# Beyond the Cognitive Map:

**Contributions to a Computational Neuroscience Theory of Rodent Navigation**

**A. David Redish**

August 1997
CMU-CS-97-166

School of Computer Science
Carnegie Mellon University
Pittsburgh PA 15213-3891

*Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy.*
*(Department of Computer Science and Center for the Neural Basis of Cognition.)*

Thesis Committee:
David S. Touretzky, Chair
James L. McClelland
David C. Plaut
Bruce L. McNaughton, University of Arizona



The work presented in this thesis was partially supported by a National Science Foundation Graduate Fellowship (1991-1994), the Neural Processes in Cognition graduate training program (NSF grant BIR-9014347), the Center for the Neural Basis of Cognition, and an NSF grant awarded to David S. Touretzky (IBN-9631336).

Keywords: Theoretical Neuroscience, Computational Neuroscience, Spatial Cognition, Place cell, Head direction cell, Path integration

# Acknowledgements

This thesis is dedicated to the memory of Hank S. Wan.

I want to take this moment to thank a number of people. First of all, my wife Laura, for her faith in me and her patience in dealing with me in these trying times; my advisor, David Touretzky, who kept me on track without letting me suspect it until it was too late; my parents for making this easier than it should have been; and my two cats for their short memories. I'd also like to thank my committee, Jay McClelland, David Plaut, and Bruce McNaughton for their high standards.

Second, I'd like to thank the Neural Processes in Cognition (NPC) program, the Center for the Neural Basis of Cognition (CNBC), and the Computer Science Department (CSD-CMU). In particular, I want to thank the students and faculty of both CNBC/NPC and CSD-CMU for all their time, comments, and encouragement, especially Marlene Behrmann, Justin Boyan, Josh Brumberg, Rick Gilmore, Geoff Gordon, Bard Ermentrout, Daniel Jackson, Angela Kennedy, John Lafferty, Greg Nelson, Joseph O'Sullivan, Mark Pimm-Smith, David Pinto, Rick Romero, Jonathon Shewchuk, Bill Skaggs, Shaun Vecera, and the reinforcement learning group. I'm especially grateful to my officemates for putting up with me all these years. And, of course, special thanks to Sharon Burks and Cindy Wood for making the departments run so smoothly.

Third, I'd like to acknowledge the other students and faculty who've contributed directly to this work including Sofyan Alyan, David Banks, Adam Elga, Mark Fuhs, Steven Gaulin, Jeremy Goodridge, Michael Pahn, Chris Reiber-Milberg, Lisa Saksida, and Hank Wan. I'd also like to thank all the undergraduates who helped with the behavioral experiments described in Sections 2.5 and 2.6: Cynthia Cole, Trish Ellis, Charles Gordon, Margaret Horner, Elizabeth Kanabe, Rolf Neill, Jay Norjeko, Leeann Walton, and Stacy Weber.

Finally, I'd like to acknowledge all the neuroscientists who've been so generous with their data, their comments, and their encouragement: Larry Abbott, Tad Blair, Neil Burgess, T.S. Collett, Howard Eichenbaum, Kati Gothard, Apostolos Georgopoulos, Michael Hasselmo, John Kalaska, Jim Knierim, Chip Levy, Etan Markus, Bartlet Mel, Horst Mittelstaedt, Bob Muller, Alex Pouget, Alexei Samsonovich, Terry Sanger, Pat Sharp, Larry Snyder, R.J. Sutherland, Jeff Taube, Misha Tsodyks, Michael Weisend, Sid Wiener, Matt Wilson, Steve Wise, Brad Wyble, and Kechen Zhang.

Last but not least: a special thanks to all the people at interlibrary loan for all their tireless work.

The work presented in this thesis was partially supported by a National Science Foundation Graduate Fellowship (1991-1994), the Neural Processes in Cognition graduate training program (NSF grant BIR-9014347), the Center for the Neural Basis of Cognition, and an NSF grant awarded to David S. Touretzky (IBN-9631336).

# Abstract

Rodent navigation is a unique domain for studying information processing in the brain because there is a vast literature of experimental results at many levels of description, including anatomical, behavioral, neurophysiological, and neuropharmacological. This literature provides many constraints on candidate theories. This thesis presents contributions to a theory of how rodents navigate as well as an overview of that theory and how it relates to the experimental literature.

In the first half of the thesis, I present a review and overview of the rodent navigation literature, both experimental and theoretical. The key claim of the theory is that navigation can be divided into two categories: *taxon/praxic navigation* and *locale navigation* (O'Keefe and Nadel, 1978), and that locale navigation can be understood as an interaction between five subsystems: *local view*, *head direction*, *path integration*, *place code*, and *goal memory* (Redish and Touretzky, 1997a). I bring ideas together from the extensive work done on rodent navigation over the last century to show how the interaction of these systems forms a comprehensive, computational theory of navigation. This comprehensive theory has implications for an understanding of the role of the hippocampus, suggesting that it shows three different modes: *storage*, *recall*, and *replay*.

In the second half of the thesis, I show specific contributions to this overall theory.

- I report the first simulation of the head direction system that can track multiple head direction speeds accurately. The simulations show that the theory implies that head direction tuning curves in the anterior thalamic nuclei should deform during rotations. This observation has been confirmed experimentally by Blair *et al.* (1997).

- By examining the computational requirements and the anatomical data, I suggest that the anatomical locus of the path integrator is in a loop comprised of the subiculum, the parasubiculum, and the superficial entorhinal cortex. This contrasts with other hypotheses of the anatomical locus of path integration (e.g. hippocampus, McNaughton *et al.*, 1996) and predicts that the hippocampus should not be involved in path integration. This prediction has been recently tested and confirmed by Alyan *et al.* (1997).

- I present simulations demonstrating the viability of the three-mode hippocampal proposal, including storage and recall of locations within single environments, with ambiguous inputs, and in multiple environments.

- I present simulations demonstrating the viability of the dual-role hippocampus (*recall* and *replay*), showing that the two modes can coexist within the hippocampus even though the two roles seem to require incompatible connection matrices.

In addition, I present simulations of specific experiments, including

- a simulation of the recent result from Barnes *et al.* (1997), showing that the model produces a bimodality in the correlations of representations of an environment in

animals with deficient LTP. These simulations show that the Barnes *et al.* (1997) result does not necessarily imply that the intra-hippocampal connections are pre-wired to form separate *charts* as suggested by Samsonovich (1997).

- a simulation of Sharp *et al.*'s (1990) data on the interaction between entry point and external cues, showing the first simulations capable of replicating all the single place field conditions reported by Sharp *et al.*
- simulations of Cheng (1986) and Margules and Gallistel (1988) showing the importance of disorientation in self-localization.
- simulations of Morris (1981), showing that the model can replicate navigation in the water maze.
- simulations of Collett *et al.* (1986) and our own gerbil navigation results, showing that the model can replicate a number of reactions to different manipulations of landmark arrays.

# Contents

# Chapter 1

# Introduction

## 1.1 Rodent navigation

Rodent navigation is a unique domain for studying information processing in the brain. There is a vast literature of experimental results at various levels of description, including anatomical, behavioral, neurophysiological, and neuropharmacological, which provides many constraints on candidate theories. In addition, there has already been a tremendous amount of work on the theoretical basis of how rodents navigate through space ranging from early psychological theories such as the *stimulus-response* theory of Hull (1943; 1952) and the *cognitive map* theory of Tolman (1948) to more recent neuroscience theories such as the *hippocampus as a cognitive map* (O'Keefe and Nadel, 1978), and others.

Third, rodent navigation is a particularly useful means of studying information processing in neural systems because it fits within the broader domains of navigation and spatial reasoning, both of which have extensive histories. Navigation has been a critically studied discipline since the earliest days of history. Early navigators developed concepts that allowed them to cross vast distances of open water. Similarly, humans have been studying the geometry of space for thousands of years, going back to the ancient Greeks (e.g. Euclid, Pythagoras, Archimedes, etc.), through Newton, Descartes, Leibniz, and others. This thesis is not a history of the study of the philosophy of space and I will not go into depth about these previous ideas. Instead, I refer the reader to excellent reviews such as those in Collinder (1955), Cotter (1968), Lanczos (1970), O'Keefe and Nadel (1978), and Gallistel (1990) to name but a few.

This thesis pulls together a comprehensive theory of rodent navigation from the extensive experimental and theoretical work done by researchers in the field over the past century and examines specific components of that comprehensive theory, including computer simulations to test complexities in those specific components. By looking at a large corpus of experiments across a variety of paradigms (see Experimental Review, Chapter 2), it is possible to bring the theoretical literature into a consistent picture (rejecting some theories, accepting some, and modifying others, see Navigation Overview (Chapter 3) and Theories of Hippocampal Function(Chapter 4)). In addition to synthesizing a comprehensive theory, I examine a number of specific components of the theory (and include simulations): the head direction system (Chapter 5), the path integration system (Chapter 6), and the role of the hippocampus (Chapters 7–9). Finally, to demonstrate that the theory truly is comprehensive, I show simulations of specific experiments (Part III), including the bimodality in the representations of an environment by old animals but not young (Barnes *et al.*, 1997, Chapter 10), the interaction between entry point and cue cards (Sharp *et al.*, 1990, Chapter 11), the differential effects of disorientation or non-disorientation (Cheng, 1986, and Margules and Gallistel, 1988, Chapter 12), self-localization in the Morris water maze (Morris, 1981, Chapter 13), and effects of manipulating the landmark array on search by gerbils (Collett *et al.*, 1986, see also experiments in Section 2.5, simulations in Chapter 14).

The specific domain of rodent navigation (as separate from either general navigation issues or the domain of spatial reasoning) has a long history itself. Early psychologists ran rats in various kinds of mazes under sensory and other manipulations (e.g. Watson, 1907; Carr and Watson, 1908; Dennis, 1932; Honzik, 1936; Tolman, 1948, see Chapter 2 for other examples). And modern researchers in a variety of fields have drastically expanded the range of experimental techniques available. Modern experiments studying rodent navigation range from behavioral experiments quantifying complex effects both with and without lesions (see Section 2.1), to correlating behavior and task parameters to EEG, single-cell, and multi-cell (simultaneous) recordings (Section 2.2). Modern experiments also include neuropharmacological manipulations (Section 2.3) and even gene-manipulation experiments (e.g. Section 2.1.2-D). And, of course, the anatomy of the rodent brain is quite well mapped out (Krieg, 1946; Zilles, 1985; Zilles, 1990).

As mentioned already, the theoretical side traces its history almost as far back; in Chapter 3, I present an overview of a comprehensive theory that draws on the theoretical and experimental work done by other researchers over the last century.

## 1.2 Neuroscience and computation

The brain can be understood on many levels. At an abstract level, one can ask what information has to be represented and how that information might be processed. For example, animals have the ability to take a circuitous path, and then, using no external information, return directly to the starting point of that path (Mittelstaedt and Mittelstaedt, 1980; Etienne, 1987; Alyan and Jander, 1994, see also Chapter 2, EX.57, page 38). This is called *path integration* or *dead reckoning* and is discussed in depth in Section 3.2.3 and Chapter 6. Because the path that animals take returning to the starting point includes both direction and distance, they must be representing a 2-dimensional spatial vector. That is, either $(x, y)$ (in Cartesian coordinates) or $(r, \phi)$ (in polar coordinates). However, this does not imply that they represent the information as two numbers $x$ and $y$ (or $r$ and $\phi$), only that the information represented is *mathematically equivalent* to the 2-D vector.

Algorithms designed to handle Cartesian or polar coordinates may not be well-suited for the representation actually used in the rodent brain. In order to understand how path integration actually occurs in the rodent brain, one needs to push further down and ask how the information is represented by populations of neurons. Location in space is represented by *place cells* in the hippocampal formation (O'Keefe and Dostrovsky, 1971, see Section 2.2.2 for a review). Each of these cells has a high probability of firing when the animal is in a small, circumscribed portion of the environment. The population of these cells represents the location of the animal in the environment. Algorithms that are appropriate for these kinds of representations are very different from those intended for Cartesian or polar coordinates.

In order to understand the computation required to accomplish a task, I will define the *functional structure* of the task which will divide the task into *subsystems*. However, there is not necessarily a one-to-one correspondence between subsystem and anatomical structure. Each subsystem is realized by an interaction between different brain structures and each brain structure plays roles in multiple subsystems.

This division into abstract functionality and detailed implementation is analogous to Marr's three levels of "understanding complex information-processing systems" (Marr, 1982, Chapter 1): *computational theory*, *representation and algorithm*, and *hardware implementation*. *Computational theory* asks what computation the system must be performing. In our path integration example, this corresponds to the recognition that animals must maintain a representation of the vector home. *Representation and algorithm* asks how the information is represented by the system and how that information gets processed. In the example of path integration, this corresponds to the recognition that place cells represent a location in a coordinate system, which is mathematically equivalent to a 2-dimensional vector (McNaughton *et al.*, 1994a; Touretzky and Redish, 1996; Samsonovich and McNaughton, 1997). (I leave in-depth discussions about how that representation is maintained and updated for Chapter 6.) The *hardware implementation* level asks how the representation and algorithm is realized physically. This concerns questions of neuronal connections, neurotransmitters, etc.

However, Marr assumed that the levels were separable, so that questions about computation could be addressed without addressing representation or algorithm, and questions about representation and algorithm could be addressed without addressing the actual implementation. I will not make this assumption.

An example of the inseparability of levels is the self-localization process described in Chapter 8. This process allows an animal to resolve ambiguous local view inputs and reset the path integrator representation to a value consistent with the local view. An abstract computation of this process would select between possible candidate locations (e.g. Collett *et al.*, 1986; Touretzky and Redish, 1996). However, the actual process occurs by a pseudo-winner-take-all network settling to a stable state. This means that the process averages nearby candidate locations, but selects between candidate locations that are more distant from each other. The pWTA process can only be understood by examining the neurophysiology (the implementational level), but this changes the computation. This is an example where the computation as understood abstractly and neurophysiologically have different behavioral effects. Thus they are not independent.

The theory, therefore, cannot address only one level, because the levels are inter-related. It is not enough to say that in order to accomplish path integration there must be a represention of a vector home; one must also ask how that vector is represented and address any data available about the representation and implementation. The key to creating a theory that addresses multiple levels is to address multiple data levels as well.

Data about the brain comes in many forms: Behavioral data gives clues to what is represented and what is not, neurophysiological data will be the key to understanding representations, and anatomical data will be crucial to understanding how the data flow occurs and how representations are processed. In addition, there is neuropharmacological data, which is very useful to help get at Marr's hardware implementation level. But there are cases in which the neuropharmacological data informs even the computational theory level (e.g. the role of acetylcholine reviewed in Section 3.2.4-F).

Finally, there is lesion data, in which one or more brain structures are destroyed. Lesion data tells little about what a brain structure does, only what the brain can do without that structure. The fact that an animal can perform a task without part of its brain does not mean that that part of the brain is not involved in the task during normal operation. An animal may even be using a different technique to solve the specific problem. Similarly, the fact that an animal cannot perform a task after a lesion does not imply that the brain structure was critically involved in the computation. For example, the brain structure may provide necessary tonic input into another brain structure that does not provide any informational content but is necessary for the second structure to perform its computations. When interpreting lesion data, it is important to be clear about what roles the structure may play in the task. I will primarily use lesion data for two purposes: as a means of identifying that certain brain structures are critically involved in a task and to confirm theories built on the foundations from other data paradigms (that is, once the theory has been built based on data from other paradigms, it will predict lesion effects that the lesion data must corroborate, but the lesion data will not generally be used to drive the specifics of the theory).

Data from all of these levels are useful constraints on a theory. In a sense, theoretical neuroscience can be thought of as a constraint satisfaction problem. However, one has to be careful about what data to include. More constraints limit the range of candidate theories and make us more confident of our predictions. But too many unrelated constraints confuse the issue.

The key to answering this problem is the concept of a *domain*. A domain is a set of experiments that a theory can be reasonably expected to address. These limits should be unrelated to the actual theory itself. A single domain is explicable by a variety of theories. Requiring all theories to address the same set of experiments allows the rejection of theories that do not fit the experimental data.

A domain should be broad enough to encompass data from multiple levels and relate them, but not so broad as to be unwieldy. The main purpose of defining a domain is to allow other people beyond the originator of the theory to decide whether an experiment is addressable by a theory or not.

This thesis presents an overview of a comprehensive theory addressing the domain of rodent navigation, and the specific contributions I have made to the theory. The theory draws heavily from the theoretical and experimental work done by other researchers over the last century, but some specific aspects of the overall theory are novel. I also make novel predictions and provide new insights into some specific experiments.

## 1.3 Structure of the thesis

The thesis is structured in three major parts: (I) Review and overview, (II) Specific contributions on selected subsystems, and (III) Simulations of specific experiments.

In the *Experimental Review*, Chapter 2, I list all of the major experimental results used in this thesis and include a short description of each. Each result is identified by a label (e.g. EX.13). This separation of the critical experimental results allows me to describe the experimental details once without repeating the complications every time the result is referred to. Readers interested in the big picture may prefer to begin with the Navigation Overview (Chapter 3) and refer to the Experimental Review (Chapter 2) only as necessary. Readers may also wish to refer to the glossary of terms included at the end of the thesis.

In Chapter 3, I present an overview of the current understanding of the field of rodent navigation, laying out the "standard model." Although there is still some debate about specific aspects of it, I will try to bring the rodent navigation literature into a coherent story (a comprehensive theory). This chapter draws heavily on the theoretical and experimental work done by other researchers over the last century (in particular over the last twenty years since the publication of O'Keefe and Nadel's *The Hippocampus as a cognitive map*). It brings together a number of disparate ideas in a novel way.

The key claim in Chapter 3 follows O'Keefe and Nadel's (1978) taxonomy dividing navigation into two major categories: *taxon/praxic navigation* and *locale navigation*. The second key claim in Chapter 3 is that locale navigation can be understood as an interaction between five subsystems: *local view*, *head direction*, *path integration*, *place code*,[1] and *goal memory* (Redish and Touretzky, 1997a).

Following explications of each of these sections, detailing computational, anatomical, and representational issues, I discuss two additional navigational issues that arise from the subsystem discussions: *reference frames* and *route navigation*. The concept of reference frames is necessary for both the explanation of a number of complex place cell properties (e.g. EX.78) and for navigation in multiple environments and under multiple goal conditions. Route navigation could be considered a third taxonomic element, but it has similarites to

[1] There has been a recent debate about whether the place code is separate from the path integration system (McNaughton *et al.*, 1996; Redish and Touretzky, 1997a; Samsonovich and McNaughton, 1997), but recent data has provided strong support for the separation of the two systems (Alyan *et al.*, 1997, see EX.59). See Chapter 6 for an in-depth discussion of this issue.

taxon/praxic navigation. It is related to memory consolidation issues often discussed in the context of hippocampal lesions (e.g. Cohen and Eichenbaum, 1993; Skaggs *et al.*, 1996).

After presenting the Navigation Overview, I present a review of the theories of the role of the hippocampus. The hippocampus is one of the most studied structures in the brain, and quite a number of theories have been proposed to explain its role. Although this thesis is not a "hippocampal model," it addresses a number of issues about the hippocampus, so, in Chapter 4, I review the major theories of hippocampal function and address how the understanding brought out in Chapter 3 relates to this question of the role of the hippocampus.

Having presented an overview of rodent navigation and a review of hippocampal theories, in Part II, I present specific contributions I have made to aspects of the general theory. Each chapter addresses a single complex element of the overall theory.

Although it would be best to simulate the entire theory described in Chapter 3 at a neural level, this is not feasible at this time. There are a number of reasons for this.

First, the components need to be understood before attempting such a complete understanding of the whole picture. There are details missing from the big picture that will require additional experiments before they can be worked out. Although the consequences of known representations can be reasoned about without completely understanding the mechanisms underlying them, this understanding does not yet enable a full simulation.

Second, there are computational limitations. Some of the simulations require thousands of neurons and up to million synapses. Even using high powered computers available today, these simulations are very slow.

Third, there are many parameters involved in each of the simulations. Many of these parameters have highly non-linear effects (see, for example, the effect of external input discussed in Appendix A). A simulation that included all of the components discussed in Part II would be so unwieldy as to become meaningless.

Instead of a single unified simulation, I present a family of simulations all using the same neural model with explicit assumptions identified for each component. Each simulation addresses a specific aspect of the overall theory. All simulations use the same neural model and all parameters and assumptions are justified in each simulation.

Each simulation is a complete, essentially modular, component of the complete navigation theory. It has its own assumptions, which means that I can make predictions based on those assumptions independent of the assumptions in the other chapters. For example, the path integration anatomical hypothesis (presented in Chapter 6) is only dependent on the existence of head direction cells in the postsubiculum (which have been well-studied, Taube *et al.*, 1990a; Sharp, 1996a, see EX.94), not on the explicit mechanism for maintenance of the head direction representation presented in Chapter 5.

The neuron model used in the simulations consisted of three equations. For an extensive discussion of the derivation, see Wilson and Cowan (1972) and Pinto *et al.* (1996).

$$V_i(t) = \gamma_i + \sum_j w_{ij} S_j(t) \tag{1.1}$$

$$F_i(t) = \frac{1 + \tanh(V_i(t))}{2} \tag{1.2}$$

$$\tau_i \frac{dS_i(t)}{dt} = -S_i(t) + F_i(t) \tag{1.3}$$

$V_i(t)$ is proportional to the voltage in neuron $i$ at time $t$, $\gamma_i$ is a tonic inhibitory input to neuron $i$ necessary for the dynamics of the system, $w_{ij}$ is the synaptic weight from neuron $j$ to neuron $i$, $F_i(t)$ is the firing rate at time $t$, and $S_i(t)$ is the *synaptic drive* of neuron $i$. $\tau_i$ is a time constant, proportional to the decay of the post-synaptic potential produced by neuron $i$ in neurons on which it synapses. Synaptic drive is a non-measurable property of a neuron and can be understood as the effect of the neuron on all the neurons on which it synapses divided by the synaptic weight across each synapse.

As has been shown by Pinto *et al.* (1996), equations 1.1–1.3 form a consistent neuronal model that can be understood as describing either a continuous approximation to a single neuron with $F_i(t)$ being the probability of firing at time $t$, or a population of neurons with $V_i(t)$ being proportional to the average voltage and $F_i(t)$ the fraction of neurons in the population firing a spike at time $t$.

In Part II, I present:

- (Chapter 5) a simulation of the head direction system that can track multiple head direction speeds accurately. From these simulations, I made an observation about the tuning curves of cells in ATN. This observation has been confirmed experimentally by Blair *et al.* (1997).

- (Chapter 6) a novel proposal for the anatomical locus of the path integrator. By examining the computational requirements and the anatomical data, I find fault with the previous hypotheses and provide a novel one. I also describe experiments that can differentiate between my proposal and previous hypotheses. One of these experiments has been recently tested by Alyan *et al.* (1997) and confirmed the prediction made by my proposal.

- (Chapter 7) a novel understanding of the role of the hippocampus in navigation: specifically that it has three modes: *storage*, *recall*, and *replay*.

- (Chapter 8) simulations demonstrating the viability of the hippocampal model from Chapter 7, including storage and recall of locations within single environments, with ambiguous inputs, and in multiple environments. I also present a novel reason for why rodents regularly return to their starting points as they explore a novel environment.

- (Chapter 9) simulations demonstrating the viability of the dual-role hippocampus, showing that the recall and replay modes can coexist within the hippocampus. (The two roles would seem to require incompatible connection matrices.)

Finally, having presented an overview of rodent navigation and contributions to general aspects of that overall theory, I present instantiations of the theory. These instantiations attempt to simulate specific experimental results.

The goal of the simulations in Part III is two-fold. First, they show that the theory applies to specific experiments and not just general ideas, and second, they allow me to address complex aspects of the theory that may not be directly obvious analytically.

As with the subsystem contributions presented in Part II, each simulation is a complete component with its own assumptions and predictions that do not depend on any assumptions not either stated explicitly in the overview or in the chapter describing the simulation itself. And as with Part II, I will be careful to explain how each simulation fits into the theory presented in Chapter 3.

In Part III, I present:

- (Chapter 10) a simulation of the recent result from Barnes *et al.* (1997), showing that the model produces a bimodality in the representation of location in animals with deficient LTP. Compare EX.83.
- (Chapter 11) a simulation of Sharp *et al.* (1990), showing the first simulations capable of replicating all the single place field conditions reported by Sharp *et al.* Compare EX.73, Figure 2.22.
- (Chapter 12), simulations of Cheng (1986) and Margules and Gallistel (1988) showing the importance of disorientation in self-localization. Compare EX.35.
- (Chapter 13), simulations of Morris (1981), showing that the model can replicate navigation in the water maze.
- (Chapter 14), simulations of Collett *et al.* (1986) and our own gerbil navigation experiments (Section 2.5), showing that the model can replicate a number of reactions to different manipulated landmark arrays.

This thesis brings together ideas from the field of rodent navigation into a comprehensive theory that is compatible with essentially all of the data available. It goes into explicit details about several components of the theory, and reports simulations showing how the theory applies to a number of rodent navigation experiments.

An important aspect of this theory is that it treats an entire domain, and thus uses a huge corpus of data with which to constrain the hypotheses. By not assuming a one-to-one mapping between structure and function, a theory has been constructed that fits data ranging from the anatomical, through the neuropharmacological and neurophysiological levels, all the way to behavior. No anatomical structure works alone. It is only through interactions between structures that behavior is accomplished.

# Part I

# Review and overview

Before presenting my specific contributions to the field of rodent navigation, I present three reviews and overviews.

- (Chapter 2) I begin with a review of the experimental literature. One of the advantages of the rodent navigation field is the vast extent of the experimental literature, including behavioral, neurophysiological, neuropharmacological, and anatomical data.
- (Chapter 3) I next present an overview of a comprehensive model of rodent navigation. It draws from theoretical work done over the last century and brings together a number of disparate ideas.
- (Chapter 4) Although this thesis is not a hippocampal model *per se*, the work presented here has strong implications for the role played by the hippocampus in learning and memory. In Chapter 4, I review the previous hippocampal theories and attempt to bring together an understanding based on the role determined in the Navigation Overview (Chapter 3).

Readers interested in the big picture may wish to start with the Navigation Overview (Chapter 3) and refer to the Experimental Review (Chapter 2) only as necessary. References to the Experimental Review are given throughout the text as appropriate.

# Chapter 2

# Experimental review

## 2.1 Behavioral and lesion data

### 2.1.1 Open field navigation.

Nearly everyone living in a modern suburb has seen squirrels caching food for the winter. How do they do this? A number of investigators have explored cache behavior in rodents and found that they used constellations of landmarks (Vander Wall, 1990; Jacobs and Liman, 1991; Jacobs, 1992; Sherry and Duff, 1996). In order to explore navigation based on local landmarks, Collett *et al.* (1986) trained gerbils to find a food reward (a sunflower seed buried in gravel) among cylindrical landmarks in an otherwise impoverished environment (see Figure 2.1). The landmarks were translated but not rotated from trial to trial and maintained the same orientation and distance from each other, thus forming a stable landmark array.

Open field navigation has also been studied by Biegler and Morris (1993; 1996), and Gothard *et al.* (1996a). Biegler and Morris examined the effects of landmark stability relative to the wall of the environment, while Gothard *et al.* recorded from hippocampal place cells while the animals performed a simplified version of one of these tasks (see discussion of neurophysiological recordings below).

Each of the labs used slightly different techniques, which I will review first. For example, Collett *et al.* used female gerbils, Biegler and Morris used male hooded rats, and Gothard *et al.* used male albino rats.

In Collett *et al.* (1986), both the gerbils' starting location and the location of the landmark array varied randomly for every trial, but the food reward was always located at the same position relative to the array and the start box always had the same orientation. The arena was surrounded by black-painted walls and was illuminated by a single incandescent bulb above its center which left the walls in shadow. The floor was covered by black granite chips. Probe trials were run without a food reward, and the time spent at each location was

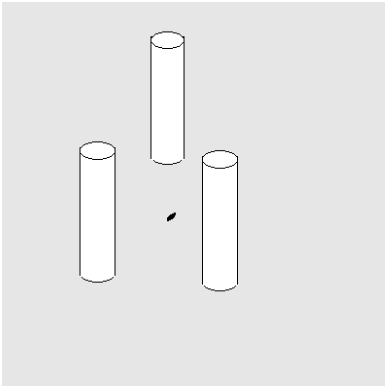


Figure 2.1: A typical open-field navigation task (Collett *et al.*, 1986). Small white cylinders arranged in an array form the local landmarks. Cylinders are translated from trial to trial within the arena, but the array configuration is not changed. Orientation of the array is also kept constant. Animals are trained to find food at a constant relation to the landmark configuration and are then tested with the landmark configuration in a novel location and no food.

histogrammed.

Gothard *et al.* (1996a) used a more limited task in which the start box was always along the southern wall and the landmarks were always placed to the north. They also did not cover the floor; it was bare black tile. They used chocolate sprinkles for food reward, which were (presumably) not visible on the black floor. This technique, however, has a distinct disadvantage in that the animals do not have to commit to a goal location. They can sweep over the sprinkles and vacuum up the reward in passing. This means that histograms of position by time do not produce strong peaks at the goal as those of Collett *et al.* do. Gothard *et al.* did include non-reinforced trials during which the animals would be expected to spend time searching near the goal, but if the animal were very accurate in its search, it might quickly recognize that there was no reward.

Biegler and Morris (1993; 1996) trained their rats to find food at one of a small number of feeders placed in their environment. Unlike the previous experiments, Biegler and Morris disoriented their rats before each trial. Their environment did include a polarizing wall cue, but the rats were disoriented prior to experiencing that cue. This may have disrupted the landmark - head direction association (see Section 3.2.2).

