body turning, i.e. arm left to the starting point (S-R version). Hippocampal lesions, as expected, impair the learning of the place version of the task. Moreover, hippocampal lesions facilitate learning in the dorsostriatal-dependent S-R version of the maze. Similar facilitation effects have been reported in the water maze task described above (Devan and White, 1999), where both hippocampal and dorsomedial striatal lesioned animals outperform control animals in the competition trial. Furthermore, ethanol administration, which produces memory impairments similar to those due to hippocampal lesions, also facilitates the acquisition of response tasks (Matthews et al., 1999).

6.4.3 Pharmacological studies

The role of neurochemical factors mediating the competition between different systems has been studied recently (see Gold (2004) for a review). Injections of glutamate in either hippocampus or striatum produce a dominance of spatial or S-R solution in a T-maze task respectively (Packard, 1999). Similarly, the preference for spatial learning changes across the oestrous cycle of female rats such that, animals are more likely to select place strategies when there are high levels of oestrogen (proestrous) (Daniel and Lee, 2004; Korol et al., 2004). A possible mechanism for oestrogen levels to affect the strategy selection is by modulation of neuromodulators such as acetylcholine (ACh) (Marriott and Korol, 2003; Korol, 2004).

Colombo and Gallagher (1998) report that increases in acetylcholynetransferase activity in the dorsolateral striatum are negatively correlated with the accuracy in hippocampal-dependent working memory tasks. Furthermore, alterations in the cholinergic septohippocampal system leading to enhanced or impaired cholinergic function in the hippocampus, seem to respectively enhance or reduce learning and memory in this structure (M. E. Ragozzino and Gold, 1995; Janis et al., 1998; Gold, 2003) (but see Cahill and Baxter (2001); Bizon et al. (2003)). Simultaneous measure of acetylcholine in the hippocampus and the striatum showed a correlation between the changes of ACh release and the expression of place-response or S-R behaviour (Chang and Gold, 2003b; McIntyre et al., 2003). Interestingly, hippocampal ACh levels prior to the experiment reliably predict which type of response will be preferred during training, suggesting a neurochemical bias toward a specific type of response depending on ACh levels. Changes in ACh in the dorsolateral striatum and the hippocampus were not negatively correlated, suggesting an indirect interaction, at least with respect to cholinergic activity, among the two systems (Gold, 2004).

6.5 Role of neuromodulators : Theoretical hypothesis

As evidenced by the studies described in Section 6.4.3, neuromodulators play an important role in learning and memory in both spatial and non-spatial tasks. Specific effects of increment of depletion of these neuromodulators in learning gives origin to several theories of the putative function of these substances as regulators of the learning process. In terms of modelling, they may correspond to learning metaparameters, i.e. they control the parameters of the learning algorithms (Doya, 2002).

In this section we review the proposed roles for Dopamine (DA), Acetylcholine (ACh), Norepinephrine[3] (NE) and Serotonin.

Dopamine role in coding the reward prediction error

Experiments in primates have shown that the activity of dopaminergic neurons in the substantia nigra pars compacta (SNc) and the ventral tegmental area (VTA) is correlated with the delivery of unexpected rewards (Schultz et al., 1997; Schultz, 1998, 2002). Monkeys were trained to press a lever after a light was turned on in order to trigger the release of liquid reward. Early in training, dopamine neurons respond to the delivery of reward, or when the liquid was released outside the task. Later on during training, those neurons respond to the light and not to the reward itself. Moreover, their activity dropped when the reward was omitted. This pattern of activity closely resemble the TD error signal (δ) used in the reinforcement learning paradigm (Montague et al., 1996; Schultz et al., 1997).

Dopamine is believed to regulate synaptic plasticity of cortical inputs to the striatum. Dopaminergic levels have been shown to reverse the direction of this plasticity (Reynolds et al., 2001; Wickens et al., 1996), in a way that may be consistent to reward-based learning algorithms (Doya, 2002).

Acetylcholine, and the balance between storing and encoding

Hasselmo and colleagues (Hasselmo, 1999; Hasselmo, Bodelon, and Wyble, 2002; Hasselmo, Hay, et al., 2002) have proposed that cholinergic activity controls the encoding of incoming information in the hippocampus, and other cortical areas. ACh has been shown to suppress excitatory recurrent collaterals in the CA3, as well as Schaffer collaterals connecting CA3 to CA1 in hippocampal slices. Consistently, cholinergic modulation also suppresses feedback connections from high-order somatosensory cortex to primary somatosensory cortex. This data suggest that changes in ACh levels

[3]Also known as noradrenaline.

may result in subsequent phases of feed-forward dominance and feedback dominance corresponding to the encoding and retrieval of memories.

These encoding and retrieval phases may correspond to high ACh levels during active waking states contrasting to low levels during quiet-awakeness and slow-wave sleep (Hasselmo, 1999). Similarly, changes in cholinergic activity during the theta cycle may correspond to the rapid transitions between encoding and retrieval phases (Hasselmo, Bodelon, and Wyble, 2002; Hasselmo, Hay, et al., 2002). This two-phase model allows the encoding of new patterns in the recurrent CA3 network without interference of previously stored patterns.

In an attempt to provide experimental support to this theory Rogers and Kesner (2003) compare the performance on rats running out a maze. They assessed the encoding of the task by comparing the performance in the first and last trials within the same day. Retrieval was measured by comparison of the performance in the first trials of one day with the last trials of the *previous* day. Consistently with the hypothesis, they found that a cholinergic antagonist impaired the encoding of new data, but had no effect on retrieval; while an acetylcholinesterase [4] inhibitor disrupted the retrieval of spatial information without affecting the encoding.

Neuromodulators as metaparameters in RL (Doya, 2002)

Complementing previously proposed roles for DA and ACh, Doya (2002) hypothesise the role of the main four neuromodulators as controlling the parameters of learning in the brain. In this framework, *dopamine* and *acetylcholine* have similar roles as the ones describe above. The former encodes global learning signals based on the prediction of reward; while the latter controls the balance between memory storage and renewal. In addition, Doya (2002) proposes that *Serotonin*, controls the balance between short-term and long-term memory, and *Norepinephrine*, controls the amount of exploration (i.e. exploration-exploitation dilemma).

These roles can be assimilated to modulating the parameters of the reinforcement learning paradigm (Sutton and Barto, 1998). Correspondingly, Dopamine would encode the TD error (δ); Serotonin, the discount factor (γ), which controls the relative importance of rewards depending how far in the future they are expected. Acetylcholine, would correspond to the learning rate (α) and Norepinephrine, the inverse temperature (β) which controls the stochasticity of the policy. (See Section 2.1.2).

Neuromodulators and uncertainty : ACh and NE

Yu and Dayan (2002, 2003, 2005), inspired by Bayesian statistical approaches, have proposed different, complementary roles for the acetylcholine and norepinephrine in

[4] The acetylcholinesterase (ACHe) is an enzyme which breaks down the ACh at the synaptic cleft. Inhibition of ACHe results in an increment of extracellular ACh levels.

signalling the uncertainty associated to observations of the environment. They propose that ACh signals the *expected* uncertainty, resulting from the known reliability of predictive cues within a context. Norepinephrine, in turn, signals the *unexpected* uncertainty associated to context switching yielding unexpected observations.

According to a statistical framework, increments in both expected and unexpected uncertainty should suppress the effect of top-down expectation-driven signals relative to bottom-up sensory driven-signals. In addition they are supposed to promote context learning from the incoming sensory input. Correspondingly, injection of both ACh and NE, enhances stimulus-evoked responses, and increases experience-dependent reorganisation of cortical maps (Gu, 2002). The specific role of ACh as coding for expected uncertainty is consistent with experimental results showing that animals learn faster about stimuli with variable consequences, and such type of learning requires cortical ACh innervation (Holland and Gallagher, 1999).

NE activity in contrast seem to be related to novelty (Yu and Dayan, 2003). Furthermore, NE in locus Coeruleus activity rapidly habituates when conditioning is expressed at the behavioural level (Sara et al., 1994). Such a transient response in NE neurons, is interpreted as a response to unexpected changes in the environment which disappears once the new context is properly predicted by top-down signals.

6.6 Models of rodent navigation

In this section several models of rodent navigation and action control are reviewed. Two types of models are presented, Section 6.6.1 reviews models of locale navigation. These models rely on a representation of space coded in a population of place cells which interacts with a neural representation of actions (i.e. directions of movement) or goal locations. The representation of actions is often proposed to reside in the ventral pathway of the BG, specifically in the NAc.

A second type of models follows the hypothesis that the neural circuit of the BG implements reward-based action selection (Houk et al., 1995; Montague et al., 1996; Doya, 1999). They propose that the BG, and more specifically its dorsal pathway, encodes different sensory-motor programs (corresponding to S-R associations). The expected outcome of the different programs is learnt through reinforcement learning, and a reward dependent competition takes place to select the appropriate actions. Some of the principles of this type of models are reviewed in Section 6.6.2.

6.6.1 Models of Hippocampal-dependent navigation

Gaussier *et al.*(1998–2002)

The navigation model proposed by Gaussier *et al.* is based on the spatial representation described in Section 3.3.4. The model proposes the hippocampal CA3 network to

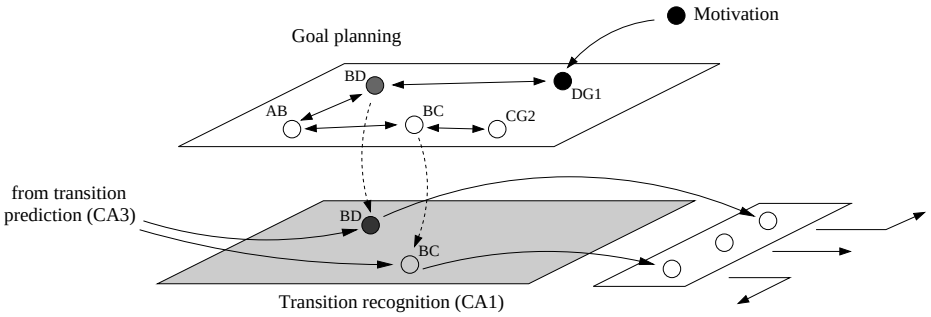


Figure 6.3: Navigation architecture proposed by Gaussier. Transition predictions are coded in the CA3 network which project onto CA1. In the CA1 layer transition selection is achieved by a winner-take-all competition. Competition is modulated by the goal planning layer which depending on motivation signals favoured transitions leading toward the selected goal.

correspond to a transition prediction layer which proposes candidate future locations. A competition mechanism selects the most active transition. Transition recognition cells are connected to motor cells trough projections updated using Hebbian learning. A schematic diagram of the model is shown in Figure 6.3.

The goal-planning layer is composed of cells coding the same transitions as the recognition layer. Cells are interconnected by constant synaptic weights $w_{ij} < 1$. In order to achieve goal-based navigation, external motivation inputs activate a node (coding for a transition leading to a goal), and the activity $A_0 = 1$ is propagated to other transitions. Cells activities, $A_i = max_j(w_{ij}Aj)$, are updated iteratively till the network reaches a stable state. At that state, cells activities are set according to the distance to the goal location. These activities propagate to the recognition layer to bias the competition toward transitions corresponding to the shortest path toward the goal.

Abbot and coworkers (1996,1997)

Abbot *et al.*(Blum and Abbott, 1996; Gerstner and Abbott, 1997) propose the recurrent network in CA3 as responsible for encoding navigation maps. The rationale of the model is that experience-dependent learning in CA3 recurrent connections shift the (initially Gaussian) receptive fields of CA3 cells toward the goal location. This shift appears after a training paradigm consisting of repeated trajectories ending at the goal location. A target-dependent modulation can be used to allow the model to store several navigation maps.

The navigation maps resulting of the model are computed as the shift of the receptive fields after the training. However, this operation is biologically implausible

as the animal would have to access explicitly information about the place field centres. Furthermore, it predicts that navigation relies solely in the hippocampal CA3 structure. Several experimental studies have shown that lesions in the nucleus accumbens or fornix also impairs animals for locale navigational tasks (Eichenbaum et al., 1990; Sutherland and Rodriguez, 1990; Packard and McGaugh, 1992). While changes in the CA3 connectivity due to experience-dependent synaptic plasticity may not be responsible for the role of action selection in navigation tasks, it can generate the replay of place cells firing sequences during quiet-awakeness and slow-wave sleep (Section 3.1.3).

Foster, Morris and Dayan (2000)

Foster et al. (2000) propose a model of navigation based on an actor-critic architecture for the temporal-difference (TD) reinforcement learning. In the model the critic is designed to estimate the *value* of each location in the environment, while the actor is responsible for action selection.

Position information is encoded in a population of place cells with perfectly tuned Gaussian receptive fields. This population projects onto a critic neuron c, whose firing rate is computed, $r_c = \sum_i w_{ci} r_i$, where r_i is the activity of a place cell i. The critic activity r_c, corresponds to the predicted value of the current location. In addition, the critic also computes the reinforcement signal δ, corresponding to the temporal difference of current and previous activities. This signal is used to update the weights from the PC population, such that $w_{ci} \rightarrow r_i \delta$.

The actor is composed of eight cells coding for corresponding actions (i.e. directions of movement). These cells are driven by the PCs, and their activity is computed as $r_a = \Sigma_i w_{ai} r_i$. Actions are selected stochastically, with highly active action cells are more likely to be selected. Weights w_{ai} are updated using a Hebbian learning modulated by the critic's reinforcement signal δ.

The system is able to solve a simulated version of the morris water maze with a hidden platform consistently located at the same position of the environment (i.e. the reference memory water maze, RMW) in about ten trials. However, if the platform is moved to a different location, the previously learnt policies interfere with the learning of the new goal location.

In order to avoid such interference and reproduce animal behaviour in the moving-platform version of the maze (i.e. delayed matching to place water maze, DMP)[5] a coordinate system (CS) was added to the model. The CS consists of three components: (i) a coordinate representation of the agent's location in allocentric coordinates

[5]In this task, the platform location is changed at the beginning of each training day and remains at the same location during the whole training block (Steele and Morris, 1999). Experiments with rats show that animals exhibit a one-trial learning after a few days of training suggesting that they are able to avoid interference between consecutive training days.

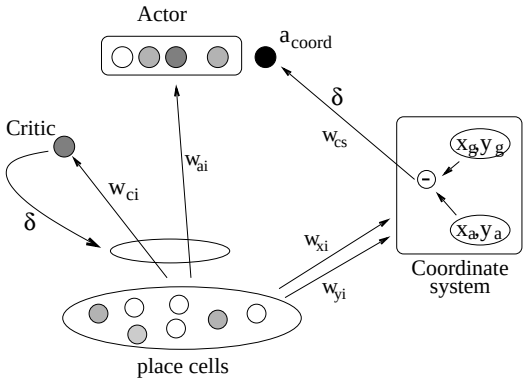


Figure 6.4: Navigation model proposed by Foster et al. (2000). An actor-critic architecture is employed to learn the navigation map. Both the actor and critic are driven by a population of place cells. The actor selects the actions to take whereas the critic evaluates the quality of performed actions, generating a reinforcement signal δ corresponding to the reward prediction error. An allocentric task-independent coordinate system (CS) is learnt from the place cell activity, and stores the location where the goal was the previously found. Such a system is used to propose goal-oriented actions (θ_g, encoded by action cell a_{coord}) computed algorithmically by vector subtraction.

(x_a, y_a). This representation is implemented by two units receiving projections from all the place cells. (ii) A goal coordinate memory, (x_g, y_g) storing the location where the goal was lastly found. (iii) A mechanism that algorithmically subtracts the agent and goal coordinates to provide the direction Θ_g leading toward the goal. The mapping of place cell activity to coordinate units is learnt using TD-learning.

An abstract action cell a_{coord} is used to encode the direction Θ_g computed by the CS. This cell also competes with other action cells for action selection, and its synaptic connection (from the CS) is updated using the reinforcement signal δ computed by the critic. The activity of a_{coord} depends on how well the CS is tuned, and whenever is selected for action-control, the agent's movement is computed algorithmically by the CS.

This model relies on an ideal place representation which does not depend on sensory inputs and is assumed to be consistent over time. Furthermore, it relies upon an explicit memory of the goal location (x_g, y_g). The coordinate system algorithmically computes the direction of movement toward the goal, and doesn't take into account local information as obstacles.

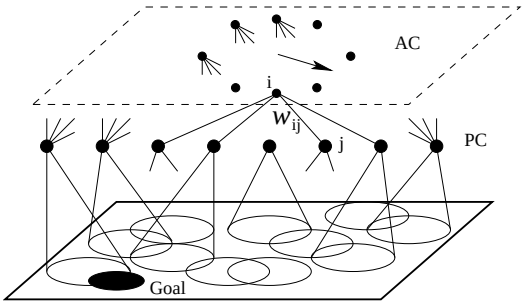


Figure 6.5: Navigation model proposed by Arleo and coworkers. A population of action encoding cells (AC) is driven by place cells (PCs). Each AC encodes a particular action (i.e. direction of movement), and its activity correspond to the estimated state-action value of taking such an action when the agent is located at a particular position in the environment (encoded by the population activity of PCs). Synaptic projections w_{ij} are modified by means of reinforcement learning.

Arleo *et al.*(2000–2001) and Strösslin *et al.*(2004)

Similar to the model of Foster et al. (2000), Arleo and Strösslin (Arleo et al., 2004; Arleo and Gerstner, 2000; Strösslin, 2004) model hippocampal based navigation using TD learning. A population of place cells (described in Section 3.3.5), projects onto a population of action cells coding for specific directions of movement. These projections are updated in order to minimise the reward prediction error δ.

In this model, instead of having a critic neuron coding for the *value* of the current position, the activity of action cells codes for the state-action value (or Q-value) of taking a specific action (corresponding to that cell) when located at the current position (indicated by the PC population activity).

