

A CONNECTIONIST MODEL OF SPATIAL LEARNING IN THE RAT

THÈSE N° 2951 (2004)

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

Thomas STRÖSSLIN

ingénieur électricien diplômé EPF
de nationalité suisse et originaire de Bâle (BS)

acceptée sur proposition du jury:

Prof. W. Gerstner, directeur de thèse
Prof. C. Brandner, rapporteur
Prof. A. Ijspeert, rapporteur
Prof. K. Pawelzik, rapporteur

Lausanne, EPFL
2004

Summary

When animals explore an environment, they store useful spatial information in their brains. In subsequent visits, they can recall this information and thus avoid dangerous places or find again a food location. This ability, which may be crucial for the animal's survival, is termed "spatial learning".

In the late 1940s, theoretical considerations have led researchers to the conclusion that rats establish a "cognitive map" of their environment. This spatial representation can then be used by the animal in order to navigate towards a rewarding location. In 1971, researchers have for the first time found direct evidence that the *hippocampus*, a brain area in the limbic system, may contain such a cognitive map. The activity of neurons in the hippocampus of rats tends to be highly correlated with the animal's position within the environment. These "place cells" have since been the target of a large body of research.

Apart from spatial learning, the hippocampus seems to be involved in a more general type of learning, namely in the formation of so-called "episodic memories". Models of hippocampal function could thus provide valuable insights for the understanding of memory processes in general.

Insights from animal navigation could also prove beneficial for the design of autonomous mobile robots. Constructing a consistent map of the environment from experience, and using it for solving navigation problems are difficult tasks. Incorporating principles borrowed from animal navigation may help building more robust and autonomous robots.

The main objective of this thesis is to develop a neural network model of spatial learning in the rat. The system should be capable of learning how to navigate to a hidden reward location based on realistic sensory input. The system is validated on a mobile robot.

Our model consists of several interconnected brain regions, each represented by a population of neurons. The model closely follows experimental results on functional, anatomical and neurophysiological properties of these regions. One population, for instance, models the hippocampal place cells. A head-direction system closely interacts with the place cells and endows the robot with a sense of direction. A population of motor-related cells codes for the direction of the next movement. Associations are learnt between place cells and motor cells in order to navigate towards a goal location.

This study allows us to make experimental predictions on functional and neurophysiological properties of the modelled brain regions. In our validation experiments, the robot successfully establishes a spatial representation. The robot can localise itself in the environment and quickly learns to navigate to the hidden goal location.

Zusammenfassung

Wenn ein Tier seine Umgebung erkundet, werden nützliche Informationen in seinem Gehirn gespeichert. Zu einem späteren Zeitpunkt können diese Informationen abgerufen werden und dem Tier helfen, gefährliche Orte zu umgehen oder eine zuvor entdeckte Nahrungsquelle wiederzufinden. Diese lebenswichtige Fähigkeit wird “räumliches Lernen” genannt.

Um 1948 haben Verhaltensforscher aus theoretischen Gründen postuliert, dass die Ratte eine “kognitive Karte” ihrer Umgebung erstellt. Diese Karte wird dann verwendet, um einen Zielort aufzusuchen. 1971 wurde zum ersten Mal ein direkter Hinweis gefunden, dass der *Hippocampus*, eine Gehirnregion im limbischen System, eine kognitive Karte enthalten könnte. Die neuronale Aktivität im Hippocampus der Ratte scheint im Zusammenhang mit dem Aufenthaltsort des Tiers zu stehen. Diese sogenannten “Ortszellen” sind seither das Ziel vieler wissenschaftlicher Studien.

Der Hippocampus scheint nicht nur am räumlichen Lernen, sondern an einer generelleren Form von Lernen beteiligt zu sein, nämlich an der Bildung von “episodischem Gedächtnis”. Mathematische Modelle dieser Hirnregion könnten wichtige Beiträge zum Verständnis von Gedächtnisfunktionen leisten.

Aus den Navigationsmechanismen bei Tieren gewonnene Einsichten könnten sich auch beim Entwurf von autonomen Robotern als nützlich erweisen. Die auf Erfahrung basierte Erstellung einer Umgebungskarte und deren Gebrauch für Navigationsprobleme sind komplizierte und schwierige Aufgaben. Das Einbinden von Prinzipien aus der Tiernavigation könnte helfen, fehlerunanfälligere und autonomere Roboter zu konzipieren.

Das Hauptziel dieser Doktorarbeit liegt in der Entwicklung eines Modells für räumliches Lernen bei der Ratte mittels eines neuronalen Netzwerks. Das System soll in der Lage sein, in einer realistischen Umgebung an einen versteckten Zielort zu navigieren. Das System wird auf einem mobilen Roboter getestet.

Unser Modell besteht aus mehreren vernetzten Gehirnregionen, wobei jede durch eine Zellpopulation repräsentiert wird. Dabei werden funktionelle, anatomische und neurophysiologische Eigenschaften dieser Regionen berücksichtigt. Eine Population modelliert zum Beispiel die Ortszellen im Hippocampus. Ein Richtungssystem interagiert ständig mit den Ortszellen und stattet den Roboter mit einem Orientierungssinn aus. Eine Population von Aktionszellen kodiert die Richtung der nächsten Bewegung. Für die zielgerichtete Navigation werden Assoziationen zwischen Ortszellen und Aktionszellen gelernt.

Die vorliegende Studie erlaubt funktionelle und neurophysiologische Voraussagen über die modellierten Gehirnregionen. In unseren Experimenten erstellt der Roboter erfolgreich Umgebungskarten, mit deren Hilfe er sich lokalisieren kann. In kurzer Zeit lernt er, zum versteckten Ziel zu navigieren.

Contents

1	Introduction	1
1.1	Motivation	1
1.2	Methods	3
1.3	Results and Contributions	4
1.4	Structure of the thesis	4
2	Spatial representations in animals	7
2.1	The rat hippocampus	8
2.1.1	Anatomy	8
2.1.2	Hippocampal EEG	12
2.1.3	When do place cells fire?	12
2.2	Path integration	15
2.3	Multimodal integration	17
2.3.1	Superior colliculus	18
2.4	Previous models of hippocampal place cells	20
2.4.1	Sharp (1991)	20
2.4.2	Burgess <i>et al.</i> (1994)	21
2.4.3	Wan, Redish and Touretzky (1994, 1996, 1997)	23
2.4.4	Gaussier and colleagues (1998, 2000, 2002)	24
2.4.5	Arleo <i>et al.</i> (2000, 2001)	25
2.5	Models of multimodal integration	27
3	Animal navigation	29
3.1	The basal ganglia	29
3.2	Memory systems	31
3.3	Navigation strategies in animals	33
3.3.1	Homing by dead-reckoning	33
3.3.2	Taxon navigation	35
3.3.3	Praxic navigation	35
3.3.4	Locale navigation	35
3.4	Common navigation tasks	36

3.4.1	Water maze	36
3.4.2	Radial arm mazes	38
3.4.3	Unrestricted arenas	39
3.4.4	Linear tracks	40
3.5	Previous models of rodent locale navigation	40
3.5.1	Burgess <i>et al.</i> (1994)	40
3.5.2	Brown and Sharp (1995)	41
3.5.3	Abbott and coworkers (1996, 1997)	42
3.5.4	Gaussier and coworkers (1998, 2000, 2002)	43
3.5.5	Foster <i>et al.</i> (2000)	43
3.5.6	Arleo <i>et al.</i> (2000, 2001)	45
4	Reinforcement learning	47
4.1	Temporal difference learning	50
4.2	Eligibility traces	52
4.3	Continuous spaces and generalisation	53
4.4	Relation to animal learning	53
5	Sensory input	55
5.1	Test environments	56
5.2	Visual processing	57
5.2.1	Gabor filters	60
5.2.2	Artificial retina	62
5.3	Tactile and vestibular input	63
6	A new hippocampus model	67
6.1	Architecture	67
6.2	Local view	68
6.2.1	Multicolumn cells	69
6.2.2	Column difference cells	73
6.3	Head direction system	76
6.4	Allothetic place code	79
6.5	Idiothetic place code	82
6.6	Combined place code	84
7	A model of superior colliculus	91
7.1	Architecture	92
7.2	Unimodal and multimodal cells	93
7.3	Learning the gating network	94

8	A new locale navigation model	99
8.1	Architecture	100
8.2	Action cells	100
8.3	Learning and eligibility trace	104
8.3.1	Neurophysiological evidence	104
8.3.2	Learning algorithm	108
8.3.3	Validation in test environments	110
8.3.4	Convergence of the algorithm	112
9	Conclusions	117
9.1	Contributions	117
9.2	Experimental predictions	121
9.3	Relation to other models	122
9.4	Limitations and perspectives	125

Acknowledgements

It is time to thank all the nice people who have helped and supported me during my studies at EPFL. First of all, my thesis advisor Prof. Wulfram Gerstner broadened my knowledge by sharing his great overview of the field. I also wish to thank the members of the thesis committee: Prof. Pascal Fua, Prof. Catherine Brandner, Prof. Auke Ijspeert and Prof. Klaus Pawelzik.

I am particularly grateful to Dr. Angelo Arleo for his work and the many enriching discussions. He has also been very patient and motivating and always brought me back to earth when I was flying away.

I had the great pleasure to work day by day with good people. I am thankful to the friendship and discussions with the people at LCN, LNMC and LAP. I am especially indebted to Silvio Borer, José del Millán, Ricardo Chavarriaga, Laurence Meylan, Denis Sheynikhovich, François Fleuret, Christophe Krebsler, Eric Sauser and Claudia Herzberg for their valuable contributions.

I was always very grateful to have Olivier Carmona and André Guignard nearby. They quickly tamed our Kheperas when they started to become too autonomous. A big thanks to K-Team for providing such a handy and easy to use robotic platform. Many thanks must also go to Stéphane Ecuyer, Benjamin Barras and Pierre-Eduard Sottas who made me forget how annoying a non-reliable computer system could be.

During my thesis, I could focus on research and never had to care about administrative matters. This was only possible because Marie-Jo Pellaud, Monique Dubois and Brigitte Ramuz kindly took care of all the paperwork—and still found time to bake tasty cakes.

A big “thank you” goes to the people who cheer me up every day: My friends who made my time in Lausanne collection of pleasant episodic memories. Finally, I wish to thank my family for their indefinite patience and encouragement.

Chapter 1

Introduction

When animals find an interesting place such as a food source, they can somehow store in their brains the location where the food was found. Later, when the animal is again near this “rewarding location”, they are able to recall this information and find the food again.

In order to study and compare animal performance, experiments often use similar environments. Probably the most frequently used setups are the *8-arm maze* and the *water maze* (figure 1.1). In the 8-arm maze, food is placed at the end of each arm. The animal is put on the central platform, from where it enters the arms and retrieves the food. After some arms have been visited, the animal is removed from the maze for a while. When it is put on the central platform again, it only enters arms that have not been visited yet. In the water maze, a small platform is placed at a fixed location just under the surface of the cylindrical swimming-pool. The animal can escape from the water by climbing onto the platform. The water is made opaque such that the platform is not visible. After a few learning trials, the animal directly swims to the platform from any location of the maze.

How can animals do that? This is the main question we address in this thesis. In the following, we argue why this is an interesting question and why research should be done in order to come closer to an answer.

1.1 Motivation

The task of memorising the locations of interesting places and finding them again later is termed “spatial learning”. Over the last twenty five years, enormous progress

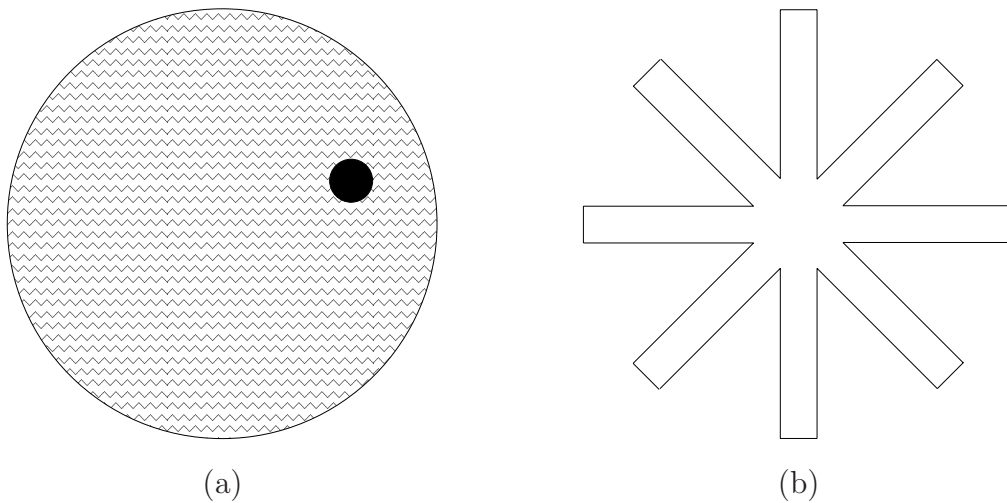


Figure 1.1: Frequently used environments for animal experiments: (a) 8-arm maze, (b) Water maze

has been made in understanding brain function in general and spatial learning in particular. For instance, it is today undisputed that the hippocampal formation, a brain area in the limbic system, is involved in spatial learning. But why is it important to know how spatial learning or the brain in general works? We propose two reasons why this could be of interest.

Understanding brain functions could help curing neurological diseases and disorders. Spatial learning seems to be a very basic function of the brain. If understood, it could provide valuable information in a much broader field of application. For instance, the hippocampus, which has such a predominant role in theories of spatial learning, is involved in much more general brain functions. Patients with hippocampal damage show severe deficits for several types of learning and memory, in particular for the so-called episodic memory.

A second reason why understanding the brain could be beneficial is that we may use the principles of brain function and incorporate them into man-made artefacts. Mobile robots, for instance, could benefit from theories of spatial learning in animals in order to become more autonomous and robust, and hence more “intelligent”.

Now, we turn to the question why there is the need for yet another model. Obviously, we are not the first to propose a neural mechanisms underlying spatial learning. Many proposals exist in which a simulated animal (which we will call “agent” from now on) localises itself in an environment and learns to navigate to a rewarding location. However, we find that in none of the existing neural spatial learning systems, the agent can localise itself in a realistic environment without the use of a compass

or a salient visual cue. Secondly, when combining information from multiple sensory systems, existing models do not adapt the importance of each sensory input to the environmental conditions, e.g. the illumination. Finally, agents in existing models learn to approach the target by moving, at each time step, in a direction drawn from a small discrete set.

The objective of this thesis is to fill these gaps. We develop a biologically plausible spatial learning system. The agent should be able to: (i) Localise itself and navigate in a realistic environment without the need of a compass or a known salient stimulus. (ii) Learn to weigh the relative importance of different sensory systems according to the environmental conditions. (iii) Efficiently learn to find a rewarding location while allowing movements in arbitrary continuous directions.

1.2 Methods

In order to achieve the aim of the thesis, we make use of several tools and methods. The two most important tools we employ are described here.

Artificial neural networks: The nerve cells in the brain are also termed “neurons”. The brain contains an incredibly large number of neurons, and each neuron is connected to many other neurons by “synapses”. Every synapse has a certain “efficacy” for transmitting information from the presynaptic neuron to the postsynaptic cell. The efficacy of a synapse is “plastic”. It can be increased or decreased, which is thought to be the underlying principle of “learning”. Neurons communicate with each other by transmitting short pulses called “spikes”. It is not known how information is coded in these trains of spikes. In this thesis, we use a simplified model of a neuron, called a “rate coded neuron”. The activity of neuron i is expressed as the mean spike rate r_i in some short time window Δt . Our network of artificial neurons is synchronised by a global clock of period Δt . This simple neuron model makes it feasible to simulate a network of thousands of interconnected neurons. Therefore, artificial neural networks are also termed “connectionist networks”.

Mobile robot platform: In order to validate our model, we implement it on an autonomous Khepera¹ mobile robot. The Khepera is equipped with a camera and other sensors (see 5.1). We also employ a simulated Khepera robot in our test experiments.

¹The Khepera mobile robot manufactured by K-Team (<http://k-team.com/>)

1.3 Results and Contributions

In this thesis, a new connectionist model of the interactions between several brain areas involved in spatial learning is proposed. Each area hosts a population of neurons which serves a particular purpose. The model is validated on a real and simulated mobile robot in three visually rich and one impoverished environment. The main results and contributions of this thesis are:

- The modelled neurons share firing properties with real neurons of the corresponding brain area. The interconnections between modelled areas are based on experimental results. The model is thus biologically plausible.
- In all test-environments, the system establishes an allocentric spatial representation of position and heading. Their spatial firing is similar to what has been reported in electrophysiological experiments in rats.
- When the agent is placed in a familiar environment and disoriented, the learnt spatial representation (position and heading) is recalibrated within four seconds using visual input. No compass or salient directional cue is necessary for this recalibration.
- The system learns to weigh visual and tactile sensory input according to the illumination [274].
- In all explored test environments, the model successfully learns to directly navigate to a hidden goal from any location after 20 learning trials. During and after learning, the robot's direction of movement is not limited to a discrete set of headings [273].
- This work allows us to make testable experimental predictions on the firing properties of place cells as well as functional and behavioral implications of lesions in the modelled brain regions.

1.4 Structure of the thesis

This thesis contains some necessary background for our model in the first four chapters. Chapter 5 describes the conditions in which our model is tested. In chapters 6 to 8, our contributions are described in detail. The topics of the chapters are:

Chapter 2 provides a review of experimental and modelling results concerning spatial representations in animals. A large part is devoted to experimental studies of the rat's hippocampal formation.

Chapter 3 is dedicated to animal navigation. In particular, we describe what kinds of strategies are used by animals. Furthermore, some typical experimental tasks are presented. An outline of previous models for animal navigation is given.

Chapter 4 is a brief introduction into the theory of *reinforcement learning*. It provides the fundamentals to the understanding of our proposed navigation model.

Chapter 5 describes the environments we used to validate our model. Four different setups which we often refer to are introduced. Furthermore, the preprocessing of the sensory input to our model is described.

Chapter 6 presents our model for the construction of a cognitive map. We describe how visual and movement-related information is combined into stable and consistent spatial representation. This chapter contains the main contributions of this thesis.

Chapter 7 provides a simple abstract model of multimodal integration. A gating network learns to weigh sensory modalities according to the environmental conditions.

Chapter 8 contains our proposal for a mechanism of animal learning. It combines our spatial representation and a reinforcement learning mechanism in order to achieve goal-directed behaviour in continuous state and action spaces.

Chapter 9 summarises the contributions of this thesis. Various experimental predictions of our model are discussed. A comparison to the previous models outlined in chapters 2 and 3 is also given. Finally, a list of limitations of our system is established and some future directions of interest are suggested.

Chapter 2

Spatial representations in animals

Animals use various navigation strategies to solve a spatial task (see section 3.3). While a simple stimulus-response behaviour is sufficient for solving easy tasks such as navigating towards a visible goal location [56], more sophisticated mechanisms are required to navigate in a complex environment. Tolman [293] first suggested that animals are using “cognitive maps” for solving such complex navigation tasks in 1948. Such a cognitive map contains information about the context of the task. It should, for instance, enable the self-localisation of the animal and support navigation by indicating the positions of interesting locations such as food sources or dangerous places. In 1971, O’Keefe and Dostrovsky [197] discovered that neurons in the hippocampus of the rat discharge as a function of the animal’s position and heading in its environment. It was quickly suspected that these “place cells” could be the neuronal substrate of a cognitive map. The rat’s hippocampus and its place cells thus have been studied extensively for the last 25 years.

In this chapter, we review the literature on cognitive maps and other representations of space that are likely to support animal navigation. The first part details the involvement of the hippocampus in spatial learning. Next, a simpler type of representation termed “path integration” is discussed. Then, we review important biological background on the multisensory maps in the superior colliculus, which is the inspiration for our model of multimodal integration (chapter 7). Finally, selected previous models are outlined.

2.1 The rat hippocampus

The startling discovery of place cells in the hippocampus [197] created a whole new field of research on the involvement of the hippocampus in spatial learning and navigation. The activity of *place cells* in the hippocampus is highly correlated with the rat's spatial location in its environment. If every location in the environment is covered by such place cells, efferent neurons could thus decode the rat's spatial location only by considering the place cells' activity [100, 99, 237, 336]. This place code may form the basis of a complex spatial learning system. This section reviews experimental results which reveal key properties of hippocampal place cells.

2.1.1 Anatomy

The hippocampal formation (HF) is a limbic brain area which occupies a considerable percentage of the rat's brain (figure 2.1 (a)). It includes the hippocampus (HPC), entorhinal cortex (EC) and the subicular complex (SbC) [7, 8].

The hippocampal formation

Inputs to HF: EC is a target of most higher cortical associative areas. HF can therefore operate on highly processed sensory information from all sensory modalities [42]. Through the fornix bundle, the hippocampal formation receives afferent connections from subcortical areas, in particular cholinergic and GABAergic projections from the medial septum. Cholinergic input targets mainly the excitatory pyramidal and granule cells, as well as inhibitory GABAergic interneurons. GABAergic septal neurons, on the other hand, selectively synapse on GABAergic interneurons only [91].

Outputs of HF: There are two main outputs of the HF: One pathway leaves the HF through the subiculum and projects to subcortical areas. It innervates thalamic nuclei, amygdala and—via the fornix fibre bundle—nucleus accumbens (NA) [329, 149]. A second pathway projects to many cortical areas through EC [126].

Internal connectivity in HF: EC is the primary sensory input area of the HPC. It consists of a medial (mEC) and a lateral (lEC) region. Figure 2.1 (d) shows a more detailed diagram of the HF internals. HPC can be further divided in: dentate gyrus (DG) and the cornu ammonis (CA). CA has four subregions, but CA1 and CA3 are the most prominent [7, 329, 8]. SbC is composed of the subiculum (Sb), the pre-subiculum (prSb) and para-subiculum (paSb) [7]. The simplified flow of sensory information can be divided into the following stages (figure 2.2): (i) EC receives sensory information from associative cortical areas and transmit them to DG, CA3, CA1 and Sb via the perforant path (pp). (ii) DG mossy cells strongly bias CA3 pyramidal cells via the mossy fibres. (iii) CA3 projects to CA1 and Sb through the Shaffer-collaterals (sc). (iv) CA1 cells synapse on Sb—as well as EC neurons. (v) Sb

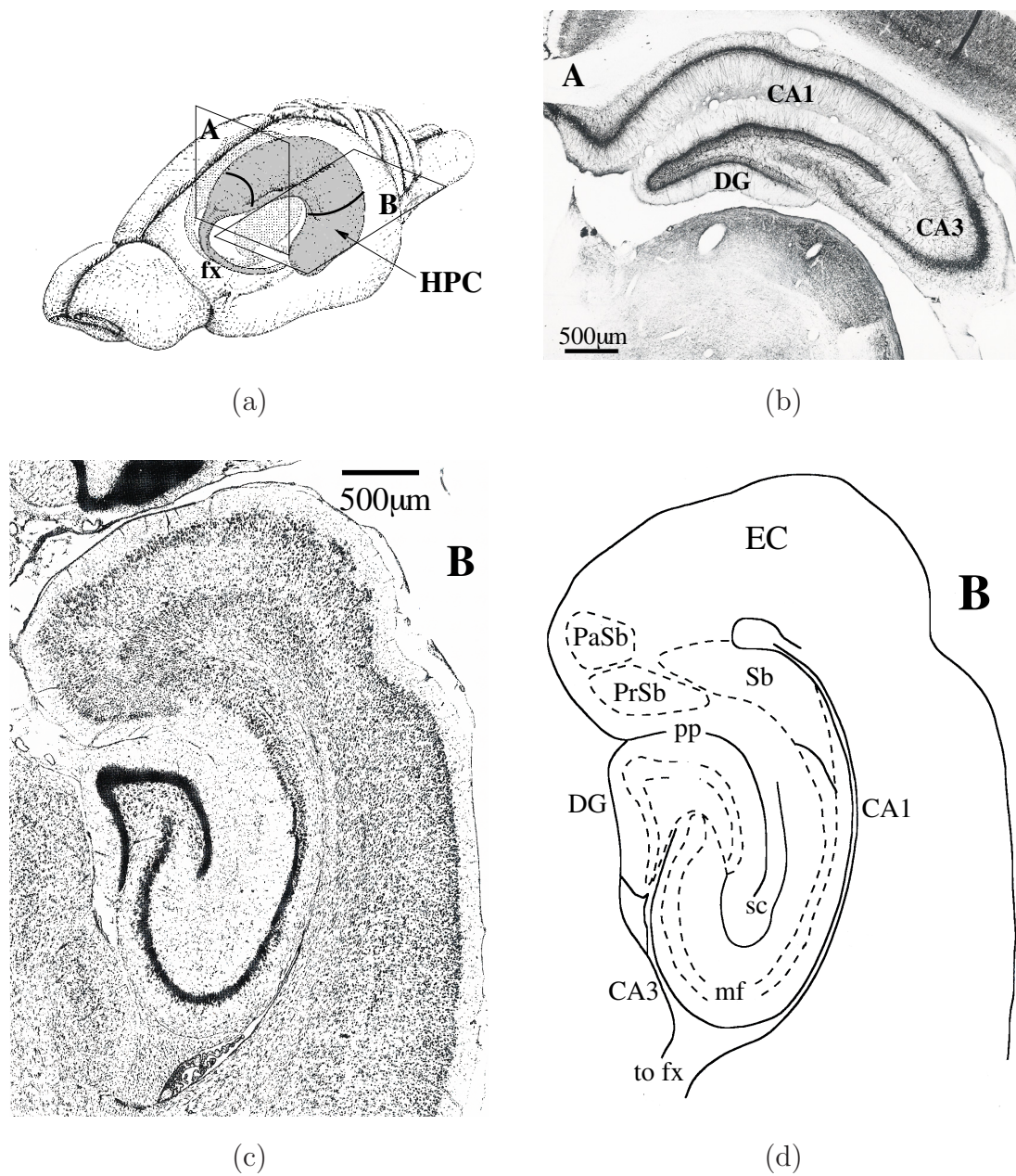


Figure 2.1: Hippocampus anatomy. All pictures adapted from [8]. (a) Schematic rat brain with highlighted hippocampus. (b) Coronal section. (c) horizontal section. (d) schematic illustration of (c). EC: entorhinal cortex. HPC: hippocampus. DG: dentate gyrus. CA1/3: cornu ammonis subregions. Sb: subiculum. PrSb: pre-subiculum. PaSb: para-subiculum. pp: perforant path. mf: Mossy fibres. sc: Shaffer collaterals. fx: fornix

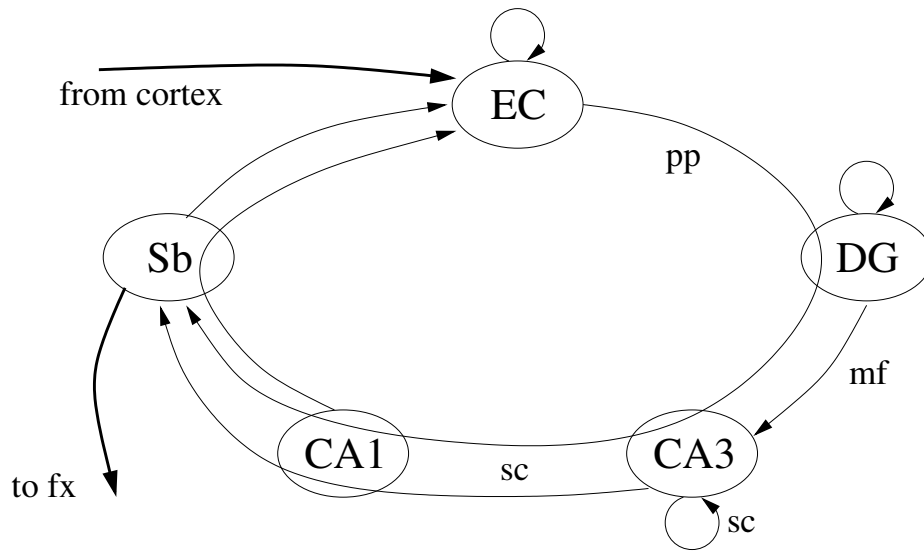


Figure 2.2: Schematic illustration of the HF internal interconnections. See figure 2.1 for a description of the labels.

sends its output via the fornix (fx) to subcortical regions, but also to EC. (vi) EC projects back to cortical areas.

Entorhinal cortex

Entorhinal cortex is the cortical gateway to the hippocampus. A distinction can be made between the medial and the lateral areas. EC follows the general neo-cortical architecture of a six-layered structure that has been extensively studied [141, 329, 126]. Nevertheless we do not make a distinction between layers of EC for reasons of simplicity and focus on the differences between the medial and the lateral areas.

Inputs to EC: 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 [329, 283, 126, 152]. Olfactory information from the olfactory bulb and piriform cortex is conveyed directly and via perirhinal cortex [141, 331, 249, 152, 30, 332, 127]. The subiculum as well as CA1 send their output mainly to mEC, but also to lEC. DG and CA3 don't innervate EC [7, 329]. Pre-subiculum also synapses on mEC.

Outputs of EC: EC cortical efferents target primarily perirhinal, orbitofrontal and piriform cortices, but parietal, temporal, frontal and occipital areas are also innervated [60, 329, 126]. Via the perforant path, EC conveys multisensory information from its cortical afferents to DG, CA3, CA1 and Sb [7, 329].

Internal connections in EC: Neurons in the deep regions connect to cells in the superficial layer of EC [7,329,8]. There is also evidence for strong synaptic innervation from the lateral to the medial region [221].

Dentate gyrus

DG granule cells receive processed sensory input from EC. Granule cells then project to mossy cells. Mossy cells laterally contact other mossy cells as well as strongly project to CA3 [55,8,115]. Throughout the entire life, neurogenesis occurs in the rat DG. Stem cells migrate into the granule layer and differentiate into fully functional and networked granule neurons [24,143,54,115].

Cornu ammonis or hippocampus proper

The hippocampus proper consists of four subregions CA1-CA4, with CA1 and CA3 being the most distinguishable. Place cells have originally been found in CA1 pyramidal cells in 1971 [197].

Inputs to CA: The CA3 region receives strong projection from DG via the mossy fibres. Both CA1 and CA3 are also innervated by EC via the perforant path [7,5,329,8,115].

Outputs from CA: CA1 and CA3 pyramidal cells connect to subiculum via the Shaffer fibre bundle. The angular bundle connects CA1 to EC (perforant path) [6].

Internal connections in CA: CA3 neurons laterally send outputs to other CA3 neurons via the Shaffer collaterals. Also via the Shaffer fibres, CA3 connects to CA1 [7,5,8].

Subiculum

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

Inputs to SC: The main input to Sb originates in CA1 and EC [6]. The paSb is innervated by the retrosplenial cortex whereas prSb is reached from areas in the temporal and parietal lobes, as well as from thalamic nuclei, which project onto poSb [40].

Outputs from SC: Via the fornix fibre bundle, Sb projects on the nucleus accumbens (NA) and the septal complex. Sb also innervates entorhinal and prefrontal cortex, amygdala and the thalamus. prSb and paSb also synapse on EC [7,329].

Internal connections in SC: Within SC, Sb projects to prSb and paSb and prSb also synapses on paSb [7,329].

2.1.2 Hippocampal EEG

The hippocampal EEG shows distinct patterns depending on the rat's behaviour. During motion (active or passive [98]), the EEG shows a 6–12Hz oscillation called *theta rhythm* [199, 41, 264]. Theta is also observable during sensory scanning (eg. sniffing) and REM-sleep [46]. On top of that, firing is synchronised to a *gamma oscillation* of 40–100Hz throughout the whole hippocampal formation [53, 59]. Subcortical cholinergic and GABA-ergic inputs from the septal region seem to be responsible for generating the hippocampal theta rhythm [326, 44, 173, 113]. When septal input is inactivated, the CA3 place fields are disrupted whereas the CA1 fields are unaffected. Simultaneously, acquisition of place learning tasks is impaired [37] and errors in working memory increase significantly [178].

While eating, drinking or awake-immobility as well as in slow-wave sleep, however, the 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 [53, 52]. It is speculated that theta/gamma waves synchronise input from cortex to hippocampus, whereas sharp waves/ripples modulate output from hippocampus back to cortex [53].

2.1.3 When do place cells fire?

Hippocampal place cells fire at rates of up to 50Hz when the rat traverses a specific portion of the environment whereas it is almost completely silent otherwise [195]. The shape of place fields varies: The firing rate can be distributed approximately like a circular Gaussian in 2d (same variance in all directions), but it may also be elongated in one dimension (especially along walls) or in rare cases have multiple peaks of activity [189, 195, 163, 324].

Sensory cues

Distal vs. proximal cues: In many experiments in various mazes (see section 3.4), rats can see local landmarks within the maze (such as corners or irregularities on the maze walls) as well as distal cues outside of the maze (ie. posters on the walls) [196, 174, 187, 189, 200, 253, 36, 221, 28, 334, 195, 58]. When the arena is rotated, the place fields generally stay relative to the *distal* cues [196, 174, 200, 253, 334, 58]. If, however, a reward is given relative to a local cue, place cells can learn to fire relative to this cue and ignore distal cues. This is particularly true for place cells near the landmark [28, 29, 108].

Directionality: Place cells fire at specific locations in an environment. However, a place cell may fire independently of the rat's heading or it may fire only when the rat is oriented in a specific direction. In general, open field environments tend

to produce omnidirectional firing [163, 186, 158, 195]. In linear tracks, radial mazes and square loops, where the rat's trajectory is stereotyped, place cells which fire at a specific location are silent when running in the opposite direction [163, 159, 107, 167, 169]. Firing is also directional in an open environment where the rat is not forced—but motivated (by food) to run in a square loop [159]. This suggests that it's the experienced trajectory or local views—not the maze type—which determines directionality.

Non-visual sensory cues: Place cells are strongly influenced by vision. However, somatosensory, olfactory and internal (path integration) cues seem to contribute to the formation and maintenance of place fields [120, 241, 242, 145, 146, 233]. When landmarks are removed from an explored environment, the place code is not necessarily disrupted [196, 120, 217, 187, 200]. Stable place fields can also be established in the dark and they persist when the light is turned on after exploration [220, 158, 242]. Rats that have been blinded shortly after birth can show normal place fields as adults [241]. Congenitally blind and deaf rats follow local cues when the maze is rotated [120] and show landmark-relative firing (but only after the first contact with a local landmark [242]). A place code established in normal lighting conditions persists for up to 8 minutes when turning off the light [220]. It is suggested that this time is bridged by path integration. Cleaning the floor during or between experiments, however, decreases this persistence to 2 minutes. Not cleaning the floor, on the other hand, strongly increases the stability and coherence of the place code even in the presence of visual cues, suggesting an important role of olfactory cues [242].

Environment geometry: Place fields have been shown to stretch if the environment is resized [187]. In rectangular arenas, fields that are centred at a certain ratio between walls are centred at the same ratio when the environment is stretched in one dimension [195].

Small or gradual changes: When cue cards are slightly rotated during a session, place fields follow the cue rotation whereas they don't if the rotation is large and quick [234]. In an environment with multiple landmarks, place cells can be stable when removing some—but not all landmarks [196, 187, 200]. This suggests that place cells learn to ignore unstable landmarks if there are sufficient other cues available.

Sequential activation

Phase precession: In the behaving rat, the firing times of place cells in CA is synchronised to the phase of theta. The relationship between place cell firing and theta phase depends on the animal's experience: When the naive rat traverses the place field of cell i , it fires at the constant phase $\theta_i = \theta_i^0$. After many repetitions of traversing the place field, θ_i starts depending on the position of the rat within the field: When the rat enters the field, the cell still fires at θ_i^0 . As the rat moves towards the centre and the far end of the field, however, the firing phase θ_i decreases systematically. This

shift is termed *phase precession*. It is a temporary effect and seems to have vanished on the following day [199, 264]. There are some interesting consequences due to phase precession: (a) The phase of firing provides additional information about the rat's location *within* a place field [41]. (b) Within one theta cycle, two place cells with adjacent fields fire in the same order as the order of activation due to the rat's movement [264, 130]. This temporal compression might be important for learning sequences on a behavioural timescale in the framework of spike timing dependent plasticity, which requires a much shorter timescale [170]. Models of how phase precession could emerge are proposed in [300, 35, 153, 168, 111]

Experience dependent Shift: If the same trajectory is repeated many times, CA1 place cells on that trajectory start to enlarge their fields backwards and become asymmetric [167, 168]. This can be viewed as a mechanism to predict future location [188]. However, this effect vanishes overnight and is not persistent across environments: A place cell with asymmetric field in a well-known environment can show a symmetric field in an unfamiliar environment. It has been proposed that this asymmetric shift is necessary to produce phase precession [168], but see also [125].

Replay during sleep: Place cells which were active during an experimental session are more likely to be activated during subsequent sleep [213, 259]. Furthermore, cells which are co-activated during a session (because their fields overlap) also show correlated firing during sleep [325]. The temporal order of firing is also correlated. This suggests that recent sessions may be replayed during sleep, possibly helping to consolidate the acquired episodic memory [263].

Context dependence

multiple environments: In any environment, only a subset of 10-25 percent of all pyramidal cells are active. All other cells are completely silent [291]. Place cells seem to be randomly attributed to these subsets and any cell may be active in several environments. If it is, however, the place fields in those environments are totally unrelated [142, 291]. Whenever an animal is returned to a familiar environment, place cells show fields consistent with previous sessions in the same environment [189, 292].

No topology: Place cells are not topologically arranged in the hippocampus. If two place cells code for a neighbouring place in an environment, they need not be anatomical neighbours in the rat's brain. Neighbours in the hippocampus needn't show neighbouring place fields either [196, 142, 187, 291]

Tuning speed: It is not clear how long it takes to establish a place code. Place cells have been observed to get tuned persistently the first time the rat traverses through the cell's receptive field [119]. However, further investigation shows that it can take from 10–30 minutes, or even hours until the place code is stable [324, 284].

Non-spatial determinants: Place cells are also sensitive to information other than the rat's location. For instance, their firing activity depends on speed, orientation,

task, reward, odour or tone discrimination and many other variables [163, 158, 159]. In general, the hippocampus seems to encode and memorise the whole context of an episode, in which spatial information is only one part.

Remappings: Entire place codes can suddenly and instantly switch to a different representation for the same environment. This event is termed a “remapping”. A different subset of place cells may be active for each task in the same environment. The subsets needn’t be disjoint and place cells seem to be randomly attributed to those subsets. Cells that are active in several tasks may have different place fields [159]. The effect largely resembles the case of multiple environments.

Place cells outside the hippocampus proper

Place cells have mainly been recorded in CA. However, the other areas of the hippocampal formation also host cells with spatial firing properties:

Dentate gyrus (DG): DG granule cells show directional place cells in the radial arm maze. Their place fields are significantly smaller than for CA pyramidal neurons [135]. Granule cells also exhibit phase precession [264, 106].

Subiculum (Sb): In the subiculum place cells show less spatial specificity than in CA [23]. Firing depends not only on place, but also on speed of motion [162]. Depending on the trajectory constraints, CA place fields may be omnidirectional (eg. in open field exploration). Subicular place fields, however, are directional even in open environments [258]. Sb place cells also preserve their topology across scaled and reshaped environments: The field of a place cell in a square environment is on the same position and has the same shape in a round arena. In a scaled environment, the place field scales with the environment [258, 255, 257]

Entorhinal cortex (EC): Similarly to Sb place cells, neurons in the medial entorhinal cortex (mEC) preserve firing topology across reshaped environments. Place cells are also likely to be active in any environment, in contrast to CA place cells, which may become silent or code for a completely different place [221]. Place fields in EC are also less spatially specific than in CA. They fire in all regions of the arena, with a peak at its dedicated position [221, 89, 90]. CA place cells, however, are silent outside of their field. EC firing is also modulated by theta [179].

2.2 Path integration

Homing by dead reckoning (HDR) is the ability to return to the starting point of a journey, eg. the nest location, without using allothetic cues [22, 177]¹. This is

¹Allothetic cues means that the source of information is outside of the animal. Visual input for instance is allothetic. In contrast, idiothetic cues are internally generated. Vestibular or proprioceptive cues are examples of idiothetic sensory input.

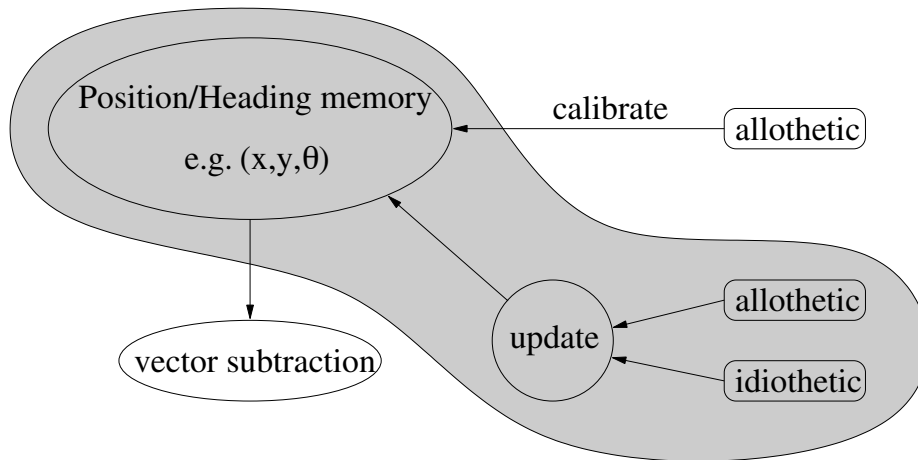


Figure 2.3: Path integration as a spatial representation (shaded area). The essential building blocks are: (i) A memory of the spatial variables to integrate, (ii) sensors which provide displacement information in various formats, and (iii) an update function which transforms from sensory specific frame into the common frame of the memory.

considered an easy task and a wide variety of species can solve it (see section 3.3.1 for a review of behavioural experiments on HDR). In this section, we focus on the spatial representation needed for HDR. In this thesis, we call this representation a *Path integrator* (PI). It is less complex than a hippocampal cognitive map. Indeed, an intact hippocampus is not necessary for HDR [4].

In its simplest form, a path integrator is a system composed of three modules (figure 2.3): (i) a memory of spatial variables (e.g. the position and heading of the agent) in an arbitrary coordinate frame. (ii) Sensory input (e.g. optic flow, vestibular, proprioceptive etc.) which provide relevant information about the change of the spatial variables in sensor-specific arbitrary coordinate frames. (iii) An update mechanism which transforms and combines all sensory information into the common frame of the spatial memory. Additionally, an external signal may recalibrate the path integrator’s memory in order to remove drifts or reset the spatial variables to a known initial state.

In our terminology, we focus on the memory aspect of PI—the storage and update of the agent’s position and heading for example, and not on the navigation part, such as *using* the stored values for homing. Other definitions of PI are more similar to what we term HDR. They include a simple navigation strategy based on vector subtraction and exclude allothetic input for the update (see section 3.3.1 and compare figures 3.2 and 2.3),

The outputs of a PI system are the spatial variables which allow the animal to

localise itself or other places (food, nest). For some HDR tasks, only the egocentric bearing of the nest needs to be integrated [260]. Let us focus on the more common view, where path integration represents the animal’s location in space with respect to some reference point (e.g. the nest). Often, a distinction is made between a head-direction and a position system. In our terminology, they both take part in path integration. The output of the PI then is the current position $\vec{p} = (x, y)$ and heading θ of the agent.

The input is a small displacement signal since the last update. This input consists of internally generated (idiothetic) cues, such as motor efferent copy, vestibular and proprioceptive input as well as external (allothetic) cues (e.g. optic flow), which can, when available, reduce the effect of systematic drifts [103, 43, 176, 164].

Independently of the origin of the input signal, path integration contains memory², in the form of the stored position and heading.

The question of how the animals perform PI is not known, but several brain areas are known to be related to PI. Candidates for the neural locus of the position system in mammals include the hippocampus [162, 238], subiculum [228], para-subiculum [228] and entorhinal cortex [228, 255, 256]. However, animals with hippocampal lesions are still capable of HDR [4], which is not consistent with [162, 238]. Brain areas suspected to be involved in a head direction system are: The postsubiculum [222, 288], anterodorsal [285, 31, 139] and laterodorsal [180, 179], nuclei of thalamus and the lateral mammillary nuclei (LMN) [150, 268].

2.3 Multimodal integration

Multimodal integration is the process of combining information from different sensory organs into a common representation. The multimodal experience and its origins have been intensely studied during the last 50 years. Psychologists and psychophysicists were the first to examine multimodal integration and its development in human infants.

Piaget [216] stresses the fact that infants explore their environment *actively* and particularly like multisensory experiences. This motivation may be important to maximise information acquisition. Intermodal correlations are then established with touch being the reference to which other modes are correlated to.

Gibson [103] puts a strong emphasis on intersensory aspects for perception of self-motion. Not only vestibular and proprioceptive information from muscles and joints, but also vision and maybe others may contribute to the percept of motion. One consequence is that touch would not necessarily be more trustworthy than any other sense (see [43] for a review of those early psychophysical experiments).

²When we just use the term “memory”, we refer to working memory, see section 3.3

From evolutionary thoughts, Marks [157] concludes that a union exists between senses already at birth and a differentiation takes place during early stages in life. All senses share some common characteristics and therefore a common representation exists for all sensory modes.

More recently, lesion studies aim at finding brain areas which are important for multisensory perception. It has been revealed that the superior colliculus (SC) seems to play a major role in spatial orientation and intersensory processing [270]. SC lesions result in severe sensory (especially visual) neglect and spatial orientation deficits. In order to find out *how* this brain region participates in multimodal integration, a series of electrophysiological experiments try to characterise the properties of individual neurons in superior colliculus [270]. The following section summarises the results of these investigations.

2.3.1 Superior colliculus

The superior colliculi (SC) are two “hills” on the dorsal surface of the midbrain. In lower animals, they are known as the optic tectum. SC is a seven-layered structure, but most commonly, only two parts are distinguished: The three most dorsal layers form the *superficial*, the four layers below are the *deep* part. In contrast to the inferior colliculi which are purely auditory, the superior colliculi receive at least visual, auditory and somatosensory inputs. The left colliculus processes input of the right sensory hemisphere and vice versa [270]. When referring to the superior colliculus (singular), we address any one of the two bumps if not specified otherwise.

Superficial layer: The superficial layer is mainly devoted to visual processing. In primates, SC is well known for its implication in saccadic eye movement. There is a large body of research on this topic, see [333] for a review. In the cat, oculomotor reflexes also seem to be influenced by SC: If a target appears in the visual field, a specific subset of SC neurons seem to initiate an orienting movement to bring the target into the view centre. The direction and amplitude of the shift seems to be coded by a population of SC neurons [191, 192]. Once the target is fixated, another set of neurons fires tonically while the target is in the centre [191]. It has also been shown that these shift-neurons form a topological map [192, 270].

Deep layer: Deep layer neurons respond to visual, auditory and somatosensory stimuli. They are innervated by ascending subcortical sensory structures as well as by descending cortical areas. The cortical afferents origin in sensory cortical areas and are therefore mainly unimodal. Some cells in deep SC respond only to a single modality. The majority of its output neurons, however, seems to be multimodal. They respond to a combination of visual, auditory and somatosensory stimuli. Efferent projections target nuclei of the thalamus, the opposite superior colliculus, and—via two pathways—brain stem and spinal cord [270, 305].

The various unimodal sensory neurons in SC form topological maps. There are

maps for at least visual, auditory and somatosensory stimuli. Neurons seem to represent the *spatial location* of important stimuli's origins. It is not always clear in which coordinate frame objects are encoded, but there are indications that visual maps are retinotopic, whereas auditory maps are head-centred and somatosensory maps are body-centred [72, 270, 171, 308]. Motor neurons which serve the purpose of orienting the sensory organs (eyes, ears, head, body) towards interesting stimuli are also organised in topological maps.

The majority of the SC's output neurons seem to be multimodal. As they respond to a combination of topologically arranged unimodal cells, multimodal neurons also form a map like structure [72, 214, 270, 171, 308]. It is not clear how the input from different frames are combined, but as the receptive fields of unimodal neurons are large, it has been suggested that a small displacement of one map (eg. due to an exocentric eye position) would not completely disrupt the multimodal map. Another possibility is that unimodal maps are coded in motor-error coordinates with respect to the dominant sense [214, 270].

An interesting property of multimodal units in SC is *multimodal enhancement*: Presenting a visual or auditory stimulus alone is much less effective in making it fire than presenting the stimuli together. In order to enhance the response of a neuron, the spatial and temporal location of the stimuli have to coincide. If a sound is not coming from the same location than the visual stimulus, however, the response is depressed [270, 306]. Recent experiments enforce the belief that cortical input to SC is necessary to produce multimodal enhancement [307, 133, 132]. The enhancement is not present at birth, but rather develops through experience [309]. The distribution of uni- and multimodal cells strongly depends on the species. In primates and cats, most unimodal cells are visual and multimodal neurons are predominantly visual-auditory, whereas in the rat, somatosensory units are the most numerous and tactile-visual are the most common multisensory cells [71, 317, 270].