#### 2.1.1-A One landmark experiments.

**Reviewed Result EX.1** *Gerbils can learn to search at a specific bearing and distance to a single circularly symmetric landmark (Collett* et al.*, 1986).*

Collett *et al.* (1986) trained gerbils to find food placed near a single circularly-symmetric

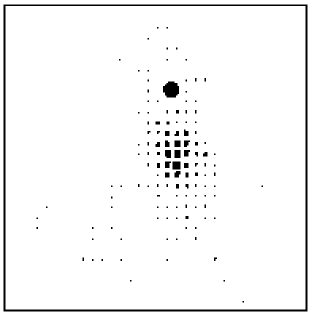


Figure 2.2: One landmark: Histogram of search time of gerbils on the one-landmark task. Solid circles indicate landmark locations and size of the blobs indicate time spent at each location, summed over all animals. From Collett *et al.* (1986), reprinted with permission of author and publisher.

landmark (Figure 2.2). The fact that the animals could learn to search at the correct bearing as well as distance implies that they have some independent means of determining bearing information. Collett *et al.* supposed that the animals were utilizing some external cue, despite the experimenters attempts to prevent this. The head direction subsystem (Section 3.2.2), once reset from the orientation of the start box (which did not change from trial to trial), would be sufficient.

**Reviewed Result EX.2** *Rats can only learn to find food near a circularly symmetric landmark if the landmark is restricted to the central portion of the environment (Biegler and Morris, 1993).*

Biegler and Morris (1993; 1996) also tried to train animals to find food relative to the location of a single landmark. They found that they could only train the animal to find food relative to the location of the landmark if the landmark was only moved to locations in a small circumscribed (central) portion of the environment.

### 2.1.1-B Two landmark experiments.

Collett *et al.* also trained animals to find food at a location specified by a pair of landmarks.

**Reviewed Result EX.3** *Gerbils can be trained to search at a point relative to a pair of landmarks (Collett* et al.*, 1986).*

The search-time histogram is shown in Figure 2.3. In contrast to a single landmark, with two landmarks, the array can be manipulated to provide ambiguous or inconsistent cues. For example one landmark can be removed or the spacing between the landmarks can be changed.

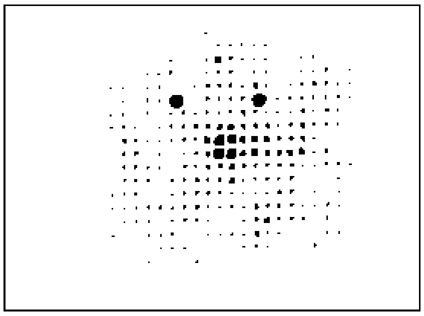


Figure 2.3: Histogram of search time on a two-landmark task. From Collett *et al.* (1986), reprinted with permission of author and publisher.

**Reviewed Result EX.4** *After being trained with two landmarks, gerbils shown only one landmark search in two locations (Collett* et al.*, 1986).*

In this case, Collett *et al.* report that the gerbils searched alternately in two locations, each at the correct distance and bearing from one of the landmarks they had observed during training (see Figure 2.4). Note that the landmark array used to probe this task is the same as in Experiment EX.1 above, but because the training regimen is different, the animals' reactions are different.

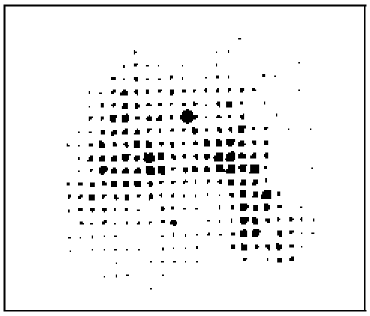


Figure 2.4: Histogram of search time of gerbils tested with one landmark, after being trained with two. From Collett *et al.* (1986), reprinted with permission of author and publisher.

**Reviewed Result EX.5** *After being trained with two landmarks, gerbils shown the two landmarks farther apart than during training search at two locations interior to the array (Collett* et al.*, 1986).*

The two-landmark array can also be stretched. When trained with two landmarks and tested with the distance between them doubled, the gerbils again searched at two locations interior to the array, each at the correct distance and bearing from the corresponding landmark, see Figure 2.5.

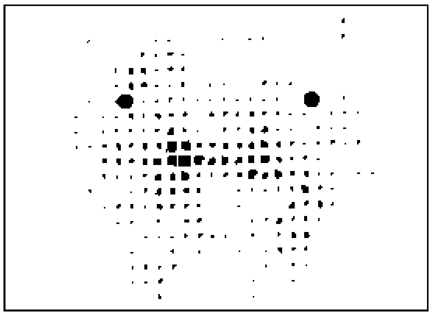


Figure 2.5: Histogram of search time of gerbils tested with two landmarks more separated than during training. From Collett *et al.* (1986), reprinted with permission of author and publisher.

**Reviewed Result EX.6** *It is easier to train rats to find food relative to separated landmarks than to clustered ones (Biegler and Morris, 1993; 1996).*

Biegler and Morris (1996) also tried to train animals to find food at locations relative to pairs of landmarks. They found that animals only followed the local landmark arrays if the arrays contained multiple landmarks not clustered together. When they tried to train animals relative to a single landmark or a cluster of landmarks, the animals searched near the landmark but at the wrong spatial relationship to it. When they trained animals to search at a position offset from a pair of spaced landmarks (70 cm separation), the animals concentrated their search time at the correct spatial relationship to the landmark pair (Biegler and Morris, 1996). It should be noted that Biegler and Morris disoriented their animals before placing them in the environment.

#### 2.1.1-C Three landmark experiments.

Collett *et al.* also tested manipulations of a third configuration: reward placed at the center of an equilateral triangular landmark array.

**Reviewed Result EX.7** *Gerbils can be trained to find food at the center of a triangle of landmarks (Collett* et al.*, 1986).*

As with the previous two configurations, well-trained animals searched for food at the center of the array, see Figure 2.6.

Again, as with the two landmark configuration, the three landmark configuration can be manipulated to produce inconsistent or ambiguous cues.

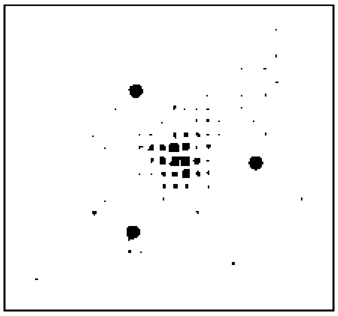


Figure 2.6: Gerbils trained to find food at the center of three landmarks. From Collett *et al.* (1986), reprinted by permission of author and publisher.

**Reviewed Result EX.8** *Gerbils tested with four landmarks after being trained with three search in the triangle at the correct orientation of three of the landmarks, ignoring the fourth (Collett* et al.*, 1986).*

If one adds a landmark to form a second triangle with opposite orientation (see Figure 2.7), the gerbils still search in the triangle with the correct orientation. Remember: the gerbils were not disoriented in any of Collett *et al.*'s experiments.

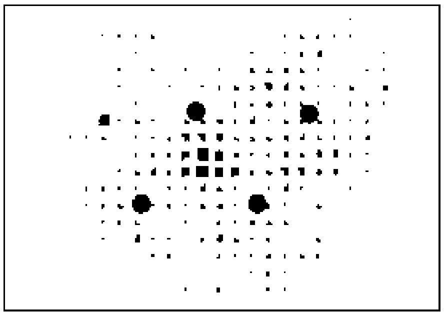


Figure 2.7: Histogram of search time of gerbils tested with four landmarks, after being trained with three. From Collett *et al.* (1986), reprinted with permission of author and publisher.

**Reviewed Result EX.9** *Gerbils tested with two landmarks after being trained with three prefer to search on one side of the pair than the other (Collett* et al.*, 1986).*

Collett *et al.* reports that when faced with only two of the three landmarks, the gerbils searched on one side and not the other (Figure 2.8). The side searched can be identified by matching the triangle with the training triangle at the same orientation. Note that this

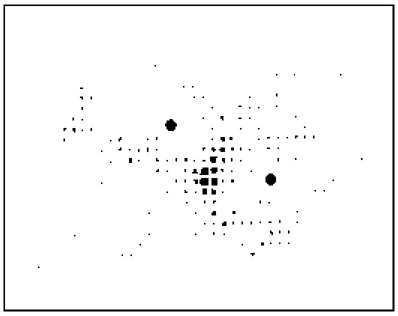


Figure 2.8: Histogram of search time of gerbils tested with two landmarks after being trained with three. From Collett *et al.* (1986), reprinted with permission of author and publisher.

is the same landmark configuration as in EX.3 above, but because the training regimen is different, the gerbils' search distribution is different.

**Reviewed Result EX.10** *Gerbils tested with one landmark after being trained with three search at three locations at the distance and bearing expected by matching the landmark to each of the training landmarks (Collett* et al.*, 1986).*

Collett *et al.* report that when trained with three landmarks and tested with one, gerbils searched at three locations corresponding to matching the landmark to each of the remembered landmarks in turn.

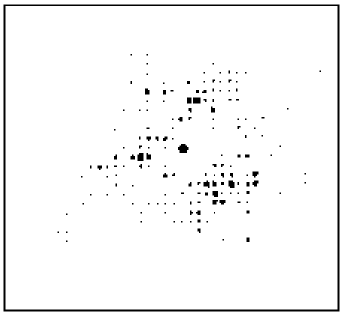


Figure 2.9: Histogram of search time of gerbils tested with one landmarks after being trained with three. From Collett *et al.* (1986), reprinted with permission of author and publisher.

**Reviewed Result EX.11** *When one landmark is moved away from the remaining two, gerbils prefer to search at the remembered distance and bearing to the non-displaced pair (Collett* et al.*, 1986).*

In addition to adding or removing landmarks, one can manipulate the positions of the three landmarks relative to each other. For example, one can stretch the triangle by moving one of the landmarks away from the other two, while leaving the distance between the other two unchanged. Collett *et al.* report that the gerbils searched closer to the two "unchanged" landmarks than to the third "stretched" one, see Figure 2.10.

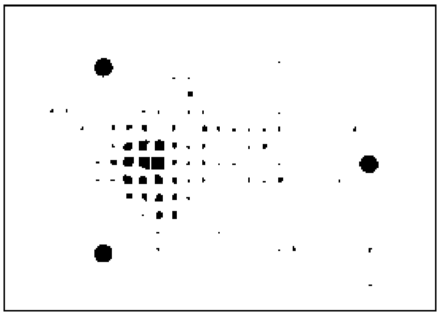


Figure 2.10: Histogram of search time of gerbils trained with three landmarks, tested with one landmark moved. From Collett *et al.* (1986), reprinted with permission of author and publisher.

**Reviewed Result EX.12** *Gerbils trained to find food at the center of a triangle of landmarks and tested with the triangle rotated by* 180° *search at four locations: the center and three "wings," at the correct bearing at distance to each of the three pairs (Collett* et al.*, 1986).*

Finally, the triangle can be inverted (or equivalently, rotated by 60°). When this is done, Collett *et al.* report that the gerbils first search the center of the array and then proceed to search three exterior points (Figure 2.11).

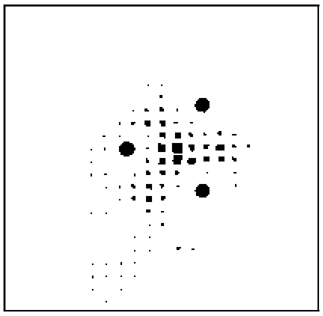


Figure 2.11: Histogram of search time of gerbils trained with three landmarks, tested with triangle rotated by 180°. From Collett *et al.* (1986), reprinted with permission of author and publisher.

### 2.1.2 The water maze.

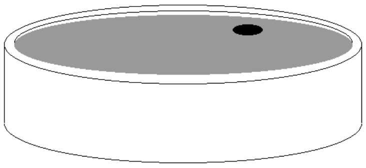


Figure 2.12: The water maze (Morris, 1981). Black circle indicates location of the platform. In the hidden-platform version, the platform is beneath the water surface and thus not visible to the animal. Typical tank diameters are 1–2 m.

Another classic navigation task is the *water maze*, first introduced by Morris (1981), see Figure 2.12. The water maze consists of a large pool of water mixed with milk, chalk, or paint so as to make the water opaque. Somewhere in the pool, there is a platform on which the rodent can stand and be out of the water. Sometimes the platform is submerged just below the surface; this is called the *hidden platform water maze*. Other times, the platform sticks out above the surface; this is called the *visible water maze*. Sometimes the location of the platform is indicated by a co-localized cue (such as a light bulb hanging directly over the platform. Even when hidden, this co-localized cue indicates the location of the platform. This version is called the *cued water maze* and is similar to the visible water maze.

The Morris water maze is one of the most used navigational tasks, not for its complex behavioral aspects, but because of its simplicity. It has been used to examine effects of lesions (Morris *et al.*, 1982; Sutherland *et al.*, 1983; Schenk and Morris, 1985; Kolb and Walkey, 1987; Sutherland *et al.*, 1988; DiMattia and Kesner, 1988; Annett *et al.*, 1989; Dean, 1990; Eichenbaum *et al.*, 1990; Morris *et al.*, 1990; Sutherland and Rodriguez, 1990; Packard and McGaugh, 1992; Taube *et al.*, 1992; Sutherland and Hoesing, 1993; Nagahara *et al.*, 1995; Whishaw *et al.*, 1995; Whishaw and Jarrard, 1996; Schallert *et al.*, 1996), neuropharmacological manipulations (Sutherland *et al.*, 1982; Whishaw, 1985), grafts (Nilsson *et al.*, 1987; Bjorklund *et al.*, 1990), and genetic mutants (Silva *et al.*, 1992; Tsien *et al.*, 1996b), as well as aging effects (Gallagher and Burwell, 1989; Gallagher *et al.*, 1993; Gallagher and Nicole, 1993; Gallagher and Colombo, 1995; Gallagher *et al.*, 1995; Barnes, 1996; Barnes *et al.*, 1997). I will review each of these manipulations in turn.

**Reviewed Result EX.13** *Normal rodents placed in the pool quickly learn to swim to a consistently located platform, even if it is hidden (Morris, 1981).*

If the platform is then removed from the pool, the animals spend most of there time near the location the platform used to be (Morris, 1981). Analogously to the open-field probe without food, this shows that the animals know the location of the platform and have an expectation of its location.

### 2.1.2-A Lesions.

**Reviewed Result EX.14** *Hippocampal or fimbria/fornix lesions impair navigation to hidden but not visible or cued platforms (Morris* et al.*, 1982; Sutherland* et al.*, 1983; Eichenbaum* et al.*, 1990; Morris* et al.*, 1990; Sutherland and Rodriguez, 1990; Packard and McGaugh, 1992).*

Morris *et al.* (1982; 1990) examined the effect of hippocampal lesions and found that lesioned rats swam in stereotyped circles the correct distance from the wall. Weisend *et al.* (1996) report that rats with ibotenic hippocampal lesions show a complete inability to remember learned water mazes out to 36 weeks (i.e. they show *retrograde amnesia*), however the lesions included large portions of dorsal subiculum and the authors did not test beyond 36 weeks. Koerner *et al.* (1996) compared (1) animals who learned the hidden-platform water maze and waited 12 weeks before receiving a hippocampal lesion, (2) animals who learned the task and received repetitions each week for 12 weeks before receiving the lesion, and (3) animals who learned 13 different hidden-platform water mazes over those 13 weeks before receiving the lesion. After hippocampal lesions (which again encroached on dorsal subiculum) only group (2) showed any residual performance above chance.

A number of authors have lesioned the fimbria/fornix as a simpler hippocampal lesion (Eichenbaum *et al.*, 1990; Sutherland and Rodriguez, 1990; Packard and McGaugh, 1992; Whishaw *et al.*, 1995). Most authors suggest that fornix lesions are similar to hippocampal lesions: they impair acquisition of a hidden but not a visible platform (Sutherland and Rodriguez, 1990; Packard and McGaugh, 1992) . Eichenbaum *et al.* (1990) and Whishaw *et al.* (1995) both report that fornix lesioned animals cannot learn the water maze with a variable start, but that they can learn the task if they are always started from the same place or if there is a cue identifying the path to take.

An important consideration is that a fimbria/fornix lesion severs the cholinergic input to the hippocampus, the connections from the subiculum to the nucleus accumbens, and the postsubicular connections to the lateral mammillary nuclei, but it does not sever the hippocampal/entorhinal connections (neither EC $\rightarrow$ HC nor HC $\rightarrow$ EC). Therefore fimbria/fornix lesions are not equivalent to hippocampal lesions.

**Reviewed Result EX.15** *Rats with colchicine lesions show impairments in hidden-platform water maze tasks if given the lesion one week after training, but not if given the lesion twelve weeks after training (Sutherland and Hoesing, 1993).*

Sutherland and Hoesing (1993) tested rats with colchicine lesions[1] in the water maze either one week after training or twelve weeks after training. If the rats were trained on the task and then given the lesion one week later, they showed profound deficits. However, the same lesion twelve weeks after training produced much smaller deficits.

McNaughton *et al.* (1989a) also examined the effect of colchicine lesions on the water

[1] Colchicine lesions destroy the dentate gyrus but leave CA3 and CA1 mostly intact (Goldschmidt and Steward, 1980, see McNaughton *et al.*, 1989a; Sutherland and Hoesing, 1993).

maze. They also found that colchicine lesions produced severe deficits in acquisition of the water maze. McNaughton *et al.* did not examine retention.

**Reviewed Result EX.16** *Rats with hippocampal or fimbria/fornix lesions can learn to perform the hidden-platform water maze under certain conditions (Morris* et al.*, 1990; Whishaw* et al.*, 1995; Schallert* et al.*, 1996; Whishaw and Jarrard, 1996).*

Whishaw *et al.* (1995) trained animals with fimbria/fornix lesions to find a visible platform and then removed the visible platform. These animals concentrated their search where the platform had been. Whishaw and Jarrard (1996) have shown the same effect with cytotoxic hippocampal lesions. Schallert *et al.* (1996) used animals with kainate-colchicine hippocampal lesions (destroying both DG and the CA3/CA1 fields). The animals were first trained with a large platform that filled almost the entire maze. Once the animals could reach that platform reliably, it was shrunk trial by trial until it was the same size as a typical platform in a water maze task. With this training regimen, the animals could learn to solve the water maze without a hippocampus.

Morris *et al.* (1990) showed a similar result in which "overtrained animals" saw 48 interspersed hidden- and cued-platform trials and found that neither hippocampal nor subicular lesions alone were sufficient to produce deficits, but that combined hippocampal and subicular lesions still produced deficits.

**Reviewed Result EX.17** *While hippocampal lesions/inactivation impair navigation to hidden-platforms, caudate nucleus lesions/inactivation impair navigation to visible platforms.*

In the water maze, lesions of the caudate nucleus disrupt navigation to cued platforms, such as visible platforms or platforms with a large black card marking the quadrant containing the goal, but not to hidden platforms (McDonald and White, 1994), while hippocampal or fimbria/fornix lesions impair acquisition of hidden platforms, but not visible or cued platforms (EX.14).

McDonald and White (1994) trained rats in the water maze for twelve days. On the first three days the animal was trained with a visible platform, but on the fourth day, the animals had to find a hidden platform at the same location. This four-day sequence was repeated three times. After these twelve days, the animals were tested with a visible platform in a new location. Half of the animals went to the new visible platform, half went to the location where the old platform had been. With hippocampal inactivation (by lidocaine), all of the animals went to the new platform, but with caudate inactivation, all of the animals went to the old location.

Packard and McGaugh (1992) compared caudate and fornix lesions on a water maze with two platforms, one stable, one unstable. If the animal tried to climb up on the unstable platform, it fell back into the water. Both platforms were hidden. Packard and McGaugh ran two versions of the experiment: In the first version, the stable platform was always in one quadrant, while the other was always in the opposite, so location of the platform denoted its stability; in the second version, the platforms alternated quadrants randomly from trial to trial, but a visual cue demarcated which of the two was stable. They found that fornix lesions impaired acquisition of the location-stable platform, but caudate lesions did not effect the performance on the location-stable platform. On the other hand, caudate lesions impaired acquisition of the cue-stable platform, but fornix lesions did not affect the ability to find the cue-stable platform.

**Reviewed Result EX.18** *Rats with subiculum lesions appeared to search randomly like naive rats (Morris* et al.*, 1990).*

While hippocampal animals circle the environment at the correct distance from the wall (implying that they may know some information about the location of the hidden platform), subicular animals wander the environment seemingly randomly.

**Reviewed Result EX.19** *Entorhinal cortex lesions show sparing of some abilities, but a sparing of others in the water maze (Schenk and Morris, 1985; Nagahara* et al.*, 1995).*

Entorhinal lesions (Schenk and Morris, 1985; Nagahara *et al.*, 1995) are much less consistent, some showing sparing of certain abilities (such as crossing the actual platform location) while showing deficits in other abilities (such as spending time in the correct quadrant). Nagahara *et al.* (1995) found that rats with entorhinal lesions were more impaired after a five minute delay relative to a 30 second delay. In contrast, Schenk and Morris (1985) found extensive spatial deficits with both acquisition and retention, but Schenk and Morris admit that their "entorhinal" lesions encroached on pre-subiculum, para-subiculum, and subiculum proper, any or all of which may have had strong effects.

**Reviewed Result EX.20** *Nucleus accumbens lesions produce deficits in naive rats on the hidden platform version, but not in the visible platform or with pre-trained rats (Annett* et al.*, 1989; Sutherland and Rodriguez, 1990).*

Nucleus accumbens lesions also produce deficits in naive rats on the hidden platform version, but not in the visible platform or with pre-trained rats (Annett *et al.*, 1989; Sutherland and Rodriguez, 1990). Annett *et al.* (1989) reports that their rats did eventually learn the task, although they were never as good as normals. However, the lesions done by Annett *et al.* were incomplete.

**Reviewed Result EX.21** *Anterior cingulate lesions have no effect on learning the water maze, but animals with posterior cingulate lesions can not learn to go directly to a hidden platform (Sutherland* et al.*, 1988; Sutherland and Hoesing, 1993).*

Sutherland and Hoesing (1993) report that when unilateral hippocampal lesions are combined with contralateral posterior cingulate lesions the deficits are as severe as bilateral hippocampal or bilateral cingulate lesions, in other words, devastating. These effects occur even if the cingulate lesions are made 12 weeks after training.

**Reviewed Result EX.22** *Lesions to the anterior thalamic nuclei (ATN) produce severe deficits in the water maze (Sutherland and Rodriguez, 1990), as do postsubicular (PoS) lesions (Taube* et al.*, 1992).*

Taube *et al.* (1992) found that although performance in the water maze was impaired relative to normals, performance did improve over time. Neither Sutherland and Rodriguez (1990) nor Taube *et al.* (1992) found errors in the cued water maze.

**Reviewed Result EX.23** *Lesions to the parietal cortex produce severe deficits in both visual and cued water maze tasks (Kolb and Walkey, 1987; DiMattia and Kesner, 1988; Kolb, 1990a).*

Animals with parietal lesions show poor initial heading errors even with visual or cued platforms (Kolb and Walkey, 1987). Kolb (1990a) reports that they never improve their initial trajectory, even with visual cues indicating the location of the platform. DiMattia and Kesner (1988) suggest that animals are reduced to random search strategies with parietal lesions.

**Reviewed Result EX.24** *Te2 lesioned rats are unable to learn to perform a visual match to sample, even with no delay, but they can learn the hidden-platform water maze (Kolb, 1990a).*

In contrast to posterior parietal lesions, Te2 lesioned rats[2] are poor at visual pattern discriminations, but not at spatial orientation discriminations (Kolb, 1990a). They are unable to learn to perform a visual match to sample, even with no delay (Kolb, 1990a). But they can learn the hidden-platform water maze normally.

### 2.1.2-B Neuropharmacological manipulations.

One of the effects fimbria-fornix lesions have is to remove the cholinergic input into the hippocampus. Some researchers have examined the role of cholinergic blockers on the water maze explicitly.

**Reviewed Result EX.25** *Rats with atropine sulfate infusions shown no improvement in finding platforms that remain in a single location over animals shown random platform locations (Sutherland* et al.*, 1982). But given enough training, animals with atropine sulfate infusions can learn the location of a single hidden platform (Whishaw, 1985).*

In addition to the lack of improvement over random platform locations, Sutherland *et al.* (1982) report the number of times the animals reared when on the platform did not decrease from trial to trial.

Whishaw (1985) notes that Sutherland *et al.* only tested their rats for one week, so it cannot be determined whether the problem is with rapid acquisition or if there is a more

[2] Te2 is a ventral visual area (Zilles, 1990).

general permanent impairment. Whishaw found that although the atropine-infused rats take longer to learn, they do eventually learn. If the start location is moved, but the platform is not, both atropine-infused and normal rats continue to show the same ability to find the platform. On the other hand, if the platform is moved, they both take a long time to find the new platform (presumably because they spend their time searching at the location it had previously been). This implies the strategy used by both normals and atropine rats is spatial. Whishaw also found that atropine given to rats who had already learned the task had no significant effect.

However, Whishaw found that atropine-infused rats could not learn a set of places, although normals could, if, for example, the platform was moved through a regular sequence of locations. When rats that had acquired a learning set were given atropine, their performance was totally disrupted.

#### 2.1.2-C Grafts and the water maze.

If the main effect of fimbria-fornix lesions is to remove the cholinergic input from the hippocampus, then it may be possible to regraft fetal tissue into the septal areas and if it were to reinnervate the hippocampus, then hippocampal function might be restored. Nilsson *et al.* (1987) have done just that.

**Reviewed Result EX.26** *Regrafted septal tissue can restore lost abilities to perform the hidden-platform water maze (Nilsson* et al.*, 1987).*

Nilsson *et al.* (1987) lesioned the fimbria, fornix, and corpus callosum and then grafted fetal septal tissue into the third ventricle near the dorsal hippocampus. They found (1) the fimbria-fornix transection disrupted spatial learning, (2) animals with grafts were able to use spatial cues, although they were impaired relative to normals both in learning time and in spatial accuracy, and (3) atropine sulfate disrupted the spatial ability of animals with grafts.

Bjorklund *et al.* (1990) showed that in fact the grafts re-innervated the hippocampus with a similar layering to that seen in normal animals. They also review data that rats with combined septal and raphe (supplying 5-HT) lesions required both septal and raphe grafts in order to see improvements in the water maze, and that the raphe transplants showed similar afferentations to that of normal animals.

#### 2.1.2-D Genetic mutants.

**Reviewed Result EX.27** *Mice with deficient NMDA receptors cannot learn the hidden platform water maze (Silva* et al.*, 1992; Tsien* et al.*, 1996b).*

There has been some amazing progress in genetic manipulation. The technology now exists to produce animals with deficits in specific areas, specific neurotransmitters, even specific receptors. I do not have the room to review the recent genetics advances here (it would be a major digression), but I want to note two recent studies examining genetic mutants in the water maze.

Silva *et al.* (1992) developed a mouse without LTP (with a malfunctioning NMDA receptor), and found that it could not learn the water maze. Tsien *et al.* (1996a) developed a mouse with malfunctioning NMDA receptors only in CA1. These mice also cannot learn the water maze (Tsien *et al.*, 1996b).

#### 2.1.2-E Aging and the water maze.

**Reviewed Result EX.28** *Old animals are deficient in the hidden platform water maze (Gallagher* et al.*, 1993; Barnes* et al.*, 1997, see Barnes, 1996 for a review).*

Normal old rats generally show deficits in all standard measures of the hidden platform water maze, but not in the cued or visible platform version (Gallagher *et al.*, 1993; Barnes *et al.*, 1997). Barnes *et al.* (1997) report that both young and old animals show a distinct bimodality in ability to solve the hidden platform water maze — on some trials, they take relatively direct paths to the goal while on others they are severely impaired. This is usually measured by the taking the total path length between the release point and the goal. In normal young animals, the "short" trials soon dominate, but in old animals, the "long" trials continue to perseverate even after extensive training.

#### 2.1.2-F Latent learning in the water maze

**Reviewed Result EX.29** *Animals can learn to navigate to a platform through a part of the environment in which they had no experience (Keith and McVety, 1988; Matthews* et al.*, 1995).*

Sutherland *et al.* (1987) tried to examine latent learning in the hidden platform water maze: they placed a barrier across the middle of the tank so that the animals could only experience half of the maze when searching for the platform. They then removed the barrier and allowed the animals to start from the other half (in which they had no experience) and found that animals were severely impaired in finding the water maze. However, Matthews *et al.* (1995) point out that the Sutherland *et al.* (1987) task is confounded by the existence (and then absence) of a very salient cue — the barrier. Matthews *et al.* (1995) repeated the experiment, but moved the barrier slowly away from the animal, so that on one trial it allowed access to half the arena and on the next it allowed access to a little more than half, and on the next, more than that, etc. At first, the animals went to examine the barrier, but they quickly learned to ignore it. After they had learned to ignore it, Matthews *et al.* tested the animals from a part of the maze in which they had never been. Matthews *et al.* found that the animals went directly to the platform. In another experiment in the water maze, Keith and McVety (1988) also showed that animals could learn to navigate to a platform through a part of the environment in which they had no experience.

### 2.1.3 The hole-board circular platform

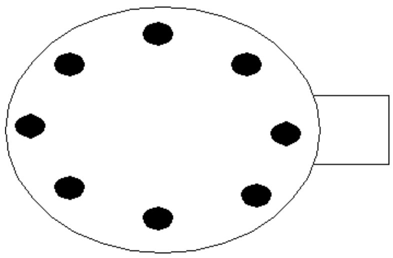


Figure 2.13: The hole-board circular platform (Barnes, 1979). Black circles indicate holes in the board. One hole leads to a hidden nest. The arena is bathed in bright light. In order to escape the bright light, rats try to find the hole with the nest where they can hide.

The hole-board arena (introduced by Barnes, 1979) is similar to the water maze in that it forces animals to use distal cues to find an escape route from an aversive stimulus (in this case bright light). It differs in that the animals have to choose among discrete holes rather than find a hidden platform in a continuum of open water. Because the holes are all around the edge of the platform, it would be sufficient to know the orientation of the escape hole.