Every action cell receives input from place cells i, and its activity corresponds to the weighted sum of the pre-synaptic activities $r_i^{\mathrm{AC}} = \sum_j w_{ij} r_j^{\mathrm{PC}}$. Strösslin (2004) propose a two-phase action selection process in which *first*, the action cell activity is computed as shown above, and the direction (Φ^{AC}) provided by the AC population vector, is chosen as the new direction of movement. The Q-value of that action is computed by linear interpolation of the activity of the two action cells coding for the closest angular direction to Φ^{AC}. *Secondly*, a Gaussian pattern of activity is enforced around Φ^{AC}, this step allows several ACs to be updated simultaneously to promote generalisation in the action space.

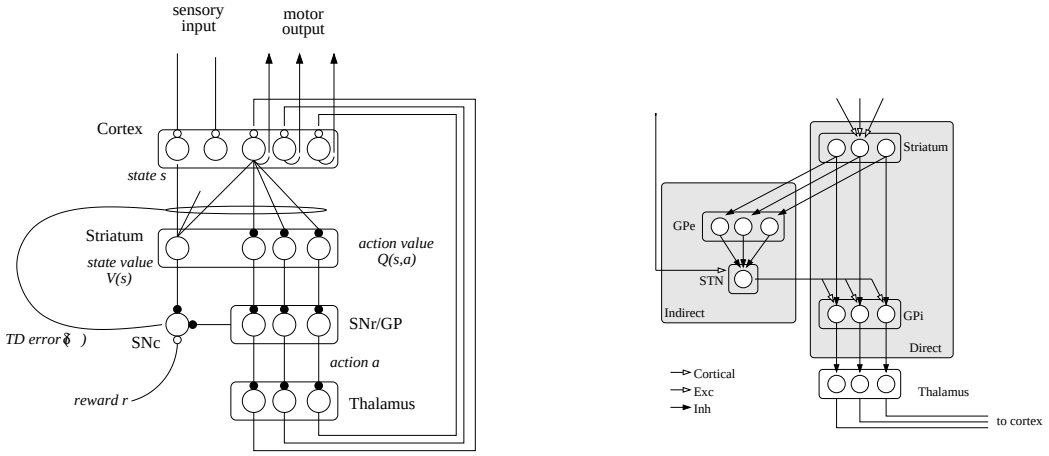


Figure 6.6: (a) Schematic diagram of a model of the basal ganglia implementing the reinforcement learning algorithm. Adapted from (Doya, 2002). (b) Direct and indirect pathways in the output of the BG as modelled by Berns and Sejnowski (1996). Thick arrows: cortical projections to BG. Solid arrowheads: inhibitory connections; open arrowheads: excitatory connections.

6.6.2 Models of Basal ganglia

The basal ganglia, and particularly its dorsal pathway, has been largely modelled as implementing the reinforcement signal in a reward-based learning paradigm. The interpretation of the dopaminergic activity as encoding the reward prediction error have been the base of such models (Montague et al., 1996; Doya, 1999; Montes-Gonzalez et al., 2000; Baldasarre, 2002). See Khamassi et al. (2005); Joel et al. (2002); Houk et al. (1995) for reviews.

In general, these models are designed to select among a repertoire of specific sensory motor programs (or S-R associations), associated to discrete actions in the reinforcement learning paradigm. Cortical projections to the striatum provide information about the state s, and dopaminergic activity represents the TD signal δ. The striatum uses this signal to learn the state value functions $V(s)$, and the action value function $Q(s, a)$ corresponding to the policy $P(a|s)$ (Doya, 2000). The dorsal striatum-patch is proposed as the neural locus for learning the state value function $V(s)$; whereas the action values $Q(s, a)$ would be encoded in the matrix area. Each independent compartment (i.e. channel) in the latter area is proposed to encode the Q-values of independent actions.

A competition mechanism dependent on the Q-values is implemented in the output

structures of the BG. Such a competition is believed to take place through a process of selected des-inhibition taking place in the so-called direct/indirect pathways to the output nuclei of the BG (Mink, 1996; Berns and Sejnowski, 1996). Different actions will be encoded in segregated channels in the dorsal pathway of the BG, and the direct pathway (Striatum → GPi), provides selective inhibition to the different channels according to their input activity levels; the indirect pathway (Striatum → GPe → GPi) generates a global excitation to all channels in the GPi. The combination of selective inhibition and global excitation, results in the des-inhibition of a single channel. This architecture corresponds to a connectionists network with lateral "on-centre off-surround" connections, known to yield a single winning neuron (or a localised patch of winning neurons).

In general, these models have been applied to biomimetic agents to implement selection among independent stereotypic actions like landmark approaching, wall following, mating (i.e. look for other agents) or battery recharge (Khamassi et al., 2005; Doya and Uchibe, 2005; Montes-Gonzalez et al., 2000). Each action is encoded in separate striatal channels.

Recently, Girard *et al.*(Girard et al., 2005; Girard, 2003) propose a model which integrates the action-selection mechanism with navigation skills. The model is composed of two modules, one selecting directions of movement and the other selecting different stereotypic actions. These modules are assimilated to the ventral and dorsal pathways in the BG, respectively (Figure 6.7).

Both modules are based upon the model proposed by Gurney et al. (2001a, 2001b). The dorsal module selects among non-locomotor stereotypic actions. The ventral module performs a competition between two types of navigation (i) A landmark approaching strategy and (ii) A trajectory planning system based on a topological representation of the environment and goals (Filliat, 2001). The output of this module encodes directions of movement in allocentric coordinates.

An asymmetric interaction between the modules is proposed. Non-locomotor actions in the dorsal module can inhibit the navigation signals generated in the ventral module (i.e. they can stop the navigation). This interaction is proposed to reside in a difference in the relative influence of sub-thalamic projections to the dorsal and ventral pathways.

Navigation in this model relies on a topological representation of space. This representation consists in a graph whose nodes encode for explored locations. Each node stores allothetic information of those locations, and the connections between nodes encodes distance and orientation of the represented places. The process of navigation is implemented algorithmically and is not biologically plausible. In addition, the selection depends on pre-defined parameters and no learning mechanism is proposed to adapt this process.

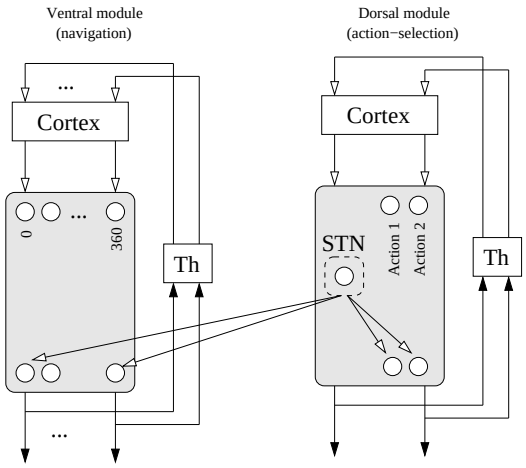


Figure 6.7: Model proposed by Girard. Shaded regions correspond to the competition processes taking place in both the ventral and dorsal modules, these processed are proposed to take place in the corresponding dorsal and ventral pathways of the BG. Selected des-inhibition in the projections from the STN to the ventral module can prevent navigation commands to be performed. Adapted from (Girard, 2003).

Chapter 7

Navigation model

According to the studies described in the previous chapter, models of animal navigation should *(i)* implement separately the two different types of response (i.e. cue-based and place-based) and *(ii)* provide a competition mechanism that selects which kind of action should be performed. The outcome of the competition may depend on the characteristics of the task and the training process (Packard and McGaugh, 1996; Oliveira et al., 1997; Chang and Gold, 2003a). Additionally, as a consequence of the competition between place and cue-response, the disruption of one of the modules should improve the performance in tasks involving the remaining, intact module.

Following these considerations, we propose a model of rat navigation behaviour which is able to develop place and cue responses and selects the appropriate action to be performed (Chavarriaga et al., 2005b, 2005a). Both place-response and cue-response associations are acquired separately using reinforcement learning, as described in Section 7.1. At any moment potential actions corresponding to both types of response are computed and a competition takes place in order to decide which action will be actually performed by the agent. The process of action-selection and learning is presented in Section 7.2. Section 7.3 summarises the steps required to update the parameters of the model, and perform reward based navigation. Finally, Section 7.4 presents the obtained results.

7.1 Navigation strategies as a RL problem

The process of learning to solve a navigational task can be interpreted as the process of developing appropriate actions for the different possible states. In goal directed

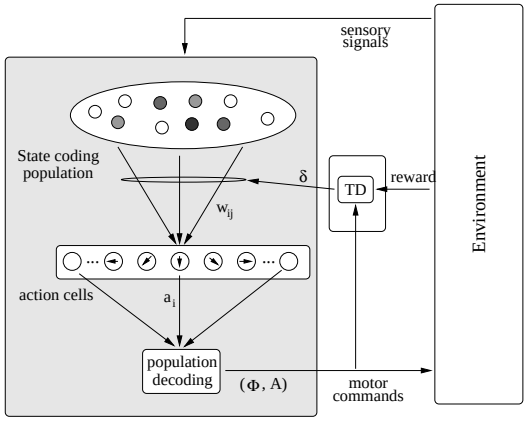


Figure 7.1: Navigation model as a reinforcement learning problem. The perceived state is encoded by the activity of a cell population. In the same way, action cells code the action value function (Q-function) of all possible directions of movement. Every action cell i, has a corresponding direction of movement ϕ_i (represented by the arrow inside the unit) and its activity correspond to the Q-value of moving in that direction. Population vector coding is used to compute an action (Φ) in a continuous action space, and its corresponding state-action value (A). Projections from the state-coding to the action-coding population are updated using TD learning.

navigation, the available actions correspond to the possible directions of movement the agent may select when attempting to reach the goal location given a particular state. In this chapter, actions are understood as a movement in a specific allocentric direction (i.e. in a world-centred coordinate system). In the case of place-based strategies, states correspond to locations in the environment, whereas in stimulus-response strategies states may correspond to sensory stimuli signalling the presence of particular cues.

In our model, populations of neurons are used to encode both the state and action space. Following previous models (Arleo et al., 2004; Strösslin, Sheynikhovich, et al., 2005), associations between states and actions can be encoded in the projections connecting the state-coding population to a set of *action cells*. A functional diagram of this approach is shown in Figure 7.1. Each action cell i, driven by a state-coding population, codes for a particular direction of movement $\phi_i = (360°)i/N_{\mathrm{AC}}$ (here N_{AC} denoted the number of action cells) and its activity a_i is computed,

$$a_i = \sum_j w_{ij} r_j^{\mathrm{pre}} \tag{7.1}$$

where r_j^{pre} is the activity of the pre-synaptic input j, and w_{ij} the synaptic weight. This activity corresponds to the predicted future reward (i.e. Q-value) if the agent

moves in direction ϕ_i, given the current state s (determined by the ensemble activity of the state-coding population). In other terms, projections from the state-coding to the action-coding population implement a parametric form of the action value function, i.e. $a_i = Q(\phi_i|s)$.

At any given moment the direction of the next movement can be computed as the population vector of action cells,

$$\Phi = \arctan\left(\frac{\sum_i a_i \sin(\phi_i)}{\sum_i a_i \cos(\phi_i)}\right) \tag{7.2}$$

Population vector decoding allows the system to provide directions of movement in a continuous space based on the finite population of ACs. The state-action value A corresponding to the continuous direction Φ can be approximated by linear interpolation of the activity of the ACs coding for the two closest directions.

As described by (Strösslin, Sheynikhovich, et al., 2005) AC activity is calculated in two stages. In the first stage, *action evaluation*, each AC computes its activity using Equation (7.1). At this stage, the activity a_i represent the estimated action value $Q(\phi_i|s)$ for moving in direction ϕ_i in the current state.

In the second stage, *generalisation*, a Gaussian AC activity profile with variance σ^2_{ac} is enforced around the selected direction of movement Φ (computed using Equation (7.2)). Let $\Delta\phi_i$ be the angular difference between Φ and ϕ_i, the AC activity r_i in this phase corresponds to,

$$r_i = exp(-\Delta\phi_i^2/2\sigma^2_{ac}) \tag{7.3}$$

During this phase, the AC activity represents which action was executed. As the parameter updating depends on the activity of action cells, the localised profile of activity allows simultaneous learning for multiple actions. This contrasts with traditional RL models in which a winner-take-all inhibits all the non-selected actions.

As mentioned, associations between state and actions are encoded in the connections w_{ij}. Following the reinforcement learning paradigm, those connections can be updated in order to minimise the reward prediction error δ (See Section 2.1.2). Taking into account that the state-action value $A(t)$ corresponds to the predicted future reward at time step t, the δ error can be computed iteratively,

$$\delta(t+1) = R_{t+1} + \gamma A(t+1) - A(t) \tag{7.4}$$

where R_t corresponds to the immediate reward received at time t, and γ corresponds to the decay factor. Weights can then be modified in order to minimise the reward prediction error, as follows,

$$\Delta w_{ij}(t) \propto \delta(t) \cdot e_{ij}(t) \tag{7.5}$$

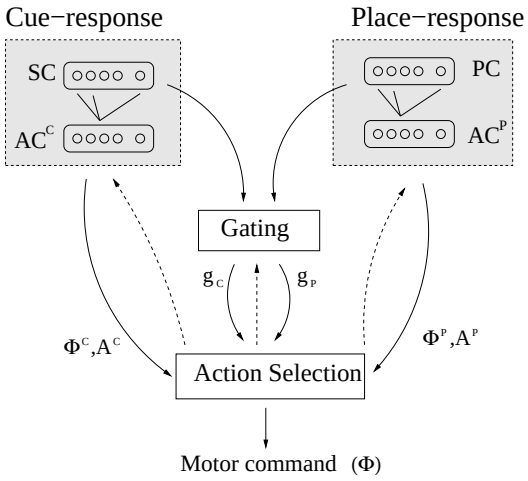


Figure 7.2: Block diagram of the model. Separate modules implement independently place and cue-responses (Φ^{P} and Φ^{C}, respectively). A competition mechanism is used to select the action to be performed, as well as to modulate learning in both modules. SC ≡ Sensory cells forming the cue representation. PC ≡ Spatial representation stored in a population of place cells. AC ≡ Action cells coding for direction of movements. g_C, g_P ≡ Gating values for the cue and place response respectively. A^{C} ≡ Action value of a movement in the direction Φ^{C} selected by the cue–response module. A^{P} ≡ Action value of a movement in the direction Φ^{P} selected by the place–response module.

where e_{ij} is the *eligibility trace*, which can be interpreted as a decaying memory of pre- and post-synaptic temporal coincidences (Sutton and Barto, 1998). The trace on the synapse from a state-coding cell j to an action cell i is computed as,

$$e_{ij}(t+1) = r_j^{\mathrm{pre}} \cdot r_i + \gamma\lambda e_{ij}(t) \tag{7.6}$$

where r_j^{pre} is the activity of the pre-synaptic cell j, and r_i is the activity of the action cell i in the generalisation phase (Equation (7.3)). The factor $\lambda \in [0,1]$ is the trace-decay factor and $e_{ij}(0) = 0$.

This approach was used to model separately place- and cue-response as described in Sections 7.1.1– 7.1.2.

7.1.1 Implementing place-response

Place-response (i.e. Locale strategy) requires a representation of space, such a representation may reside in place-sensitive cells in the rat hippocampus as reviewed in Chapter 3; this hypothesis is supported by experimental studies showing that the ability of solving this type of task is hippocampal-dependent (Morris, 1981; Redish,

1999).

Following the framework presented in the previous section, a place-based strategy is encoded by state-action associations where the state space corresponds to possible locations within the environment. The spatial representation described in Chapter 5 is therefore used as state-representation. In other words, the modelled PC population projects onto a population of action cells AC^{P}, coding for responses associated to spatial information.

Equations (7.1) and (7.2) can then be used to select an action based on place-response, and Equations (7.4)-(7.6), to update the weight connections encoding the place-action associations.

The superscript P will be used in this document to denote variables corresponding to the place-response strategy, i.e. Φ^{P} and A^{P} are the direction of movement selected by this strategy and its corresponding action value.

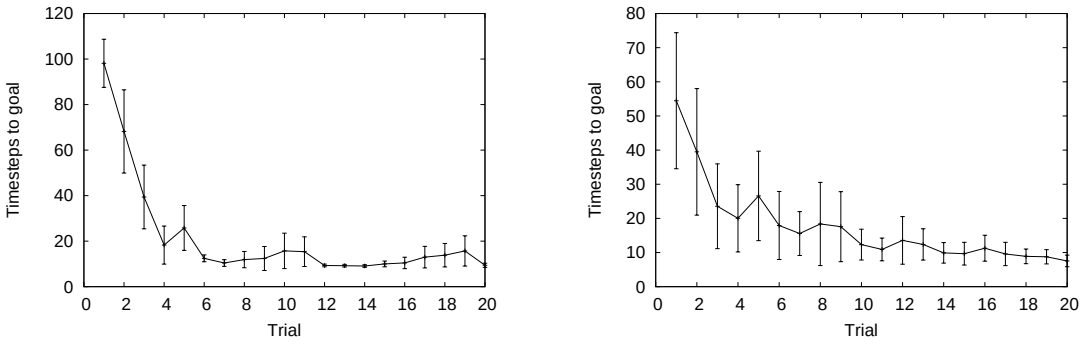
The performance of this module alone was tested by simulating the conditions of the hidden version of the water maze task. After random exploration of a square environment (120cm x 120cm) in order to establish a place representation (See Chapters 4-5), goal-directed behaviour was acquired and tested. At each time step the agent moves in the direction Φ^{P} computed using Equation (7.2). Whenever the agent reaches a circular goal area (diameter = 12 cm) a positive reward $R = 15$ is provided. In order to promote obstacle-avoidance behaviour, a negative reward $R_w = -5$ is given every time the agent hits a wall. No reward is given in other locations within the environment.