Recent experiments also show that the basal ganglia—to which SC is reciprocally connected—use similar enhancement and depression processing than SC and thus modulates sensory activity in the superior colliculus [131]. The basal ganglia are thought to be involved in selecting an appropriate motor response in conflicting behavioural situations. In sections 3.3.2 and 4.4, the relation of the striatum (a region of the basal ganglia) to animal navigation is further discussed. In particular, during “taxon navigation” (section 3.3.2), rats seem to make use of superior colliculus to orient themselves towards a goal location.

In the rat, over 80 percent of 93 cells measured in superior colliculus show theta field activity temporally coincident with spatial firing in hippocampus (see section 2.1.2). It is suggested that superior colliculus may be linked to hippocampal theta-generating structures [194]. In rats performing a spatial memory task, spatial firing patterns (tuned to place, orientation, or both) have been found in SC [57].

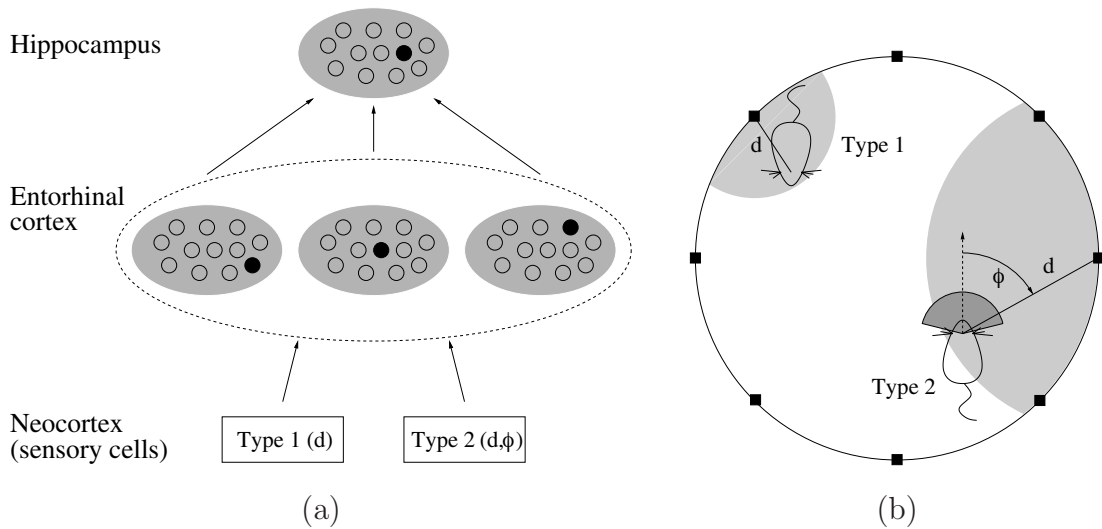


Figure 2.4: Place cell model proposed by Sharp (1991): (a) system architecture. Shaded areas mean winner-take-all competition. (b) Environment with eight landmarks. Sensory cells encode the view considering only the distance d (type 1), or the distance d as well as bearing ϕ to dedicated landmarks (type 2).

2.4 Previous models of hippocampal place cells

2.4.1 Sharp (1991)

The model proposed by Sharp [254] builds on visual cells that encode the agent’s distance and bearing to several landmarks in the environment. Three layers of simulated rate-code neurons and Hebbian-type learning between them results in a spatial representation similar to hippocampal place cells. The circular environment contains eight landmarks evenly spaced on the wall (figure 2.4).

In the first layer, two types of neurons encode the agent’s view of the simulated environment. Each cell is “responsive” to one or several landmarks. Type 1 neurons fire whenever the agent is within a certain range of its landmark, i.e. distance d to the landmark is smaller than a neuron-specific maximum. The size of this range varies among cells and can extend over the centre point of the arena. Type 2 cells also only fire when the distance d to the landmark is within a given characteristic limit. An additional requirement, however, is that the agent’s bearing ϕ with respect to its landmark lies within a cell-specific limit. The size of these receptive bearing angles varies from 80° to 170° . The tuning of these cells (which landmarks at what distance and bearing range activate a given cell) of both types of cells is random and remains

fixed throughout the experiments.

The intermediate layer is comprised of three populations of 20 neurons. Each cell j receives input $h_j = \sum_i w_{ji}$ from all sensory neurons i of the first layer. In each cluster, the cell j^* with the largest input $h_j^* = \max_j h_j$ fires with a rate of h_j^* . All other cells are silent [235]. A neuronal implementation of this winner-take-all (WTA) scheme could be lateral interconnections within each cluster. All three active cells of this layer are eligible for Hebbian type learning on their input synapses w_{ji} . Sharp proposes entorhinal cortex as location for this layer.

The three active cells of the second layer project their activity onto the third layer, which consists of one cluster of 20 neurons. The same WTA mechanism as in the entorhinal layer is applied here, and again, the winner-neuron adapts its weights using Hebbian learning. In this layer, which is proposed to be located in the hippocampus, place cells with omnidirectional place fields are reported by Sharp for simulations in the circular environment. When the movement of the agent is restricted to follow paths like in an eight-arm-maze, simulated place cells are unidirectional, which is consistent with recordings in rats.

This model relies on an abstract visual system where the exact distances to the eight landmarks are available. All eight landmarks are also perfectly distinguishable by the sensory cells. In the absence of light, this model can not sustain spatial firing due to the lack of a path integrator. Rats can, however, maintain place fields in the dark [220, 158, 242]. Finally, a winner-take-all mechanism is applied to the modelled entorhinal cortex and hippocampus. Experimental data however suggests a distributed and redundant coding of place in the hippocampal formation.

2.4.2 Burgess *et al.* (1994)

Burgess *et al.* [41] offer a model of place cells consisting of four layers of neurons, as depicted in figure 2.5 (a). Visual cells are based on distances to landmarks placed near the walls. The effect of theta-phase precession is reproduced by the system.

A set of 15 sensory cells is attributed to each landmark of the environment. Per theta cycle, sensory cell i fires a number of spikes n_i which depends on the difference between the actual distance to the landmark and its specific preferred distance d_i . The tuning curves are large and of triangular shape, and the preferred distances uniformly cover the environment.

One layer above, each cell in entorhinal cortex (EC) receives input from two predefined sensory cells i and j and fires $\lfloor n_i \cdot n_j / 2 \rfloor$ spikes. The two afferent sensory cells are chosen such that each is coding for a different landmark and the location of their peak firing activities coincides with the centre of the entorhinal cell's receptive field. The angle α between the agent's heading and the egocentric orientation of the place-field centre determines the phase (with respect to the theta rhythm) at which spikes are fired (figure 2.5 (b)): If $|\alpha| < 60^\circ$, the cell fires at a "late" phase, if $60^\circ < |\alpha| < 120^\circ$, the phase is "middle", and else the cell fires "early" in the theta

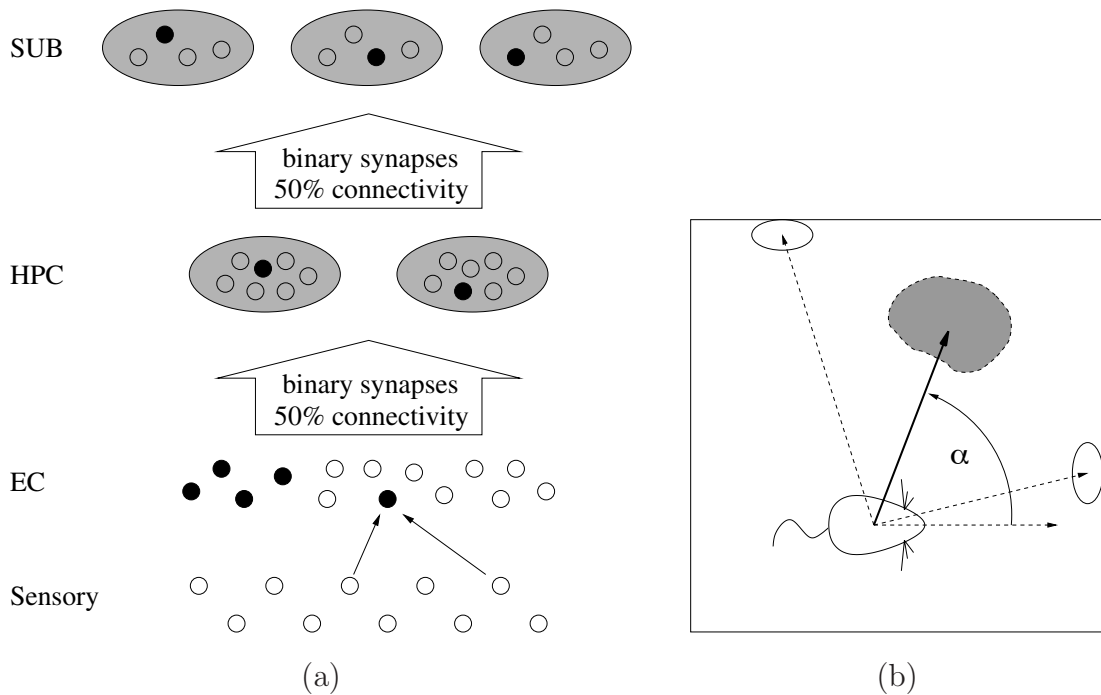


Figure 2.5: Place cell model proposed by Burgess *et al.* (1994): (a) system architecture. Shaded areas mean winner-take-all competition. EC: Entorhinal cortex, HPC: Hippocampus, SUB: subiculum. (b) θ Phase precession: Firing of EC neurons is “late” when the angle $|\alpha|$ between agent heading and place field centre is $< 60^\circ$, “middle” when $60^\circ < |\alpha| < 120^\circ$ and “early” when $|\alpha| > 120^\circ$. The shaded area is the place-field centre, which is, by construction, located between the two landmarks.

cycle. Burgess *et al.* postulate that phase precession as seen in hippocampal place cells is generated in EC, which forwards this information to its target structure.

Each neuron in the EC layer connects to half of the cells in the hippocampus (HPC). The synaptic weight is binary (0 or 1). Initially, most connections are turned off. A Hebbian-type learning rule allows these synapses to be switched on if the pre- and postsynaptic cells are both maximally active. The input to each place cell in HPC is proportional to the sum of presynaptic spikes at active synapses. Neurons in the HPC layer are clustered into five groups of 50 neurons. In each group, only the four cells with largest inputs are allowed to fire spikes [235].

Each HPC place cell projects to half of the cells in the subicular layer (SUB) of the model. The same Hebbian learning procedure as between EC and HPC is implemented here. The only difference is that in SUB, cells are arranged in ten

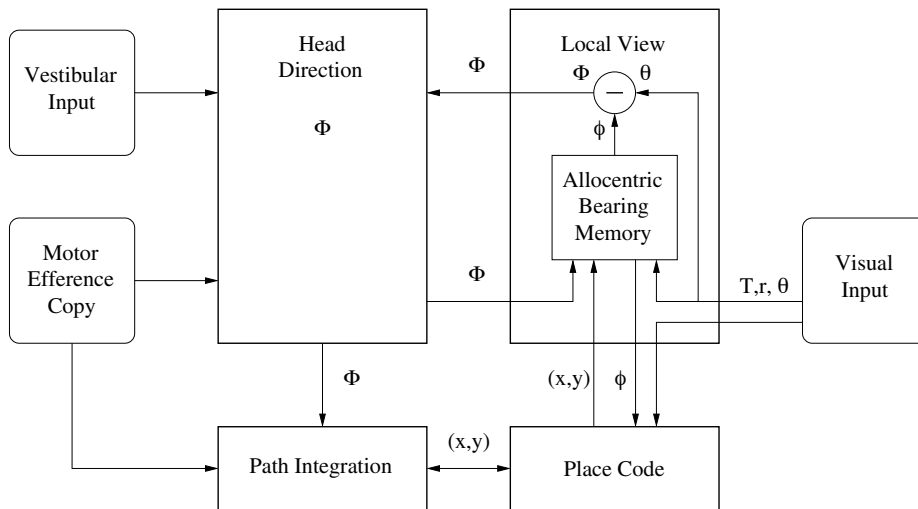


Figure 2.6: Architecture of the model proposed by Wan, Redish and Touretzky. During exploration, landmark information (type T , distance r and egocentric bearing θ) is combined with head direction Φ to produce allothetic landmark bearing ϕ . The allothetic bearing memory stores (T, r, ϕ) . Path integration updates position (x, y) . Place cells use all this information to tune their receptive fields.

groups of 25 neurons. As a consequence, each cell has to compete with less cells, which results in larger place fields as in HPC.

The allothetic input to this model is algorithmic. It consists of the exact distance to the arena walls. Furthermore, the mechanism which determines the phase of firing with respect to the theta rhythm depends on the bearing to the cell's centre of receptive field. It is not clear how the rat can compute this bearing. This model does not include a path integrator and is thus incapable of producing spatial firing in the dark.

2.4.3 Wan, Redish and Touretzky (1994, 1996, 1997)

The system by Wan, Redish and Touretzky [310,296,228,229] consists of separate populations for the local view, head direction, path integrator and place code (figure 2.6). All populations interact with each other in order to form a consistent representation of space.

The visual input provides the system with information about landmarks. In their simulated environment, the type T_i , distance r_i and bearing angle θ_i to each landmark i enters the local view system. The compass bearing ϕ_{ik} of landmark i , viewed from place (x, y) is then calculated and stored. This calculation also requires information about the agent's current heading Φ which is provided by the head direction

system.

The head direction system keeps track of the compass bearing Φ of the agent by integrating angular velocity signals from vestibular cues and efferent motor copies of motor commands. If the agent is disoriented, the current heading can be reset by the local view system by comparing the egocentric and compass bearing of landmarks.

The path integrator updates the agent's position (x, y) within the environment by summing up motor efference copies. When reentering a familiar environment, the internal state of the path integration system may be incorrect. It can, however, be recalibrated using visual input. This is done via the place code module.

The place code population combines information about the local view and the internal path integration system. Each newly recruited place unit tunes to the following parameters: (i) Type T , distance d and compass bearing ϕ of two randomly chosen landmarks, (ii) retinal angle difference $\alpha = \theta_i - \theta_j$ between two (possibly different) randomly selected landmarks, (iii) position information (x, y) given by path integration. Place units compute a “fuzzy conjunction” of their inputs in which terms that are unavailable or thought incorrect drop out.

This model relies on an abstract allothetic input which features a perfect measure of the landmark type, distance and bearing. The computations in the model are performed in an algorithmic instead of a neuronal way.

2.4.4 Gaussier and colleagues (1998, 2000, 2002)

The models by Gaussier and colleagues [96,95,97] relies on the detection of landmarks in real panoramic camera images. Spatial information is coded by the transition between places (figure 2.7).

In a first step, visual features are extracted from the panoramic view. For each detected landmark in turn, its type and compass bearing (the agent has a built-in compass) are represented in a merged “what” and “where” matrix. The type neurons form a winner-take-all network, whereas the compass bearing network supports generalisation by “spreading” activity to neighbouring neurons.

When a place seems interesting (e.g. close to a goal location), a place cell of the place recognition layer is recruited and units from the view-matrix connect to it. At each time step, the activity of place cells is calculated in two steps: First, the initial activation act_i of place cell i is determined according to

$$act_i = 1 - \frac{1}{\pi N_i} \cdot \sum_{k=1}^N V_{ik} \cdot [v_k |\Theta_{ik} - \theta_{ik}|_{\pi} + \pi(1 - v_k)] \quad (2.1)$$

where N is the number of detected landmarks in the current time step, N_i is the number of visible landmarks when place cell i was recruited, Θ_{ik} is the compass bearing of landmark k viewed from place i , θ_{ik} is the compass bearing of the same

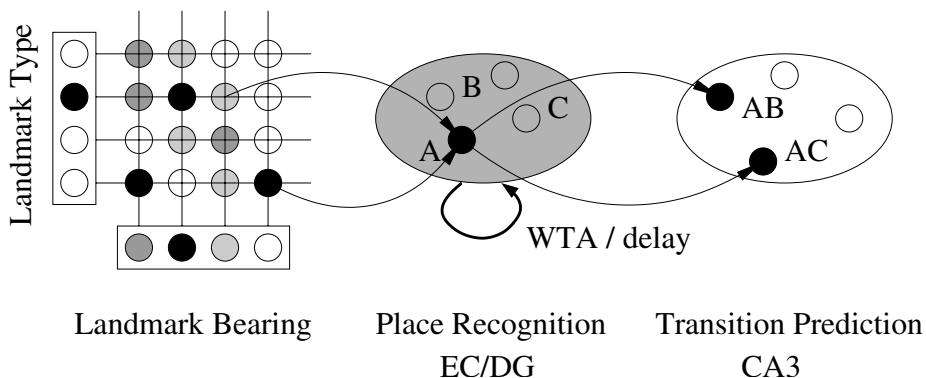


Figure 2.7: Architecture of the model proposed by Gaussier and colleagues. Landmark type and bearing are extracted and merged in a recognition matrix. Place cells compete with each other implementing a winner-take-all (WTA) mechanism. Place cells allow the following layer to predict potential future places by associating the delayed place of the previous time step (e.g. place A) with the current place. After transition learning, possible future places (e.g. B or C) can be predicted.

landmark viewed from the current position. V_{ik} is set to one if landmark k was visible when cell i was recruited or zero otherwise. v_k is set to one if landmark k is visible from the current location and zero otherwise. $|\cdot|_{\pi}$ is the absolute value modulo π . Second, a winner-take-all mechanism resets the activities of all but the winning cells to zero. The place recognition module is attributed to cells in entorhinal cortex (EC) and dentate gyrus (DG).

A delay in the place recognition layer allows the next layer to learn place transitions: Previously active cells connect to transition cells using a Hebbian-type learning rule.

Allothetic input to this system is provided by a real camera. However, these camera images are aligned using an artificial compass. In the place recognition layer, a winner-take-all mechanism suppresses all but one neuron. This is in contradiction with experimental data which reveals the distributed and redundant nature of the hippocampal place code. In the absence of visual input, this model does not work due to the lack of a path integration component.

2.4.5 Arleo *et al.* (2000, 2001)

Arleo *et al.* [13,14,16,15] propose a spatial learning system based on low-level feature extraction of real camera images from a miniature robot. Idiothetic representations are calibrated using visual stimuli (figure 2.8).

The feature extraction module transforms the high-dimensional camera image

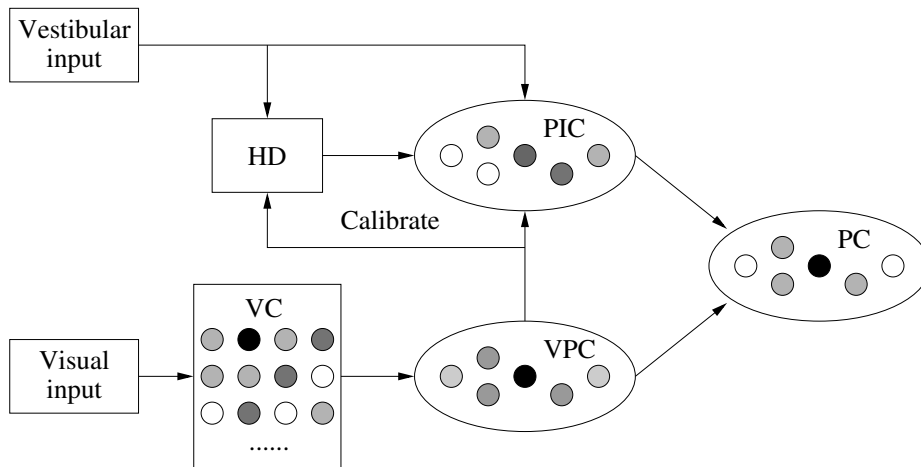


Figure 2.8: Architecture of the model proposed by Arleo *et al.* Low-level visual features are extracted and stored in view cells (VC) which drive visual place cells (VPC). Vestibular cues update the head direction (HD) system as well as a population of path integration cells (PIC). VPC and PIC converge onto a population of place cells (PC). Synapses are modified using a Hebbian learning. HD and PIC are constantly recalibrated by VPC to keep the representations consistent.

into a filter-based representation. Two Alternatives are described. (i) Linear vision camera: Walsh-like filters are tuned to various patterns and spatial frequencies. (ii) 2d camera: A set of modified Gabor filters are tuned to different orientations and spatial frequencies. A log-polar retinotopic sampling grid is placed on the image and the filter set applied to each “retina” point.

At each time step, the agent takes four “snapshots”, one in each cardinal direction. For each orientation, the filter activities are stored in “view cells” (VCs).

An unsupervised growing network scheme is employed which recruits visual place cells (VPCs) when necessary. Synapses from VCs to VPCs are initialised and modified using a Hebbian-type learning rule. VPCs are suggested to occupy superficial lateral entorhinal cortex.

Vestibular input drives populations of head direction (HD) and path integration (PIC) cells. PIC is postulated to be located in superficial medial entorhinal cortex, whereas the HD system is distributed across ante- and laterodorsal thalamic nuclei, lateral mammillary nuclei and postsubiculum. Position information from VPC is used to recalibrate path integration. Similarly, the bearing angle to a salient landmark (a lamp) in conjunction with VPCs are used to recalibrate the head direction system.

PIC and VPC project to place cells in the hippocampus proper (PC). The same unsupervised growing network scheme and Hebbian learning is applied to PCs. Re-

alistic place fields are reported for both visual systems.

In this model, four real camera snapshots provide the visual stimulus for a place. However, the images must always be taken in the four cardinal directions. Even a small error in the estimation of the heading impairs the construction of a representation and the localisation of the agent. Furthermore, the sensory preprocessing assumes a foveal vision. The rat, in contrast, does not seem to have a fovea [124].

2.5 Models of multimodal integration

Most robotic systems make use of explicit probabilistic sensor models to represent measurements and their uncertainty [302, 160].

As representation of the environment, occupancy grids [77, 302, 160], stochastic maps [48] or biologically inspired hippocampal place cells [13, 14] are successfully used. The hippocampal place code differs from the other representations in that it doesn't contain the information of free vs. occupied space, but rather associates the agent's location directly with the corresponding sensorial measurement.

Some systems use sensor fusion techniques to enhance the quality of the spatial representation. Probabilistic methods [48, 302] as well as neural networks [160] are employed to integrate multimodal information into one coherent representation. Those techniques aim at reducing the uncertainty present in the probabilistic map. Most of these systems rely on representations and operations which are not easily mapped to neuronal properties or brain areas.

Biologically inspired spatial learning systems [96, 95, 14, 16, 97], on the other hand, mostly focus on a single allothetic modality and thus don't propose a solution to the integration of multiple allothetic sensory information.

Chapter 3

Animal navigation

In this chapter, we review experimental and modelling studies of animal navigation. Animals show various behaviours when solving navigational tasks. The selection of an appropriate strategy depends on the task's complexity and on the available sensory input. The basal ganglia seem to play an important role in navigation. In particular, they are suspected to be involved in the process of selecting an appropriate strategy. Therefore, it is necessary to briefly review the basal ganglia in order to understand the following sections. One factor which determines the complexity of a task is its memory requirements. An appropriate navigation strategy thus relies on a suitable memory system. For this reason, we present an overview of the properties of various types of memories before the introduction of a common taxonomy for navigation strategies. Then, we review experimental data on behavioural, neurophysiological, and lesion studies which aim at discerning those strategies and finding their neural substrates. The final section presents previous models of animal navigation relevant to our work.

3.1 The basal ganglia

The basal ganglia (BG) is a brain area comprised of several nuclei in the fore- and midbrain (figure 3.1). BG is known to be involved in motor-disorders such as Parkinson or Tourette. It is suggested that the BG resolves the conflict of competing motor programs by inhibition of all but the selected [116, 122, 175].

The striatum is the main input structure of BG. It is innervated by almost all cortical areas and the hippocampus. The substantia nigra pars reticulata (SNr) as

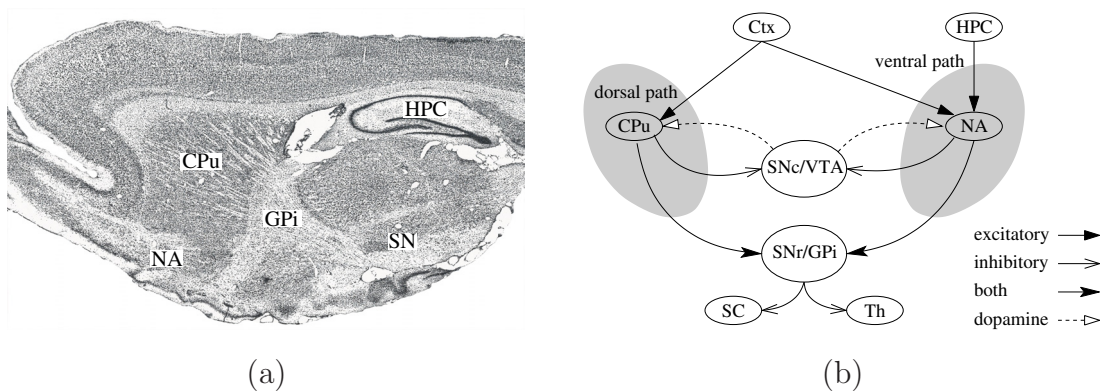


Figure 3.1: Basal ganglia. (a) para-sagittal view (adapted from [116]). (b) Coarse functional diagram of the basal ganglia: The dorsal path is thought to be involved in orienting behaviours whereas the ventral stream is important for more complex spatial navigation. Ctx: cortex. HPC: hippocampus. SNr/SNc substantia nigra pars reticulata/compacta. VTA: ventral tegmental area. GPi: Globus pallidus pars interna. NA: nucleus accumbens. CPu: caudate putamen. Th: Thalamus. SC: superior colliculus.

well as the internal part of the globus pallidus (GPi) are the main outputs of the BG. They project mainly to the thalamus (and thus indirectly to cortex) as well as to the superior colliculus (SC) [320, 116, 8, 175].

Dopaminergic neurons in the BG—namely the substantia nigra pars compacta (SNc) as well as the ventral tegmental area (VTA)—both project back to the striatum. They tend to synapse on the same spines than cortical and hippocampal afferents and seem to be involved in the processing of reward signals which could possibly modulate synaptic plasticity in the striatum [92, 251, 266, 122, 248]. In particular, they might signal errors in the prediction of rewards, a component which is at the core of reinforcement learning mechanisms [122, 248] (cf. chapter 4).

Two almost separate circuits can be distinguished: Caudate-putamen (CPu) receives its input from cerebral cortex and is the entry point to the dorsal path, whereas the nucleus accumbens (NA), which is innervated by hippocampus and cortical areas, is the first stage of the the ventral pathway. The paths are termed dorsal and ventral because CPu and NA occupy the dorsal and ventral areas of the striatum. Both pathways contain a direct inhibitory and an indirect excitatory pathway from input to output [3, 265, 267]. These two pathways could reflect two different navigation strategies which compete for the execution of motor programs [116, 175]. The dorsal path could, for instance, be responsible for taxon navigation while the ventral path supports a locale navigation strategy [211, 161, 210, 321] (cf. section 3.3).

3.2 Memory systems

Each of the navigation strategies discussed in sections 3.3.1–3.3.4 makes use of particular computations and operations relevant for the task. For instance, visual input must be encoded, stored and recalled in order to compare the current to an already experienced situation. These operations must be supported by some kind of *memory*.

Depending on the research area, different aspects of memory systems have been investigated and led to different, but overlapping, terminologies. Our taxonomy of navigation strategies, like many others, focuses on the underlying neural substrates. However, their memory requirements are of equal importance. For each strategy, we state what kind of memory is involved and what other type is not needed. It is therefore important to first define what we mean by “memory” and relate the different types to existing terminologies of memory. Thus we start with *our* definition of different types of memory before turning to navigation strategies. We distinguish three types of memory. For each of them, we’ll summarise behavioural or functional properties, theoretical considerations, and neural substrates and mechanisms.

Long-term memory (LTM): A long term memory system can store a very large amount of experiences. Those memories are relatively stable over years or even lifetimes. There are three types of LTMs: *Episodic* memory refers to a rich multi-sensory personal experience anchored in time and space, as for example the first date with your beloved. *Semantic* memory concerns facts of the world, as for example knowing that Bern is the capital of Switzerland. Episodic and semantic are also termed *declarative* memory. *Procedural*, or non-declarative memory refers to skills, like playing squash or riding a snowboard. The process of storing something in LTM is called *learning*.

The learning speed seems to vary between different brain areas. In hippocampus, for instance, it is extremely fast: One exposure is sufficient to permanently memorise the experience. This type of memory is also termed *reference memory* [202,203,129,236,226]. In cortex, on the other hand, learning is suggested to be slow [136,137]. This slowness might be important for generalisation: Unimportant and variable details are averaged out and only the underlying principles are learnt [328,26,327].

The neuronal correlate of LTM is a stable modification of the properties of neurons and, in particular, their synapses. It has been suggested that LTM is the effect of long-term potentiation and –depression (LTP/LTD) of synaptic efficacy [17,18,32]. Experiments show that if two connected neurons are stimulated such that presynaptic action potentials are repeatedly followed by postsynaptic spikes, the synapse can be potentiated. Conversely, the synapse’s efficacy is decreased if the postsynaptic neuron fires just before the presynaptic neuron.

This effect is termed *spike-timing-dependent plasticity* (STDP) [156,27,151] and may be a mechanism to form LTM. The temporal window for STDP is 20–40ms. Once a potentiation is induced, during a period of minutes or hours following LTP induction, the synapse undergoes a process of *consolidation* in which the modification is fixated [147]. Retention of such a potentiation is beyond hours or days [62,73,74]. In this thesis, we follow this interpretation of LTM and assume that it involves modifying synaptic efficacies.

Short-term memory (STM): Short term memory is a storage of recent events for short time spans. An example could be to remember whether or not you already added sugar to your coffee, or, what the contents of the last paragraph was. Some short-term memories may become part of long-term memory by a process called *consolidation*, but others just vanish. A variety of conditions, including normal aging, can diminish or destroy short-term memory, while leaving long-term memory intact.

It is not clear what neuronal mechanisms are responsible for STM. There is also no clear distinction between STM and working memory (see below). Some definitions group them together as one class. Here, however, we make a distinction based on theoretical considerations. We define STM as a gradually decaying trace of past events. The quality and strength of the trace is slowly fading away, and more recent events may overwrite this memory at any time. The timescale should be in the order of seconds or minutes. This memory may be related to LTP in that it could “bridge” the temporal gap until the synaptic modification is consolidated.

Potential neuronal mechanisms for our type of STM are slowly varying concentrations of chemical agents in the synapses. For instance, a burst of incoming spikes has been shown to facilitate synaptic transmission for short periods of time [303,67,337]. It has been suggested that the presynaptic neurotransmitter release probability is transiently increased, resulting in this short-term-facilitation of the synapse [299,250]. Postsynaptic concentrations of calcium [322,243,316] only decay slowly and might implement a short-term memory trace postsynaptically. Calmodulin-dependent protein kinase II [122,205] and neuromodulators such as dopamine [207] have been shown to modulate short-term, as well as long-term potentiation.

Working memory (WM): Working memory, like short term memory, is limited to a short period of time. The difference between WM and STM is that WMs are thought to be symbolic instances that can readily be manipulated. They contain everything that is important for immediate processes, such as thinking, calculating, etc. Classical experiments have shown that in humans, WM is limited to “seven plus minus two” items, irrespective of the type and complexity

of items used [172]. WM seems to be correlated with intelligence (as assessed by intelligence tests)

Theoretically, we can think of WM as “states”. Each item to be remembered is locked in such a state. In contrast to STM, we define WM to be a non-decaying memory. Its retention time is in the order of seconds or minutes. Working memory operations are quick: Information may be stored, recalled or overwritten in a switching rather than gradual manner.

WM may be implemented by sustained activity as proposed by discrete or continuous attractor network models [9, 10, 335, 227, 238, 272]. WM relies on LTP to tune the synaptic efficacies for attractor networks. Once tuned, no further modification of synapses is necessary for WM.

Clearly, these types of memory are not always distinguishable. Theoretically, it is mainly the timescale which separates them. Nevertheless, we think it useful to ask which type of memory is involved for each of the navigation strategies presented in the following sections. In particular, working memory seems to be a key distinguishing factor.

3.3 Navigation strategies in animals

There are several taxonomies for navigation strategies proposed in the literature [293, 198, 94, 297, 321]. We present here a mixture of these classifications, with the aim of simplicity and relevance for this thesis.

3.3.1 Homing by dead-reckoning

The ability of an animal to return to the starting point of a journey, eg. the nest location, without using external cues is often termed *path integration* or *dead reckoning* in the literature [22, 177, 313, 315, 80, 79, 314]. It implies that during the entire journey, a displacement and heading vector is continuously updated. In this thesis, however, we will call this ability *homing by dead reckoning* (HDR), because it explicitly suggests that not only the updating and storing of spatial information, but also its *use* is included in this term. We reserve the notion of *path integration* to the capability of storing spatial information and continuously adding small displacement information to this representation of space (see section 2.2). Figure 3.2 schematically illustrates (gray area) what we mean by HDR. It is important to note that the update of the spatial variable relies only on idiothetic (internal) cues, such as motor efferent copy, vestibular cues or proprioception—but not, e.g. optic flow, because it relies on the illumination of the environment.

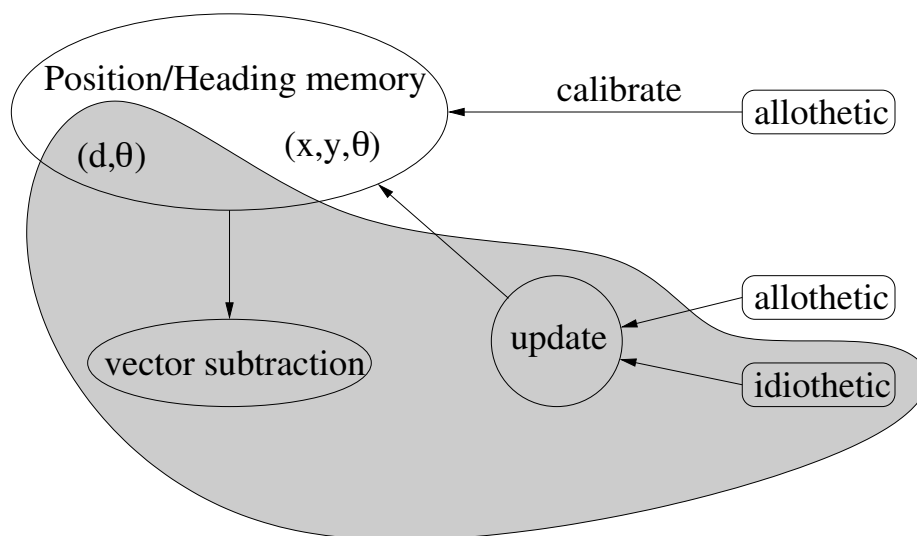


Figure 3.2: Homing by dead reckoning: The capability of returning to the initial location of a journey by only using idiothetic cues. This requires storing and updating two quantities: egocentric bearing θ of the target as well as its distance d .

The distance and bearing of the nest location continuously needs to be updated and stored. This operation requires working memory. A subset of the more general path integrator (section 2.2) might be used, where only idiothetic cues are considered for the update.

Transforming the idiothetic input into a distance and bearing update needs to be learnt. This transformation might be stored in a long term memory. It is, however, environment independent and needs to be learnt only once. After the transformation has been learnt, it is usable in all environments and tasks.

Cataglyphis desert ants may travel thousands of body lengths in a journey for food—and still return to the nest location in a straight line [313, 315, 314]. If the straight return path is blocked by newly inserted large barriers, the ant avoids the obstacle and, at the edge of the obstacle, directly head towards the nest [314]. They seem to use the sun as polarising orientation cue [313, 314]. Other invertebrates, e.g. spiders or bees are also HDR-capable.

More recently, homing capabilities are also studied in mammals. Gerbils [177], hamsters [78, 81, 80, 79, 260], rats [293, 319, 25] and many other species show HDR ability. The homing capability is usually tested in the dark, by luring the animal in a curvy path away from its nest, and then startle the animal as a motivation to quickly return to the “save” home. Depending on the experimental design, the animal only needs to keep track of the egocentric bearing of the nest. Indeed, it is

demonstrated in [260] that homing performance may increase significantly if the nest is at the periphery, rather than the centre of an enclosed arena.

When the animal is allowed a brief view of the environment at the end of the outward journey, the homing is more accurate. This and other similar experiments indicate that the path integrator can be recalibrated using visual information [81,79].

3.3.2 Taxon navigation

Taxon navigation is the strategy of orienting and moving towards a salient stimulus. For instance, finding a visible food source can be achieved by taxon navigation. This category includes the case where the food is hidden, but a distinct landmark is placed at its location. All the animal has to do is learn to detect the landmark, rotate until it is straight ahead, and move forward.

Such a mechanism is often termed *stimulus-response* navigation, because a proper action can be associated to each percept [293,198]. No working memory is required for this type of navigation. A long-term memory, however, is needed to learn and store the correct response to the possible signals. In this thesis, we may use the terms "stimulus-response", or "memoryless" interchangeably for a mechanism that doesn't involve working memory.

The following brain-areas seem to be involved in taxon navigation: Superior colliculus [270], caudate putamen [211,161,210,321] and posterior parietal cortex [11,271]

3.3.3 Praxic navigation

Some tasks can be solved by just executing a sequence of motor actions. For instance, if the relation between entry-point and goal location in an environment never changes between trials, the goal can be found without taking external cues into account. In contrast to HDR, where this motor program is trivial (namely a straight line), the term *praxic navigation* is used for complex motor responses.

Suppose that the complex motor response can be divided into a sequence of straight trajectories. Storing the order of the components may be seen as a long-term memory process. But in order to know the current position within the component (or the position within the entire sequence), the animal has to use path integration. Therefore, praxic navigation requires working memory.

There is evidence that posterior cingulate cortex [164,278,50,49,230] and caudate putamen [161,210] contribute to praxic navigation.

3.3.4 Locale navigation

While taxon navigation is sufficient to navigate towards a landmark, *locale navigation* enables the animal to reach any point in an environment. The animal combines many

spatial and temporal aspects of the environment or the task and stores them in the form of a cognitive map [293, 198].

Such a spatial representation allows the animal to localise itself in the environment and plan routes from one place to another. It therefore supports navigation to an invisible goal. As in taxon navigation, a long-term memory may store the associations of place and direction of the goal. This memory is often termed *reference memory* [226] in the context of locale navigation. Working memory is not required for this task.

However, this is only one aspect of locale navigation. The cognitive map needn't be purely spatial. More generally, a *context* is represented. This includes temporal relationships between events. Consider the following task: An animal starts at place *A* and in a first stage has to go left to reach place *B*. Then, it has to return to *A* and in a second stage go right to place *C*. The choice of going left or right at *A* thus depends on the stage of the task. This task can be solved by locale navigation and it requires working memory.

The locale system is the most complex of those described here. Not surprisingly, many brain areas contribute in building and using the cognitive map. The most relevant brain areas for locale navigation are: Hippocampus [197, 196, 163, 200, 189, 187, 165, 259, 135, 324, 199, 158, 139, 159, 195, 263, 169, 168], entorhinal cortex [23, 221, 179], subiculum [23, 190, 258, 255] and nucleus accumbens [76, 279, 320, 39].

3.4 Common navigation tasks

In this section, we present some frequently used experimental setups for spatial navigation tasks. We review some of the results which lead to the taxonomy of rodent navigation discussed in section 3.3.

3.4.1 Water maze

The water maze (figure 3.3(a)) was introduced by Morris in 1981 [183]. It is a cylindrical arena of 1–2m diameter, filled with water. Milk or paint is added to the water in order to make it opaque. A platform is located somewhere in the pool where the animal can escape from the water. If the platform is slightly submerged so that the animal can't see it, the setup is termed *hidden water maze*, in contrast to the *visible water maze*, where the platform sticks out of the water. Sometimes the position of the (maybe hidden) platform is indicated by a strong visual landmark directly above the platform; this is termed *cued water maze*.

Rats with hippocampal lesions are unable to find the platform in the hidden water maze task [184, 185]. Instead, they circle the arena at a fixed, but correct distance from the wall. Lesions of the fimbria fornix fibres, which connect the hippocampus to

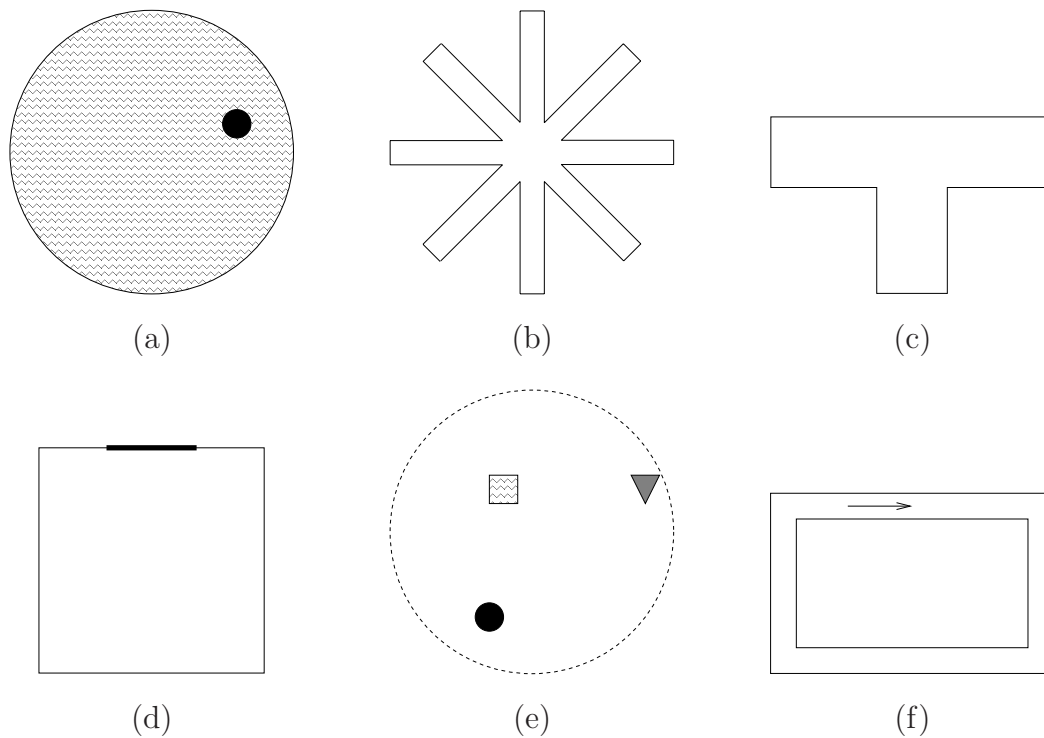


Figure 3.3: Frequently used environments for navigation experiments: (a) Water maze, (b) 8-arm maze, (c) T maze, (d) small enclosed environment with cue card, (e) open environment with local objects, (f) linear track

the nucleus accumbens, also impair the rats in the hidden, but not the visible water maze task [76, 209]. If, however, the rats always start from the same location, even fornix-lesioned animals can learn to find the hidden platform. These results suggest that the hippocampus is needed for locale, but not for praxic navigation.

Whereas hippocampal lesions leave the animal with some information (e.g. the distance to the wall), lesions to the subiculum, post-subiculum or anterior thalamic nuclei, which are all part of the head-direction system, produce severe deficits if the platform is hidden. If the platform is visible, however, performance remains normal [185, 279, 287]. Hence, the head-direction system seems to play a role in locale, but not in taxon navigation.

Learning to swim directly to the hidden platform takes the rat in the order of 20 trials. After this task has been learnt, relocating the platform results in much shorter training time of around 5 trials [88]. If the platform is moved every day, but left at a fixed position during all sessions of a day, the rat learns to escape to the platform after the first trial from day 6 onwards [269]. This suggests that some

knowledge or experience can be reused when the location of the platform changes. More precisely, the rat may build a spatial representation which is independent of the goal location.

Caudate-putamen lesions, in contrast, impair navigation in the cued or visible water maze task [209, 161]. If well-trained rats in the visible or hidden platform task are tested with a visible platform at a new location, hippocampal lesions make rats escape on the new platform whereas caudate lesioned rats swim to the location where the platform had been during training [161, 66]. When the hidden platform is moved from trial to trial, but always cued with a salient landmark, fornix¹ lesions don't decrease performance whereas caudate lesions do. Conversely, caudate lesions don't affect navigation to a location-stable hidden platform, whereas fornix lesions do [209]. This suggests that caudate putamen is involved in taxon, but not locale navigation. Conversely, the hippocampus and nucleus accumbens are important for locale, but not for taxon navigation.

3.4.2 Radial arm mazes

Several types of radial mazes, most notably the eight-arm maze shown in figure 3.3(b), plus-maze (basically a 4-arm radial maze) and the T-maze (figure 3.3(c)), are employed to investigate rodent navigation. Popular experiments in radial mazes include working memory versus long-term reference memory tasks².

In the non-rebaited 8-arm maze, reward is given in some arms, but the reward is not refilled after consumption. This task is also termed win-shift task. In most win-shift task, a delay is imposed to the rat after some arms are visited. This reduces the possibility that the animals use a praxic strategy to solve this task, i.e. just execute a stored motor program. On the average, rats visit more than seven different arms in their first eight choices. This is a sign of working memory, since the animals have to remember which arms they have recently visited [204]. Arms that are not baited in any of the training trials are not visited in probe trials [202]. The ability to avoid never baited arms does not require working memory, but it depends on long-term (reference) memory only.

If the arms are rotated during the imposed delay after the third arm visit, it is revealed that the rats seem to identify the remaining arms by the distal (extra-maze) cues [204]. Hippocampus lesioned rats, however, make working memory and reference memory errors, i.e. they reenter already visited arms and search the unbaited arm [129].

Rats can also be trained to honour local cues such as floor texture and ignore distant landmarks [202]. After rotating the arms, rats don't enter the cued unbaited

¹the fornix fibre bundle connects the hippocampus to the nucleus accumbens, see section 2.1.1

²reference memory in our terminology is a form of long term memory which stores information related to spatial contexts

arm even with hippocampal lesions. They do, however, make working memory errors (reenter already visited arms) [129]. These results show that an intact hippocampus is required for spatial working memory as well as for memorising a spatial context (reference memory), whereas for locally cued mazes that can be solved by taxon navigation, hippocampal lesions have no effect.

In a plus-maze, rats easily learn to go to a specific place (eg. the arm that points to the west) regardless of their starting arm. More learning is required to make them always turn in a certain direction independently of the starting arm [294]. Indeed, when rats are trained for eight days to go from the south arm to the west arm, they will also go to the west arm when starting from the north in a test trial on the ninth day. After another eight days of south to west training, however, they turn left (into the east arm) in a test trial from the north. Caudate lesioned animals, in contrast, always go to the west arm, while hippocampal lesions produce a left-turning behaviour [208, 210]. This nicely shows that hippocampal-dependent locale navigation is learnt first, but can be overruled by caudate-dependent praxic navigation after further training.

3.4.3 Unrestricted arenas

Small environments surrounded by walls (figure 3.3(d)) or open field environments like in figure 3.3(e) where the animal's movement is unrestricted have been used to study the properties of place cell firing as well as the behavioural response to various manipulations. Rectangular as well as circular environments of various sizes are employed. In some experiments, local objects are placed in the arena.

In a closed rectangular environment where all corners are distinguishable, rats are trained to run from any of eight starting positions to a specific corner to find food. When the rats are disoriented, the rats run to the opposite corner almost 50% of the time [51]. When the rats are not disoriented, however, they choose the right direction in more than 75% of the cases [155]. In other experiments, it has been shown that head direction cells do not follow external cues if the cue is unstable over trials, whereas it strongly influences head direction cell firing if it is [164, 262, 286, 139]. These findings suggest that the head-direction system is mainly influenced by idiothetic input but can be reset by a stable allothetic cue, but not by an unstable allothetic signal.

In a circular arena, place cells rotate with local landmarks only if they are pushed against the walls. If they are put in the interior of the arena, place fields do not follow landmark rotations [58]. However, in an open arena with an arrangement of distinguishable local landmarks, gerbils or rats can learn to find food relative to the landmarks even if the whole arrangement is moved in the environment and the animals start from a different location in each trial. They search at the correct distances and

bearings³ from each landmark if one or two of the landmarks are removed [56, 29]. Indeed, electrophysiological recordings show that hippocampal place fields can follow nearby local landmarks instead of distal cues if the landmark is relevant for finding a reward [108].

3.4.4 Linear tracks

A linear track is an environment where the animal can only move along one dimension. In a loop environment (figure 3.3(f)), for instance, the rat can only move in one direction. In other variants, the animal moves back and forth between two endpoints of a short straight track. Its movement is thus extremely restricted and controllable. For this reason, numerous experiments study place cell activity under those conditions.

If a rat shuttles between a movable start box and a fixed goal location, a conflict situation between the local view and internally generated movement signals can be created. When the start box is moved closer to the goal location, the sequence of place cell activity is accelerated to compensate for the shorter travel distance [107]. In the dark, however, the place cells fire with respect to the moved starting box [106]. This suggests that under normal lighting conditions, visual cues and path integration continuously interact to keep a consistent representation whereas in the dark, place cells are driven by path integration [106]. This is consistent with experiments where a place representation established in the dark persists when the light is turned on [158].