**Reviewed Result EX.30** *Old animals show deficits in finding a hidden location in the hole-board circular platform (Barnes, 1979). Young animals show better retention of the location of the escape hole even after a week's delay (Barnes* et al.*, 1980).*

Barnes (1979) found that both old and young animals showed three basic behavioral patterns: (1) examination of many holes with many center-crossings, (2) systematic exploration of the holes, going from one hole to the next, and (3) taking a direct path to the escape hole. Early on rats showed the first pattern, followed eventually by the second, and finally (after enough trials) the third. Old animals were impaired on this task but did show a spatial strategy in that when the tunnel was moved, their latency to goal increased significantly.

Like the water maze, after switching the location of the escape tunnel, animals with a nonspatial strategy (such as check all holes randomly) will not change their time-to-goal, while animals who prefer to keep trying to use a spatial strategy will (Barnes, 1979).

**Reviewed Result EX.31** *Colchicine-induced lesions produced deficits in acquisition of the hole-board circular platform (McNaughton* et al.*, 1989a).*

As reviewed above (EX.15), colchicine destroys the dentate granule fields while sparing the CA3 and CA1 hippocampal fields (Goldschmidt and Steward, 1980, see McNaughton *et al.*, 1989a; Sutherland and Hoesing, 1993). McNaughton *et al.* (1989a) report that animals with colchicine lesions showed many more errors than controls and were relatively less impaired on "reversal trials" than normals (changing which hole led to the escape tunnel) implying they were using a non-spatial strategy to find the escape tunnel.

**Reviewed Result EX.32** *Mice genetically engineered to have deficient LTP cannot learn the circular platform escape task (Bach* et al.*, 1995).*

Bach *et al.* (1995) tried to train mice who had been genetically engineered to be deficient in LTP under certain conditions. (LTP was ineffective when the shocks given to the slice occurred at a frequency in the 5–10 Hz, but LTP was still normal when the stimulus was in the 100 Hz range). They found that animals could not learn the location of the escape hole based on the distal cues.

### 2.1.4 Small, enclosed arenas.

Figure 2.14: Small, enclosed arenas (Muller *et al.*, 1987). Food is scattered on the floor of a small cylinder or rectangle. Typical cylinder diameters are 50–100 cm.

**Reviewed Result EX.33** *In a odor-match-to-place task, lateral entorhinal cortex (LEC) lesions produce odor discrimination errors, but LEC + medial entorhinal cortex (MEC) lesions produce spatial errors as well (Otto* et al.*, 1996).*

Otto *et al.* (1996) examined a task which paired an odor with a location (given a specific odor, go to a specific location), LEC lesions produced occasional odor discrimination errors (animal went to the location indicated by the wrong odor). When the lesions encroached on MEC, the animals also occasionally made spatial errors (animal went to a location not associated with any odor).

**Reviewed Result EX.34** *Hippocampal lesions combined with caudate lesions produced a drastic increase in locomotor behavior, while hippocampal lesions combined with accumbens lesions produce a drastic decrease in locomotor behavior (Whishaw and Mittleman, 1991).*

A number of early experiments (see Isaacson, 1974) found that hippocampal lesions produced hyperactivity in open arenas. Whishaw and Mittleman tested hippocampal lesions with 6-OHDA-induced damage to the dopaminergic inputs to the caudate or nucleus accumbens. They found that hippocampal + caudate lesions produced decreased stereotypical movements but increased locomotion while hippocampal + accumbens lesions produced decreased locomotion with increased stereotypic movements.

**Reviewed Result EX.35** *Disoriented rats cannot learn to differentiate two geometrically similar but identifiably different corners (Cheng, 1986); non-disoriented rats have no trouble differentiating these two corners (Margules and Gallistel, 1988).*

Cheng, Margules, and Gallistel (Cheng, 1986; Margules and Gallistel, 1988; Gallistel, 1990) found a strong effect of disorientation (or a lack of it) on behavior in a rectangular arena.

Cheng (1986) tried to train rats to find food at one corner of a rectangular arena (Figure 2.14). In order to make the corners as distinct as possible, he placed a panel at each one, covered with a different type of material. In addition, the panels had different numbers of pinholes through which light was visible, and two of the panels had unique odorants behind them. Cheng disoriented the rats before placing them in the arena at a random location, and found that although the rats were able to distinguish one pair of diagonally opposed corners from the other, they could not distinguish the two corners in each pair. Cheng reports that the animals chose the correct corner in approximately 50% of the trials, and in the other 50%, they chose the corner opposite it. This suggests that the rats are sensitive to the geometric structure of the environment, and were ignoring other cues that could distinguish between the two corners.[3] Margules and Gallistel (1988) replicated the experiment without disorientation and found that most animals had no difficulty selecting the correct corner over 75% of the time; some achieved better than 90% success rates.

### 2.1.5 T, Y, and plus mazes

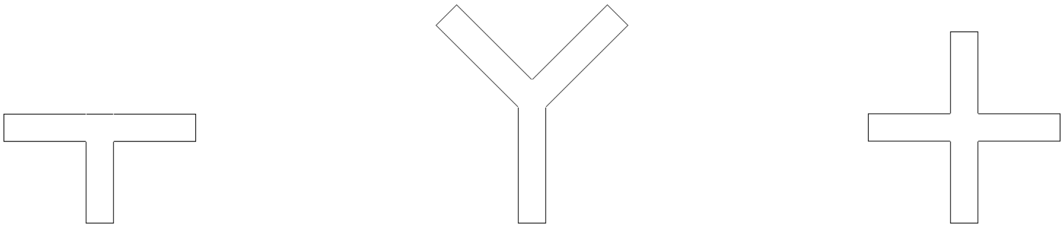


Figure 2.15: The T, Y, and plus mazes. Typical arm lengths are 50 cm.

**Reviewed Result EX.36** *Rats trained to take an* L*-shaped path proceed directly towards the goal when the normal path is blocked (Tolman* et al.*, 1946a; Tolman, 1948).*

[3]Humans also make similar errors: Hermer and Spelke (1994) did an equivalent experiment on young children: a toy was hidden (in the child's presence) in one of two differently colored boxes in two corners of a rectangular room. The child was then spun and asked to find the toy. Even human children chose the wrong box 50% of the time. But if the boxes were moved out of the corners of the room (so the children could no longer use spatial strategies), they always chose the correct box. This even works on adults (Hermer and Spelke, 1994). While normal adults do not show any problems, even when spun, if they are given a distractor task which interferes with linguistic memory, they show these same errors.

Tolman *et al.*'s (Tolman *et al.*, 1946a; Tolman, 1948) experiments showing that rats trained to take an L-shaped path would proceed directly towards the goal when the normal path was blocked showed conclusively that rats were not learning a *habit* (i.e. make response $x$ to stimulus $y$), but were learning to go to a place or to approach a stimulus. Recently, some have taken this task to be indicative of path integrative abilities, but a light bulb above the goal indicated the direction, so this task could be solved by learning to approach the light bulb.

**Reviewed Result EX.37** *Rats can be more easily trained to go to a place than to engage in a response (Tolman* et al.*, 1946b)*

Tolman *et al.* (1946b) tested rats on a plus maze, training one group to always turn left, whether started from the north or south arm of the plus, and training the other group to go to a place (e.g. turn left from the north arm and right from the south, always going east). They found place-learning rats to take from 11–18 trials to reach a criterion of 10 perfect trials. In contrast, the response-learning rats took 25 or more trials and four of the seven never reached criterion in 72 trials.

**Reviewed Result EX.38** *Animals on the Y-maze show latent learning (Tolman, 1948).*

The Y-maze used to demonstrate latent learning includes a food reward at the end of one arm of the Y and a water reward at the other. Animals are allowed to explore this environment but are neither hungry or thirsty. Half of the animals are then made hungry and the other half thirsty. The hungry animals immediately run to the food source and thirsty animals to the water source (Tolman, 1948). This demonstrates that animals do not have to be strongly rewarded for them to learn goal locations.

**Reviewed Result EX.39** *Fimbria/fornix lesions impair rats' abilities to reach two goals on alternate days (Hirsh* et al.*, 1978).*

Hirsh *et al.* (1978) ran rats on the same Y-maze used by Tolman *et al.* (see Tolman, 1948, EX.38) with food at one end and water at the other. Animals were alternately hungry and thirsty (on alternate days). Normal animals learned this task easily, going to the food source when hungry and to the water source when thirsty. However, animals with fornix-lesions could not learn both tasks. They learned one task completely and then unlearned that task as they learned the other. In other words, they could learn to go to one of the arms, but could not remember two rewards on two arms.

**Reviewed Result EX.40** *Old animals prefer to follow cues or perseverate motor responses rather than to use spatial strategies (Barnes* et al.*, 1980).*

Barnes *et al.* (1980) tested old and young animals on a T-maze with differing textures on the floor of the two goal-arms. Animals were trained to always make a right turn with the T always in the same orientation. This allows the animals to use 3 strategies (see Navigation Overview, Chapter 3): a cue strategy (turn based on floor texture), a response strategy (always turn right), or a spatial strategy (go to a place). These can be differentiated by (1) flipping the textures, (2) rotating the T or (3) moving the cues. Both young and old animals were most likely to use response strategies, but young animals used place strategies much more than old. One problem with this experiment is that animals may switch strategies depending on the amount of training received. This was explicitly explored by Packard and McGaugh (1996, see next Reviewed Result).

**Reviewed Result EX.41** *When trained to turn left from the south to the west arms of a plus maze, and then tested from the north arm, rats with less than a week of training turn right (to the west arm), but rats with two weeks of training turn left (to the east arm) (Packard and McGaugh, 1996). This second strategy is mediated by the caudate nucleus (Packard and McGaugh, 1996).*

Packard and McGaugh (1996) tested rats on a plus-maze: rats were always started on the south arm and trained to turn left (to the west arm). The animals were given seven days of four trials per day. On the eighth day, they were placed on the north arm. They all turned right (to the west arm). This demonstrates that they were using a place or *locale* strategy (see Navigation Overview, Chapter 3). The animals were then trained for seven more days to turn left from the south arm to the west arm. Then, on the sixteenth day, they were again placed on the north arm. They all turned left (to the east arm), demonstrating a response or *praxic* strategy (see Navigation Overview, Chapter 3). Packard and McGaugh then inactivated caudate (with lidocaine) and tested the animals on the north arm. They all turned right, indicating the place strategy again.

### 2.1.6 The radial arm maze

The radial maze, shown in Figure 2.16, consists of a central platform with thin arms radiating out from it equally spaced around the circle. A four-arm radial maze is essentially equivalent to the plus maze (Figure 2.15), but most radial mazes have eight or more arms. This task is one of the most popular navigation tasks studied and has been studied under a number of different conditions, including after cue manipulations (Olton and Samuelson, 1976; Suzuki *et al.*, 1980), while recording extracellularly (McNaughton *et al.*, 1983a; Mizumori and Williams, 1993; Chen *et al.*, 1994a; Chen *et al.*, 1994b; McNaughton *et al.*, 1994a; Lavoie and Mizumori, 1994; Young *et al.*, 1994; Markus *et al.*, 1995; Mizumori and Cooper, 1995; Dudchenko and Taube, 1997), after neuropharmacological manipulations (Decker and McGaugh, 1991; Ohta *et al.*, 1993; Shen *et al.*, 1996), after lesions (Olton *et al.*, 1980; Taube *et al.*, 1992; Jarrard, 1993; Floresco *et al.*, 1997), and grafts (Shapiro *et al.*, 1989; Hodges *et al.*, 1991a; Hodges *et al.*, 1991b), and in aged animals (Barnes *et al.*, 1980; Barnes, 1988; Mizumori *et al.*, 1996; see Barnes, 1996 for a review).

Olton and Samuelson (1976) first introduced the radial maze as a spatial *working memory* task: Each arm of the maze was baited. After the reward was taken from the end of an arm,

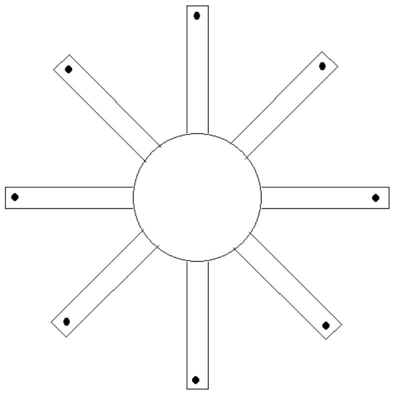


Figure 2.16: The radial arm maze (Olton and Samuelson, 1976). Black ovals indicate food locations. Each arm is baited only once. Some mazes also include gates at the entrances to each arm. In some paradigms, some arms are never baited. Typical arm lengths are 50–100 cm.

the arm was not rebaited. This meant that if an animal returned to an arm it had already visited, it did not receive any additional reward. Returning to an unbaited arm was scored as a *working memory error*. In the *delayed working memory* version of this task, the animal is allowed to enter three of the arms, then a delay was imposed, after which the rats could enter any arm. In the *reference memory version*, some of the arms are never baited (Olton *et al.*, 1979). This forces the animals to remember those arms across all trials.

**Reviewed Result EX.42** *Rats choose an average of more than seven different arms in their first eight choices (Olton and Samuelson, 1976).*

This suggests that the animals are able to recognize the different arms. They don't use an easily recognizeable pattern but are still able to differentiate the arms. By interchanging arms, Olton and Samuelson (1976) were able to show that animals were going to spatial locations not of arms the maze.

Rats could also be trained to use a cued version of the task in which the arms had distinctive floor textures and the arms were rotated after each choice (Olton *et al.*, 1979). Unfortunately, this task includes a strong delay component after each choice as the arms are rotated, and there is no way to remove that delay component from the cued version of the task.

**Reviewed Result EX.43** *In the delayed version, rats skip the previously visited arms and only choose the remaining arms once (Olton and Samuelson, 1976).*

Olton and Samuelson (1976) showed that accuracy is not dependent on time of delay but only on the number of choices experienced before the delay.

**Reviewed Result EX.44** *If the constellation of distal landmarks is rotated by* 180° *during the delay, rats rotate their remaining arm choices by* 180° *(Suzuki* et al.*, 1980). However, if the constellation of landmarks is permuted instead of rotated, rats behave as if the environment were unfamiliar (Suzuki* et al.*, 1980).*

The behavior of rats is sensitive to the constellation of external cues, not to individual ones. Suzuki *et al.* (1980) tested rats on the eight-arm radial maze: they allowed the animals to retrieve food from three of the eight arms and then gave the rats a delay. During the delay the landmarks were rotated as a group by 180°. When the animals were returned to the maze, they looked for food in the 180° rotational equivalent of the arms they had not been in before. In contrast, when landmarks are permuted rather than rotated, rats behave as if the environment were unfamiliar (Suzuki *et al.*, 1980). This implies that rats are sensitive to combinations of distal landmarks.

**Reviewed Result EX.45** *Rats with hippocampal and fimbria-fornix lesions show both reference memory and working memory "place" errors (Jarrard, 1993).*

Jarrard (1993) used two different tasks, a "place" task in which which arms were never baited depended on extramaze cues and a "cue" task in which the never-baited arms depended on intramaze cues. They found that animals with ibotenic hippocampal lesions made both reference memory and working memory errors on the place task but not on the cue task. However, Jarrard also found that animals trained on the place task before surgery did not show the same errors.

**Reviewed Result EX.46** *Transient lesions of the ventral hippocampus and subiculum produce errors in a spatial working memory task (Floresco* et al.*, 1997).*

Floresco *et al.* (1997) tested both a non-delayed and delayed versions of the radial maze with bilateral infusions of lidocaine into the ventral aspect of CA1 and subiculum. They found no effect when the lidocaine was injected before training, but if it was injected after training, they found a significant effect, in particular, rats with lesions made more across-delay errors.

**Reviewed Result EX.47** *Transient unilateral lesions of ventral hippocampus and subiculum combined with transient contralateral unilateral lesions of nucleus accumbens produces errors in the non-delayed but not in the delayed versions, while similar lesions of ventral hippocampus and subiculum with prelimbic cortex produce errors on the delayed but not the non-delayed versions (Floresco* et al.*, 1997).*

Floresco *et al.* (1997) also tested disconnection lesions of the ventral CA1 and subiculum with nucleus accumbens and ventral CA1 and subiculum with prelimbic frontal cortex. They found no effect of sub-PL lesions on the non-delayed version, but found a strong effect on the delayed version. Conversely, they found no effect of sub-NAcb lesions on the delayed version, but found a strong effect on the non-delayed version.

**Reviewed Result EX.48** *Lidocaine lesions of nucleus accumbens produce errors in spatial win-shift but not cued win-stay tasks (Seamans and Phillips, 1994).*

Seamans and Phillips (1994) tested a delayed version of the radial maze and found that lidocaine injected into nucleus accumbens affected the abilities of animals when infused during the delay. It also affected abilities when infused prior to the pre-delay phase, but much less so.

Seamans and Phillips (1994) also tested a cued version where animals had to find food on the same four arms of an eight-arm maze before and after the delay. Lidocaine in nucleus accumbens does not affect this second task.

**Reviewed Result EX.49** *Caudate lesions impair certain navigation tasks on radial mazes (Potegal, 1982).*

Potegal (1982) also showed that caudate lesions impair certain navigation tasks on radial mazes. Animals were trained on a 12-arm radial maze in which the food was located on a single arm at a constant angle from the starting arm. This meant that the only viable strategy to find the food was to go to the center of the maze and make a turn at that angle. Caudate lesions impaired the animal's ability to find the food.

**Reviewed Result EX.50** *Animals with colchicine lesions are impaired on reference memory in the radial arm maze (McNaughton* et al.*, 1989a).*

McNaughton *et al.* (1989a) baited one arm of the maze and tried to train an animal to find the food. A different arm was baited each day, but the same arm was baited all that day. This is reference memory, not working memory. Animals with colchicine infusions (which destroys the dentate gyrus granule cells) made more errors than normals.

**Reviewed Result EX.51** *Post-training entorhinal lesions produce both working memory and reference memory errors in spatial but not cued tasks (Rasmussen* et al.*, 1989).*

Rasmussen *et al.* (1989) examined rats on a collection of tasks on the radial maze that purport to examine working memory and reference memory across the cue-spatial axis. Non-spatial reference memory: one floor-textured-cued arm was baited. Spatial reference memory: one externally-cued arm was baited. Non-spatial working memory: four arms of the eight were baited; after each choice the arms were rearranged (forcing this to be a delayed version). Spatial working memory: rat was allowed to visit four arms, a delay was imposed, and the rat was allowed to visit any of the eight arms. Rasmussen *et al.* found that after entorhinal lesions, rats took longer to reach criterion on the spatial reference memory and working memory tasks, but not on the non-spatial tasks.

**Reviewed Result EX.52** *Postsubiculum lesions produce deficits in the radial maze (Taube* et al.*, 1992).*

Taube *et al.* (1992) found postsubicular lesions to affect the radial maze, but lesioned animals did improve with additional training (although they never reached the levels of normals even with 15 days of training.

**Reviewed Result EX.53** *Radial maze and aging: Old animals make many working memory errors on the radial maze (Barnes* et al.*, 1980; Mizumori* et al.*, 1996).*

Barnes *et al.* (1980) found that old animals repeated 30% more sequences than young animals. Mizumori *et al.* (1996) found that old animals make almost three times as many errors as young animals on the radial maze.

**Reviewed Result EX.54** *Scopolomine infusions impair performance on the radial maze (Eckerman* et al.*, 1980; Wirsching* et al.*, 1984; Buresova* et al.*, 1986; Hodges* et al.*, 1991a; Hodges* et al.*, 1991b).*

Buresova *et al.* (1986) tested rats in a water-tank modification of the radial maze (each arm was a track underwater which the animals could swim in; goals were small platforms at the ends of each arm which could be dropped to the bottom as a means of de-baiting the arm). They found that scopolomine (a cholinergic antagonist) affected this task, making animals repeat more arms. Not surprisingly, the proportion of errors increased with the number of arms already visited. Buresova *et al.* (1986) also tested a delayed version of this task (rats were made to wait for 40 minutes on a central platform). Buresova *et al.* found that (1) rats made about twice as many errors in the delayed version than in the uninterrupted version, (2) scopolomine produced a dramatic increase in the number of errors (quadrupling the effect of scopolomine on the uninterrupted version). Other researchers (Eckerman *et al.*, 1980; Wirsching *et al.*, 1984; Hodges *et al.*, 1991b) also found that cholinergic antagonists disrupted working memory on the radial maze. Eckerman *et al.* (1980) only tested working memory, but Wirsching *et al.* (1984) tested both working and reference memory and found that at low doses, working memory was much more impaired than reference memory, but did find a trend to show increased reference memory errors at higher doses. Hodges *et al.* (1991b) also report that low doses of cholinergic antagonists show more marked effects on working memory than reference memory.

### 2.1.7 More complex mazes.

Many of the early experiments in rodent navigation were done in extremely complex mazes (Watson, 1907; Carr and Watson, 1908; Carr, 1917; Dennis, 1932; Honzik, 1936; Tolman, 1948). Although the environments in which rodents are tested have become simplified over recent years (this allows better control of the variables involved), there is still important data to be gleaned from these early experiments.

**Reviewed Result EX.55** *Well-trained rats can navigate complex mazes under extreme sensory-deprivation (Watson, 1907; Carr and Watson, 1908; Carr, 1917; Honzik, 1936).*

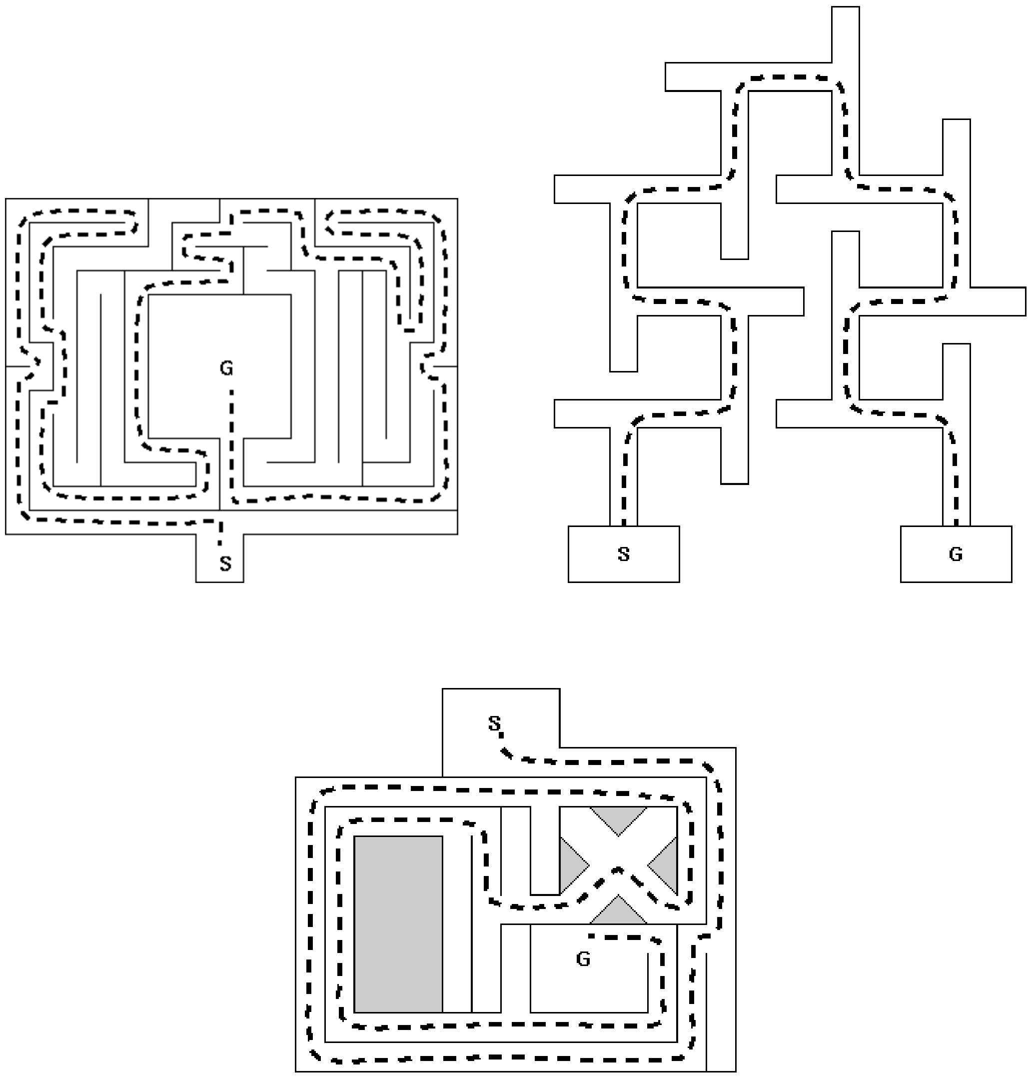



Figure 2.17: Some complex mazes used in the early rodent navigation literature. S indicates the start and G indicates the goal. Dashed line indicates successful path. (top left) The Hampton-court maze used by Watson (1907). (top right) The 14-dead end maze used by Honzik (1936). (bottom) The maze used by Carr and Watson (1908).

Honzik (1936) found that rats could navigate complex mazes even with strong sensory deficits (i.e. after being blinded or made anosmic). Watson (1907) found that animals could navigate a Hampton-Court maze without ever touching the walls, even when made blind, deaf, anosmic, and lacking vibrisae.

**Reviewed Result EX.56** *Well-trained rats run into walls when the lengths of the corridors are shortened (Carr and Watson, 1908, see also Dennis, 1932).*

Carr and Watson (Watson, 1907; Carr and Watson, 1908; Carr, 1917) report that well-trained rats run the maze very fast and "with confidence". They note that the rats do not seem to be using external sensory cues to guide them. When placed in the maze at a point along the route (i.e. not the start), the rat seems confused at first but then once the rat "gets his cue," it runs at top speed to the goal (Carr and Watson, 1908). Unfortunately, this means that if a corridor was shortened or lengthened the rats ran full tilt into the wall (Carr and Watson, 1908). Analogously, Dennis (1932) reports that when dead ends on an elevated maze are changed, rats run right off the edge, sometimes barely catching themselves with their hind claws.

### 2.1.8 Path integration

**Reviewed Result EX.57** *Gerbils can directly return to their starting point after an apparent random walk (Mittelstaedt and Mittelstaedt, 1980; Etienne* et al.*, 1986; Etienne, 1987; Etienne* et al.*, 1988).*

Mittelstaedt and Mittelstaedt (1980) showed that a female gerbil searching for a missing pup via an apparent random walk could execute a straight-line return to the nest once the pup was found. The experiment was performed in the dark to rule out visual homing. Displacement of the animal during its search at speeds below the vestibular detection threshold caused the return path to be offset by a comparable amount, eliminating the possibility that auditory or olfactory cues guided the trajectory.

Etienne (1987) showed similarly that golden hamsters trained to find food at the center of a circular arena used path integration to return to the nest. When the environment was rotated 90° or 180° while the animal was at the center of the arena, the animal returned to where the nest had been originally, ignoring the rotation of the arena (Etienne *et al.*, 1986; Etienne, 1987; Etienne *et al.*, 1988). The fact that animals are going to the center of the arena, only implies angular integration; linear integration is not necessary if the animals always go to the center (B. McNaughton, personal communication). However, Etienne does note that the hamsters could return accurately from locations other than the center (cited as "unpub. res." in Etienne, 1987), but does not present the data in detail.

**Reviewed Result EX.58** *Path integration shows systematic errors (Seguinot* et al.*, 1993, see also Maurer and Seguinot, 1995).*

Müller and Wehner (1988) examined path integration in desert ants (*cataglyphis fortis*) and found that they made systematic errors. Following on this work, Seguinot *et al.* (1993) tested hamsters on simple 1–5 stage paths and measured the direction in which they started their journey home. Seguinot *et al.* found systematic errors depending on specifics of the path taken. Other species make similar errors (Maurer and Seguinot, 1995).

**Reviewed Result EX.59** *Rats can show path integration abilities even after hippocampal lesions (Alyan* et al.*, 1997).*

Alyan *et al.* (1997) tested rats in two tasks — (a) rats were lured via a circuitous route to a location in the arena from which they were allowed to return back to the nest. They went directly to the nest. In probe trials, animals were lured to the center. A rotation of the arena (with no corresponding rotation of the animal) ruled out intramaze cues driving the return journey.[4], and (b) rats were trained to take an L-shaped path and then the path was blocked, forcing the rats to try another route back. Rats started back along the shortest path, turning directly toward the other end of the L. Both tasks were performed in the dark to rule out visual homing. In both cases there were no significant differences between normals and lesioned animals, implying that even the lesioned animals had intact path integrative abilities.

### 2.1.9 Local landmarks and exploration

**Reviewed Result EX.60** *Animals return to "home bases" throughout exploration*

From the earliest examinations of exploration (e.g. Chance and Mead, 1955), it was clear that animals spent large portions of their time in specific locations in an environment as they explore that environment. These locations are called *home bases* (Chance and Mead, 1955). A similar behavior has been reported by Leonard and McNaughton (1990), who report that animals begin exploration by making small excursions from the initial entry point. More recently, Golani *et al.* (Eilam and Golani, 1989; Golani *et al.*, 1993) have shown that not only do animals spend more time at these home bases, but they also visit the home base more than any other site in the environment. In addition, animals rear more and spend more time grooming themselves at these home bases then elsewhere in the environment (Eilam and Golani, 1989; Golani *et al.*, 1993).

**Reviewed Result EX.61** *Other exploration results (Poucet* et al.*, 1986; Thinus-Blanc* et al.*, 1987; Thinus-Blanc* et al.*, 1991; Thinus-Blanc* et al.*, 1992).*

[4] Alyan *et al.* (1997) is a published in abstract form
only and only describes paths that lead to the center which would imply that the task could be solved using angular integration only, however, the actual experiment performed included paths to other points in the environment (S. Alyan, personal communication; B. McNaughton, personal communication), which implies that the task does indeed require both linear and angular integration and can be considered a true path integration task.