Each trial starts with the agent at a random location at least 30 cm away from the goal location. The trial ends when the agent reaches the goal, if it has not reached that location after 100 time-steps, it is algorithmically guided to the goal location. The mean escape latency over 100 repetitions is shown in Figure 7.3a. Consistently with previously reported results using similar approaches (Strösslin, Sheynikhovich, et al., 2005), the escape latency decreases with the number of trials. This indicates that the module is able to establish the correct associations between locations and actions, i.e. it develops a place-response strategy.

7.1.2 Implementing stimulus-response

Stimulus-response relies on associating a sensory cue to a specific response. A simple representation of the cues was used in this work. A horizontal one dimensional greyscale image (I) was used to drive a population of sensory cells (SC). Each sensory cell i has a narrow receptive field pointing in direction φ_i and its activity corresponds to the normalised greyscale value at that direction, $r_i^{\mathrm{SC}} = I(\varphi_i)$.

In simulations of visible goals we use a dark cue on light background to signal the goal location. As described above, SCs (coding for states) project onto a population of action cells AC^{C} coding for actions based on cue-information. The population activity



(a) (b)

Figure 7.3: Test of navigation modules. (a) place-response module. Mean escape latency in a simulated hidden water maze. (b) cue-response module. Mean escape latency in a simulated visible water maze. In both cases, the mean latency is computed over 100 repetitions.

of actions cells in this module provides a direction of movement Φ^{C} corresponding to a cue-response strategy, and its corresponding action value A^{C}.

Whenever the task requires cue approaching behaviour, the direction Φ^{C} should correspond to the allocentric bearing of the cue, allowing the agent to orient itself toward the cued goal and move in that direction.

The visible version of the water maze was used to test the navigation capabilities of this module. The same environment described in Section 7.1.1 was used, but a visible cue was put at the goal location. Moreover, the goal location was changed at every trial. As shown in Figure 7.3b, the escape latency decreases with the experience reflecting the acquisition of a cue-response behaviour.

7.2 Strategy selection and learning

The modules presented in Sections 7.1.1 and 7.1.2 implement place- and stimulus-response, respectively. Depending on the place representation and the sensory input, each module computes a direction of movement Φ^{P} and Φ^{C}, corresponding to place-based and cue-based strategies respectively. These modules also provide the corresponding Q-values (i.e. expected future rewards) of these actions. In fact, the action-values are given by the magnitude of the activation of the action cells.

The navigation system has to select one the two proposed actions. The action Φ to be performed by the agent is selected through a competition mechanism in which the probability $P(\Phi = \Phi^k)$ of moving in the direction Φ^k computed by a module

$k \in [\mathrm{P}, \mathrm{C}]$ depends in its action value A^k, and a gating factor g_k. Specifically,

$$P(\Phi = \Phi^k) = \frac{g_k A^k}{(g_{\mathrm{P}} A^{\mathrm{P}}) + (g_{\mathrm{C}} A^{\mathrm{C}})} \tag{7.7}$$

The gating values g_k (g_P and g_C, for place- and cue-response respectively) depend on the hippocampal activity (PC), as well as the sensory input coding for the presence of cues (SC).

$$g_k = \sum_{j \in \mathrm{PC}} (z_{kj}^{\mathrm{PC}} r_j^{\mathrm{PC}}) + \sum_{j \in \mathrm{SC}} (z_{kj}^{\mathrm{SC}} r_j^{\mathrm{SC}}) \tag{7.8}$$

where z_{kj}^{PC} and z_{kj}^{SC} are the weights of a connection from a pre-synaptic cell j to a gating unit k.

At each time step both modules propose a direction of movement (Φ^{C}, Φ^{P}) according to their perceived state; gating values are computed (Equation (7.8)), and used to select the direction of the next movement. To allow a chosen response to be applied during several time steps, competition among the modules (Equation (7.7)) is performed only when the accumulated prediction error (since the moment of the last competition) reaches a predefined threshold.

The parameters of the model should be updated in such a way that (i) *both* modules can improve their performance in the current task, but (ii) only the module with better reward prediction should increase its probability of being selected (i.e. increase its gating value).

In the TD(λ) algorithm (Sutton and Barto, 1998) weights are updated proportional to the reward prediction error δ_k (Equation (7.4)). Each module compute its corresponding error,

$$\delta_{\mathrm{P}}(t+1) = R_{t+1} + \ \gamma A^{\mathrm{P}}(t+1) - A^{\mathrm{P}*}(t) \tag{7.9}$$

$$\delta_{\mathrm{C}}(t+1) = R_{t+1} + \ \gamma A^{\mathrm{C}}(t+1) - A^{\mathrm{C}*}(t) \tag{7.10}$$

where δ_{P} and $\delta_{\mathrm{C}}(t)$ correspond to the reward prediction error for the place- and stimulus-response modules respectively. $A^{\mathrm{P}*}(t)$ and $A^{\mathrm{C}*}(t)$ correspond to the action values for the *selected* direction of movement $\Phi(t)$, as estimated by the place-response and cue-response modules respectively.

In our model the signal δ_k is scaled by a factor h_k depending on both the gating value and the prediction error for the module k (Baldasarre, 2002).

$$h_k = \frac{g_k c_k}{\sum_i (g_i \ c_i)} \tag{7.11}$$

where $c_k = \exp(-\rho\delta_k^2)$ ($\rho > 0$). For the module k weights will be updated according to

$$\Delta w_{ij}^k = \eta^k \ \delta_k \ h_k \ e_{ij}^k \tag{7.12}$$

Here η^k is the learning rate, for expert k, e_{ij}^k is the eligibility trace (as defined in Equation (7.6)). The scaling factor h_k ensures that weight modification is most significant in the modules with high gating values (more likely to be selected), and best reward prediction ($c_k \to 1$). It should be noticed that a module with a small gating value or a big reward prediction error will not be updated significantly. This allows the agent to preserve previously learnt responses which, if not useful in the present task might be required in the future. In contrast, in situations with no conflict (i.e. when both tasks are equally good solving the task), both modules will tend to have small prediction errors, leading to comparable gating values for the two strategies, and therefore both modules will be updated similarly.

The gating network weights are updated such that $g_k \to h_k$,

$$\Delta z_{kj}^{\mathrm{PC,SC}} = \xi(h_k - g_k) r_j^{\mathrm{PC,SC}} \tag{7.13}$$

with a learning rate ξ. This will assign higher gating values to the modules that consistently have small prediction error (Chavarriaga et al., 2005b; Baldasarre, 2002).

7.3 Learning algorithm

In this section we describe the procedure used to perform goal-directed navigation in the model. Initially the agent updates its state representations, that is the activity of both the place cell (PC) and the sensory cell (SC) populations is computed. Accordingly to the reinforcement learning paradigm, the agent then updates its estimation of the action value function as it interacts with the environment.

Based on the perceived state $s(t)$, the action value estimations (A^k) are updated in both modules, and the gating values g_k are computed. Both the gating values and the action values are used to select the direction of the next movement $\Phi(t)$. Once the agent has moved, the new state $s(t+1)$ is perceived and the reward prediction error (δ_k) is computed by subtracting received reward R_t and the reward predicted by the difference of the previous and current action values ($A(t) - \gamma A(t+1)$). The error signal δ_k, is then used to update the parameters of both the place- and cue-response modules, as well as the gating network.

In the following we describe the sequence of steps required to evaluate the action value estimations for both modules, perform the action selection process and to update all the model parameters.

1. *Calculate action values in both place- and cue-response modules. Action cells activity is computed using Equation (7.1). Population activity is used to compute the directions of movement* Φ^{P} *and* Φ^{C} *, and their corresponding Q-values,* $A^{\mathrm{P}}(t)$ *and* $A^{\mathrm{C}}(t)$*.*
2. *Compute the gating values,* g_{P} *and* g_{C}*, using Equation (7.8).*
3. *Select a direction of movement* Φ *according to Equation (7.7).*
4. *Generalise in action space: Enforce an activity profile* $r_i(t)$ *(Equation (7.3)) around the selected action* Φ*, in the action cell populations of both place- and cue-response modules.*
5. *Update eligibility trace* $e_{ij}(t)$ *in both modules.*
6. $t \rightarrow t+1$*. Perform the movement and update the state representations.*
7. *Calculate action values* $A^{\mathrm{P}}(t+1)$ *and* $A^{\mathrm{C}}(t+1)$*.*
8. *Calculate reward prediction error* δ_{P}*,* δ_{C} *for both modules (Equation (7.9)).*
9. *Compute scale factor* h_k *(Equation (7.11)), and update synaptic strengths (Equation (7.12)).*
10. *Update gating network weights (Equation (7.13)).*

The set of parameters used in the simulations presented in the next section is shown in Table 7.1.

Place- and Cue- response modules

Parameter	Value	Description (Eq.)
N_{AC}	36	Number of action cells
σ_{ac}	22.5°	SD of the enforced activity profile (7.3)
$\eta^{\mathrm{P,C}}$	0.01	Learning rate (7.12)
$\gamma^{\mathrm{P,C}}$	0.8	Reward decay factor (7.9)
$\lambda^{\mathrm{P,C}}$	0.95	Trace-decay factor (7.6)

Combined place coding

Parameter	Value	Description (Eq.)
ξ	0.3	Learning rate of the gating network (7.13)

Table 7.1: Main parameters of the navigation model.

7.4 Results

We test the navigation model using the same simulation environment as described before. At each time-step the agent moves 6 cm in the direction Φ provided by the action controller. The simulation environment provides external sensory input (linear vision and distances to the walls) and odometry information from the agent. Before each experiment all the weights are randomly initialised. All the experiments were simulated in a square environment of 120×120 cm, with a target having a diameter of 12 cm. Three types of experiments were performed testing the model capabilities to *(i)* solve a task using a taxon strategy (i.e. the visible water maze); *(ii)* solve a task requiring a locale strategy (i.e. hidden water maze); and *(iii)* develop both strategies simultaneously (i.e. combined cue-place learning).

All the experiments consists of 10 blocks, with 4 trials per block. In the combined cue-place learning task, the 10th block corresponds to the competition trials as described in (Devan and White (1999); see Section 6.4.1). Each trial starts with the agent located at a random position at a minimum distance of 70 cm from the centre of the goal (requiring the robot to make at least 10 steps to get to the goal location). A trial is finished when the agent reaches the target location and at that moment a positive reward is given. If the agent is not able to reach the goal within 100 time-steps, it is guided to the platform, analogous to the experimental procedure with rats (Devan and White, 1999).

In the first experiment, the visible version of the water maze is modelled by placing the agent in an environment with white walls and a dark, visible cue in the place were a reward is provided. The location of the platform is changed in every trial such that spatial information cannot be used to solve the task. The navigation maps for both the place- and the cue-response modules at different stages of training are presented in Figure 7.4. At each point the lines shows the selected direction Φ^k for that specific module, the length of the line is proportional to the action value (A^k) associated to a movement in that direction. They show that the cue-response module is able to learn the association between the sensory input and the appropriate action to solve the task. As the spatial information is not useful to reach the visible goal, the place-response module is not able to build a proper association between location and action.

The second type of experiments test the ability of the system to develop place-based strategies in order to find a hidden goal. The goal is solely defined by the fixed location of the reward. In particular there were no visual cues signalling its position. Navigation maps for both modules are presented in Figure 7.5. Opposite to the previous case, the place-response module is able to learn the correct action required to take the agent toward the invisible goal. As no cue is available at the location of the goal, the cue-response strategies cannot solve the task.

In order to test the simultaneous development of both place- and cue-response, we simulate the experimental paradigm described in Section 6.4.1 (Devan and White,

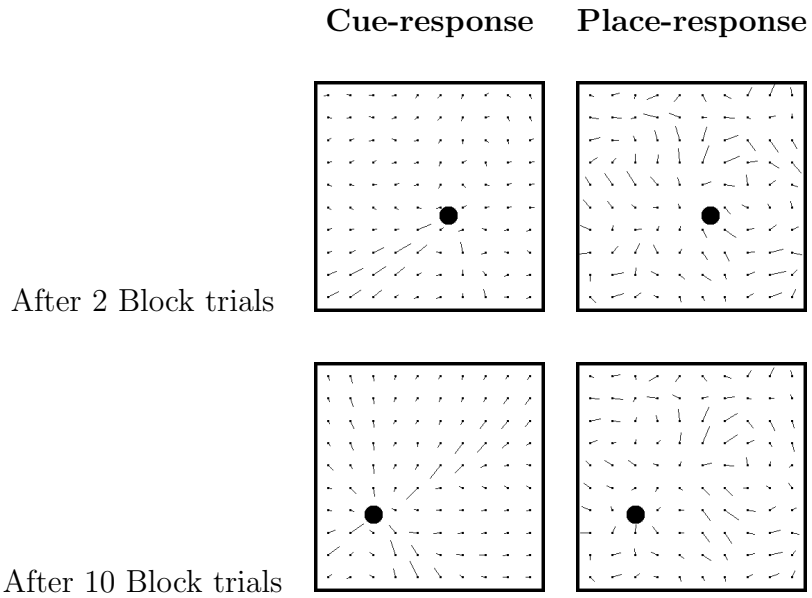


Figure 7.4: Navigation maps after training in the *visible* version of the water maze. *Left.* Cue-response module. *Right.* Place-response module.

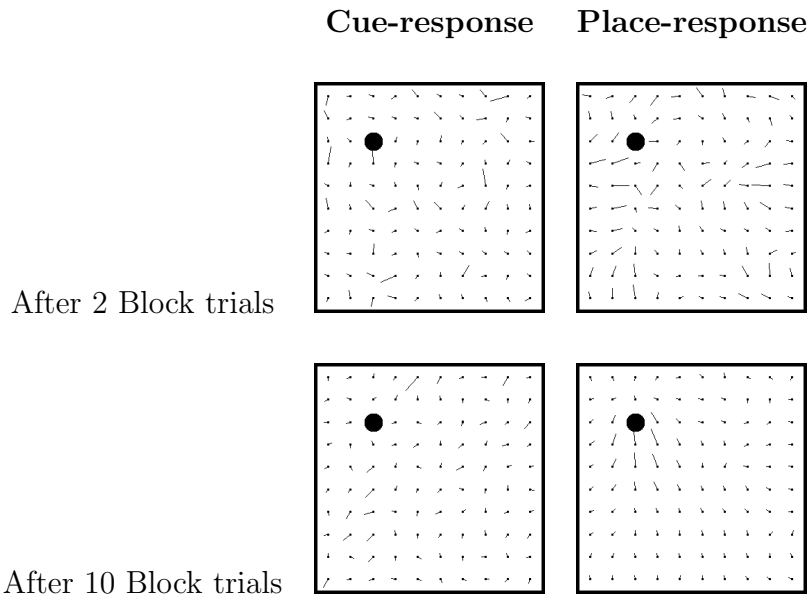


Figure 7.5: Navigation maps after training in the *hidden* water maze. *Left.* Cue-response module. *Right.* Place-response module.

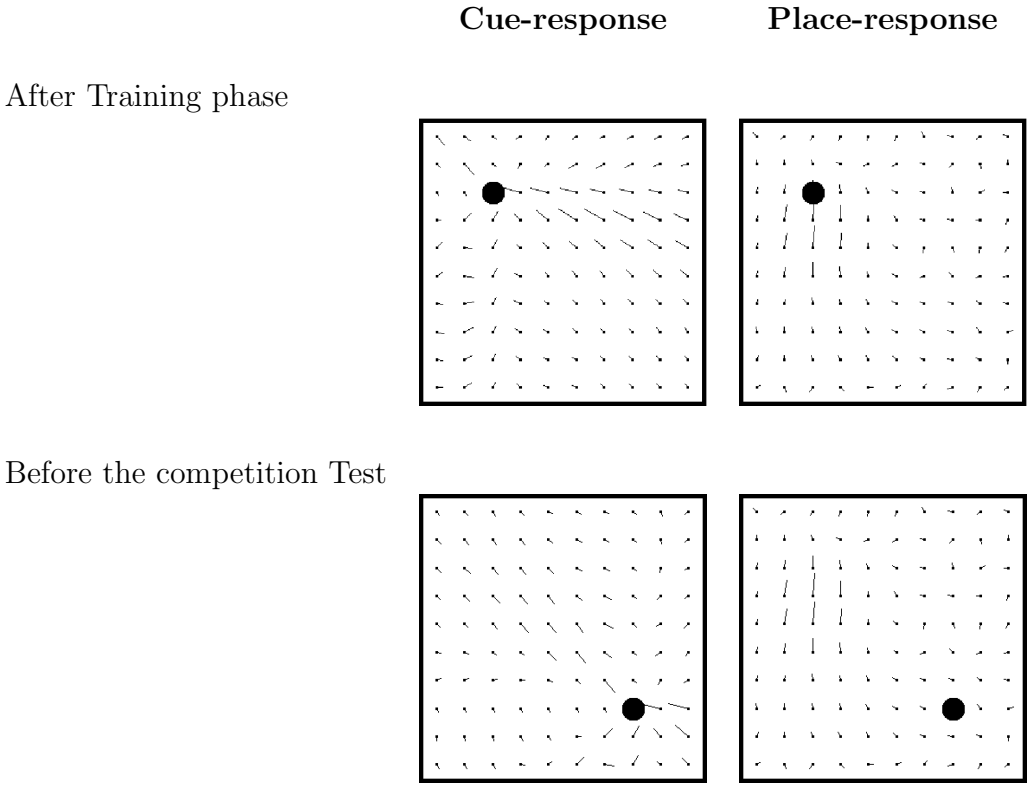


Figure 7.6: Navigation map at the end of the training phase in the combined cue-place learning task (Devan and White, 1999). *Top.* The filled circle marks the location of the platform in both the hidden and visible training trials. *Bottom.* Competition trial. The visible target has been moved to a new location. *Left.* Cue-response module. *Right.* Place-response module.