Place fields are directional in a linear track: A cell active on the outward journey generally isn't active on the way back [107, 166, 106]. It has been suggested that a different reference frame is used for the two paths [296, 228, 238].

Well trained rats show asymmetric place fields in linear tracks. The field enlarges against the direction of movement. This shift results in a representation that predicts future locations [188, 167, 166, 169].

3.5 Previous models of rodent locale navigation

Several models have been proposed in the literature to emulate the reviewed animal navigation strategies. In this chapter, we present a brief overview of models which are related to our proposal (chapter 8). They all concern rodent locale navigation, for it is this navigation strategy which is also addressed by our model.

3.5.1 Burgess *et al.* (1994)

Here we describe how Burgess *et al.* [41] use the spatial representation summarised in section 2.4.2 for navigation. The place cells in the subiculum (SUB) project to a

³Here, the bearing is defined with respect to the direction given by a vector between landmarks

population of eight goal cells (GC). A head direction (HD) system which also contains eight cells with fixed one-to-one connections to GCs attributes a direction to each GC. These connections are, however, gated by a reward signal, such that goal cells are only activated if a HD cell is active *and* a reward is delivered.

Every time the agent reaches a rewarding location, it looks in all eight directions. In each of them, a reward is delivered at the late phase of the θ cycle. Hebbian-type learning is then applied to the binary synapses from SUB to GC. Place cells firing at the late phase of θ tend to have receptive fields located *ahead* of the present position. The goal cell coding “north”, for instance, is thus contacted by subicular cells located to the north of the goal. As the receptive fields of subicular place cells are very large, the fields of GCs are also large. This allows the system to estimate its bearing and distance with respect to the goal location from almost any place of the environment after only one trial.

This model postulates goal cells in the subiculum which are sufficient locale navigation. This is in contradiction to experimental data where fornix lesions impair rats in the hidden water maze [76, 279, 209]. Furthermore, the goal cells in this model create a *global* basin of attraction towards the goal. Local information such as obstacles are not taken into account. Finally, this learning mechanism also suffers a “distal reward” problem because only those place cells whose fields contain the goal may learn place to action associations.

3.5.2 Brown and Sharp (1995)

The model by Brown and Sharp [39] (figure 3.4) is based on a simplified version of the spatial representation outlined in section 2.4.1. A population of “motor” cells in nucleus accumbens receives spatial information from the hippocampal place cells. Together with a head direction system, the model learns to perform movement commands which lead to a rewarding location. Nucleus accumbens cells are separated in two clusters of 60 motor neurons and the same amount of inhibitory interneurons. Each of the 60 place cells is connected to a unique interneuron of each cluster. Each interneuron in turn projects to a unique subset of 59 motor cells, which form the output of the navigation system. These synapses are fixed.

Every time the agent encounters a rewarding position in the environment, head direction cells modify their synaptic weights to the active motor unit using a reward-modulated Hebbian-type learning rule. An exponentially decaying memory trace of pre- and postsynaptic coactivation enables the agent to propagate reward information along its trajectory. Most recently active synapses are strengthened most, while the change in synaptic efficacy for previously active synapses is smaller.

In test trials, the two active motor units compete for action selection. A left turn is performed if the “left” motor unit is more active than the right and vice versa. The model is able to solve the hidden water maze task.

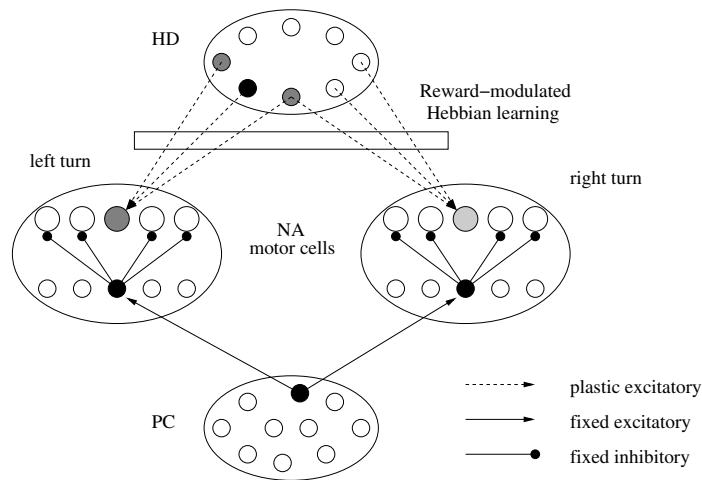


Figure 3.4: Navigation architecture proposed by Brown and Sharp (1995): Nucleus accumbens (NA) consists of two populations; one for right turns, the other for left turns. Each place cell (PC) has a fixed connection to an inhibitory neuron in each NA cluster. Each interneuron is connected to a unique subset of excitatory output cells, containing all but one “motor” cells. Head direction (HD) cells learn to select left or right turn motor cells by reinforcement-modulated Hebbian learning.

This model updates its place-action mapping only when a rewarding location is encountered. Places that are far away from the goal are only associated to actions using a global temporal activity trace. The existence of such long memory traces in animals is still an open question. Temporal difference learning rules 4.1 don’t suffer this limitation.

3.5.3 Abbott and coworkers (1996, 1997)

The work by Abbott and coworkers [34, 101] suggests that hippocampal region CA3 is the neural substrate for navigation maps. Learning on the recurrent CA3 connections using spike timing dependent plasticity results in a shift of receptive fields towards the goal location. Initially, CA3 place cells have perfectly Gaussian receptive fields with high overlap. Training consists of repeated trials, ending when the agent reaches the goal location. This procedure introduces an inhomogeneity with respect to the experienced trajectories: No path can lead from the target to other places because trials end at the goal location.

Navigation maps to multiple targets can be represented simultaneously: Place cell activity is modulated by the distance to the target location. After learning, navigation maps to multiple targets can be recalled. Routes to novel target location can also be

generated by superposition of the learned maps.

The learning rule used in this model produces a shift of the place fields in the direction opposite to the goal location. In order to use this shift for navigation, the original place field centres are accessed explicitly to calculate the agent’s next movement, which is a biologically implausible operation. Furthermore, this model of locale navigation is entirely concentrated in the hippocampus. It is not consistent with the impairments reported in the hidden water maze following fornix or nucleus accumbens lesions [76, 279, 209].

3.5.4 Gaussier and coworkers (1998, 2000, 2002)

The navigation model by Gaussier and coworkers [96, 95, 97] is based on the spatial representation reviewed in section 2.4.4. A transition prediction network based on place cells “proposes” candidate future places (see figure 3.5). Competition within the transitions recognition layer selects the most active transition (e.g. node BD, leading from place B to place D). This competition is biased by a goal planning layer. Motor actions are associated to transitions using Hebbian-type learning.

The goal-planning layer contains units which code for the same transitions as the recognition layer. Here, however, the transition units are interconnected with constant synaptic weights $w_{ij} < 1$. Whenever a motivation input activates a node (e.g. DG1, the transition from place D to goal G1), this activity $A_0 = 1$ propagates back to all other transitions. Activities $A_i = \max_j(w_{ij}A_j)$ are calculated iteratively until the network settles in a stable state. Once stabilised, node activities are set according to their distance to the goal location. These activities bias the competition in the transition recognition layer, such that transitions which lead to the goal on the shortest path are favoured.

According to this model, locale navigation is implemented in the hippocampus, which is in contradiction with experimental data, suggesting that the fornix projection to nucleus accumbens is necessary to solve the hidden water maze task [76, 279, 209]. The goal planning layer operates on the symbolic place transition nodes. The shortest path to the goal in this layer needn’t correspond to the shortest path in the real environment.

3.5.5 Foster *et al.* (2000)

The model proposed by Foster *et al.* [88] (figure 3.6) is based on an actor-critic architecture for temporal-difference (TD) reinforcement learning (see section 4.1). A layer of place cells (PCs) with perfectly tuned Gaussian receptive fields provides the navigation system with the agent’s position within its environment.

A “critic” neuron c receives input from each PC i . Its firing rate $r_c = \sum_i w_{ci}r_i$ represents the estimated “value” of the current agent position. The critic also outputs

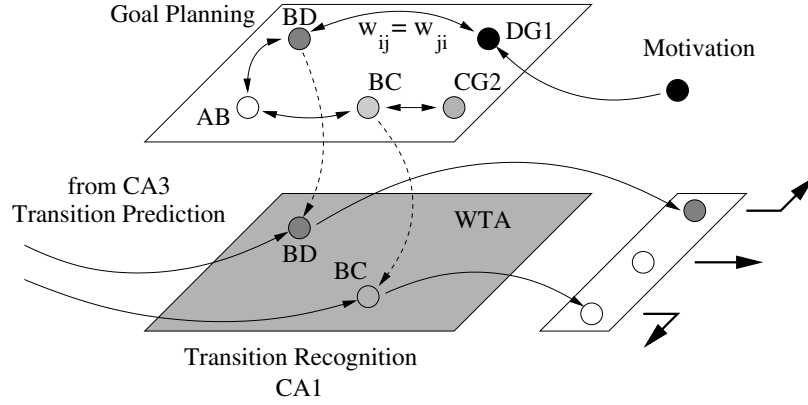


Figure 3.5: Navigation architecture proposed by Gaussier and coworkers: Transition predictions are made in CA3 (not shown) and propagated to the CA1 network. There, winner-take-all (WTA) competition selects the most active transition and executes its attributed motor command. Competition is biased by transition units in the goal planning system which, according to motivation, back-propagates transitions to goal locations (e.g. DG1) through the network.

a reinforcement signal δ in the form of a temporal difference of current and previous activities (i.e. position value prediction error). This signal is used to improve value estimation by modifying afferent connection weights w_{ci} towards $\delta \cdot r_i$ and thus reduce the error δ .

An actor network consisting of eight neurons is responsible for selecting actions. Each cell a codes for a direction of movement and receives afferent connections of strengths w_{ai} from each PC i . Actions are selected stochastically, but cells with a high firing rate $r_a = \sum_i w_{ai} r_i$ are favoured over cells with low firing rates. The weights w_{ai} are modified using Hebbian-type learning, modulated by the critic’s error signal δ .

Using this mechanism, navigation to a stable hidden goal location can be learnt in about ten trials. When a learnt target is relocated, however, relearning the new location takes longer due to interference with the previous goal location. To overcome this difficulty, a coordinate system (CS) is added to the model: The CS consists of five neurons: two of them represent the current position of the agent (x, y) , the next two code for a goal location (x_g, y_g) to be learnt, and the remaining neuron is an action neuron a_{coord} . The mapping of place cell activity to coordinates (x, y) is learnt using TD-learning. Each time the reward is found (i.e. at the end of each trial), the current estimated coordinates (x, y) are copied into the goal memory (x_g, y_g) . The “abstract” action a_{coord} neuron competes with all other actor cells for action selection. Its activity, however, doesn’t depend on the agent location, but on how

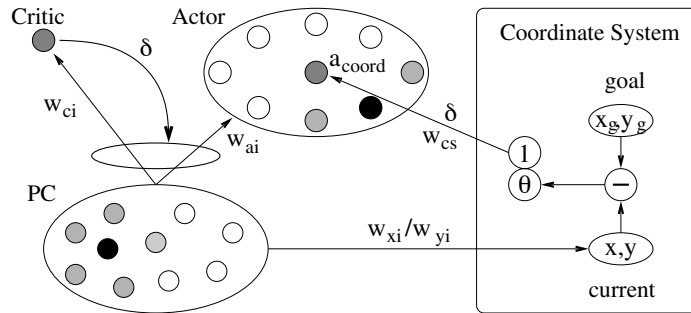


Figure 3.6: Navigation architecture proposed by Foster *et al.*: An actor-critic architecture is employed to learn a navigation map. Places are coded by a place cell population (PCs). The actor selects actions whereas the critic estimates the quality of these actions. The critic also generates a reinforcement signal δ which guides learning of better actions and quality estimations. A task-independent coordinate system (CS) is learnt from place cell activity. Once learnt, CS can store the goal location and propose goal-oriented actions a_{coord} by vector subtraction.

well the coordinate system is tuned. Modification of its weight w_{cs} is similar to other actor neurons, i.e. modulated by δ . Whenever a_{coord} is selected, the direction of movement θ is given by vector subtraction of goal–and current coordinates.

During learning, the agent’s movement is restricted to eight predefined headings. Furthermore, learning does not generalise to neighbouring directions. The coordinate-system module, once adapted, creates a global basin of attraction. Local information such as obstacles are then completely ignored. The direction of the next movement is algorithmically calculated by explicitly accessing the coordinates of the goal location, which is not biologically plausible.

3.5.6 Arleo *et al.* (2000, 2001)

Arleo *et al.* [14, 16] propose a locale navigation system using reinforcement learning [280]. It is based on the spatial representation outlined in section 2.4.5. Each place cell projects to four “action cells”, coding for north, south, east and west respectively (figure 3.7). The synaptic strengths represent a “navigation map” and are modified using a reward-based learning method: Suppose that the agent is at place A . The activity $r_i(s) = \sum_j w_{ij} r_j$ of action cell i depends on place cell activities r_j and synaptic weights w_{ij} . It estimates the “value” of action $a_i(s)$. When taking action a_i , the agent reaches place B and gains access to action value estimates of the new place. The weights w_{ij} are adapted to correct for bad estimates at the previous place A . In particular, Watkins $Q(\lambda)$ (see section 4.1) is used. This model can learn to navigate from any place in the environment to a hidden goal location. During learning, only

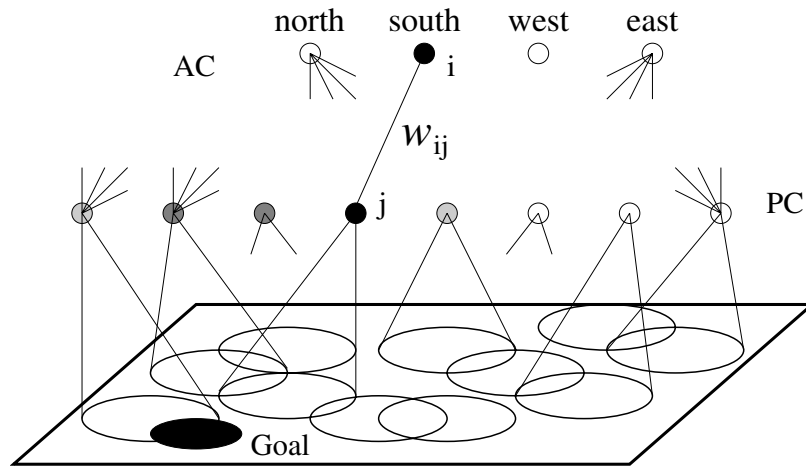


Figure 3.7: Navigation architecture proposed by Arleo *et al.*: Place cells (PCs) drive a set of four action cells (ACs). Connections strengths w_{ij} are modified by a reinforcement learning rule.

one of the four actions can be taken (winner-take-all mechanism). Once the navigation map is learnt, however, generalisation to continuous actions can be achieved by interpolating between the discrete actions.

A “reward expectation cell” (REC) learns to associate place cell activity with the goal location G_0 using a Hebbian-type learning rule. REC is highly active at G_0 before learning. When the location of the goal has been learnt, the reward is expected and the cell is silent at G_0 . If the reward is relocated to place G_1 , however, REC is strongly depressed at G_0 . This depression triggers the relearning of the navigation map. Goal location G_0 will then be forgotten.

During learning, the agent’s movement is restricted to four predefined headings. Because there is no generalisation mechanism in action space, the learning time would increase if more headings were allowed.

Chapter 4

Reinforcement learning

In this chapter, we review a learning method termed “reinforcement learning” (RL). Our model of rodent locale navigation (chapter 8) employs a form of RL, the bases of which are given here. Instead of giving a complete review, we focus on some aspects and problems of RL that are important for our model.

Reinforcement learning (RL) [280], sometimes also called trial-and-error learning aims at achieving a goal by continuous interaction between an *agent* and an *environment*. The agent perceives some, but not all properties of the environment. Based on this information, it takes an *action*. The environment responds to the agent’s action by moving it to a different situation. The environment sometimes responds with a special *reward* signal to tell the agent whether or not its actions are good. RL is a class of methods which aims at optimising the agent’s behaviour in order to maximise the reward.

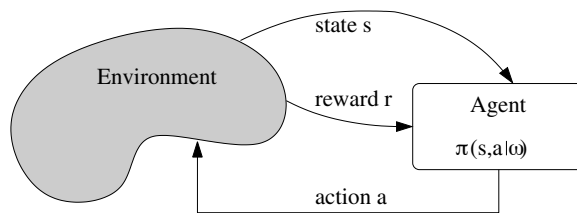


Figure 4.1: Agent-environment interaction in reinforcement learning (adapted from [280]). The task is to maximise the return by tuning the parameters ω of the policy $\pi(s, a | \omega)$.

States, actions and policies

At each discrete¹ time step t , the agent perceives the environment via the *state* $s_t \in \mathcal{S}$ where \mathcal{S} is the set of all possible states. It then selects and executes an *action* $a_t \in \mathcal{A}(s_t)$ where $\mathcal{A}(s_t)$ denotes all possible actions available in state s_t . At the beginning of the next time step, it receives a *reward* $r_{t+1} \in \mathcal{R}$ and ends up in state s_{t+1} . Note that the reward is a scalar, while the states and actions are vectors. The agent has to decide which action to execute at each time step, given the perceived state. This can be described as a mapping of the states to actions. This mapping is called a *policy* $\pi(s, a \mid \omega)$ where ω is a vector of tunable parameters:

$$\pi(s, a \mid \omega) = \Pr\{a_t = a \mid s_t = s, \omega_t = \omega\} \quad (4.1)$$

It specifies the probability of selecting action a when in state s , given the parameters ω . A policy which always selects the action which seems best according to the current information is called a *greedy policy*. It tries to maximise rewards according to the current knowledge even if this would prevent access to more abundant rewards in the future. An ϵ -*greedy policy* selects the greedy action most of the time, but with small probability ϵ randomly chooses between other available actions. This general scheme of reinforcement learning is depicted in figure 4.1.

Rewards and returns

We mentioned earlier that the objective of learning was to maximise the reward. Here we focus on tasks that have a terminal state, which means that at some time step t_T the *trial* or *episode* terminates. The *expected return* $R(s_t)$ is then defined as the total reward from time t up to the end of the trial:

$$R(s_t) = r_{t+1} + r_{t+2} + r_{t+3} + \dots + r_T = \sum_{k=t+1}^T r_k \quad (4.2)$$

If the task only provides a reward at the terminal state, but several paths with different lengths lead to that “goal state”, $R(s_t)$ would equal r_T for all states s_t . Thus $R(s_t)$ of equation 4.2 would not be very informative. One way to solve this problem is to give a negative at each time step. Another approach is to give more importance to immediate with respect to future rewards. Equation 4.3 takes this into account and defines the *discounted expected return*:

$$R(s_t) = r_{t+1} + \gamma r_{t+2} + \gamma^2 r_{t+3} + \dots + \gamma^{(T-t-1)} r_T = \sum_{k=0}^{T-t-1} \gamma^k r_{t+1+k} \quad (4.3)$$

¹We only consider the time-discrete case here, but generalisations to continuous time have been proposed [68, 69]

where $\gamma \in [0, 1]$ is termed *discount rate*. It determines how important a future return is if it were received at the present time. The higher γ the more far-sighted is the agent. If $\gamma = 0$, future rewards are worth nothing. For the remainder of this thesis, we always refer to the discounted version when speaking of return. The RL task now consists in tuning the parameters ω of policy $\pi(s, a|\omega)$ such as to maximise the discounted expected return.

Markov property

The agent’s state vector is a summary of what the agent knows about the environment at the present time. It usually consists at least of preprocessed immediate sensory input. But it may also be constructed by taking a sequence of sensory inputs into account, ie. contain a form of working memory. From the state variable, the agent can then calculate the probabilities of future states and rewards for potential actions, as formulated in equation 4.4:

$$\Pr\{s_{t+1} = s', r_{t+1} = r' | s_t, a_t, r_t, s_{t-1}, a_{t-1}, r_{t-1}, \dots, r_1, s_0, a_0\} \quad (4.4)$$

This “prediction” of how the environment responds takes the whole sequence of states, actions and rewards into account. If, however, the state itself includes all relevant information, the agent is said to have a *Markov state representation*. It allows the prediction of future states without knowing how the current state came about:

$$\Pr\{s_{t+1} = s', r_{t+1} = r' | s_t, a_t\} \quad (4.5)$$

In other words, the Markov property means that equation 4.4 equals equation 4.5. State transitions are memoryless: Once a new state is reached, the sequence of all previous states is forgotten. This does not mean, however, that the state vector itself may not contain a memory of the past. Whether or not a representation is Markov depends also on the task and on the environment’s dynamics.

Value functions and Bellman equation

In order to plan an appropriate action, the agent has to estimate how good it is to be in some state, and how much an action is worth in that state. These expected returns are called *value functions*. They depend on what actions the agent will take in the future, ie. on the agent’s policy π . The *state value function* $V^\pi(s)$ defined in equation 4.6 tells the agent how much return to expect when in state s :

$$V^\pi(s) = E^\pi\{R_t | s_t = s\} \quad (4.6)$$

where $E^\pi\{.\}$ denotes the expected value when following policy π . Similarly, an *action value function* $Q^\pi(s, a)$ is defined for taking action a at state s , following policy π thereafter:

$$Q^\pi(s, a) = E^\pi\{R_t | s_t = s, a_t = a\} \quad (4.7)$$

Probably the most important property of value functions is that they follow a recursive relationship to their previous or successor value. For policy π , state s and action a , action values satisfy the following equality relation:

$$Q^\pi(s, a) = \sum_{s'} \mathcal{P}_{s \rightarrow s'}^a \left[\mathcal{R}_{s \rightarrow s'}^a + \gamma \cdot \underbrace{\sum_{a'} \pi(s', a') Q^\pi(s', a')}_{V^\pi(s')} \right] \quad (4.8)$$

where $\mathcal{P}_{s \rightarrow s'}^a$ denotes the probability that state s' is reached when taking action a in state s . $\mathcal{R}_{s \rightarrow s'}^a$ is the expected value for the reward received when action a leads from state s to s' . Equation 4.8 is called the Bellman equation for Q^π . The state value function satisfies a similar Bellman equation. These recursive relationships are important because they can propagate a known (or well estimated) value back to a previous state (or state-action pair).

Value functions can serve as quality measure for policies: if a policy π yields higher expected return than another policy π' for all states, π is a better policy than π' . More formally: $\pi > \pi'$ if $V^\pi(s) > V^{\pi'}(s) \quad \forall s \in \mathcal{S}$. Policies that are better or equal than all other policies are termed *optimal policies* π^* . They share the same optimal state and action value functions $V^*(s)$ and $Q^*(s)$ respectively:

$$\begin{aligned} V^*(s) &= \max_{\pi} V^\pi(s) \quad \forall s \in \mathcal{S} \\ Q^*(s) &= \max_{\pi} Q^\pi(s) \quad \forall s \in \mathcal{S} \end{aligned} \quad (4.9)$$

The goal of reinforcement learning is to estimate the optimal policy by estimating the optimal value functions. The Bellman equation forms the basis of many approaches for approximating these functions. In the following, we will focus on one such method: Temporal difference-learning.

4.1 Temporal difference learning

One of the most important methods to estimate the value functions is *temporal difference* (TD) learning. It learns directly from experience and iteratively updates estimates based on other learnt estimates without waiting for the end of the trial. Several variants of TD-learning have been proposed.

Sarsa and TD(0)

Sarsa is a TD learning method that estimates the action value function Q from experience.

At every time step, the estimation $Q(s_t, a_t)$ is updated according to equation 4.10. η is the step size or learning rate of the system and δ_t is the *reward prediction error* or TD-error (equation 4.11). It corresponds to the difference between the actual reward r_{t+1} and the predicted reward $Q(s_t, a_t) - \gamma Q(s_{t+1}, a_{t+1}^\pi)$.

$$Q(s_t, a_t) = Q(s_t, a_t) + \Delta Q(s_t, a_t) = Q(s_t, a_t) + \eta \cdot \delta_t \quad (4.10)$$

$$\delta_t = [r_{t+1} + \gamma Q(s_{t+1}, a_{t+1}^\pi) - Q(s_t, a_t)] \quad (4.11)$$

where $Q(s_{t+1}, a_{t+1}^\pi)$ equals zero if s_{t+1} is a terminal state.

Sarsa is an *on-policy* method because it updates Q^π by actually following policy π . This is manifested in equation 4.11 in the term $Q(s_{t+1}, a_{t+1}^\pi)$ where a_{t+1}^π corresponds to the action selected by policy π . For small learning rates η , Sarsa converges to an optimal policy π^* and action value function Q^* if all state-action pairs are visited an infinite number of times and if the policy converges to the greedy policy [261]. TD(0) is like Sarsa except that the state values are estimated. TD(0) converges under the same conditions as Sarsa [282, 63, 128].

Q-learning

Q-learning, in contrast to Sarsa is an *off-policy* algorithm. It optimises action values by using the *best*, rather than the *selected* action (equation 4.12) for the update (equation 4.10):

$$\delta_t = \left[r_{t+1} + \gamma \max_a Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t) \right] \quad (4.12)$$

Q-learning directly estimates Q^* while following an arbitrary policy. Like Sarsa, it also converges if all state-action pairs are tried indefinitely. There is, however, no need for the action selection to converge to the greedy policy [311, 312, 128].

Actor-critic architectures

Sarsa and Q-learning both estimate the action value function only. Actor-critic systems separate the estimation of the state value function and the policy optimisation. The critic receives the reward feedback from the environment and estimates the state value function. The actor adjusts its policy to maximise the return. Learning in both actor and critic is driven by a TD error signal generated by the critic. This architecture is shown in figure 4.2. It has been suggested that the brain implements an actor-critic architecture [122].

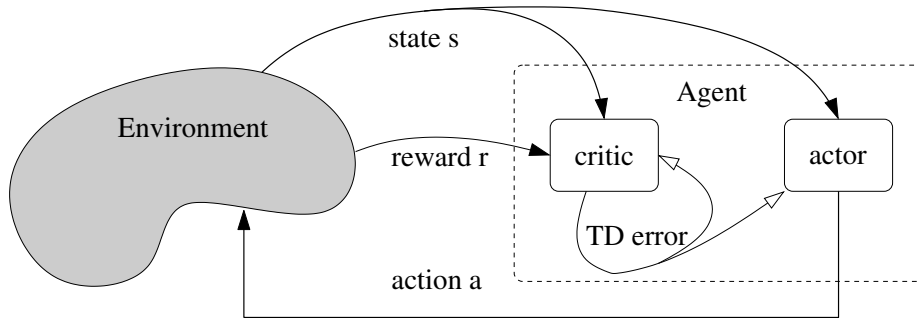


Figure 4.2: Actor-critic architecture. The critic generates the TD error signal that drives learning in both actor and critic (adapted from [280]).

4.2 Eligibility traces

So far the value functions have been calculated on the basis of a neighbouring state or state-action pair only. *Eligibility traces* (ETs) extend this idea to benefit from estimates that lie further away in time. ETs can be combined with almost any TD learning variant. An ET is a memory of previously occurred states or state-action pairs. The update of the value function estimates can then be done for all states and actions *eligible* for learning. For Q-learning, an eligibility trace $e_t(s, a)$ at time t can be defined as follows:

$$e_t(s, a) = \begin{cases} \gamma\lambda e_{t-1}(s, a) + 1 & \text{if } (s, a) = (s_t, a_t) \\ \gamma\lambda e_{t-1}(s, a) & \text{else} \end{cases} \quad (4.13)$$

For undiscounted returns ($\gamma = 1.0$), this trace decays exponentially with a *trace decay factor* of λ . When the future rewards are discounted, however, the ET decays at least with the discount rate. With eligibility traces, the update $\Delta Q(s_t, a_t)$ of equation 4.10 extends to:

$$\Delta Q(s_t, a_t) = \eta \cdot \delta_t \cdot e_t(s, a) \quad (4.14)$$

Equations 4.13 and 4.14 make all previously visited state-action pairs eligible for learning. Most recent actions get more “credit” for the current estimate of expected return and their values are modified to a greater extent than for decisions taken far in the past. TD(λ) is a natural extension of TD(0) in that it uses an eligibility trace. It has been shown that TD(λ) converges to the optimal policy under constraints similar to TD(0) [63, 64, 128]. For Q-learning, eligibility traces are valid only until a non-greedy action is taken [311]. This considerably reduces the benefit of ETs for Q-learning. Other variants of off-policy TD methods seem to work well in practise but their convergence is still an open question [215], but see also [219].

4.3 Continuous spaces and generalisation

One of the problems in reinforcement learning is that the learning speed highly depends on the dimensionalities of the state and action spaces. A related issue is the case of continuous spaces. Both problems can be solved by using function approximation for the state and action values.

The key idea is that updating an estimate for a specific state also affects the estimates of similar states.

This idea thus applies to the problem of continuous states/actions as well as to generalisation. In special cases, convergence has been proved, whereas other cases are known to diverge [105, 20, 298, 281, 219, 218]. In our model (chapter 8), we implicitly use function approximation to generalise in continuous state and action spaces.

4.4 Relation to animal learning

There is increasing evidence that dopamine is involved in reward-related learning. Dopaminergic neurons (DNs) have been found in the substantia nigra pars compacta (SNc) and in the ventral tegmental area (VTA). They project to the dorsal striatum (caudate putamen), the ventral striatum (nucleus accumbens) and most parts of neocortex [245]. DNs show regular or tonic firing patterns, as well as transient or phasic firing activity. DNs phasic activity is related to primary rewards [121], predicted reward [244, 12, 247], as well as novel stimuli which could be implicated in attention shifting [225, 245], but not to aversive stimuli [245].

In a conditioning learning task, DNs are first active at to onset of Unconditioned reward stimulus (US). During learning, this activity shifts from the US to the conditioned stimulus (CS) which predicts the US (reward) irrespective of the sensory modality of the used stimuli [248, 121]. If a fully learnt CS is shown in combination with a new stimulus, DN activity remains at the onset of the CS and behavioural learning of the association of the new stimulus with reward fails. This is called “blocking” and suggests that a fully predicted reward inhibits learning of a new stimulus-reward association and that the DN activity plays the role of a reward predictor [304, 245].

More specifically, phasic dopamine release could be the biological implementation of the TD-error of reinforcement learning [122, 248, 275, 276, 277, 70] and code for the difference of actual minus expected reward.

In the striatum, dopaminergic neurons tend to synapse on the same spines than afferent axons from cortical areas. It is therefore possible that DNs influence or modulate either synaptic transmission or synaptic plasticity of cortico-striatal connections [92, 251, 266]. There are at least two possibilities how DNs can influence learning. First, dopamine has been shown to focalise cortico-striatal transmission, which only allows the strongest signals to pass [207, 244, 225, 245]. This can be seen

as a gating signal which only enables the most important inputs to reach the post-synaptic neuron. If a behavioural response is required after the CS in order to receive the reward, the shift from US to CS in the dopaminergic neurons could serve as an advance information which results in an attentional focus and more precise learning of the response [245]. Alternatively, phasic dopamine release might directly influence plasticity in the target area. In striatum, prefrontal cortex and hippocampus, dopamine agonists enhance synaptic potentiation whereas antagonists impair potentiation [206, 207, 109, 138]. Recent experiments have electrically stimulated dopamine neurons in substantia nigra to simulate a reward prediction error. This resulted in potentiation of cortico-striatal [323, 231] projections. Prefrontal and auditory cortex also show potentiation if a phasic dopamine release is simulated [21, 33].

Other neuromodulators have been suspected to correspond to reinforcement learning variables. Serotonin may regulate the discount rate of future rewards (equation 4.3). Indeed, rats with depleted serotonin levels tend to impulsively favour small immediate over larger, but delayed rewards [69, 181, 70]

Noradrenaline seems to be involved in the control of arousal and relaxation and is suspected to govern the exploration-exploitation tradeoff (parameter of the policy π in 4.1) [69, 70]. Activity of noradrenergic neurons is correlated with the accuracy of action selection, especially in urgent situations [19, 301]

Acetylcholine may regulate the speed of learning (η in equation 4.10) [69, 70]. In hippocampus, striatum and cortex, acetylcholine modulates synaptic plasticity [224, 212].

Chapter 5

Sensory input

In this chapter, the sensory input to our model of spatial learning presented in the remainder of this thesis is described.

Perceptual models for spatial learning largely fall in two categories based on how the sensory input is generated: (i) Relatively simple mathematical models [254, 41, 296, 228, 229] and (ii) real-world input [96, 13, 14, 95, 16, 97]. A great advantage of mathematical models of input data is that there is no limitation as to what aspects of the world-agent relation is perceivable by the agent. Furthermore, perfect control can be exerted over the noise level in the input.

On the other hand, this freedom can have disadvantages as well: One problem is that the artificial agent might be allowed to access variables that are not perceivable by an animal. Or, conversely, the agent might not receive a relevant stimulus necessary to produce an essential property of cognitive maps. Secondly, a major drawback of mathematical models lies in the fact that it is impossible to recreate a realistic setup which is directly comparable to animal studies. Finally, applications of the model other than understanding brain functions, for example a commercial robotic application, only makes sense if the system can operate in a real environment. For these reasons, we choose to simulate our model in realistic environments and propose a biologically plausible computational mechanism to process sensory input.

First, the environments we used for testing the model are presented. Then, it is shown how the visual input acquired by the agent is processed. Finally, a description of how we emulate tactile and vestibular sensory input is given.

5.1 Test environments

The experiments of chapter 6 are run in four different setups. One of them includes a real Khepera¹ mobile robot [182,86,85], and the remaining three are simulations of the Khepera robot. All of them emulate a rat in an experimental arena. Experiments in chapter 7 are only run in the real robot setup, but not in the simulated environment. We use the term “agent” when speaking of similarities between the robot, simulated robot or sometimes even the rat.

The simulated robot has the advantage that it is more reliable and easier to handle. For instance, it takes no time to move the simulated robot, but the speed of the real robot is of course limited. On the other hand, a simulation, even when carefully designed, never captures all aspects of the real world. It is therefore useful to test the principles on the real robot, and supplement the results using simulation.

Before discussing the various setups in detail, some common features are outlined. The Khepera is equipped with a CCD camera with adaptive gain control, eight proximity sensors and two odometers—one for each wheel (cf. figure 5.1 (a)). Vision is the most important input to our model, but we also use the odometer and the proximity sensors. The simulator mimics the real Khepera robot by emulating its sensory input and performing similar movements. The rat’s field of view spans 270°–320° and depends on the rat’s head angle [1,124]. In our system, this view of the world is projected onto a cylindrical screen around the agent and transformed into a digital image of 800×316 pixels with 256 grey-values per pixel. For reasons of simplicity, the agent’s movement is decomposed into two primitives: (i) In-place rotation and (ii) forward movement. At each time step, the agent first rotates to the desired heading and then moves forward. After this movement, a sensory “snapshot” of the environment is acquired.

“Office”: The real robot setup

A Khepera miniature mobile robot is placed in an 80cm×60cm arena on a table in a normal office. The arena is surrounded by borders of 3cm height. The Khepera’s camera has a view-angle of approximately 60° (see figure 5.1). Four pictures in directions separated by 60° are merged into a single image \mathcal{I} of approximately $\theta = 240^\circ$ by rotating the robot in-place. Note that this rotation is only performed in order to acquire a panoramic view. In order to achieve an approximation of a cylindrical projection, the camera-lens distortion is first compensated for. Then, the image, which is recorded by a flat CCD sensor, is transformed such as to simulate a projection onto a cylinder. An example view is shown in figure 5.2. For the experiments in

¹The Khepera mobile robot manufactured by K-Team (<http://k-team.com/>) is a modular platform popular for research and education.

chapter 7, a large rectangular-shaped object is placed in the arena to increase the amount of tactile input to the system.

“Buildings”: Artificial structures

This and the following environments emulate the Khepera robot in a 77×77 cm square arena. Images are pasted to virtual walls placed outside of the arena. The view of the virtual world is projected onto a cylindrical screen which covers a view angle of 280° using a standard computer graphics algorithm. In the “Buildings” setup, four walls are placed in a square around the environment and decorated with pictures of buildings and other man-made structures. An example of what the agent sees is given in figure 5.3.

“Davos”: Natural environment

Mountains may not be a natural habitat for rats, but buildings probably aren’t either. Nevertheless, we also test our model in a natural scene which contains less structure than the man-made objects of the previous two setups. Here, a panoramic view of the Swiss mountain village of Davos is pasted onto a big cylinder surrounding the arena. Thus, unlike the previous setup, the walls are not flat. This results in a more beautiful view of the alps (figure 5.4).

“Minimal”: Impoverished sensory input

The previous environments all provide rich visual stimuli. In most animal experiments, however, the view is restricted to a small number of well-defined cues. In order to emulate such an impoverished environment, four walls are placed in a square around the arena. In the centre of each wall, one simple geometrical object is placed. The objects are a filled black square, a filled white circle, a triangle and a double cross. The background of each wall is covered with noise of low contrast (figure 5.5).

5.2 Visual processing

Place fields of hippocampal neurons in the rat respond to external, most notably visual, as well as internal sensorial cues [196, 174, 58, 220, 158, 80]. In models of spatial cognition, artificial sensors are employed to provide the neural system with perceptions of the world. This sensory input can not be used directly. It needs to be preprocessed before it is passed on to our neural network model. Here we present how the visual input from the camera is decomposed and its dimensionality reduced. The processing is inspired by neuronal properties in primary visual cortex.

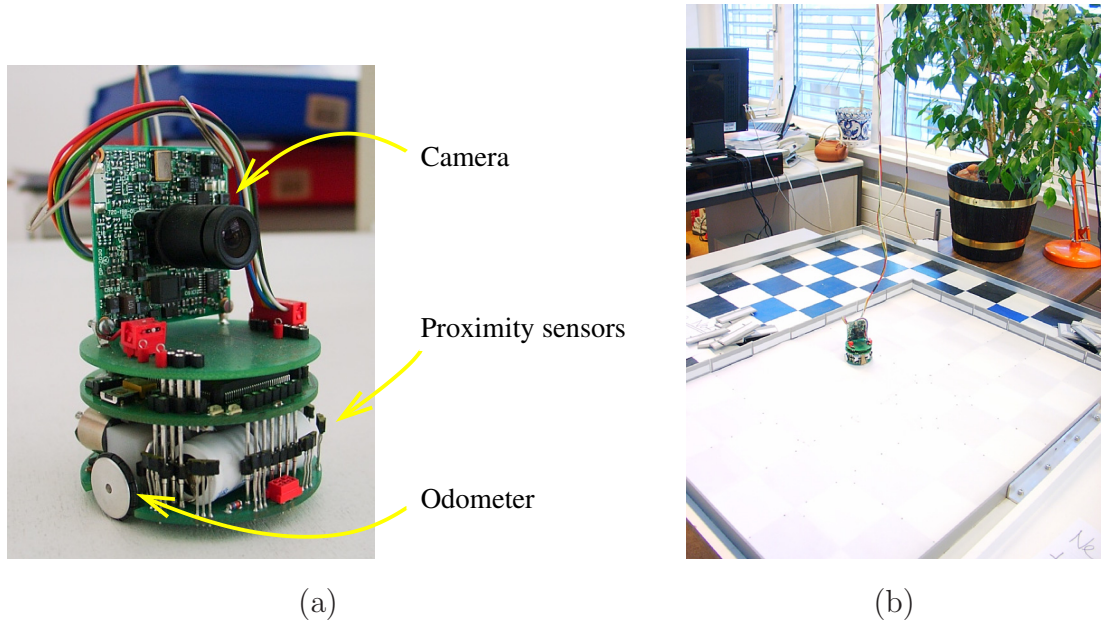


Figure 5.1: (a) The Khepera robot and its sensors. (b) View of the “Office” environment with the Khepera robot.



Figure 5.2: A view taken by the Khepera in the “Office” environment. Four camera images are merged into one panoramic view, covering about 240° .



Figure 5.3: A view taken by the virtual Khepera in the “Buildings” environment. Pictures of buildings are pasted to four walls placed around the arena.



Figure 5.4: A panoramic view of the alps around Davos is pasted onto a cylinder around the arena.

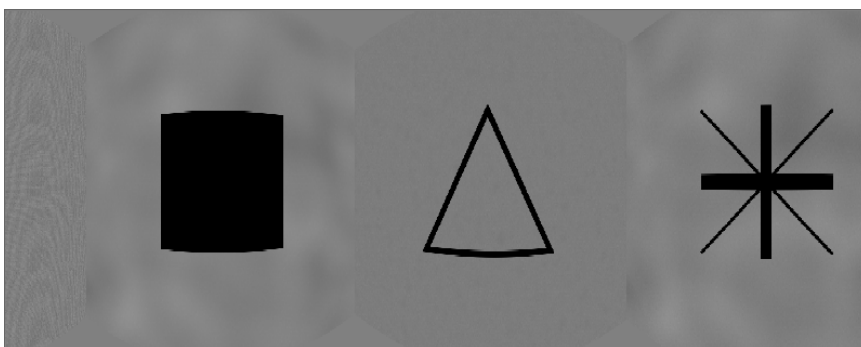


Figure 5.5: A simple object is placed in the centre of each of the four walls surrounding the environment. The walls are also covered by low contrast noise. Here, at a viewing angle of 270°, only three objects are visible.

5.2.1 Gabor filters

The normalised image data $\mathcal{I}(x, y) \in [-1.0, 1.0]$ is represented using a set of Gabor wavelets G . These filters are sinusoidal waves modulated by a Gaussian window function [93, 148]. They respond to edges in the image much like *complex cells* in visual cortex area V1 [123, 61, 134]. The 2-dimensional complex Gabor wavelet in space domain is defined as:

$$G(x, y | \sigma_x, \sigma_y, \omega) = e^{-x^2/2\sigma_x^2} \cdot e^{-y^2/2\sigma_y^2} \cdot e^{i\omega x} \quad (5.1)$$

where σ_x^2 is the variance of the Gaussian window in the direction of the wave propagation, σ_y^2 is the variance in the orthogonal direction and ω is the frequency² of the sinusoidal wave. Depending on the phase of the sine wave, a distinction is made between the real part $\Re(G)$ and the imaginary part $\Im(G)$ (figure 5.6). For an illustration of the 2d-version of $\Re(G)$, see figure 5.7.

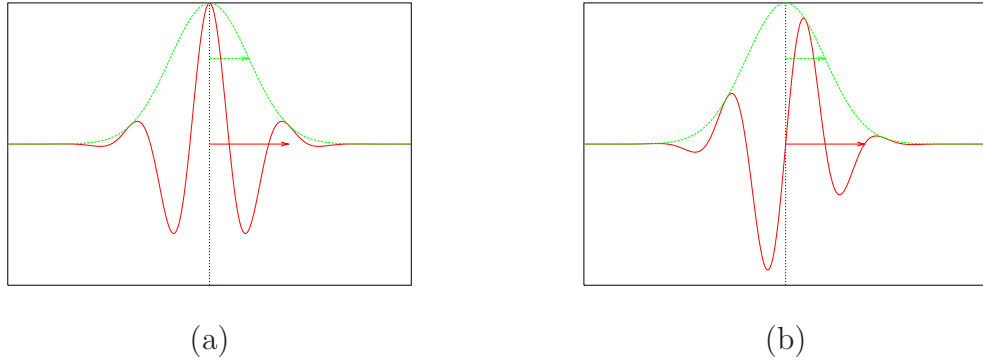


Figure 5.6: Gabor filters in 1 dimension. Red arrow indicates wavelength λ of: (a) real part $\Re(G)$ (b) imaginary part $\Im(G)$. Green arrows indicate standard deviation σ of the Gaussian envelope (in green).

In the Fourier domain, a Gabor filter with frequency ω and variances σ_x^2 and σ_y^2 is a Gaussian centred around the frequency ω and with variances $1/\sigma_x^2$ and $1/\sigma_y^2$:

$$G(x, y | \sigma_x, \sigma_y, \omega) \xrightarrow{\mathcal{F}} \mathcal{G}(\omega_x, \omega_y | \sigma_x, \sigma_y, \omega) = A \cdot e^{-\sigma_x^2(\omega_x - \omega)^2/2} \cdot e^{-\sigma_y^2(\omega_y - \omega)^2/2} \quad (5.2)$$

The filter in equation 5.1 responds best to a vertical bar of width $2\pi/\omega$ (in [pixels]). In order to detect larger and smaller bars with arbitrary orientation, a family of filters

²In this example, it is actually the angular velocity, but for the sake of simplicity, we'll use the term "frequency" to refer to either frequency or angular velocity

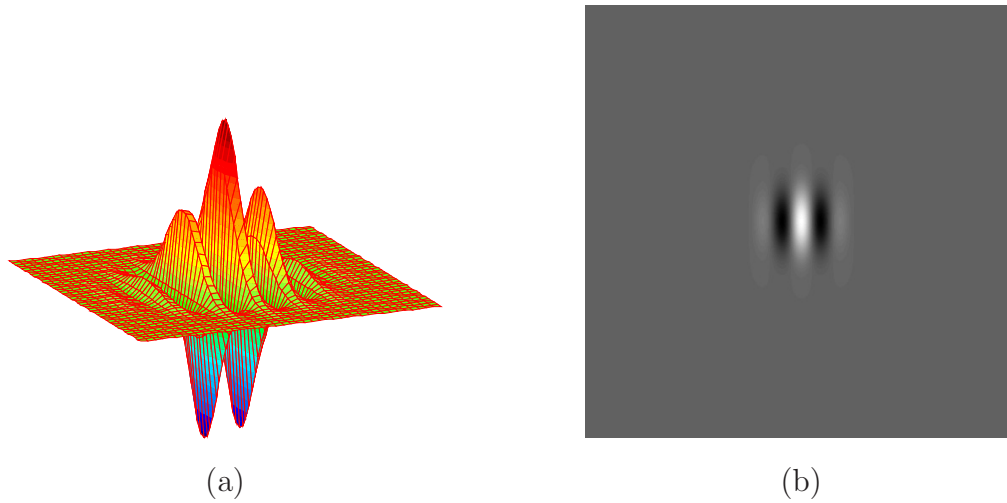


Figure 5.7: Illustration of Gabor filters in 2 dimensions

with varying frequencies and orientations has to be applied to the image $\mathcal{I}(x, y)$. This filter set is created by scaling and rotating the mother wavelet of equation 5.1. A schematic illustration of such a set of filters in the 2d-frequency domain is given in figure 5.8. The ellipses represent the iso-value curves for a cut at $e^{-0.5}A$ (see equation 5.2). The principal axes of the ellipses then correspond to the variances of the Gaussians.

We use three wavelengths ($\lambda \in \{75, 50, 25\}$ [pixels]) and eight orientations which are uniformly distributed in $[0, \pi]$. As we use circular Gaussians ($\sigma_x = \sigma_y = \sigma$), the ellipses are circles in our case. Furthermore, we fix the window size to half the wavelength ($\sigma = \lambda/2$).

Gabor filters have an infinite support because of the Gaussian window. We cut the filters at a distance of 3σ from the centre which cuts off values smaller than $\approx 1.1\%$ of the maximal value A . All filters of the same frequency are cut to the same dimensions. Then, all negative values are scaled such as to achieve a mean of zero. This makes the filters insensitive to shifts in luminance. Finally, all values are scaled in order to get L2-normalised filters ($\sum_{x,y} G(x, y) \cdot \overline{G(x, y)} = 1$, where \overline{x} is the complex conjugate of x).

The filter value $f_{jk}(x, y)$ for frequency j and orientation k is given by the magnitude of the complex filter response to the image \mathcal{I} at position (x, y) :

$$f_{jk}(x, y) = \sqrt{(\Re(G_{jk}(x, y)) * \mathcal{I})^2 + (\Im(G_{jk}(x, y)) * \mathcal{I})^2} \quad (5.3)$$

where $*$ denotes the convolution operator. Combining the two complex parts guarantees a response to an edge in the whole range of the filter window.

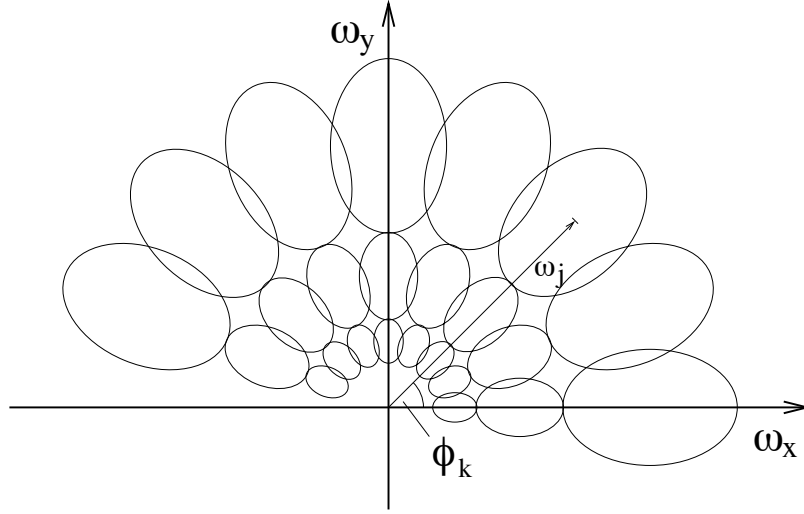


Figure 5.8: The set of Gabor filters in Fourier domain, where they are Gaussians. The ellipses are iso-value curves for a cut at $e^{-0.5A}$ (see equation 5.2). Filter G_{jk} is tuned to frequency ω_j and orientation ϕ_k ($j \in \{1, \dots, n_{freq}\}, k \in \{1, \dots, n_{orie}\}$). We use $n_{freq} = 3$ and $n_{orie} = 8$

5.2.2 Artificial retina

The mobile agent constantly acquires panoramic views of its environment during test experiments. The image resolution is $r_x \times r_y = 800 \times 316$ gray-level pixels in the range $[-1.0, 1.0]$. This high-dimensional space needs to be transformed into a more convenient representation for further processing.