The most extensive work done on determining the role of local landmarks and exploration is that of Poucet, Thinus-Blanc, and their colleagues (Poucet *et al.*, 1986; Thinus-Blanc *et al.*, 1987; Thinus-Blanc *et al.*, 1991; Thinus-Blanc *et al.*, 1992) examining hamsters in a cylinder with a striped cue card subtending 90° and three or four objects placed in a regular arrangement around the environment. They found that moving a landmark into or out of the array produced additional exploration. But they also found that expanding or contracting the square formed by the four landmarks did not produce additional exploration. Removing a landmark entirely also produced extensive exploration.

## 2.2 Neurophysiology

### 2.2.1 EEG

**Reviewed Result EX.62** *The hippocampus show two major modes of activity differentiated by characteristic EEG signals (Vanderwolf, 1971; Vanderwolf and Leung, 1983; Buzsáki, 1989; Stewart and Fox, 1990; Vanderwolf, 1990).*

During motion and REM sleep, in the presence of acetylcholine and serotonin, the hippocampal EEG shows a 7–12 Hz rhythm called *theta*; during rest and slow-wave sleep, in the absence of ACh and 5-HT, the hippocampal EEG shows irregular activity, called *LIA* (Large-amplitude Irregular Activity), characterized by short-duration *sharp waves*.

During LIA, hippocampal pyramidal cells tend to fire during sharp waves, but are then mostly all silent (Buzsáki, 1989; Ylinen *et al.*, 1995). During theta, each cell fires only when the animal is in the corresponding place field. Since each cell has a different place field, cells fire a few at a time (see Section 2.2.2-A).

Chrobak and Buzsáki (1994) have shown that during theta, cells in superficial layers of EC fire in a pattern correlated to the theta rhythm, while cells in deep layers of EC do not. In contrast, during LIA, cells in deep layers of EC fire in a pattern correlated to the sharp waves, while cells in superficial EC do not.

### 2.2.2 Place Cells

#### 2.2.2-A Hippocampal place cells

**Reviewed Result EX.63** *For spatial tasks, the first-order correlate of spikes fired by hippocampal CA3 and CA1 pyramidal cells is the location of the rat: each cell fires when the animal is in a specific place (called the "place field" of the cell) (O'Keefe and Dostrovsky, 1971; see Muller* et al.*, 1991a, for a review).*

The typical place field is that shown in Figure 2.18, a continuous, compact field with a single peak that falls off smoothly in all directions.

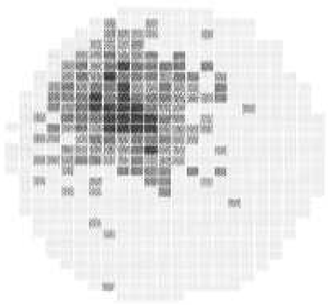


Figure 2.18: A typical place field. This plot shows a histogram of the firing rate of a single place cell from CA1 at each position in a small cylindrical arena (Figure 2.14). Darker colors indicate high firing, lighter colors, low firing. From Muller *et al.* (1991a), reprinted with permission of author and publisher.

Because *place cells* show such a clear correlate between firing rate and a spatial variable, many experiments have been done to explore how they react to environmental manipulations. As shown in Figure 2.18, a typical place field covers an area a few times the size of the rodent. Place fields have been recorded from

- small enclosed arenas (Figure 2.14) (Muller *et al.*, 1987; Muller and Kubie, 1987; Sharp *et al.*, 1990; Wilson and McNaughton, 1993; Knierim *et al.*, 1995; Markus *et al.*, 1995; O'Keefe and Burgess, 1996),
- open fields (Figure 2.1) (Gothard *et al.*, 1996b),
- radial mazes (Figure 2.16) (Miller and Best, 1980; Olton *et al.*, 1978; McNaughton *et al.*, 1983a; Pavlides and Winson, 1989; Thompson and Best, 1989; Thompson and Best, 1990; Markus *et al.*, 1994; Markus *et al.*, 1995; Mizumori *et al.*, 1996),
- T- and plus-mazes (Figure 2.15) (O'Keefe, 1976; O'Keefe and Conway, 1978; Pico *et al.*, 1985; O'Keefe and Speakman, 1987; Young *et al.*, 1994; Markus *et al.*, 1995),
- linear tracks (Figure 2.19) (O'Keefe and Recce, 1993; Markus *et al.*, 1994; Gothard *et al.*, 1996a),
- elevated tracks, such as triangular (Skaggs *et al.*, 1996; Skaggs and McNaughton, 1996) or rectangular loops (Mehta and McNaughton, 1996; Barnes *et al.*, 1997),

Place fields have also been recorded under a number of manipulations, including

- in the dark (O'Keefe, 1976; McNaughton *et al.*, 1989b; Quirk *et al.*, 1990; Markus *et al.*, 1994),
- during multiple and complex tasks (Eichenbaum *et al.*, 1987; Eichenbaum and Cohen, 1988; Wiener *et al.*, 1989; Cohen and Eichenbaum, 1993; Hampson *et al.*, 1993; Markus *et al.*, 1995),
- with environmental manipulations (O'Keefe and Conway, 1978; O'Keefe and Speakman, 1987; Kubie and Ranck, 1983; Muller and Kubie, 1987; Sharp *et al.*, 1990; Bostock *et al.*, 1991; Sharp *et al.*, 1995; Sharp, 1996b).
- in old rats (Barnes *et al.*, 1983; Markus *et al.*, 1994; Mizumori *et al.*, 1996; Barnes *et al.*, 1997; Shen *et al.*, 1996; Tanila *et al.*, 1997a; Tanila *et al.*, 1997b),
- after sensory deprivations (Hill and Best, 1981; Save *et al.*, 1996), lesions (Miller and Best, 1980; McNaughton *et al.*, 1989a; Shapiro *et al.*, 1989; Mizumori *et al.*, 1989; Mizumori *et al.*, 1994; Dudchenko and Taube, 1997), grafts (Shapiro *et al.*, 1989), and genetic manipulations (Rotenberg *et al.*, 1996; McHugh *et al.*, 1996).

**Reviewed Result EX.64** *Place cells are directional when an animal traverses repeated paths, but not when wandering randomly over open arenas.*

Figure 2.19: The linear track used by Gothard *et al.* (1996a). Two boxes sit on a linear track. The left box can be moved to any position along the track, varying the effective length of the track.

When the animal traverses repeated paths (such as back and forth on a linear track Gothard *et al.*, 1996a, along the arms of a radial arm maze, McNaughton *et al.*, 1983a, or in a fixed trajectory on an open field maze, Markus *et al.*, 1994), place fields tend to be directionally dependent. However, in open environments, place cells are not directionally dependent, i.e. their firing rate is independent of head direction (Muller *et al.*, 1994).

In order to examine this issue in detail, Markus *et al.* (1995) examined four cases:

1. food thrown into an open field with a random uniform distribution which produces non-repeated trajectories around the environment,
2. food placed at the four corners of a square (in sequence) on the same environment, which produces a highly repeated trajectory around that square,
3. food placed at random on a plus-maze,

4. food placed at the ends of the arms of a plus-maze.

This examines the question along two complementary axes: whether the environment produces limited trajectories (thin arms of the plus-maze vs. the more open space of the open field) and whether the food distribution is uniform or not. Markus *et al.* found that cells in the plus maze were very directional, but in the open field some cells were more directional than others. They found a higher proportion of directional cells in task 2 than in task 1, suggesting that it may be the trajectories taken that determines directionality.

**Reviewed Result EX.65** *Some place fields are crescent-shaped (Muller* et al.*, 1987); they hug the arena walls, see Figure 2.20.*

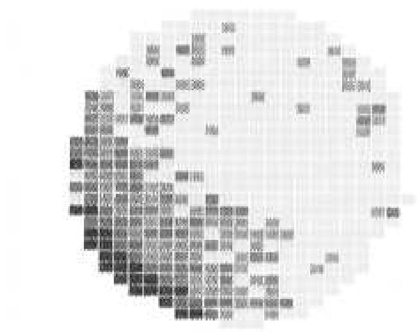


Figure 2.20: A crescent-shaped field that hugs the wall. From Muller *et al.* (1991a), reprinted with permission of author and publisher.

**Reviewed Result EX.66** *Some place cells show multiple subfields within a single environment, see Figure 2.21.*

Although many early reports of place fields suggested that cell had multiple subfields (e.g. O'Keefe and Conway, 1978; O'Keefe and Speakman, 1987), it should be noted that place cells recorded with a stereotrode (McNaughton *et al.*, 1983b) show fewer subfields than those recorded with a single-wire electrode, and those recorded with a tetrode (Wilson and McNaughton, 1993) show even fewer. However, even with tetrodes, some cells still show multiple fields.

**Reviewed Result EX.67** *Moving distal landmarks produces corresponding movements in place fields (Muller and Kubie, 1987; O'Keefe and Speakman, 1987; McNaughton* et al.*, 1994a; Knierim* et al.*, 1995).*

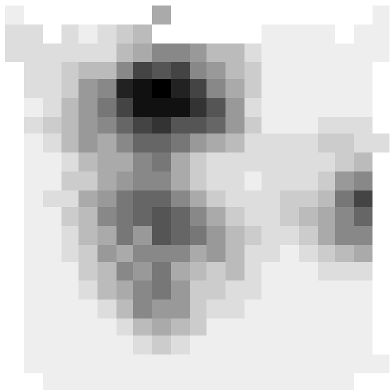


Figure 2.21: A place field with multiple subfields. Black areas indicate regions of high firing, grey low. Figure courtesy Bill Skaggs.

Many early experiments used T-, plus-, or radial mazes with cues available on the walls of the surrounding room (e.g. O'Keefe and Conway, 1978; Miller and Best, 1980; O'Keefe and Speakman, 1987; Shapiro *et al.*, 1989; Young *et al.*, 1994). Sometimes there were differing textures on each arm of the maze, but even if there weren't one would expect some sort of local intra-maze cues. The local and distal cues can then be separated by rotating the maze. Place cells generally followed the distal cues (Miller and Best, 1980; Shapiro *et al.*, 1989).

Similar results have been found by Cressant *et al.* (1997) in which they showed that place fields only rotate with landmarks inside a small circular arena if the landmarks are pushed all the way against the wall, i.e. if they are *orienting* or *distal* stimuli. If the landmarks (three colored cylinders) were clumped together and on the interior of the arena, the place fields were not tied to the landmarks, but if the landmarks were pushed to the walls of the arena, the place fields followed them. The task used by Cressant *et al.* (1997) was independent of the actual locations of the landmarks. In tasks in which the animals must attend to the landmarks in order to find reward, they can learn to do so and then the place fields are tied to the landmarks when the animal is near the landmarks (within 70 cm) (Gothard *et al.*, 1996b).

**Reviewed Result EX.68** *Place cells can be controlled by nonvisual landmarks (Hill and Best, 1981; Save* et al.*, 1996).*

Hill and Best (1981) examined place fields in blind and deaf rats on a 6-arm maze. They found that cells continued to show place fields, but most of the cells followed local cues when the maze was rotated.

Save *et al.* (1996) showed that even rats who had been blinded one week post-natally continued to show normal place fields as adults. Their place fields were sensitive to the rotation of prominent somatosensory landmarks. The landmarks were placed at the edge of a circular arena so in order to determine their location, the rats first went to the wall and then made a loop around the circumference of the arena. When the landmarks were rotated, the place fields also rotated.

**Reviewed Result EX.69** *Barriers produce changes in place fields that cross the barrier (Muller and Kubie, 1987).*

Muller and Kubie (1987) found that when a barrier was added to a cylinder, the firing rates of 9 out of 10 cells whose fields intersected the barrier diminished. The remaining cell increased its firing rate. Substituting a transparent barrier didn't affect the results. However, the effect only occurred if what was placed interfered with the rat's motion. They also experimented with a small heavy object that did not interfere with travel (the base of the barrier). If only this base was put into the place field, there was no effect on firing rate. Barriers outside the place field also had no effect.

**Reviewed Result EX.70** *Place cells continue to show place fields when landmarks are removed (Muller and Kubie, 1987; O'Keefe and Speakman, 1987, see also Pico* et al.*, 1985; O'Keefe and Conway, 1978).*

As mentioned above, place fields can change when landmarks are moved (EX.67). Place fields continue to show compact fields when visual landmarks are removed (Muller and Kubie, 1987; O'Keefe and Speakman, 1987). When the cue card was removed from the cylindrical environment shown in Figure 2.14, the place fields sometimes rotated around the center of the arena, but they did not change shape (Muller and Kubie, 1987), suggesting that the cue card serves mainly as an orienting stimulus. O'Keefe and Speakman (1987) trained rats on a plus maze and then removed the landmarks. Both O'Keefe and Conway (1978) and Pico *et al.* (1985) found that while removing some of the visual cues did not disrupt place fields, removing all of the cues did. The difference between the two former and two latter experiments is that in the former, the animals were first placed in the environment with the cues present.

**Reviewed Result EX.71** *Place cells continue to show compact fields in the dark (O'Keefe, 1976; McNaughton* et al.*, 1989b; Quirk* et al.*, 1990; Markus* et al.*, 1994).*

Quirk *et al.* (1990) report that when the animal is placed into the arena in the light and the lights are extinguished, the place fields rarely change. But when they do change, they continue to show the same shape and distance from the arena wall; it is only their orientation that varies.

Markus *et al.* (1994) report that more cells had fields in the light and that place fields were more reliable in the light than in the dark. Markus *et al.* found in particular that if the first trial in a session occurred in the light then cells were much more likely to be stable (measure by the correlation to the mean place field in the light) than when sessions started with a dark trial.

**Reviewed Result EX.72** *Place fields are seen on initial entry into an environment (Hill, 1978; Austin* et al.*, 1990; Wilson and McNaughton, 1994; Tanila* et al.*, 1997b).*

Hill (1978) reports that place cells show pretty clean fields the first time an animal passes through the field, but later studies have shown that fields can require 10–30 minutes (Wilson and McNaughton, 1993) or as much as 4 hours (Austin *et al.*, 1993) to "tune up". Although they may show initial place fields, the place fields become more reliable and more stable over that initial exploration period (Austin *et al.*, 1993; Wilson and McNaughton, 1993). Tanila *et al.* (1997b) report that some cells showed strong place fields in the first few minutes while others took more than 30 minutes to build up their fields, during which time the mean selectivity increased.

**Reviewed Result EX.73** *Place fields show interactions between consistent entry locations and external landmarks (Sharp* et al.*, 1990).*

Sharp *et al.* (1990) examined rats in a small cylindrical arena with a cue card subtending 90° (see Figure 2.14). During training the animals were not disoriented before entering the arena, and they always entered at the same location (the northwest corner). Sharp *et al.* then tested the animals with (1) the original configuration, (2) the cue card on the opposite side from its original location, (3) both cue cards from the previous configurations, and (4) two cue cards rotated by ±30°. They tested each of these cases in two conditions: with the animal entering the arena at the northwest corner and with the animal entering the arena at the southeast corner. See Figure 2.22.

When a second cue card was added opposite the first (Figure 2.22c), most place fields did not double.[5] Instead, the cells continued to fire at their original locations. However, if the rat was introduced into the double-card environment at the *southeast* corner (Figure 2.22f), the place fields rotated by 180°. But rotation did not occur in probe trials with the original configuration and a southeast entry point (Figure 2.22b). When tested with cue cards rotated by ±30°, Sharp *et al.* observed that place field locations were controlled by an interaction of the choice of entry point with the cue card positions (Figure 2.22g and 2.22h.)

**Reviewed Result EX.74** *The specific timing of spikes fired by place cells as an animal traverses a place field precesses along the theta rhythm (O'Keefe and Recce, 1993; Skaggs* et al.*, 1996).*

Recent work by O'Keefe and Recce (1993) and Skaggs *et al.* (1996) has shown that the timing of action potentials fired by place cells has an interesting interaction with the theta rhythm: when an animal enters a place field, the cell generally fires its spikes approx 90°–120° after the peak of the theta cycle (Skaggs *et al.*, 1996). (Skaggs *et al.* define the peak of the theta cycle as the phase of "maximal pyramidal cell population activity".) As

[5] Five of the 18 cells recorded by Sharp *et al.* changed their place fields over the various recording sessions, including three that showed doubled (symmetric) place fields at one time or another.

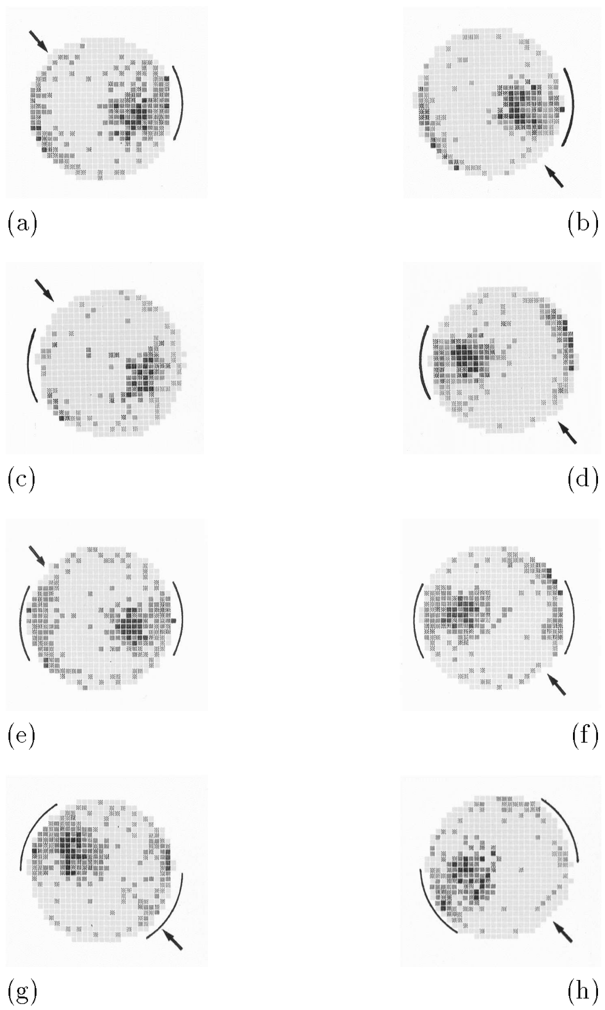



Figure 2.22: Results from Sharp *et al.* (1990), reprinted with permission of author and publisher. Arrow indicates entry point; arcs indicate cue cards. Darker areas show regions of higher firing. Arena was a cylinder; grey areas outside the circle are artifacts of the scanning process used to replicate the figures.

the animal then moves through the place field, the cell generally fires its spikes earlier and earlier in the theta cycle. Although this effect is most easily seen on linear tracks, it also occurs in two dimensions (Skaggs *et al.*, 1996).

**Reviewed Result EX.75** *Place fields shift with experience (Mehta* et al.*, 1996; Mehta and McNaughton, 1996).*

Mehta and McNaughton (1996) recently showed that the place fields of cells shift backwards along a much-repeated path as the animal runs along that path. They trained a rat to run in a loop on a rectangular elevated maze and recorded place cells in CA1. They found that the place fields shifted backwards over the course of a single session. However, they also showed that while the area covered by the place field shifted by almost 50% (approx. 5–7 cm), the center of mass of the field actually shifted very little (20%, approx. 2–3 cm).

**Reviewed Result EX.76** *Place cells disrupted by fimbria/fornix lesions are restored by septal grafts (Shapiro* et al.*, 1989).*

Shapiro *et al.* (1989) studied place cells in normal rats, in rats with lesions to the fimbria/fornix (thus lacking ACh) and in rats with fetal basal forebrain tissue grafted in (with ACh restored). They recorded from place cells in a radial arm maze and then examined the effect of rotating the maze. Shapiro *et al.* found that place cells in fimbria/fornix-lesioned rats followed local cues, but were partially influenced by distal cues, as evidenced by the lower reliability of place fields in lesioned rats during rotated trials (but not during non-rotated trials). In rats with grafts, place fields followed distal cues, as they do in normal rats (Miller and Best, 1980; Shapiro *et al.*, 1989).

**Reviewed Result EX.77** *Septal inactivation produces significant reduction in location specific firing of CA3 but not CA1 cells on a radial maze (Mizumori* et al.*, 1989).*

Mizumori *et al.* (1989) tested medial septal inactivation (with tetracaine) while recording from place cells in both CA3 and CA1. They found that the CA3 fields were disrupted but the CA1 were not. They also found a significant increase in working memory errors on the radial maze during the septal inactivation.

**Reviewed Result EX.78** *Place cells are sensitive to more than place.*

The firing rates of hippocampal pyramidal ("place") cells are also correlated with information other than location of the animal: speed, direction, and turning angle (McNaughton *et al.*, 1983a; Wiener *et al.*, 1989; Markus *et al.*, 1994), texture underfoot (Young *et al.*, 1994), odor (Eichenbaum *et al.*, 1987; Eichenbaum and Cohen, 1988; Otto and Eichenbaum, 1992; Cohen and Eichenbaum, 1993), task (Markus *et al.*, 1995), and stage of task (Eichenbaum *et al.*, 1987; Otto and Eichenbaum, 1992; Hampson *et al.*, 1993).

In these cases, if a cell has a place field under one condition, it may or may not show a place field under the other, and if two cells both show place fields under both conditions, then the spatial relationships between them may change drastically from one condition to the other. Thus many of these effects can be said to be second-order effects because they determine how the place fields change. Essentially, a cell's place field (in fact whether it has a place field at all) is independent from one condition to another.

**Reviewed Result EX.79** *The spatial relationship between place cells changes from environment to environment (O'Keefe and Conway, 1978; Kubie and Ranck, 1983; Thompson and Best, 1989; Muller and Kubie, 1987).*

Kubie and Ranck (1983) recorded from place cells in rats in three different environments (a radial maze, an operant-conditioning chamber, and the animal's home cage). They found that when cells have fields in multiple environments, the topology of the fields is unrelated. Thompson and Best (1989) also recorded from place cells in three environments (a radial maze, a circular drum, and a small box). They first identified cells during barbiturate anesthesia (since hippocampal cells are much more active under barbiturate anesthesia) and then measured them in each of the environments. They found that the numbers of cells that showed place fields in each environment was compatible with the hypothesis that 10–25% of the total CA1 population show a place field in any environment[6] and that, for any specific environment, the subpopulation of cells with place fields is random and independent of that seen in any other environment.

An important point to make is that all of these experiments show that when the animal is returned to an environment, the place fields return to the representation encoding that environment. In other words, place fields are stable from session to session (Muller *et al.*, 1987). Thompson and Best (1990) report recording a stable place field for months.

**Reviewed Result EX.80** *Changing the cue card from white to black eventually produces change in place field topology (Bostock* et al.*, 1991).*

Bostock *et al.* (1991) recorded from place cells in a cylindrical arena (Figure 2.14), first with a white cue card, and then with a black cue card. Sometimes the place fields were similar and sometimes they were unrelated (as if the two situations were encoded as different environments). However, once a place field changed when the cue card was changed then all other place fields recorded subsequently from the same animal changed with the cue cards but when the white cue card was returned, the place field returned to its original configuration.

**Reviewed Result EX.81** *A single environmental manipulation half-way through a recording session eventually produces a change in place field topology (Sharp* et al.*, 1995).*

[6] This number must be taken as a ballpark number only, determining the exact proportions is a complex experiment requiring understanding issues of rescaling, multiple reference frames within a single environment, etc.

Sharp *et al.* (1995) saw a similar effect in their experiment examining the effects of vestibular cues on place cells. In their experiment, they rotated the environment (or part of the environment) 20 minutes into a 40 minute recording session. They found that sometimes the fields rotated (or not, depending on the manipulation) with the manipulation, but in some sessions (independent of the manipulation), the field changed location dramatically (or disappeared completely) when the manipulation occurred. The probability of a radical change/disappearance occurring increased over the sessions.

**Reviewed Result EX.82** *Old animals do not show place fields transitions with environmental manipulations that produce novel transitions in young animals (Tanila* et al.*, 1997b).*

Using a plus maze and rotating it relative to the external world, Tanila *et al.* (1997b) report that old and young animals react differently to this manipulation. Young rats tend to create a new map, while old rats do not. Both animals had extensive experience with the original environment.

**Reviewed Result EX.83** *Older animals show unstable representations of the environment (Barnes* et al.*, 1997).*

Barnes *et al.* (1997) allowed animals to run a rectangular loop track for 25 minutes, while a few dozen cells were simultaneously recorded. The animals were then removed from the track for one hour, after which they were returned to the track for another 25 minutes. For young animals, the ensemble correlation between the place fields seen during the first and second 25 minute experiences showed a unimodal distribution (around 0.7, indicative of a similar representation between experiences). But for older animals, the ensemble correlation was bimodal (around 0, indicative of a complete remapping, and around 0.7, indicative of a similar representation between experiences). Within a single 25 minute run, the ensemble correlation (taken between two halves of the run) was always high (around 0.8).

**Reviewed Result EX.84** *Mice that have been genetically engineered to have LTP deficits in CA1 show instability in their place fields between sessions (Rotenberg* et al.*, 1996).*

Rotenberg *et al.* (1996) measured place fields in mice that were genetically engineered to be deficient in LTP when stimulation was in the 5–10 Hz range. These mice show severe spatial deficits (Bach *et al.*, 1995, EX.32). Like Barnes *et al.* (1997, see previous Reviewed Result), Rotenberg *et al.* found that place fields in these mice were stable within a single session but unstable across sessions.

**Reviewed Result EX.85** *Place cells are sensitive to task within a single environment (Markus* et al.*, 1995).*

When rats were trained to search for food on a large elevated platform either randomly or at the corners of a diamond, different subsets of place cells were active for each task, and some cells that were active for both tasks had different place fields, as if the animals were encoding the tasks as different environments (Markus *et al.*, 1995). When the animal switched between these two tasks, the change between representations was rapid, suggesting a shift in a property encompassing the entire system.

**Reviewed Result EX.86** *Place cells are sensitive to odor cues in certain tasks (Eichenbaum* et al.*, 1987).*

In more complex tasks than simply finding food scattered on the floor of the arena, place cells do not always fire when the animal is in the place field. Eichenbaum *et al.* (Eichenbaum *et al.*, 1987; Eichenbaum and Cohen, 1988; Cohen and Eichenbaum, 1993) tested rats in an odor-detection task and found that some place cells were dependent on whether the rat was going to the reward location or not. In a similar task, Eichenbaum *et al.* (Otto and Eichenbaum, 1992; Cohen and Eichenbaum, 1993) found that cells responded when two odors matched in a delayed match-to-sample task, but not when they didn't. Animals in the Eichenbaum *et al.* (1987) task had been trained to go to a reward location given one set of odors ($S^+$) but not given another ($S^-$). This meant that the animals were learning to take two different paths to reward depending on the whether the odor was in the $S^+$ or $S^-$ set. Eichenbaum *et al.* (1987) report that hippocampal pyramidal cells ("place" cells) are odor sensitive. However, they are not sensitive to different odors, they are really sensitive to the different reward conditions, indicated in this experiment by different *odor sets.* Equivalently, the Otto and Eichenbaum (1992) experiment showed place cells sensitive to $S^+$ and $S^-$ odor-pairs.

**Reviewed Result EX.87** *Place cells are sensitive to stages of complex tasks (Eichenbaum* et al.*, 1987; Eichenbaum and Cohen, 1988; Cohen and Eichenbaum, 1993; Hampson* et al.*, 1993; Gothard* et al.*, 1996b).*

In addition to sensitivity to reward availability, Eichenbaum *et al.* (Eichenbaum *et al.*, 1987; Eichenbaum and Cohen, 1988; Cohen and Eichenbaum, 1993) report that some cells are correlated with location during different stages of the task. For example, one cell might show a place field when the animal is approaching a sniff port to sample the odor (to determine whether it is an $S^+$ or an $S^-$ odor), but not when the animal leaves the sniff port to either go to the reward location or back to the starting point.

Hampson *et al.* (1993) trained rats to do a multiple lever-pressing delayed-match-to-sample task and found place cells dependent on whether a lever had already been pressed or not. They also found cells sensitive to which lever had been pressed (and therefore which lever had to be pressed to receive reward). They report that when the animal made mistakes, the place code indicated the incorrect lever.

Gothard *et al.* (1996b) also report that place cells are sensitive to components (or stages) of a open-field landmark task (see Section 2.1.1): they trained animals to leave a start box when a door was opened and proceed across an open floor to a pair of landmarks. Forming a triangle with the pair of landmarks was a small food reward. After the animal found the goal location, it was to return to the start box for another trial. They found cells correlated (within this task) to leaving the start box, to passing through the open room, to the goal location itself, and returning to the start box.

**Reviewed Result EX.88** *Hippocampal representations are replayed during sleep (Pavlides and Winson, 1989; Wilson and McNaughton, 1994; Skaggs and McNaughton, 1996; Kudrimoti* et al.*, 1996).*

Pavlides and Winson (1989) showed that cells with recently visited place fields were more active during REM sleep than other cells whose fields had not been recently visited. Wilson and McNaughton (1994) showed that during slow wave sleep (SWS) cells that showed correlated firing during a session in an environment (because their place fields overlapped) also showed a stronger correlation during sleep immediately after the session.

Skaggs and McNaughton (1996) explicitly examined the temporal nature of replay during sharp waves in slow wave sleep. They defined the *temporal bias* $B_{ij}$ between two cells $i$ and $j$ to be the difference between the integrated cross-correlation for the 200 ms after each spike of cell $j$ and the integrated cross-correlation for the 200 ms before each spike of cell $j$. Thus if cell $i$ generally fires after cell $j$ rather than before, $B_{ij}$ will be greater than 0. They report that the temporal bias during sleep after running on a linear track is strongly correlated with the temporal bias seen while the animal was running on the track.