1999). During training in an environment with white walls, in blocks 1, 2, 4, 5, 7, and 8, the goal location was signalled by a dark visible cue, such that the task may be solved by either strategy. On blocks 3, 6, and 9, no visible cue was available, requiring the use of spatial information to solve the task.

Figure 7.6 shows the navigation maps for both the modules. It can be seen that, after training, both modules are able to guide the agent toward the goal location (filled circle). It should be noticed that they have simultaneously learnt to solve the task by using a different type of information. This becomes clear once the visible goal is moved to another location (Figure 7.6 bottom) and the cue-response strategy guides the agent to the new location, whereas the place-response module still points toward the location of the goal in the training phase.

The averaged escape latency (time-steps required to reach the goal) over 10 sim-

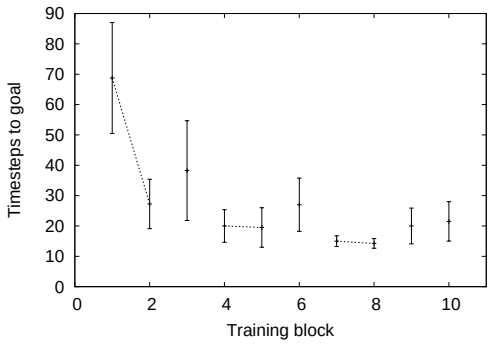


Figure 7.7: Escape latency (time steps taken by the agent to reach the platform). Each training block consists of 4 trials, and each trial starts with the agent located at a random position (selected such that it takes at least 9 steps to reach the goal). *Disconnected* points correspond to the trials in the hidden water maze (blocks 3,6, and 9) and the competition test (block 10).

ulations is showed in Figure 7.7. The simultaneous development of place- and cue-response strategies can be observed as the model improves its performance in both visible (blocks 1,2,4,5,7, and 8) and hidden (blocks 3,6, and 9) trials. This result qualitatively reproduces those reported by Devan and White (1999), as shown in Figure 6.2b.

This experiment was designed to dissociate the response to the different kinds of information (i.e. the use of different navigational strategies). Representative trajectories at the competition trial for two runs of the simulation are shown in Figure 7.8. As observed in intact animals (Figure 6.2), in some cases the agent goes first to the place where the platform was during training (place-response), and then changes the strategy and then moves toward the cued platform, whereas in other cases, the agent adopts follows a cue-response from the beginning and swims directly toward the visible goal.

The predicted future reward (normalised to one) for both modules at different stages of training is shown in Figure 7.9. It corresponds to the action value (A^k) of the action selected at different points in the environment (it shows the action value corresponding to the directions of movement shown in the navigation map Figure 7.6). At the end of block 8 (the last block with the visible platform), both modules successfully predict the location where the reward is delivered. As shown above, before the competition trial, spatial information still leads the agent toward the previous location of the platform. The cue-response, in contrast, points toward the landmark (and goal) location. If training continues in the competition situation (and the goal remains at the same location), the place-response module will start

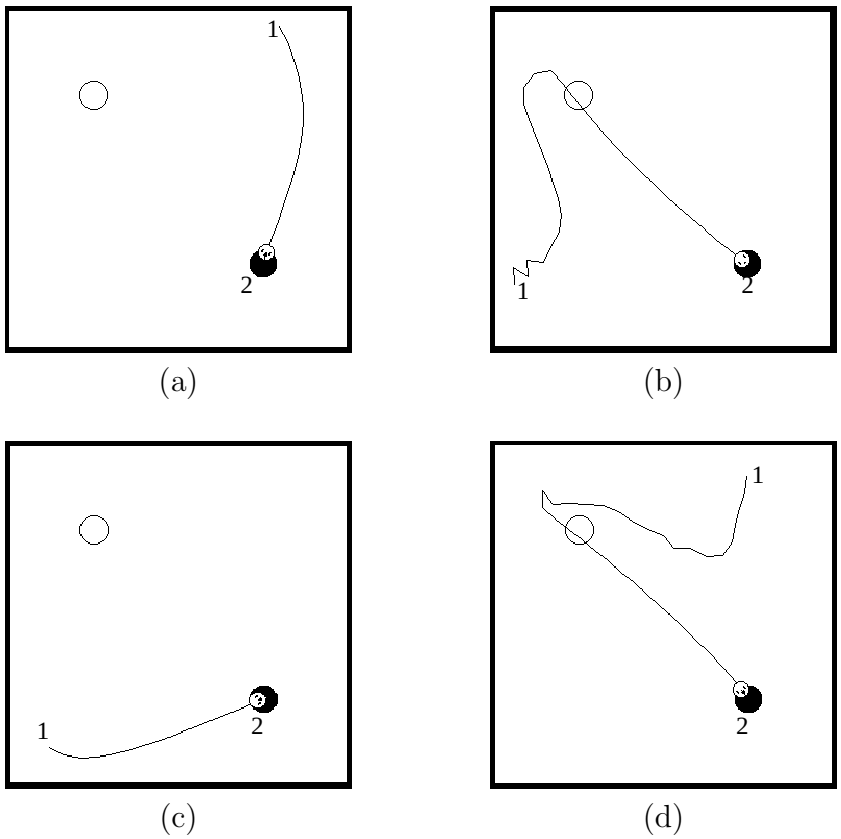


Figure 7.8: Examples of trajectories observed during the competition trial. In some cases the agent goes directly toward the visible goal (a) and (c) by applying cue–response behaviour. In (b) and (d), a place–based strategy is selected leading the agent to the place were the platform was located during training (open circle). A cue-response is then adopted allowing the animal to reach the goal (filled circle). 1. Starting position. 2. Final position of the agent.

learning the new location of the goal.

Figure 7.10 presents the probability after training for every module to be selected in the experiments shown above (average over 10 simulations). When the agent is solving the visible maze, the gating system is more likely to select cue-responses over place-responses. As opposed to the hidden maze condition, where the locale strategy is preferred. In the case of the combined cue-place task, after training (block 9) both strategies can solve the task. Hence, at that point of training both modules have about equal probabilities of being selected. Once the agent is exposed to the conflict condition (competition test), the place-response is no longer suitable to solve the task. This causes a change of strategy in the middle of the task in those cases

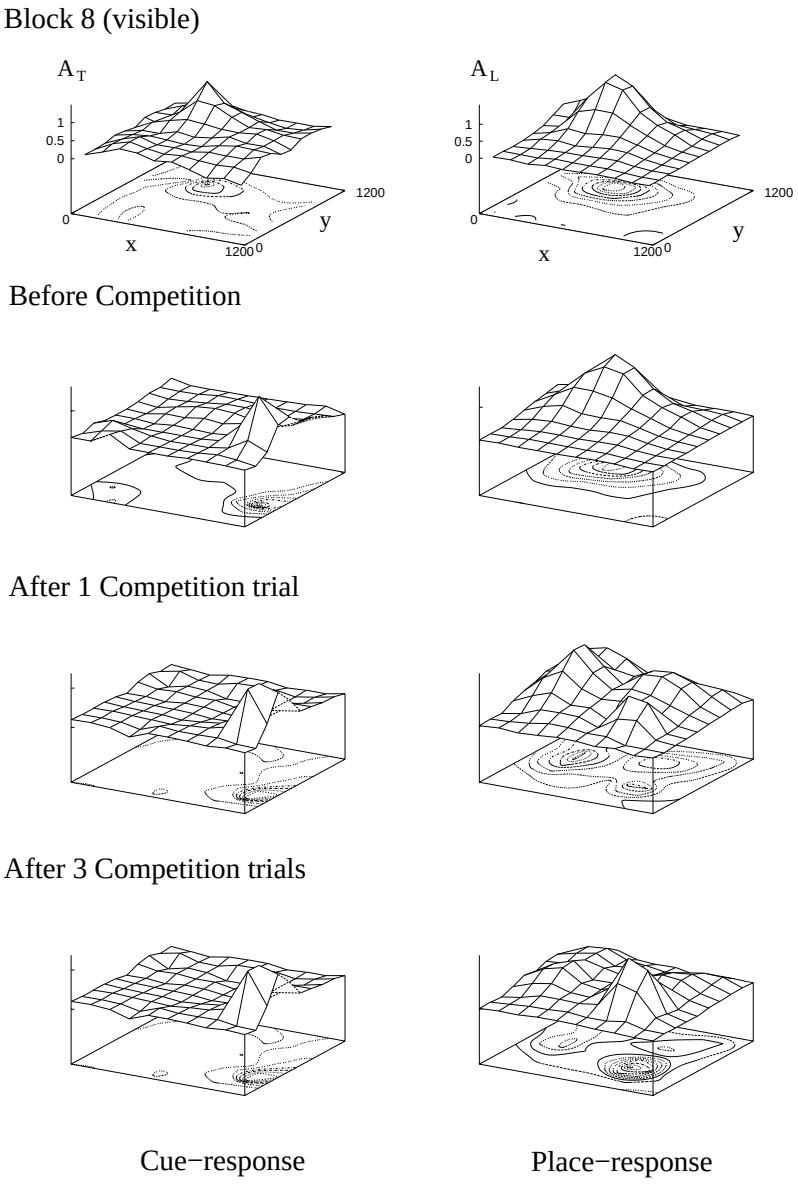


Figure 7.9: Future reward predicted by each module at different stages of training. From top to bottom: *End of block 8*, last training block with the visible goal. Both modules predict the maximum reward at the location of the goal (x=300,y=900). *Before competition*, the taxon module predicts the maximum reward at the new location of the landmark (x=900,y=300). The locale strategy still leads to the same location as before. *After 1 and 3 competition trials*, the locale strategy starts to learn the new location but still predicts a high reward for the former location of the platform.

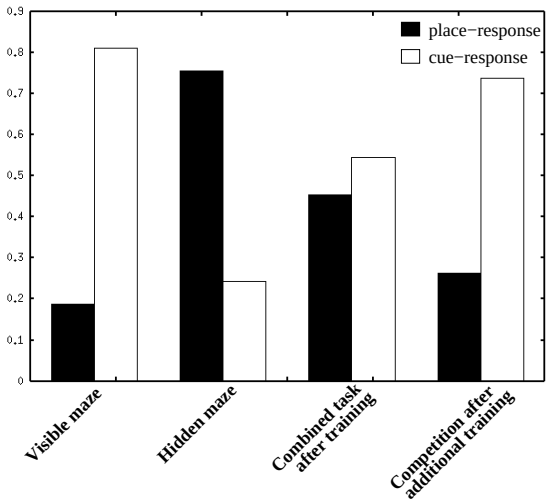


Figure 7.10: Probabilities for the different modules to be selected. The cue-response module has a higher probability of being selected when the agent solves the visible maze and, conversely, place-response is preferred in the case of the hidden maze. After 9 blocks of training in the combined cue-place task, both modules have rather equal probabilities of being selected, as both strategies are able to solve the task. The last two columns show the probability of selecting place- or cue-based strategies after three blocks of trials with the *competition* set up, i.e. the platform is visible but at a location different from the one encountered during the previous training phase. Since the learnt place response does not solve the task, the probability of selecting cue-responses increases.

where a locale strategy was chosen at the beginning of the trial (Figure 7.8). Finally, if the agent is exposed during several trials (three extra training blocks, in this case) to the new condition, and the locale strategy persistently fails in solving the task, the probability of selecting cue-responses will increase.

A facilitation effect has been proposed to be a consequence of the competition among different memory systems. That is, lesions in one system improve the performance in a task associated with the unlesioned system. It has been observed that hippocampal lesions improve the performance in S-R tasks, presumably by releasing the striatum from competition (Packard et al., 1989; Packard and McGaugh, 1996; Devan and White, 1999). According to this, an agent following exclusively a taxon strategy should have smaller latencies in the test block than an agent controlled by the competition mechanism. In the Figure 7.11 we compare the performance in the competition trial (as showed in Figure 7.7) to the performance obtained if a cue-response is always selected. As expected, the deactivation of the place-response results in shorter latencies in the competition test (a S-R task).

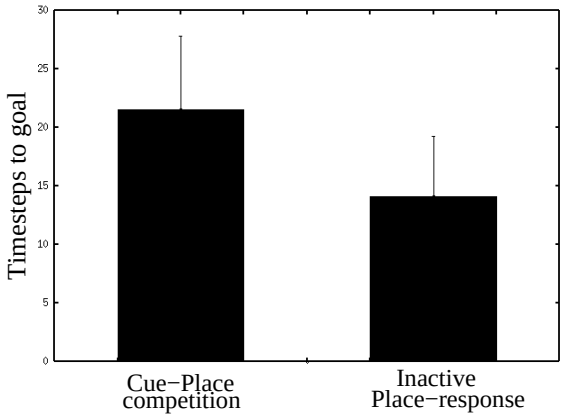


Figure 7.11: Deactivation of the place-response module improves the performance in the competition trial. When both cue and place-response modules compete for action selection, the agent takes on average more time to reach the visible goal than when only the cue-response module is functional.

In another experiment, Pearce et al. (1998) train rats to search for a platform located at a fixed distance and direction from a single landmark. The training consists of eleven blocks of four trials. At the beginning of each block both the landmark and the platform were moved to a different location, and remained in the same place for the entire block session (Figure 7.12a). In this experiment the landmark gives correct, although not precise information about the location of the goal; whereas place information will be disrupted every time a new block starts. Consistently with this, animals with hippocampal lesions perform better in the first trial of each block than control animals, presumably guided by the landmark location. At every new block, intact animals tend to explore the region of the pool where the platform was previously located before swimming to the new location indicating the use of spatial information to solve the task (i.e. locale strategy). Lesioned animals, in contrast take a more direct path to the platform (and landmark) location.

Despite their poor performance in the first trial, control animals significantly improve their performance during each block of training whereas rats with damage to the hippocampus do not. This suggests that lesioned animals rely on the less precise information from the landmark location (a stimulus-response behaviour) while control rats also use spatial information to solve the task. Every time a new block starts, as spatial information is no longer suitable to reach the platform, the rat changes its navigation strategy and uses the information provided by the landmark location, similar to the place responder behaviour reported by Devan and White (1999).

Figure 7.12b shows the results of our model using this paradigm. We compare the

performance of the competition model (*Control*) with the results of using stimulus-response exclusively (*S-R*) as described in section 7.1.2. The figure shows the average latency over ten simulations for the first and last trial of each block. The performance of the competition mechanism is consistent with the results reported by Pearce et al. (1998) for control animals; it yields longer latencies for the first trial of each block but latencies significantly decrease during the three subsequent trials. The performance of a pure taxon strategy is not negatively affected by the start of a new trial, but it does not improve during the training block.

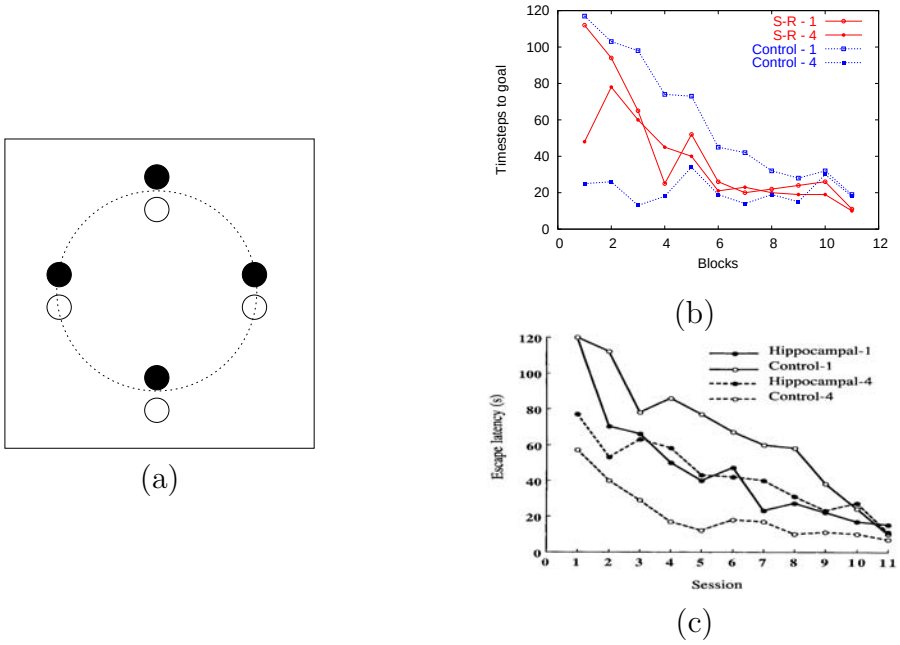


Figure 7.12: Experimental paradigm analogous to the one used by Pearce et al. (1998). (a) The experiment takes place in a square arena (120x120 cm) with a single landmark. The landmark can be located in one out of four possible positions (filled circles). The goal (open circles) is always located at a fixed distance and direction from the landmark. (b) Mean latency in the first (*Control-1* and *S-R-1*) and fourth (*Control-4* and *S-R-4*) trials of each training block. Control: actions are selected by the competition of locale and taxon strategies. S-R: Actions are selected using exclusively a stimulus-response strategy (as described in section 7.1.2). (c) Experimental results reported by Pearce et al. (1998).

7.5 Discussion

We have presented a system-level computational model for navigation based on the competition among different modules implementing independently place- and cue-response. Consistent with experimental data, the modules are able to develop simultaneously the proper strategy to solve navigational tasks. Moreover, the model develops a pure cue-response, if the location of a cued platform varies from one trial to the next; and a pure place response if the location of the hidden platform remains stable.

Each module implements a navigation strategy by learning associations between states and action (encoded in neuronal populations) using reinforcement learning. One module, implementing place response, is driven by the place coding population (PCs) described in Chapter 5, while the module implementing cue-response is driven by a population coding for simple sensory stimulus (SC population).