We employ the set of Gabor filters described in section 5.2.1 and apply them to a small set of sampling points in the image \mathcal{I} . These sampling points form an artificial rectangular retina. On each retinal point, the filter vector \vec{f} , composed of all responses to frequencies j and orientations k , is calculated according to equation 5.3.

The distance (δ_x) between adjacent retinal points in x must be in relation to the filter envelope size, parametrised by σ_x . First, we note that a rotation of the agent produces a translation of the image in x . Suppose that an agent rotation translates the image such that a feature which was located at a sampling point is now exactly between this and the neighbouring point. The distance between retina points in x direction has to be small enough to still “see” this landmark. In the y direction, there is no such constraint because the tilt-angle of the robot camera is constant.

We require that the sum of envelopes placed on the retinal points of one row remains approximately constant over the width of the image. This is achieved if δ_x doesn’t exceed $2\sigma_x$ (see figure 5.9). Because the sum of all filter envelopes is ap-

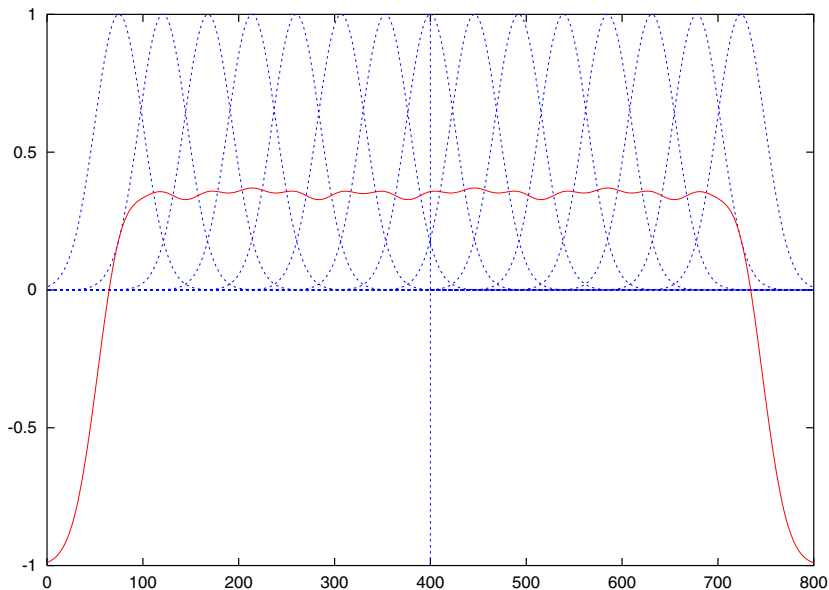


Figure 5.9: Approximately constant sum of Gabor filter envelopes displaced by 2σ

proximately constant, a landmark between two retina points can be located with high accuracy if a population decoding scheme is used [100,324,237,336,239]. Furthermore, all filters must completely lie inside the image.

An example of the retinal response is shown in figure 5.10 for the “Buildings” environment. The retina consists of 15 columns and 3 rows. On each point, the filter vector \vec{f}_j for the lowest wavelength $\lambda_j = 50[\text{pixels}]$ is represented by a blue line. It indicates the direction and “strength” (line length) of edges in its neighbourhood.

5.3 Tactile and vestibular input

Tactile input

The rat’s whiskers are highly sensitive organs. On each side, there are ≈ 30 “macro-vibrissae” of of 10–60mm length, arranged in 5 rows of 4–7 whiskers positioned between eye and mouth. In each row, the length of the vibrissae increases from rostral to caudal [38]. They are mainly involved in spatial tasks and object shape recognition [38]. The 40–70 “micro-vibrissae” are located near the mouth and mostly point downwards. Their small length ($< 7\text{mm}$) and high spatial density makes them well suited for their implication in fine texture discrimination [47, 38]. When exploring objects, rats rhythmically move their whiskers with a frequency of $\approx 8\text{Hz}$. The maximal excursion is around 30° . This active whisking increases the sensory range and

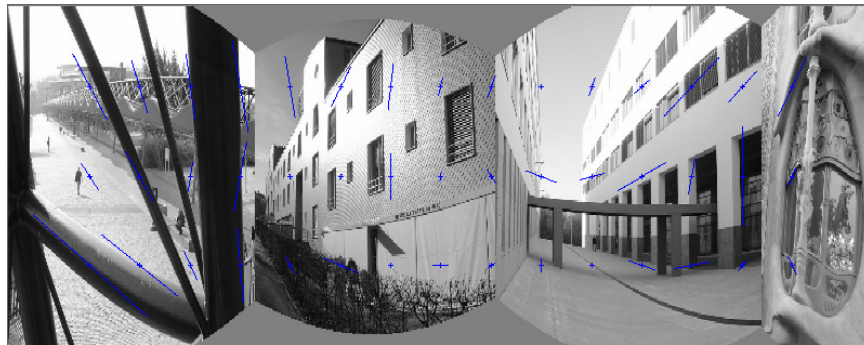


Figure 5.10: Response of the rectangular retina with 15 columns and 3 rows applied to a view from the “Buildings” environment. The blue lines indicate the direction and “strength” (line length) of edges near each retinal point.

may facilitate the detection or discrimination of objects [47,82].

The Khepera robot is equipped with eight infrared light sensors which measure the ambient light as well as proximity to an object. The range of the proximity sensors depend on the surface material of the object. For painted wood (which is used in our experiments), the detectable proximity ranges from around 3 to 5cm. Below 3cm, the sensor is saturated and above, it is unable to detect the object (figure 5.11). Between those bounds, the response is roughly linear. If an object is not placed exactly in front of the robot, but at an angular displacement ϕ , the sensor value decreases almost linearly with ϕ . At $\phi \approx 70^\circ$, the response is zero.

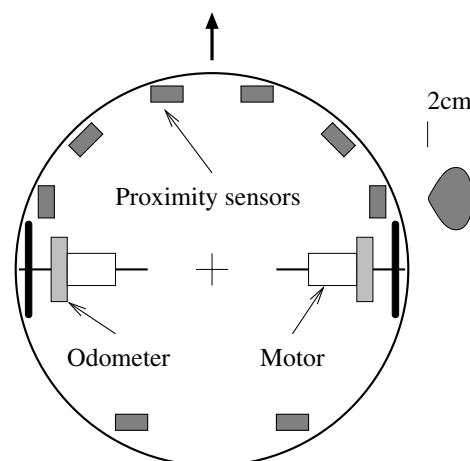


Figure 5.11: Schematic illustration of the Khepera’s sensors. Proximity sensors detect objects from around 2 to 5cm.

“Tactile” input from each of the eight infrared proximity sensors is scaled to $[0, 1]$. 0 means no detectable object, 1 means an object very close to the robot is detected. A vector $\vec{s}(t_k)$ containing all sensor values represents the agent’s tactile input at time t_k . Tactile cells (TCs) are recruited to store and translate this input to neuronal activity. A TC’s activity $r^{tc}(\vec{s}(t_j)|\vec{s}(t_k))$ depends on the mean distance between the current sensor values at time t_j and it’s stored value over all eight sensors, as expressed in equation 5.4.

$$r^{tc}(\vec{s}(t_j)|\vec{s}(t_k)) = e^{-\frac{(\|\vec{s}(t_k) - \vec{s}(t_j)\|)^2}{2k_{tc}\sigma_{tc}^2}} \quad (5.4)$$

where $k_{tc} = 8$ is a normalisation factor and $\sigma = 0.8$ determines the sensitivity of TCs. In practise, the sensors are unfortunately almost binary. Recently, other groups have proposed robotic systems equipped with real rat whisker arrays for object detection and discrimination [82, 110, 83].

Vestibular input

The Khepera’s odometers are used to emulate vestibular or proprioceptive input to our system. After each movement, we use standard trigonometric formulas to calculate the incremental distance $(\Delta x, \Delta y)$ and heading $\Delta\phi$ in an external Cartesian frame (figure 5.12):

$$\begin{aligned} \Delta\phi &= \frac{dR - dL}{B} \\ R &= \frac{dL + dR}{2\Delta\phi} \end{aligned}$$

$$\begin{aligned} \Delta x' &= R \sin \Delta\phi \\ \Delta y' &= R(1 - \cos \Delta\phi) \end{aligned}$$

$$\begin{aligned} \Delta x &= \Delta x' \cos \alpha - \Delta y' \sin \alpha \\ \Delta y &= \Delta x' \sin \alpha + \Delta y' \cos \alpha \end{aligned} \quad (5.5)$$

where dL and dR denote the distance travelled by the left and the right wheel respectively and B is the distance between the two wheels.

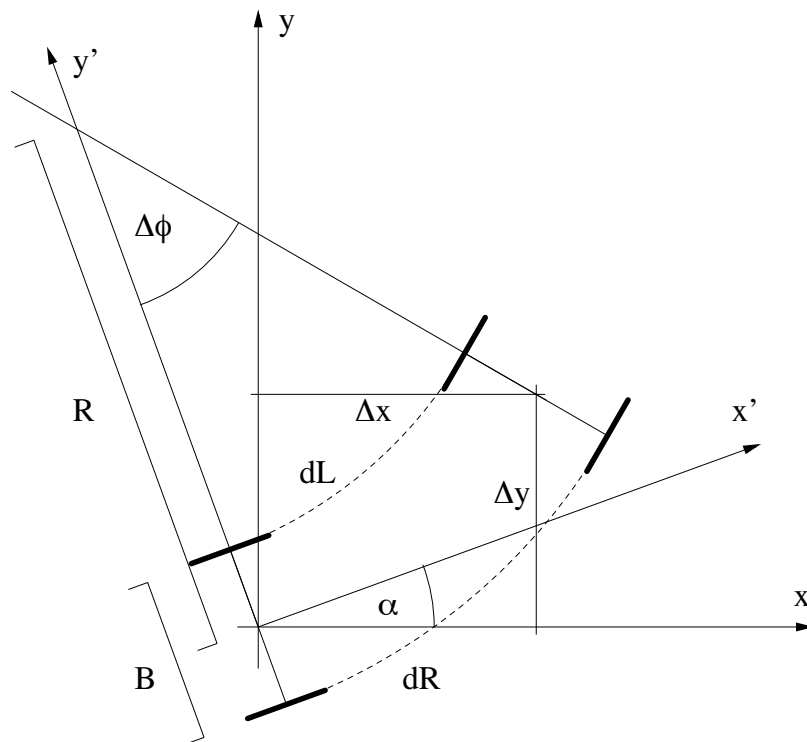


Figure 5.12: Calculating the idiothetic update.

Chapter 6

Spatial representation in the hippocampus: A new model

Neurons in all parts of the hippocampal formation seem to be important for solving spatial learning tasks. In particular, granule cells in dentate gyrus (DG) as well as pyramidal cells in the CA1 and CA3 region of the hippocampus (HPC) fire in strong correlation with the rat's position in the environment [196, 174, 187, 189, 200, 253, 36, 221, 28, 334, 195, 58]. Entorhinal cortex (EC) is the “gateway” to HPC. Highly processed multimodal (visual, olfactory, vestibular, tactile, somatosensory, auditory) information from EC reach HPC via the perforant path [232, 330, 221].

Activity in postsubiculum (poSb) neurons is correlated with the animal's head direction [222, 288, 289, 31]. Damage to the poSb results in severe deficits in spatial behaviours [287].

In this chapter, a new connectionist model of the relation between EC, poSb and HPC is presented. It is able to learn spatial representations based on combining allothetic and idiothetic information. Experimental results are obtained by implementation of the model on a real and simulated mobile robot platform.

6.1 Architecture

The architecture of our model is inspired by the anatomical findings reviewed in 2.1.1. Nevertheless, a functional, rather than an anatomical, terminology for the components of the model is used. Each component is described in turn, results are presented and a mapping to anatomical structures in the brain is proposed.

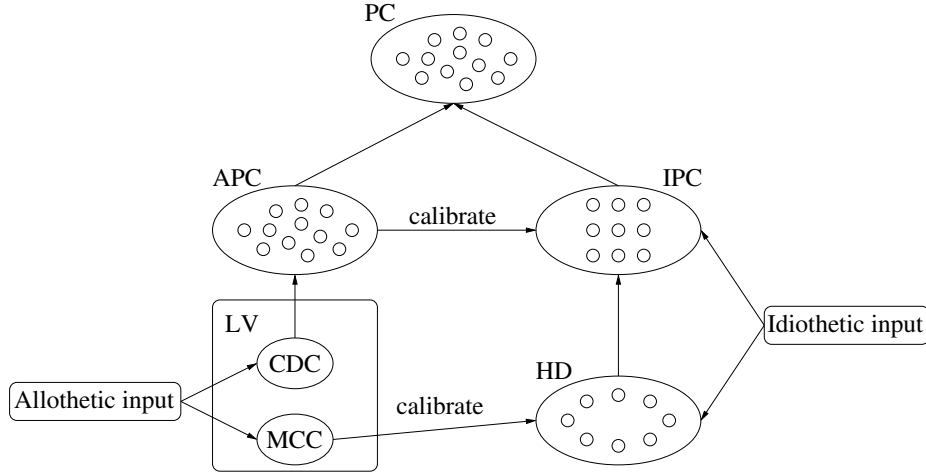


Figure 6.1: Architecture for learning spatial representations. Local view (LV) cells process, store and compare visual stimuli. We model two types of local view cells: Multicolumn cells (MCC) and column-difference cells (CDC). CDCs drive cells in the allothetic place code (APC) module and MCCs calibrate the head direction system (HD). Internal cues drive the idiothetic place code (IPC) and HD. APC calibrates IPC and they both project to the combined place code (PC).

The spatial representation consists of five interconnected modules (figure 6.1): (i) The local view module stores and compares sensory input. (ii) The head direction system continuously updates the agent’s sense of orientation. (iii) The allothetic place code estimates the agent’s position within the environment based on the local view. (iv) The idiothetic place code keeps track of the agent’s position by integration of internal motion cues. (v) The combined place code links the allothetic and idiothetic codes and forms the output of the system.

The head direction system together with the idiothetic place code forms a *path integrator (PI)* (see section 2.2).

6.2 Local view

The local view module receives visual input from the columns of the artificial retina described in section 5.2.2. Its purpose is to extract information which is relevant for the construction of a spatial representation. A previous view stored at time t_k can be compared to the current view at time t_j by defining a similarity measure between two sets of filter activities. This measure is implemented by visual neurons described in this section.

The choice of a comparison function depends on what transformations the image is subject to. The agent’s local view in an environment depends on two variables, namely position and heading. Accordingly, we will call an agent’s movement a “rotation” if it only changes the heading. If only the agent’s position is changed, and that only by a small amount, we call it a “step”. Independently of the environment, a rotation produces a translation along the horizontal axis of the view. A step, however, causes a complicated and environment-dependent transformation. We assume that translations along the vertical axis are not possible, i.e. the agent does not look up or down, only left or right. While this holds for the mobile robot, it is only an approximation for the rat.

One of the problems is that if the visual cues are far away, a step produces only a small change in the view, whereas a rotation always produces a big, but predictable change. For a spatial representation, however, the visual system must permit the discrimination of both heading and position. We therefore “read” the retinal information in two ways: One neural population tries to discriminate well between headings but not position, and vice versa for a second population. Both should have broad tuning curves in order to provide good generalisation of known samples to new views.

Neural substrate: In our model, only low-level features are used to drive the spatial learning system. There is no notion of objects, landmarks or other high-level structures. This is a strongly simplified view of the input system to the hippocampus. An anatomical locus of this processing stage is therefore hard to determine. We just note here that the posterior parietal as well as perirhinal and postrhinal cortices are the main inputs to entorhinal cortex, the gateway to the hippocampus. Hence these regions may be involved in processing (storing and comparing) the local view.

6.2.1 Multicolumn cells

Multicolumn cells (MCCs) aim at discriminating headings regardless of position. The receptive fields should span a large range of headings, i.e. translations of the image should not cause a drastic change in activity.

This is achieved by combining information from neighbouring retinal columns. For each retinal column i positioned at x_i , the weighted sum of filter activities $\vec{h}(x_i, t_k)$ at time t_k , not the activities $\vec{f}(x_i, t_k)$ themselves, are stored/compared:

$$\vec{h}(x_i, t_k) = c_0 \cdot \vec{f}(x_i, t_k) + \sum_{j=1}^{\lceil n_{cols}/2 \rceil} c_j \cdot [\vec{f}(x_{left(i,j)}, t_k) + \vec{f}(x_{right(i,j)}, t_k)] \quad (6.1)$$

where $left(i, j) = |i - j|$ and $right(i, j) = (n_{cols} - 1) - |i + j - (n_{cols} - 1)|$ are the j^{th} column indices to the left and right of the current column i . In the centre, $left(i, j) = i - j$ and $right(i, j) = i + j$. Near the borders, however, columns which

would lie outside of the image are mirrored on the left and right borders. Note that vector \vec{f} in equation 6.1 is composed of the filter responses from all rows of column i . The weights c_j are sampled from a Gaussian \mathcal{N}_c , the standard deviation σ_c [pixels] of which determines the amount of translation invariance. Figure 6.2 shows the weighted sum of the Gaussian Gabor filter envelopes. Similarly to figure 5.9, the receptive field of the filter vector at column x_i is represented by its envelope. For the sum, however, each envelope is weighted by a factor c_j according to its distance j to the central column. The resulting weighted sum illustrates the receptive field of one element of the vector \vec{h} , placed at the central column of the image at pixel 400.

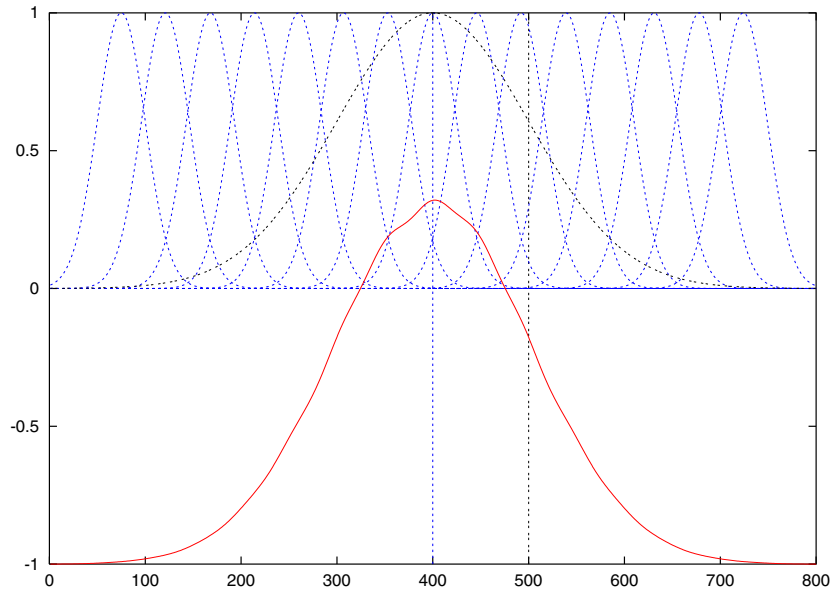


Figure 6.2: The receptive field is enlarged by combining neighbouring columns in a weighted sum. *blue:* Plain filter envelopes. *black:* Gaussian weighing window \mathcal{N}_c (standard deviation $\sigma_c = 100$ [pixels] indicated by line). *red:* Enlarged receptive field by weighted sum of plain filter envelopes (shifted vertically by -1)

When a small stimulus is translated on the image, the plain Gabor filters can only see it in a small range defined by their envelopes. The weighted sum described above enlarges the receptive field without losing discrimination capability. If $h(x_i)$ is tuned to detect a stimulus at position x_i , the shifted stimulus can be detected over a broad range, but only when it is in the central location x_i , $h(x_i)$ is at its maximum.

A translation of the image is caused by a rotation of the agent. It is therefore useful to test the filter response and the weighted sum with respect to the agent's heading. To keep everything as simple as possible, a virtual environment with an "optimal" stimulus as the only visual cue at a heading of 0° is used. The retina

only contains one row and each retina point only contains one Gabor filter f . The stimulus, shown in figure 6.3 (a) is optimal in the sense that it activates most strongly the vertical Gabor filter f of the central column when the agent's heading is 0° .

Three responses are shown versus the agent's heading in figure 6.3 (b): (i) The plain response of Gabor filter $f(x_i)$ at the central column x_i is sharply tuned but does not respond to the stimulus if the bearing angle is greater than $\pm 25^\circ$. (ii) The uniform sum $\sum_j f_j$ of all columns j . Due to the relation between column distance and filter width discussed in section 5.2.2, the sum is constant within the agent's field of view Ψ . Here, $\Psi = 180^\circ$ so the stimulus can't activate the filters for headings greater than $\approx 90^\circ$ because the the image is just gray for those headings. (iii) The weighted sum $h(x_i)$ is broadly tuned, but still discriminates with respect to heading. Here, the width $\sigma_c = \sigma_\phi \cdot r_x / \Psi$ of the Gaussian neighbourhood function for an image width of $r_x = 800$ pixels is expressed in degrees of agent rotation and is set to $\sigma_\phi = 30^\circ$.

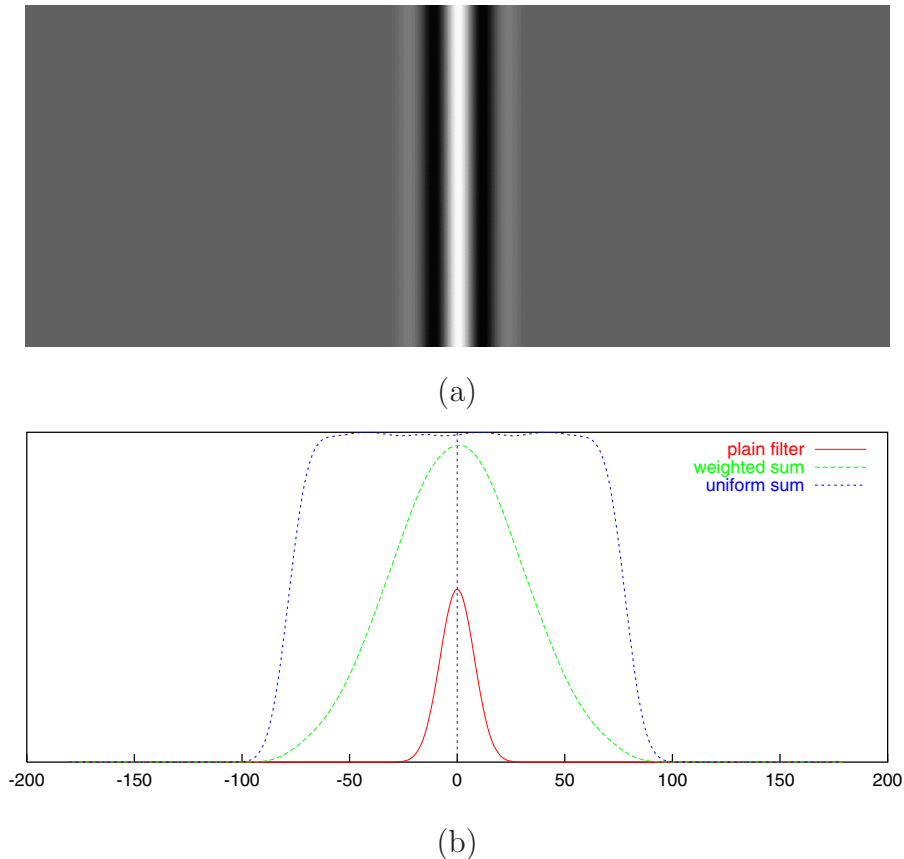


Figure 6.3: (a) "Optimal" stimulus. (b) Rotation experiment. Retina response to optimal stimulus versus agent heading: (a) *red*: filter response f . *green*: Weighted sum h of responses f . *blue*: uniform (unweighted) sum of responses f .

In order to compare two views, a similarity measure must be defined. At each time step t_k , the agent takes a visual snapshot. The image is processed as described above and encoded by a set of MCCs. For each column i , a newly recruited MCC stores the feature vector $\vec{h}(x_i, t_k)$. The activity of an MCC represents the similarity of a column at time t_j with respect to what the cell stored at time t_k . The activation function is:

$$r^{mcc}(t_j | x_i, t_k) = e^{-\frac{(\|\vec{h}(x_i, t_k) \ominus \vec{h}(x_i, t_j)\|_1)^2}{2k \cdot \sigma_{mcc}^2}} \quad (6.2)$$

where $\vec{z} = \vec{x} \ominus \vec{y}$ is the *relative* difference: Each element l of the normal differences is element-wise divided by \vec{x} , i.e. $z_l = (x_l - y_l)/x_l$. $\|\cdot\|_1$ denotes the L1-norm and $k = (n_{rows} \cdot n_{freq} \cdot n_{orie})$ is a normalisation factor (number of rows per retina column, number of gabor frequencies and orientations). Thus r^{mcc} depends on the average relative distance between stored and current filter activity. In figure 6.4, the same rotation experiment as in figure 6.3 shows, for the optimal stimulus, how MCC activity depends on σ_{mcc} . The distance $d = \|\vec{h}(x_i, t_k) \ominus \vec{h}(x_i, t_j)\|_1$ does not increase linearly with the agent’s heading. In particular, d saturates for headings greater than 90° . For this reason, r^{mcc} is not zero for large headings. σ_{mcc} thus determines the spontaneous activity of MCC neurons. In all other experiments, a value of $\sigma_{mcc} = 0.25$ is used.

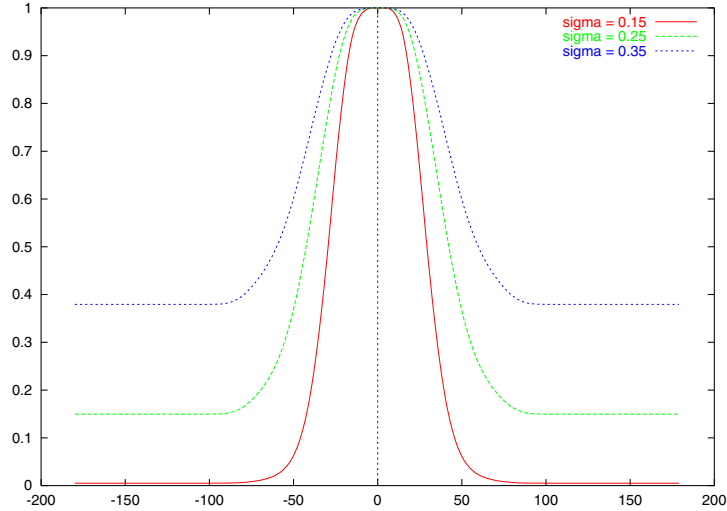


Figure 6.4: MCC activity for an ideal stimulus. The parameter σ_{mcc} determines how much background activity there is.

The receptive fields (RFs) of 50 cells has been visually inspected for all test environments. In figure 6.5 (a), the RFs of three MCCs are visualised. They are taken after 1000 time steps of exploration in the “Buildings” environment (see section 5.1).

A cell's firing rate r in the contour-plots is coded by colour-temperature. For instance, dark red corresponds to $r = 1.0$ and dark blue means $r = 0.0$. For each cell, a block of nine contour plots is shown. Each of the small squares represents the environment. The eight peripheral images illustrate the receptive field when the agent is oriented towards the corresponding direction. For instance, the top-right image shows the receptive field when the agent is facing north-east. The central image is the average of all directional plots.

The standard deviation of the Gaussian neighbourhood function of equation 6.1 (which determines the factors c_j) is expressed in degrees of agent rotation and is set to $\sigma_\phi = 30^\circ$. The range of sensitivity therefore spans almost 180° (see figure 6.3 (b)). For realistic visual input, this also holds. Nevertheless, a position-sensitive component in the MCC activity has been observed in nearly all RFs we visualised. In about 50% of the cases, the centre of the directional RFs are not at the same place for different directions, as is shown for the cell at the top row of figure 6.5 (a). In almost all cases, the mean RF (averaged over the eight directions and shown in the centre of the RF presentation) is very flat. This does not allow a clean localisation of the agent. The cell's preferred heading, however, is in most cases clearly distinguishable. In around a third of the cells observed, a peak of considerable activity is present at a wrong heading. Such a cell is shown at the bottom.

6.2.2 Column difference cells

The second way of reading out the retinal activation should help to discriminate positions, regardless of the agent's heading. Suppose that two distinguishable landmarks l_1 and l_2 are the only visual cues in an environment. Let us define $d(\mathbf{p})$ as the difference in the bearing angle to landmarks l_1 and l_2 when the agent is at location \mathbf{p} . If the agent has a field of view of 360° , the two landmarks are visible for all agent headings and the difference $d(\mathbf{p})$ is independent of this heading. However, the bearing difference depends on the agent's *position*, i.e. for most positions \mathbf{p}_1 and \mathbf{p}_2 , $d(\mathbf{p}_1) \neq d(\mathbf{p}_2)$. This property is exploited by the *column difference cells (CDCs)*.

At every time step t_k , a variable number of CDCs is recruited: We consider two retinal columns s and $s + \delta$. If the filters of both columns are sufficiently active, their filter vector difference $\vec{d}(\delta, s, t_k)$ is stored:

$$\vec{d}(\delta, s, t_k) \text{ stored if } \delta \in \{3, \dots, 6\} \text{ and } \|\vec{f}(x_s, t_k)\|_1 > \theta_{cdc} \text{ and } \|\vec{f}(x_{s+\delta}, t_k)\|_1 > \theta_{cdc}$$

$$\vec{d}(\delta, s, t_k) = \vec{f}(x_s, t_k) - \vec{f}(x_{s+\delta}, t_k) \quad (6.3)$$

We set $\theta_{cdc} = 1.0$. Such an activity threshold is particularly useful in impoverished environments, where most filter columns just see plain walls. It prevents recruitment of cells which store no relevant information.

The column distance δ should ideally correspond to big angle differences because—at least in enclosed environments—big landmark bearing differences depend more on the agent’s position than small ones, which makes position discrimination easier. However, if the agent’s field of view Ψ is smaller than 360° , a big value for δ reduces the probability that both landmarks are visible at the same time. In our experiments, Ψ equals 280° . We allow a range of empirically determined column distances $\delta \in \{3, \dots, 6\}$ which corresponds to bearing differences between approximately 50° and 100° .

CDCs should detect at time t_j if the difference of column filter activities of a given distance δ has already been seen and stored at time t_k . The comparison should be independent of the agent’s heading, i.e. the absolute position of the two columns spaced by δ on the retina should not matter. To achieve this translation invariance in the retinal image, $\vec{d}(\delta, s, t_k)$ is compared with the *most similar* column difference at time t_j . The firing rate $r^{cdc}(t_j | \delta, s, t_k)$ of a CDC which stored the difference vector $\vec{d}(\delta, s, t_k)$ at time t_k is given by:

$$r^{cdc}(t_j | \delta, s, t_k) = e^{-\frac{(\min_i \|\vec{d}(\delta, s, t_k) \ominus \vec{d}(\delta, i, t_j)\|_1)^2}{2k\sigma_{cdc}^2}} \quad (6.4)$$

where k is the same normalisation factor as for MCCs (equation 6.2) and σ_{cdc} determines the baseline firing rate. In our experiments, we set σ_{cdc} to 0.1.

The receptive fields (RFs) of 50 cells has been visually inspected for all test environments. In figure 6.5 (b), the RFs of three CDCs are visualised. They are taken after 1000 time steps of exploration in the “Buildings” environment (see section 5.1). A cell’s firing rate r in the contour-plots is coded by colour-temperature. For instance, dark red corresponds to $r = 1.0$ and dark blue means $r = 0.0$. For each cell, a block of nine contour plots is shown. Each of the small squares represents the environment. The eight peripheral images illustrate the receptive field when the agent is oriented towards the corresponding direction. For instance, the top-right image shows the receptive field when the agent is facing north-east. The central image is the average of all directional plots.

Even though CDCs are translation invariant in the retina frame, their firing rates depend on the agent’s heading, i.e. CDCs are directional. The reason is that the agent’s field of view (280°) is smaller than 360° . Sometimes, only one of the two columns is visible and the other is hidden. Additionally, the difference $\vec{d}(\delta, i, t_j)$ of column x_i and $x_{i+\delta}$ in equation 6.4 is calculated for all $i \in \{1, \dots, n_{cols} - \delta\}$ where n_{cols} is the number of retina columns. In other words, columns in the retina are not wrapped around. Nevertheless, rotation invariance of at least 90° was present in all observed cells. The activity in directions opposite to the main peak is low in all cases observed. The two cells in the top and central row show rather clean position tuning. The bottom row displays a cell with a doubly peaked tuning curve. Such ambiguous RFs has been observed in roughly 30% of the cells.

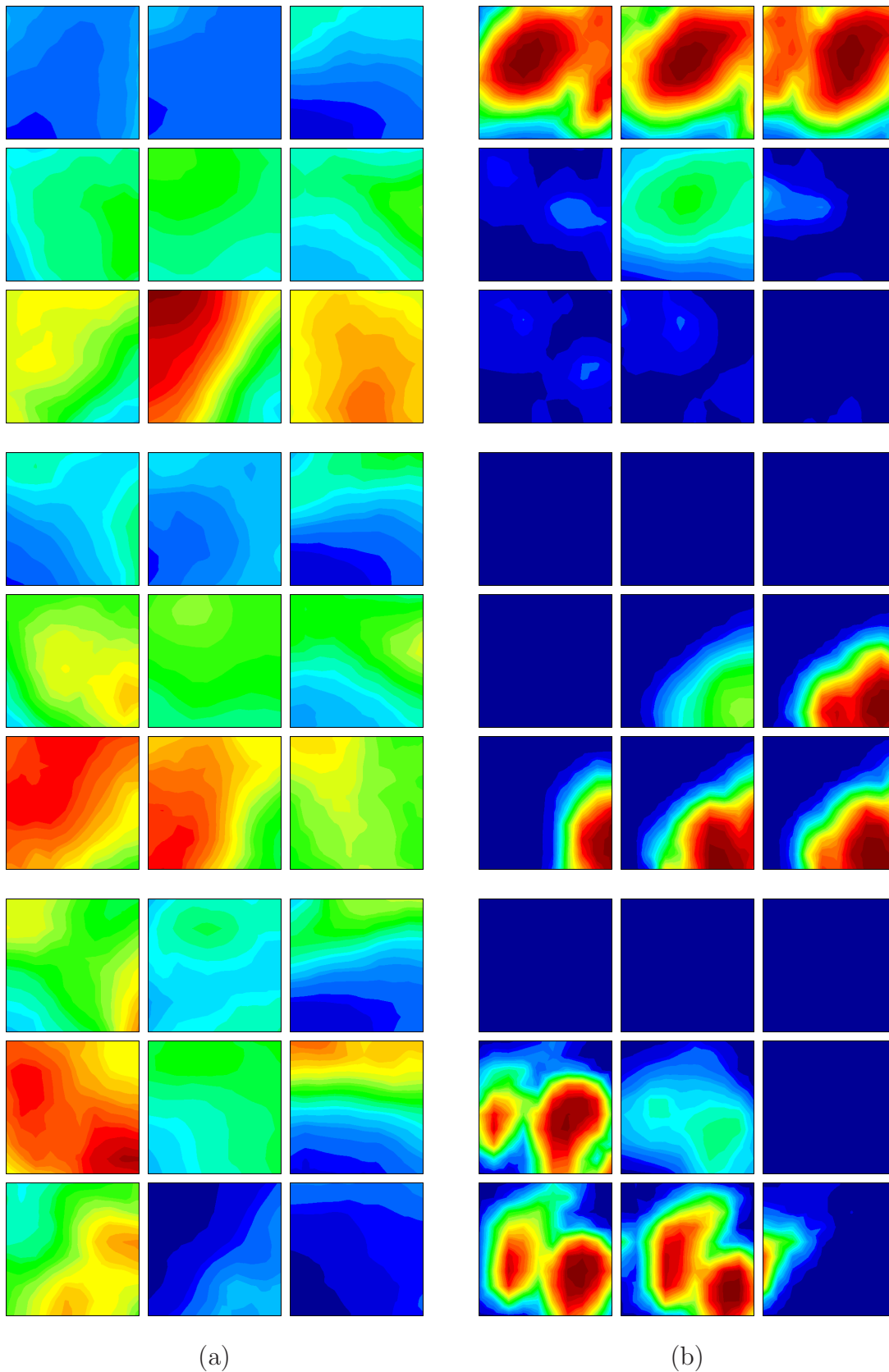


Figure 6.5: Receptive fields. See text for interpretation. (a) Multicolumn cells (MCCs). (b) Column difference cells (CDCs)

6.3 Head direction system

The head direction (HD) module forms the most important part of the path integrator system. Small rotational errors can totally impair the system’s ability to build and maintain a spatial representation. This section describes how the HD system combines idiothetic and allothetic information in order to produce a stable (non-drifting) estimate of the agent’s heading.

A population of $N_{hd} = 120$ directional neurons codes for the agent’s heading Φ with respect to an arbitrary fixed compass bearing. Each HD neuron i represents the heading $\phi_i = i \cdot 360^\circ / N_{hd}$.

The firing rate of HD cells is calculated in two stages. First, the current heading Φ is estimated. Then, a large activity profile around Φ is enforced in the HD system. Lateral interconnection between HD cells could be the neuronal substrate for such activity profiles. Here, we emulate lateral interactions by enforcing a Gaussian activity profile around Φ . Formally, the firing rate r_i^{hd} of HD cell i is:

$$r_i^{hd} = e^{-\frac{(\|\Phi - \phi_i\|_\varphi)^2}{2\sigma_{hd}^2}} \quad (6.5)$$

where $\|\cdot\|_\varphi \in [0, 2\pi]$ is the angular distance and σ_{hd}^2 is the angular variance of the Gaussian profile. In our experiments, a value of $\sigma = 60^\circ$ is used. An example of the HD activity profile is illustrated in figure 6.6 (a) for a heading of $\approx 105^\circ$.

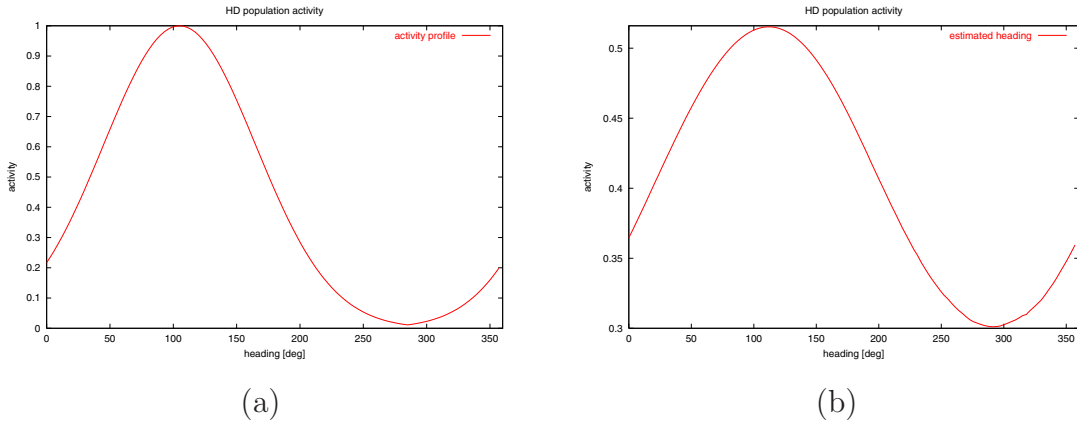


Figure 6.6: Activity of the head direction (HD) cell population. (a) Enforced Gaussian activity profile driven by idiothetic cues for a heading of $\approx 105^\circ$. (b) HD activity at the same location but driven by allothetic cues only.

When the agent moves¹, its angular displacement $\delta\Phi_i$ is estimated by dead-

¹It is always assumed that the agent performs “uniform” movements. The movement is uniform in the sense that the angular velocity ω remains constant throughout the movement.

reckoning (odometry) and made available to the HD system when the movement is completed. The new heading, estimated by idiothetic cues, is then $\Phi_i = \Phi + \delta\Phi_i$, where Φ is the previous heading estimation.

Any real-world path or heading integration system based on dead-reckoning is subject to noise, and more importantly, to a systematic drift. In order to keep the HD representation accurate, the local views encountered during exploration are continuously associated to their heading direction. In particular, all multicolumn cells (MCCs) are connected to all HD cells and synapses are activated/modified at each time step as follows: If a synapse has zero strength and its pre- and postsynaptic activities r_j and r_i are above a threshold θ , the synapse is activated and its strength initialised using a one-shot Hebbian type learning rule:

$$w_{ij} = r_i \cdot r_j \quad \text{if } r_i > \theta \text{ and } r_j > \theta \quad (6.6)$$

Once activated, the synaptic efficacy is modified at each time step:

$$\Delta w_{ij} = \eta \cdot r_i (r_j - w_{ij}) \quad (6.7)$$

where η is a small learning rate. After each learning step, weights are renormalised: $\tilde{w}_{ij} = w_{ij} / (\sum_k w_{ik})$. The activities of all MCCs j connected to HD cell i produce an input potential $h_i = \sum_j \tilde{w}_{ij} r_j$ at the HD cell i . The allothetic heading estimate Φ_a is then given as the circular population vector of the input potentials h_i :

$$\Phi_a = \arctan \left(\frac{\sum_i h_i \cdot \sin(2\pi i / N_{hd})}{\sum_i h_i \cdot \cos(2\pi i / N_{hd})} \right) \quad (6.8)$$

In figure 6.6 (b), an example input potential distribution is shown. It is recorded in the “Buildings” environment after 1000 exploration time steps when the agent faces a heading of $\approx 105^\circ$. The activity is indeed centred around the correct heading, with a broad but clearly discriminative distribution.

At each time step, both idiothetic and allothetic heading estimations are determined. The new estimated heading is then calculated as:

$$\Phi = \Phi_i - \alpha \cdot (\Phi_i - \Phi_a) \quad (6.9)$$

$\alpha \in [0, 1]$ determines the influence of the allothetic estimate. A value of zero means no allothetic influence at all whereas a value of one would completely ignore idiothetic information. Note that although we explicitly access angular information in equations 6.5, 6.8 and 6.9 which seems at first sight biologically implausible. However, if continuous attractors using lateral connections [9, 10, 335, 227, 238, 272] or a probabilistic activity transition matrix [117, 118] between HD cells was used for implementing equation 6.5, the recalibration could be performed without explicitly accessing angular information.

The advantage of the idiothetic update is that it is always smooth. There are no “jumps” if the noise in the dead reckoning system is not too big. This smoothness is not present in the allothetic heading estimation. At each time step, the previous heading is forgotten and a new independent estimation is performed. In other words: Idiothetic estimation contains memory and allothetic estimation does not. For this reason, α in equation 6.9 should be small. We use $\alpha = 0.1$. As an example, the distributions of Φ_a and Φ for the “Buildings” environment are shown in figure 6.7. While the distribution of Φ_a is centred around zero, the drifting odometry produces a shift into the distribution of Φ . The idiothetic noise, however, is smaller than for the allothetic estimation. This is the reason why the recalibrated heading error distribution is narrower.

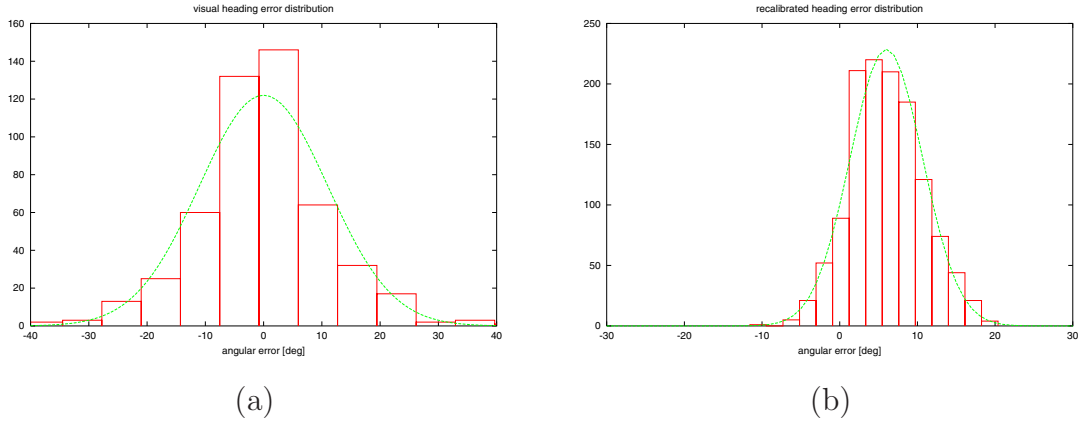


Figure 6.7: Heading error distributions in the explored “Buildings” environment. (a) Pure allothetic estimation. (b) combined allothetic and idiothetic estimations.

The tracking capability of the HD system is tested to ensure that calibration works. Figure 6.8 (a) illustrates for the explored “Buildings” environment that the non-calibrated heading error ($\alpha = 0$) drifts away whereas the error for the continuously calibrated system ($\alpha = 0.1$) remains bounded.

The distribution of allothetic heading estimation errors is experimentally estimated for all fully explored test environments. The agent is placed 500 times at random positions and headings. For each trial, the allothetic heading estimation error $\Delta\Phi_a$ is recorded. Gaussians are fitted to the distributions. The mean values are below one degree for all test environments. The standard deviations for all environments are shown in figure 6.8 (b). In a second series of experiments, the agent navigates for 500 time steps in each environment, continuously recalibrating its HD system according to equation 6.9. The recalibrated heading estimation error $\Delta\Phi$ is

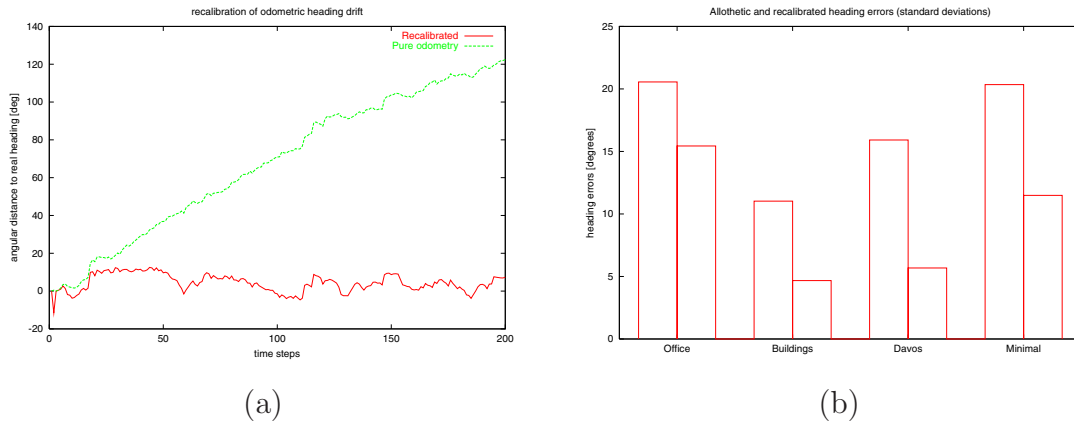


Figure 6.8: Calibrating the head direction (HD) system using allothetic information. (a) Heading error over time for pure idiothetic ($\Delta\Phi_i$) and recalibrated ($\Delta\Phi$) HD system. The systematic drift is successfully removed. (b) The standard deviations of the heading estimation errors for all test environments. The higher values are the allothetic estimation errors $\Delta\Phi_a$, whereas the lower values are for the combined allothetic and idiothetic estimates $\Delta\Phi$.

recorded. Its standard deviation is also shown in figure 6.8 (b). It is always smaller than in the pure allothetic case, which confirms that idiothetic information is extremely useful.

Neural substrate: Head direction cells have been found in several regions of the brain. However, the postsubiculum (poSb), together with anterodorsal (adT) as well as lateral mammillary (lmT) nuclei of the thalamus seem to be key brain areas for spatial orientation [222, 288, 289, 31]. Damage to poSb, for instance, results in severe deficits in spatial behaviours [287]. Neurons in the subiculum (Sb) show environment-independent firing which is correlated to position *and* heading [258, 255, 257]. We therefore propose that adT, lmT, poSb and possibly Sb form the neural substrate of the HD module.