Kudrimoti *et al.* (1996) measured the increase in correlation of cells whose place fields overlapped during a recent exploration of an environment during three blocks of LIA in SWS. Blocks 1 and 2 occurred before the first bout of REM sleep and block 3 occurred soon after it. They found that the correlation in block 1 was higher than the correlation in block 2, but that the correlation during block 3 was larger than block 2 and comparable to block 1.

**Reviewed Result EX.89** *There are differences between place fields in dorsal and ventral hippocampus (Jung* et al.*, 1994; Poucet* et al.*, 1994).*

All of the cells previously described in this section were recorded from dorsal hippocampus. Poucet *et al.* (1994) did not find any differences between dorsal and ventral place cells, but their methodology pre-selects cells that show place fields that look similar to Figure 2.18, so their result can only be taken as an existence proof of place cells in ventral hippocampus. Jung *et al.* (1994) found that fewer ventral hippocampal pyramidal cells had place fields and that the average spatial selectivity was lower for ventral cells compared to fields of dorsal cells.

### 2.2.2-B Extrahippocampal place cells.

**Reviewed Result EX.90** *Dentate gyrus cells also show place fields (Jung and McNaughton, 1993).*

Jung and McNaughton (1993) found that granule cells also showed clear place fields on a radial maze. Like CA3 and CA1 place fields, these cells showed a strong directionality on the radial maze.

**Reviewed Result EX.91** *Cells in medial entorhinal cortex show place fields (Barnes* et al.*, 1990; Mizumori* et al.*, 1992; Quirk* et al.*, 1992).*

Quirk *et al.* (1992) report that MEC place cells were larger and more noisy than hippocampal place fields, but they did show strong spatial signals reminiscent of hippocampal place fields. These fields rotated with a cue card, but continued to show place fields when the card was removed. These cells do not change their topology between two similar environments: a cylinder and a rectangle in both of which a cue card subtended 90° (see Figure 2.14). MEC place fields "stretched" topologically between the environments. All cells that had place fields in one environment had a corresponding place field in the other. These two environments produce dramatic changes in hippocampal place field topologies (Compare EX.79).

Mizumori *et al.* (1992) also report MEC place fields (recorded from a radial maze, Figure 2.16). They also found broad place fields but noted a weak directional component to the place fields.

**Reviewed Result EX.92** *Cells in subiculum show place fields (Barnes* et al.*, 1990; Muller* et al.*, 1991b; Sharp and Green, 1994; Sharp, 1996b).*

Like entorhinal place cells (EX.91), but unlike hippocampal place cells (EX.79), dorsal subicular place cells show similar place fields across different environments (Sharp, 1996b). Sharp tested dorsal subicular cells in four variations each of the cylinder and square environments (see Figure 2.14). She found that subicular cells always showed similar place fields (allowing for translation and rotation) between the two environments.

Unlike entorhinal or hippocampal cells, subicular cells do show a (weak) directional signal (Sharp and Green, 1994).

Phillips *et al.* (1996) report that ventral subicular cells do not show place fields in complex environments. It is not known whether there is a difference between dorsal and ventral subicular cells as there seems to be between dorsal and ventral hippocampal cells (Jung and McNaughton, 1993, EX.90)

**Reviewed Result EX.93** *Cells in parasubiculum show place fields (Taube, 1996).*

Taube (1996) report that 10.3% of parasubicular cells recorded showed place fields in a cylinder, however, because of the methodology used (pre-screening for spatial signals), this number should be taken as a lower bound only. Because they tested these cells in only one environment, it is not known whether parasubicular place cells show the same topology-preservation between environments (like medial entorhinal place cells, Quirk *et al.*, 1992, Reviewed Result EX.91, and like subicular place cells, Sharp, 1996b, EX.92, but unlike hippocampal place cells, Kubie and Ranck, 1983; Muller and Kubie, 1987; Thompson and Best, 1989, Reviewed Result EX.79). Taube notes that the fields in parasubiculum are larger and qualitatively more like subicular place fields than like hippocampal.

### 2.2.3 Head Direction Cells

**Reviewed Result EX.94** *Cells in postsubiculum (PoS) are tuned to the orientation of the animal in space (Ranck, 1984; Taube* et al.*, 1990a).*

Head direction cells in the postsubiculum (PoS, also known as dorsal presubiculum) were first described by Ranck (1984). In subsequent work, Taube *et al.* (1990a) characterized these cells as having triangular tuning curves: the firing rate drops off linearly from a peak at the preferred direction until it reaches a baseline value. Taube *et al.* (1990a) report that PoS cells typically have baseline-to-baseline tuning curve widths of 100°. However, these curves can also be modeled very closely by Gaussians with an average standard deviation of approximately 66° (Blair and Sharp, 1995; Zhang, 1996a).

**Reviewed Result EX.95** *Cells in the anterior dorsal nucleus of the thalamus (AD) are tuned to the orientation of the animal in space (Blair and Sharp, 1995; Knierim* et al.*, 1995; Taube, 1995).*

Early reports of ATN cells (Blair and Sharp, 1995; Knierim *et al.*, 1995; Taube, 1995), suggested that they showed very similar tuning curves to PoS cells, but more recent reports have suggested a number of differences (see EX.101 and EX.102).

**Reviewed Result EX.96** *Cells in the lateral mammillary nuclei (LMN) are tuned to the orientation of the animal in space (Leonhard* et al.*, 1996).*

PoS and ATN are both interconnected with the lateral mammillary nuclei (LMN) (van Groen and Wyss, 1990; Bentivoglio *et al.*, 1993) in which similar head direction cells have been found (Leonhard *et al.*, 1996).

**Reviewed Result EX.97** *Head direction tuning curves rotate in sync (Taube* et al.*, 1990b; Goodridge and Taube, 1995; Taube and Burton, 1995).*

Whenever multiple head direction cells have been recorded from ATN, PoS, or LMN the difference between their preferred directions is a constant across all recording sessions in all environments (Taube *et al.*, 1990b; Goodridge and Taube, 1995; Taube and Burton, 1995, Taube, personal communication).

**Reviewed Result EX.98** *PoS, LMN, and AD cells are sensitive to rotation of distal cues (Ranck, 1984; Taube* et al.*, 1990b; Taube, 1995; Goodridge and Taube, 1995; Knierim* et al.*, 1995; Leonhard* et al.*, 1996).*

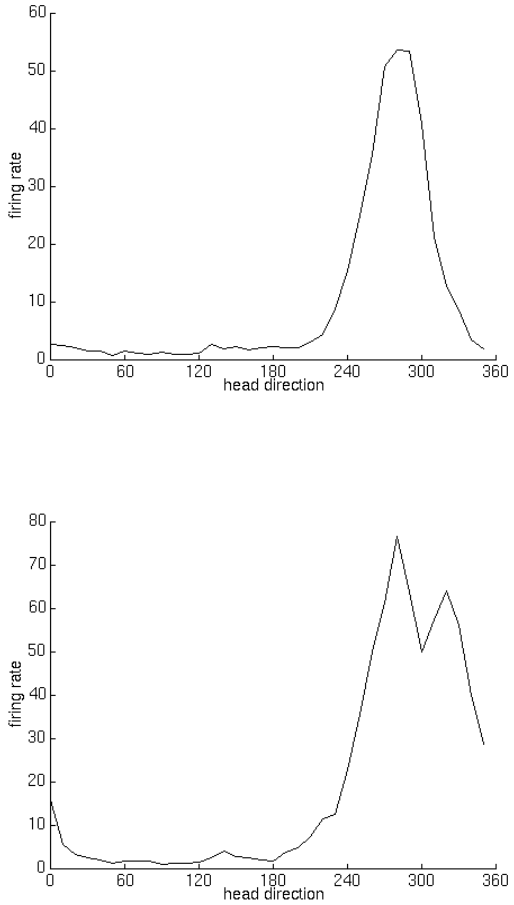



Figure 2.23: Sample head direction tuning curves from postsubiculum (top) and ATN (bottom). Data courtesy T. Blair and P. Sharp.

In all of these cases, head direction cells were recorded from a small cylinder with a cue card subtending 90° (see Figure 2.14). When the animal was returned to the arena, the head direction tuning curve was oriented identically relative to the cue card (Taube *et al.*, 1990b; Taube, 1995; Goodridge and Taube, 1995; Leonhard *et al.*, 1996).

However, Knierim *et al.* (1995) found that when animals were disoriented prior to each experience, the tuning curves in anterior dorsal thalamus (AD nucleus of the ATN) no longer followed the cue card as strongly. Even in rats who had been trained without disorientation, after a number of sessions before which they had been disoriented, the cue card lost its effect on the initial tuning curve of the cell (Knierim *et al.*, 1995).

Taube and Burton (1995) also examined a similar issue: they allowed rats to explore a cylinder and rectangle connected by a U-shaped passage. After exploration, the animals were removed from the environment and the cue card in the cylinder rotated by 90°. When the animal was returned to the cylinder, the head direction tuning curve rotated proportionally. When the animal traversed the U-shaped passage, it encountered a cue-conflict situation: the internal cues (from the self-motion via the passage from the cylinder) and the external cues (from the cue card in the rectangular arena) were conflicting. Taube and Burton found that in some animals the tuning curves rotated and in others they didn't, but Taube and Burton found that the shift was generally constant within each animal, even though it differed between animals.

**Reviewed Result EX.99** *PoS, LMN, and AD cells show normal head direction tuning curves in the dark (see Taube* et al.*, 1996 for a review).*

PoS, AD head direction cells do not require visual input to show a strong directional signal (see Taube *et al.*, 1996 for a review). Although no controlled studies have been done, preliminary results suggest that LMN cells show normal HD tuning curves in the dark (Taube, personal communication). If a rat is brought into a maze in the dark, the head direction is carried over from its previous environment, most likely by vestibular input, but probably also by motor efferent information if available (Ranck, 1984; Goodridge and Taube, 1995; Taube *et al.*, 1996).

**Reviewed Result EX.100** *LMN cell activity is strongly correlated with angular velocity as well as direction (Leonhard* et al.*, 1996).*

McNaughton *et al.* (1991) report that some cells in parietal cortex are correlated with both head direction and angular velocity. Taube *et al.* (1990b) and Sharp (1996a) have also reported that some PoS cells are correlated with both as well. But these cells are rare, while most of the cells in LMN are strongly correlated with both head direction and angular velocity (Leonhard *et al.*, 1996).

**Reviewed Result EX.101** *While PoS head direction cells are best correlated with current head direction, AD cell activity is best correlated with head direction approximately 20–40 ms in the future (Blair and Sharp, 1995; Taube and Muller, 1995).*

ATN cell activity is best correlated not with current head direction, but with head direction approximately 20–40 ms in the future, while PoS head direction cells are best correlated with the animal's current (or recent) head direction (Blair and Sharp, 1995; Taube and Muller, 1995). Both Blair and Sharp (1995) and Taube and Muller (1995) report an optimal correlation of ATN activity with future head direction and PoS activity with current head direction, but both have recently revised their estimates (Taube *et al.*, 1996; Blair and Sharp, personal communication), suggesting that although ATN activity still anticipates future head directions (by 25 ms), PoS activity may lag the current head direction (by 15 ms).

Leonhard *et al.* also report that LMN cells are correlated with future direction (by as much as 83 ms), at least in one direction, but it is not yet clear whether this anticipation occurs for both clockwise and counter-clockwise turns. Because LMN activity is also correlated with angular velocity (see Reviewed Result EX.100, below), it might seem to anticipate future head direction with angular velocity in one direction but not the other.

**Reviewed Result EX.102** *AD head direction cells show multiple peaks in the tuning curve (Blair* et al.*, 1997).*

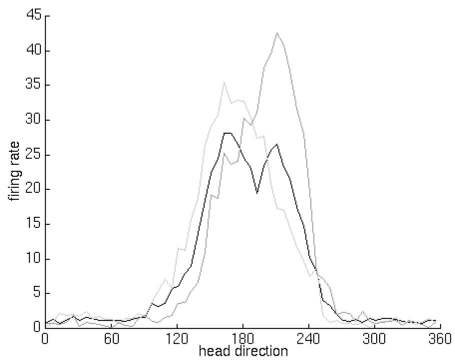



Figure 2.24: ATN cells at three different angular velocities. Note the two peaks during non-rotations (black line). With clockwise rotations (light grey line), the left peak rises (the cell is anticipating its preferred head direction when the animal is not yet at that direction). Likewise, with counter-clockwise rotations (medium grey line), the right peak increases. Data courtesy T. Blair and P. Sharp.

Blair *et al.* (1997) have recently reported that the tuning curves of ATN cells change during rotations.[7] Figure 2.24 shows an AD cell recorded during left-, right-, and non-rotations. During rotations one of the peaks increases while the other decreases.

[7] This experiment tested a prediction from the model presented in Chapter 5, originally published in Redish *et al.* (1996). As will be seen in Chapter 5, the experimental results strongly support the model.

#### 2.2.3-A Other related areas

Cells with firing rates correlated to head direction have also been found in the lateral dorsal nucleus of the thalamus (LDN: Mizumori and Williams, 1993), the parietal and cingulate cortices (McNaughton *et al.*, 1994b; Chen *et al.*, 1994b; Chen *et al.*, 1994a), and the dorsal striatum (the caudate nucleus: Wiener, 1993; Mizumori *et al.*, 1996), but these cells have different characteristics from cells in the three primary head direction areas (PoS, ATN, and LMN).

**Reviewed Result EX.103** *Cells in the lateral dorsal nucleus of the thalamus (LDN) show correlations to head direction on the radial maze (Mizumori* et al.*, 1992).*

Neurons in the lateral dorsal nucleus of the thalamus (LDN) are sensitive to head direction of the animal, but require exposure to the environment in the light in order to show directional sensitivity (Mizumori *et al.*, 1992). With a 30 second exposure, Mizumori *et al.* saw directionality only about 47% of the time, but with a 60 second exposure, they always saw directionality, and the directionality was consistent with prior exposures to the environment. Mizumori *et al.* also report that when the lights were turned off, directionality was maintained for approximately two to three minutes, after which it began to rotate regularly left or right.

**Reviewed Result EX.104** *Cells in the dorsal striatum (caudate nucleus) show correlations to head direction under specific conditions (Wiener, 1993; Mizumori and Cooper, 1995).*

Cells correlated with head direction have also been found in the dorsal striatum (Wiener, 1993) in a rectangular arena (see Figure 2.14), but in the task used to study these head direction cells, a number of other factors such as approach to a corner are also correlated with head direction. Any of these factors could produce head direction sensitivity. Wiener reports that the directional preferences of these cells rotated with rotation of the arena and were independent of extra-maze room cues and a light which did not rotate with the arena.

When Mizumori and Cooper (1995) recorded from caudate cells on the 8-arm radial maze, they found that most cells were tuned to single directions along pairs of arms, but also found that some cells changed their preferred directions suddenly and drastically within a single trial. Others were tuned, not to direction, but to a path along a pair of arms.

**Reviewed Result EX.105** *Cells in parietal and cingulate cortex show correlations to head direction on the radial maze (Chen, 1989; Chen, 1991; McNaughton* et al.*, 1994b; Chen* et al.*, 1994a; Chen* et al.*, 1994b).*

McNaughton *et al.* (1994b) report that the firing rates of cells recorded from rodent posterior parietal cortex while the animal traversed a radial maze were correlated with whether the animal was turning (some with left turns, some with right turns, some with straight-ahead movements), as well as whether the animal was progressing inwards (towards the center) or outwards (away from the center) of the maze. Chen *et al.* (1994b) found that approximately 5-10% of the cells were sensitive to direction on the radial arm maze, and that of those, half required a cue in order to show a directional tuning. These cells only fired if the cue was present. But if the cue was present and then removed, some of the cells continued to show a tuning to direction, as if the cell remembered the direction the cue had been. They tended to show a broader tuning to direction than postsubicular head direction cells (Chen *et al.*, 1994b). They also did not show the clean sensitivity to distal cues reported for ATN and PoS head direction cells (EX.98, above); some cells showed bimodal head direction tuning during cue-manipulation trials (Chen *et al.*, 1994a). Some cells also showed a strong modulation by behavior, as reported by McNaughton *et al.* (1994b).

#### 2.2.3-B Head direction lesion studies

**Reviewed Result EX.106** *Lesions to AD cause a disruption of directional selectivity in the PoS head direction population (Goodridge and Taube, 1994).*

**Reviewed Result EX.107** *After vestibular lesions, AD HD cells are no longer correlated with head direction (Stackman and Taube, 1997).*

Because Taube *et al.* only record from one neuron at a time, it is not clear whether these two results (EX.106 and EX.107) are showing a population that still represents a coherent representation of head direction that has been decoupled from the real world, or whether the representation itself is disrupted. Multi-unit recordings could differentiate these two possibilities.

**Reviewed Result EX.108** *Postsubiculum lesions do not disrupt the directional selectivity of AD cells (Goodridge and Taube, 1994; Taube* et al.*, 1996).*

**Reviewed Result EX.109** *After PoS lesions, AD cells are not sensitive to external sensory cues, they are only driven by vestibular cues (Taube, personal communication).*

**Reviewed Result EX.110** *Lesions of the lateral dorsal thalamus (LDN) do not affect the directional selectivity of PoS head direction cells (Golob and Taube, 1994).*

### 2.2.4 Nucleus accumbens

**Reviewed Result EX.111** *Nucleus accumbens cells are correlated to reward expectation, place, and movement parameters (Lavoie and Mizumori, 1994)*

Lavoie and Mizumori (1994) recorded from accumbens cells in a standard working memory task on the radial-arm maze. They found three major correlations to firing rate: place, reward, and movement. Some of the reward-correlated cells were actually correlated to reward expectation, i.e. they reduced their firing on finding reward, and some of them were also correlated with the magnitude of the reward.

### 2.2.5 Long-term potentiation (LTP)

The specifics of LTP are not within the scope of this thesis, and I refer the reader to reviews such as McNaughton and Morris (1987); Bliss and Lynch (1988); Landfield and Deadwyler (1988); Brown *et al.* (1991); McNaughton (1993); Barnes (1995); Malenka (1995); Abbott and Blum (1996) and Levy (1996). Instead, I will take note of two results which will eventually be used to inform the theory.

**Reviewed Result EX.112** *LTP is correlational (see McNaughton and Morris, 1987; Bliss and Lynch, 1988; Brown* et al.*, 1991; McNaughton, 1993; and Malenka, 1995 for reviews).*

As originally suggested by Hebb (1949), the connection between two neurons is potentiated only when spikes in the presynaptic neuron are combined with a depolarization of the postsynaptic neuron (as would happen when the postsynaptic neuron fires its own spike). Thus the increase in connection strength between two neurons $a$ and $b$ is proportional to the product of their firing rates $F_a$ and $F_b$.

**Reviewed Result EX.113** *LTP is preferentially asymmetric (see Abbott and Blum, 1996, and Levy, 1996 for reviews).*

If neuron $a$ fires shortly before neuron $b$ then it is the $a \rightarrow b$ connection that is potentiated, not the $b \rightarrow a$ connection.

## 2.3 Neuropharmacology

### 2.3.1 Acetylcholine (ACh)

Acetylcholine (ACh) enters the hippocampus from the septum via the fimbria (see Haas, 1983; Cooper *et al.*, 1986; Hasselmo and Bower, 1993; Mizumori *et al.*, 1992; Stewart and Fox, 1990; Gallagher and Colombo, 1995; Gallagher *et al.*, 1995 for reviews). It seems to be correlated to the presence of the theta rhythm (see Vanderwolf, 1990 for a review).

**Reviewed Result EX.114** *ACh influences the Schaffer collaterals more than the perforant path (Hasselmo and Schnell, 1994).*

Hasselmo and Schnell (1994) have shown that carbachol (a cholinergic agonist) infused into hippocampal slices reduces the size of the EPSP in CA1 produced by stimulating Schaffer collaterals by 90%, while only reducing that produced by stimulating the perforant path by 40%. Since the Schaffer collaterals are the same axons which form the CA3 excitatory feedback pathway, this suggests that ACh may shut off these recurrent connections. ACh also has an effect on learning: although it suppresses synaptic transmission, ACh enhances LTP in DG, CA1, and other structures (Hasselmo, 1995).

**Reviewed Result EX.115** *Rats show normal habituation even under scopolomine (Buhot and Naili, 1995).*

Buhot and Naili (1995) found that rats showed normal habituation (measured as a decrease in locomotor activity over time) to an environment under scopolomine (an ACh antagonist), although there was a general increase in locomotor activity compared to normals. However, rats injected with scopolomine could only recognize major changes to an environment. Adding objects to an empty environment produced exploration, but displacing one object or replacing one object with a dissimilar one did not. Normal animals examine the novel object extensively.

### 2.3.2 Norepinephrine (NE)

Norepinephrine[8] (NE) enters the hippocampus from the locus coeruleus (LC) (see Gray, 1982b; Haas, 1983; Cooper *et al.*, 1986; Sara and Segal, 1991; Aston-Jones *et al.*, 1991 for reviews). I will not review the effect of NE on cellular physiology and refer the reader to Hasselmo (1995) for a review.

**Reviewed Result EX.116** *NE infused into the hippocampus increases locomotor activity and exploratory nosepokes.*

Flicker and Geyer (1982) studied the effect of hippocampal microinfusions of NE on exploration. They found that increasing NE produced significant increases in both locomotor activity and exploratory nosepokes into holes in the environment. Plaznik *et al.* (1983) also found similar results with infusions of NE into the dentate gyrus of hippocampus.

### 2.3.3 Dopamine (DA)

Dopamine enters the nucleus accumbens from the ventral tegmental area (VTA) (Sesack and Pickel, 1990). There hippocampal and catecholaminergic terminals converge on single spines (Sesack and Pickel, 1990) I am not going to review the complex effects of dopamine on cellular physiology (I refer the reader to Hasselmo, 1995, for a review), nor will I review the anatomical structure of the basal ganglia or the involvement of dopamine therein. For this and other related issues, I refer the reader to Houk *et al.* (1995). Here, I will note two key results which will inform our understanding of the goal memory (see Section 3.2.5).

**Reviewed Result EX.117** *Dopamine neurons in the primate ventral tegmental area (VTA) respond to unexpected reward (Schultz* et al.*, 1995; Schultz, 1997).*

Schultz *et al.* (1995) report that dopamine neurons in VTA respond most vigorously to an unexpected reward, and, conversely, that they decrease their firing rates when an expected

[8] Norepinephrine is also known as noradrenaline.

reward is not given. Although this result is from primate and has not been replicated in the rodent, it will help inform our understanding of the role of dopamine in the goal memory (Section 3.2.5).

**Reviewed Result EX.118** *Dopamine levels in rodent nucleus accumbens increases gradually in a lever press task until the animal actually presses the lever and then there is a sharp phasic decrease in extracellular dopamine (Kiyatkin and Gratton, 1994).*

Kiyatkin and Gratton (1994) measured the level of dopamine in rodent accumbens during a lever press task. They found that extracellular dopamine increased gradually until the animal pressed the lever. At the lever press, they found a sharp phasic decrease in extracellular dopamine. When the animal was rewarded with more food, animals showed a larger phasic decrease, and when the animal was rewarded with less food, the animals showed a smaller phasic decrease.

These measurements are the inverse of those reported by Schultz *et al.* (1995) (Kiyatkin and Gratton: DA increase when expected reward exceeds actual reward; Schultz *et al.*: DA neurons fire with unexpected reward), but Kiyatkin and Gratton were measuring extracellular dopamine while Schultz *et al.* were measuring the firing rate of dopaminergic neurons.

### 2.3.4 Serotonin (5-HT)

There are two different types of theta rhythm: atropine-sensitive and atropine-resistant (Vanderwolf, 1990). Atropine-sensitive theta (hereafter ACh-theta) is dependent on inputs from the septal nuclei, while atropine-resistant theta (hereafter 5-HT-theta) is dependent on serotonergic inputs from the raphe nuclei. Blocking the input from the raphe nuclei whether by serotonergic blockers such as parachlorophenylalanine (which blocks the synthesis of 5-HT), 5,7-Dihydroxytryptamine (which kills 5-HT producing neurons), or by surgical lesions which remove input from the raphe nuclei, eliminates 5-HT-theta (Vanderwolf, 1990). Stimulation of the raphe nuclei produce theta like rhythms.

**Reviewed Result EX.119** *Rats with combined septal and raphe nuclei lesions require both septal and raphe grafts in order to see water maze improvements (Bjorklund* et al.*, 1990).*

**Reviewed Result EX.120** *CP-93,129 (a 5-HT agonist) produces a drastic decrease in exploratory behavior (Buhot and Naili, 1995).*

**Reviewed Result EX.121** *8-OH-DPAT (a 5-HT agonist) impairs water maze performance, but increases exploratory behavior (Buhot and Naili, 1995).*

Buhot and Naili (1995) report data that CP-93,129 (a 5-HT agonist) produced a drastic decrease in exploration early in the experiment. The animals seemed to treat the environment as familiar from the moment they entered it. This implies that CP-93,129 may impair the recognition of a novel environment. In agreement with this, Plaznik *et al.* (1983) report that serotonin injections inhibited exploratory behavior in an open-field test.

Although rats injected with CP-93,129 were able to recognize changes to objects and spent more time with novel or repositioned objects, they were significantly impaired relative to normals. 8-OH-DPAT (also a 5-HT agonist) impaired water maze performance, but the animals showed more extended exploratory behavior.

There are many different types of 5-HT receptors (Cooper *et al.*, 1986; Buhot, 1997) and each may play a different role. For example, CP-93,129 is a 5-HT1A agonist, stimulating one receptor subtype, while 8-OH-DPAT is a 5-HT1B agonist, stimulating another. It is not clear at this point what the role is of each of these subtypes, or the role serotonin plays in navigation. Further studies are clearly needed, particularly those that might combine infusions of 5-HT agonists and antagonists with multi-cell recordings which might elucidate the effect of 5-HT on hippocampal place cells.

## 2.4 Anatomy

The anatomy of the component structures involved in all of the experiments described above is extremely complex and not completely known. In addition, the relation of anatomical data to the theory may be complex: the existence of an anatomical connection does not imply that it must be required for the theory, although the lack of an anatomical connection can be troubling for a theoretical hypothesis. I will review the anatomy of each structure as it is required by the theory.

## 2.5 Contributions: Open-field navigation

I have also contributed directly to the experimental rodent navigation literature. Because this is a thesis and I need to explicitly show my contributions, I have separated this experiment and the next one from the previous EX.121 reviewed results.

We[9] have also studied open field navigation in tasks similar to those studied by Collett *et al.* (1986), Biegler and Morris (1993; 1996), and Gothard *et al.* (1996a). Figure 2.25 shows a gerbil searching for food in one of our tasks.

**Methods** We used the same training technique as Collett *et al.* (1986) with the following exceptions: We used both male and female gerbils. The environment was cue-rich; no attempt was made to control the external landmarks in the surrounding room. The gerbils were started from one of eight locations around the arena and they were always started

[9] This work was done in collaboration with Sofyan Alyan, David Banks, Steven J. C. Gaulin, Michael Pahn, Chris Reiber-Milberg, Lisa Saksida, and David S. Touretzky and has been reported in abstract form (Saksida *et al.*, 1995).

Figure 2.25: A close-up view of the arena used in our lab to study open-field navigation. Small white cylinders arranged in an array form the local landmarks. Cylinders are translated from trial to trial within the arena, but the array configuration is not changed. Orientation of the array is also kept constant. The arena was open to the outside room and many distal cues were available. The floor of the arena was covered in woodchips. Animals are trained to find food at a constant relation to the landmark configuration and are then tested with the landmark configuration in a novel location and no food.

facing the wall. The floor of the arena was covered with wood chips instead of gravel, but we also used a single sunflower seed as the food reward.

Animals were given 15 training trials per day (in five blocks of three) until they reliably dug for and found the food. If a gerbil did not eat the seed within two minutes of the beginning of a training trial, it was shown the food and allowed to eat the seed before being returned to its home cage. Gerbils were maintained at 85% ad lib feeding weight throughout the experiment. Water was always available. Early training trials began with the food placed in a dish at the goal location. Over the subsequent trials, the dish was removed, and the food buried deeper and deeper until the food was on the tiled floor under three to four inches of wood chips. In agreement with Collett *et al.* (1986), we found training took approximately 150 trials.

After training was complete, probe trials were intermixed with additional training trials in a (TTT)P(TTT)(TTT)P(TTT) sequence. The order of the probes was randomly determined. Probes lasted for 60 seconds, did not include food reward, and included the landmark array centered at a novel position. The wood chips were raked thoroughly between each probe trial.

The animals' position during all training and probe trials was continuously recorded at 18 frames/second by a video tracking system (Poly-Track Video Tracking System, San Diego Instruments).

**Analysis methods** The amount of time the gerbil spends in a particular spot is assumed to reflect its hypotheses about the goal location; each animal is presumed to divide its search time among a set of *goal predictions*. We modeled the distribution of search time spent by the gerbils as a mixture of 2D Gaussian generators: each position recorded by the tracker is assumed to have been generated by one of the goal predictions, and the distribution around each prediction is assumed to show a Gaussian fall-off.

These statistical tests require independent samples. However, when animal positions are sampled at 18 frames/second, there is a high correlation between the coordinates observed in a one frame and the frames that precede or follow it, due to the limited speed at which the animal moves. If we wait long enough between measurements (or use only every $n$th measurement in the series), the animal could in principle be anywhere in the arena, and the samples will be independent. In order to determine the minimum necessary subsampling rate, we calculated autocorrelation values for the time series of $x$ and $y$ coordinates, independently. The autocorrelation was found to become insignificant after approximately 100 samples (approximately 5.5 secs). We therefore subsampled all data used in subsequent analyses at a ratio of 1:100.