As a result of the competition, the agent is able to change from one type of response to another when it is no longer suitable to solve the task (Figure 7.8). Additionally, experiments have shown that the performance on a given task can be improved by lesioning unrelated competing neural systems (Packard et al., 1989; Devan and White, 1999; Chang and Gold, 2003a). In our model, deactivation of the place-response module yield smaller escape latencies in the competition test of the experimental paradigm proposed by Devan and White (1999) (Figure 7.11).

Both modules apply reinforcement learning to learn associations between state (place or cue) representations and actions (Sutton and Barto, 1998). The update of the parameters of the system is modulated by the gating system (Equations (7.11–7.12)) such that modules with better performance (small prediction error, δ) and high gating values (g_k) will be more strongly modified.

Simulations of experimental paradigms were performed to test the model. These experiments were aimed to,

- Assess independently, the capability of the place- and cue-response modules to learn appropriate associations between states (location and presence of cues, respectively) and actions (Figure 7.3).
- Test the ability of the whole navigation model to select the appropriate response in the hidden and visible versions of the water maze (Figures 7.4,7.5 and 7.10).
- Reproduce the behaviour in combined place-cue learning tasks (Figures 7.6 – 7.12). In one of these tasks (Devan and White, 1999) the model selects the appropriate strategy (either cue- or place-response) depending on whether the platform is visible or not (Figure 7.6).
- Show that the system can change the strategy in the middle of the task if the chosen strategy does not lead it to the goal location (Figure 7.8).

- Show that the both strategies interfere with each other in the competition for action control. Accordingly, inactivation of the module implementing one type of response improves the agent's performance in tasks associated to the remaining, intact strategy (Figure 7.11).

7.5.1 Putative neural substrates

Neural locus of navigation strategies

A substantial body of experimental results supports the theory of multiple brain systems acting in parallel and mediating different forms of learning (See review in Chapter 6). Based on these results, we propose that the hippocampus and the ventral striatum (NAc) are responsible for spatial learning (place-response) (Morris, 1981; Redish, 1999), and the dorsolateral striatum is responsible for S-R behaviour (cue-response).

Hippocampal neurons project onto the NAc via the fornix fibre bundle. Lesions of the fornix or nucleus accumbens impair rat navigation in the hidden version of the water maze, but do not affect navigation toward a visible platform (Eichenbaum et al., 1990; Sutherland and Rodriguez, 1990; Packard and McGaugh, 1992; Whishaw et al., 1995). Lesions in the dorsal striatum have the opposite effect, they decrease performance when the platform is visible (Packard and McGaugh, 1992; McDonald and White, 1994; Devan and White, 1999).

This contrasts with the predictions of Girard (2003), which proposes that all navigation strategies are implemented in the ventral pathway of the basal ganglia, and the dorsal pathway is in charge of non-locomotor actions. This prediction is inconsistent with studies showing that lesions in the fornix or nucleus accumbens do not prevent the animal to navigate based on available sensory cues (Eichenbaum et al., 1990; Sutherland and Rodriguez, 1990; Whishaw et al., 1995).

System interactions

How different systems in the rat brain interact in the competition for action selection is still an open question. One possibility for the competition is that a brain structure, outside the striatum and hippocampus, receives the information processed by those systems and controls the expression of learnt responses (Gold, 2004; Mizumori et al., 2004). Also an indirect way of communication between systems via the cerebral cortex, has been proposed by White (2004). He suggests that output coming from one of the systems could change the representation of information stored in the cortex that will be, in turn, sent to all other competing systems. Currently experimental support for these hypotheses is still missing.

In the water maze the medial region of the dorsal striatum seems to be involved

in the competition process (Whishaw et al., 1987; Devan and White, 1999) as lesions in this area did not prevent spatial learning, but produced a preference for cue-response when cue information was available. This suggests a form of interaction via the dorsomedial striatum, in which hippocampal-mediated behaviour can override cue responses. The role of the basal ganglia in action selection has been modelled as a competition among different sensory-motor programs (implemented by parallel cortico-striatal loops) (Doya, 2002; Gurney et al., 2001a). This competition may include actions corresponding to place-responses as one of the motor programs participating in the competition. Spatial information from the hippocampus may be transferred via the ventral and dorsomedial striatum, and compete with stimulus-response programs encoded in the dorsolateral striatum.

Role of the neuromodulators

Dopamine

As reviewed in Section 7, dopaminergic activity in both ventral and dorsal striatum may provide the reward error signal (δ) used to update the parameters of the model. In particular, the phasic activity of dopaminergic neurons in SNc and the VTA correlates with the difference between actual and expected reward. Suggesting that these areas may implement the temporal difference computation used in the reinforcement learning paradigm (Houk et al., 1995; Schultz et al., 1997; Schultz, 1998, 2002).

Dopaminergic neurons project onto both the ventral and dorsal striatum (Whishaw and Mittleman, 1991; Schultz et al., 1992; Legault et al., 2000). They tend to synapse on the basal part of spines which host synapses from cortical (Freund et al., 1984; Sesack and Pickel, 1992; Smith et al., 1994) as well as hippocampal (Totterdell and Smith, 1989; Floresco et al., 2001) afferent projections. Therefore, dopaminergic activity may be able to modulate synaptic plasticity of cortical and hippocampal afferents to the striatum (Houk et al., 1995; Schultz et al., 1997; Schultz, 1998).

Acetylcholine

Even though the biological mechanism for the competition has yet to be clarified, recent experiments suggest that cholinergic activity might convey information related to the strategy selection mechanism (alike to the gating signals in our model). Several experiments provide evidence for the cholinergic role in modulating learning and memory (M. E. Ragozzino and Gold, 1995; M. E. Ragozzino et al., 1998; Hasselmo, 1999; Chang and Gold, 2003b). See Section 6.4.3.

In addition, it has been proposed that the role of acetylcholine (ACh) is to balance the contribution of different neural systems in learning a given task (Gold, 2003, 2004). Simultaneous measure of acetylcholine in the hippocampus and the striatum showed a correlation between the changes of ACh release and the expression of place-response or S-R behaviour (Chang and Gold, 2003b; McIntyre et al., 2003). Moreover, relative

levels of ACh in the hippocampus and striatum *prior* to training reliably predict the preference for place-response over stimulus-response (McIntyre et al., 2003). These results give further support to the hypothetic role of ACh as biasing the preference for a given strategy, which may be related to the gating values g_k in our model. Consistently, Dayan *et al.* (Dayan et al., 2000; Yu and Dayan, 2002, 2003) have proposed a statistical framework in which cholinergic levels act as a gate to learning in a specific neural system by reflecting the (expected) uncertainty in its predictions.

In addition, ACh has been proposed as regulating the speed of memory update (Hasselmo and Bower, 1993; Hasselmo, 1999), which in modelling terms may correspond to the learning rate (Doya, 2002). This dual role of ACh biasing the strategy selection and modulating the learning speed in the different systems implementing each strategy is consistent with the role the scale factor h_k (Equation (7.11)) plays in the learning rules for the gating system (Equation (7.13) updates the gating network such that $g_k \rightarrow h_k$), and the separate navigation modules (Equation (7.12)).

The question of what mechanism regulates the relative levels of ACh in the different structures involved in learning remains unsolved. One possibility is that an input system (to be identified) controls the ACh release in the forebrain. However, neuroanatomical differences of the cholinergic system in the hippocampus and striatum[1] suggest that this regulation is more likely due to pre-synaptic mechanisms of release within each neural structure (Gold, 2004).

7.5.2 Limitations

The proposed model focuses on the computation underlying the strategy selection in navigational tasks. This selection is accomplished as a competition among different modules, where each module implements a different strategy. The likelihood of a module to be selected for action control depends on the future reward it predicts and the accuracy of this prediction. The strategy selection process is computed in an algorithmic way. Therefore, it does not attempt to implement a biologically plausible mechanism for this process and cannot be used to test hypothesis about the putative role of different brain structures in the selection between cue-response and place-response.

The selection between place and cue-responses in the model depends on the performance of those strategies during training, as indicated by errors in their predictions of reward. According to this, if both strategies successfully solve the task they will have equal probability of being selected. However, animals gradually shift from place to cue response when trained extensively (Packard and McGaugh, 1996; Chang and Gold, 2003b). Such a transition from hippocampal-dependent place-responses to

[1] ACh in the striatum is derived from intrinsic cholinergic neurons, whereas hippocampal ACh is derived from projecting neurons from the basal forebrain cholinergic neurons.

dorsostriatal-dependent stimulus-responses is not reproduced by the model.

After training, despite the preference for cue-responses, the hippocampal structure seem to keep engaged in the task as it is indicated by the high levels of ACh release in the hippocampus (Chang and Gold, 2003b), and the fact that place-response can be restated by reversible striatal lesions (Packard and McGaugh, 1996).

The assumption that the strategy selection depends exclusively on the expected performance of each system implies that after extensive training the striatal system will consistently yield smaller prediction errors than the hippocampal system (Daw et al., 2005). There is no experimental evidence supporting such an assumption. An alternative hypothesis is that animals are biased toward cue-responses when both responses are equally good solving the task. This bias can be due to the fact that the dorsolateral striatum can take advantage of direct connections to motor and somatosensory cortices, while the hippocampus does not have such connections. In consequence, after extensive training, when both responses accurately solve the task, rapidly elicited stimulus-responses will overcome hippocampal-dependent place-responses, producing the shift from place to cue-responses.

In modelling terms, after extensive training, the competition described in this chapter will yield approximately equal gating values (analogous to asymptotic values of cholinergic levels reported by Chang and Gold (2003b)), and an additional mechanism will privilege the execution of cue-responses. In his model, Girard (2003) proposes that stimulus-response associations encoded in the dorsal striatum may prevent the execution of navigational actions coded in the ventral striatum via the subthalamic nucleus. However, this interaction depends on a fixed parameter and no learning mechanism is proposed. Therefore, it is not able to reproduce the gradual change from place- to stimulus-responses observed in animals.

Chapter 8

Conclusions

As stated at the beginning of this document, the work here presented is aimed to address two fundamental questions: (i) How can a representation of space, alike to the one found in the hippocampal place cells be constructed based on both internal and external sensory information; and (ii) How can a bio-inspired agent select among different navigational strategies depending on the characteristics of the task to be solved and its own experience. This chapter resumes the main contributions of the present work in resolving those questions and presents some of the limitations and possible improvements of the proposed model.

8.1 Contributions

8.1.1 Learning spatial representations

This work presents a model, based on the artificial neural network paradigm, which captures some of the physiological features of place-sensitive cells found in the rat hippocampus. The model follows the rationale of previous models developed in our laboratory (Arleo et al., 2004; Arleo and Gerstner, 2000). Two pathways coding for external (i.e. allothetic) and internal (i.e. idiothetic) sensory input converge onto a distributed place representation encoded in a population of rate-coded cells.

The *allothetic* pathway builds a vision based place representation based on storage and recognition of previously experienced local views (VC population). Information from several, corresponding local views is further combined in a second population where a more stable representation is built (APC, for Allothetic place cells). A population of cells driven by path integration composes the idiothetic pathway. These

cells, arranged according to predefined spatial preferences are algorithmically updated according to the integration of odometry signals.

The model is built incrementally as the agent explores the environment, and interactions between the two pathways are provided to allow re-calibration of path integration, and to disambiguate visually similar information corresponding to different locations (i.e. perceptual aliasing).

In this thesis we have studied several properties of the resulting place code which deserve to be highlighted. First, the sparseness of the place code changes according to the presence of external stimuli. When exploring open environments, place fields at the centre of the arena tend to be broader than those in the periphery (Section 4.3.1). This is in agreement of experimental data comparing the sparseness of the hippocampal code in cue-rich vs. cue-poor environments (Battaglia et al., 2004).

Place cells in the new model change their directionality depending on the exploratory behaviour of the agent. Random exploration of open environments yields directional independent activity, whereas directed trajectories or exploration of linear tracks results on directionally dependent place cells (See Section 4.3.2, 5.4.1 and 5.4.2). Changes in the directional selectivity of modelled place cells reproduce experimental findings in rats, resolving an open issue of previous models developed in our laboratory.

A whole set of experimental paradigms was applied to test the performance of the model. Besides the aforementioned exploration of radial mazes, changes in the place code as a result of geometrical manipulation of the environment was also assessed. Moreover, realignment of the place code upon mismatch of internal and external cues; and reset and calibration of the path integrator based on the allothetic representation were also tested.

The model proposes that path integration takes place in an extra-hippocampal structure, the recently found *grid cells* in the mEC being a possible locus for this process. A direct connection from allothetic coding to the path integrator is also proposed (Section 5.2). This predicts that changes in the activity pattern of place-sensitive cells in the mEC as a result of cue manipulation may take place even after disruption of hippocampal afferents to the EC.

There is no conclusive experimental evidence on whether or not place fields are intrinsically directional. Our model proposes that they are, and that the directionality may be lost as the animal explores the environment through Hebbian modification of the VC $\rightarrow$ APC connections. Non-directional cells in the mEC reported by Fyhn et al. (2004) support the model prediction that changes in the directionality properties of place-sensitive cells may take place upstream of the hippocampus, opposite to the predictions of the model proposed by Brunel and Trullier (1998).

8.1.2 Navigation model

We have proposed a system-level model of navigation able to reproduce the behaviour of animals in several situations, using a model of multiple experts with reinforcement learning. It relies on the theory of parallel navigating systems competing for selection of the most appropriate action. Based on the results of neurobehavioural experiments, we propose those systems to involve the hippocampus and the ventral striatum for place-based strategies (Morris, 1981; Redish, 1999), and the dorsolateral striatum implementing stimulus-response strategies (Packard and Knowlton, 2002). The strategy selection depends on the characteristics of the task to be solved and the training process (White and McDonald, 2002).

Navigational strategies are learnt as a mapping between states and actions acquired by means of reinforcement learning. Each module is composed of two populations of rate coded cells. One of these populations codes for the perceived state. In the case of the place-response module, the state (i.e. spatial information) is encoded by the PC population described in the first part of this thesis. In the cue-response module, a population of sensory cells code for the presence of cues (potentially signalling a reward site). Each *state coding* population projects onto a population of *action cells*. Each action cell corresponds to one direction of movement, and its activity codes the estimated Q-value of performing that action. Thus, connections from the state-coding population to action cells code a parametric action value function (Q-function). A temporal difference (TD) algorithm is used to update the synaptic weights in order to improve the estimated returns (and, in consequence, improve the agent's behaviour).

The model selects among the actions proposed by the place-response and cue-response modules using a gating mechanism. Gating values are learnt such that modules showing small reward prediction error will have a higher probability of being selected for action control. Furthermore, the gating value also modulates the learning rate of each module, such that a more significant update will be done in modules with high gating values, and consistently small prediction errors.

Different experimental paradigms were applied to test the performance of the model. Simulations of the *visible* and *hidden* version of the water maze show that the model is able to develop cue- and place-response respectively. Furthermore, the gating network assigns a higher probability of selection to appropriate strategy in each case, i.e. place-response is used when looking for hidden goals, while cue-response in used to reach the visible platform.

Moreover, two experimental procedures in which both strategies develop simultaneously were simulated. The model qualitatively reproduces the acquisition of a learning paradigm using interleaved hidden and visible trials in the water maze (Devan and White, 1999). Besides this, in certain conditions it is able to change the strategy during the task in order to reach the goal.

When compared to the performance of a pure cue-response strategy, the navigation model reproduces the effect of hippocampal lesions in conflict conditions (i.e. competition test in (Devan and White, 1999), and inter-block differences in (Pearce et al., 1998)). Such effects are believed to be the consequence of the competition of different brain systems for action control.

8.2 Limitations and perspectives

In the same way as this work is based on previous modelling efforts, several improvements to the present model can be developed in order to enhance its biological soundness and its ability to reproduce neurophysiological and behavioural data. This Section presents some of these improvement

8.2.1 Learning spatial representations

Role of other hippocampal structures. The feed-forward architecture of the model is proposed to correspond to the entorhinal-CA1 pathway in the rodent hippocampus. While this pathway has been shown to be sufficient to maintain a stable place code and show the firing properties reproduced by the model (Brun et al., 2002; Fyhn et al., 2004), it remains the question of what is the role of the other hippocampal structures not included here.

Dentate gyrus and CA3 are characterised by their recurrent connectivity. Attractor dynamics in these structures may be responsible of the process of orthogonalisation and completion of incoming patterns. These processes may contribute to enhance the spatial and directional selectivity of place cells in the CA1 and mEC.

Realistic sensory input. The current model has been tested using simple, abstract signals as external sensory input (c.f. distances to surrounding walls). A natural next step consists in testing the principles proposed here in a system dealing with more complex sensory input. Currently, the calibration of the path integrator, and the realignment of the place code described in Section 5.4.3 have already been implemented using realistic visual inputs (Sheynikhovich et al., 2005; Strösslin, Chavarriaga, et al., 2005).

The use of a more powerful visual system will allow to further test differences in the resulting place code when the animal explore cue-rich vs cue-poor environments. In the present work, such differences were observed by comparing the place code sparseness at the centre and periphery of the environment (Figure 4.4).

Alignment of local views. The alignment of local views as presented here, requires knowledge of the heading direction corresponding to the stored and present view. Although a head direction system can provide such signals, information already present in the image can be used to perform the alignment. Recently, an alignment mecha-

nism based on the correlation of the two local views has been proposed and tested with two-dimensional visual inputs (Sheynikhovich et al., 2005). The main drawback of this approach is its computational cost.

8.2.2 Navigation model

A parallel has been drawn between the classification of locale and non-locale strategies and the division between episodic (or declarative) and procedural memory (Redish, 2001). Both locale strategies and episodic memories are hippocampal-dependent and imply the flexible use of complex sensory stimuli. Non-locale strategies and procedural memory, in contrast, depend on the dorsal striatum, are less flexible and require time to be developed. This analogy highlights the potential importance of studying navigation strategies in rodents as a model of multiple memory systems in mammals and its interactions.