6.4 Allothetic place code

In our allothetic place code module, a spatial representation based on visual information is constructed from experience. In each time step, a new neuron i is recruited if the current place is not yet well represented by a sufficient number of active neurons:

$$\text{new cell } i \text{ recruited if } \sum_k \mathcal{H}(r_k - \theta_{act}) < C \quad (6.10)$$

where k runs over all APC neurons, C controls the density of coverage, $\mathcal{H}(\cdot)$ is the Heaviside function and θ_{act} is a firing rate threshold above which neurons are considered active. We set $\theta_{act} = 0.8$, which means 80% of the maximum activity. If the place is already well represented, no new cell is recruited.

Suppose that a new cell i has been recruited. Neurons coding for the current local view synapse on the new place cell i . In particular, a difference cell (CDC) j connects with weight w_{ij} to place cell i if its firing rate r_j is high:

$$w_{ij} = \begin{cases} r_j \cdot \underbrace{r_i}_{=1} & \text{if } r_j > \theta_{act} \\ 0 & \text{else} \end{cases} \quad (6.11)$$

This is a thresholded one-shot Hebbian-type rule with learning rate one. The newly recruited cell should represent the current place. Therefore, it should be maximally active ($r_i = 1$) for the current afferent CDC projection. This is achieved by tuning the parameters of the neuron's piecewise linear activation function:

$$r_i = \begin{cases} 0 & \text{if } \kappa_i h_i < \theta_{low} \\ 1 & \text{if } \kappa_i h_i > 1 \\ \frac{\kappa_i h_i - \theta_{low}}{1 - \theta_{low}} & \text{else} \end{cases} \quad (6.12)$$

where $h_i = \sum_j w_{ij} r_j$ is the input potential to APC neuron i , $\theta_{low} = 0.2$ is the minimal input to activate the neuron and $\kappa_i = 1/h_i^0$ determines the saturation potential of neuron i , with h_i^0 standing for the input potential at the time when neuron i was recruited. At the moment when the cell i is recruited, we have $\kappa \cdot h_i = 1$, and hence $r_i = 1$. For this reason, r_i may be omitted in equation 6.11

The resulting place code represents the agent's position in the environment. The agent's encoded position is interpreted by calculating the population vector [100, 189, 324, 237, 34, 336, 239]:

$$\mathbf{P}_a = \frac{\sum_i r_i \cdot \mathbf{x}_i}{\sum_i r_i} \quad (6.13)$$

where \mathbf{x}_i is the agent's position where APC i was recruited. The position error vector $\Delta \mathbf{P}_a = \mathbf{P} - \mathbf{P}_a$ is the difference between the real position and the allothetic position estimate.

To assess the precision of the allothetic place code, the position error distribution is estimated experimentally for all test environments. First, the agent explores the environment in order to establish a population of APCs. After exploration, the agent is placed 500 times at random positions and headings. For each trial, the position error vectors $\Delta \mathbf{P}_a$ are recorded. Circular two-dimensional Gaussians are fitted to the $\Delta \mathbf{P}_a$. The mean values for all environments are below 1cm. The standard deviations of the

Gaussian fits are plotted in figure 6.9 (a). As an example, the distribution of $|\Delta\mathbf{P}_a|$ is shown in figure 6.9 (b) for the “Buildings” environment.

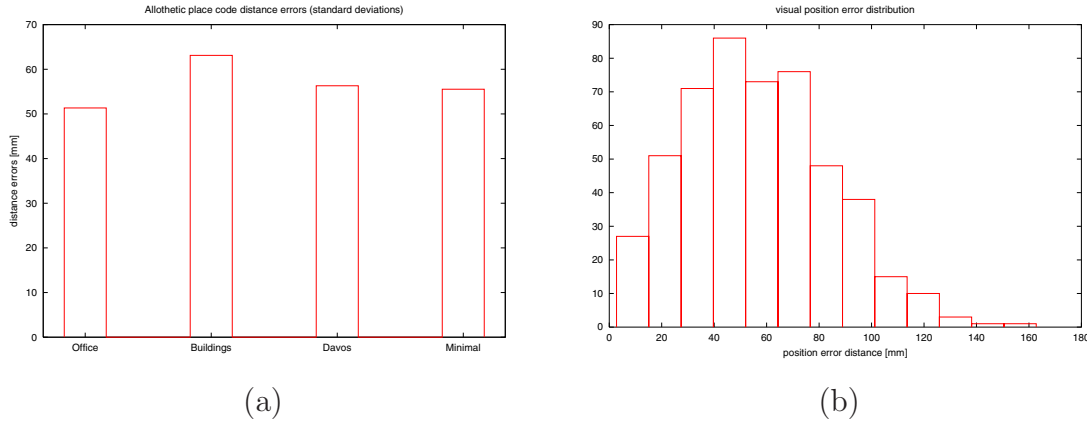


Figure 6.9: Allothetic position error distribution. (a) For all test environments, the standard deviations of circular Gaussians fitted to the position estimation error vectors $\Delta\mathbf{P}_a$ are plotted. (b) The distribution of the lengths of the error vectors $|\Delta\mathbf{P}_a|$ for the “Buildings” environment.

Examples of the resulting place code are visualised in figure 6.12 (b) for the “Buildings” environment after 1000 time steps of exploration. The square represents the environment. Each dot represents an APC neuron i . Cell i is drawn at the location where the agent was when i was recruited. Its firing rate r_i is coded by colour-temperature. For instance, dark red corresponds to $r = 1.0$ and dark blue means $r = 0.0$. The population vector is indicated by a cross. The agent’s heading is not indicated in the plots. Not all APCs which code for a certain position are active when the agent is near it. The reason is that APCs are directional. They only fire when the agent’s heading is near the cell’s preferred heading. This becomes clearer when the receptive fields are discussed in the next paragraph. For most agent positions, the code is rather disperse. The population vector, however, reliably codes for the real agent’s location, which is indicated in the left column (figure 6.12 (a)). In some places (most notably in the corners), the code is sparse. This is probably due to the fact that close walls make nearby places look more different than if the walls are far away. The last row shows an example where the allothetic estimate deviates more than usual from the correct value.

The receptive fields (RFs) of 50 cells has been visually inspected for all test environments. In figure 6.13 (a), the RFs of three APC neurons are visualised. They are taken after 1000 time steps of exploration in the “Buildings” environment (see section 5.1). A cell’s firing rate r in the contour-plots is coded by colour-temperature.

For instance, dark red corresponds to $r = 1.0$ and dark blue means $r = 0.0$. For each cell, a block of nine contour plots is shown. Each of the small squares represents the environment. The eight peripheral images illustrate the receptive field when the agent is oriented towards the corresponding direction. For instance, the top-right image shows the receptive field when the agent is facing north-east. The central image is the average of all directional plots.

RFs of APC neurons are rather stereotypic. The cells are broadly tuned around their preferred position, like the cell in the first row of figure 6.13 (a). All APC cells observed are directional, and they are all activated in a range of about 180° around the preferred agent heading. Cells that code for positions near the corners tend to have smaller RFs, as shown for the cell in the centre. The cell shown in the bottom row has an unusually large and flat RF.

Neural substrate: Entorhinal cortex (EC) is the gateway to the hippocampus (HPC). It receives processed multisensory information from cortex and projects to HPC. Environment-independent place cells have been found in the medial EC. An allothetic place code, however, must depend on the environment. To our knowledge, no published studies investigate spatial firing properties in the lateral EC. As this is the only other pathway to HPC, we postulate that the lateral entorhinal cortex contains a broadly tuned, environment-dependent allothetic place code.

6.5 Idiothetic place code

The HD system described above keeps track of the agent’s current compass bearing. The idiothetic place code (IPC) module presented in this section implements the memory for the agent’s current position. Together, they work as a path integrator, i.e. an environment-independent spatial representation.

A population of $N_{ipc} = 400$ simulated neurons encode the agent’s estimated position \mathbf{P}_i in a Cartesian coordinate frame. Each IPC neuron j is assigned a predefined preferred position \mathbf{p}_j such that a square region of space is uniformly covered. The firing rate r_j of cell j is a two-dimensional Gaussian with variance σ_{ipc}^2 over the euclidian distance $\|\mathbf{P}_i - \mathbf{p}_j\|_2$:

$$r_j = e^{-\frac{(\|\mathbf{P}_i - \mathbf{p}_j\|_2)^2}{2\sigma_{ipc}^2}} \quad (6.14)$$

As with the HD system, such an activity profile may result from lateral interactions between the IPC neurons. The estimated agent position \mathbf{P}_i is updated using dead-reckoning (amount of displacement) as well as HD (direction of movement) information. Figure 6.12 (a) illustrates the population activities at various places in the square “Buildings” environment.

In order to prevent the estimated agent position \mathbf{P}_i from drifting away, the representation needs to be recalibrated using allothetic information. At each time step, the idiothetic and allothetic position estimates \mathbf{P}_i and \mathbf{P}_a are determined. The new recalibrated position estimate \mathbf{P} is then calculated as:

$$\mathbf{P} = \mathbf{P}_i - \beta \cdot (\mathbf{P}_i - \mathbf{P}_a) \quad (6.15)$$

where $\beta \in [0, 1]$ determines the influence of the allothetic estimate. Similar to the HD system, a value of zero means no allothetic influence and a value of one would completely ignore idiothetic information. We used $\beta = 0.1$ in our experiments. As in the HD system, a small value of β keeps the position estimate smooth, while still removing systematic drifts.

In order to calculate the allothetic position estimate \mathbf{P}_a , we use the position of allothetic place field centres (\mathbf{x}_i in equation 6.13). This is not a biologically plausible way of recalibrating the IPC population. Similarly to the HD system, however, we could associate allothetic place cell activity to the firing of idiothetic place cells using unsupervised Hebbian learning and use an attractor network for implementing equation 6.14. Then, no explicit spatial information would be needed for recalibration.

This calibration method removes the systematic drift inherent in a pure idiothetic system. This is illustrated in figure 6.10 (a) for the explored “Buildings” environment: The uncalibrated ($\beta = 0$) position error vector length drifts away whereas the error for the recalibrated system ($\beta = 0.1$) remains bounded.

The distribution of the recalibrated position errors is experimentally estimated for all fully explored test environments. The agent navigates for 500 time steps in each environment, continuously recalibrating its IPC system according to equation 6.15. As for the allothetic place code, circular two-dimensional Gaussians are fitted to the position vector errors $\Delta\vec{P}$. The standard deviations of $\Delta\vec{P}_a$ (from figure 6.9 (a)) and $\Delta\vec{P}$ are plotted in figure 6.10 (b). As expected, the recalibrated estimate yields a narrower distribution than the pure allothetic estimate.

Examples of the resulting place code are visualised in figure 6.12 (a) for the “Buildings” environment after 1000 time steps of exploration. The square represents the environment. Each dot represents an IPC neuron i . Cell i is drawn at the location where the agent was when i was recruited. Its firing rate r_i is coded by colour-temperature. For instance, dark red corresponds to $r = 1.0$ and dark blue means $r = 0.0$. The population vector is indicated by a cross. The agent’s heading is not indicated in the plots. Unlike the APC population code in the central column (figure 6.12 (b)), IPC cells are only tuned to position, and not to heading. The tuning curves of IPC neurons is determined before exploration starts, and cells are uniformly distributed over the environment.

Neural substrate: The medial entorhinal cortex (mEC) contains neurons with spatial firing correlates. The spatial selectivity of an mEC place cell is lower than in

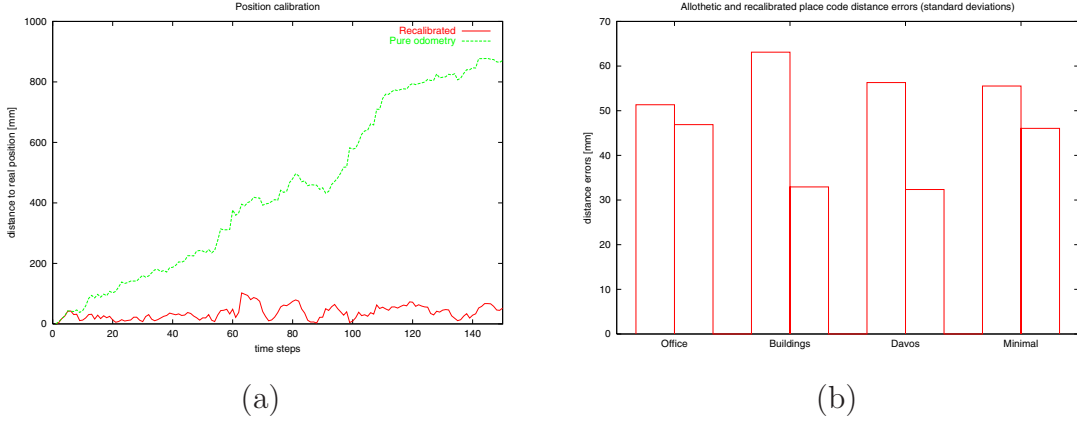


Figure 6.10: Calibrating the idiothetic place code using allothetic information. (a) Position error over time for pure and recalibrated idiothetic place code in the “Buildings” environment. The systematic drift is successfully removed. (b) The standard deviations of the position estimation errors for all test environments. The higher values are for pure allothetic estimates \mathbf{P}_a , whereas the lower values are for the combined allothetic and idiothetic estimates \mathbf{P} .

the hippocampus, but there is a clear peak activity at the cell’s preferred position. mEC Neurons are likely to be active in any environment, regardless of shape and size. The shape of the place fields is also preserved across environments [221, 89, 90]. Place fields in the subiculum (Sb) are also environment-independent. They can, however, be directional [23, 258].

We propose that mEC and possibly Sb are the neural substrates for an idiothetic place code. mEC is innervated by presubicular neurons [329], which could provide heading information to update the current position estimate in mEC. The lateral entorhinal cortex strongly projects to mEC [221], which could serve as a positional calibration signal using allothetic cues.

6.6 Combined place code

The idiothetic and allothetic place cells project to a layer of combined place code (PC) neurons (see figure 6.1). At each time step, place cells are recruited according to equation 6.10. Synapses originating in the APC and IPC layers are recruited and initialised as defined by equation 6.11. The firing rate of PC neuron i is given by equation 6.12. The input threshold $\theta_{low} = 0.3$ in the PC layer is higher than for APC neurons. This results in a sparser representation and smaller receptive fields.

Additionally, at each time step, weights w_{ij} of APC→PC synapses are modified using a Hebbian-type learning rule:

$$\Delta w_{ij} = \eta r_i(r_j - w_{ij}) \quad (6.16)$$

where r_i is the firing rate of PC neuron i and r_j is the firing rate of APC neuron j . η is the learning rate.

Example population activities are shown in figure 6.12 (c) for the “Buildings” environment after 1000 time steps of exploration. The square represents the environment. Each dot represents a PC neuron i . Cell i is drawn at the location where the agent was when i was recruited. Its firing rate r_i is coded by colour-temperature. For instance, dark red corresponds to $r = 1.0$ and dark blue means $r = 0.0$. The population vector is indicated by a cross. The agent’s heading is not indicated in the plots. The environment is densely covered by place cells. Most place cells near the encoded position are active. This suggests that most cells are non-directional, unlike the APC neurons in the central column (figure 6.12 (b)). The code is also more sparse than the allothetic representation.

The receptive fields (RFs) of 50 cells has been visually inspected for all test environments. In figure 6.13 (b), the RFs of three PC neurons are visualised. They are taken after 1000 time steps of exploration in the “Buildings” environment (see section 5.1). A cell’s firing rate r in the contour-plots is coded by colour-temperature. For instance, dark red corresponds to $r = 1.0$ and dark blue means $r = 0.0$. For each cell, a block of nine contour plots is shown. Each of the small squares represents the environment. The eight peripheral images illustrate the receptive field when the agent is oriented towards the corresponding direction. For instance, the top-right image shows the receptive field when the agent is facing north-east. The central image is the average of all directional plots. The majority of place cells are heading independent, like the cell shown at the top. Place fields are more compact than in the APC layer. Around 20% of the observed cells are to some degree directional. Such a cell is shown in the central row. We observed one cell only which does not respond in all headings. This cell is shown at the bottom.

Neural substrate: The hippocampal areas are candidate locations for the combined place representation. All regions (dentate gyrus, CA3 and CA1) show clear spatial firing properties as well as sparse coding. The model PC activity is also sparser than in both APC and IPC layers, consistent with experimental data. Our model of PC is purely feed-forward. As DG and CA3 contain lateral interconnections whose function is not yet resolved, our model best fits the hippocampal region CA1.

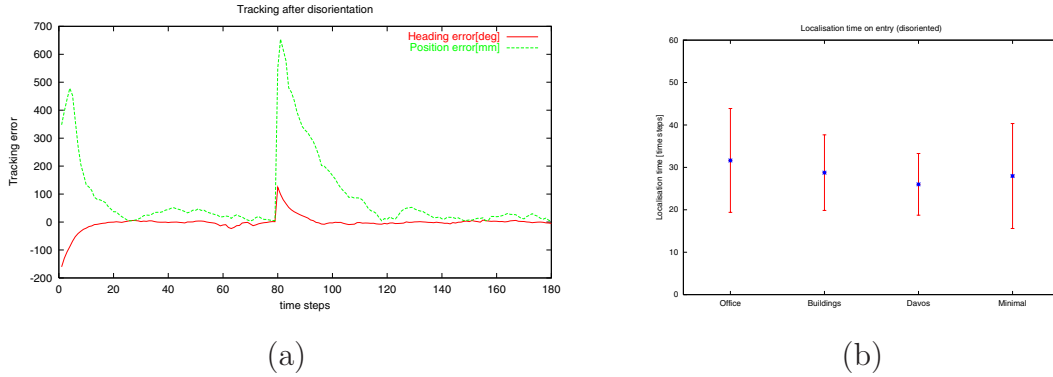


Figure 6.11: (a) Position and heading error evolution when the agent is disoriented and inserted into the explored “Buildings” environment. At time step 80, the agent is again disoriented. (b) Mean number of time steps until the agent is localised for all test environments.

Entering a familiar environment

In order to assess the performance of the full system, we test if the model is capable of localising itself autonomously. We use the following procedure:

The agent is put at a random position and heading into a familiar environment. Then, its IPC and HD systems are randomly initialised. This corresponds to a disorientation process often used in animal experiments. Next, the agent moves randomly in the environment and tries to localise itself. Using the described calibration scheme, the agent may recalibrate its path integrator using allothetic cues. At each time step, the allothetic representation “pulls” the path integrator towards the allothetic estimation of position and heading. An example of the temporal evolution of the recalibrated path integrator (HD and IPC modules) estimation errors $\Delta\Phi$ and $\Delta\mathbf{P}$ is shown in figure 6.11 (a) for the “Buildings” environment.

As can be seen in the first 10 time steps, the position estimation error heavily depends on the correct heading direction: it may even continue to rise for some time until the heading error is sufficiently low. At time step 80, the agent is again “disoriented”, but the calibration process continues to “draw” the idiothetic representation towards the allothetic estimate in order to produce a consistent place code.

How many time steps does it take until the estimations are “locked” to the correct values? This question is answered by the next series of experiments. In each of the explored test environments, 100 “disorientation” experiments are performed: The agent is placed at a random position with random heading. Its IPC and HD systems are randomly initialised, and the agent starts to move randomly, trying to localise itself. Whenever the heading and position error estimates $\Delta\Phi$ and $\Delta\mathbf{P}$ fall below the

standard deviations for the allothetic estimates (figures 6.8 (b) and 6.10 (b)) for five consecutive time steps, we conclude that the agent has localised itself and the trial is ended. The mean value and standard deviations of the trial lengths are shown in figure 6.11 (b). If we assume that the rat “samples” its sensory input with a rate of 8Hz [254, 41], which is in the frequency range of the theta rhythm, it takes our simulated rat around four seconds to localise itself.

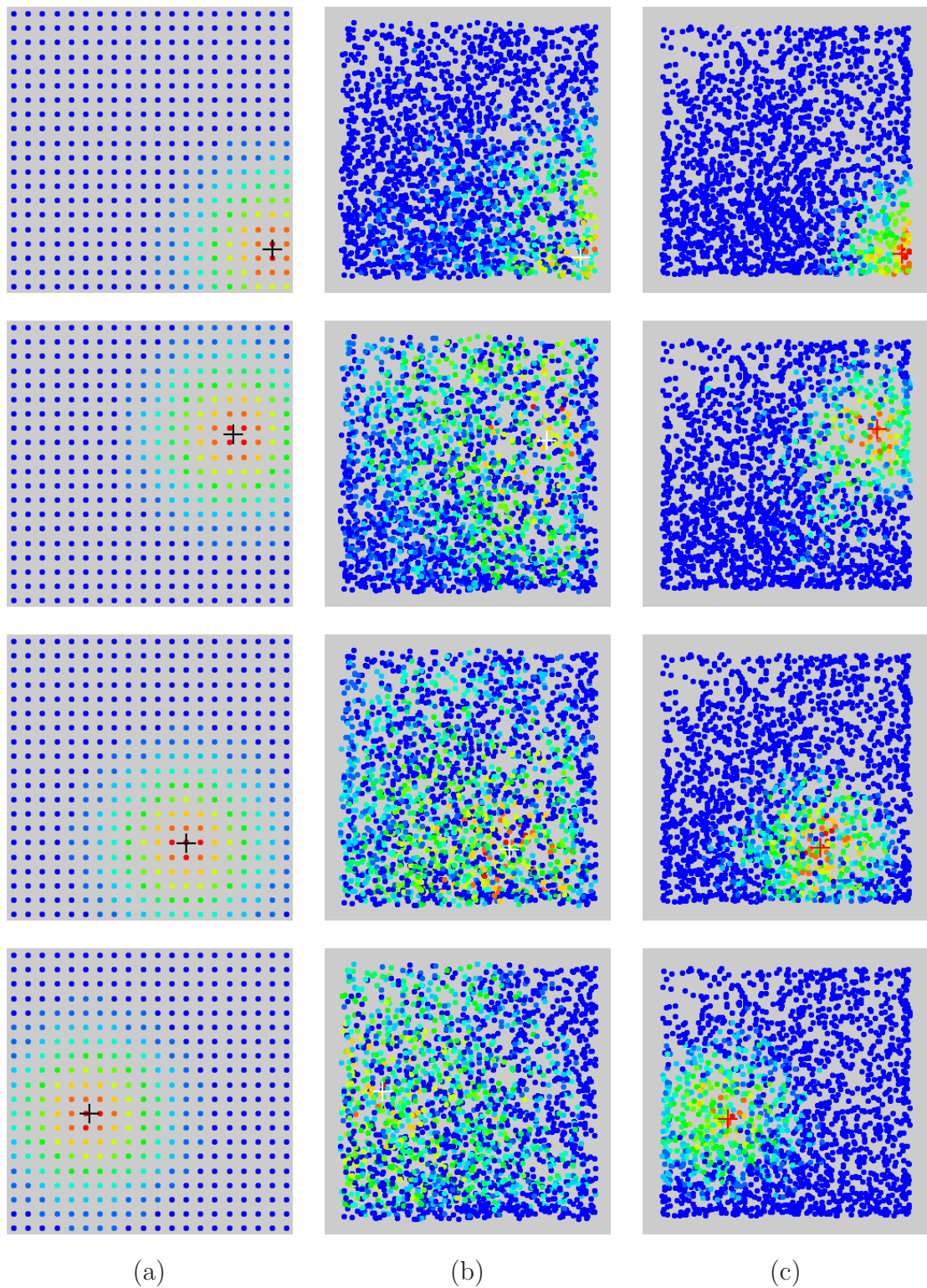


Figure 6.12: Population activities in a 770×770 mm environment. Each row shows the population activity at a specific time step. Crosses indicate the population vectors. (a) Idiopathic place code, (b) allothetic place code, (c) combined place code

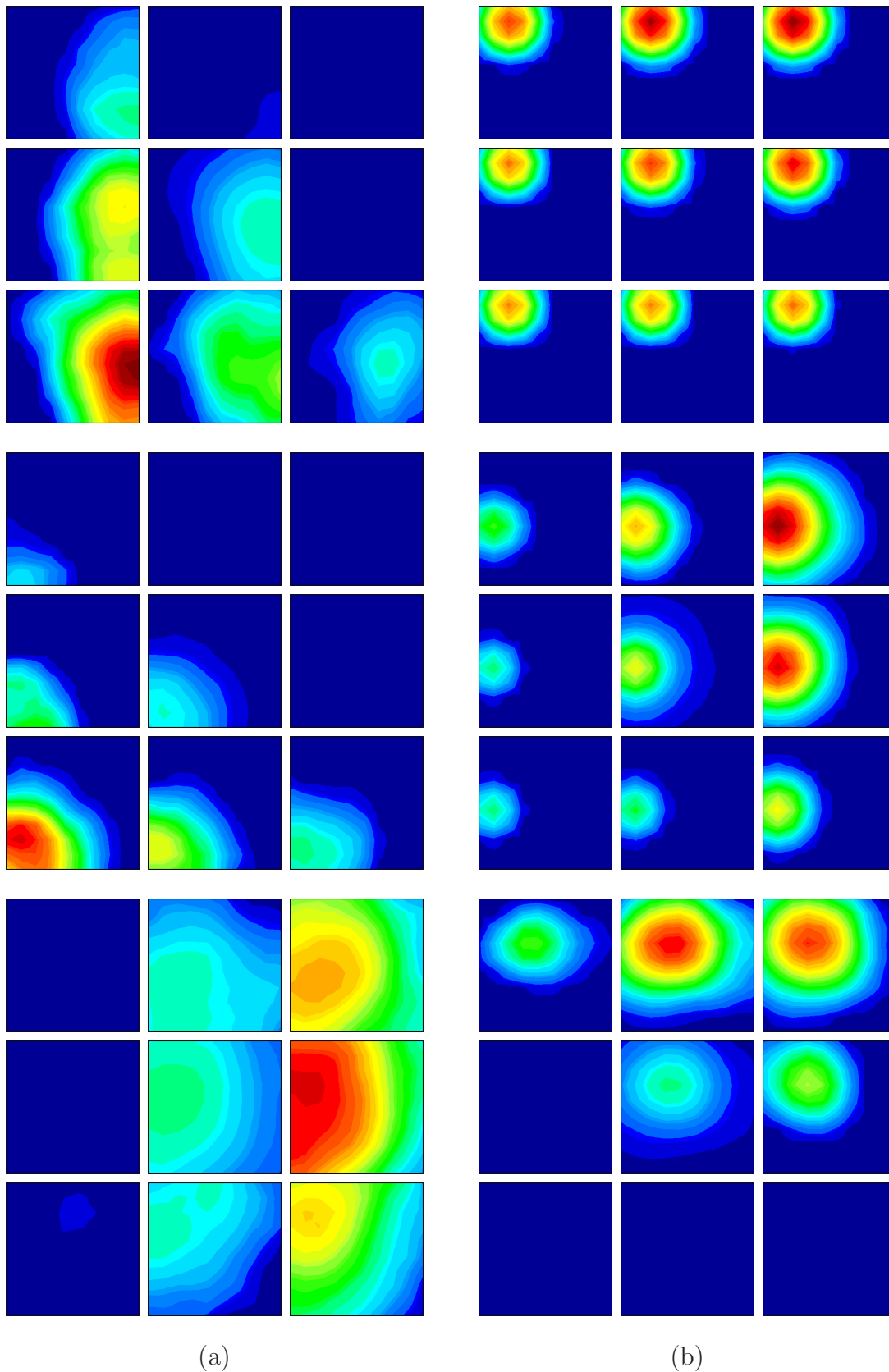


Figure 6.13: Directional receptive field contour plots. (a) Allothetic place cells (APCs). (b) Combined place cells (PCs)

Chapter 7

Towards a model of multimodal integration in the superior colliculus

When combining information from several sensory organs into a common spatial representation, each sensory modality must be assigned a certain level of influence on the combined representation. However, the relevance of information from a specific sensory organ depends on the current environmental conditions. An adaptive weighting scheme is needed to solve this multimodal integration task. In this chapter, a model is proposed to combine visual and “tactile” information into a common representation. It is an extension to the hippocampal place cell model described in chapter 6. It is implemented and tested on a Khepera mobile robot platform. The model is inspired by neuronal properties of the superior colliculus (SC), a layered midbrain structure. Its superficial areas are involved in oculomotor responses and the deep layers (dSC) process multimodal information (visual, tactile and auditory). In section 2.3.1, some single neuron and population properties, as well as behavioural implications of SC are reviewed. To summarise, it is noted that:

- dSC contains separate topological maps for each sensory modality.
- These maps are combined into multimodal sensory-motor maps.
- Most dSC input neurons are unimodal, whereas the majority of output cells show multimodal enhancement.

- Some SC neurons show spatial firing properties.
- The majority of SC neurons show theta field activity in situations where hippocampus also does.

We use some of these aspects in a simplified model of the deep layers of the superior colliculus. The model is rather abstract. It focuses on the “gating” mechanism which learns to assign appropriate weights to each sensory modality at each time step.

7.1 Architecture

The different modalities are combined in the allothetic place code layer (see section 6.1). As an example, we address the integration of visual and tactile signals, but the concept could easily be extended to other modalities. The weight of each sense is modulated by a gating network which learns to adapt the importance of each modality to the current environmental condition. Intermodal correlations are established using uni- and multimodal units inspired by neurons in the superior colliculus. Figure 7.1 shows the architecture of the system.

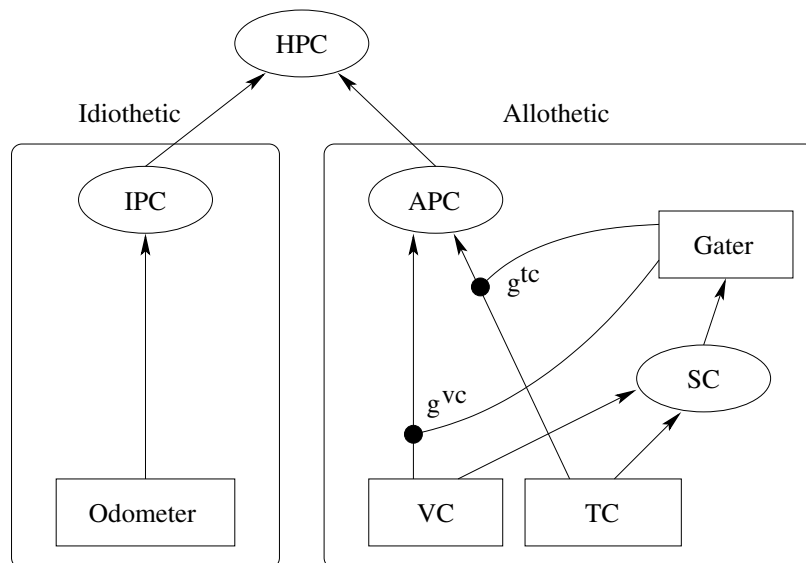


Figure 7.1: Architecture of our system for multimodal integration. Visual (VC) and tactile (TC) cells are combined in the allothetic place code (APC). Their influence is modulated by the output of a gating network. The superior colliculus (SC) population stores and compares intermodal correlations.

We use the simpler visual processing proposed by Arleo *et al.* [16] rather than the full system described in sections 5.2 and 6.2¹: Instead of the more or less heading-insensitive column difference cells, four orientation-sensitive view cells (VCs) code for the current place. In order to produce a heading independent representation, the four views are always taken in the four cardinal directions (i.e. north, south, west and east). This eliminates the problem of orientation versus position sensitivity in the visual processing. The idiothetic pathway is similar to chapter 6.

At each time step, the agent turns to each of the four cardinal directions. It recruits two sensory cells (a visual (VC) as well as a tactile (TC) cell) which store the preprocessed sample of the visual [16] or tactile (section 5.3) input. The recruitment method for APC neurons and VC→APC synapses follows equations 6.10 and 6.11. Synapses from sensory cells to APC neurons evolve using the Hebbian learning rule of equation 6.16. The same procedure is implemented one layer downstream in the PC population.

For reasons of convenience, we define $\tilde{w}_{ij} = w_{ij}/(\sum_k w_{ik})$ as the normalised weight from the presynaptic neuron j to the postsynaptic cell i . The firing rate r_i of HPC neurons i is given by the weighted mean activity of its presynaptic neurons j :

$$r_i = \sum_j \tilde{w}_{ij} r_j \quad (7.1)$$

In the APC population, however, the inputs from all presynaptic neurons of the same modality are modulated by a gating factor g^{vc} or g^{tc} which depends on the environmental conditions. This weighing factor is the output of the gating network (section 7.3):

$$r_i = g^{vc} \cdot \left(\sum_{j \in VC} \tilde{w}_{ij} r_j \right) + g^{tc} \cdot \left(\sum_{j \in TC} \tilde{w}_{ij} r_j \right) \quad (7.2)$$

7.2 Unimodal and multimodal cells

In the deep layers of superior colliculus (dSC), most of the sensory input cells are unimodal, whereas the majority of output cells responds to multiple sensory modes [270]. Our abstract model of dSC contain also follows this principle. Sensory cells (VCs and TCs) project to the input layer of dSC which consists of unimodal visual (UVs) and tactile (UTs) cells. Those unimodal cells project to multimodal cells (MMs) in the output layer of dSC. The architecture of the dSC model is shown in Figure 7.2 (a).

Unimodal input cells: Whenever the agent receives strong visual and tactile input simultaneously, a tactile and a visual unimodal cell are recruited. Connections

¹ This is mainly due to historical reasons: The results presented in this chapter are the earliest part of this work. We focus on the gating mechanism and not on the visual processing

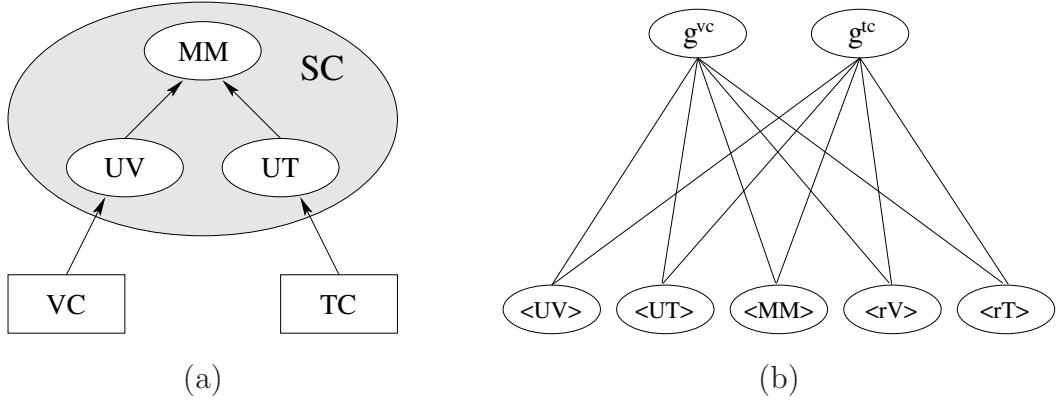


Figure 7.2: (a) Architecture of the deep superior colliculus model. VC: Visual cells, TC: Tactile cells, UV/UT: Unimodal cells, MM: Multimodal cells. (b) Architecture of the gating network. $\langle XY \rangle$: average firing rate of population XY . $\langle rV \rangle$: average pixel brightness of the input image. $\langle rT \rangle$: average activation of the proximity sensors. g^{vc}/g^{tc} : output gating neurons.

between active TCs and UTs are established and their weight fixed using the Hebbian learning rule of equation 6.16. Equivalently, VCs are connected to UVs. The firing rate of unimodal cells is given by the weighted mean activity of its presynaptic neurons (equation 7.1).

Multimodal output cells: The most active unimodal cells connect to a newly recruited multimodal cell and synaptic weights are again fixed according to equation 6.16. The firing rate r_i of multimodal cell i differs from equation 7.1 for unimodal cells in that both UTs and UVs need to be active to trigger the firing of a multimodal cell. This nonlinear interaction simulates the effect of *multimodal enhancement*, to which the majority of multimodal dSC cells seems to be subject [270, 306, 132]:

$$r_i = \tanh \left[k \left(\sum_{j \in UV} \tilde{w}_{ij} r_j \right) \cdot \left(\sum_{j \in UT} \tilde{w}_{ij} r_j \right) \right] \quad (7.3)$$

where k governs the amount of input drive necessary to saturate the neuron.

7.3 Learning the gating network

During the initial exploration of the environment, locations where multimodal stimuli are available are detected and the corresponding visual and tactile input is stored. Due to the multiplicative interaction between visual and tactile input in the activation

function of MM cells, inconsistencies can be detected. For instance, if a low obstacle which cannot be seen by the visual model but sensed by the tactile input is associated to a particular view, the absence of the obstacle activates corresponding UV cell, but both UT and MM cells are silent. In general, a missing feature of a learnt multimodal stimulus produces activity in unimodal, but not in multimodal cells.

In a second phase, a gating network learns to modulate the importance of vision and touch according to the current environmental conditions. The gating network consists of five input neurons which are fully connected to two output neurons as shown in figure 7.2 (b). Each input neuron codes for the average activity of a population. For example, $\langle UV \rangle$ is the average activity of UV cells and so on. $\langle rV \rangle$ is the average pixel brightness of the raw (unprocessed) camera image. $\langle rT \rangle$ is the average proximity sensor input. A high value of $\langle rT \rangle$ means that many obstacles are very close.

The output neurons g^{vc} and g^{tc} provide the gating values of equation 7.2. Their firing rate is calculated in two steps: First, the weighted sum h^{vc} and h^{tc} of their input activity is determined according to equation 7.1. In order to keep the total input strength constant, a soft-max operation follows which keeps the sum of the firing rates g^{vc} and g^{tc} constant:

$$g^{vc} = \frac{h^{vc}}{h^{vc} + h^{tc}} \quad \text{and} \quad g^{tc} = \frac{h^{tc}}{h^{vc} + h^{tc}} \quad (7.4)$$

The task is to adapt the synapses of the network such that the sensory modality which is most useful for spatial coding is enhanced while the other is suppressed.

The weights between input neuron j with firing rate r_j and gating neuron i with firing rate i are updated according to a Hebbian-type rule, but the synaptic modification Δw_{ij} is also modulated by a reward signal q :

$$\Delta w_{ij} = q \cdot \eta \cdot r_i (r_j - w_{ij}) \quad (7.5)$$

The reward q depends on two properties of the allothetic place code: (a) variance σ around centre of mass and (b) the number $n = \sum_k \mathcal{H}(r^{apc} - \theta_a)$ of APC neurons which are more active than $\theta_a = 0.8$. Positive reward is given if many cells are active and the place code is compact (i.e. small variance σ). However, if the variance is high, the reward still needs to be positive, because the tactile sensory input always produces high variances due to the high ambiguity in the sensory input. An insufficient number of active cells, however, always results in a negative reward. The equation for the reward q is:

$$q = \tanh[\alpha_a \cdot (n - n_0)] \cdot s_\sigma \tanh[-\alpha_\sigma (\sigma - \sigma_0)] \quad (7.6)$$

The first factor evaluates amount of activity in the APC layer. If there are more than $n_0 = 3$ active APC, the reward is positive, else it is negative. $\alpha_a = 0.5$ determines

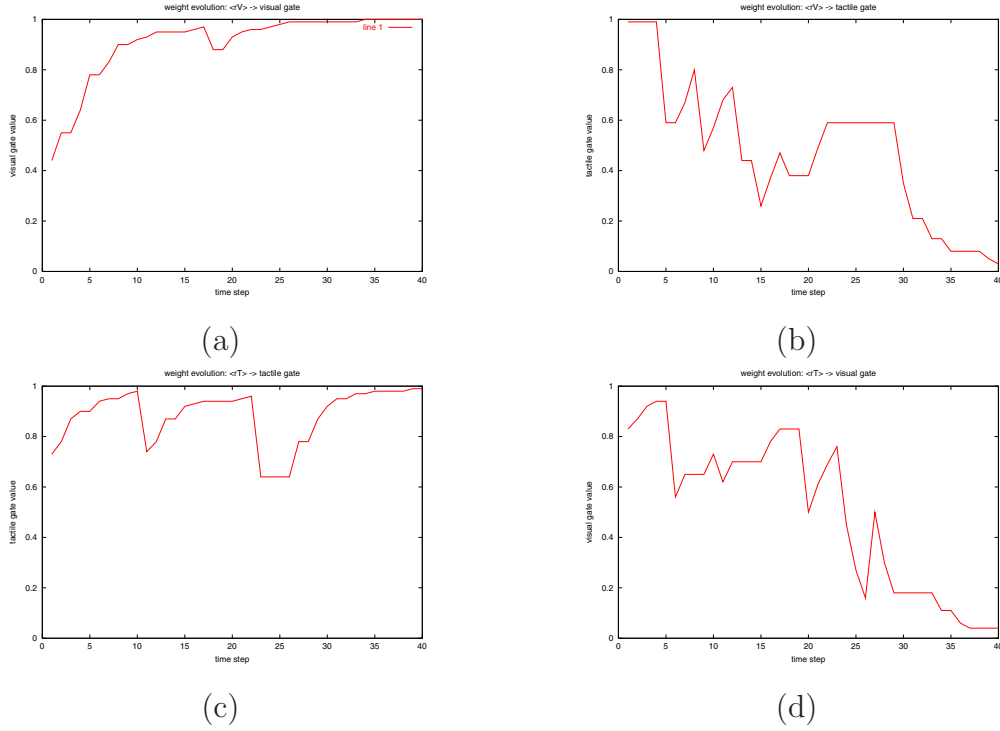


Figure 7.3: Evolution of the gating weights for the raw sensory inputs $\langle rV \rangle$ and $\langle rT \rangle$ to the output gating neurons g^{vc} and g^{tc} . (a): $\langle rV \rangle \rightarrow g^{vc}$, (b): $\langle rV \rangle \rightarrow g^{tc}$, (c): $\langle rT \rangle \rightarrow g^{tc}$, (d): $\langle rT \rangle \rightarrow g^{vc}$.

the slope of the nonlinear curve. It is set so as to saturate to a reward of ≈ 1.0 if there are more than 10 active APC neurons.

The second factor evaluates the variance σ of the allothetic place code around its centre of mass. The parameters $s_\sigma = 0.25$, $\alpha_\sigma = 0.1$ and $\sigma_0 = 50\text{cm}$ are set such that a normal APC variance of 20cm produces a reward of ≈ 1.0 whereas if a variance of 80cm, which is unusually high, results in a value of ≈ 0.5 .

Experiments are conducted on a Khepera mobile robot. An $60 \times 60\text{cm}$ boarded arena placed on a table in a normal office serves as environment. A large rectangular-shaped object is placed in the arena to increase the amount of tactile input to the system. After the environment is explored, a learning phase for the gating network begins. The weights of the network are randomly initialised prior to learning.

During learning, the agent moves randomly in the explored environment and tries to localise itself. At each time step, the illumination in the environment is either artificially turned down to 20% with probability P_L or left at its normal value otherwise.

The pixels of the raw camera image are scaled down, but Gaussian noise is also added to simulate the reduced quality of visual input in the dark. The Gabor filters used to process the image are much more sensitive to edges than they are to luminosity. This means that with the light turned down, the filter activity might still be high, but it now only inaccurately represents the image due to the increased noise.

At each time step, the weights are then updated according to equation 7.5. Figure 7.3 shows the evolution of the weights between $\langle rV \rangle$ as well as $\langle rT \rangle$ and the output gating neurons g^{vc} and g^{tc} . As expected, high luminance in the input image activates the visual gate, whereas high values of the proximity sensors contributes to the tactile gate. Conversely, the luminance does not influence the tactile gate and proximity sensors don't contribute in activating the visual gate.

Once the gating neurons are tuned, two test-runs are performed: One in normal lighting conditions, the other with the light dimmed down to 20%. The gating values are for both tests are plotted against time.

Figure 7.4 (a) shows the gating values for the visual and tactile senses in the lit environment. More than half the time, visual input is the only activated modality. Each time the robot is near an obstacle however, the tactile sense is assigned a slightly higher importance than vision. The abrupt changes are due to the binary nature of the tactile sensors.

Figure 7.4 (b) shows the gate values when the illumination is reduced by 80%. Most of the time, vision and tactile senses receive equal importance. Due to the soft-max operation, both values are set to 0.5. Whenever an obstacle is near, however, the agent relies mostly on its tactile input. Again, the abrupt changes show that the tactile sensors either don't sense an obstacle or they are saturated by a close object. Indeed, as illustrated in figure 5.11 on page 64, the proximity sensors only have an effective range of only $\approx 2\text{cm}$.

Learning has been performed by dimming the light down to 20%. In figure 7.5, it is illustrated that the activity of the visual gating neuron g^{vc} also generalises to other illumination conditions.

The main difficulty in learning the importance of sensory input lies in determining the reliability and uncertainty of a percept. We use the average place cell activity and the activity variance around the centre of mass as a quality measure to change the weights of the gating network. However, accessing the variance in spatial representations might be difficult to motivate biologically.

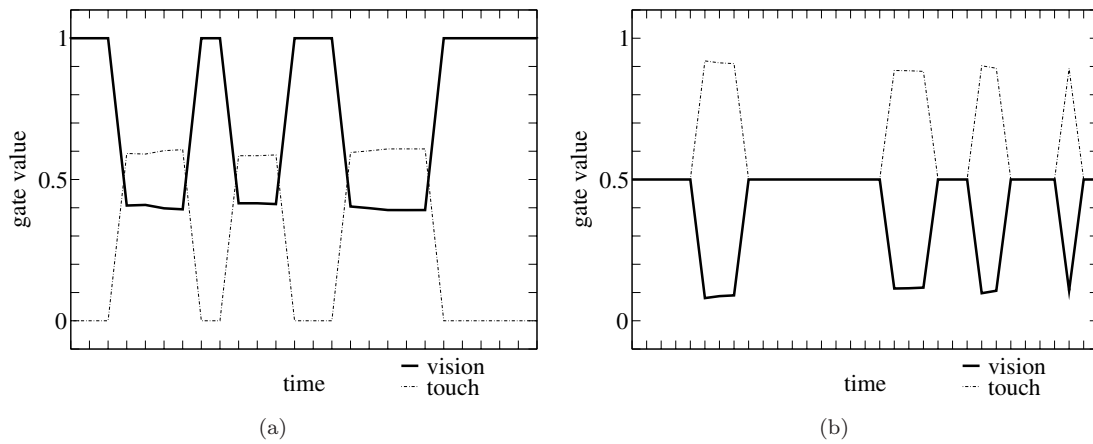


Figure 7.4: Gate values in open-field and border positions. (a) good illumination. (b) almost no light

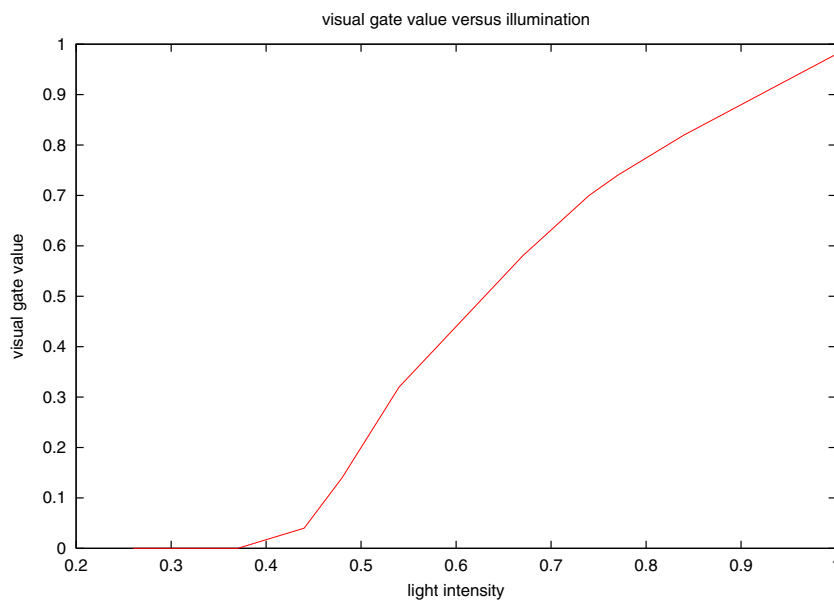


Figure 7.5: Activity of the visual gating neuron g^{vc} versus illumination. During learning, the illumination was either normal (100%) or reduced to 20%. However, g^{vc} generalises to intermediate illumination conditions.

Chapter 8

A new locale navigation model

In this chapter, a new model of rodent navigation is presented. In particular, we propose a *locale navigation* model according to the taxonomy introduced in section 3.3. This type of navigation allows the animal to learn to find a stable but hidden reward location in an environment. The navigation system makes use of the spatial representation learnt according to chapter 6.

The hippocampus (HPC) is one of the main afferent structures of the nucleus accumbens (NA), the ventral part of the striatum (see basal ganglia review in section 3.1). HPC neurons project to NA via the fornix fibre bundle. Lesions of the fornix or nucleus accumbens impair the rats in the hidden, but not the visible water maze task [76, 279, 209, 318], whereas lesions in the caudate-putamen, which forms the dorsal part of the striatum, don't show an effect. If the platform is visible or cued, however, the reverse is true: NA lesions don't show an effect and caudate lesions result in a decrease in performance [209, 161, 66]. These results suggest that NA is involved in locale navigation and caudate-putamen is not.