The EM (*Expectation-Maximization*) algorithm is a general-purpose algorithm for fitting general mixtures of models to data (Dempster *et al.*, 1977). We used the EM algorithm to fit mixtures of Gaussians here. For stability reasons, the variances of all the Gaussians was assumed to be equal, that is, there was a single parameter $\sigma$ which was fit by the EM algorithm.

We found in practice that the data from our experiment was very noisy. We therefore added a variation to the EM algorithm to allow it to throw out outliers. This thresholding is similar to the *Shortest-half* estimator (Shorth, Encyclopedia of Statistical Sciences, 1988), which is a very reliable estimator of the mean and is very robust against noise. In the Expectation step, any point which contributed less than $10^{-5}$ to all current generators was not included in calculations for the following Maximization step. Outliers were re-included in the following Expectation step, so a data sample could be an outlier in one iteration and not in another.

The quality of the fit of the 2D Gaussians to the data sample can be measured by log likelihood (Dempster *et al.*, 1977). We determined whether $K'$ Gaussians produced a significantly better fit than $K$ Gaussians by comparing the log likelihoods. If the increase in log likelihood was more significant than $P < 0.05$, we used $K'$ Gaussians instead of $K$.

In order not to fit the noise, we discounted Gaussians with weights less than 0.1 (i.e. those that accounted for less than 10% of the data.) In practice, a few Gaussians had weights on the order of 0.4–0.9 and the rest had weights $\ll 0.1$.

In order to differentiate theories, it would be useful to know the probability that the actual data sample could have been seen if the hypothesis were correct. We therefore determined the 95% confidence interval of the means of the Gaussians fit by the EM algorithm. Hypotheses that predict a goal location outside of the 95% confidence interval can be rejected. The confidence interval can be found by the *bootstrap algorithm* (Efron, 1982).

**Display details.** Figures 2.26–2.33 show data and analyses from our open field navigation experiment. Filled (grey) circles indicate landmark locations. Dark empty circles indicate predictions of the search theories (such as *vector voting* or *center of mass*, see Chapter 4). Light grey circles with a cross inside them indicate best fit Gaussians: center of the cross indicates mean; size indicates 95% confidence interval.

### 2.5.1 Two landmark experiments

We trained the animals to find food at the midpoint of two landmarks, placed 40 cm apart.

#### 2.5.1-A T2LM

**Reviewed Result EX.122** *Gerbils can learn to search at a single location relative to local landmarks even in cue-rich environments.*

Well-trained animals searched at a single location, with the goal position well within the 95% confidence interval for the single Gaussian, see Figure 2.26.

#### 2.5.1-B T2M1

**Reviewed Result EX.123** *In cue-rich environments, gerbils trained to search at the midpoint of a pair of landmarks, search at two points when faced with a single landmark.*

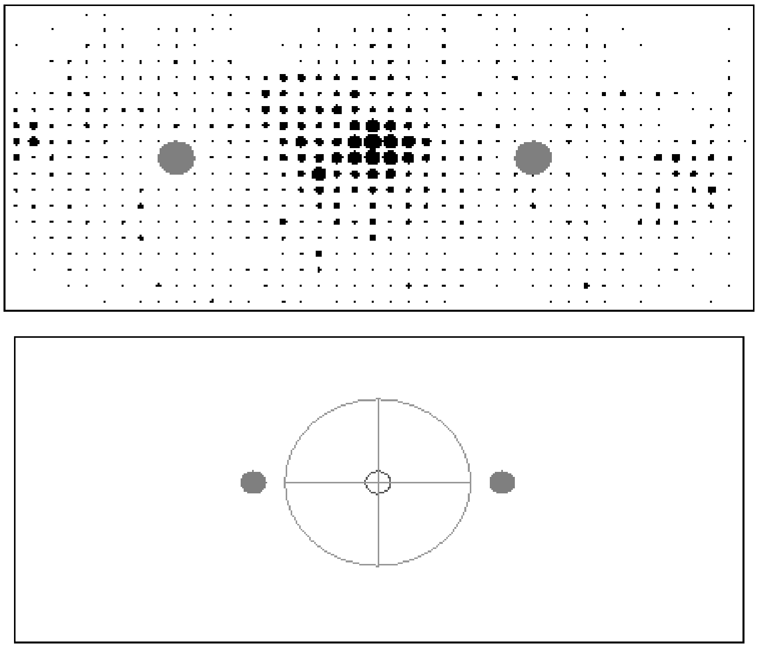


Figure 2.26: Two landmarks. (top) Histogram of search time of gerbils. (bottom) Analysis of search time showing best fit by Gaussians.

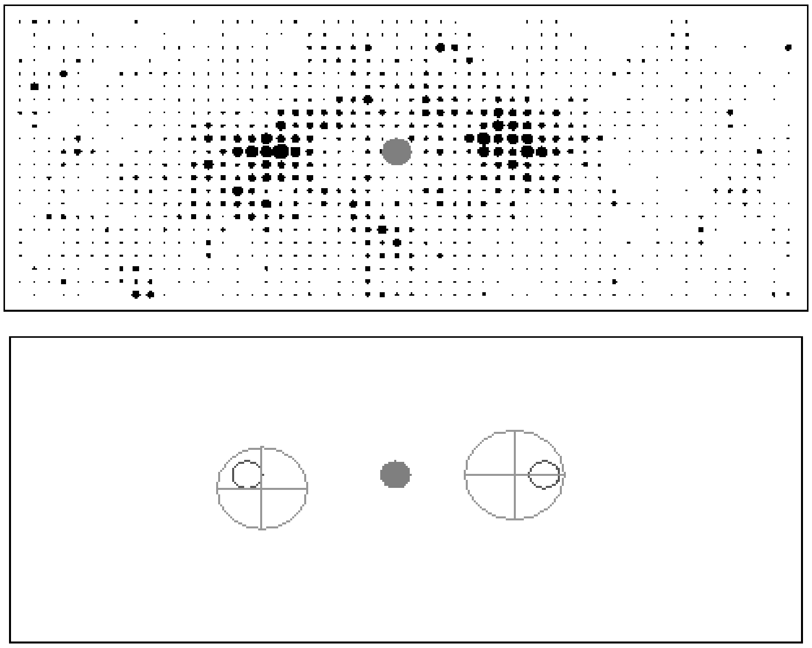


Figure 2.27: Tested with one landmark, after being trained with two. (top) Histogram of search time of gerbils. (bottom) Analysis of search time showing best fit by Gaussians.

With two landmarks, the array can be manipulated to provide ambiguous or inconsistent cues. For example one landmark can be removed (see Figure 2.27). In this case, Collett *et al.* report that the gerbils searched alternately in two locations, each at the correct distance and bearing from one of the landmarks they had observed during training. We found that when faced with this transformed array, the animals usually did not dig at any location, but they did go to specific locations, rear and then return to the wall. They repeated this sequence throughout the 60 second probe trial. Although it was difficult to determine whether the locations they went to were the expected locations while observing the gerbils directly, the peaks show up in the histograms because the gerbils regularly returned to these locations, and because they reared at these locations, spending a significant proportion of their time there.

This histogram is best modeled by two Gaussians each accounting for 50% of the data, with the expected locations well within the 95% confidence intervals.

### 2.5.1-C T2X2

**Reviewed Result EX.124** *In cue-rich environments, gerbils trained to search at the mid-point of a pair of landmarks search at two interior points when faced with landmarks more separated than during training.*

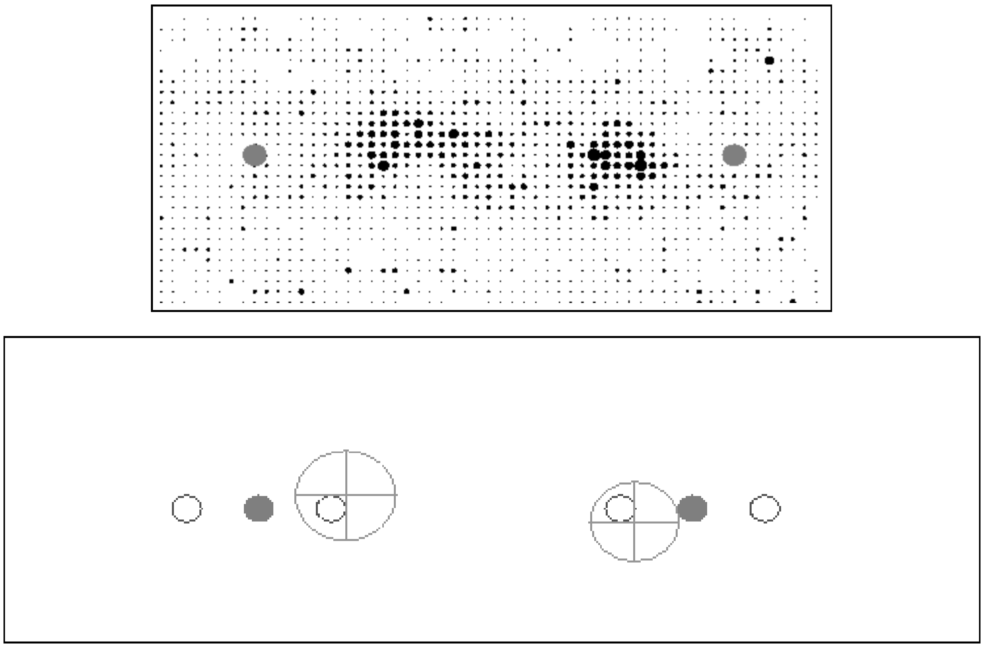


Figure 2.28: Tested with two landmarks more separated than during training. (top) Histogram of search time of gerbils. (bottom) Analysis of search time showing best fit by Gaussians.

The two-landmark array can also be stretched, see Figure 2.28. When trained with two landmarks and tested with the distance between them doubled, the gerbils again searched at two locations, each at the correct distance and bearing from the corresponding landmark. From eyeballing the data, it is clear that the animals did not search at the two exterior locations (same distance from landmark, but outside the pair) as well as the internal ones, and the analysis definitively finds that the data is best explained by two Gaussians, each accounting for 50% of the data, centered at the expected locations.

### 2.5.2 Three landmark experiments.

We also tested manipulations of a three-landmark configuration: reward placed at the center of an equilateral triangular array.

#### 2.5.2-A T3LM

**Reviewed Result EX.125** *Gerbils can be trained to find food at the center of a triangle of local landmarks even in cue-rich environments.*

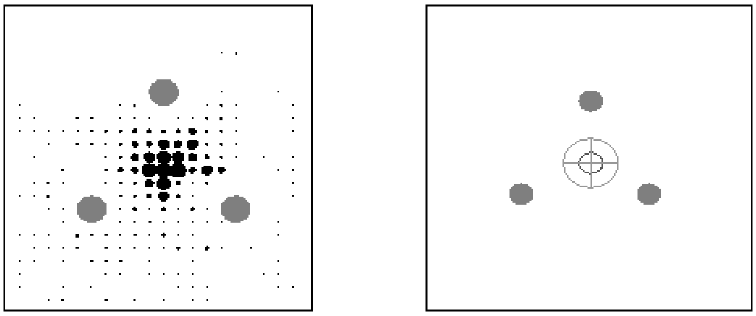


Figure 2.29: Three landmarks. (left) Histogram of search time of gerbils. (right) Analysis of search time showing best fit by Gaussians.

As with the previous two configurations, well-trained animals searched for food at the center of the array. The analysis shows that this search distribution is best explained by a single Gaussian centered over the goal.

#### 2.5.2-B T3M1

**Reviewed Result EX.126** *In cue-rich environments, gerbils trained to find food at the center of a triangle of local landmarks and faced with a pair of landmarks, search on both sides of the pair, but spend most of their search time on the pair predicted by the orientation of the triangle.*

In our experiments, removing one landmark produced search on both sides of the line formed by the remaining landmarks. The analysis finds two Gaussians, one at the expected location (treating the two landmarks as part of a correctly-oriented triangle) and one at the mirror location. The Gaussian at the expected location accounts for 70% of the data, while the Gaussian at the mirror accounts for 30% of the data. Collett *et al.* report that the animals search on only one side of the landmarks, however, re-examination of their figure (reprinted as Figure 2.8) shows that there is some search on the opposite side of the line.

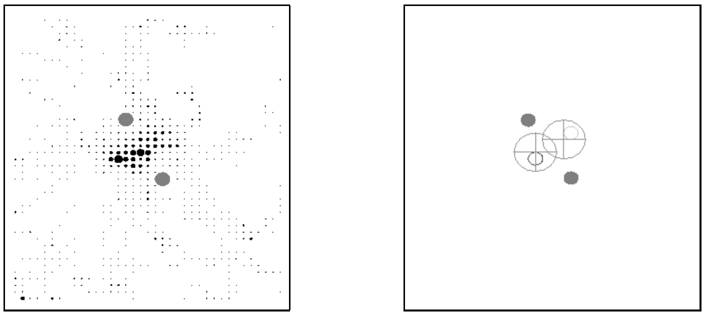


Figure 2.30: Tested with two landmarks after being trained with three. (left) Histogram of search time of gerbils in our experiments. (right) Analysis of search time from our experiments showing best fit by Gaussians.

### 2.5.2-C T3M2

**Reviewed Result EX.127** *In cue-rich environments, gerbils trained to find food at the center of a triangle of local landmarks, and faced with a single landmark, do not distribute their search time among three distinct locations as described by Collett* et al. *(1986).*

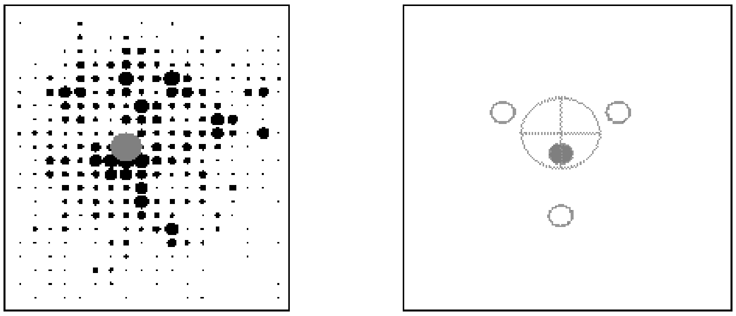


Figure 2.31: Tested with one landmarks after being trained with three. (left) Histogram of search time of gerbils in our experiments. (right) Analysis of search time from our experiments showing best fit by Gaussians.

Collett *et al.* report that when trained with three landmarks and tested with one, gerbils searched at three locations corresponding to matching the landmark to each of the remembered landmarks in turn. We were unable to replicate this result in our lab.

We do not know why our experimental results are so different from Collett *et al.*'s, but one possibility is that Collett *et al.* performed much more stringent cue controls than we did. Our arena was completely open to the room and there were plenty of distal cues available to the animals at all times. Biegler and Morris (1996) report similar effects while trying to train animals to clustered vs. separated landmark pairs.

### 2.5.2-D T3SA

**Reviewed Result EX.128** *Even in cue-rich environments, gerbils trained to find food at the center of a triangle of local landmarks and faced with two landmarks at the same relation to each other as during training (the "non-displaced" landmarks) and one landmark farther away from the center (the "displaced" landmark) prefer to search near the non-displaced landmarks. They do not search at the center of the stretched triangle.*

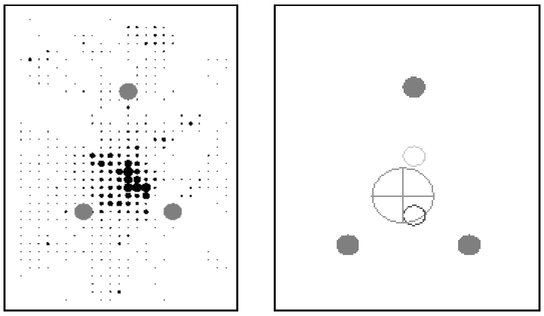


Figure 2.32: Trained with three landmarks, tested with one landmark moved. (left) Histogram of search time of gerbils. (right) Analysis of search time showing best fit by Gaussians. Light empty circle shows center-of-mass goal prediction.

In addition to adding or removing landmarks, one can manipulate the positions of the three landmarks relative to each other. For example, one can stretch the triangle by moving one of the landmarks away from the other two, while leaving the distance between the other two unchanged. The search distribution is best explained by a single Gaussian centered at the correct distance from the two landmarks with the correct inter-landmark distance, but too far from the third landmark. We can use the analysis to disprove a hypothesis about the gerbils' navigation strategy: one possible strategy is that the gerbils are simply going directly to the center of the landmark array. However, this point lies outside the 95% confidence interval for the Gaussian used to explain the search distribution. Thus we conclude that the animal is searching closer to the two "unchanged" landmarks than to the third "stretched" one.

### 2.5.2-E T3RT

**Reviewed Result EX.129** *In cue-rich environments, gerbils trained to find food at the center of a triangle of local landmarks search at the center of a similar triangle that has been rotated by* 180°.

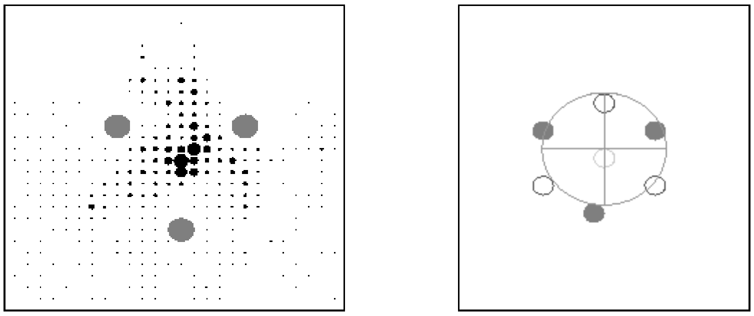


Figure 2.33: Trained with three landmarks, tested with triangle rotated by 180°. (left) Histogram of search time of gerbils in our experiments . (right) Analysis of search time showing best fit by Gaussians. Light empty circle shows center-of-mass goal prediction.

Finally, the triangle can be inverted (or equivalently, rotated by 60°). When this is done, Collett *et al.* report that the gerbils first search the center of the array and then proceed to search three exterior points. Because we did not observe digging throughout the manipulated experiments, we cannot say that the gerbils first search the center and then the exterior points.

Our thresholded search histogram looks extremely similar to that reported by Collett *et al.* (compare Figures 2.33 and 2.11), however, our analysis finds that the search distribution is best modeled by a single Gaussian.

## 2.6 Contributions: Exploration

We[10] have also tested whether animals return regularly to their home base during exploration. Compare EX.60, above.

**Reviewed Result EX.130** *Gerbils exploring a novel cue-rich environment regularly return to their starting points.*

**Methods.** We allowed naive gerbils to explore a large open field arena (the same arena used in Section 2.5, but with no local landmarks). The surroundings were very cue-rich and no attempt was made to control any of the cues, except to ensure that they did not move during a single session. Sessions lasted for fifteen minutes while position was sampled at 18 frames/sec from a picture taken by a camera mounted above the arena.

[10] This work was done in collaboration with Sofyan Alyan, David Banks, Steven J. C. Gaulin, Michael Pahn, Chris Reiber-Milberg, Lisa Saksida, and David S. Touretzky, and has been reported in abstract form (Touretzky *et al.*, 1996).

We tested six gerbils (three male and three female). No significant differences were found between them. I will therefore present a canonical example.

Figure 2.34 shows the track taken by one gerbil during a complete fifteen minute session (its first in the environment). Figure 2.35 shows a two-dimensional histogram of the time spent by the gerbil at each location. The point at which it spent the most time is declared the *home base*.

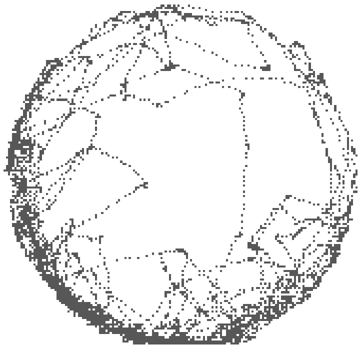


Figure 2.34: Track of a gerbil exploring a large circular open field (10 ft diameter) for 15 minutes.

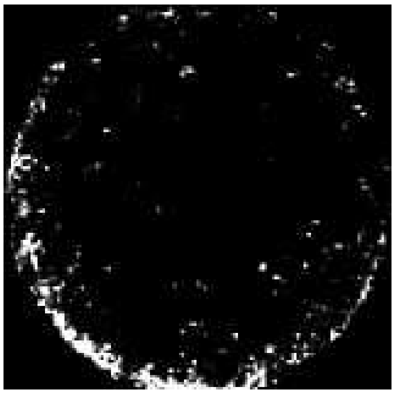


Figure 2.35: Histogram of time spent by a gerbil exploring a large circular open field (10 ft diameter) for 15 minutes.

The gerbil spent most of its time in the octant containing the home base, and it also entered that octant more than any of the others. This demonstrates that the reason the gerbil spent most of its time in the home octant was not because it just sat there, but because

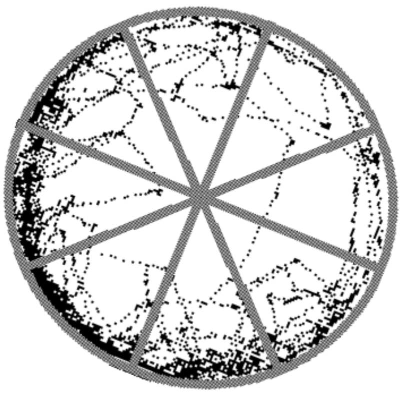


Figure 2.36: The eight octants we used overlaid on the track from Figure 2.34.

it kept returning there. Figure 2.36 shows how the octants were laid out (overlaying the X-Y track from Figure 2.34). Figure 2.37 shows a histogram of the time spent in each octant and Figure 2.38 a histogram of the number of times the animal entered the octant. As can be clearly seen in Figures 2.37 and 2.38, the gerbil kept returning to the home base.

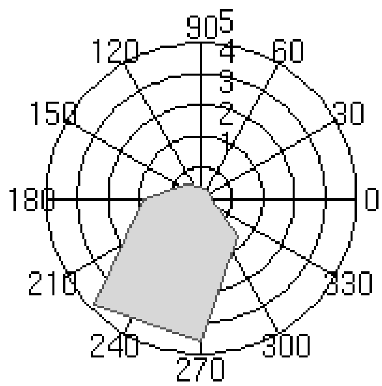



Figure 2.37: Histogram of time spent in each octant.

Although we can determine the existence of a home base from the previous figures, we wanted to see if the animal would *regularly* return to its home base. We test this by plotting the distance between the animal and the putative home base by time. Call this $D(t)$. As can be seen in Figure 2.39, $D(t)$ regularly approaches 0, suggesting that not only does the gerbil often return home, but it regularly returns home.

We can formalize this by calculating the Fourier transform of $D(t)$, to look for cycles.

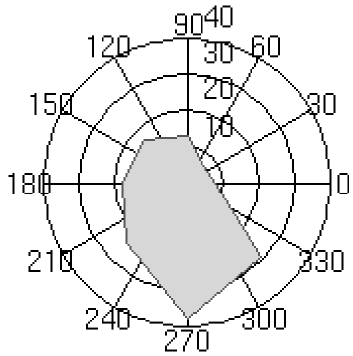



Figure 2.38: Histogram of number of visits to each octant.

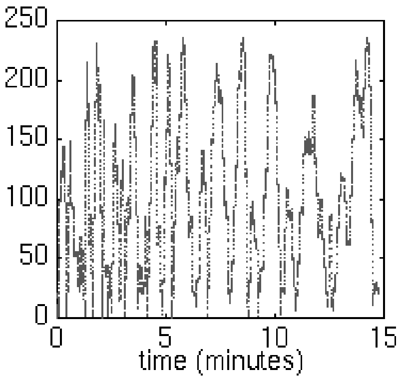



Figure 2.39: Plot of *distance from home* $D(t)$ as a function of time. The gerbil is regularly returning to its *home base*.

Figure 2.40 shows the first 100 components of the Fourier transform. As can be seen, the only significant cycles occur at 15-30 cycles per 15 minutes, i.e. with periods of about 15-60 seconds.

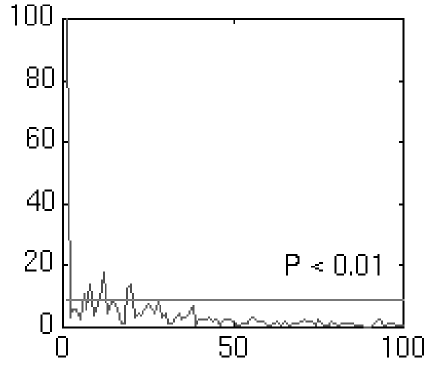



Figure 2.40: The first 100 components of the Fourier transform of the function $D(t)$ plotted in Figure 2.39. The line indicates significance of $P < 0.01$.

Alternatively, we can look at a histogram of the *inter-home* times, which we can define in the following way: Let the gerbil be *at home* when it is within 10 cm of its starting point. Then we can define the *inter-home* time as the time between successive samples that are *at home*. We expect a huge peak at 0 seconds (which we crop short in Figure 2.41). There is another broad peak between 30 and 60 seconds.

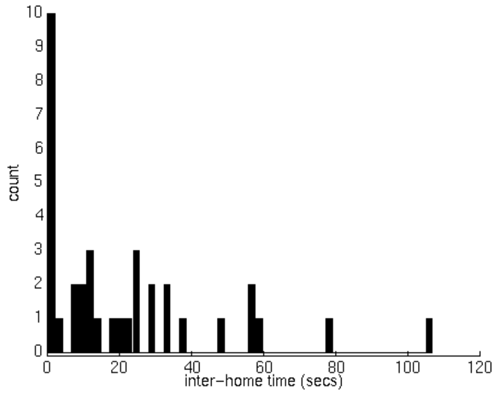



Figure 2.41: Inter-home time histogram. The *interhome time* is defined as the time between two successive samples both of which are within 25 cm of the *home base*. The 0 entry has been truncated for display purposes.

# Chapter 3

# Navigation overview

Having reviewed the experiments forming the domain of rodent navigation (Chapter 2), I will now bring together a comprehensive theory of navigation. This theory is built from ideas pulled from the extensive work done by other researchers over the last hundred years with some ideas of my own. I will be careful to identify the sources of each of the ideas, but one contribution I have made in this thesis is to bring these ideas together.

Let us begin our discussion of navigation with a look at a classic navigation task: the *water maze*, see Experimental Review, Section 2.1.2, Figure 2.12.

There are four strategies an animal can take to find the platform (O'Keefe and Nadel, 1978; Sutherland *et al.*, 1983; Whishaw and Mittleman, 1986):

- **Taxon navigation.** The animal can find a cue towards which it can always swim. For example, if the platform is visible, it can simply "swim towards the platform."

- **Praxic navigation.** The animal can execute a constant motor program. For example, if the animal always starts at the same location, in the same orientation, and the platform does not move, it can use praxic navigation to reach the platform.

- **Locale navigation.** The animal can learn the location of the platform relative to the constellation of cues.

- **Random navigation.** If the animal has no information about the location of the platform, it must search randomly for it.

Each of these strategies can be used by different animals at different times. Most researchers use a hidden platform and start the animal from a variety of starting points. This forces the animals to use a locale strategy (Morris, 1981). However, if the animals are given a cue or a visible platform, they can learn to use a taxon strategy (Morris, 1981), and if they are always started from a limited number of positions, they can learn a collection of praxic strategies (Whishaw and Mittleman, 1986).

Animals can also execute combinations of these strategies. For example, animals with hippocampal lesions often spend most of their time circling the environment at the correct distance from the wall (Morris *et al.*, 1982; DiMattia and Kesner, 1988; Morris *et al.*, 1990; Sutherland and Rodriguez, 1990, EX.14). This is an example of taxon (near to the wall) and praxic navigation (circle the environment).

The caudate nucleus is critical for taxon and praxic strategies, while the hippocampus is critical for locale strategies. This can be most cleanly seen in the following three experiments:

In the water maze, lesions of the caudate nucleus disrupt navigation to cued platforms, such as visible platforms or platforms with a large black card marking the quadrant containing the goal, but not to hidden platforms (McDonald and White, 1994), while hippocampal lesions impair acquisition of hidden platforms, but not visible or cued platforms (Morris *et al.*, 1982; Eichenbaum *et al.*, 1990; Morris *et al.*, 1990; Sutherland and Rodriguez, 1990; White, 1997).

McDonald and White (1994, EX.17) trained rats in the water maze, alternating visible and hidden platforms always in the same location. After twelve days, the animals were tested with a visible platform in a new location. Half of the animals went to the new visible platform (taxon strategy), half went to the location where the old platform had been (locale navigation). With hippocampal inactivation (by lidocaine), all of the animals went to the new platform, but with caudate inactivation, all of the animals went to the old location. With hippocampal inactivation, the animals showed a taxon strategy, but with caudate inactivation, they showed a locale strategy.

Packard and McGaugh (1992, EX.17) compared caudate and fornix lesions on a water maze with two platforms, one stable, one unstable. When location of the platform was correlated with its stability, fornix lesions impaired acquisition, and caudate lesions had no effect; but when a visual cue demarcated which of the two was stable, caudate lesions impaired acquisition of the cue-stable platform, and fornix lesions had no effect.

In another experiment by Packard and McGaugh (1996, EX.41), on a plus-maze, rats were always started on the south arm and trained to turn left (to the west arm). After eight days of training, they were placed on the north arm. They all turned right (to the west arm). This demonstrates that they were using a locale strategy. After sixteen days of training, they were again placed on the north arm. They all turned left (to the east arm), demonstrating a praxic strategy. Packard and McGaugh then inactivated caudate (with lidocaine) and tested the animals on the north arm. They all turned right, demonstrating the locale strategy again.

Thus taxon and praxic navigation strategies require the caudate nucleus while locale strategies require the hippocampus. In the following sections, I will address the computations and neurophysiology of taxon, praxic, and locale navigation. I will not address random navigation in this thesis.

## 3.1 Taxon and praxic navigation

Taxon navigation strategies identify a cue and move the animal towards or away from it, while praxic navigation strategies execute a specific motor program. Both of these are stimulus-response strategies, i.e. given a single stimulus, there is a single response which can be easily learned. They are sometimes called *route* strategies (O'Keefe and Nadel, 1978) because traversing a route requires one or both.