Development of habits. The acquisition of stimulus-response strategies (alike to procedural memories) can be related to the development of habits. In navigational tasks, animals gradually shift from place-response (hippocampal dependent) to stimulus-response (striatum dependent) strategies when trained extensively (Packard and McGaugh, 1996; Chang and Gold, 2003b). Such a transition has been interpreted as the basis for habit formation (Packard and McGaugh, 1996; Packard and Knowlton, 2002). As described in Section 7.5.2 extensions to the present model can be made to capture this phenomenon.

Role of other brain areas. In this work we have only referred to the putative roles the hippocampus and the basal ganglia may have in goal-directed navigation. However other brain areas are also implied in this type of behaviour. Particularly the Prefrontal cortex has been associated to the development of habits (Killcross and Coutureau, 2003), the execution of taxon strategies (Bruin et al., 2001), and learning and retrieval of goal information (Hasselmo, 2005).

The determine what is the role of such area in the type of tasks we have modelled here and how can we integrate it to the model, constitutes an interesting extension of this work.

Multiple S-R . We have modelled a competition between actions generated by a place-response strategy and a stimulus-response strategy. However, in the general situation, the animal may have learnt several S-R associations (i.e. approach food, freeze after hearing a tone, etc.). At any moment the animal may be required to choose among the multiple S-R associations and place-dependent strategies. This selection may be implemented in several ways.

On one hand, all the responses (including both S-R and place responses) can be implemented as different channels in the output structures of the basal ganglia. A single competition mechanism will be in charge of choosing the winning action. On the other hand, a hierarchical competition may be implemented as in (Girard,

2003). A dorsal module (coding for S-R) and a ventral module (coding for place-based actions) selects one action each, and these two interact to choose which action will be finally performed. These implies an interaction mechanism between place- and stimulus-response additional to the competition used within each module.

Experimental and anatomical data should be gathered to determine which structure reflects better animal behaviour and neurophysiological properties of the involved brain structures.

Other factors modulating navigation strategies. Endogenous factors like stress (Kim and Baxter, 2001), hormonal status (Marriott and Korol, 2003), or motivation (White and McDonald, 2002; Mizumori et al., 2004) also affect the selection of learning strategies. In the general case, these factors may modulate levels of the different neuromodulators, which in turn have been proposed to be biological substrates for learning metaparameters (Doya, 2002).

A mechanism of meta-modulation may change the model parameters (Table 7.1), accordingly to reported changes in animal behaviour. For instance, high oestrogen levels may increase the learning rate of the place-response model (η^{P}), promoting the use of spatial-based strategies as reported by Korol et al. (2004).

Complex stimuli. Simple sensory stimuli was used for the implementation of cue-response behaviour. Complementary to the extension of the spatial learning model to use realistic 2D vision, a visual system can be added to the present model. This system can detect pre-defined visual cues and provide a measure of the probability that a specific cue c_i is located at the egocentric direction φ. That information can be used to drive the action cells in the cue-response module.

Neurorobotics implementations. The use of realistic sensory stimuli will allow the test of the proposed model in robotic platforms. Previously, the locale navigation strategy has been tested using realistic 2D visual input and odometry information from a Kephera robot (Arleo et al., 2004; Strösslin, Sheynikhovich, et al., 2005). The use of real embodiment and agent-environment interactions constitutes a powerful testbed for the robustness of the model.

In our model, by having modules driven by different input spaces (place and cue information) we are able to solve tasks requiring different navigation strategies. Bio-inspired robots can use the same approach by maintaining different mappings of the perception-action relations. Action control in autonomous robots can be performed by choosing among reactive behaviour in egocentric coordinates (e.g. Braitenberg-like obstacle avoidance, approaching a beacon), or trajectory planning in an allocentric frame of reference. The selection can be based on a competition mechanism like the one we proposed in this model, which takes into account how well each representation has performed in the past and the predicted outcome of the proposed actions.

References

Alexander, G. E., and Crutcher, M. D. (1990, July). Functional architecture of basal ganglia circuits: neural substrates of parallel processing. *Trends Neurosci*, *13*(7), 266-71. 75

Alexander, G. E., DeLong, M. R., and Strick, P. L. (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annu Rev Neurosci*, *9*, 357-81. 75

Alyan, S. H., Paul, B. M., Ellesworth, E., White, R. D., and McNaughton, B. L. (1997). Is the hippocampus required for path integration? *Society for Neuroscience Abstracts*, *23*, 504. 71

Amaral, D. G. (1993). Emerging principles of intrinsic hippocampal organization. *Current Opinion in Neurobiology*, *3*, 225–229. 18

Amaral, D. G., Dolorfo, C., and Alvarez-Royo, P. (1991). Organization of Ca1 projections to the subiculum: a PHA-L analysis in the rat. *Hippocampus*, *1*, 415–435. 18

Amaral, D. G., and Witter, M. P. (1989). The three-dimensional organization of the hippocampal formation: A review of anatomical data. *Neuroscience*, *31(3)*, 571–591. 17, 18

Amaral, D. G., and Witter, M. P. (1995). Hippocampal formation. In G. Paxinos (Ed.), *The rat nervous system* (Second ed., pp. 443–493). Academic Press. 16, 18, 75

Arleo, A. (2000). *Spatial learning and navigation in neuro-mimetic systems, modeling the rat hippocampus*. Phd thesis, Ecole polytechnique fédérale de Lausanne (EPFL), Lausanne, Switzerland. 2

Arleo, A., and Gerstner, W. (2000). Spatial cognition and neuro-mimetic navigation: A model of hippocampal place cell activity. *Biological Cybernetics, Special Issue on Navigation in Biological and Artificial Systems*, *83*, 287–299. 29, 75, 87, 115

Arleo, A., Smeraldi, F., and Gerstner, W. (2004, May). Cognitive navigation based on nonuniform gabor space sampling, unsupervised growing networks, and reinforcement learning. *IEEE Transactions on Neural Networks*, *15*(3), 639–652. 29, 30, 54, 87, 92, 115, 120

Baldasarre, G. (2002, Mar). A modular neural-network model of the basal ganglia's

role in learning and selecting motor behaviours. *Cognitive Systems Research*, *3*(1), 5-13. 88, 97, 98

Banquet, J., Gaussier, P., Quoy, M., Revel, A., and Burnod, Y. (2005). A hierarchy of associations in hippocampo-cortical systems: Cognitive maps and navigation strategies. *Neural Computation*, *17*(6), 1339–1384. 28

Battaglia, F. P., Sutherland, G. R., and McNaughton, B. L. (2004). Local sensory cues and place cell directionality: Additional evidence of prospective coding in the hippocampus. *J. Neurosci.*, *24*(19), 4541-4550. 9, 23, 38, 69, 116

Bayer, S. A. (1982). Changes in the total number of dentate granule cells in juvenile and adult rats: a correlated volumetric and 3h-thymidine autoradiographic study. *Exp Brain Res*, *46*(3), 315-23. 18

Berns, G., and Sejnowski, T. (1996). How the basal ganglia make decision. In A. Damasio, H. Damasio, and Y. Christen (Eds.), *The neurobiology of decision making* (p. 101-113). Berlin: Springer-Verlag. 88, 89

Biegler, R., and Morris, R. G. M. (1993). Landmark stability is a prerequisite for spatial but not discrimination learning. *Nature*, *361*, 631–633. 21

Biegler, R., and Morris, R. G. M. (1996). Landmark stability: Studies exploring whether the perceived stability of the environment influences spatial representation. *Journal of Experimental Biology*, *199(1)*, 187–193. 21

Biella, G., and De Curtis, M. (2000, April). Olfactory inputs activate the medial entorhinal cortex via the hippocampus. *Journal of Neurophysiology*, *83*(4), 1924-1931. 17

Bizon, J. L., Han, J. S., Hudon, C., and Gallagher, M. (2003). Effects of hippocampal cholinergic deafferentation on learning strategy selection in a visible platform version of the water maze. *Hippocampus*, *13*(6), 676-684. 77, 80

Bliss, T. V., and Collingridge, G. L. (1993, Jan 7). A synaptic model of memory: long-term potentiation in the hippocampus. *Nature*, *361*(6407), 31-9. 31

Blum, K. I., and Abbott, L. F. (1996). A model of spatial map formation in the hippocampus of the rat. *Neural Computation*, *8*, 85–93. 84

Brandner, C., and Schenk, F. (1998, March). Septal lesions impair the acquisition of a cued place navigation task: Attentional or memory deficit? *Neurobiology of Learning and Memory*, *69*(2), 106-125. 20

Breese, C. R., Hampson, R. E., and Deadwyler, S. A. (1989, April). Hippocampal place cells: stereotypy and plasticity. *J Neurosci*, *9*(4), 1097-111. 21

Bruin, J. P. de, Moita, M. P., Brabander, H. M. de, and Joosten, R. N. (2001, March). Place and response learning of rats in a morris water maze: differential effects of fimbria fornix and medial prefrontal cortex lesions. *Neurobiol Learn Mem*, *75*(2), 164-78. 119

Brun, V. H., Otnass, M. K., Molden, S., Steffenach, H. A., Witter, M. P., Moser, M. B., et al. (2002, Jun 21). Place cells and place recognition maintained by direct entorhinal-hippocampal circuitry. *Science*, *296*(5576), 2243-6. 19, 31,

69, 71, 72, 118

Brunel, N., and Trullier, O. (1998). Plasticity of directional place fields in a model of rodent ca3. *Hippocampus*, *8*(6), 651-65. 31, 69, 116

Burgess, N., Jeffery, K. J., and O'Keefe, J. (1999). Integrating hippocampal and parietal functions: a spatial point of view. In K. J. J. N. Burgess and J. O'Keefe (Eds.), *The hippocampal and parietal foundations of spatial cognition* (pp. 3–29). Oxford University Press. 18

Burgess, N., and O'Keefe, J. (1996). Neuronal computations underlying the firing of place cells and their role in navigation. *Hippocampus.*, *6*(6), 749-62. 26

Burgess, N., Recce, M., and O'Keefe, J. (1994). A model of hippocampal function. *Neural Networks*, *7*, 1065–1081. 23, 25, 26, 75

Buzsáki, G. (1984). Feed-forward inhibition in the hippocampal formation. *Progress in Neurobiology*, *22*, 131–153. 16, 20

Cahill, J. F., and Baxter, M. G. (2001, December). Cholinergic and noncholinergic septal neurons modulate strategy selection in spatial learning. *Eur J Neurosci*, *14*(11), 1856-1864. 80

Chang, Q., and Gold, P. E. (2003a, Sep 15). Intra-hippocampal lidocaine injections impair acquisition of a place task and facilitate acquisition of a response task in rats. *Behav Brain Res*, *144*(1-2), 19-24. 77, 79, 80, 91, 109

Chang, Q., and Gold, P. E. (2003b, Apr 1). Switching memory systems during learning: changes in patterns of brain acetylcholine release in the hippocampus and striatum in rats. *J Neurosci*, *23*(7), 3001-3005. 77, 80, 111, 112, 113, 119

Chavarriaga, R., and Gerstner, W. (2004). Combining visual and proprioceptive information in a model of spatial learning and navigation. In *Proceedings 2004 international joint conference on neural networks (ijcnn'2004)* (pp. 603–608). IEEE Press. 31, 56

Chavarriaga, R., Strösslin, T., Sheynikhovich, D., and Gerstner, W. (2005a). Competition between cue response and place response : A model of rat navigation behaviour. *Connection Science*, *17*(1–2), 167–183. 91

Chavarriaga, R., Strösslin, T., Sheynikhovich, D., and Gerstner, W. (2005b). A computational model of parallel navigation systems in rodents. *Neuroinformatics*, *3*(3), 223–242. 91, 98

Chrobak, J. J., and Buzsáki, G. (1996). High-frequency oscillations in the output networks of the hippocampal-entorhinal axis of the freely moving rat. *Journal of Neuroscience*, *16*(9), 3056–3066. 20

Chrobak, J. J., Lörincz, A., and Buzsáki, G. (2000). Physiological patterns in the hippocampo-entorhinal cortex system. *Hippocampus*, *10*, 457–465. 20

Ciaroni, S., Cuppini, R., Cecchini, T., Ferri, P., Ambrogini, P., Cuppini, C., et al. (1999, Aug 30). Neurogenesis in the adult rat dentate gyrus is enhanced by vitamin e deficiency. *J Comp Neurol*, *411*(3), 495-502. 18

Claiborne, B. J., Amaral, D. G., and Cowan, W. M. (1986, Apr 22). A light and

electron microscopic analysis of the mossy fibers of the rat dentate gyrus. *J Comp Neurol*, *246*(4), 435-58. 18

Colombo, P. J., and Gallagher, M. (1998, November). Individual differences in spatial memory and striatal chat activity among young and aged rats. *Neurobiol Learn Mem*, *70*(3), 314-327. 77, 80

Da Cunha, C., Wietzikoski, S., Wietzikoski, E. C., Miyoshi, E., Ferro, M. M., Anselmo-Franci, J. A., et al. (2003, May). Evidence for the substantia nigra pars compacta as an essential component of a memory system independent of the hippocampal memory system. *Neurobiol Learn Mem*, *79*(3), 236-242. 77

Daniel, J. M., and Lee, C. D. (2004, September). Estrogen replacement in ovariectomized rats affects strategy selection in the morris water maze. *Neurobiol Learn Mem*, *82*(2), 142-149. 80

Daw, N., Niv, Y., and Dayan, P. (2005). Recents breakthroughs in basal ganglia research. In Bezard (Ed.), (chap. Actions, policies, values, and the basal ganglia). Nova science publishers. 113

Dayan, P., Kakade, S., and Montague, P. R. (2000, November). Learning and selective attention. *Nat Neurosci*, *3 Suppl*, 1218-23. 112

Devan, B. D., and White, N. M. (1999, Apr 1). Parallel information processing in the dorsal striatum: relation to hippocampal function. *J Neurosci*, *19*(7), 2789-98. 2, 77, 79, 80, 100, 102, 103, 106, 107, 109, 110, 111, 117, 118

Doya, K. (1999, October). What are the computations of the cerebellum, the basal ganglia and the cerebral cortex? *Neural Networks*, *12*(7-8), 961-974. 83, 88

Doya, K. (2000, December). Complementary roles of basal ganglia and cerebellum in learning and motor control. *Curr Opin Neurobiol*, *10*(6), 732-9. 88

Doya, K. (2002). Metalearning and neuromodulation. *Neural Networks*, *15*(4-6), 495–506. 6, 81, 82, 88, 111, 112, 120

Doya, K., and Uchibe, E. (2005). The Cyber Rodent Project: Exploration of Adaptive Mechanisms for Self-Preservation and Self-Reproduction. *Adaptive Behavior*, *13*(2), 149-160. 89

Eichenbaum, H., Stewart, C., and Morris, R. G. (1990, November). Hippocampal representation in place learning. *J Neurosci*, *10*(11), 3531-3542. 77, 85, 110

Etienne, A. S. (1998, December). Mammalian navigation, neural models and biorobotics. *Connection Science*, *10*(3-4), 271-289. 73

Etienne, A. S., and Jeffery, K. J. (2004). Path integration in mammals. *Hippocampus*, *14*(2), 180-92. 22

Etienne, A. S., Maurer, R., and Seguinot, V. (1996, January). Path integration in mammals and its interaction with visual landmarks. *J Exp Biol*, *199*(Pt 1), 201-9. 21, 22

Filliat, D. (2001). *Cartographie et estimation globale de la position pour un robot mobile autonome*. Phd thesis, LIP6/AnimatLab, Université Pierre et Marie Curie, Paris, France. (Spécialité Informatique) 89

Floresco, S. B., Blaha, C. D., Yang, C. R., and Phillips, A. G. (2001, Apr 15). Modulation of hippocampal and amygdalar-evoked activity of nucleus accumbens neurons by dopamine: cellular mechanisms of input selection. *J Neurosci*, *21*(8), 2851-60. 111

Foster, D. J., Morris, R. G. M., and Dayan, P. (2000). A model of hippocampally dependent navigation, using the temporal difference learning rule. *Hippocampus*, *10(1)*, 1–16. 75, 85, 86, 87

Freund, T. F., Powell, J. F., and Smith, A. D. (1984, December). Tyrosine hydroxylase-immunoreactive boutons in synaptic contact with identified striatonigral neurons, with particular reference to dendritic spines. *Neuroscience*, *13*(4), 1189-215. 111

Fyhn, M., Molden, S., Witter, M. P., Moser, E. I., and Moser, M. B. (2004, Aug 27). Spatial representation in the entorhinal cortex. *Science*, *305*(5688), 1258-64. 8, 19, 69, 70, 71, 72, 116, 118

Gaussier, P., Revel, A., Banquet, J. P., and Babeau, V. (2002, January). From view cells and place cells to cognitive map learning: processing stages of the hippocampal system. *Biol Cybern*, *86*(1), 15-28. 28, 29

Georgopoulos, A. P., Schwartz, A. B., and Kettner, R. E. (1986). Neuronal population coding of movement direction. *Science*, *233*, 1416–1419. 12, 13

Gerstner, W., and Abbott, L. F. (1997, January). Learning navigational maps through potentiation and modulation of hippocampal place cells. *J Comput Neurosci*, *4*(1), 79-94. 84

Girard, B. (2003). *Intégration de la navigation et de la sélection de l'action dans une architecture de contrôle inspirée des ganglions de la base.* Phd thesis, LIP6/AnimatLab, Université Pierre et Marie Curie, Paris, France. (Spécialité Informatique) 89, 90, 110, 113, 120