The ventral tegmental area (VTA), which—like NA—is a part of the basal ganglia, contains dopaminergic neurons (DNs). On top of a regular tonic firing pattern, DNs also show transient phasic activity. This phasic firing seems to be related to the processing of reward signals. In particular, there is evidence that DNs code for the difference between actual and expected reward, and thus implement the temporal difference error of the reinforcement learning paradigm (chapter 4) [122, 248, 244, 12, 247, 304, 245].

DNs of VTA also project to nucleus accumbens [320, 247, 149]. They tend to synapse on the basal part of spines which host synapses from cortical [92, 252, 266] as well as hippocampal [295, 251, 265, 87] afferent projections. It is therefore possible

that DNs modulate either synaptic transmission or synaptic plasticity of cortical and hippocampal afferent connections to striatum [122, 248, 244, 275, 276, 277].

Our model of locale navigation is based on reinforcement learning (RL, cf. chapter 4) applied to HPC→NA synapses. In particular, we focus on a mechanism by which states as well as actions are treated as continuous variables. Learning quickly generalises in both spaces. RL has previously been used to solve navigation tasks for autonomous mobile agents [39, 14, 88, 16, 65]. Some models operate in continuous state and/or action spaces using function approximation [14, 240, 69, 65]. However, we are not aware of any neural model of locale navigation where both state and action spaces are continuous.

In most RL-based models, an eligibility trace [280] is used. While such traces of past state-action pairs result in a considerable learning speed-up, it is not clear if and how such a memory is available to the brain. We discuss some possible mechanisms for eligibility traces, which allow rewarding events to generalise back in time.

8.1 Architecture

In this model, a spatial representation based on place cells and constructed according to chapter 6 serves as state space. It consists of a population of hippocampal place cells (PCs) with highly overlapping receptive fields. Here we focus on the use of this representation for navigation. PCs project onto a population of action cells (ACs). As the nucleus accumbens (NA) is (i) a target of the hippocampus, (ii) receives dopaminergic projections and (iii) is related to motor behaviour, we postulate NA as the neuronal substrate for ACs. A navigation map is constructed using reinforcement learning on PC→AC synapses.

Reinforcement learning has been used for problems where a small discrete set of actions is available at each state. The number of states is usually discrete and finite. The population vector of PCs, however, can be interpreted as the continuous state variable s which represents the agent's location $\mathbf{P} \in \mathbb{R}^2$ in the environment [14]. Similarly, the population vector Φ^{ac} of the AC layer stands for the direction of the next movement and represents a continuous action a . The model is tested on a real and simulated mobile robot platform (see chapter 5). The architecture of the navigation system is shown in figure 8.1.

8.2 Action cells

A population of $N_{ac} = 120$ action cells (ACs) code for the motor-commands of the agent's next movement. Each AC i represents a particular compass heading ϕ_i , which are uniformly distributed between 0 and 2π . In each time step, the agent gathers

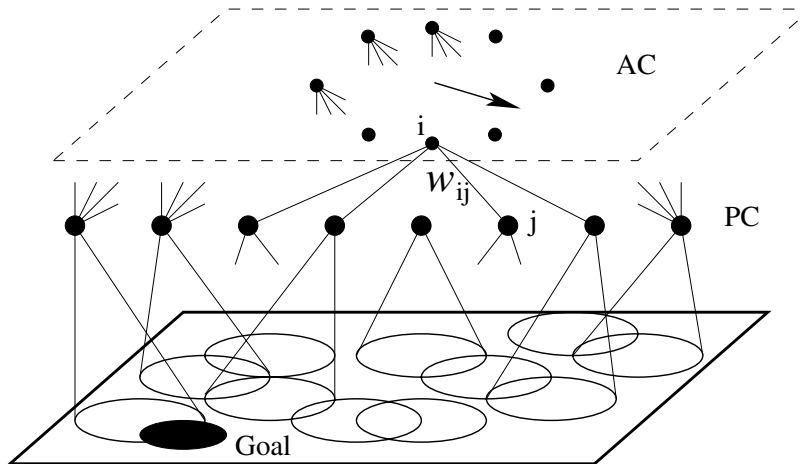


Figure 8.1: Architecture of our navigation system: A layer of hippocampal place cells (PCs) represent the environment. Each PC is active in a region of the environment and their receptive fields overlap. In order to learn to navigate to the goal, PCs are connected to action cells (ACs) which code for the direction of the next movement. The population vectors of both layers allow a continuous interpretation of position (PCs) and direction of movement (ACs).

sensory information which activates the simulated hippocampal place code (PC) layer. Based on PC activity, the AC module selects the next action.

Each action consists of two movement commands: (i) An in-place rotation by an angle θ which orients the agent towards the allocentric orientation indicated by the population vector Φ^{ac} . (ii) A forward movement of a fixed distance d or until an obstacle is blocking the way. The step size d is chosen so as to emulate a running rat. The rate at which the rat processes spatial information is assumed to be related to the theta rhythm (6–12Hz, see section 2.1.2). At each cycle, the rat “senses” the world and reacts appropriately. We interpret one time step of the model as one theta cycle, corresponding to a sampling rate of 8Hz [254, 41, 290, 14]. For a running speed of 48cm/s, we get $d = 48/8 = 6\text{cm}$.

The AC population vector Φ^{ac} is calculated according to equation 6.8. It determines the allocentric direction of the next movement. The interpretation of this heading is tied to the population vector of the head direction system (section 6.3), which defines the allocentric angular frame of reference. The rotation angle θ is the angular difference between the population vectors of the HD and AC layers:

$$\theta = \Phi^{hd} - \Phi^{ac} \quad (8.1)$$

The activity of ACs is calculated in two stages. In each stage, ACs code for a different property:

Action-evaluation: First, each AC i receives state information from all PCs j and learns to attribute a value to each action. This tells the agent which actions are good in the current state s . The input potential

$$h_i = \sum_j w_{ij} \cdot r_j \quad (8.2)$$

to AC i represents the estimated value $Q(s, a_i)$ for the current state s and action a_i . In traditional Q-learning, an optimal action is a discrete action whose Q-value is bigger or equal than all other action values. Here, we take a slightly different approach: In contrast to most other models, we don't use the *max*-operator to determine the optimal action. Instead, we use the direction of the AC population vector Φ_h^{ac} (equation 6.8). Φ_h^{ac} represents the continuous action a^o which supposedly maximises the return (i.e. the total discounted future reward), given the current estimation of Q-values. This is called the "greedy action" because it exploits the current estimation of Q-values instead of trying to improve the estimations by exploration. Therefore, we sometimes need to take non-greedy actions. This action selection method is described in section 8.3.2. The optimal action a^o is a continuous variable. The Q-values, however, are only estimated for the discrete set of ACs, ie. $Q(s, a^o)$ is not directly accessible. It is calculated by linear interpolation of the Q-values of the two nearest discrete actions.

Generalisation: As soon as an action is selected, a generalisation mechanism is applied: A Gaussian AC activity profile with variance σ_{ac}^2 is enforced around the selected action a^x (direction Φ^{ac}). If $\Delta\phi_i$ stands for the angular distance between Φ^{ac} and ϕ_i , the profile can be expressed as

$$r_i = \exp(-\Delta\phi_i^2/2\sigma_{ac}^2) \quad (8.3)$$

The firing rates r_i of the AC layer then represent the action which was selected for execution. Recurrent connections within nucleus accumbens or via another population could be responsible for the formation of this "blob" of activity. In most neuronal RL models, the update of the Q-value estimation is proportional to the activity of action cells. Traditional RL models employ a winner-take-all mechanism which inhibits all non-selected actions. Only the winner neuron may refine the estimation of its Q-value. In our model, however, many action cells are active and thus eligible for learning. This results in a generalisation of learning in action space.

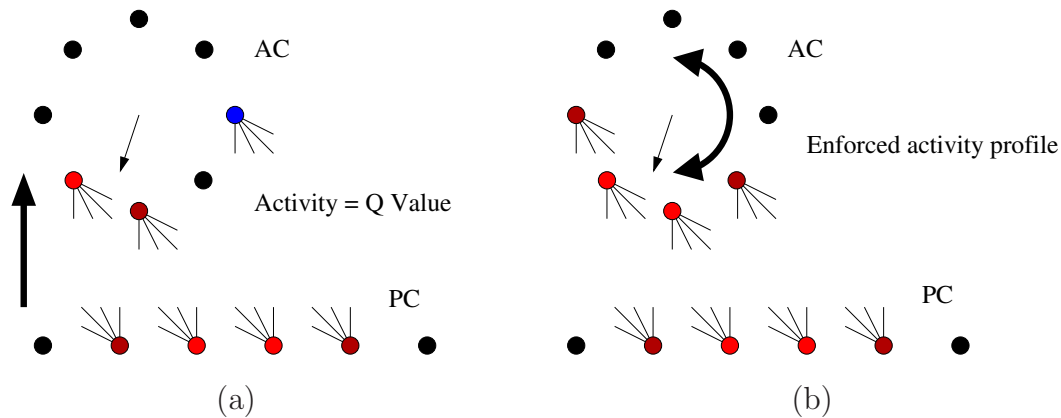


Figure 8.2: (a) Action evaluation stage: The activity of action cells represent the Q-values of the corresponding directional movement. It only depends on the current place cell activities and is purely feed-forward. The population vector (thin arrow) points in the direction of the estimated optimal action. (b) Generalisation stage: An action has been taken (in this case, the optimal action). Then, a Gaussian activity profile is enforced around the selected action. This profile allows not only the selected—but also similar actions to learn and forms the basis of the generalisation capability in action space

Figure 8.2 illustrates the two stage process. In the evaluation stage, which is a purely feed-forward operation, only the current place cell activities drive ACs. The current estimation of action values, stored in the synaptic efficacies of the PC→AC connections, supports the selection of an appropriate next action. Once this decision is taken, the lateral interaction becomes effective and drive ACs above firing threshold. In the example, the “greedy” action is taken.

The separation of the two stages of AC processing is subject to the following constraints: First, an action selection system has to be given enough time to read the optimal action Φ_h^{ac} in the evaluation stage. Secondly, in an implementation of the learning rule with spiking neurons, the firing of ACs must closely follow action potentials emitted by PCs such that the pre- and postsynaptic spikes fall into the timing window of spike timing dependent plasticity (see next section).

These constraints require a precise timing of the system. It has been suggested that different phases of theta activate different processing stages in hippocampus [140, 114, 112]. Similarly, we propose that the theta-rhythm could provide a separation of the two stages in nucleus accumbens: First, during the late phase, low theta-activity allows hippocampal place cells to fire and thus pass spatial information to nucleus accumbens. The pathway of a recurrent loop within NA or involving another area

is disabled at this time, so as not to interfere with the estimation of Q-values. At a later phase, when an action is selected, the recurrent loop shapes the activity profile of the generalisation stage.

8.3 Learning and eligibility trace

Rats quickly learn to navigate to a rewarding location from any place in the environment. If, in the framework of reinforcement learning, the reward is solely given in the goal state, the learning procedure is slow because only the last state before the delivery of the reward may update its estimation of action values. Thus, information only slowly propagates back to states further away from the goal state. As a work-around, a decaying trace of recently active state/action pairs can be kept. Rewards are then not attributed only to the most recent state/action pair, but to the whole trace. In reinforcement learning, such a memory¹ is termed *eligibility trace* (ET), because it keeps track of all state/action pairs which are eligible for learning. It enables a generalisation of knowledge back in time and tremendously speeds up reinforcement learning algorithms [280].

8.3.1 Neurophysiological evidence

The quick learning often observed in animals is an indication that they are equipped with effective generalisation mechanisms. It is, however, unclear, whether or not animals use ETs. Synaptic plasticity, and in particular long term potentiation (LTP) and -depression (LTD) has been suggested to be the neuronal mechanism for learning [17, 18, 32]. Experiments show that repeated application of a spike pairing protocol, where presynaptic action potentials are closely followed by postsynaptic spikes result in a potentiation of the synapse. Conversely, the synapse's efficacy is decreased if the postsynaptic neuron repeatedly fires just before the presynaptic neuron [156, 27, 151, 250, 223, 102]. This spike-timing-dependent plasticity (STDP) may be the neuronal protocol to induce LTP/LTD.

If the neuronal mechanisms underlying ETs were also located at synapses, three requirements must be fulfilled: (i) They must contain a decaying short-term memory which stores recent activation, (ii) they induce LTP or LTD the magnitude of which is modifiable by some reward signal, (iii) the modulation of the induced LTP/LTD must be possible *after* the synapse has been activated. In the following, experimental results that are consistent with the idea that synapses can implement ETs are discussed.

Synaptic memory: During LTP/LTD induction, Ca^{2+} ions enter the postsynaptic cell through NMDA channels [17, 32, 18, 27]. It has been proposed that the pro-

¹In the terminology introduced in section 3.3, this is a short-term memory

longed change in calcium concentration after synaptic transmission could serve as ET [322, 244]. More recently, it has been shown that dendritic depolarisation after incoming spikes can be sustained for hundreds of milliseconds and remains local to the dendritic branch [243, 316]. This sustained depolarisation could also mark the synapse as “recently active”. Calmodulin-dependent protein kinase (CaM-k)² is activated by Ca²⁺ and is suspected to remain in an activated state for several hundreds of milliseconds [122], which could implement the required STM. Finally, a short-term facilitation has been observed after synaptic stimulation. This facilitation also lasts for ≈ 100 ms and is suggested to be caused by a higher presynaptic neurotransmitter release [303, 299, 67, 337].

modulation of LTP/LTD: Dopamine seems to be an important factor which determines the magnitude of synaptic potentiation in LTP induction experiments. In striatum, prefrontal cortex and hippocampus, dopamine agonists enhance synaptic potentiation whereas antagonists impair potentiation [206, 109, 138]. In most of these experiments, the static level of dopamine concentration is altered several minutes before the LTP induction protocol begins. This protocol therefore emulates the *tonic* activity of dopaminergic neurons. However, it is the transient or *phasic* activity which seems to correspond to the error in reward prediction [246, 248, 244, 84]. In recent experiments, dopamine neurons in substantia nigra are electrically stimulated in order to simulate a phasic activity. This results in potentiation of cortico-striatal [323, 231] projections. Prefrontal and auditory cortex also show potentiation if a phasic dopamine release is simulated [21, 33].

modulation of previous LTP/LTD induction: The process of *consolidation* is responsible for fixating a memory trace. At a neuronal level, it requires the expression of genes and the synthesis of proteins [62, 73, 147, 74]. Consolidation operates at a timescale of minutes or hours [147]. During this period, the synaptic potentiation is vulnerable. A cascade of molecular events change the properties of the synapse and a disruption of any stage may block the consolidation of LTP [104, 2, 201, 73]. In most dopamine-dependent plasticity experiments, the concentration of dopaminergic agents is altered several minutes *before* the LTP induction protocol begins. Nevertheless, dopamine administration *after* LTP induction has no effect on LTP [206]. This seems to be in contradiction to the hypothesis of a delayed rewarding signal in reinforcement learning. However, low-frequency stimulation after LTP induction—which normally results in a depotentiation of the synapse—is effectively blocked if dopamine is applied after LTP induction [207, 144, 154]. Thus, many synapses could be potentiated during an experiment, but only those that receive later a dopaminergic “confirmation” signal remain strengthened.

Instead of being a synaptic process, the ET could also be the result of a network effect. Sustained neural discharge could store the trace of recent state/action pairs.

²A postsynaptic protein involved in the induction (but not maintenance) of LTP [205]

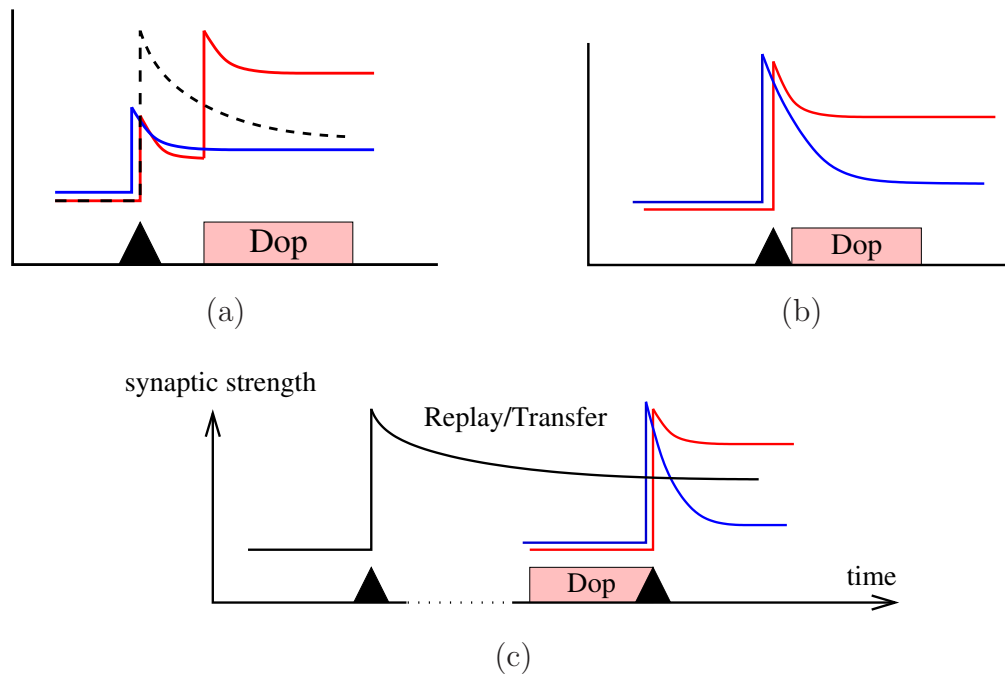


Figure 8.3: Schematic illustration of possible mechanisms for eligibility traces. Black triangle: LTP induction. Red curves: LTP with dopamine. Blue curves: LTP without dopamine. (a) *Synaptic memory*: A synaptic short-term memory “flags” recently active synapses (dashed curve). Dopamine enhances LTP at flagged synapses. (b) *Inhibition of depotentiation*: Per default, all active synapses are potentiated. However, if dopamine is not administered after induction, the synapse is depotentiated. (c) *Replay/Transfer*: Dopamine might act on a “replay” of the original synaptic activity, which may transfer relevant memories to other networks. Dopamine is then available *before* the replay is activated and thus standard dopamine dependent plasticity applies.

Graded sustained activity—proposed to implement short term memory—has been found in entorhinal cortex [75, 140] and in striatum [246]. Its time-span is in the order of seconds. For this implementation of short term memory, it is not a problem that dopamine does not modulate LTP if applied *after* initial LTP induction, because the neuronal activity is still present when the reinforcement occurs. Finally, it is also suggested that learning does not occur on the synapses of the original activity, but rather on a “replay” of this activity which transfers the memory from one network to another [45, 207]. As this replay may *follow* the reward administration, this idea is consistent with the dopamine-dependent plasticity experiments.

To summarise, there are three main proposals on how an eligibility trace can be implemented in the brain. (i) Biochemical markers may “flag” recently active synapses.

When a reward-related learning signal arrives within about 100ms, the synapse may be potentiated more than without the reward signal (figure 8.3 (a)). Instead of a chemical marker, sustained decaying neuronal activity may serve as short term memory which could span seconds. (ii) Synapses are potentiated by default. If, however, a dopamine signal fails to arrive within a vulnerable period of seconds or minutes, the synapse is likely to be depotentiated (figure 8.3 (b)). (iii) Reward-based learning does not operate on the synapses that were initially active, but rather on a “replay” of this activity, which may also transfer the memory to another network (figure 8.3 (c)).

In our model of locale navigation, an eligibility trace is assumed to operate on PC→AC synapses. We propose that ETs are a synaptic process rather than a network effect. The time scale is in the order of seconds. This is not consistent with the prolonged postsynaptic calcium Ca^{2+} or CaM-k concentration, which only lasts for about 100ms. We therefore propose that the ET in our system is the result of dopamine-dependent depotentiation inhibition outlined above [207,144,154]. Furthermore, we extend the effect on LTD: If the dopamine level following synaptic activation is decreased, the synapse will be depressed rather than potentiated. The trace p_{ij} on the synapse from PC j to AC i decays exponentially in time. It is formally expressed as:

$$p_{ij}(t) = \alpha \cdot p_{ij}(t - 1) + r_i(t) \cdot r_j(t)$$

where r_j and r_i are the pre- and postsynaptic rates and $\alpha = \gamma \cdot \lambda$ is the time constant which is composed of a future discount rate $\gamma = 0.95$ and a trace decay factor $\lambda = 0.88$ (see equation 4.3. If γ is set too small, the system could suffer from a distant reward problem. The rate we use makes sure that Q-values are never discounted more than 70% during the traversal of the entire test environment. λ is set such that the time scale of the ET is one second, i.e. eight time steps or theta cycles.

Eligibility traces can be interpreted as a generalisation mechanism in time. Unexpected rewards induce an update of the value functions also for previously visited states. Due to the distributed coding scheme for states and actions in our model, learning is also generalised in space, even if ETs were not used. The large tuning curves of place cells (figure 6.13 (b)) and dense coverage result in highly overlapping receptive fields. At each position, many place cells are activated (figure 6.12 (c)) and thus eligible for learning. The large Gaussian activity profile enforced in the action cell layer of nucleus accumbens serves the same purpose. With each selected action, all similar actions are also eligible for learning. While the high overlap in state space is due to similarities in the acquired sensory input, i.e. it is purely feed-forward, the overlap in action space requires lateral interactions between action cells.

This temporal and spatial generalisation is illustrated in a visualisation of the eligibility trace (figure 8.4). For this graph, a set of perfectly tuned Gaussian place

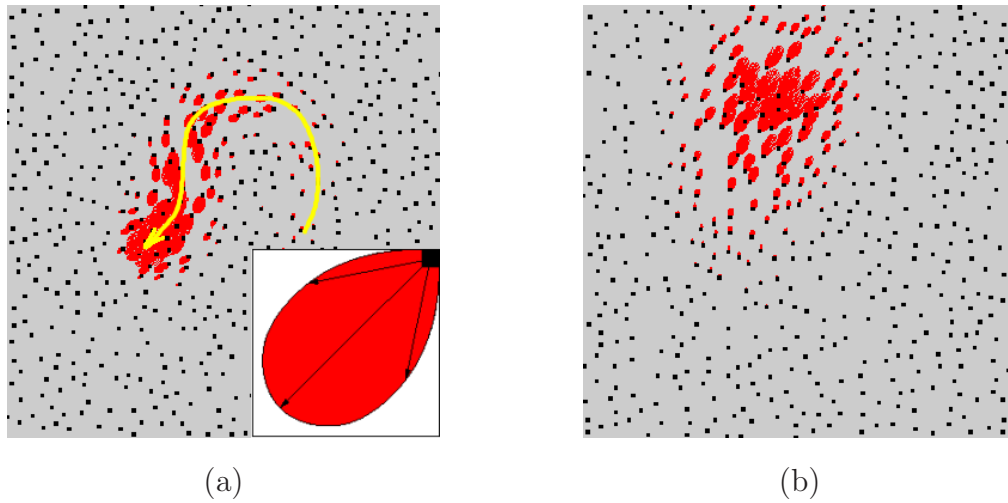


Figure 8.4: Two examples of the modelled eligibility trace. The yellow curvy arrow shows the agent’s path. Each black point represents the receptive field centre of a perfectly Gaussian place cell (variance σ^2). The red drop-shaped form visualises the eligibility trace (time constant α) memorised for each place cell. The inset shows a schematic example of a memorised south-west movement. Due to the generalisation in action space, similar actions like south-southwest are also eligible for learning. (a) long and narrow trace ($\sigma = 4\text{cm}$, $\alpha = 0.88$). (b) short but wide trace ($\sigma = 8\text{cm}$, $\alpha = 0.56$).

cells is generated. Each place cell is represented by a black dot at the centre of its receptive field. The synaptic memory from place cell j to action cell i is visualised as a red line, starting at the place cell. The line direction corresponds to the direction of movement ϕ_i associated with cell i . The inset is a schematic enlargement of the trace for a place cell which memorises a south-west movement. The drop-shaped ET means that, for instance, the action cell coding for south-south-west is also eligible for learning. The line length represents the strength p_{ij} of the memory trace. In the example to the left, the agent’s path is drawn as a yellow curve. Here, the place fields are small ($\sigma = 4\text{cm}$) but the trace is long ($\alpha = 0.88$). To the right, the width of the Gaussian place fields is enlarged ($\sigma = 8\text{cm}$) and the trace shortened ($\alpha = 0.56$). This results in a wide generalisation in space, but not time.

8.3.2 Learning algorithm

This section assembles the ideas from above and describes the procedure we use to learn a navigation map. Similar to animal experiments, the agent is first allowed to explore the arena. During this period, no reward is given. Once the environment is

fully explored, i.e. place cells densely cover the whole surface, a set of reward training trials follows. An invisible reward is placed at a fixed location. At the beginning of each trial, the agent is inserted into the environment at a random location and with random heading. The trial ends when the agent finds the reward. During these trials, the agent learns a navigation map.

As in all temporal difference learning rules, the environment must be sensed before and after each action in order to update the estimation of value functions. These sensory input values stem from two adjacent time steps. Here we present the algorithm from the beginning of a learning iteration to its end, not from one time step to the next:

First, the action values $Q(s(t-1), a_i)$, i.e. the input potentials h_i to all action cells i are calculated. Next, a continuous action $a^x(t-1)$ is selected. Most of the time, the optimal action $a^o(t-1)$, i.e. Φ_h^{ac} is chosen. With a small probability $\epsilon = 0.2$, however, the new direction of movement is drawn randomly from a Gaussian distribution with variance $\sigma_x = 30^\circ$ around the current heading. This ϵ -greedy policy with constant ϵ balances exploration versus exploitation [280]. The action selection process is assumed to operate on a slower time scale than the processing of sensory input. The decision to either explore or exploit is only taken every fourth time step, i.e. twice per second. Then, the AC activity profile is enforced around the selected action $a^x(t-1)$ and the eligibility trace $p_{ij}(t-1)$ is updated. Here, the time step ends ($t-1 \rightarrow t$). In the beginning of the new time step, the agent receives the immediate reward $R(t)$. The agent processes its input, i.e. the place cell activities and $h_i(t)$ are updated. The standard reward prediction error $\delta(t)$ for Q-learning is calculated. Finally, the PC \rightarrow AC synaptic weights w_{ij} from place cells j to action cells i are modified using the standard RL update rule with learning rate $\eta = 0.001$. The following list briefly summarises these steps:

1. Calculate action values: $Q(s(t), a_i) = h_i(t)$.
2. Select action: $a^x(t) = a^o(t)$ with probability $1 - \epsilon$ (exploitation) or randomly select action with probability ϵ (exploration).
3. Generalise in action space: Lateral connections impose activity profile $r_i(t)$ around the selected action $a^x(t)$.
4. Update eligibility trace $p_{ij}(t)$.
5. $t \rightarrow t + 1$. Calculate $h_i(t + 1)$.
6. Calculate reward prediction error:

$$\delta(t + 1) = R(t + 1) + \gamma \cdot Q(s(t + 1), a^o(t + 1)) - Q(s(t), a^x(t)).$$
7. Update synaptic strengths: $\Delta w_{ij}(t + 1) = \eta \cdot \delta(t + 1) \cdot p_{ij}(t)$.

The reward prediction error $\delta(t)$ can be interpreted as the output of dopaminergic neurons in the ventral tegmental area. One problem which is not addressed here is how dopaminergic neurons can process information coming from different time steps. We assume that dopamine neurons receive action values via a direct and an indirect pathway, which is consistent with basal ganglia anatomy [3, 265, 122, 175, 267] (cf. section 3.1).

One problem of RL is “the curse of dimensionality”. When the state- and action spaces are large, learning all parameters is very slow. In our case, we have ≈ 1000 place cells and ≈ 100 action cells. However, these state variables are not uncorrelated. Due to the high overlap, learning quickly generalises in both spaces. The size of a place cell’s receptive field has a physical meaning (a portion of the environment) whereas in action space, the enforced activity profile has a width $\sigma = 30^\circ$ which represents a range of headings. Both sizes are independent of the number of cells, and therefore, the learning speed is independent of the number of cells.

8.3.3 Validation in test environments

The model is tested in all simulated test environments. A circular goal area of 7cm diameter and value $R_g = 15$ is the only rewarding location. However, each time the agent hits the wall, it receives a punishment of $R_w = -5$. This results in an obstacle-avoidance behaviour.

Each trial starts with a random heading at a random location whose distance to the goal is at least 20cm. The number of time steps per trial is a measure of how well the task has been learnt. Nevertheless, this “escape latency” does not only depend on the quality of the navigation map. It heavily depends on the starting point, especially during early learning. As a second factor, the ϵ -greedy navigation strategy adds noise, especially because the choice of exploration or exploitation is only made once every four time steps. In order to get a good estimation of the navigation map’s quality, the escape latency is therefore measured several times before each trial. These test-trials use the same ϵ -greedy navigation strategy, but learning is disabled.

Figure 8.5 (a) presents the average number of time steps to find the goal versus the number of learning trials for the “Buildings” environment. After about 20 trials, the task is learnt. In order to assess the model for navigation independently of the model of spatial representation, the same experiment is simulated, but instead of the learnt place code, a set of perfectly Gaussian place cells serves as state space. The escape latency for this perfect place code is shown in figure 8.5 (b). It is averaged over 100 instead of only 10 test trials. As expected, the mean escape latency and its variance decrease with the number of trials. The variance, however, starts with a very large value. The reason is that our ϵ -greedy policy does not perform well when there is no information about the goal yet. For instance, the agent could visit some places several times even if they are apparently not rewarding locations. Real animals show

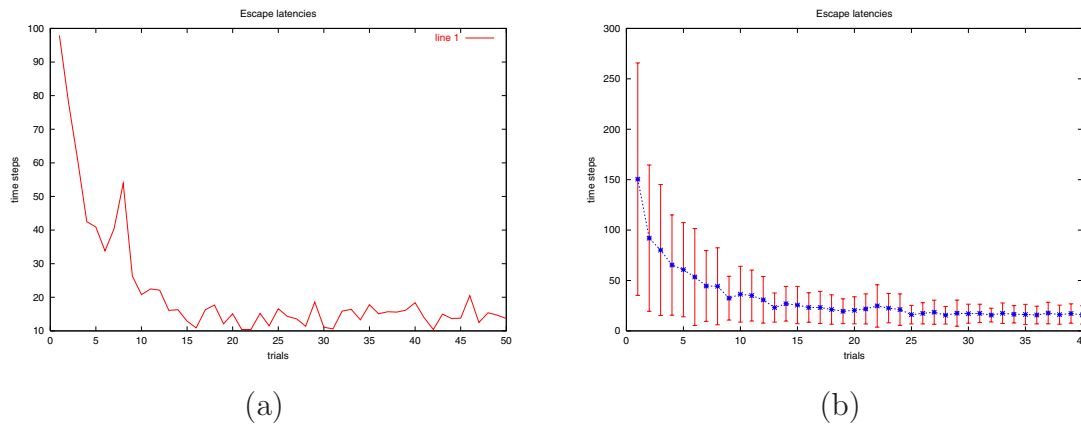


Figure 8.5: Escape latency (in time steps) versus number of trials. (a) For the “Buildings” environment, the average of 10 test trials is plotted. (b) The mean escape latency over 100 test trials in an artificial simulated environment with perfectly Gaussian ($\sigma = 6\text{cm}$) tuned place cells as spatial representation.

much more sophisticated behaviours and therefore don’t show such a high variability.

The learning algorithm improves the estimations of action values. The input potential h_i to action cell i represents the Q-value for action i . Because the state and action spaces are continuous, the values can be estimated for any location and heading in the environment. In order to visualise the action values, the agent is placed at regularly spaced sampling points (figure 8.6). At each point, a line is drawn in the direction of each action cell’s preferred heading. Red lines represent negative values, blue stands for positive values. The line length is proportional to the magnitude of the Q-value. Before learning (on the left), the values are random due to the random initialisation of the place cell to action cell connection strengths. After 10 learning trials (on the right), blue lines near the goal indicate that the goal location is attributed a high value.

An example navigation map before learning is shown in figure 8.7 (a). At each sample location, a line points in the direction of the population vector Φ_h^{ac} which represents the action which is considered optimal. The green area represents the goal location.

The navigation maps in figure 8.8 are taken for each of the four test environments (see section 5.1) after 20 learning trials. The task has been learnt in all environments. In the region below the goal, the navigation map does not point towards the goal in most cases. One reason is that the starting points of all trials have been set at a minimum distance of 20cm. It means that the agent never starts beneath the goal. In general, the probability that the agent is located at place \mathbf{p} at any time

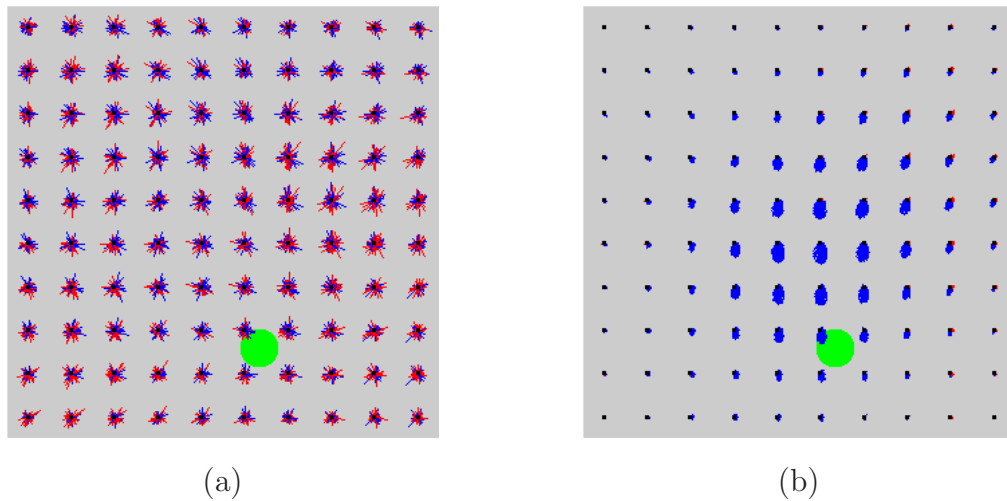


Figure 8.6: The Q-values as estimated by the input potential of action cells. At each sample point of the environment (Black points), the Q-values are plotted for all actions (directions). Blue means positive, red negative. The green area indicates the goal location. (a) Before learning. The values are random. (b) After 10 trials, regions near the goal get much higher values than location further away.

step t is higher for central positions than for locations near the walls. Therefore, the agent approaches the target more often from the north than from the south. As a consequence, the estimates of Q-values are better for central locations because they have been adapted more often. This is clearly visible also in figure 8.6 (b), where the action values below the goal are underestimated for northward movements. Secondly, the generalisation in state space, i.e. the large tuning curves of place cells, reinforce the “go-south” actions in places to the south of the goal even when reaching the goal from the north. This is also visible in figure 8.6 (b), where the area south of the goal falsely shows high positive Q-values for the south-direction, even though going south would lead the agent near the “dangerous” wall.

In all four test environments, the escape latencies after 20 learning trials is around 15 time steps, as shown in figure 8.7 (b). This corresponds to around four seconds if one time step is assumed to correspond to a theta cycle. The fluctuations are due to the randomness of the starting positions and the ϵ -greedy policy.

8.3.4 Convergence of the algorithm

In all our experiments, we observe that the navigation map and especially escape latencies are stable long before the Q-values have converged. Navigation maps after

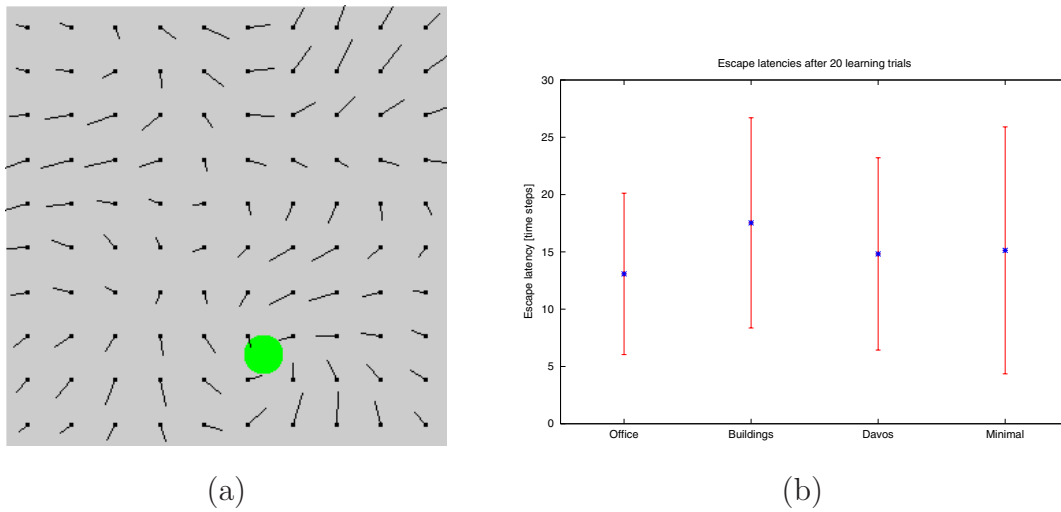


Figure 8.7: (a) Navigation map before learning. At each sample location marked with a black dot, a line points in the direction of the optimal action. The green area is the goal-location. Before learning, the map is random. (b) Average escape latencies after 20 learning trials for all environments. From left to right: “Office”, “Buildings”, “Davos” and “Minimal”.

10 and 100 trials look rather similar, but the changes in Q-values are non negligible. This is illustrated in figure 8.9.

The convergence of $Q(\lambda)$, i.e. Q-learning with eligibility traces, is only proved to converge if ETs are reset at each exploratory action [311,312]. As it is questionable if the brain implements such resets, we prefer not to include them in our model. Other online updates with eligibility traces for off-policy learning have been proposed [215], but no proof of convergence exists. Recently, convergence of new algorithms have been proved [219]. These are however only valid for offline updating, which means that all estimates are updated only at the end of each trial. Such an offline update rule does not seem to be biologically plausible.

Temporal difference learning in continuous spaces using function approximation is in general not proved to converge [105]. For Q-learning, explicit divergence has even been demonstrated [20]. Recently, a new off-policy algorithm with function approximation and eligibility trace has been proved to converge, but only with offline updates [218].

Nevertheless, our model seems to work with both eligibility trace and function approximation, i.e. generalisation in time and space. No proof of convergence, however, can be presented.

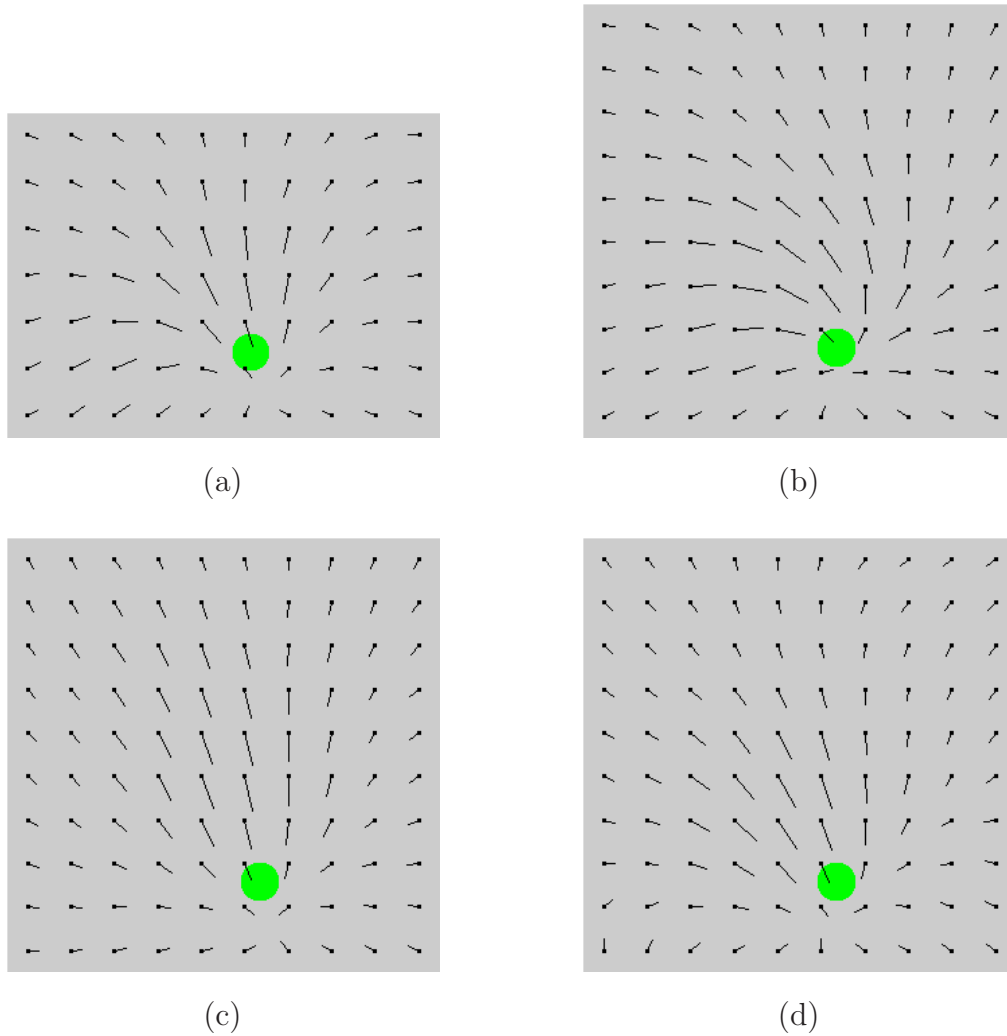


Figure 8.8: Navigation maps acquired after 20 trials for all test environments (see figure 8.7 on how to interpret navigation maps). In all environments, the task has been solved. The map has errors in the south of the goal. These are due to the large receptive fields of the place cells and the inhomogeneous sampling (the agent spends more time to the north than to the south of the goal). (a) “Office”. (b) “Buildings”. (c) “Davos”. (d) “Minimal”.

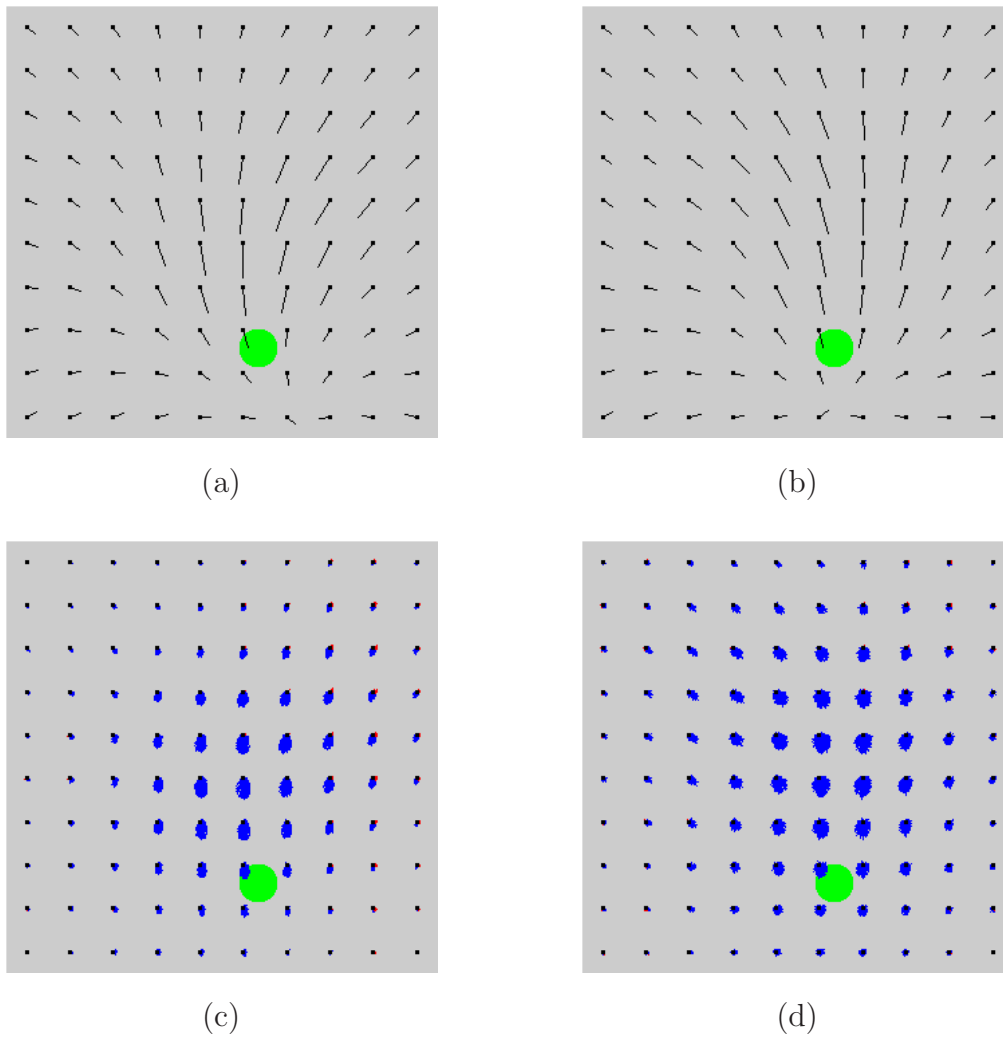


Figure 8.9: The navigation map converges well before the Q-values (see figures 8.7 and 8.6 on how to interpret the figures). (a,b) Navigation maps after 10 and 100 trials. The map does not change much. (c,d) Q-values after 10 and 100 trials. The estimates for action values continues to be improved, but it does not influence the navigation map.

Chapter 9

Conclusions

Now, at the end of this thesis, we would like to take the time to look back and summarise what has been presented and put it into a larger context. First, we outline the contributions of this work to the field of spatial learning. As any model of animal learning, our proposal is based on certain assumptions. These, together with our results, allow us to make some experimental predictions. Other models of spatial learning in the rat have been proposed in the literature. We discuss the differences and similarities to those models which are most closely related to our approach. Finally, we turn to the limitations of our system and propose future directions to improve our model. We also formulate some essential general questions in the area of spatial learning that are still open.

9.1 Contributions

In this thesis, we develop a connectionist model of the interactions between several brain areas involved in spatial learning. Each area hosts a population of neurons which serves a particular purpose. Figure 9.1 illustrates the architecture of the full system. To the left of each population, its functional name is indicated. To the right, we indicate the proposed neural substrate. We will briefly discuss how these populations solve their tasks. First, the construction of cognitive maps presented in chapter 6 is summarised. Then, we turn to the simple mechanism for multimodal integration of chapter 7 and finally to the locale navigation model detailed in chapter 8.

In order to evaluate our model, we test it on a real mobile robot platform in an office environment. The cognitive map formation and locale navigation model are also

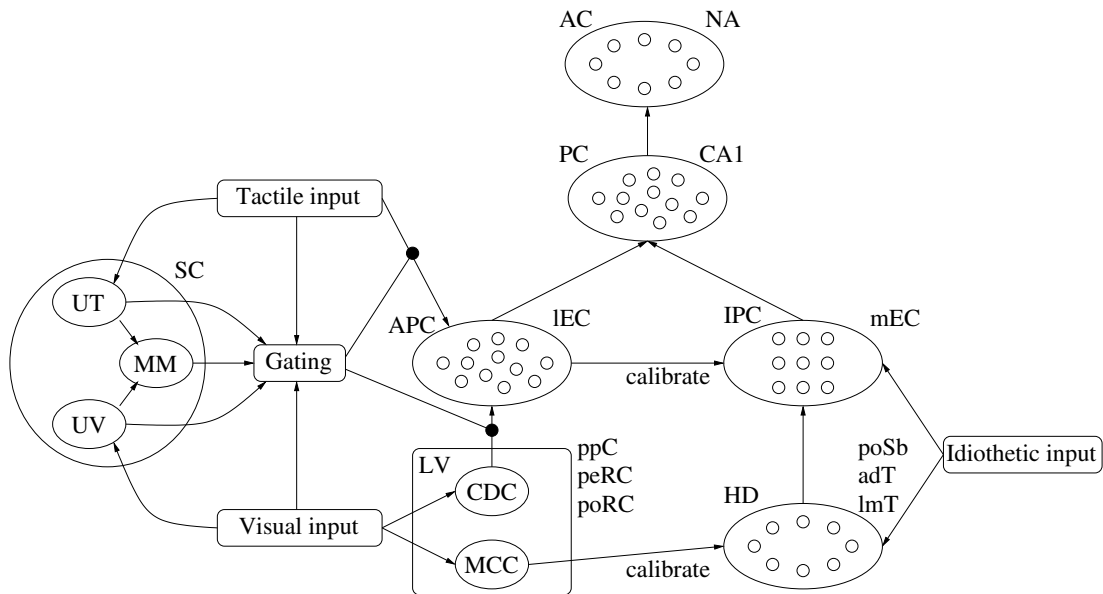


Figure 9.1: The architecture of the full model. The labels to the left of each population indicates its function. To the right, the proposed neural substrate is indicated. See text for details.

validated on a simulated agent in three different virtual environments. Two of them provide rich realistic visual input whereas the last is an impoverished environment similar to what is often used in animal experiments. In all environments, the agent successfully establishes a spatial representation. It also quickly learns to directly navigate to a hidden goal location in all cases.