As discussed above, animals with caudate lesions cannot learn visibly-cued water maze tasks, and caudate lesions unmask locale navigation strategies on over-learned tasks on the plus-maze. In addition, Potegal (1982, EX.49) showed that caudate lesions impair praxic navigation strategies. Animals were trained on a 12-arm radial maze in which the food was located on an arm at a constant angle from the starting arm. This meant that the only viable strategy to find the food was to go to the center of the maze and make a turn at that angle. This is a praxic strategy. Neither locale nor taxon strategies will help the animal solve the task. Caudate lesions impaired the animal's ability to find the food.

Wiener (1993, EX.104) recorded from caudate cells and found cells correlated with direction, location, and behavioral sequence. The arena was a small square with food cups in each corner. Mizumori and Cooper (1995, EX.104) also recorded from caudate cells, this time on the 8-arm radial maze. They found that most cells were tuned to single directions along pairs of arms, but also found that some cells changed their preferred directions suddenly and drastically within a single trial. Others were tuned, not to direction, but to a path along a pair of arms.

Parietal cortex is also involved taxon strategies. In the hidden-platform water maze, parietal lesions produce random search (DiMattia and Kesner, 1988, EX.23). Even when a visible cue is available (whether the platform itself is visible or there is a visible cue card on the wall indicating the quadrant with the platform), animals with parietal lesions reverted to circling at the appropriate distance from the wall (Kolb and Walkey, 1987; Kolb, 1990b, EX.23). Parietal cortex may be involved in the representation of the local view, which represents spatial aspects of local and distal landmarks (see Section 3.2.1 below for discussion and references). Obviously, in order for an animal to use taxon strategies (say, approach to a landmark), the animal must be able to represent spatial parameters such as the egocentric bearing to the landmark.[1]

Taxon strategies can be understood using simple stimulus-response models. Given a single stimulus, the animal simply orients towards it and proceeds. Models of taxon strategies have been extensively studied, see Schöne (1984) for a review, and will not be reviewed here. Figure 3.1 shows some of the structures involved in taxon navigation.

Praxic strategies can also be understood using simple stimulus-response models, but instead of a simple orienting response, the response is a more complex sequence of motor

[1] Strategies typically described as taxon usually only require simple spatial parameters such as egocentric bearing, not complex spatial parameters such as distance and allocentric bearing. Egocentric bearing is the angle between some aspect of the animal's body (such as its midline) and the cue, while allocentric bearing is the angle between the landmark and an external reference direction ("global north").

actions. These kinds of motor sequences have been well-studied also, under the rubrics of motor pattern generators (Gallistel, 1980), motor programs (Alexander *et al.*, 1992), sequence learning (Blum and Abbott, 1996; Levy, 1996; Gerstner and Abbott, 1997, see also examples in Houk *et al.*, 1995), and reinforcement learning (Sutton, 1992) and will not be reviewed here.

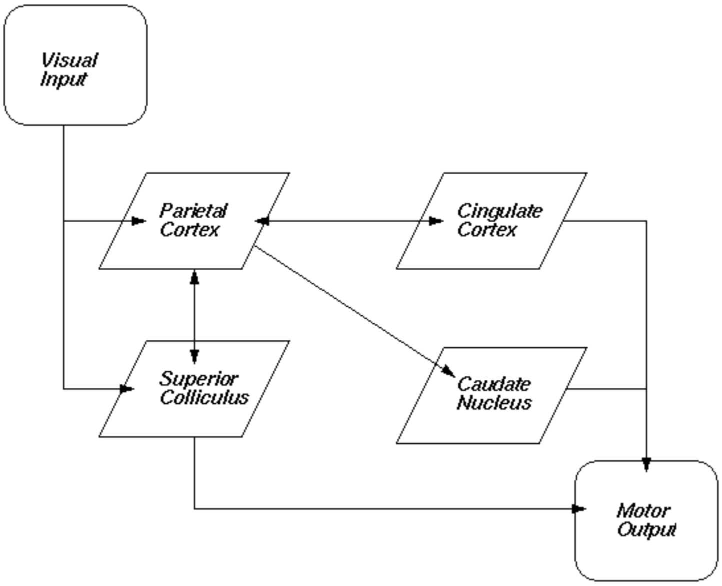



Figure 3.1: Some of the areas involved in taxon navigation.

## 3.2 Locale navigation

Locale navigation is the ability to navigate to a target that is not identified by any direct cues (O'Keefe and Nadel, 1978; Morris, 1981). I will argue in this overview that a rodent's ability to perform locale navigation tasks such as the hidden platform water maze depends on an interaction between five different spatial representations (see also Redish and Touretzky, 1997a):

- a representation of the animal's relationship to landmarks in its environment called the **local view**,
- a metric representation of position that accommodates vector arithmetic (the **path integrator**),
- a representation of orientation in space (**head direction**),

- a distributed representation of position that ties local views to path integrator coordinates (the **place code**), and
- a **goal memory** that associates motivational and spatial inputs and accommodates trajectory planning.

These abstract subsystems should not be expected to be anatomically localized; their functions will be distributed across several brain structures.

### 3.2.1 Local view ($\mathcal{LV}$)

The local view subsystem represents the sensory aspects of landmarks. It should at a minimum include representations of spatial aspects of the various landmarks (such as distance and bearing). The term "local view" was first introduced to mean what can be seen from a particular viewing position (McNaughton, 1989, see Leonard and McNaughton, 1990 for a review), however even the earliest theories of local view did not require the landmarks to be solely visual (see Leonard and McNaughton, 1990). Rodents can navigate based on any number of non-visual cues, such as auditory, olfactory, and somatosensory cues (Schöne, 1984; Goodale and Carey, 1990; Kelly, 1990; Stein and Meredith, 1993), and rats with severe sensory deficits can still navigate accurately using external cues (Watson, 1907; Carr and Watson, 1908; Carr, 1917; Dennis, 1932; Honzik, 1936; Zoladek and Roberts, 1978; Hill and Best, 1981; Save *et al.*, 1996).

Thus *local view* consists of the sensory input available from a particular viewing position and orientation in a specific environment. With a sufficiently rich set of cues and no pathological symmetries in the environment, local views describe unique places, however, in the real world, some cues move from time to time, and others appear and disappear.

#### 3.2.1-A Spatial aspects

Since the point of the local view system is to allow the animal to determine its position from external cues, we can make some computational claims about the spatial aspects included in the local view. The local view must include enough spatial cues to uniquely define a point. For example, three distance measurements are sufficient, but two are not. As can be seen in Figure 3.2, having only two distance measurements leaves an ambiguity of two points. The actual representation of landmark spatial aspects is likely to be noisy, so the local view can be expected to overdetermine position in order to compensate. So what aspects can be included in local view?

**Distance.** Distance to landmarks is correlated with aspects such as size on the retina and position relative to the horizon. Both of these aspects vary regularly with respect to the distance an object is from the animal. Thus a cell tuned to these aspects will appear to be tuned to distance.

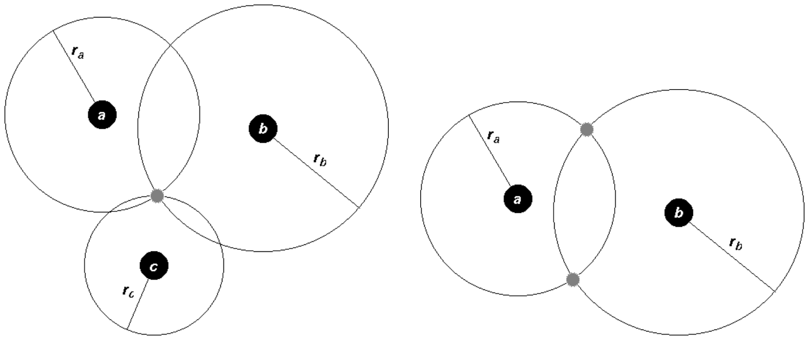



Figure 3.2: Knowing the distances to three unique, non-co-linear landmarks allows localization to a single point (left), but knowing distances to only two does not (right).

Primates have forward facing eyes which allow them to use stereo vision to determine distance. Rodents are prey animals and in general do not have forward facing eyes,[2] but the overlap of the visual fields of the nasal retinae does allow about 50° to 80° of binocularity (Goodale and Carey, 1990). Both Goodale and Carey (1990) and Dean (1990) suggest that head movements seen during rearing in discrimination-learning tasks may be designed to bring discriminants into the temporal retinae so that they can be viewed stereoscopically. Gallistel (1990) suggests that the purpose of these movements is to produce parallax which could also provide distance information.

**Egocentric and allocentric bearing.** In addition to distance, orientation to landmarks must be a key component of local view. Orientation can be encoded *egocentrically* or *allocentrically.* Egocentric bearing measures the angle between the landmark and the animal's heading ($\theta_i$ in Figure 3.3), while allocentric bearing measures the angle between the landmark and a reference direction ($\phi_i$ in Figure 3.3). Allocentric bearing can be calculated from egocentric bearing given the angle between the animal's heading and the reference direction ($\Phi_H$ in Figure 3.3).

$$\phi_i = \theta_i + \Phi_H \tag{3.1}$$

[2]Of course, most rodents are not generally visual creatures; they are generally nocturnal and their visual acuity is much less than ours (gerbils: 1.75 cycles per degree, albino rats: 0.38 cycles/deg, pigmented rats: 1.2 cycles/deg, compare with humans: 60 cycles/deg) (Collett *et al.*, 1986; Leonard and McNaughton, 1990), but rodents are sensitive to distal cues which are only detectable visually (e.g. EX.44, EX.67, etc.). An interesting question is whether all these measurements hold for diurnal, arboreal creatures such as squirrels who need to calculate distance to a branch very accurately. For the purposes of this thesis, I will accept that rodent vision is poor but useful, and remind the reader that "local view," irrespective of its name, includes non-visual cues as well as visual.

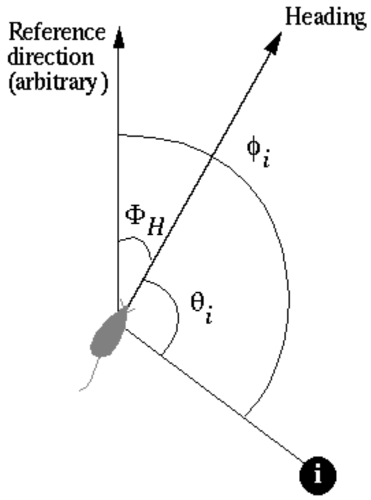



Figure 3.3: Ego- and allocentric bearings: $\Phi_H$: Angle between head direction and reference orientation; $\theta_i$: egocentric bearing to landmark $i$; $\phi_i$: allocentric bearing to landmark $i$.

Allocentric bearings have the advantage that they are independent of the animal's orientation, but they require a representation of the animal's heading relative to the reference direction ($\Phi_H$ in Equation 3.1).[3]

Taxon navigation is generally based on egocentric bearing: orient towards a stimulus and approach it. Evidence that animals can use egocentric bearing for navigation is extensively reviewed by Schöne (1984).

Evidence that rodents have access to allocentric bearing is more subtle, but also exists. As reviewed in the experimental review, rodents can learn to search for food at a specific distance and bearing to a single landmark; they do not search at an annulus around the landmark (Collett *et al.*, 1986, EX.1). Collett *et al.* attempted to control the available external directional cues: they used a circularly symmetric landmark and the walls of the room were surrounded by a black curtain. Access to allocentric bearing can explain the search distribution seen by Collett *et al.* Biegler and Morris (personal communication) had difficulties training an animal to search at a single point relative to a circularly symmetric landmark, but they disoriented their animals before each trial. If the allocentric bearing computation is dependent on the representation of head direction, then the disorientation used by Biegler and Morris would be like putting the food at a different orientation from

[3]This is exactly what is provided by the representation of head direction, see Section 3.2.2 below.

the landmark with each trial. This issue has also been examined by Knierim *et al.* (1995).

#### 3.2.1-B Cue types

**Point landmarks.** There are many examples of rodents navigating based on small objects (e.g. the cylinders in Collett *et al.*, 1986), but whether these are seen as single "objects" by the rodents is an open question. However, most models treat landmarks as *points* in space. This gives the distinct advantage that distance and bearing are easily defined. Of course, no object is a true point (i.e. all objects have extent), but whether one considers navigation to an object or a point on an object is irrelevant. What matters is that rodents can measure spatial parameters of point-like landmarks.

**Surface landmarks.** Rodents with hippocampal damage solve the water maze by swimming at a set distance from the wall (EX.14), strongly implying they have a representation of their relation to the wall, which can then be used in taxon navigation. An important question is how far away from the surface an animal can detect the surface orientation. If rodents use their vibrisae to detect local surface orientation, then they would only be able to use the wall as a sensory cue when close.

Another indication that the local view might include some representation of local surface orientation is that changes in the local surface orientation produce changes in place fields (Muller *et al.*, 1987; Muller and Kubie, 1987, see EX.65 and EX.69). Muller *et al.* (1987) report cells with place fields along the wall of a cylindrical arena (see Figure 2.20). The interior edges of these fields are concave (i.e. the fields "hug" the arena wall). This implies that the cells must be sensitive to the arena wall;[4] the fact that these cells have only been seen near the wall may imply that the detection range of surface orientation might be very local. The fact that that only barriers that intersected a cell's place field changed the field (Muller and Kubie, 1987) also supports the hypothesis that the detection range of surface orientation might be local.

**Compass points** If a landmark is very distant to the environment, the egocentric bearing to it won't change much as the animal moves around an arena. In addition, accurate distances would be hard to determine (beyond that it is very far). These cues are sometimes called *compass points*, because they can be used to inform the head direction system directly: an association between the egocentric bearing to a distal cue and a head direction is valid no matter where an animal is in the environment. Like the north star, the distal cue can serve as a direct orienting cue.

Whether local view preferentially includes information from distal or local cues has been measured behaviorally by examining whether rodents follow intra-maze or extra-maze cues (e.g. Watson, 1907; Honzik, 1936; Dennis, 1932; Olton and Samuelson, 1976; Suzuki

[4] There are two alternative hypotheses explaining these place fields. They are described in the discussion of crescent shaped place cells, page 100.

*et al.*, 1980). Early experiments tested this by training rats to solve complex mazes (see Figure 2.17) made of elevated tracks and then interchanging or replacing the tracks (without changing the basic shape of the maze) (Honzik, 1936; Dennis, 1932). Presumably, the distinctive features of the tracks (such as odor or texture) would then conflict with the extra-maze cues. They found that generally this manipulation had little or no effect, suggesting that animals followed distal cues.

Other attempts at this question rotated or translated the maze (Watson, 1907; Carr and Watson, 1908; Honzik, 1936). Unfortunately, rotating the maze has the unintended consequence of changing the allocentric bearing of intramaze cues as well as the relation to extramaze cues. Watson (1907) found that rotating the maze produced severed deficits while translating the maze did not.

More modern behavioral attempts to get at this question have looked at the radial maze (e.g. Suzuki *et al.*, 1980, EX.44). By rotating the radial maze, internal and external cues are in conflict. Again, animals generally follow extra-maze cues, but can be trained to follow the internal cues.

Whether the local view preferentially incorporates information from distal or local cues has generally been tested by measuring whether place cells prefer to follow distal or local cues. Many early experiments used T-, plus-, or radial mazes with cues available on the walls of the surrounding room (e.g. O'Keefe and Conway, 1978; Miller and Best, 1980; O'Keefe and Speakman, 1987; Shapiro *et al.*, 1989; Young *et al.*, 1994; see EX.67). Sometimes there were differing textures on each arm of the maze, but even if there weren't one would expect some sort of local intra-maze cues. The local and distal cues can then be separated by rotating the maze. Place cells generally followed the distal cues (Miller and Best, 1980; Shapiro *et al.*, 1989). Similar results have been found by Cressant *et al.* (1997) in which they showed that inside a small circular arena place fields only follow the rotated cues if the landmarks were pushed all the way against the wall, i.e. if they are *orienting* or *distal* stimuli.

On the other hand, animals can be trained to perform tasks based on local cues, such as searching for food at a position relative to an array of local landmarks (Collett *et al.*, 1986; Biegler and Morris, 1993; Saksida *et al.*, 1995; Biegler and Morris, 1996; Gothard *et al.*, 1996a; Gothard *et al.*, 1996b; see EX.1–EX.12 and EX.122–EX.129) However, training can take a long time.

This implies that rodents are initially sensitive to distal landmarks, but can learn to attend to local landmarks if they are stable relative to the animal's internal representations and provide spatially constant clues to reward location.

### 3.2.1-C Non-spatial aspects

Traditionally, the term local view has been used to refer to spatial parameters of landmarks (both local and distal); most place cell models allow each landmark to have a unique identity $T$ (see, for example, Sharp, 1991; Hetherington and Shapiro, 1993; Touretzky and Redish, 1996). Although these labels do not have to be unique for each landmark (in many models they are), they are assumed to identify landmarks perfectly. That is, each landmark is associated with a type and these types are not confusable. For example, if there are three white cylinders in an environment and an arena wall, the three cylinders might have a type $T_1 = T_2 = T_3 = A$, while the arena wall might have type $T_4 = B$.

Rodents are also sensitive to non-spatial aspects of the environment, such as luminosity of a cue (Bostock *et al.*, 1991). How these aspects inform the local view is still unknown, but they should probably be included in the concept.

#### 3.2.1-D Anatomy

In the primate, there is strong evidence that there are two representational pathways (or *streams*) for the local view: one representing spatial aspects (the dorsal stream, passing through parietal cortex) and the other representing identity aspects (the ventral stream, passing through inferotemporal cortex). The extensive evidence supporting this distinction in the primate will not be reviewed here, I refer the reader to the many excellent reviews such as Mishkin *et al.* (1983), Ungerleider and Haxby (1994), and Goodale *et al.* (1994).

There is evidence that this same distinction occurs in rodents (Kolb, 1990a; Kolb *et al.*, 1994). The discussion of the two streams generally occurs in the visual mode (discussing how visual information is processed), but it should be remembered that the local view is not solely visual. In fact neurons in posterior parietal cortex (PPC) of the rodent are multi-modal (Chen, 1989; Chen, 1991; Chen *et al.*, 1994a; McNaughton *et al.*, 1994b) (as they are in the primate, Hyvarinen, 1982; Colby and Duhamel, 1991). We know of no recordings of neurons in rodent inferotemporal cortex (Te2), but in the primate inferotemporal cortex neurons are primarily visual (Desimone *et al.*, 1985; Tanaka *et al.*, 1991). This suggests that while spatial parameters may be encoded in a supra-modal representation, object identity might not be.

Evidence that the spatial aspects of local view are represented in posterior parietal cortex comes from anatomical, lesion, and neurophysiological data. More extensive neurophysiology work on parietal cortex has been done in the monkey (see Andersen *et al.*, 1993, and Colby *et al.*, 1995, for reviews) and extensive neurology work has been done in humans with parietal lesions (see Bisiach and Vallar, 1988, and Ungerleider and Haxby, 1994, for reviews). In both the human and non-human primate, parietal cortex seems to be critically involved in representing nearby space (Stein, 1991; 1992). There is extensive cytoarchitectonic and neurophysiological evidence that primate parietal cortex consists of a variety of sub-areas (Andersen, 1988; Colby and Duhamel, 1991); rodent parietal cortex is also likely to consist of a variety of sub-areas, but more work needs to be done to decode what these areas are and the roles they play in representing the local view. Zilles (1990) identifies four parietal areas as well as three secondary visual areas. Zilles identifies the secondary visual areas with Krieg's (1946) area 7, which Krieg suggests as the homolog of primate parietal cortex. Zilles notes that even these areas are not necessarily cytoarchitectonically homogeneous and may need to be subdivided even further by additional anatomical studies. For now, I will continue to treat PPC as a single unified brain structure, keeping in mind the caveat that it is actually quite diverse.

Posterior parietal cortex in the rat receives input from the lateral dorsal and lateral posterior thalamic nuclei (LDN and LPN, respectively), as well as from primary visual cortex (Kolb and Walkey, 1987; Chandler *et al.*, 1992). The superior colliculus sends information to the posterior parietal cortex via the LPN (Dean, 1990). I suspect that PPC also receives input from auditory representations (as is evidenced in the primate by auditory responses of parietal neurons, Stricanne *et al.*, 1996), but have found no data supporting or refuting this hypothesis. PPC sends output to the posterior cingulate cortex (Chandler *et al.*, 1992), the postsubiculum (van Groen and Wyss, 1990), the superior colliculus (Kolb, 1990a), and to the postrhinal cortex (Burwell *et al.*, 1995), which sends robust efferents to the entorhinal cortex (Burwell *et al.*, 1995). Some cells in parietal cortex show directional and behavioral correlations that may indicate they play a role in local view processes (Chen, 1989; McNaughton *et al.*, 1991; Chen, 1991; Chen *et al.*, 1994a; Chen *et al.*, 1994b; McNaughton *et al.*, 1994b; EX.105). LDN cells also show directional correlates but are dependent on intial availability of light (Mizumori *et al.*, 1992, EX.103) indicating that they may also play a role in local view processes.

In contrast to posterior parietal lesions, Te2 lesioned rats are poor at visual pattern discriminations, but not at spatial orientation discriminations (Kolb, 1990a). They are unable to learn to perform a visual match to sample, even with no delay (Kolb, 1990a). But they can learn the hidden-platform water maze normally. I know of no recordings from Te2 neurons in rats, but in the primate, temporal cortex neurons are sensitive to visual patterns and to object identities (Desimone *et al.*, 1985; Tanaka *et al.*, 1991).

Te2 and PPC project to the perirhinal and postrhinal cortices respectively which both project to entorhinal cortex (EC) (Burwell *et al.*, 1995). The perirhinal and postrhinal cortices receive direct afferents from a host of other cortical and subcortical sources, including (among others) the somatosensory, auditory, and medial prefrontal cortices as well as the anterior and posterior cingulate cortices, amygdala, nucleus accumbens and caudate nucleus (Burwell *et al.*, 1995). Te2 also sends direct projections to the lateral EC (Staubli *et al.*, 1984; Myrher, 1991). Both medial and lateral EC send projections into the hippocampus via the perforant path. This suggests that the dorsal/ventral separation may be combined in EC.

However, Otto *et al.* (1996) report data that suggests that the dorsal/ventral separation is maintained into lateral/medial EC (EX.33). They found that in a task which paired an odor with a location (given a specific odor, go to a specific location), LEC lesions produced occasional odor discrimination errors (animal went to the location indicated by the wrong odor). When the lesions encroached on MEC, the animals also occasionally made spatial errors (animal went to a location not associated with any odor). We will see in our discussion of *reference frames* (Section 3.3, below) that while one EC input pathway may inform the animal of its location, the other may inform it of the environment (reference frame) it is in.

The local view subsystem represents the animal's relationship to landmarks. These cues can be differentiated from cues like vestibular cues, which can only tell the animal how quickly it is turning or accelerating. Cues can be categorized as being either *external* or

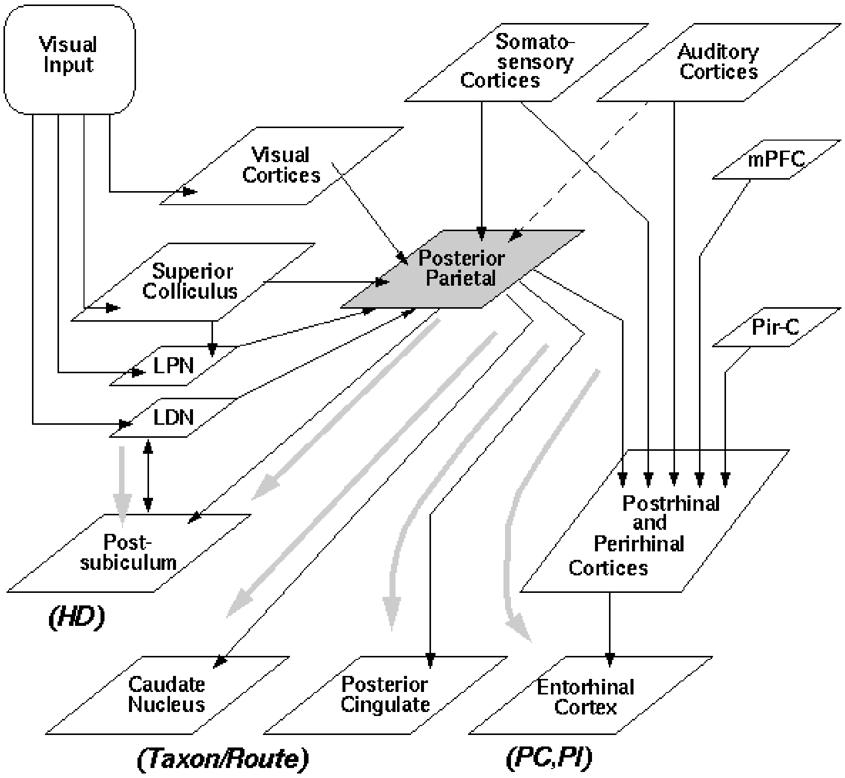



Figure 3.4: LV anatomy. Visual input enters the system through the primary visual cortex, the superior colliculus, and the lateral dorsal and lateral posterior thalamic nuclei (LDN, LPN). Visual, somatosensory, and auditory input is combined in posterior parietal cortex which distributes the information through to the other navigation systems (HD : head direction system, via Post-subiculum; PC,PI: place code, path integrator, via entorhinal cortex; Taxon/Route: route system, via posterior cingulate). Postrhinal and perirhinal cortices receive afferents from posterior parietal, medial prefrontal (mPFC), piriform (Pir-C) and other cortical and subcortical structures not shown here.

*internal*.[5] External cues allow one to calculate one's position given general knowledge about the environment, but requiring no prior knowledge about one's position; internal cues allow one to calculate one's position at time $t$ given a known position at $t - \Delta t$, but requiring no knowledge of the environment. Thus external and internal cues complement each other. Internal cues will be addressed in the following two sections: Head direction (HD) and Path integration (PI).

### 3.2.2 Head direction ($\mathcal{HD}$)

Because animals can use allocentric bearing in navigation (e.g. Collett *et al.*, 1986, EX.1, see also Section 3.2.1, above), they must have a means of calculating allocentric bearing. Allocentric bearing can be computed by measuring egocentric bearing and factoring out the animals own head direction (Equation 3.1).

Cells with firing rates reflecting head direction have been discovered in a number of structures in the rodent brain:[6] postsubiculum (PoS: Taube *et al.*, 1990a, EX.94), the anterior thalamic nuclei (ATN: Blair and Sharp, 1995; Knierim *et al.*, 1995; Taube, 1995, EX.95), the lateral mammillary nuclei (LMN: Leonhard *et al.*, 1996, EX.96), the lateral dorsal nucleus of the thalamus (LDN: Mizumori and Williams, 1993, EX.103), and to a lesser extent parietal and cingulate cortex (Chen *et al.*, 1994a; Chen *et al.*, 1994b, EX.105). See Taube *et al.* (1996) for a review.

Many of these areas are tightly linked anatomically: The AD nucleus of the ATN and PoS are directly interconnected (van Groen and Wyss, 1990); PoS also sends a strong projection to LMN (van Groen and Wyss, 1990), which in turn sends a strong projection to the AD nucleus (Bentivoglio *et al.*, 1993). LDN and PoS are also interconnected (van Groen and Wyss, 1990). The posterior parietal cortex receives input from LDN and primary visual cortex and sends efferents to PoS, as well as the cingulate cortex (van Groen and Wyss, 1990; Kolb, 1990a), while posterior cingulate is interconnected with both ATN and PoS directly (van Groen and Wyss, 1990; Wyss and van Groen, 1992; Bentivoglio *et al.*, 1993).

There are differences between these areas which can be used to narrow down the possible roles each structure might play in the head direction system. As reviewed in the Experimental Review, PoS, LMN, and ATN head direction cells are sensitive to rotation of distal cues (Taube *et al.*, 1990b; Taube, 1995; Goodridge and Taube, 1995; Leonhard *et al.*, 1996, EX.98), and when two head direction cells have been simultaneously recorded, their tuning curves have always rotated in sync (Taube *et al.*, 1990b; Goodridge and Taube, 1995, EX.97). Cells in LDN, parietal and cingulate cortex, and striatum do not, however, show such a clean tuning to directional cues.

[5] Note that I am only referring to spatial internal cues (such as motor efferent copy or vestibular signals). Spatial internal cues are sometimes called *idiothetic* cues (Mittelstaedt and Mittelstaedt, 1980; Mittelstaedt and Glasauer, 1991).

[6] The firing rates of these cells are not correlated with the angle between the head and the body, but with the orientation of the head relative to the surrounding environment.

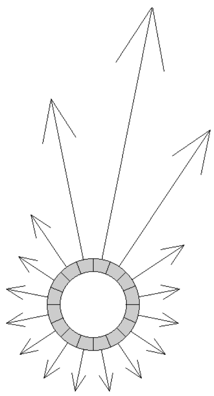


Figure 3.5: Head direction population as a ring. Each shaded segment corresponds to a cell with a preferred direction pointing away from the center. Lengths of the arrows are proportional to firing rate. This is a hand-drawn figure and is meant to be indicative; it does not show neurophysiological recordings or simulations. After a figure from Skaggs *et al.* (1995).

Head direction cells in LDN are not sensitive to the movement of single cues, although they are sensitive to movement of the entire visual world (Mizumori and Williams, 1993, EX.103). They are also dependent on the presence of visual input (unlike ATN and PoS HD cells, compare EX.99). Similarly, cells in parietal and cingulate cortices are sensitive to behavioral cues as well as directional cues (EX.105). They are sensitive, for example, to whether the animal is turning left or right or moving straight, and whether the animal is proceeding inward or outward on a radial maze (Chen *et al.*, 1994b, EX.105). As reviewed in the Experimental Review, some parietal and cingulate cells showed tunings to head direction that had multiple peaks during cue-manipulation tasks (Chen, 1991; Chen *et al.*, 1994a), something never seen in ATN or PoS recordings. LDN and parietal cortex may be better understood as part of the local view subsystem (Section 3.2.1, above) than the head direction subsystem.