Girard, B., Filliat, D., Meyer, J.-A., Berthoz, A., and A. Guillot, A. (2005). Integration of navigation and action selection functionalities in a computational model of cortico-basal ganglia-thalamo-cortical loops. *Adaptive Behavior*, *13*(2), 115–130. 89

Gold, P. E. (2003, November). Acetylcholine modulation of neural systems involved in learning and memory. *Neurobiol Learn Mem*, *80*(3), 194-210. 80, 111

Gold, P. E. (2004, November). Coordination of multiple memory systems. *Neurobiol Learn Mem*, *82*(3), 230-42. 2, 79, 80, 110, 111, 112

Gothard, K. M., Skaggs, W. E., and McNaughton, B. L. (1996). Dynamics of mismatch correction in the hippocampal ensemble code for space: Interaction between path integration and environmental cues. *Journal of Neuroscience*, *16(24)*, 8027–8040. 8, 22, 23, 55, 62, 64, 65, 66, 70

Gothard, K. M., Skaggs, W. E., Moore, K. M., and McNaughton, B. L. (1996). Binding of hippocampal CA1 neural activity to multiple reference frames in a landmark-based navigation task. *Journal of Neuroscience*, *16(2)*, 823–835. 21

Graybiel, A. M. (1998, Jul-Sep). The basal ganglia and chunking of action repertoires. *Neurobiol Learn Mem*, *70*(1-2), 119-36. 75

Gu, Q. (2002). Neuromodulatory transmitter systems in the cortex and their role in cortical plasticity. *Neuroscience*, *111*(4), 815-35. 83

Guillot, A., and Meyer, J.-A. (2001). The animat contribution to cognitive systems. *Journal of Cognitive Systems Research*, *2*(2), 157–165. 7

Gurney, K., Prescott, T. J., and Redgrave, P. (2001a, June). A computational model of action selection in the basal ganglia. i. a new functional anatomy. *Biol Cybern*, *84*(6), 401-10. 89, 111

Gurney, K., Prescott, T. J., and Redgrave, P. (2001b, June). A computational model of action selection in the basal ganglia. ii. analysis and simulation of behaviour. *Biol Cybern*, *84*(6), 411-23. 89

Hafting, T., Fyhn, M., Molden, S., Moser, M. B., and Moser, E. I. (2005, Jun 19). Microstructure of a spatial map in the entorhinal cortex. *Nature*. 1, 19, 71, 72

Hargreaves, E. L., Rao, G., Lee, I., and Knierim, J. J. (2005, Jun 17). Major dissociation between medial and lateral entorhinal input to dorsal hippocampus. *Science*, *308*(5729), 1792-4. 71

Hartley, T., Burgess, N., Lever, C., Cacucci, F., and Keefe, J. O. (2000). Modeling place fields in terms of the cortical inputs to the hippocampus. *Hippocampus*, *10*, 369-379. 33

Hasselmo, M. E. (1999, September). Neuromodulation: acetylcholine and memory consolidation. *Trends Cogn Sci*, *3*(9), 351-359. 81, 82, 111, 112

Hasselmo, M. E. (2005, July). A model of prefrontal cortical mechanisms for goal-directed behavior. *J Cogn Neurosci*, *17*(7), 1115-29. 119

Hasselmo, M. E., Bodelon, C., and Wyble, B. P. (2002, April). A proposed function for hippocampal theta rhythm: separate phases of encoding and retrieval enhance reversal of prior learning. *Neural Comput*, *14*(4), 793-817. 20, 81, 82

Hasselmo, M. E., and Bower, J. M. (1993). Acetylcholine and memory. *Trends in Neurosciences*, *16(6)*, 218–222. 20, 112

Hasselmo, M. E., Hay, J., Ilyn, M., and Gorchetchnikov, A. (2002, Jun-Jul). Neuromodulation, theta rhythm and rat spatial navigation. *Neural Networks*, *15*(4-6), 689-707. 20, 81, 82

Hastings, N. B., Seth, M. I., Tanapat, P., Rydel, T. A., and Gould, E. (2002, Oct 28). Granule neurons generated during development extend divergent axon collaterals to hippocampal area ca3. *J Comp Neurol*, *452*(4), 324-33. 18

Haykin, S. (1994). *Neural networks: A comprehensive foundation*. Englewood Cliffs NJ: MacMillan. 5

Heimer, L., Zahn, D. S., and Alheid, G. F. (1995). Basal ganglia. In G. Paxinos (Ed.), *The rat nervous system* (Second ed., pp. 579–628). Academic Press. 75

Hill, A. J., and Best, P. J. (1981). Effects of deafness and blindness on the spatial

correlates of hippocampal unit activity in the rat. *Experimental neurology*, *74*, 204–217. 22

Holland, P. C., and Gallagher, M. (1999, February). Amygdala circuitry in attentional and representational processes. *Trends Cogn Sci*, *3*(2), 65-73. 83

Houk, J. C., Adams, J. L., and Barto, A. G. (1995). A model of how the basal ganglia generate and use neural signals that predict reinforcement. In J. C. Houk, J. L. Davis, and D. G. Beiser (Eds.), *Models of information processing in the basal ganglia* (pp. 249–270). Cambridge, Massachusetts, USA: MIT Press. 83, 88, 111

Insausti, R., Herrero, M. T., and Witter, M. P. (1997). Entorhinal cortex of the rat: Cytoarchitectonic subdivision and the origin and distribution of cortical efferents. *Hippocampus*, *7*, 146–83. 17

Insausti, R., Marcos, P., Arroyo-Jimenez, M. M., Blaizot, X., and Martinez-Marcos, A. (2002). Comparative aspects of the olfactory portion of the entorhinal cortex and its projection to the hippocampus in rodents, nonhuman primates, and the human brain. *Brain Research Bulletin*, *57*(3–4), 557–560. 17

Jackson, J., Johnson, A., and Redish, A. D. (2005). Hippocampal sharp wave events increase during behavior with experience within session. In *14th annual computational neuroscience meeting, cns 2005.* 21

Janis, L. S., Glasier, M. M., Fulop, Z., and Stein, D. G. (1998, January). Intraseptal injections of 192 igg saporin produce deficits for strategy selection in spatial-memory tasks. *Behav Brain Res*, *90*(1), 23-34. 77, 80

Joel, D., Niv, Y., and Ruppin, E. (2002, Jun-Jul). Actor-critic models of the basal ganglia: new anatomical and computational perspectives. *Neural Netw*, *15*(4-6), 535-47. 88

Kali, S., and Dayan, P. (2000, Oct 1). The involvement of recurrent connections in area ca3 in establishing the properties of place fields: a model. *J Neurosci*, *20*(19), 7463-77. 69

Khamassi, M., Lachèze, L., Girard, B., Berthoz, A., and Guillot, A. (2005). Actor-critic models of reinforcement learning in the basal ganglia: From natural to artificial rats. *Adaptive Behavior, Special Issue Towards Artificial Rodents*, *13*(2), 131–148. 88, 89

Killcross, S., and Coutureau, E. (2003, April). Coordination of actions and habits in the medial prefrontal cortex of rats. *Cereb Cortex*, *13*(4), 400-8. 119

Kim, J. J., and Baxter, M. G. (2001, June). Multiple brain-memory systems: the whole does not equal the sum of its parts. *Trends Neurosci*, *24*(6), 324-330. 120

Korol, D. L. (2004, November). Role of estrogen in balancing contributions from multiple memory systems. *Neurobiol Learn Mem*, *82*(3), 309-323. 80

Korol, D. L., Malin, E. L., Borden, K. A., Busby, R. A., and Couper-Leo, J. (2004, May). Shifts in preferred learning strategy across the estrous cycle in female

rats. *Horm Behav*, *45*(5), 330-338. 80, 120

Kosel, K. C., Hoesen, G. W. V., and West, J. R. (1981). Olfactory bulb projections to the parahippocampal area of the rat. *Journal of Comparative Neurology*, *198*, 467–482. 17

Kubie, J. L., and Ranck, J. B. (1983). Sensory-behavioral correlates in individual hippocampus neurons in three situations: Space and context. In W. Seifert (Ed.), *Neurobiology of the hippocampus* (p. 433-447). New York: Academic Press. 19, 21

Kudrimoti, H. S., Barnes, C. A., and McNaughton, B. L. (1999, May 15). Reactivation of hippocampal cell assemblies: effects of behavioral state, experience, and eeg dynamics. *J Neurosci*, *19*(10), 4090-101. 20

Kuhn, H. G., Dickinson-Anson, H., and Gage, F. H. (1996, Mar 15). Neurogenesis in the dentate gyrus of the adult rat: age-related decrease of neuronal progenitor proliferation. *J Neurosci*, *16*(6), 2027-33. 18

Lavenex, P., and Schenk, F. (1996, jan-feb). Integration of olfactory information in a spatial representation enabling accurate arm choice in the radial arm maze. *Learning & Memory*, *2*(6), 299-319. 21

Legault, M., Rompré, P.-P., and Wise, R. A. (2000). Chemical stimulation of the ventral hippocampus elevates nucleus accumbens dopamine by activating dopaminergic neurons of the ventral tegmental area. *Journal of Neuroscience*, *20*(4), 1635–1642. 111

Liu, P., and Bilkey, D. K. (1997). Parallel involvement of perirhinal and lateral entorhinal cortex in the polysynaptic activation of hippocampus by olfactory inputs. *Hippocampus*, *7*(3), 296-306. 17

Maaswinkel, H., and Whishaw, I. Q. (1999, March). Homing with locale, taxon, and dead reckoning strategies by foraging rats: sensory hierarchy in spatial navigation. *Behav Brain Res*, *99*(2), 143-52. 21

Markus, E. J., Barnes, C. A., McNaughton, B. L., Gladden, V. L., and Skaggs, W. E. (1994). Spatial information content and reliability of hippocampal ca1 neurons: Effects of visual input. *Hippocampus*, *4*, 410–421. 21, 22

Markus, E. J., Qin, Y., Leonard, B., Skaggs, W. E., McNaughton, B. L., and Barnes, C. A. (1995). Interactions between location and task affect the spatial and direction firing of hippocampal neurons. *Journal of Neuroscience*, *15*, 7079–7094. 8, 12, 21, 23, 40, 44, 46, 48, 61

Marriott, L. K., and Korol, D. L. (2003, November). Short-term estrogen treatment in ovariectomized rats augments hippocampal acetylcholine release during place learning. *Neurobiol Learn Mem*, *80*(3), 315-322. 77, 80, 120

Matthews, D. B., Ilgen, M., White, A. M., and Best, P. J. (1999, November). Acute ethanol administration impairs spatial performance while facilitating nonspatial performance in rats. *Neurobiol Learn Mem*, *72*(3), 169-179. 80

McClelland, J. L., McNaughton, B. L., and O'Reilly, R. C. (1995, July). Why

there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychol Rev*, *102*(3), 419-57. 20

McDonald, R. J., and White, N. M. (1994). Parallel information processing in the water maze: Evidence for independent memory systems involving dorsal striatum and hippocampus. *Behavioral and Neural Biology*, *61*, 260–270. 110

McIntyre, C. K., Marriott, L. K., and Gold, P. E. (2003, March). Patterns of brain acetylcholine release predict individual differences in preferred learning strategies in rats. *Neurobiol Learn Mem*, *79*(2), 177-183. 77, 80, 111, 112

McNaughton, B. L., Barnes, C. A., Gerrard, J. L., Gothard, K., Jung, M. W., Knierim, J. J., et al. (1996, January). Deciphering the hippocampal polyglot: the hippocampus as a path integration system. *J Exp Biol*, *199 (Pt 1)*, 173-85. 19

McNaughton, B. L., Barnes, C. A., Meltzer, J., and Sutherland, R. J. (1989). Hippocampal granule cells are necessary for normal spatial learning but not for spatially-selective pyramidal cell discharge. *Exp Brain Res*, *76*(3), 485-96. 19

McNaughton, B. L., Barnes, C. A., and O'Keefe, J. (1983). The contributions of position, direction, and velocity to single unit activity in the hippocampus of freely-moving rats. *Experimental Brain Research*, *52*, 41–49. 1, 21

McNaughton, B. L., Leonard, B., and Chen, L. (1989). Cortical-hippocampal interactions and cognitive mapping: A hypothesis based on reintegration of the parietal and inferotemporal pathways for visual processing. *Psychobiology*, *17*(3), 230–235. 22

Mehta, M. R., Lee, A. K., and Wilson, M. A. (2002, Jun 13). Role of experience and oscillations in transforming a rate code into a temporal code. *Nature*, *417*(6890), 741-6. 20

Mehta, M. R., Quirk, M. C., and Wilson, M. A. (2000). Experience-dependent, asymmetric shape of hippocampal receptive fields. *Neuron*, *25*, 707–715. 19, 23

Miller, R. (1991). *Cortico-Hippocampal interplay and the representation of contexts in the brain*. Springer-Verlag. 20

Mink, J. W. (1996, November). The basal ganglia: focused selection and inhibition of competing motor programs. *Prog Neurobiol*, *50*(4), 381-425. 74, 75, 89

Mittelstaedt, H. (1983). The role of multimodal convergence in homing by path integration. *Fortschritte der Zoologie*, *28*, 197–212. 22

Mittelstaedt, M. L., and Mittelstaedt, H. (1980). Homing by path integration in a mammal. *Naturwissenschaften*, *67*, 566–567. 22

Mizumori, S. J., McNaughton, B. L., Barnes, C. A., and Fox, K. B. (1989, November). Preserved spatial coding in hippocampal ca1 pyramidal cells during reversible suppression of ca3c output: evidence for pattern completion in hippocampus. *J Neurosci*, *9*(11), 3915-28. 19, 20

Mizumori, S. J., Yeshenko, O., Gill, K. M., and Davis, D. M. (2004, November). Parallel processing across neural systems: Implications for a multiple memory system hypothesis. *Neurobiol Learn Mem*, *82*(3), 278-298. 110, 120

Montague, P. R., Dayan, P., and Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive hebbian learning. *Journal of Neuroscience*, *16*(5), 1936–1947. 81, 83, 88

Montes-Gonzalez, F., Prescott, T., Gurney, K., Humphries, M., and Redgrave, P. (2000). An embodied model of action selection mechanisms in the vertebrate brain. In J.-A. Meyer, A. Berthoz, D. Floreano, H. Roitblat, and S. W. Wilson (Eds.), *From animals to animats 6: Proceedings of the sixth international conference on simulation of adaptive behaviour.* Cambridge MA: MIT Press. 88, 89

Morris, R. G. M. (1981). Spatial localization does not require the presence of local cues. *Learning and Motivation*, *12*, 239–260. 1, 2, 95, 110, 117

Muller, R. U., Bostock, E., Taube, J. S., and Kubie, J. L. (1994). On the directional firing properties of hippocampal place cells. *Journal of Neuroscience*, *14(12)*, 7235–7251. 8, 23, 41, 44

Muller, R. U., and Kubie, J. L. (1987). The effects of changes in the environment on the spatial firing of hippocampal complex-spike cells. *Journal of Neuroscience*, *7*, 1951–1968. 19, 21, 22

Muller, R. U., Kubie, J. L., and Ranck, J. J. B. (1987). Spatial firing patterns of hippocampal complex-spike cells in a fixed environment. *Journal of Neuroscience*, *7*, 1935–1950. 8, 21, 37

O'Keefe, J., and Burgess, N. (1996). Geometric determinants of the place fields of hippocampal neurons. *Nature*, *381*, 425–428. 8, 9, 22, 23, 27, 47, 69, 70

O'Keefe, J., and Conway, D. H. (1978). Hippocampal place units in the freely moving rat: why they fire where they fire. *Experimental Brain Research*, *31*, 573–590. 19, 21, 22

O'Keefe, J., and Dostrovsky, J. (1971). The hippocampus as a spatial map. preliminary evidence from unit activity in the freely-moving rat. *Brain Research*, *34*, 171–175. 1, 15, 18

O'Keefe, J., and Nadel, L. (1978). *The hippocampus as a cognitive map.* Oxford: Clarendon Press. 1

O'Keefe, J., and Recce, M. (1993). Phase relationship between hippocampal place units and the EEG theta rhythm. *Hippocampus*, *3*, 317–330. 20, 23

O'Keefe, J., and Speakman, A. (1987). Single unit activity in the rat hippocampus during a spatial memory task. *Experimental Brain Research*, *68*, 1–27. 22

Oliveira, M. G., Bueno, O. F., Pomarico, A. C., and Gugliano, E. B. (1997, July). Strategies used by hippocampal- and caudate-putamen-lesioned rats in a learning task. *Neurobiol Learn Mem*, *68*(1), 32-41. 77, 91

Packard, M. G. (1999, Oct 26). Glutamate infused posttraining into the hippocampus

or caudate-putamen differentially strengthens place and response learning. *Proc Natl Acad Sci U S A*, *96*(22), 12881-12886. 80

Packard, M. G., Hirsh, R., and White, N. M. (1989, May). Differential effects of fornix and caudate nucleus lesions on two radial maze tasks: evidence for multiple memory systems. *J Neurosci*, *9*(5), 1465-72. 2, 77, 79, 106, 109

Packard, M. G., and Knowlton, B. J. (2002). Learning and memory functions of the basal ganglia. *Annu Rev Neurosci*, *25*, 563-93. 2, 77, 117, 119

Packard, M. G., and McGaugh, J. L. (1992). Double dissociation of fornix and caudate nucleus lesions on acquisition of two water maze tasks: Further evidence for multiple memory systems. *Behavioral Neuroscience*, *106(3)*, 439–446. 85, 110

Packard, M. G., and McGaugh, J. L. (1996, January). Inactivation of hippocampus or caudate nucleus with lidocaine differentially affects expression of place and response learning. *Neurobiol Learn Mem*, *65*(1), 65-72. 77, 78, 91, 106, 112, 113, 119