Cognitive map

In chapter 6, we propose that a cognitive map is comprised of five interacting representations, each coding for a different aspect of the agent-environment relation. These aspects are: (i) the local view (LV), (ii) head direction (HD), (iii) allothetic place code (APC), (iv) idiothetic place code (IPC) and (v) a combined place code (PC).

LOCAL VIEW (LV): Two populations of neurons process the input of our artificial retina (section 5.2.2). *Multicolumn cells (MCCs)* combine information from several retina points to achieve a graded response to rotations of the agent. These cells are not very sensitive to place. *Column difference cells (CDCs)* respond to the difference in activity of two retinal columns. This difference depends on the agent’s position, but it is rotation insensitive within some limits. We propose that posterior parietal cortex (ppC), as well as perirhinal (peRC) and postrhinal (poRC) cortices code for the local view.

HEAD DIRECTION (HD): A population of *head direction cells* continuously integrates idiothetic information in order to track the agent's heading. During exploration of an environment, multicolumn cells associate the current local view with the current heading using an unsupervised Hebbian type learning rule. Due to these associations, idiothetic drift can be eliminated. Postsubiculum (poSb), anterodorsal (adT) and lateral mammillary (lmT) nuclei of thalamus are the proposed neural substrate for the head direction system.

ALLOTHETIC PLACE CODE (APC): At each unknown location, the system recruits an *allothetic place cell*. This cell is then contacted by active column difference cells using unsupervised learning. Thus, the current local view is associated to the current place. This results in broadly tuned directional place cells. We postulate that the lateral entorhinal cortex (IEC) contains such an allothetic spatial representation.

IDIOTHETIC PLACE CODE (IPC): A population of *idiothetic place cells* continuously integrates positional increments using internal cues. The head direction system provides the direction of the increment. However, the representation continuously incorporates information from the allothetic place code. This results in a stable nondirectional representation. Idiothetic drifts are effectively removed. We suggest the medial entorhinal cortex (mEC) as the biological locus of this representation.

COMBINED PLACE CODE (PC): The allothetic and idiothetic place representations converge onto a layer of combined place cells. Hebbian type unsupervised learning is used to update synaptic connections. The resulting place fields are mildly direction sensitive and more compact than in the purely idiothetic or allothetic cells. This layer is suggested to represent place cells in the CA1 region of the hippocampus proper.

Our model implements a spatial representation based on realistic visual input. Indeed, the results for the real and simulated environments are comparable. The relevant visual cues are associated to the correct positions and headings by an unsupervised learning rule. We are not aware of any other neural model which can construct a cognitive map and localise itself in a real environment without a compass or a strong polarising visual cue.

Multimodal integration

In chapter 7, we present a simple model for multisensory integration. It is inspired by neurophysiological properties of the deep layers of superior colliculus (SC). This brain area is involved in the alignment of visual, somatosensory and auditory topological maps. We propose a gating mechanism which learns to weigh visual and tactile sensory input for the hippocampal place code.

UNIMODAL AND MULTIMODAL CELLS (UV, UT, MM): An abstract model of neurons in the deep layers of superior colliculus is proposed. During an exploration stage, unimodal input cells store the visual and tactile stimuli separately. Unimodal

cells contact a newly recruited multimodal cell according to an unsupervised Hebbian-type learning rule.

GATING NETWORK: After the locations of multimodal coactivation has been acquired, a gating network learns to weigh visual and tactile input according to environmental conditions.

While sensor-fusion mechanisms in classical robotics systems are quite common, few biologically inspired solutions have been proposed. Most robotics applications employ probabilistic methods in the context of “occupancy grids”, which seem to be difficult to relate to brain functions. We do not know of any other biologically plausible model of spatial representations which learns to weigh sensory modalities according to the environmental conditions.

Locale navigation

In chapter 8, a model of locale navigation is developed based on hippocampal place cells and action cells in nucleus accumbens. Reinforcement learning in continuous state and action spaces is employed.

ACTION CELLS (AC): The place cells of our cognitive map project to a layer of directional action cells thought to be located in the nucleus accumbens (NA). These action cells estimate the total future reward for each direction of movement. The large activity profile enforced around the selected action results in a quick generalisation in action space. A population vector decoding scheme enables the interpretation of continuous actions.

REINFORCEMENT LEARNING: In order to learn the navigation map to a hidden target, a reinforcement learning method is used for adapting the synaptic connections from place cells to action cells. The system learns to navigate to a rewarding location from anywhere in the environment in around 20 trials.

ELIGIBILITY TRACE: Dopamine has been shown to enhance the long term potentiation of synapses. It has also been related to the reward prediction error used in reinforcement learning. It remains unclear, however, how dopamine can modulate the potentiation of synapses which were active *before* the animal finds a reward. A memory trace is required. We discuss several biologically plausible mechanisms for such an eligibility trace.

Our navigation system neurally implements a reinforcement learning mechanism in continuous state and action spaces. We are not aware of any other neural model of locale navigation which is continuous in state and action spaces.

9.2 Experimental predictions

In this section, we discuss some predictions that this work allows us to make. The predictions are based on the assumptions and results of each component of our system.

Cognitive map

Local view: we suggest that the posterior parietal cortex (ppC) as well as perirhinal (peRC) and postrhinal (poRC) cortices process the local view. These neurons should therefore show firing properties selective to the current allothetic sensory input in general, and visual input in particular. In conflict situations between idiothetic and allothetic cues, ie. when the rat moves but the external input remains stable or vice versa, these areas should follow the allothetic stimulus.

Head direction system: The proposed neural substrates of our head direction (HD) system include postsubiculum (poSb), anterodorsal (adT) and lateral mammillary (lmT) nuclei of thalamus. Our results suggest that lesions in ppC, peRC and poRC should produce drifting idiothetic head direction cells because recalibration using visual input would be disabled. Lesions in the idiothetic pathway should cause broadly tuned HD cells in known environments and impair acquisition of HD tuning in new environments.

Allothetic place code: Our model proposes the lateral entorhinal cortex (IEC) as neural substrate for an allothetic place code. It receives sensory input from the local view module and extracts information relevant to encode the rat's *position*. Therefore, lesions in ppC, peRC and poRC, which convey the local view in our model, should abolish spatial firing in lateral entorhinal cortex. In conflict situations between idiothetic and allothetic cues, IEC cells should follow the allothetic cue. Lesions in the idiothetic pathway should not affect this area. Finally, we predict that place cells in IEC are directional.

Idiothetic place code: the medial entorhinal cortex (mEC) and the subiculum (Sb) are the suggested neural substrates of our idiothetic spatial representation. Place cells in mEC should be mostly non-directional. Disrupting the allothetic pathway and in particular lesioning IEC→mEC connections should produce a drifting idiothetic place code. Lesions in the idiothetic pathway should abolish firing in mEC. Lesions in poSb, adT and ldT, which provide mEC with the animal's heading, should produce severe inconsistencies in the idiothetic place code.

Combined place code: The idiothetic and allothetic representations combine into a spatial map the locus of which is predicted to be the hippocampal region CA1. Lesions in mEC should leave CA1 with more broadly tuned, purely allothetic, directional place cells. Lesions of the head direction system should severely disrupt spatial firing, except if mEC is also lesioned, which should produce similar results than disrupting mEC alone.

Multimodal integration

In the dark, it should be possible to establish spatial representations in IEC using non-visual allothetic cues (olfactive, somatosensory etc.). If, however, the same environment is illuminated but contains no visual cues, it should not be possible to tune place fields in IEC because vision would be given a high weight and suppress the other modes, even if the visual input contains no information.

Locale navigation

Nucleus accumbens: Our model postulates the existence of a population of allocentric directional action cells in nucleus accumbens (NA). We predict that lesions in CA1 disrupts directional firing in NA and impairs locale navigation. Damaging NA itself should also impair locale navigation while leaving the hippocampal place code intact. We propose that the firing properties of cells in nucleus accumbens predict the future direction of movement in an allocentric coordinate frame. We also suggest that lateral interactions exist within NA. Finally, reward-driven learning should not be impaired when blocking plasticity in the hippocampus after the spatial representation has been learnt.

Dopamine: Learning a navigation map is suggested to be driven by a dopaminergic error signal from the ventral tegmental area (VTA). Lesions in VTA should therefore impair acquisition of a new task but not affect the navigation to an already learnt goal location. Stimulating VTA whenever the rat is at a certain location should resemble the case where a reward is given at this location.

Eligibility trace: If eligibility traces are implemented by dopamine-dependent inhibition of depotentiation due to low-frequency stimulation, then the blocking of presynaptic firing after reward delivery should disrupt learning, because with a lack of low-frequency stimulation, none of the synapses get depotentiated.

9.3 Relation to other models

Here we discuss the similarities and differences to existing models of spatial learning. In the case of our simple superior colliculus (SC) model for multimodal integration, no comparable models exist. Most experimental work on SC is done in the superficial layers in monkeys or cats. Existing models of deep SC are on a neuronal level with the aim of finding the mechanisms of multimodal enhancement. We therefore focus our comparison on the spatial representation and locale navigation parts of this thesis.

Brown and Sharp (1991, 1995)

Brown and Sharp's model [254, 39] relies on an artificial local view system where the distance to all uniquely identifiable landmarks is available. In contrast, our system uses realistic camera images without a high-level object recognition mechanism. Whereas our model results in a highly redundant and distributed place code consistent with experimental data, their winner-take-all mechanism prevents all but one place cell from firing. Their representation, unlike ours, has no path integration component and therefore doesn't support location-specific firing in the dark.

Our navigation system is similar to theirs. We also rely on reward-driven learning in nucleus accumbens. Their output neurons, however, code for egocentric directions whereas our action cells code for an allocentric direction. While their model needs a long temporal trace in order to overcome a distal reward problem, our approach does not suffer such a limitation.

Burgess *et al.* (1994)

Burgess *et al.* [41] use a mathematical similarity measure, namely the euclidian distance to the arena walls, as allothetic stimulus. Our system relies on real camera images. Their proposal includes competition in the hippocampus and subiculum. Our place cell layer does not include lateral connections or any other competition mechanism. In the absence of allothetic input, their model can not sustain place cell firing, whereas our path integrator can drive place cells in the dark. In their model, the effect of phase precession is produced in entorhinal cortex by an abstract geometrical mechanism. Our model does not address this issue.

The goal-learning mechanism is also reward-driven in their proposal. However, their approach suffers a "distal reward" problem because only those place cells whose fields contain the goal may learn place to action associations. Our approach is based on temporal difference learning which does not suffer this limitation. Locale navigation is impaired in our model if the connection between hippocampus and nucleus accumbens is lesioned [76, 279, 209]. Their system is not consistent with these experimental findings.

Wan, Redish and Touretzky (1994, 1996, 1997)

The model proposed by Wan, Redish and Touretzky [310, 296, 228, 229] encodes allothetic signals mathematically by type, distance and bearing to a set of landmarks and does not work in realistic environments. Another difference is that there is no purely allothetic place code, like our IEC population. In general, their neuronal model is more abstract.

Gaussier and colleagues (1998, 2000, 2002)

The system proposed by Gaussier and colleagues [96,95,97] like our proposal extracts information from real panoramic camera images. However, their model relies on a compass for the construction of a spatial representation whereas ours does not suffer this limitation. Their hippocampus model is more detailed than ours in that it distinguishes between a place recognition layer in entorhinal cortex and dentate gyrus, a transition prediction module in CA3 and a transition recognition population in CA1. Nevertheless, their representation does not include a path integration component and can, unlike our system and the rat, only operate when allothetic input is available. A winner-take-all mechanism is applied in several of their populations whereas our approach supports a distributed place code consistent with experimental data.

According to their model, locale navigation is implemented in the hippocampus, which is in contradiction with experimental data [76,279,209]. Our model, on the other hand, postulates the implication of nucleus accumbens. The symbolic nature of place transitions might reduce the robustness and the flexibility of their system. Our distributed approach is at the same time robust and flexible. Finally, the output of their navigation system is an egocentric direction, whereas we use an allocentric representation of actions.

Arleo *et al.* (2000, 2001)

The model by Arleo *et al.* [14,16,15] also uses real camera images in order to establish a place code. At each visited location, they take four snapshots in the four cardinal directions, whereas our system handles images taken in any direction. Their head direction system is more detailed in that it models the interaction between different populations involved in the processing of head direction. Recalibration using allothetic cues is also performed differently. Whereas we calibrate the HD system using unsupervised learning between the local view representation and HD cells, Arleo *et al.* use a stable light source at the border of the arena as a reference. As this directional cue is not location independent, the association between the allothetic place code in entorhinal cortex and the egocentric bearing of the light source is learnt in an offline supervised learning stage. Lesions in entorhinal cortex thus disrupts their heading calibration. This is in contrast to our system where the calibration depends on the local view only. In their work, the place cells in lateral entorhinal cortex are predicted to be non-directional whereas we predict directional firing.

Finally, our action learning mechanism differs from theirs in that it supports continuous action selection and generalisation in action space. Their system uses four discrete actions, which means that during learning, the agent can only move in four different directions in their proposal.

Abbot and colleagues (1996, 1997)

Hippocampal place cells in Abbott and colleagues' work [34, 101] have perfectly Gaussian tuned receptive fields prior to action learning. They do not depend on allothetic cues whereas we learn the place fields from real images. After navigation learning, place cell firing is sensitive to the goal location in their model whereas our place cells are task independent.

They propose a goal-oriented navigation mechanism which only depends on the hippocampus. A navigation map is stored in the CA3 collaterals. This conflicts with experimental data showing that lesioning the fornix or nucleus accumbens impairs locale navigation [76, 279, 209]. Our model is consistent with these findings because a navigation map is learnt in nucleus accumbens, which receives spatial information via the fornix. In order to generate movement commands, the original place field centres are accessed explicitly, which is a biologically implausible operation. In our model, positions and actions are implicitly coded.

Foster, Morris and Dayan (2000)

The model proposed by Foster, Morris and Dayan [88] relies a population of perfectly Gaussian tuned place cells with no allothetic component. Our place fields are learnt by experience.

Their navigation model postulates a set of eight action neurons which code, like our action cells, for the allothetic direction of movement. Unlike our approach, learning does not generalise to similar actions and does not allow for continuous directions of movement. Their navigation learning rule is based on an actor-critic temporal difference learning algorithm whereas we use Q-learning. In order to overcome interference with previously learnt goal locations, their model features a coordinate system module which learns to transform place code activity into Cartesian coordinates. Once learnt, algorithmic vector subtraction replaces the action selection based on the local view. Local obstacle avoidance is no longer possible. Our model does not address the quick relearning of a goal location.

9.4 Limitations and perspectives

Here we discuss some limitations of our approach and propose future directions which could improve the performance or biological plausibility of our model. We also address some essential general questions in the area of spatial learning that are still unresolved.

Path integration: Our path integration system (both IPC and HD system) consists of neurons with artificially tuned receptive fields. Furthermore, the idiothetic update and the new firing rates are algorithmically calculated. Instead, lateral interconnections and a plausible update mechanism should be implemented in order to

improve the biological plausibility of our model.

Calibration of IPC: In order to calibrate the idiothetic place code, we rely on the APC population vector. This may not be biologically plausible. Direct associations between local view or APC cells and IPC neurons should be established during exploration. These connections, however, would then need to be weighted appropriately in order to efficiently calibrate the idiothetic representation.

Multiple environments: A hippocampal place cell which is active in a particular environment is not necessarily active in second environment. If it is, however, the cell does not generally code for a similar location within the second environment [142,291]. The change in the firing properties occurs suddenly and is termed a “remapping” [159]. Our model does not have the capability to perform a “remapping”. Indeed, the underlying principle of this effect is still an open question.

Sequence learning: Sequence learning is the ability to store and retrieve sequential activation of place cells. We suggest that this is implemented in the CA3 region of the hippocampus. The association between places might be stored in the lateral interconnections present in CA3. Our model, however, doesn’t feature sequence learning.

Directionality of place fields: Place cells in our combined place code layer are mainly non-directional. However, hippocampal place cells can have directional fields when the rat’s movement is restricted, e.g. in a linear track. We suggest that the hippocampal place fields are initially driven mainly by path integration, and are therefore non-directional at an early stage of exploration. If, however, a position is always paired with a particular view, which is the case for restricted movement but not in free exploration, allothetic information might become more important than path integration. As our allothetic place code in IEC is directional, we propose that the IEC→CA1 projections become stronger than mEC→CA1 connections.

Consolidation of cognitive maps: Locale navigation tasks depend on the hippocampus shortly after learning. After a few weeks, however, rats can perform the task even if their hippocampus is lesioned. It is suggested that the cognitive map is *consolidated* and transferred to a different brain region [193]. The mechanism of consolidation is, however, an open question and our model does not propose a solution.

Working memory: Our model can not solve working memory tasks. We find that path integration (PI) is extremely useful in order to “smooth” the cognitive map in time and reproduce the clean place fields of the hippocampus. PI is a form of working memory in the sense we interpret it in this thesis (section 3.2). However, it is not sufficient to solve the win-shift working memory task in the radial arm maze (section 3.4). The task we addressed in this thesis, namely finding a hidden goal location, can also be solved by a memoryless stimulus-response mechanism. In our model, this would correspond to disabling the PI components.

Relevance of a sensory mode: Our simple model of multimodal integration contains a gating network which learns to adapt the weight of each sensory mode according to environmental conditions. Learning is driven by a reward function which

assesses the quality of the place code. An approach based on the information content of a sensory signal could provide a better estimate of the relevance of a sensory signal.

Multiple goal locations: Our model can only store one navigation map. A natural extension would provide each target with its own population of action cells. A form of working memory could then specify which target is currently the most interesting, based on motivation for instance. This extension could solve the win-shift working memory task in the radial arm maze mentioned above. However, the working memory would be modelled outside of the hippocampus. A different mechanism could include working memory in the form of goal-dependent cells in the hippocampal model and use only one action cell population. Such an architecture could also solve the win-shift task.

Different navigation strategies: We have modelled a locale navigation mechanism. An interesting extension would be to incorporate other modes of navigation such as homing by dead reckoning or taxon navigation. A model of the basal ganglia could then select an appropriate strategy according to motivation, sensory input and complexity of the task.

Action selection: We employ an ϵ -greedy policy to select the next action in our navigation model. A sophisticated method which adapts the action selection to the current conditions would be an important improvement. For instance, the selection could depend on a quality estimation of the navigation map, or it could avoid already visited places etc.

Finding a sound navigation learning algorithm: We use Q-learning with eligibility traces and function approximation in order to learn a navigation map. The soundness of such methods is no longer an open question: It has been shown that this approach may lead to divergence [20]. Nevertheless, our model and many others seem to work in practice. Obviously, it would be better to find sound solutions. So far, we are not aware of any proposal of a sound online temporal difference learning algorithm which uses function approximation.

Eligibility trace: We discuss some potential neuronal mechanism which could implement an eligibility trace. They are based on neuronal properties which can not be captured by our rate-coded neurons. A possible extension of our work is to transpose the entire model to the case of spiking neurons. The mechanism for eligibility traces proposed in section 8.3 could then be validated.

Other mazes: Finally, we only test our model in small open-field environments. It would be interesting to run experiments in all environments of figure 3.3 (page 37) and compare the results to animal experiments.

Bibliography

- [1] Adams A.D. and Forrester J.M.: “The projection of the rat’s visual field on the cerebral cortex.” *Q J Exp Physiol Cogn Med Sci*, **53**(3): pp. 327–36, **1968**. 56
- [2] Alberini C.M., Ghirardi M., Metz R. and Kandel E.R.: “C/EBP is an immediate-early gene required for the consolidation of long-term facilitation in Aplysia.” *Cell*, **76**(6): pp. 1099–114, **1994**. 105
- [3] Albin R.L., Young A.B. and Penney J.B.: “The functional anatomy of basal ganglia disorders.” *Trends Neurosci*, **12**(10): pp. 366–75, **1989**. 30, 110
- [4] Alyan S.H., Paul B.M., Ellesworth E., White R.D. and McNaughton B.L.: “Is the hippocampus required for path integration?” *Society for Neuroscience Abstracts*, **23**: p. 504, **1997**. 16, 17
- [5] Amaral D.G.: “Emerging principles of intrinsic hippocampal organization.” *Current Opinion in Neurobiology*, **3**: pp. 225–229, **1993**. 11
- [6] Amaral D.G., Dolorfo C. and Alvarez-Royo P.: “Organization of Ca1 projections to the subiculum: a PHA-L analysis in the rat.” *Hippocampus*, **1**: pp. 415–435, **1991**. 11
- [7] Amaral D.G. and Witter M.P.: “The three-dimensional organization of the hippocampal formation: A review of anatomical data.” *Neuroscience*, **31**(3): pp. 571–591, **1989**. 8, 10, 11
- [8] Amaral D.G. and Witter M.P.: “Hippocampal Formation.” In Paxinos G. (editor) “The Rat Nervous System,” chapter 21, pp. 443–493. Academic Press, second edition, **1995**. 8, 9, 11, 30
- [9] Amari S.I.: “Dynamics of pattern formation in lateral-inhibition type neural fields.” *Biological Cybernetics*, **27**: pp. 77–87, **1977**. 33, 77
- [10] Amit D.J.: *Modeling Brain Function. The world of attractor neural networks*. Cambridge University Press, Cambridge, **1989**. 33, 77

- [11] Andersen R.A., Essick G.K. and Siegel R.M.: “Encoding of spatial location by posterior parietal neurons.” *Science*, **230**(4724): pp. 456–8, **1985**. 35
- [12] Apicella P., Scarnati E., Ljungberg T. and Schultz W.: “Neuronal activity in monkey striatum related to the expectation of predictable environmental events.” *J Neurophysiol*, **68**(3): pp. 945–60, **1992**. 53, 99
- [13] Arleo A. and Gerstner W.: “Modeling rodent head-direction cells and place cells for spatial learning in bio-mimetic robotics.” In Meyer J.A., Berthoz A., Floreano D., Roitblat H.L. and Wilson S.W. (editors) “From Animals to Animats VI,” pp. 236–245. MIT Press, Cambridge MA, **2000**. 25, 27, 55
- [14] Arleo A. and Gerstner W.: “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**: pp. 287–299, **2000**. 25, 27, 45, 55, 100, 101, 124
- [15] Arleo A. and Gerstner W.: “Spatial Orientation in Navigating Agents: Modeling Head-direction Cells.” *Neurocomputing*, **38–40**: pp. 1059–1065, **2001**. 25, 124
- [16] Arleo A., Smeraldi F., Hug S. and Gerstner W.: “Place Cells and Spatial Navigation Based on 2D Visual Feature Extraction, Path Integration, and Reinforcement Learning.” In Leen T.K., Dietterich T.G. and Tresp V. (editors) “Advances in Neural Information Processing Systems 13,” pp. 89–95. MIT Press, **2001**. 25, 27, 45, 55, 93, 100, 124
- [17] Artola A. and Singer W.: “Long-term potentiation and NMDA receptors in rat visual cortex.” *Nature*, **330**(6149): pp. 649–52, **1987**. 31, 104
- [18] Artola A. and Singer W.: “Long-term depression of excitatory synaptic transmission and its relationship to long-term potentiation.” *Trends Neurosci*, **16**(11): pp. 480–7, **1993**. 31, 104
- [19] Aston-Jones G., Rajkowski J., Kubiak P. and Alexinsky T.: “Locus coeruleus neurons in monkey are selectively activated by attended cues in a vigilance task.” *J Neurosci*, **14**(7): pp. 4467–80, **1994**. 54
- [20] Baird III L.C.: “Residual Algorithms: Reinforcement Learning with Function Approximation.” In “International Conference on Machine Learning,” pp. 30–37. **1995**. 53, 113, 127
- [21] Bao S., Chan V.T. and Merzenich M.M.: “Cortical remodelling induced by activity of ventral tegmental dopamine neurons.” *Nature*, **412**(6842): pp. 79–83, **2001**. 54, 105

- [22] Barlow J.S.: “Inertial navigation as a basis for animal navigation.” *Journal of Theoretical Biology*, **6**: pp. 76–117, **1964**. 15, 33
- [23] Barnes C.A., McNaughton B.L., Mizumori S.J.Y., Leonard B.W. and Lin L.H.: “Comparison of spatial and temporal characteristics of neuronal activity in sequential stages of hippocampal processing.” In Storm-Mathisen J., Zimmer J. and Ottersen O.P. (editors) “Understanding the Brain through the Hippocampus: The Hippocampal Region as a Model for Studying Brain Structure and Function,” volume 83 of *Progress in Brain Research*, pp. 287–300. Elsevier Science Publishers B. V., New York, **1990**. 15, 36, 84
- [24] Bayer S.A.: “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): pp. 315–23, **1982**. 11
- [25] Benhamou S.: “Path integration by swimming rats.” *Anim Behav*, **54**(2): pp. 321–7, **1997**. 34
- [26] Berkes P. and Wiskott L.: “Slow feature analysis yields a rich repertoire of complex-cell properties.” *Cognitive Sciences EPrint Archive (CogPrints)* 2804, <http://cogprints.ecs.soton.ac.uk/archive/00002804/>, **2003**. 31
- [27] Bi G.Q. and Poo M.M.: “Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type.” *J Neurosci*, **18**(24): pp. 10464–72, **1998**. 32, 104
- [28] Biegler R. and Morris R.G.M.: “Landmark stability is a prerequisite for spatial but not discrimination learning.” *Nature*, **361**: pp. 631–633, **1993**. 12, 67
- [29] Biegler R. and Morris R.G.M.: “Landmark stability: Studies exploring whether the perceived stability of the environment influences spatial representation.” *Journal of Experimental Biology*, **199**(1): pp. 187–193, **1996**. 12, 40
- [30] Biella G. and De Curtis M.: “Olfactory inputs activate the medial entorhinal cortex via the hippocampus.” *Journal of Neurophysiology*, **83**(4): pp. 1924–1931, **2000**. 10
- [31] Blair H.T. and Sharp P.E.: “Anticipatory head direction signals in anterior thalamus: Evidence for a thalamocortical circuit that integrates angular head motion to compute head direction.” *Journal of Neuroscience*, **15**(9): pp. 6260–6270, **1995**. 17, 67, 79
- [32] Bliss T.V. and Collingridge G.L.: “A synaptic model of memory: long-term potentiation in the hippocampus.” *Nature*, **361**(6407): pp. 31–9, **1993**. 31, 104

- [33] Blond O., Crepel F. and Otani S.: “Long-term potentiation in rat prefrontal slices facilitated by phased application of dopamine.” *Eur J Pharmacol*, **438**(1-2): pp. 115–6, **2002**. 54, 105
- [34] Blum K.I. and Abbott L.F.: “A model of spatial map formation in the hippocampus of the rat.” *Neural Computation*, **8**: pp. 85–93, **1996**. 42, 80, 125
- [35] Bose A., Booth V. and Recce M.: “A temporal mechanism for generating the phase precession of hippocampal place cells.” *Journal of Computational Neuroscience*, **9**: pp. 5–30, **2000**. 14
- [36] Bostock E., Muller R.U. and Kubie J.L.: “Experience-dependent modifications of hippocampal place cell firing.” *Hippocampus*, **1**(2): pp. 193–206, **1991**. 12, 67
- [37] Brandner C. and Schenk F.: “Septal lesions impair the acquisition of a cued place navigation task: Attentional or memory deficit?” *Neurobiology of Learning and Memory*, **69**(2): pp. 106–125, **1998**. 12
- [38] Brecht M., Preilowski B. and Merzenich M.M.: “Functional architecture of the mystacial vibrissae.” *Behav Brain Res*, **84**(1-2): pp. 81–97, **1997**. 63
- [39] Brown M.A. and Sharp P.E.: “Simulation of spatial-learning in the Morris water maze by a neural network model of the hippocampal-formation and nucleus accumbens.” *Hippocampus*, **5**: pp. 171–188, **1995**. 36, 41, 100, 123
- [40] Burgess N., Jeffery K.J. and O’Keefe J.: “Integrating hippocampal and parietal functions: a spatial point of view.” In N. Burgess K.J.J. and O’Keefe J. (editors) “The Hippocampal and Parietal Foundations of Spatial Cognition,” chapter 1, pp. 3–29. Oxford University Press, **1999**. 11
- [41] Burgess N., Recce M. and O’Keefe J.: “A model of hippocampal function.” *Neural Networks*, **7**: pp. 1065–1081, **1994**. 12, 14, 21, 40, 55, 87, 101, 123
- [42] Burwell R.D. and Amaral D.G.: “Cortical afferents of the perirhinal, postrhinal, and entorhinal cortices of the rat.” *J Comp Neurol*, **398**(2): pp. 179–205, **1998**. 8
- [43] Butterworth G.: “The Origins of Auditory-Visual Perception and Visual Proprioception in Human Development.” In Pick H.L. Jr. and Walk R.D. (editors) “Intersensory Perception and Sensory Integration,” Perception and Perceptual Development, chapter 2, pp. 37–70. Plenum Press, New York, **1981**. 17

- [44] Buzsáki G.: “Feed-forward inhibition in the hippocampal formation.” *Progress in Neurobiology*, **22**: pp. 131–153, **1984**. 12
- [45] Buzsáki G.: “Two-stage model of memory trace formation: a role for ”noisy” brain states.” *Neuroscience*, **31**(3): pp. 551–70, **1989**. 106
- [46] Buzsáki G.: “Theta oscillations in the hippocampus.” *Neuron*, **33**(3): pp. 325–340, **2002**. 12
- [47] Carvell G.E. and Simons D.J.: “Biometric analyses of vibrissal tactile discrimination in the rat.” *J Neurosci*, **10**(8): pp. 2638–48, **1990**. 63, 64
- [48] Castellanos J.A., Neira J. and Tardós J.D.: “Multisensor Fusion for Simultaneous Localization and Map Building.” *IEEE Transactions on Robotics and Automation*, **17**(6): pp. 908–914, **2001**. 27
- [49] Chen L.L., Lin L., Barnes C.A. and McNaughton B.L.: “Head-direction cells in the rat posterior cortex. II. Contributions of visual and idiothetic information to the directional firing.” *Experimental Brain Research*, **101**: pp. 23–34, **1994**. 35
- [50] Chen L.L., Lin L., Green E.J., Barnes C.A. and McNaughton B.L.: “Head-direction cells in the rat posterior cortex. I. Anatomical distribution and behavioral modulation.” *Experimental Brain Research*, **101**: pp. 8–23, **1994**. 35
- [51] Cheng K.: “A purely geometric module in the rat’s spatial representation.” *Cognition*, **23**(2): pp. 149–78, **1986**. 39
- [52] Chrobak J.J. and Buzsáki G.: “High-Frequency Oscillations in the Output Networks of the Hippocampal-Entorhinal Axis of the freely moving Rat.” *Journal of Neuroscience*, **16**(9): pp. 3056–3066, **1996**. 12
- [53] Chrobak J.J., Lörincz A. and Buzsáki G.: “Physiological Patterns in the Hippocampo-Entorhinal Cortex System.” *Hippocampus*, **10**: pp. 457–465, **2000**. 12
- [54] Ciaroni S., Cuppini R., Cecchini T., Ferri P., Ambrogini P., Cuppini C. and Del Grande P.: “Neurogenesis in the adult rat dentate gyrus is enhanced by vitamin E deficiency.” *J Comp Neurol*, **411**(3): pp. 495–502, **1999**. 11
- [55] Claiborne B.J., Amaral D.G. and Cowan W.M.: “A light and electron microscopic analysis of the mossy fibers of the rat dentate gyrus.” *J Comp Neurol*, **246**(4): pp. 435–58, **1986**. 11

- [56] Collett T.S., Cartwright B.A. and Smith B.A.: “Landmark learning and visuospatial memories in gerbils.” *Journal of Comparative Physiology A*, **158**: pp. 835–851, **1986**. 7, 40
- [57] Cooper B.G., Miya D.Y. and Mizumori S.J.Y.: “Superior Colliculus and Active Navigation: Role of Visual and Non-Visual Cues in Controlling Cellular Representations of Space.” *Hippocampus*, **8**: pp. 340–372, **1998**. 19
- [58] Cressant A., Muller R.U. and Poucet B.: “Failure of centrally placed objects to control firing fields of hippocampal place cells.” *Journal of Neuroscience*, **17(7)**: pp. 2531–2542, **1997**. 12, 39, 57, 67
- [59] Csicsvari J., Jamieson B., Wise K. and Buzsáki G.: “Mechanisms of gamma oscillations in the hippocampus of the behaving rat.” *Neuron*, **37(2)**: pp. 311–322, **2003**. 12
- [60] da Silva F.H.L., Witter M.P., Boeijinga P.H. and Lohman A.H.M.: “Anatomical organization and physiology of the limbic cortex.” *Physiological Reviews*, **70**: pp. 453–511, **1990**. 10
- [61] Daugman J.G.: “Two-dimensional spectral analysis of cortical receptive field profiles.” *Vision Research*, **20**: pp. 847–856, **1980**. 60
- [62] Davis H.P. and Squire L.R.: “Protein synthesis and memory: a review.” *Psychol Bull*, **96(3)**: pp. 518–59, **1984**. 32, 105
- [63] Dayan P.: “The Convergence of TD(λ) for General λ .” *Machine Learning*, **8(3-4)**: pp. 341–362, **1992**. 51, 52
- [64] Dayan P. and Sejnowski T.J.: “TD(λ) Converges with Probability 1.” *Machine Learning*, **14(3)**: pp. 295–301, **1994**. 52
- [65] del R. Millán J., Posenato D. and Dedieu E.: “Continuous-Action Q-Learning.” *Machine Learning*, **49**: pp. 247–265, **2002**. 100
- [66] Devan B.D. and White N.M.: “Parallel information processing in the dorsal striatum: relation to hippocampal function.” *J Neurosci*, **19(7)**: pp. 2789–98, **1999**. 38, 99
- [67] Dittman J.S., Kreitzer A.C. and Regehr W.G.: “Interplay between facilitation, depression, and residual calcium at three presynaptic terminals.” *J Neurosci*, **20(4)**: pp. 1374–85, **2000**. 32, 105

- [68] Doya K.: “Temporal Difference Learning in Continuous Time and Space.” In Touretzky D.S., Mozer M.C. and Hasselmo M.E. (editors) “Advances in Neural Information Processing Systems,” volume 8, pp. 1073–1079. The MIT Press, **1996**. 48
- [69] Doya K.: “Reinforcement learning in continuous time and space.” *Neural Computation*, **12**: pp. 219–245, **2000**. 48, 54, 100
- [70] Doya K.: “Metalearning and neuromodulation.” *Neural Netw*, **15**(4-6): pp. 495–506, **2002**. 53, 54
- [71] Dräger U.C. and Hubel D.H.: “Responses to visual stimulation and relationship between visual, auditory, and somatosensory inputs in mouse superior colliculus.” *J Neurophysiol*, **38**(3): pp. 690–713, **1975**. 19
- [72] Dräger U.C. and Hubel D.H.: “Topography of visual and somatosensory projections to mouse superior colliculus.” *J Neurophysiol*, **39**(1): pp. 91–101, **1976**. 19
- [73] Dudai Y.: “Consolidation: fragility on the road to the engram.” *Neuron*, **17**(3): pp. 367–70, **1996**. 32, 105
- [74] Dudai Y.: “Molecular bases of long-term memories: a question of persistence.” *Curr Opin Neurobiol*, **12**(2): pp. 211–6, **2002**. 32, 105
- [75] Egorov A.V., Hamam B.N., Franssen E., Hasselmo M.E. and Alonso A.A.: “Graded persistent activity in entorhinal cortex neurons.” *Nature*, **420**(6912): pp. 173–8, **2002**. 106
- [76] Eichenbaum H., Stewart C. and Morris R.G.M.: “Hippocampal representation in place learning.” *Journal of Neuroscience*, **10**(11): pp. 3531–3542, **1990**. 36, 37, 41, 43, 99, 123, 124, 125
- [77] Elfes A.: “Using Occupancy Grids for Mobile Robot Perception and Navigation.” *IEEE Computer*, **22**(6): pp. 46–57, **1998**. 27
- [78] Etienne A.S.: “The control of short-distance homing in the golden hamster.” In Ellen P. and Thinus-Blanc C. (editors) “Cognitive Processes and Spatial Orientation in Animals and Man,” volume 1 (Experimental Animal Psychology and Ethology), pp. 233–251. Martinus Nijhoff Publishers, Boston, **1987**. 34
- [79] Etienne A.S., Boulens V., Maurer R., Rowe T. and Siegrist C.: “A brief view of known landmarks reorientates path integration in hamsters.” *Naturwissenschaften*, **87**(11): pp. 494–8, **2000**. 33, 34, 35

- [80] Etienne A.S., Maurer R., Berlie J., Reverdin B., Rowe T., Georgakopoulos J. and Seguinot V.: “Navigation through vector addition.” *Nature*, **396**(6707): pp. 161–4, **1998**. 33, 34, 57
- [81] Etienne A.S., Maurer R. and Seguinot V.: “Path integration In mammals and its interaction with visual landmarks.” *Journal of Experimental Biology*, **199**(1): pp. 201–209, **1996**. 34, 35
- [82] Fend M., Bovet S., Yokoi H. and Pfeifer R.: “An active artificial whisker array for texture discrimination.” In “IEEE/RSJ International Conference on Intelligent Robots and Systems,” **2003**. To appear. 64, 65
- [83] Fend M., Yokoi H. and Pfeifer R.: “Optimal Morphology of a Biologically-Inspired Whisker Array on an Obstacle-Avoiding Robot.” In Banzhaf W., Christaller T., Dittrich P., Kim J.T. and Ziegler J. (editors) “Advances in Artificial Life - Proceedings of the 7th European Conference on Artificial Life (ECAL),” volume 2801 of *Lecture Notes in Artificial Intelligence*. Springer Verlag Berlin, Heidelberg, **2003**. In print. 65
- [84] Fiorillo C.D., Tobler P.N. and Schultz W.: “Discrete coding of reward probability and uncertainty by dopamine neurons.” *Science*, **299**(5614): pp. 1898–902, **2003**. 105
- [85] Floreano D. and Mondada F.: “Evolutionary neurocontrollers for autonomous mobile robots.” *Neural Networks*, **11**: pp. 1461–1478, **1998**. 56
- [86] Floreano D. and Mondada F.: “Hardware solutions for evolutionary robotics.” In Husbands P. and Meyer J. (editors) “Proc. of the First European Workshop on Evolutionary Robotics,” Springer, **1998**. 56
- [87] Floresco S.B., Blaha C.D., Yang C.R. and Phillips A.G.: “Modulation of hippocampal and amygdalar-evoked activity of nucleus accumbens neurons by dopamine: cellular mechanisms of input selection.” *J Neurosci*, **21**(8): pp. 2851–60, **2001**. 99
- [88] Foster D.J., Morris R.G.M. and Dayan P.: “A model of hippocampally dependent navigation, using the temporal difference learning rule.” *Hippocampus*, **10**(1): pp. 1–16, **2000**. 37, 43, 100, 125
- [89] Frank L.M., Brown E.N. and Wilson M.A.: “Trajectory Encoding in the Hippocampus and Entorhinal Cortex.” *Neuron*, **27**: pp. 169–178, **2000**. 15, 84
- [90] Frank L.M., Brown E.N. and Wilson M.A.: “A Comparison of the Firing Properties of Putative Excitatory and Inhibitory Neurons From CA1 and the Entorhinal Cortex.” *J Neurophysiol*, **86**(4): pp. 2029–2040, **2001**. 15, 84

- [91] Freund T.F. and Antal M.: “GABA-containing neurons in the septum control inhibitory interneurons in the hippocampus.” *Nature*, **336**(6195): pp. 170–3, **1988**. 8
- [92] Freund T.F., Powell J.F. and Smith A.D.: “Tyrosine hydroxylase-immunoreactive boutons in synaptic contact with identified striatonigral neurons, with particular reference to dendritic spines.” *Neuroscience*, **13**(4): pp. 1189–215, **1984**. 30, 53, 99
- [93] Gabor D.: “Theory of communication.” *Journal of the IEE*, **93**: pp. 429–457, **1946**. 60
- [94] Gallistel C.R.: *The organization of learning*. MIT Press, Cambridge, MA, **1990**. 33
- [95] Gaussier P., Joulain C., Banquet J.P., Leprêtre S. and Revel A.: “The visual homing problem: An example of robotics/biology cross fertilization.” *Robotics and Autonomous Systems*, **30**(1-2): pp. 155–180, **2000**. 24, 27, 43, 55, 124
- [96] Gaussier P., Leprêtre S., Joulain C., Revel A., Quoy M. and Banquet J.P.: “Animal and robot learning: Experiments and models about visual navigation.” In “7th European Workshop on Learning Robots,” Edinburgh, UK, **1998**. 24, 27, 43, 55, 124
- [97] Gaussier P., Revel A., Banquet J.P. and Babeau V.: “From view cells and place cells to cognitive map learning: processing stages of the hippocampal system.” *Biol Cybern*, **86**(1): pp. 15–28, **2002**. 24, 27, 43, 55, 124
- [98] Gavrillov V.V., Wiener S.I. and Berthoz A.: “Whole body rotations enhance hippocampal theta rhythm slow activity in awake rats passively transported on a mobile robot.” *Annals of the New York Academy of Sciences*, **781**: pp. 385–398, **1996**. 12
- [99] Georgopoulos A.P., Kettner R.E. and Schwartz A.: “Primate motor cortex and free arm movements to visual targets in three-dimensional space. II. Coding of the direction of movement by a neuronal population.” *Neuroscience*, **8**: pp. 2928–2937, **1988**. 8
- [100] Georgopoulos A.P., Schwartz A. and Kettner R.E.: “Neuronal population coding of movement direction.” *Science*, **233**: pp. 1416–1419, **1986**. 8, 63, 80
- [101] Gerstner W. and Abbott L.F.: “Learning navigational maps through potentiation and modulation of hippocampal place cells.” *J Comput Neurosci*, **4**(1): pp. 79–94, **1997**. 42, 125

- [102] Gerstner W. and Kistler W.: *Spiking Neuron Models: Single Neurons, Populations, Plasticity*. Cambridge University Press, **2002**. 104
- [103] Gibson J.J.: *The Senses Considered as Perceptual Systems*. Houghton Mifflin Company, Boston, USA, **1966**. 17
- [104] Goelet P., Castellucci V.F., Schacher S. and Kandel E.R.: “The long and the short of long-term memory—a molecular framework.” *Nature*, **322**(6078): pp. 419–22, **1986**. 105
- [105] Gordon G.J.: “Stable function approximation in dynamic programming.” In Frieditis A. and Russell S. (editors) “Proceedings of the Twelfth International Conference on Machine Learning,” pp. 261–268. Morgan Kaufmann, San Francisco, CA, **1995**. 53, 113
- [106] Gothard K.M., Hoffman K.L., Battaglia F.P. and McNaughton B.L.: “Dentate Gyrus and CA1 Ensemble Activity during Spatial Reference Frame Shifts in the Presence and Absence of Visual Input.” *J. Neurosci.*, **21**(18): pp. 7284–7292, **2001**. 15, 40
- [107] Gothard K.M., Skaggs W.E. and McNaughton B.L.: “Dynamics of mismatch correction in the hippocampal ensemble code for space: Interaction between path integration and environmental cues.” *Journal of Neuroscience*, **16**(24): pp. 8027–8040, **1996**. 13, 40
- [108] Gothard K.M., Skaggs W.E., Moore K.M. and McNaughton B.L.: “Binding of hippocampal CA1 neural activity to multiple reference frames in a landmark-based navigation task.” *Journal of Neuroscience*, **16**(2): pp. 823–835, **1996**. 12, 40
- [109] Gurden H., Tassin J.P. and Jay T.M.: “Integrity of the mesocortical dopaminergic system is necessary for complete expression of in vivo hippocampal-prefrontal cortex long-term potentiation.” *Neuroscience*, **94**(4): pp. 1019–27, **1999**. 54, 105
- [110] Hafner V.V., Fend M., Lungarella M., Pfeifer R., König P. and Körding K.P.: “Optimal coding for naturally occurring whisker deflections.” In “Artificial Neural Networks / Neural Information Processing - ICANN/ICONIP,” pp. 805–812. Springer, **2003**. 65
- [111] Harris K.D., Henze D.A., Hirase H., Leinekugel X., Dragoi G., Czurko A. and Buzsáki G.: “Spike train dynamics predicts theta-related phase precession in hippocampal pyramidal cells.” *Nature*, **417**(6890): pp. 738–41, **2002**. 14

- [112] Hasselmo M.E., Bodelon C. and Wyble B.P.: “A proposed function for hippocampal theta rhythm: separate phases of encoding and retrieval enhance reversal of prior learning.” *Neural Comput*, **14**(4): pp. 793–817, **2002**. 103
- [113] Hasselmo M.E. and Bower J.M.: “Acetylcholine and memory.” *Trends in Neurosciences*, **16**(6): pp. 218–222, **1993**. 12
- [114] Hasselmo M.E., Hay J., Ilyn M. and Gorchetchnikov A.: “Neuromodulation, theta rhythm and rat spatial navigation.” *Neural Netw*, **15**(4-6): pp. 689–707, **2002**. 103
- [115] Hastings N.B., Seth M.I., Tanapat P., Rydel T.A. and Gould E.: “Granule neurons generated during development extend divergent axon collaterals to hippocampal area CA3.” *J Comp Neurol*, **452**(4): pp. 324–33, **2002**. 11
- [116] Heimer L., Zahn D.S. and Alheid G.F.: “Basal Ganglia.” In Paxinos G. (editor) “The Rat Nervous System,” chapter 23, pp. 579–628. Academic Press, second edition, **1995**. 29, 30
- [117] Herrmann J.M., Pawelzik K. and Geisel T.: “Self-localization of autonomous robots by hidden representations.” *Autonomous Robots*, **7**(1): pp. 31–40, **1999**. 77
- [118] Herrmann J.M., Pawelzik K. and Geisel T.: “Learning predictive representations.” *Neurocomputing*, **32**: pp. 785–791, **2000**. 77
- [119] Hill A.J.: “First occurrence of hippocampal spatial firing in a new environment.” *Experimental Neurology*, **62**: pp. 282–297, **1978**. 14
- [120] Hill A.J. and Best P.J.: “Effects of deafness and blindness on the spatial correlates of hippocampal unit activity in the rat.” *Experimental neurology*, **74**: pp. 204–217, **1981**. 13
- [121] Hollerman J.R. and Schultz W.: “Dopamine neurons report an error in the temporal prediction of reward during learning.” *Nat Neurosci*, **1**(4): pp. 304–9, **1998**. 53
- [122] Houk J.C., Adams J.L. and Barto A.G.: “A Model of How the Basal Ganglia Generate and Use Neural Signals That Predict Reinforcement.” In Houk J.C., Davis J.L. and Beiser D.G. (editors) “Models of Information Processing in the Basal Ganglia,” chapter 13, pp. 249–270. MIT Press, Cambridge, Massachusetts, USA, **1995**. 29, 30, 32, 51, 53, 99, 100, 105, 110

- [123] Hubel D.H. and Wiesel T.N.: “Receptive fields, binocular interaction and functional architecture in the cat’s visual cortex.” *Journal of Physiology*, **160**: pp. 106–154, **1962**. 60
- [124] Hughes A.: “The Topography of Visuon in Mammals of Contrasting Life Style: Comparative Optics and Retinal Organisation.” In Crescitelli F. (editor) “The Visual System in Vertebrates,” volume 7/5 of *Handbook of Sensory Physiology*, chapter 11, pp. 613–756. Springer-Verlag, Berlin, **1977**. 27, 56
- [125] Huxter J., Burgess N. and O’Keefe J.: “Independent rate and temporal coding in hippocampal pyramidal cells.” *Nature*, **425**(6960): pp. 828–32, **2003**. 14
- [126] Insausti R., Herrero M.T. and Witter M.P.: “Entorhinal cortex of the rat: Cytoarchitectonic subdivision and the origin and distribution of cortical efferents.” *Hippocampus*, **7**: pp. 146–83, **1997**. 8, 10
- [127] Insausti R., Marcos P., Arroyo-Jimenez M.M., Blaizot X. and Martinez-Marcos A.: “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): pp. 557–560, **2002**. 10
- [128] Jaakkola T., Jordan M.I. and Singh S.P.: “On the convergence of stochastic iterative dynamic programming algorithms.” *Neural Computation*, **6**(6): pp. 1185–1201, **1994**. 51, 52
- [129] Jarrard L.E.: “On the role of the hippocampus in learning and memory in the rat.” *Behavioral Neural Biology*, **60**(1): pp. 9–26, **1993**. 31, 38, 39
- [130] Jensen O. and Lisman J.E.: “Hippocampal CA3 region predicts memory sequences: accounting for the phase precession of place cells.” *Learn Mem*, **3**(2-3): pp. 279–87, **1996**. 14
- [131] Jiang H., Stein B.E. and McHaffie J.G.: “Opposing basal ganglia processes shape midbrain visuomotor activity bilaterally.” *Nature*, **424**(6943): pp. 982–6, **2003**. 19
- [132] Jiang W. and Stein B.E.: “Cortex controls multisensory depression in superior colliculus.” *J Neurophysiol*, **90**(4): pp. 2123–35, **2003**. 19, 94
- [133] Jiang W., Wallace M.T., Jiang H., Vaughan J.W. and Stein B.E.: “Two cortical areas mediate multisensory integration in superior colliculus neurons.” *J Neurophysiol*, **85**(2): pp. 506–22, **2001**. 19

- [134] Jones J.P. and Palmer L.A.: “An evaluation of the two-dimensional gabor filter model of simple receptive fields in cat strata cortex.” *Journal of Neurophysiology*, **58(6)**: pp. 1233–1258, **1987**. 60
- [135] Jung M.W. and McNaughton B.L.: “Spatial selectivity of unit activity in the hippocampal granular layer.” *Hippocampus*, **3(2)**: pp. 165–182, **1993**. 15, 36
- [136] Karni A. and Bertini G.: “Learning perceptual skills: behavioral probes into adult cortical plasticity.” *Curr Opin Neurobiol*, **7(4)**: pp. 530–5, **1997**. 31
- [137] Karni A., Meyer G., Rey-Hipolito C., Jezard P., Adams M.M., Turner R. and Ungerleider L.G.: “The acquisition of skilled motor performance: fast and slow experience-driven changes in primary motor cortex.” *Proc Natl Acad Sci U S A*, **95(3)**: pp. 861–8, **1998**. 31
- [138] Kerr J.N. and Wickens J.R.: “Dopamine D-1/D-5 receptor activation is required for long-term potentiation in the rat neostriatum in vitro.” *J Neurophysiol*, **85(1)**: pp. 117–24, **2001**. 54, 105
- [139] Knierim J.J., Kudrimoti H.S. and McNaughton B.L.: “Place cells, head direction cells, and the learning of landmark stability.” *Journal of Neuroscience*, **15**: pp. 1648–1659, **1995**. 17, 36, 39
- [140] Koene R.A., Gorchetchnikov A., Cannon R.C. and Hasselmo M.E.: “Modeling goal-directed spatial navigation in the rat based on physiological data from the hippocampal formation.” *Neural Netw*, **16(5-6)**: pp. 577–84, **2003**. 103, 106
- [141] Kosel K.C., Hoesen G.W.V. and West J.R.: “Olfactory bulb projections to the parahippocampal area of the rat.” *Journal of Comparative Neurology*, **198**: pp. 467–482, **1981**. 10
- [142] Kubie J.L. and Ranck J.B.: “Sensory-behavioral correlates in individual hippocampus neurons in three situations: Space and context.” In Seifert W. (editor) “Neurobiology of the Hippocampus,” pp. 433–447. Academic Press, New York, **1983**. 14, 126
- [143] Kuhn H.G., Dickinson-Anson H. and Gage F.H.: “Neurogenesis in the dentate gyrus of the adult rat: age-related decrease of neuronal progenitor proliferation.” *J Neurosci*, **16(6)**: pp. 2027–33, **1996**. 11
- [144] Kulla A. and Manahan-Vaughan D.: “Depotentiation in the dentate gyrus of freely moving rats is modulated by D1/D5 dopamine receptors.” *Cereb Cortex*, **10(6)**: pp. 614–20, **2000**. 105, 107

- [145] Lavenex P. and Schenk F.: “Integration of olfactory information in a spatial representation enabling accurate arm choice in the radial arm maze.” *Learning & Memory*, **2**(6): pp. 299–319, **1996**. 13
- [146] Lavenex P. and Schenk F.: “Olfactory traces and spatial learning in rats.” *Animal Behaviour*, **56**: pp. 1129–1136, **1998**. 13
- [147] Lechner H.A., Squire L.R. and Byrne J.H.: “100 years of consolidation–remembering Muller and Pilzecker.” *Learn Mem*, **6**(2): pp. 77–87, **1999**. 32, 105
- [148] Lee T.S.: “Image representation using 2D gabor wavelets.” *Ieee Transactions On Pattern Analysis and Machine Intelligence*, **18**(10): pp. 959–971, **1996**. 60
- [149] Legault M., Rompré P.P. and Wise R.A.: “Chemical Stimulation of the Ventral Hippocampus Elevates Nucleus Accumbens Dopamine by Activating Dopaminergic Neurons of the Ventral Tegmental Area.” *Journal of Neuroscience*, **20**(4): pp. 1635–1642, **2000**. 8, 99
- [150] Leonhard C.L., Stackman R.W. and Taube J.S.: “Head direction cells recorded from the lateral mammillary nuclei in rats.” *Society for Neuroscience Abstracts*, **22**: p. 1873, **1996**. 17
- [151] Linden D.J.: “The return of the spike: postsynaptic action potentials and the induction of LTP and LTD.” *Neuron*, **22**(4): pp. 661–6, **1999**. 32, 104
- [152] Liu P. and Bilkey D.K.: “Parallel involvement of perirhinal and lateral entorhinal cortex in the polysynaptic activation of hippocampus by olfactory inputs.” *Hippocampus*, **7**(3): pp. 296–306, **1997**. 10
- [153] Magee J.C.: “Dendritic mechanisms of phase precession in hippocampal CA1 pyramidal neurons.” *J Neurophysiol*, **86**(1): pp. 528–32, **2001**. 14
- [154] Manahan-Vaughan D. and Kulla A.: “Regulation of depotentiation and long-term potentiation in the dentate gyrus of freely moving rats by dopamine D2-like receptors.” *Cereb Cortex*, **13**(2): pp. 123–35, **2003**. 105, 107
- [155] Margules J. and Gallistel C.R.: “Heading in the Rat: Determination by Environmental Shape.” *Animal Learning & Behavior*, **16**(4): pp. 404–410, **1988**. 39
- [156] Markram H., Lubke J., Frotscher M. and Sakmann B.: “Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs.” *Science*, **275**(5297): pp. 213–5, **1997**. 32, 104

- [157] Marks L.E.: *The Unity of the Senses: Interrelations among the Modalities*. Academic Press, New York, USA, **1978**. 18
- [158] Markus E.J., Barnes C.A., McNaughton B.L., Gladden V.L. and Skaggs W.E.: “Spatial information content and reliability of hippocampal CA1 neurons: Effects of visual input.” *Hippocampus*, **4**: pp. 410–421, **1994**. 13, 15, 21, 36, 40, 57
- [159] Markus E.J., Qin Y., Leonard B., Skaggs W.E., McNaughton B.L. and Barnes C.A.: “Interactions between location and task affect the spatial and direction firing of hippocampal neurons.” *Journal of Neuroscience*, **15**: pp. 7079–7094, **1995**. 13, 15, 36, 126
- [160] Martens S., Carpenter G.A. and Gaudio P.: “Neural Sensor Fusion for Spatial Visualization on a Mobile Robot.” In Schenker P.S. and McKee G.T. (editors) “Proceedings of SPIE, Sensor Fusion and Decentralized Control in Robotic Systems,” Proceedings of SPIE, **1998**. 27
- [161] McDonald R.J. and White N.M.: “Parallel information processing in the water maze: Evidence for independent memory systems involving dorsal striatum and hippocampus.” *Behavioral and Neural Biology*, **61**: pp. 260–270, **1994**. 30, 35, 38, 99
- [162] McNaughton B.L., Barnes C.A., Gerrard J.L., Gothard K., Jung M.W., Knierim J.J., Kudrimoti H., Qin Y., Skaggs W.E., Suster M. and Weaver K.L.: “Deciphering the hippocampal polyglot: the hippocampus as a path integration system.” *J Exp Biol*, **199**(1): pp. 173–85, **1996**. 15, 17
- [163] McNaughton B.L., Barnes C.A. and O’Keefe J.: “The contributions of position, direction, and velocity to single unit activity in the hippocampus of freely-moving rats.” *Experimental Brain Research*, **52**: pp. 41–49, **1983**. 12, 13, 15, 36
- [164] McNaughton B.L., Chen L.L. and Markus E.J.: “Dead reckoning, landmark learning, and the sense of direction: A neurophysiological and computational hypothesis.” *Journal of Cognitive Neuroscience*, **3**: p. 190, **1991**. 17, 35, 39
- [165] McNaughton B.L., Leonard B. and Chen L.: “Cortical-hippocampal interactions and cognitive mapping: A hypothesis based on reintegration of the parietal and inferotemporal pathways for visual processing.” *Psychobiology*, **17**(3): pp. 230–235, **1989**. 36
- [166] Mehta M.R.: “Neuronal dynamics of predictive coding.” *Neuroscientist*, **7**(6): pp. 490–5, **2001**. 40

- [167] Mehta M.R., Barnes C.A. and McNaughton B.L.: “Experience-dependent, asymmetric expansion of hippocampal place fields.” In “Proceedings Natl. Acad. Science USA,” volume 94, pp. 8918–8921. **1997.** 13, 14, 40
- [168] Mehta M.R., Lee A.K. and Wilson M.A.: “Role of experience and oscillations in transforming a rate code into a temporal code.” *Nature*, **417**(6890): pp. 741–6, **2002.** 14, 36
- [169] Mehta M.R., Quirk M.C. and Wilson M.A.: “Experience-dependent, asymmetric shape of hippocampal receptive fields.” *Neuron*, **25**: pp. 707–715, **2000.** 13, 36, 40
- [170] Melamed O., Gerstner W., Maass W., Tsodyks M. and Markram H.: “Coding and learning of behavioral sequences.” *Trends in Neurosciences*, **27**: pp. 11–14, **2004.** 14
- [171] Meredith M.A. and Stein B.E.: “Spatial determinants of multisensory integration in cat superior colliculus neurons.” *J Neurophysiol*, **75**(5): pp. 1843–57, **1996.** 19
- [172] Miller G.A.: “The magical number seven plus or minus two: some limits on our capacity for processing information.” *Psychol Rev*, **63**(2): pp. 81–97, **1956.** 33
- [173] Miller R.: *Cortico-Hippocampal interplay and the representation of contexts in the brain.* Springer-Verlag, **1991.** 12
- [174] Miller V.M. and Best P.J.: “Spatial correlates of hippocampal unit activity are altered by lesions of the fornix and entorhinal cortex.” *Brain Research*, **194**: pp. 311–323, **1980.** 12, 57, 67
- [175] Mink J.W.: “The basal ganglia: focused selection and inhibition of competing motor programs.” *Prog Neurobiol*, **50**(4): pp. 381–425, **1996.** 29, 30, 110
- [176] Mittelstaedt H.: “The role of multimodal convergence in homing by path integration.” *Fortschritte der Zoologie*, **28**: pp. 197–212, **1983.** 17
- [177] Mittelstaedt M.L. and Mittelstaedt H.: “Homing by path integration in a mammal.” *Naturwissenschaften*, **67**: pp. 566–567, **1980.** 15, 33, 34
- [178] Mizumori S.J., McNaughton B.L., Barnes C.A. and Fox K.B.: “Preserved spatial coding in hippocampal CA1 pyramidal cells during reversible suppression of CA3c output: evidence for pattern completion in hippocampus.” *J Neurosci*, **9**(11): pp. 3915–28, **1989.** 12

- [179] Mizumori S.J.Y., Ward K.E. and Lavoie A.M.: “Medial septal modulation of entorhinal single unit activity in anesthetized and freely moving rats.” *Brain Research*, **570**: pp. 188–197, **1992**. 15, 17, 36
- [180] Mizumori S.J.Y. and Williams J.D.: “Directionally selective mnemonic properties of neurons in the lateral dorsal nucleus of the thalamus of rats.” *Journal of Neuroscience*, **13**: pp. 4015–4028, **1993**. 17
- [181] Mobini S., Chiang T.J., Ho M.Y., Bradshaw C.M. and Szabadi E.: “Effects of central 5-hydroxytryptamine depletion on sensitivity to delayed and probabilistic reinforcement.” *Psychopharmacology (Berl)*, **152**(4): pp. 390–7, **2000**. 54
- [182] Mondada F., Franzi E. and Ienne P.: “Mobile robot miniaturization: A tool for investigation in control algorithms.” In Yoshikawa T. and Miyazaki F. (editors) “Proceedings of the Third International Symposium on Experimental Robotics,” pp. 501–513. Springer Verlag, Tokyo, **1994**. 56
- [183] Morris R.G.M.: “Spatial localization does not require the presence of local cues.” *Learning and Motivation*, **12**: pp. 239–260, **1981**. 36
- [184] Morris R.G.M., Garrud P., Rawlins J.N.P. and O’Keefe J.: “Place navigation impaired in rats with hippocampal lesions.” *Nature*, **297**: pp. 681–683, **1982**. 36
- [185] Morris R.G.M., Schenk F., Tweedie F. and Jarrard L.E.: “Ibotenate lesions of the hippocampus and/or subiculum: Dissociating components of allocentric spatial learning.” *European Journal of Neuroscience*, **2**: pp. 1016–1028, **1990**. 36, 37
- [186] Muller R.U., Bostock E., Taube J.S. and Kubie J.L.: “On the directional firing properties of hippocampal place cells.” *Journal of Neuroscience*, **14**(12): pp. 7235–7251, **1994**. 13
- [187] Muller R.U. and Kubie J.L.: “The effects of changes in the environment on the spatial firing of hippocampal complex-spike cells.” *Journal of Neuroscience*, **7**: pp. 1951–1968, **1987**. 12, 13, 14, 36, 67
- [188] Muller R.U. and Kubie J.L.: “The firing of hippocampal place cells predicts the future position of freely moving rats.” *Journal of Neuroscience*, **9**: pp. 4101–4110, **1989**. 14, 40
- [189] Muller R.U., Kubie J.L. and Ranck J.J.B.: “Spatial firing patterns of hippocampal complex-spike cells in a fixed environment.” *Journal of Neuroscience*, **7**: pp. 1935–1950, **1987**. 12, 14, 36, 67, 80

- [190] Muller R.U., Kubie J.L. and Saypoff R.: "The hippocampus as a cognitive graph." *Hippocampus*, **1(3)**: pp. 243–246, **1991**. 36
- [191] Munoz D.P. and Guitton D.: "Control of orienting gaze shifts by the tectoreticulospinal system in the head-free cat. II. Sustained discharges during motor preparation and fixation." *J Neurophysiol*, **66(5)**: pp. 1624–41, **1991**. 18
- [192] Munoz D.P., Pelisson D. and Guitton D.: "Movement of neural activity on the superior colliculus motor map during gaze shifts." *Science*, **251(4999)**: pp. 1358–60, **1991**. 18
- [193] Nadel L. and Bohbot V.: "Consolidation of memory." *Hippocampus*, **11(1)**: pp. 56–60, **2001**. 126
- [194] Natsume K., Hallworth N.E., Szgatti T.L. and Bland B.H.: "Hippocampal Theta-Related Cellular Activity in the Superior Colliculus of the Urethane-Anesthetized Rat." *Hippocampus*, **9**: pp. 500–509, **1999**. 19
- [195] O'Keefe J. and Burgess N.: "Geometric determinants of the place fields of hippocampal neurons." *Nature*, **381**: pp. 425–428, **1996**. 12, 13, 36, 67
- [196] O'Keefe J. and Conway D.H.: "Hippocampal place units in the freely moving rat: why they fire where they fire." *Experimental Brain Research*, **31**: pp. 573–590, **1978**. 12, 13, 14, 36, 57, 67
- [197] O'Keefe J. and Dostrovsky J.: "The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat." *Brain Research*, **34**: pp. 171–175, **1971**. 7, 8, 11, 36
- [198] O'Keefe J. and Nadel L.: *The Hippocampus as a Cognitive Map*. Clarendon Press, Oxford, **1978**. 33, 35, 36
- [199] O'Keefe J. and Recce M.: "Phase relationship between hippocampal place units and the EEG theta rhythm." *Hippocampus*, **3**: pp. 317–330, **1993**. 12, 14, 36
- [200] O'Keefe J. and Speakman A.: "Single unit activity in the rat hippocampus during a spatial memory task." *Experimental Brain Research*, **68**: pp. 1–27, **1987**. 12, 13, 36, 67
- [201] O'Leary F.A., Byrne J.H. and Cleary L.J.: "Long-term structural remodeling in Aplysia sensory neurons requires de novo protein synthesis during a critical time period." *J Neurosci*, **15(5 Pt 1)**: pp. 3519–25, **1995**. 105
- [202] Olton D.S., Becker J.T. and Handelmann G.E.: "Hippocampus, Space, and Memory." *Behavioral and Brain Sciences*, **2(3)**: pp. 313–322, **1979**. 31, 38

- [203] Olton D.S. and Papas B.C.: “Spatial Memory and Hippocampal Function.” *Neuropsychologia*, **17**(6): pp. 669–682, **1979**. 31
- [204] Olton D.S. and Samuelson R.J.: “Remembrance of Places Passed: Spatial Memory in Rats.” *Journal of Experimental Psychology-animal Behavior Processes*, **2**(2): pp. 97–116, **1976**. 38
- [205] Otmakhov N., Griffith L.C. and Lisman J.E.: “Postsynaptic Inhibitors of Calcium/Calmodulin-Dependent Protein Kinase Type II Block Induction But Not Maintenance of Pairing-Induced Long-Term Potentiation.” *Journal of Neuroscience*, **17**(14): pp. 5357–5365, **1997**. 32, 105
- [206] Otmakhova N.A. and Lisman J.E.: “D1/D5 Dopamine Receptor Activation Increases the Magnitude of Early Long-Term Potentiation at CA1 Hippocampal Synapses.” *Journal of Neuroscience*, **16**(23): pp. 7478–7486, **1996**. 54, 105
- [207] Otmakhova N.A. and Lisman J.E.: “D1/D5 Dopamine Receptors Inhibit Depotentiation at CA1 Synapses via cAMP-Dependent Mechanism.” *Journal of Neuroscience*, **18**(4): pp. 1270–1279, **1998**. 32, 53, 54, 105, 106, 107
- [208] Packard M., Hirsh R. and White N.: “Differential effects of fornix and caudate nucleus lesions on two radial maze tasks: evidence for multiple memory systems.” *J. Neurosci.*, **9**(5): pp. 1465–1472, **1989**. 39
- [209] Packard M.G. and McGaugh J.L.: “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): pp. 439–446, **1992**. 37, 38, 41, 43, 99, 123, 124, 125
- [210] Packard M.G. and McGaugh J.L.: “Inactivation of hippocampus or caudate nucleus with lidocaine differentially affects expression of place and response learning.” *Neurobiology of Learning and Memory*, **65**(1): pp. 65–72, **1996**. 30, 35, 39
- [211] Packard M.G. and White N.M.: “Dissociation of Hippocampus and Caudate-Nucleus Memory-Systems by Posttraining Intracerebral Injection of Dopamine Agonists.” *Behavioral Neuroscience*, **105**(2): pp. 295–306, **1991**. 30, 35
- [212] Partridge J.G., Apparsundaram S., Gerhardt G.A., Ronesi J. and Lovinger D.M.: “Nicotinic acetylcholine receptors interact with dopamine in induction of striatal long-term depression.” *J Neurosci*, **22**(7): pp. 2541–9, **2002**. 54
- [213] Pavlides C. and Winson J.: “Influence of hippocampal place cell firing in the awake state on the activity of these cells during subsequent sleep episodes.” *Journal of Neuroscience*, **9**(8): pp. 2907–2918, **1989**. 14

- [214] Peck C.K., Baro J.A. and Warder S.M.: “Sensory integration in the deep layers of superior colliculus.” *Prog Brain Res*, **95**: pp. 91–102, **1993**. 19
- [215] Peng J. and Williams R.J.: “Incremental Multi-Step Q-Learning.” *Machine Learning*, **22**(1-3): pp. 283–290, **1996**. 52, 113
- [216] Piaget J.: *The Construction of Reality in the Child*. Basic Books Inc., New York, USA, **1954**. 17
- [217] Pico R.M., Gerbrandt L.K., Pondel M. and Ivy G.: “During stepwise cue deletion, rat place behaviors correlate with place unit responses.” *Brain Research*, **330**: pp. 369–373, **1985**. 13
- [218] Precup D., Sutton R.S. and Dasgupta S.: “Off-Policy Temporal-Difference Learning with Function Approximation.” In “Proc. 18th International Conf. on Machine Learning,” pp. 417–424. Morgan Kaufmann, **2001**. 53, 113
- [219] Precup D., Sutton R.S. and Singh S.: “Eligibility Traces for Off-Policy Policy Evaluation.” In “Proc. 17th International Conf. on Machine Learning,” pp. 759–766. Morgan Kaufmann, San Francisco, CA, **2000**. 52, 53, 113
- [220] Quirk G.J., Muller R.U. and Kubie J.L.: “The Firing of Hippocampal Place Cells in the Dark depends on the Rat’s Recent Experience.” *Journal of Neuroscience*, **10**(6): pp. 2008–2017, **1990**. 13, 21, 57
- [221] Quirk G.J., Muller R.U., Kubie J.L. and Ranck Jr. J.B.: “The positional firing properties of medial entorhinal neurons: Description and comparison with hippocampal place cells.” *Journal of Neuroscience*, **12**(5): pp. 1945–1963, **1992**. 11, 12, 15, 36, 67, 84
- [222] Ranck Jr. J.B.: “Head-direction cells in the deep cell layers of dorsal presubiculum in freely moving rats.” *Society for Neuroscience Abstracts*, **10**: p. 599, **1984**. 17, 67, 79
- [223] Rao R.P. and Sejnowski T.J.: “Spike-Timing-Dependent Hebbian Plasticity as Temporal Difference Learning.” *Neural Computation*, **13**: pp. 2221–2237, **2001**. 104
- [224] Rasmusson D.D.: “The role of acetylcholine in cortical synaptic plasticity.” *Behav Brain Res*, **115**(2): pp. 205–18, **2000**. 54
- [225] Redgrave P., Prescott T.J. and Gurney K.: “Is the short-latency dopamine response too short to signal reward error?” *Trends Neurosci*, **22**(4): pp. 146–51, **1999**. 53

- [226] Redish A.D.: *Beyond the Cognitive Map, From Place Cells to Episodic Memory*. MIT Press-Bradford Books, London, **1999**. 31, 36
- [227] Redish A.D., Elga A.N. and Touretzky D.S.: “A coupled attractor model of the rodent head direction system.” *Network*, **7(4)**: pp. 671–685, **1996**. 33, 77
- [228] Redish A.D. and Touretzky D.S.: “Cognitive maps beyond the hippocampus.” *Hippocampus*, **7(1)**: pp. 15–35, **1997**. 17, 23, 40, 55, 123
- [229] Redish A.D. and Touretzky D.S.: “Navigating with landmarks: Computing goal locations from place codes.” In Ikeuchi K. and Veloso M. (editors) “Symbolic Visual Learning,” chapter 12, pp. 325–351. Oxford University Press, **1997**. 23, 55, 123
- [230] Redish A.D. and Touretzky D.S.: “The role of the hippocampus in solving the Morris water maze.” *Neural Comput*, **10(1)**: pp. 73–111, **1998**. 35
- [231] Reynolds J.N., Hyland B.I. and Wickens J.R.: “A cellular mechanism of reward-related learning.” *Nature*, **413(6851)**: pp. 67–70, **2001**. 54, 105
- [232] Rolls E.T.: “A model of the operation of the hippocampus and entorhinal cortex in memory.” *International Journal of Neural Systems*, **6**: pp. 51–70, **1995**. 67
- [233] Rossier J., Haerberli C. and Schenk F.: “Auditory cues support place navigation in rats when associated with a visual cue.” *Behavioural Brain Research*, **117(1-2)**: pp. 209–214, **2000**. 13
- [234] Rotenberg A. and Muller R.U.: “Variable place cell coupling to a continuously viewed stimulus: Evidence that the hippocampus acts as a perceptual system.” *Royal Society Philosophical Transactions: Biological Sciences*, **352(1360)**: pp. 1505–1514, **1997**. 13
- [235] Rumelhart D.E. and Zipser D.: “Feature discovery by competitive learning.” In Rumelhart D.E. and McClelland J.L. (editors) “Parallel distributed processing 1,” MIT Press, Cambridge, Massachusetts, **1986**. 21, 22
- [236] Sakurai Y.: “Involvement of auditory cortical and hippocampal neurons in auditory working memory and reference memory in the rat.” *Journal of Neuroscience*, **14(5)**: pp. 2606–2623, **1994**. 31
- [237] Salinas E. and Abbott L.F.: “Vector reconstruction from firing rates.” *Journal of Computational Science*, **1**: pp. 89–107, **1994**. 8, 63, 80

- [238] Samsonovich A. and McNaughton B.L.: “Path integration and cognitive mapping in a continuous attractor neural network model.” *Journal of Neuroscience*, **17(15)**: pp. 5900–5920, **1997**. 17, 33, 40, 77
- [239] Sanger T.D.: “Neural population codes.” *Curr Opin Neurobiol*, **13(2)**: pp. 238–49, **2003**. 63, 80
- [240] Santamaría J.C., Sutton R.S. and Ram A.: “Experiments with reinforcement learning in problems with continuous state and action spaces.” *Adaptive Behavior*, **6(2)**: pp. 163–218, **1998**. 100
- [241] Save E., Cressant A., Thinus-Blanc C. and Poucet B.: “Spatial Firing of Hippocampal Place Cells in Blind Rats.” *Journal of Neuroscience*, **18(5)**: pp. 1818–1826, **1998**. 13
- [242] Save E., Nerad L. and Poucet B.: “Contribution of Multiple Sensory Information to Place Field Stability in Hippocampal Place Cells.” *Hippocampus*, **10**: pp. 64–76, **2000**. 13, 21
- [243] Schiller J., Major G., Koester H.J. and Schiller Y.: “NMDA spikes in basal dendrites of cortical pyramidal neurons.” *Nature*, **404(6775)**: pp. 285–9, **2000**. 32, 105
- [244] Schultz W.: “Predictive Reward Signal of Dopamine Neurons.” *Journal of Neurophysiology*, **80**: pp. 1–27, **1998**. 53, 99, 100, 105
- [245] Schultz W.: “Getting formal with dopamine and reward.” *Neuron*, **36(2)**: pp. 241–63, **2002**. 53, 54, 99
- [246] Schultz W., Apicella P., Romo R. and Scarnati E.: “Context-dependent Activity in Primate Striatum Reflecting Past and Future Behavioral Events.” In Houk J.C., Davis J.L. and Beiser D.G. (editors) “Models of Information Processing in the Basal Ganglia,” chapter 2, pp. 11–27. MIT Press, Cambridge, Massachusetts, USA, **1995**. 105, 106
- [247] Schultz W., Apicella P., Scarnati E. and Ljungberg T.: “Neuronal activity in monkey ventral striatum related to the expectation of reward.” *J Neurosci*, **12(12)**: pp. 4595–610, **1992**. 53, 99
- [248] Schultz W., Dayan P. and Montague P.R.: “A neural substrate of prediction and reward.” *Science*, **275**: pp. 1593–1599, **1997**. 30, 53, 99, 100, 105
- [249] Schwerdtfeger W.K., Buhl E.H. and Germroth P.: “Disynaptic Olfactory Input to the Hippocampus mediated by Stellate Cells in the Entorhinal Cortex.” *Journal of Comparative Neurology*, **292(2)**: pp. 163–177, **1990**. 10

- [250] Senn W., Markram H. and Tsodyks M.: “An algorithm for modifying neurotransmitter release probability based on pre- and post-synaptic spike timing.” *Neural Computation*, **13**(1): pp. 35–68, **2001**. 32, 104
- [251] Sesack S.R. and Pickel V.M.: “In the rat medial nucleus accumbens, hippocampal and catecholaminergic terminals converge on spiny neurons and are in apposition to each other.” *Brain Res*, **527**(2): pp. 266–79, **1990**. 30, 53, 99
- [252] Sesack S.R. and Pickel V.M.: “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): pp. 145–60, **1992**. 99
- [253] Shapiro M.L., Simon D.K., Olton D.S., Gage F.H., Nilsson O. and Bjorklund A.: “Intrahippocampal grafts of fetal basal forebrain tissue alter place fields in the hippocampus of rats with fimbria-fornix lesions.” *Neuroscience*, **32**(1): pp. 1–18, **1989**. 12, 67
- [254] Sharp P.E.: “Computer simulation of hippocampal place cells.” *Psychobiology*, **19**(2): pp. 103–115, **1991**. 20, 55, 87, 101, 123
- [255] Sharp P.E.: “Subicular cells generate similar spatial firing patterns in two geometrically and visually distinctive environments: Comparison with hippocampal place cells.” *Behavioral and Brain Research*, **85**: pp. 71–92, **1997**. 15, 17, 36, 79
- [256] Sharp P.E.: “Complimentary roles for hippocampal versus subicular/entorhinal place cells in coding place, context, and events.” *Hippocampus*, **9**: pp. 432–443, **1999**. 17
- [257] Sharp P.E.: “Subicular place cells expand/contract their spatial firing pattern to fit the size of the environment in an open field, but not in the presence of barriers: Comparison with hippocampal place cells.” *Behavioral Neuroscience*, **1999**. 15, 79
- [258] Sharp P.E. and Green C.: “Spatial correlates of firing patterns of single cells in the subiculum of freely moving rat.” *Journal of Neuroscience*, **14**(4): pp. 2339–2356, **1994**. 15, 36, 79, 84
- [259] Shen J., Kudrimoti H.S. and McNaughton B.L.: “Reactivation of neuronal ensembles in hippocampal dentate gyrus during sleep after spatial experience.” *Journal of Sleep Research*, **7**: pp. 6–16, **1998**. 14, 36

- [260] Siegrist C., Etienne A.S., Boulens V., Maurer R. and Rowe T.: “Homing by path integration in a new environment.” *Animal Behaviour*, **65**: pp. 185–194, **2003**. 17, 34, 35
- [261] Singh S.P., Jaakkola T., Littman M.L. and Szepesvári C.: “Convergence Results for Single-Step On-Policy Reinforcement-Learning Algorithms.” *Machine Learning*, **38**(3): pp. 287–308, **2000**. 51
- [262] Skaggs W.E., Knierim J.J., Kudrimoti H.S. and McNaughton B.L.: “A model of the neural basis of the rat’s sense of direction.” In Tesauro G., Touretzky D.S. and Leen T.K. (editors) “Advances in Neural Information Processing Systems 7,” pp. 173–180. MIT Press, Cambridge, MA, **1995**. 39
- [263] Skaggs W.E. and McNaughton B.L.: “Replay of neuronal firing sequences in rat hippocampus during sleep following spatial experience.” *Science*, **271**: pp. 1870–1873, **1996**. 14, 36
- [264] Skaggs W.E., McNaughton B.L., Wilson M.A. and Barnes C.A.: “Theta phase precession in hippocampal neuronal populations and the compression of temporal sequences.” *Hippocampus*, **6**(2): pp. 149–173, **1996**. 12, 14, 15
- [265] Smith A.D. and Bolam J.P.: “The neural network of the basal ganglia as revealed by the study of synaptic connections of identified neurones.” *Trends Neurosci*, **13**(7): pp. 259–65, **1990**. 30, 99, 110
- [266] Smith Y., Bennett B.D., Bolam J.P., Parent A. and Sadikot A.F.: “Synaptic relationships between dopaminergic afferents and cortical or thalamic input in the sensorimotor territory of the striatum in monkey.” *J Comp Neurol*, **344**(1): pp. 1–19, **1994**. 30, 53, 99
- [267] Smith Y., Bevan M.D., Shink E. and Bolam J.P.: “Microcircuitry of the direct and indirect pathways of the basal ganglia.” *Neuroscience*, **86**(2): pp. 353–87, **1998**. 30, 110
- [268] Stackman R.W. and Taube J.S.: “Firing properties of rat lateral mammillary single units: head direction, head pitch, and angular head velocity.” *J Neurosci*, **18**(21): pp. 9020–37, **1998**. 17
- [269] Steele R.J. and Morris R.G.M.: “Delay-dependent impairment of a matching-to-placetask with chronic and intrahippocampal infusion of the NMDA-antagonist D-AP5.” *Hippocampus*, **9**: pp. 118–136, **1999**. 37
- [270] Stein B.E. and Meredith M.A.: *The Merging of the Senses*. MIT Press, **1993**. 18, 19, 35, 93, 94

- [271] Stricanne B., Andersen R.A. and Mazzone P.: “Eye-centered, head-centered, and intermediate coding of remembered sound locations in area LIP.” *J Neurophysiol*, **76**(3): pp. 2071–6, **1996**. 35
- [272] Stringer S.M., Trappenberg T.P., Rolls E.T. and de Araujo I.E.: “Self-organizing continuous attractor networks and path integration: one-dimensional models of head direction cells.” *Network*, **13**(2): pp. 217–42, **2002**. 33, 77
- [273] Strösslin T. and Gerstner W.: “Reinforcement Learning in Continuous State and Action Space.” In “Artificial Neural Networks - ICANN 2003,” **2003**. 4
- [274] Strösslin T., Krebsler C., Arleo A. and Gerstner W.: “Combining Multimodal Sensory Input for Spatial Learning.” In Dorronsoro J.R. (editor) “Artificial Neural Networks - ICANN 2002,” pp. 87–92. Springer LNCS, **2002**. 4
- [275] Suri R.E. and Schultz W.: “Learning of sequential movements by neural network model with dopamine-like reinforcement signal.” *Exp Brain Res*, **121**(3): pp. 350–4, **1998**. 53, 100
- [276] Suri R.E. and Schultz W.: “A neural network model with dopamine-like reinforcement signal that learns a spatial delayed response task.” *Neuroscience*, **91**(3): pp. 871–90, **1999**. 53, 100
- [277] Suri R.E. and Schultz W.: “Temporal difference model reproduces anticipatory neural activity.” *Neural Comput*, **13**(4): pp. 841–62, **2001**. 53, 100
- [278] Sutherland R.J. and Hoising J.M.: “Posterior cingulate cortex and spatial memory: A microlimnology analysis.” In Vogt B.A. and Gabriel M. (editors) “Neurobiology of Cingulate Cortex and Limbic Thalamus: A Comprehensive Handbook,” pp. 461–477. Birkhauser, Boston, **1993**. 35
- [279] Sutherland R.J. and Rodriguez A.J.: “The role of the fornix/fimbria and some related subcortical structures in place learning and memory.” *Behavioral and Brain Research*, **32**: pp. 265–277, **1990**. 36, 37, 41, 43, 99, 123, 124, 125
- [280] Sutton R. and Barto A.G.: *Reinforcement Learning - An Introduction*. MIT Press, **1998**. 45, 47, 52, 100, 104, 109
- [281] Sutton R., McAllester D., Singh S. and Mansour Y.: “Policy Gradient Methods for Reinforcement Learning with Function Approximation.” In “Advances in Neural Information Processing Systems 12,” pp. 1057–1063. MIT Press, **1999**. 53
- [282] Sutton R.S.: “Learning to predict by the methods of temporal differences.” *Machine Learning*, **3**(1-3): pp. 9–44, **1988**. 51

- [283] Suzuki W.A. and Amaral D.G.: “Topographic organization of the reciprocal connections between the monkey entorhinal cortex and the perirhinal and parahippocampal cortices.” *J Neurosci*, **14**(3 Pt 2): pp. 1856–77, **1994**. 10
- [284] Tanila H., Sipilä P., Shapiro M. and Eichenbaum H.: “Brain aging: Impaired coding of novel environmental cues.” *Journal of Neuroscience*, **17**(13): pp. 5167–5174, **1997**. 14
- [285] Taube J.S.: “Qualitative analysis of head-direction cells recorded in the rat anterior thalamus.” *Society for Neuroscience Abstracts*, **18**: p. 708, **1992**. 17
- [286] Taube J.S. and Burton H.L.: “Head direction cell activity monitored in a novel environment and during a cue conflict situation.” *Journal of Neurophysiology*, **74**: pp. 1953–1971, **1995**. 39
- [287] Taube J.S., Klessak J.P. and Cotman C.W.: “Lesions of the rat postsubiculum impair performance on spatial tasks.” *Behavioral and Neural Biology*, **5**: pp. 131–143, **1992**. 37, 67, 79
- [288] Taube J.S., Muller R.I. and Ranck Jr. J.B.: “Head direction cells recorded from the postsubiculum in freely moving rats. I. Description and quantitative analysis.” *Journal of Neuroscience*, **10**: pp. 420–435, **1990**. 17, 67, 79
- [289] Taube J.S., Muller R.I. and Ranck Jr. J.B.: “Head direction cells recorded from the postsubiculum in freely moving rats. II. Effects of environmental manipulations.” *Journal of Neuroscience*, **10**: pp. 436–447, **1990**. 67, 79
- [290] Tchernichovski O., Benjamini Y. and Golani I.: “The dynamics of long-term exploration in the rat. Part I. A phase-plane analysis of the relationship between location and velocity.” *Biol Cybern*, **78**(6): pp. 423–32, **1998**. 101
- [291] Thompson L.T. and Best P.J.: “Place cells and silent cells in the hippocampus of freely-behaving rats.” *Journal of Neuroscience*, **9**(7): pp. 2382–2390, **1989**. 14, 126
- [292] Thompson L.T. and Best P.J.: “Long-term stability of the place-field activity of single units recorded from the dorsal hippocampus of freely behaving rats.” *Brain Res*, **509**(2): pp. 299–308, **1990**. 14
- [293] Tolman E.C.: “Cognitive maps in rats and men.” *Psychological Review*, **55**: pp. 189–208, **1948**. 7, 33, 34, 35, 36
- [294] Tolman E.C., Ritchie B.F. and Kalish D.: “Studies in Spatial Learning: II. Place Learning versus Response Learning.” *Journal of Experimental Psychology*, **36**(1): pp. 221–229, **1946**. 39

- [295] Totterdell S. and Smith A.D.: “Convergence of hippocampal and dopaminergic input onto identified neurons in the nucleus accumbens of the rat.” *J Chem Neuroanat*, **2**(5): pp. 285–98, **1989**. 99
- [296] Touretzky D.S. and Redish A.D.: “A theory of rodent navigation based on interacting representations of space.” *Hippocampus*, **6**(3): pp. 247–270, **1996**. 23, 40, 55, 123
- [297] Trullier O., Wiener S.I., Berthoz A. and Meyer J.A.: “Biologically-based artificial navigation systems: Review and prospects.” *Progress in Neurobiology*, **51**: pp. 483–544, **1997**. 33
- [298] Tsitsiklis J.N. and Roy B.V.: “Feature-Based Methods for Large Scale Dynamic Programming.” *Machine Learning*, **22**(1-3): pp. 59–94, **1996**. 53
- [299] Tsodyks M.V. and Markram H.: “The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability.” *Proc Natl Acad Sci U S A*, **94**(2): pp. 719–23, **1997**. 32, 105
- [300] Tsodyks M.V., Skaggs W.E., Sejnowski T.J. and McNaughton B.L.: “Population dynamics and theta rhythm phase precession of hippocampal place cell firing: a spiking neuron model.” *Hippocampus*, **6**(3): pp. 271–80, **1996**. 14
- [301] Usher M., Cohen J.D., Servan-Schreiber D., Rajkowski J. and Aston-Jones G.: “The Role of Locus Coeruleus in the Regulation of Cognitive Performance.” *Science*, **283**(5401): pp. 549–554, **1999**. 54
- [302] van Dam J.W., Kröse B.J. and Groen F.C.: “Adaptive Sensor Models.” In “Proceedings of 1996 International Conference on Multisensor Fusion and Integration for Intelligent Systems,” pp. 705–712. Washington D.C., USA, **1996**. 27
- [303] Varela J.A., Sen K., Gibson J., Fost J., Abbott L.F. and Nelson S.B.: “A quantitative description of short-term plasticity at excitatory synapses in layer 2/3 of rat primary visual cortex.” *J Neurosci*, **17**(20): pp. 7926–40, **1997**. 32, 105
- [304] Waelti P., Dickinson A. and Schultz W.: “Dopamine responses comply with basic assumptions of formal learning theory.” *Nature*, **412**(6842): pp. 43–8, **2001**. 53, 99
- [305] Wallace M.T., Meredith M.A. and Stein B.E.: “Converging influences from visual, auditory, and somatosensory cortices onto output neurons of the superior colliculus.” *J Neurophysiol*, **69**(6): pp. 1797–809, **1993**. 18

- [306] Wallace M.T., Meredith M.A. and Stein B.E.: “Multisensory integration in the superior colliculus of the alert cat.” *J Neurophysiol*, **80**(2): pp. 1006–10, **1998**. 19, 94
- [307] Wallace M.T. and Stein B.E.: “Cross-modal synthesis in the midbrain depends on input from cortex.” *J Neurophysiol*, **71**(1): pp. 429–32, **1994**. 19
- [308] Wallace M.T. and Stein B.E.: “Sensory organization of the superior colliculus in cat and monkey.” *Prog Brain Res*, **112**: pp. 301–11, **1996**. 19
- [309] Wallace M.T. and Stein B.E.: “Development of multisensory neurons and multisensory integration in cat superior colliculus.” *J Neurosci*, **17**(7): pp. 2429–44, **1997**. 19
- [310] Wan H.S., Touretzky D.S. and Redish A.D.: “Towards a computational theory of rat navigation.” In Mozer M., Smolensky P., Touretzky D.S., J.Elman and Weigend A. (editors) “Proceedings of the 1993 Connectionist Models Summer School,” pp. 11–19. Hillsdale, NJ, Erlbaum, **1994**. 23, 123
- [311] Watkins C.J.C.H.: *Learning from delayed rewards*. Ph.D. thesis, University of Cambridge, England, **1989**. 51, 52, 113
- [312] Watkins C.J.C.H. and Dayan P.: “Q-Learning.” *Machine Learning*, **8**(3/4): pp. 279–292, **1992**. 51, 113
- [313] Wehner R.: “Himmelsnavigation bei Insekten. Neurophysiologie und Verhalten.” *Neujahrsblatt der Naturforschenden Gesellschaft Zürich*, **184**: pp. 1–132, **1982**. 33, 34
- [314] Wehner R.: “Desert ant navigation: how miniature brains solve complex tasks.” *Journal of Comparative Physiology a-neuroethology Sensory Neural and Behavioral Physiology*, **189**(8): pp. 579–588, **2003**. 33, 34
- [315] Wehner R. and Wehner S.: “Path integration in desert ants: Approching a long-standing puzzle in insect navigation.” *Monitore Zool. Ital.*, **20**(3): pp. 309–331, **1986**. 33, 34
- [316] Wei D.S., Mei Y.A., Bagal A., Kao J.P., Thompson S.M. and Tang C.M.: “Compartmentalized and binary behavior of terminal dendrites in hippocampal pyramidal neurons.” *Science*, **293**(5538): pp. 2272–5, **2001**. 32, 105
- [317] Weldon D.A. and Best P.J.: “Changes in sensory responsivity in deep layer neurons of the superior colliculus of behaving rats.” *Behav Brain Res*, **47**(1): pp. 97–101, **1992**. 19

- [318] Whishaw I.Q., Cassel J.C. and Jarrard L.E.: “Rats with fimbria-fornix lesions display a place response in a swimming pool: A dissociation between getting there and knowing where.” *Journal of Neuroscience*, **15(8)**: pp. 5779–5788, **1995**. 99
- [319] Whishaw I.Q. and Maaswinkel H.: “Absence of dead reckoning in hippocampal rats.” *Society for Neuroscience Abstracts*, **23**: p. 1839, **1997**. 34
- [320] Whishaw I.Q. and Mittleman G.: “Hippocampal modulation of nucleus accumbens: Behavioral evidence from amphetamine-induced activity profiles.” *Behavioral and Neural Biology*, **55**: pp. 289–306, **1991**. 30, 36, 99
- [321] White N.M. and McDonald R.J.: “Multiple parallel memory systems in the brain of the rat.” *Neurobiology of Learning and Memory*, **77(2)**: pp. 125–184, **2002**. 30, 33, 35
- [322] Wickens J. and Kötter R.: “Cellular Models of Reinforcement.” In Houk J.C., Davis J.L. and Beiser D.G. (editors) “Models of Information Processing in the Basal Ganglia,” chapter 10, pp. 187–214. MIT Press, Cambridge, Massachusetts, USA, **1995**. 32, 105
- [323] Wickens J.R., Begg A.J. and Arbuthnott G.W.: “Dopamine reverses the depression of rat corticostriatal synapses which normally follows high-frequency stimulation of cortex in vitro.” *Neuroscience*, **70(1)**: pp. 1–5, **1996**. 54, 105
- [324] Wilson M.A. and McNaughton B.L.: “Dynamics of the hippocampal ensemble code for space.” *Science*, **261**: pp. 1055–1058, **1993**. 12, 14, 36, 63, 80
- [325] Wilson M.A. and McNaughton B.L.: “Reactivation of hippocampal ensemble memories during sleep.” *Science*, **265**: pp. 676–679, **1994**. 14
- [326] Winson J.: “Loss of hippocampal theta rhythm results in spatial memory deficits in the rat.” *Science*, **201**: pp. 160–163, **1978**. 12
- [327] Wiskott L.: “How Does Our Visual System Achieve Shift and Size Invariance?” In van Hemmen J.L. and Sejnowski T.J. (editors) “Problems in Systems Neuroscience,” Oxford University Press, **2004**. (to appear). 31
- [328] Wiskott L. and Sejnowski T.: “Slow Feature Analysis: Unsupervised Learning of Invariances.” *Neural Computation*, **14(4)**: pp. 715–770, **2002**. 31
- [329] Witter M.P.: “Organization of the entorhinal-hippocampal system: A review of current anatomical data.” *Hippocampus*, **3**: pp. 33–44, **1993**. 8, 10, 11, 84

- [330] Witter M.P. and Groenewegen H.J.: “Laminar origin and septotemporal distribution of entorhinal and perirhinal projections to the hippocampus in the cat.” *Journal of Comparative Neurology*, **224**: pp. 371–385, **1984**. 67
- [331] Witter M.P., Groenewegen H.J., da Silva F.H.L. and Lohman A.H.M.: “Functional organization of the extrinsic and intrinsic circuitry of the parahippocampal region.” *Progress in Neurobiology*, **33**: pp. 161–253, **1989**. 10
- [332] Witter M.P., Naber P.A., van Haeften T., Machielsen W.C., Rombouts S.A., Barkhof F., Scheltens P. and Lopes da Silva F.H.: “Cortico-hippocampal communication by way of parallel parahippocampal-subicular pathways.” *Hippocampus*, **10**(4): pp. 398–410, **2000**. 10
- [333] Wurtz R.H. and Albano J.E.: “Visual-motor function of the primate superior colliculus.” *Annu Rev Neurosci*, **3**: pp. 189–226, **1980**. 18
- [334] Young B.J., Fox G.D. and Eichenbaum H.: “Correlates of hippocampal complex-spike cell activity in rats performing a nonspatial radial maze task.” *Journal of Neuroscience*, **14**(11): pp. 6553–6563, **1994**. 12, 67
- [335] Zhang K.: “Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: A theory.” *Journal of Neuroscience*, **16**(6): pp. 2112–2126, **1996**. 33, 77
- [336] Zhang K., Ginzburg I., McNaughton B.L. and Sejnowski T.J.: “Interpreting neuronal population activity by reconstruction: A unified framework with application to hippocampal place cells.” *Journal of Neurophysiology*, **79**: pp. 1017–1044, **1998**. 8, 63, 80
- [337] Zucker R.S. and Regehr W.G.: “Short-term synaptic plasticity.” *Annu Rev Physiol*, **64**: pp. 355–405, **2002**. 32, 105

Curriculum Vitæ

Thomas Strösslin

33, av. de Morges, 1004 Lausanne, Switzerland
email: thomas.stroesslin@epfl.ch
Born March 5, 1973 in Basel, Switzerland

Education

- **Current:** Ph.D. student with Wulfram Gerstner at the Laboratory of Computational Neuroscience (LCN), EPFL, Switzerland.
- **October 1999:** Graduation in Electrical Engineering at ETH Zürich.
- **Feb. - May 1999:** Diploma project at Ecole Nationale Supérieure des Télécommunications Paris (ENST Paris).
- **Sept. 1997 - March 1998:** Studies at Bristol University (UK) with ERASMUS exchange program.

Working Experience

July 1998

Company: Supercomputing Systems Ltd. Zürich, Switzerland
Responsibilities: Implementation and test of a multi-DSP and FPGA board for stereo image processing.

June 1997 - August 1997

Company: Philips Semiconductors AG. Zürich, Switzerland.
Responsibilities: (i) Design of peripheral blocks for an integrated GSM set using VHDL, (ii) Implementation of a MC68k to ARM cross-assembler in Perl.

Languages

German: Mother tongue.
English: Fluent (total of 10 months stay in English speaking countries).
French: Fluent (living in French part of Switzerland since April 2000).
Spanish: Basic knowledge.

Publications

Strösslin T., Gerstner, W.: Reinforcement Learning in Continuous State and Action Space. In: Artificial Neural Networks - ICANN 2003,

Strösslin, T., Krebsler, C., Arleo, A., Gerstner, W.: Combining multimodal sensory input for spatial learning. In Dorransoro, J.R., ed.: Artificial Neural Networks - ICANN 2002, Springer LNCS (2002) 87–92