Pearce, J. M., Roberts, A. D., and Good, M. (1998, Nov 5). Hippocampal lesions disrupt navigation based on cognitive maps but not heading vectors. *Nature*, *396*(6706), 75-7. 2, 77, 107, 108, 118

Pico, R. M., Gerbrandt, L. K., Pondel, M., and Ivy, G. (1985). During stepwise cue deletion, rat place behaviors correlate with place unit responses. *Brain Research*, *330*, 369–373. 22

Pouzet, B., Zhang, W. N., Feldon, J., and Rawlins, J. N. (2002, Jul 18). Hippocampal lesioned rats are able to learn a spatial position using non-spatial strategies. *Behav Brain Res*, *133*(2), 279-291. 77

Praxinos, G., and Watson, C. (1998). *The rat brain in stereotaxic coordinates* (fourth edition ed.). Academic Press. 16

Quirk, G. J., Muller, R. U., and Kubie, J. L. (1990). The firing of hippocampal place cells in the dark depends on the rat's recent experience. *Journal of Neuroscience*, *10*(6), 2008–2017. 22, 25

Quirk, G. J., Muller, R. U., Kubie, J. L., and Ranck Jr., J. B. (1992). The positional firing properties of medial entorhinal neurons: Description and comparison with hippocampal place cells. *Journal of Neuroscience*, *12*(5), 1945–1963. 1, 19, 22

Ragozzino, K. E., Leutgeb, S., and Mizumori, S. J. (2001, August). Dorsal striatal head direction and hippocampal place representations during spatial navigation. *Exp Brain Res*, *139*(3), 372-6. 2

Ragozzino, M. E., and Gold, P. E. (1995, October). Glucose injections into the medial septum reverse the effects of intraseptal morphine infusions on hippocampal acetylcholine output and memory. *Neuroscience*, *68*(4), 981-988. 77, 80, 111

Ragozzino, M. E., Pal, S. N., Unick, K., Stefani, M. R., and Gold, P. E. (1998, Feb 15). Modulation of hippocampal acetylcholine release and spontaneous alternation scores by intrahippocampal glucose injections. *J Neurosci*, *18*(4), 1595-601. 111

Redish, A. D. (1999). *Beyond the Cognitive Map, From Place Cells to Episodic Memory*. London: MIT Press-Bradford Books. 2, 73, 77, 95, 110, 117

Redish, A. D. (2001, Dec 14). The hippocampal debate: are we asking the right questions? *Behav Brain Res*, *127*(1-2), 81-98. 77, 119

Redish, A. D., Rosenzweig, E. S., Bohanick, J. D., McNaughton, B. L., and Barnes, C. A. (2000, Dec 15). Dynamics of hippocampal ensemble activity realignment: time versus space. *J Neurosci*, *20*(24), 9298-309. 22, 55, 62, 66

Reynolds, J. N., Hyland, B. I., and Wickens, J. R. (2001, Sep 6). A cellular mechanism of reward-related learning. *Nature*, *413*(6851), 67-70. 81

Rogers, J. L., and Kesner, R. P. (2003, nov). Cholinergic modulation of the hippocampus during encoding and retrieval. *Neurobiol Learn Mem*, *80*(3), 332-42. 20, 82

Rolls, E. T. (1999). Spatial view cells and the representation of place in the primate hippocampus. *Hippocampus*, *9*(4), 467-80. 69

Salinas, E., and Abbott, L. F. (1994). Vector reconstruction from firing rates. *Journal of Computational Science*, *1*, 89–107. 13

Sara, S. J., Vankov, A., and Herve, A. (1994). Locus coeruleus-evoked responses in behaving rats: a clue to the role of noradrenaline in memory. *Brain Res Bull*, *35*(5-6), 457-65. 83

Save, E., Cressant, A., Thinus-Blanc, C., and Poucet, B. (1998). Spatial firing of hippocampal place cells in blind rats. *Journal of Neuroscience*, *18*(5), 1818–1826. 22, 25

Save, E., Nerad, L., and Poucet, B. (2000). Contribution of multiple sensory information to place field stability in hippocampal place cells. *Hippocampus*, *10*, 64–76. 21, 22, 55

Schroeder, J. P., Wingard, J. C., and Packard, M. G. (2002). Post-training reversible inactivation of hippocampus reveals interference between memory systems. *Hippocampus*, *12*(2), 280-284. 77, 79

Schultz, W. (1998). Predictive Reward Signal of Dopamine Neurons. *Journal of Neurophysiology*, *80*, 1–27. 81, 111

Schultz, W. (2002, Oct 10). Getting formal with dopamine and reward. *Neuron*, *36*(2), 241-63. 75, 81, 111

Schultz, W., Apicella, P., Scarnati, E., and Ljungberg, T. (1992, December). Neuronal activity in monkey ventral striatum related to the expectation of reward. *J Neurosci*, *12*(12), 4595-610. 111

Schultz, W., Dayan, P., and Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*, 1593–1599. 75, 81, 111

Schwerdtfeger, W. K., Buhl, E. H., and Germroth, P. (1990, feb 8). Disynaptic olfactory input to the hippocampus mediated by stellate cells in the entorhinal cortex. *Journal of Comparative Neurology*, *292*(2), 163-177. 17

Sesack, S. R., and Pickel, V. M. (1992, Jun 8). Prefrontal cortical efferents in the

rat synapse on unlabeled neuronal targets of catecholamine terminals in the nucleus accumbens septi and on dopamine neurons in the ventral tegmental area. *J Comp Neurol*, *320*(2), 145-60. 111

Sharp, P. E. (1991). Computer simulation of hippocampal place cells. *Psychobiology*, *19(2)*, 103–115. 24, 25, 26

Sharp, P. E. (1997). Subicular cells generate similar spatial firing patterns in two geometrically and visually distinctive environments: Comparison with hippocampal place cells. *Behavioral and Brain Research*, *85*, 71–92. 19

Sharp, P. E. (1999). Subicular place cells expand/contract their spatial firing pattern to fit the size of the environment in an open field, but not inthe presence of barriers: Comparison with hippocampal place cells. *Behavioral Neuroscience*, *113*(4), 643–62. 19

Sharp, P. E., and Green, C. (1994). Spatial correlates of firing patterns of single cells in the subiculum of freely moving rat. *Journal of Neuroscience*, *14(4)*, 2339–2356. 19

Sharp, P. E., Kubie, J. L., and Muller, R. U. (1990, September). Firing properties of hippocampal neurons in a visually symmetrical environment: contributions of multiple sensory cues and mnemonic processes. *J Neurosci*, *10*(9), 3093-105. 22

Sheynikhovich, D., Chavarriaga, R., Strösslin, T., and Gerstner, W. (2005). Biomimetic neural learning for intelligent robots. In S. Wermter, G. Palm, and M. Elshaw (Eds.), (pp. 245–264). Berlin Heidelberg: Springer-Verlag. 118, 119

Silva, F. H. L. da, Witter, M. P., Boeijinga, P. H., and Lohman, A. H. M. (1990). Anatomical organization and physiology of the limbic cortex. *Physiological Reviews*, *70*, 453–511. 17

Skaggs, W. E., and McNaughton, B. L. (1996). Replay of neuronal firing sequences in rat hippocampus during sleep following spatial experience. *Science*, *271*, 1870–1873. 20

Skaggs, W. E., McNaughton, B. L., Gothard, K., and Markus, E. (1992). An information-theoretic approach to deciphering the hippocampal code. In S. Hanson, J. Cowan, and L. Giles (Eds.), *Advances in neural information processing systems 5* (pp. 1030–1036). Morgan-Kaufmann. 10, 11

Skaggs, W. E., McNaughton, B. L., Wilson, M. A., and Barnes, C. A. (1996). Theta phase precession in hippocampal neuronal populations and the compression of temporal sequences. *Hippocampus*, *6*(2), 149-72. 20

Smith, Y., Bennett, B. D., Bolam, J. P., Parent, A., and Sadikot, A. F. (1994, Jun 1). Synaptic relationships between dopaminergic afferents and cortical or thalamic input in the sensorimotor territory of the striatum in monkey. *J Comp Neurol*, *344*(1), 1-19. 111

Steele, R. J., and Morris, R. G. M. (1999). Delay-dependent impairment of a

matching-to-place task with chronic and intrahippocampal infusion of the nmda-antagonist d-ap5. *Hippocampus*, *9*(2), 118-136. 85

Strösslin, T. (2004). *A connectionist model of spatial learning in the rat*. Phd thesis, Ecole polytechnique fédérale de Lausanne (EPFL), Lausanne, Switzerland. 2, 26, 29, 30, 31, 56, 87

Strösslin, T., Chavarriaga, R., Sheynikhovich, D., and Gerstner, W. (2005). Modelling path integrator recalibration using hippocampal place cells. In W. Duch, J. Kacprzyk, and E. Oja (Eds.), *Artificial neural networks: Biological inspirations - icann 2005, part i* (pp. 51–56). Berlin Heidelberg: Springer-Verlag. 118

Strösslin, T., Sheynikhovich, D., Chavarriaga, R., and Gerstner, W. (2005). Modelling robust self-localisation and navigation using hippocampal place cells. *Neural Networks*, *18*, 1125–1140. 54, 56, 92, 93, 95, 120

Sutherland, R. J., and Rodriguez, A. J. (1990). The role of the fornix/fimbria and some related subcortical structures in place learning and memory. *Behavioral and Brain Research*, *32*, 265–277. 85, 110

Sutton, R., and Barto, A. G. (1998). *Reinforcement learning - an introduction*. Cambridge, MA: MIT Press. 6, 75, 82, 94, 97, 109

Suzuki, W. A., and Amaral, D. G. (1994, March). Topographic organization of the reciprocal connections between the monkey entorhinal cortex and the perirhinal and parahippocampal cortices. *J Neurosci*, *14*(3 Pt 2), 1856-77. 17

Taube, J. S., Muller, R. U., and Ranck, J. B. (1990, February). Head-direction cells recorded from the postsubiculum in freely moving rats. i. description and quantitative analysis. *J Neurosci*, *10*(2), 420-35. 41

Thompson, L. T., and Best, P. J. (1989). Place cells and silent cells in the hippocampus of freely-behaving rats. *Journal of Neuroscience*, *9(7)*, 2382–2390. 19, 21

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review*, *55*, 189–208. 1, 15, 74

Totterdell, S., and Smith, A. D. (1989, Sep-Oct). Convergence of hippocampal and dopaminergic input onto identified neurons in the nucleus accumbens of the rat. *J Chem Neuroanat*, *2*(5), 285-98. 111

Tsodyks, M. V., Skaggs, W. E., Sejnowski, T. J., and McNaughton, B. L. (1996). Population dynamics and theta rhythm phase precession of hippocampal place cell firing: a spiking neuron model. *Hippocampus*, *6*(3), 271-80. 23

Wan, H. S., Touretzky, D. S., and Redish, A. D. (1994). Towards a computational theory of rat navigation. In M. Mozer, P. Smolensky, D. S. Touretzky, J.Elman, and A. Weigend (Eds.), *Proceedings of the 1993 connectionist models summer school* (pp. 11–19). Hillsdale, NJ, Erlbaum. 27, 28

Whishaw, I. Q., Cassel, J. C., and Jarrad, L. E. (1995, August). Rats with fimbria-fornix lesions display a place response in a swimming pool: a dissociation be-

tween getting there and knowing where. *J Neurosci*, *15*(8), 5779-5788. 77, 110

Whishaw, I. Q., and Mittleman, G. (1991). Hippocampal modulation of nucleus accumbens: Behavioral evidence from amphetamine-induced activity profiles. *Behavioral and Neural Biology*, *55*, 289–306. 111

Whishaw, I. Q., Mittleman, G., Bunch, S. T., and Dunnett, S. B. (1987, May). Impairments in the acquisition, retention and selection of spatial navigation strategies after medial caudate-putamen lesions in rats. *Behav Brain Res*, *24*(2), 125-138. 77, 111

White, N. M. (2004, November). The role of stimulus ambiguity and movement in spatial navigation: A multiple memory systems analysis of location discrimination. *Neurobiol Learn Mem*, *82*(3), 216-229. 110

White, N. M., and McDonald, R. J. (2002, March). Multiple parallel memory systems in the brain of the rat. *Neurobiol Learn Mem*, *77*(2), 125-84. 2, 77, 117, 120

Wickens, J. R., Begg, A. J., and Arbuthnott, G. W. (1996, January). Dopamine reverses the depression of rat corticostriatal synapses which normally follows high-frequency stimulation of cortex in vitro. *Neuroscience*, *70*(1), 1-5. 81

Wilson, M. A., and McNaughton, B. L. (1994). Reactivation of hippocampal ensemble memories during sleep. *Science*, *265*, 676–679. 20

Winson, J. (1978). Loss of hippocampal theta rhythm results in spatial memory deficits in the rat. *Science*, *201*, 160–163. 20

Witter, M. P. (1993). Organization of the entorhinal-hippocampal system: A review of current anatomical data. *Hippocampus*, *3*, 33–44. 17, 18

Witter, M. P., Groenewegen, H. J., Silva, F. H. L. da, and Lohman, A. H. M. (1989). Functional organization of the extrinsic and intrinsic circuitry of the parahippocampal region. *Progress in Neurobiology*, *33*, 161–253. 17

Witter, M. P., Naber, P. A., Haeften, T. van, Machielsen, W. C., Rombouts, S. A., Barkhof, F., et al. (2000). Cortico-hippocampal communication by way of parallel parahippocampal-subicular pathways. *Hippocampus*, *10*(4), 398-410. 17

Yu, A. J., and Dayan, P. (2002, Jun-Jul). Acetylcholine in cortical inference. *Neural Netw*, *15*(4-6), 719-30. 82, 112

Yu, A. J., and Dayan, P. (2003). Expected and unexpected uncertainty: Ach & ne in the neocortex. In S. T. S. Becker and K. Obermayer (Eds.), *Advances in neural information processing systems 15* (pp. 157–164). Cambridge, MA: MIT Press. 82, 83, 112

Yu, A. J., and Dayan, P. (2005, May 19). Uncertainty, neuromodulation, and attention. *Neuron*, *46*(4), 681-92. 82

Zipser, D. (1985, October). A computational model of hippocampal place fields. *Behav Neurosci*, *99*(5), 1006-18. 24, 33

Ricardo A. Chavarriaga

3, Av. Avant-Poste, 1005 Lausanne
Tel: 021.3119687/076.5030895
E-Mail: ricardo.chavarriaga@a3.epfl.ch
Web: `http://diwww.epfl.ch/~chavarri`

Age: 31
Married
Nationality: Colombian
Permit: B

Education

Current: PhD student, Laboratory of Computational Neuroscience
École Polytechnique Fédérale de Lausanne, Switzerland.
Thesis advisor: Prof. Wulfram Gerstner

2001: Graduate school in Computer, Communication and Information Sciences
École Polytechnique Fédérale de Lausanne, Switzerland.

1998: BS Degree in electrical engineering
Facultad de Ingeniería - Pontificia Universidad Javeriana, Cali, Colombia.

Experience

2001 – 2005: École Polytechnique Fédérale de Lausanne, Switzerland
Postgraduate research - Laboratory of Computational Neuroscience
• Modelling of vision-based self-localisation and navigation using neural networks (unsupervised and reinforcement learning).
• Supervision of undergraduate and masters student projects and teaching assistantship in courses on object-oriented programming and machine learning

1999 – 2000: Pontificia Universidad Javeriana, Cali, Colombia
Lecturer
• Courses on linear circuits, instrumentation and machine learning
• Supervision of undergraduate student projects

1998: Universidad de Murcia, Cartagena, Spain
Undergraduate research assistant
• Development and simulation of a kinematic controller of a robotic arm using neural networks and visual feedback

1996 - 1997: Pontificia Universidad Javeriana, Cali, Colombia
Undergraduate research assistant
• Design of digital hardware to control a 6-DoF robotic arm
• Implementation of parallel software for neural network simulation

Languages

Spanish: Mother tongue
English: Fluent
French: Fluent

Publications

Journals

Chavarriaga R., Strösslin T., Sheynikhovich D.V. and Gerstner W. Competition between cue response and place response: A model of rat navigation behaviour. *Connection Science*, 17(1-2):167–183. 2005.

Chavarriaga R., Strösslin T., Sheynikhovich D.V. and Gerstner W. A Computational model of parallel navigation systems in rodents. *Neuroinformatics*, 3(3): 223–242. 2005.

Strösslin T., Sheynikhovich D.V., Chavarriaga R. and Gerstner W. Robust self-localisation and navigation based on hippocampal place cells. *Neural Networks*, 18(9):1125-1140. 2005.

Sheynikhovich D.V., Chavarriaga R., Strösslin T. and Gerstner W. Adaptive sensory processing for efficient place coding. *Neurocomputing*. In Press.

Book chapters

Sheynikhovich D.V., Chavarriaga R., Strösslin T. and Gerstner W. Spatial representation and navigation in a bio-inspired robot. In: Wermter S., Palm G. Wermter S. and Elshaw M. (Eds.), *Biomimetic neural learning for intelligent robots*, LNAI 3575, Springer-Verlag, Berlin Heidelberg. pp. 245–264, 2005.

Peer-reviewed proceedings

Strösslin T., Chavarriaga R., Sheynikhovich D.V. and Gerstner W. Modelling path integrator recalibration using hippocampal place cells. In: W. Duch et al. (Eds), *Proc. 2005 International Conference on Artificial Neural Networks, (ICANN 2005)*, LNCS 36965, Springer-Verlag, Berlin Heidelberg, pp. 51–56.

Chavarriaga R. and Gerstner W. Combining visual and proprioceptive information in a model of spatial learning and navigation. *Proc. 2004 International Joint Conference on Neural Networks, (IJCNN'04)*, Budapest, Hungary, IEEE press, I (2004) pp. 603–608.