When multiple head direction cells have been recorded from ATN, PoS, or LMN the difference between their preferred directions is a constant across all environments (Taube *et al.*, 1990b; Goodridge and Taube, 1995, Taube, personal communication, EX.97). If the population of head direction cells were imagined as a one-dimensional ring with each cell placed in a position on that ring corresponding to its preferred direction (Figure 3.5), then at any time the population would always encode a single direction. One can therefore talk about the *precession of the head direction representation as a whole.*

One way of interpreting the activity of these cells is as a distributed representation of the rat's current head direction. A population of HD cells with preferred directions $\tilde{\phi}_i$ evenly distributed through 360° represents the direction of the weighted vector sum $\sum_i F_i \cdot \vec{v}_i$, where $F_i$ is the normalized firing rate of cell $i$ and $\vec{v}_i$ is a unit vector pointing in direction

$\tilde{\phi}_i$. This is the weighted circular mean (Mardia, 1972), and is also known as a *population vector* encoding (Georgopoulos *et al.*, 1983). Of course, real head direction cells are not necessarily evenly distributed, but as long as the cells are distributed approximately evenly, this interpretation is still valid (Georgopoulos *et al.*, 1988; Redish and Touretzky, 1994).

Recently, important distinctions have been made between each of the three key areas (ATN, PoS, and LMN): As reviewed in the Experimental Review, ATN cell activity is best correlated not with current head direction, but with future head direction, while PoS head direction cells are best correlated with the animal's recent head direction (Blair and Sharp, 1995; Taube and Muller, 1995, EX.101). Also, LMN cell activity is strongly correlated with angular velocity as well as direction (Leonhard *et al.*, 1996, EX.96).

Tracking head direction accurately requires three components: a means of maintaining a stable representation (Skaggs *et al.*, 1995; Redish *et al.*, 1996; Zhang, 1996a), a means of updating the representation from vestibular input (McNaughton *et al.*, 1991; Skaggs *et al.*, 1995; Redish *et al.*, 1996; Zhang, 1996a), and a means of updating the representation from local view input (McNaughton *et al.*, 1991; Knierim *et al.*, 1995). The key to understanding these three components lies in properties of *attractor networks*, reviewed (with simulations) in Appendix A. The key properties of these networks are (1) that a coherent representation of head direction is a stable state, (2) slightly offset excitatory input forces the representation to precess, and (3) strong, distally offset excitatory input forces the representation to reset to match the offset input.

### 3.2.2-A Maintaining a stable representation

The suggestion that a key component of the head direction subsystem is an attractor network which maintains the stable representation can be traced to Skaggs *et al.* (1995) and the first simulation to Zhang (1996a). Redish *et al.* (1996) suggested the postsubiculum as a likely candidate for the attractor network.

It is not known if the real head direction system includes an attractor network — this is a hypothesis that forms part of the standard model of head direction, see Chapter 5. The attractor nature of the system could be demonstrated with multi-unit recordings: if one records from multiple head direction cells simultaneously (say 20 cells), then the population encoding (Figure 3.5) can be seen directly. If noise was injected into the system by microstimulation, then the representation should be transiently disrupted. Even in darkness, without external or self-motion cues, the population's firing rates should return to a well-formed representation of head direction.

### 3.2.2-B Resetting head direction from local view

Within any environment, the head direction system must represent a consistent orientation — if a cell's preferred direction corresponds to (say) northwest at one point in the environment, then it should always correspond to northwest when the animal experiences that environment. Even though it may correspond to a totally different direction in another environment, it must not change in this environment. That means (1) the head direction representation must not drift during a single session, and (2) it must be reset upon re-entry into the environment. The fact that these are both possibilities can be seen by evidence that (1) the head direction system can drift in animals with disorientation training (Knierim *et al.*, 1995) and (2) that in the absence of an orienting cue card, the head direction system does not reset to the same orientation upon re-entry into a familiar environment (Goodridge and Taube, 1995; Taube and Burton, 1995). One possible solution is an association between local view inputs and head direction orientations (McNaughton *et al.*, 1991).

Postsubiculum is interconnected with both parietal cortex and LDN. I have already discussed the evidence that these structures support the local view subsystem (Section 3.2.1). By correlational LTP occurring between representations of head direction based purely on distal cues ($\theta_i$ in Figure 3.3) and head direction cells, external information about head direction can be associated with actual head direction (McNaughton *et al.*, 1991; Knierim *et al.*, 1995; Skaggs *et al.*, 1995).

External orientation cues are associated with internal (idiothetic) cues, and not the other way around (Knierim *et al.*, 1995). Knierim *et al.* found that the head direction of rats with disorientation training did not follow the cue card as well as control rats. They also found that even in control rats, after a number of probes with rotated cues, the head direction representations did not follow the cues as well as they had previously (EX.98).

As pointed out by Knierim *et al.*, this data implies that the internal cues are primary and the external cues are associated with them. If the external cues were primary and head direction representations were associated with them then one would expect to see head direction representations following the cue card in both disoriented and control rats. This is the opposite of what Knierim *et al.* found. The most parsimonious explanation for Knierim *et al.*'s data is that external cues are associated with head direction representations. Knierim *et al.* suggest that this means that, to the disoriented rats, the cue card was an unstable cue (its orientation changed from trial to trial). Because stable cues are associated with the same directions but unstable cues are not, only stable cues can reset head direction representations.

The mechanism by which this reset can occur was first suggested by McNaughton *et al.* (1991) and clarified by Skaggs *et al.* (1995). The first simulations of this effect were shown by Zhang (1996a).

As shown in Figures A.2 and A.4, sufficiently strong tonic excitatory input added into an attractor network will either slowly rotate the representation until it matches the tonic input or will shift it suddenly. Which of these two processes occurs depends on the angular difference between the current representation and the input direction. Note that, according to this theory, the input into PoS from local view is not "orientation of the distal cue" but rather "head direction in the environment as informed by the distal cue."

An important question immediately arises as to what happens when an animal enters a cue-conflict situation. As reported by Taube and Burton (1995, EX.98) animals exposed to a conflict between internal and external representations sometimes reset head direction on entry into the new environment and sometimes do not. Sometimes, they shift part-way towards the reset, but not all the way. Note that whenever more than one cell was recorded simultaneously, the representation shifted by the same amount. Thus the reference direction shifted but the system still maintained a coherent representation of head direction.

#### 3.2.2-C Updating the representation

From the earliest recordings of head direction cells, researchers believed that internal cues (i.e. vestibular and self-motion cues) were being used to update the head direction representation (Ranck, 1984; Taube *et al.*, 1990a; Taube *et al.*, 1990b). The first hypothesis for how this could happen was suggested by McNaughton *et al.* (1991), clarified by Skaggs *et al.* (1995), simulated by Zhang (1996a), and shown to be able to track accurately and understood anatomically by Redish *et al.* (1996).

The first key hypothesis is that there is a population of cells which receives input from both the head direction representation and vestibular input and represents the product of angular velocity and head direction (McNaughton *et al.*, 1991). I shall follow Blair (1996) and refer to these as AVHD cells. (They were called $\mathcal{HH}'$ cells by McNaughton *et al.*, 1991.) The second key hypothesis is that these AVHD cells project to cells in the head direction representation with slightly offset connections, where the direction of offset is dependent on the preferred angular velocity (clockwise or counter-clockwise) (Skaggs *et al.*, 1995, see also Elga *et al.*, 1996; Redish *et al.*, 1996; Zhang, 1996a).

Given the recent data from Leonhard *et al.* (1996, EX.100), the most parsimonious theory is that the AVHD cells are in two head direction populations in LMN, differentiated by whether they increase their firing with left rotations or with right rotations. A left-offset cell with preferred direction $\tilde{\phi}_i$ would send projections to ATN cells with preferred direction counter-clockwise to $\tilde{\phi}_i$, while a right-offset cell with preferred direction $\tilde{\phi}_i$ would send projections to ATN cells with preferred direction clockwise to $\tilde{\phi}_i$.

This architecture entails that the locations of the represented directions PoS, LMN, and ATN will be synchronized during periods of no rotation, but during rotations ATN will lead PoS (Redish *et al.*, 1996). Furthermore, the amount of lead in ATN will depend on the angular velocity of the rotation (Redish *et al.*, 1996).

A prediction of this model, originally pointed out in Redish *et al.* (1996) is that the tuning curves of ATN cells will change with angular velocity. As reviewed in the Experimental Review, (EX.102, see Figure 2.24), this prediction has been recently confirmed by Blair *et al.* (1997), who found that during non-rotations, ATN cells show two peaks, and during left- and right- rotations, one of those peaks grows. Blair *et al.* hypothesize that the two peaks seen in the non-rotation period are indicative of the left- and right-offset connections.

#### 3.2.2-D Relation to the lesion data

Lesions to ATN cause a disruption of directional selectivity in the PoS head direction population (Goodridge and Taube, 1994, EX.106). The theory argued for here is compatible with this result because without the offset connections, the representation of head direction in PoS can not be updated.

Stackman and Taube (1997, EX.107) have shown that with vestibular lesions, the ATN HD cells are no longer correlated with head direction. However, it is not clear whether the ATN HD population still shows a coherent representation of head direction that has been decoupled from the real world, or whether the representation itself is disrupted. Multi-unit recordings could differentiate these two possibilities. According to this theory, as long as PoS and LMN were still intact, the ATN population should still contain a coherent representation of head direction, but the representation should be decoupled from rotations made by the animal.

Unfortunately, postsubiculum lesions do not seem to disrupt the directional selectivity of AD cells (Goodridge and Taube, 1994; Taube *et al.*, 1996, EX.108). This issue will be discussed in depth in Chapter 5.

Golob and Taube (1994, EX.110) examined lesions of the lateral dorsal thalamus (LDN) and found that they did not affect the directional selectivity of PoS head direction cells. One possibility is that although LDN is interconnected with PoS (van Groen and Wyss, 1990), the local view information can also enter the system via parietal cortex which receives input from primary visual cortex, LDN, LPN, and superior colliculus. This means that a combined LDN/parietal lesion may be necessary to affect PoS sensitivity to visual cues.

There is no published data on LMN lesions or on how LDN lesions affect ATN cells or how parietal or cingulate lesions affect the head direction system.

### 3.2.3 Path integration ($\mathcal{PI}$)

Path integration is the ability to return directly to a starting point from any location in an environment, even in the dark or after a long circuitous route (Barlow, 1964; Gallistel, 1990; Maurer and Seguinot, 1995). Sometimes called "dead reckoning," this ability has been shown in gerbils (Mittelstaedt and Mittelstaedt, 1980; Mittelstaedt and Glasauer, 1991), hamsters (Etienne, 1987; Etienne, 1992; Chapuis and Scardigli, 1993), house mice (Alyan and Jander, 1994), rats (Tolman, 1948; Alyan *et al.*, 1997), birds (Mittelstaedt and Mittelstaedt, 1982; von Saint Paul, 1982), and even insects (Wehner and Srinivasan, 1981) and arthropods (Mittelstaedt, 1983), as well as dogs, cats, and humans (Beritashvili, 1965). See EX.57.

The path integration ability has been the subject of argument for more than a century, including a notable debate in 1873 between Alfred Wallace and Charles Darwin in which Wallace suggested that animals find their way back via sequences of smells and Darwin argued that animals must be using dead reckoning. See Wallace (1873a; 1873b), Darwin (1873a; 1873b), Nature (1873), Forde (1873), and Murphy (1873). The carefully controlled experiments of Mittelstaedt and Mittelstaedt (1980) and Etienne (1987) have demonstrated conclusively that this ability is a consequence of integrating internal cues from vestibular signals and motor efferent copy.

Path integration has been used in ship-board navigation for thousands of years. Polynesian navigators used path integration techniques to cross the pacific ocean over distances

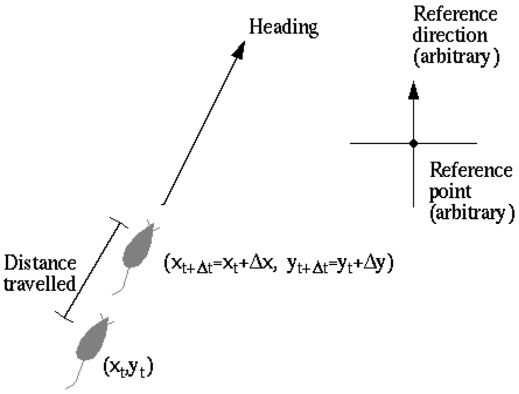



Figure 3.6: Path Integration: If the animal has a representation of its position at time $t$, $(x_t, y_t)$, and representations of its speed and heading, then it can calculate its position at time $t + \Delta t$.

of thousands of miles with no land in sight (Gladwin, 1970; Oatley, 1974). Even as late as the 18th century, European navigators were still using dead reckoning to determine longitude (Navigation, Encyclopedia Britannica, 1994). (Latitude can be determined by azimuth of the North star in the northern hemisphere and by azimuth of known southern stars in the southern hemisphere.) The basic idea of path integration is shown in Figure 3.6: if one knows where one's location at time $t$ as well as one's speed and direction at time $t$ then position at time $t + \Delta t$ can be calculated. Modern submarines can navigate under the polar ice cap using dead reckoning with errors of less than a mile per week (Navigation, Encyclopedia Britannica, 1994). The main problem with path integration is that if your measurement of speed and direction are wrong, your representation of position will be increasingly inaccurate.

Note that because errors in the path integrator can be corrected from local view information, there may be some systematic error in the path integrator which is corrected by external cues. Maurer and Seguinot (1995) review data showing that systematic errors in path integration do exist in a variety of species. However, a path integrator that drifts too much will be useless.

A number of models of path integration have been introduced (Jander, 1957; Mittelstaedt, 1962; Mittlestaedt, 1985; Müller and Wehner, 1988; McNaughton and Nadel, 1990; Touretzky *et al.*, 1993; McNaughton *et al.*, 1994a, see Maurer and Seguinot, 1995 for a review), but the currently most reasonable model of path integration is a two-dimensional analogue of the one-dimensional head direction system (McNaughton *et al.*, 1996; Zhang, 1996a; Samsonovich and McNaughton, 1997; Samsonovich, 1997): Each cell has a *preferred location* and shows a monotonic Gaussian fall-off in firing rate as distance from that location increases. A second population represents the product of location $\times$ speed $\times$ direction and a third population receives offset connections from the second population. I will review all of the models in depth in Chapter 6.

There has been a recent debate about the anatomical instantiation of the path integrator, particularly about whether the hippocampus is directly involved or not. McNaughton *et al.* (1996) have suggested that it is, while Sharp (1996b) has suggested that the key anatomical structure is the subiculum, and Redish and Touretzky (1997a) have suggested a three stage loop of subiculum, parasubiculum, and superficial entorhinal cortex. Recent lesion data from Alyan *et al.* (1997) showing that animals can perform path integration tasks even after ibotenic lesions of the hippocampus suggests that the hippocampus is not necessary for path integration (EX.59). I will discuss these anatomical hypotheses in detail in Chapter 6.

### 3.2.4 Place Code ($\mathcal{PC}$)

In order to navigate within a familiar environment, an animal must use a consistent representation of position from session to session. Although visual cues can serve to inform the animal of its initial position, they may be ambiguous or incomplete. Therefore there must be a mechanism to settle on a consistent representation of location, even given an ambiguous or incomplete local view. The evidence suggests that the *place cells* of the hippocampus (see Experimental Review, Section 2.2.2-A) are well-suited for this role.

#### 3.2.4-A Anatomy

The hippocampus is a bilateral C-shaped structure, with one end at the septal regions dorsally (just ventral to parietal cortex) and the other at the amygdala ventrally (just medial to the entorhinal cortex) (Amaral, 1993, see Figure 3.7). This makes it a very easy target for recording electrodes; it is both large and accessible. However, it is the internal structure of the hippocampus that has made it such a treasure trove for experimental work.

The structure of the hippocampus is most easily understood by looking at a laminar transverse slice. Under a Nissl stain (showing cell bodies, see Figure 3.8), the hippocampus shows two interlocking Cs, the tightly packed granule cells of the dentate gyrus (DG) and the pyramidal cells of the CA3 and CA1 regions.

Figure 3.8 shows only a transverse, laminar slice, but the hippocampus is a three-dimensional structure and its intrinsic connections follow a complex three-dimensional pattern (Amaral, 1993). From a gross anatomical viewpoint, the three major subareas of the hippocampus (DG, CA3, CA1) form a straightforward feed-forward loop. The hippocampus receives its primary input from the superficial layers of the entorhinal cortex (via the *perforant path*) and contributes its primary output via two pathways: the *angular bundle* projecting to the subiculum and the deep layers of entorhinal cortex, and the *fornix*, which projects to a variety of subcortical structures, most notably the nucleus accumbens. The fornix takes most of its input from subiculum, but some CA1 cells also contribute to it, see Amaral and Witter (1989) and Witter (1993) for reviews.

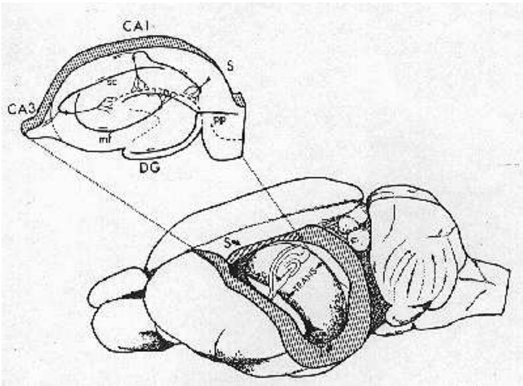



Figure 3.7: The location of the hippocampus in the rodent brain. DG: dentate gyrus; S: subiculum; pp: perforant path; mf: mossy fibers; sc: Schaffer collaterals. From Amaral and Witter (1989). Reprinted by permission of the author.

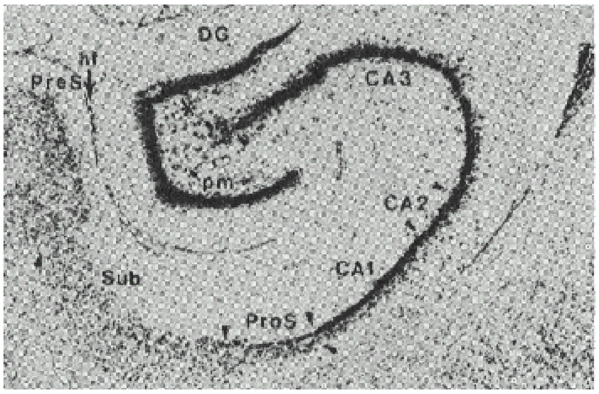



Figure 3.8: Nissl staining of the hippocampus. DG: dentate gyrus; ProS: prosubiculum, Sub: subiculum, PreS: subiculum; hf: hippocampal fissure. From Cohen and Eichenbaum (1993), reprinted with permission of author and publisher.

The perforant path consists of fibers projecting from the superficial layers (layers II and III) of the entorhinal cortex to the dentate gyrus, the CA3, CA1, and subicular fields, see Jones (1993) and Witter (1993) for reviews. DG projects (via the *mossy fibers*) to CA3 and CA3 projects (via the *Schaffer collaterals*) to CA1. CA3 also has recurrent connections, in which collaterals of the projection to CA1 also synapse on CA3 cells. CA1 cells project to the subiculum and to the deep layers of entorhinal cortex, see Amaral and Witter (1989) and Witter (1993) for reviews. Subiculum projects to a variety of structures, including the post- and parasubiculum, the septal regions, nucleus accumbens, and the cingulate and frontal cortices (Witter, 1993).

The hippocampus also receives non-glutamatergic input from a variety of subcortical structures, including cholinergic (Ach) and GABA-ergic (GABA) input from the septal regions (see Haas, 1983; Cooper *et al.*, 1986; Hasselmo and Bower, 1993; Mizumori *et al.*, 1992; Stewart and Fox, 1990; Gallagher and Colombo, 1995; Gallagher *et al.*, 1995 for reviews), noradrenergic (NA) input from the locus coeruleus (see Haas, 1983; Cooper *et al.*, 1986; Decker and McGaugh, 1991 for reviews), and serotonergic (5-HT) input from the raphe nucleus (see Haas, 1983; Cooper *et al.*, 1986; Decker and McGaugh, 1991; Freund and Buzsáki, 1996 for reviews).

### 3.2.4-B Neurophysiology

As described in the Experimental Review (EX.63), the first-order correlate of spikes fired by hippocampal CA3 and CA1 pyramidal cells is the location of the rat: each cell fires when the animal is in a specific place (called the *place field* of the cell). These *place cells* are some of the most-studied neurons in the rodent brain; they have been examined in a wide variety of environmental manipulations. The hippocampus is also probably the most extensively modeled system in the rodent brain. In this section, I will review the key experimental results and their explanations. I will then put together an amalgamated model compatible with each of the previous suggestions before moving on to discuss the role of the place code in navigation.

**Place fields move when distal landmarks move (Muller and Kubie, 1987; O'Keefe and Speakman, 1987; McNaughton *et al.*, 1994a; Knierim *et al.*, 1995, EX.67).** The typical explanation of this effect is that place fields are partially generated by integrating local view information (see, for example, the models described in Section 7.2.1). However, this cannot be the whole story because of other properties of place cells (see below).

**Some place cells show crescent-shaped fields along the wall of circular arenas (Muller *et al.*, 1987, EX.65).** There are three possible explanations for this result:

1. These cells might be sensitive to distal landmarks, such that the center of the place field is external to the maze and only a small edge of the place field is accessible to the animal (Sharp, 1991).

2. Place fields could be small, compact (Gaussian-shaped) fields, but they could drift in orientation (due to drift in the head direction system) (B. McNaughton, personal communication).

3. Place cells could be sensitive to the surface orientation of the wall (Touretzky and Redish, 1996).

The first mechanism is unlikely to be correct because it cannot generate place fields with concave sides — only convex, and the interior edges of these cells are concave (see e.g. Figure 2.20).

The second mechanism implies (1) that the place fields should appear as small, compact, convex fields when the time over which the measurements are integrated is small, and (2) that simultaneously recorded place fields with locations near the wall should all "stretch" in synchrony. These experiments have not yet been done. However, the second mechanism is unlikely to be correct because it also predicts that cells away from the wall would occasionally form arcs at a constant distance from the wall. No such cells have ever been reported.

The most parsimonious explanation is the third which fits the available data. An alternative description of the third mechanism has been suggested by McNaughton (B. McNaughton, personal communication): the sensory cue provided by the wall could become associated with place cells representing a specific location and then the cue "drags" the activity packet, however, this is just another way of saying that the cells at that location are "sensitive" to the wall.

**Place cells are sensitive to displaced, but not to missing, landmarks (O'Keefe and Conway, 1978; Pico *et al.*, 1985; Muller and Kubie, 1987; O'Keefe and Speakman, 1987, EX.70).** There are three possible explanations for this result:

1. Cells are tuned to many different landmarks (O'Keefe and Nadel, 1978; Sharp, 1991). If a place cell was tuned to more than one landmark then even if some of the landmarks were removed, the others might be sufficient to drive the place cells.

2. The hippocampus shows associative-memory properties (Marr, 1971; McNaughton and Morris, 1987; McNaughton, 1989; McNaughton *et al.*, 1989b; Rolls, 1989; Leonard and McNaughton, 1990; McNaughton *et al.*, 1991; Hetherington and Shapiro, 1993; Shapiro and Hetherington, 1993; Recce and Harris, 1996; Rolls, 1996). If there is enough information to inform some of the place cells of their location, and the hippocampus includes associative memory properties, then place cells normally "tuned" to the missing landmarks can be driven by the internal connections.

3. The place code is partially informed by path integrator information (O'Keefe, 1976; Wan *et al.*, 1994c; McNaughton *et al.*, 1996; Touretzky and Redish, 1996; Redish and Touretzky, 1997a; Samsonovich and McNaughton, 1997). According to these theories, place cells are driven by internal not external cues during normal navigation. It is

the role of external cues to reset the internal cues. This means that if a landmark is moved, it will reset the internal path integration representation, but if a landmark is missing entirely, the place cells will still be driven to show place fields by the internal cues.

I will argue in Chapter 7 that the third mechanism is the most parsimonious in light of the other data.

**Place cells continue to show compact fields in the dark (O'Keefe, 1976; McNaughton *et al.*, 1989b; Quirk *et al.*, 1990; Markus *et al.*, 1994, EX.71).** One possible explanation is that these cells are sensitive to non-visual cues. (Even in the dark, auditory, olfactory, and somatosensory cues are still available.) Alternatively, place cell firing in the dark may be a consequence of input from the path integrator. We can differentiate between these two hypotheses because the non-visual hypothesis implies that place fields will not change at all, but the path integrator hypothesis implies that place fields may drift in the dark.

Quirk *et al.* (1990) report that when the animal is placed into the arena in the light and the lights are extinguished, the place fields rarely change. But when they do change, they continue to show the same shape and distance from the arena wall; it is only their orientation that varies. This suggests that the cells are dependent on input from the head direction system to orient themselves. Since some of these cells are centered far from the arena wall, the wall itself is not always an available cue in the dark. Therefore, place cells in the dark must be driven at least partially by path integration information.

**Place cells are directional when an animal traverses repeated paths, but not when wandering randomly over open arenas (Muller *et al.*, 1987; McNaughton *et al.*, 1983a; Markus *et al.*, 1994; Muller *et al.*, 1994; Gothard *et al.*, 1996a, EX.64).** There are two important issues here: (1) why do cells show directional place fields when traversing repeated paths, and (2) why do cells not show directional place fields when wandering randomly over open fields?

An early suggestion and model was presented by Sharp (1991) who suggested that competitive learning combined with a limited local view (covering only 300°) would produce directional place fields on linear tracks and radial mazes but not while wandering randomly on open fields. However, this model predicted that place fields would not be directional on the central dais of a radial maze. Markus *et al.* (1994) report that they are. In addition, the Sharp model requires directionality to develop over time, but place cells can show directional tuning from the first traversal of a linear track (B. McNaughton, personal communication).

Another suggestion was made by McNaughton *et al.* (McNaughton *et al.*, 1994a; Markus *et al.*, 1995): that on linear tracks, animals attended to specific landmarks which were different for each direction, but when wandering randomly on open, animals shifted attention between many landmarks continuously, so the place fields appeared non-directional.

Finally, a suggestion was made by Wan *et al.* (1994c, see Touretzky and Redish, 1996; Redish and Touretzky, 1997a), that different *reference frames* were being used for each direction on the linear track, but on the open field, a single reference frame was being used. See Section 3.3 for an in-depth discussion of reference frames. A similar proposal has now been put forward by McNaughton *et al.* (McNaughton *et al.*, 1996; Samsonovich and McNaughton, 1997; Samsonovich, 1997).

I will argue that this last explanation is the most parsimonious with the other data.

**Place fields are seen on initial entry into an environment (Hill, 1978; Austin *et al.*, 1990; Wilson and McNaughton, 1994; Tanila *et al.*, 1997b, EX.72).** The fact that place fields appear on initial entry into an environment implies that some of the connections in the hippocampus are "pre-wired" (i.e. present before entry). Early models assumed that local view input was pre-wired and then learning might modify those connections slightly (but not strongly) (e.g. Zipser, 1985; Zipser, 1986; Sharp, 1991). Samsonovich and McNaughton (1997) have argued that the data implies that the intra-hippocampal connections must be pre-wired, but I will show (Chapter 10) that this is not necessary. I will posit pre-wired connections into the dentate gyrus and between DG and CA3, with learning occurring between EC and CA3 and within CA3.

**Place cells are sensitive to more than place, EX.78.** As discussed in the Experimental Review, place cells in a number of tasks have been found to be correlated with non-spatial aspects including speed, direction, and turning angle (McNaughton *et al.*, 1983a; Wiener *et al.*, 1989; Markus *et al.*, 1994), texture underfoot (Young *et al.*, 1994), odor (Eichenbaum *et al.*, 1987; Eichenbaum and Cohen, 1988; Otto and Eichenbaum, 1992; Cohen and Eichenbaum, 1993), task (Markus *et al.*, 1995), and stage of task (Eichenbaum *et al.*, 1987; Otto and Eichenbaum, 1992; Hampson *et al.*, 1993).

However, in all experiments where place cells showing non-spatial correlations but have also been tested for spatial correlations, place cells continue to show spatial correlations. In these cases, if a cell has a place field under one condition, it may or may not show a place field under the other, and if two cells both show place fields under both conditions, then the spatial relationships between them may change drastically from one condition to the other. Thus many of these effects can be said to be second-order effects because they determine how the place fields change.

The main discussion of this will be left for Section 3.3, but it is an extremely important result that must be explained by any hippocampal model.

### 3.2.4-C A place code model

I will not review the large number of theories addressing the role of the hippocampus here (see Chapter 4), nor will I review the large number of place cell models (see Chapter 7). In Chapter 7, I will lay out a model of the place code and the role of the hippocampus that draws selected aspects of previous models and incorporates some novel ideas, and in the subsequent chapters, I will show simulations of specific aspects of the model.

The essentials of the model are

- that the hippocampus receives both path integrator and local view input (O'Keefe, 1976; McNaughton, 1989; McNaughton *et al.*, 1989b; McNaughton *et al.*, 1991; Muller *et al.*, 1991a; McNaughton *et al.*, 1994a; Wan *et al.*, 1994c; McNaughton *et al.*, 1996; Touretzky and Redish, 1996; Redish and Touretzky, 1997a),
- that random, strong connections enter the dentate gyrus, which *orthogonalizes* the input (Marr, 1969; McNaughton and Morris, 1987; Rolls, 1989; Rolls, 1996; O'Reilly and McClelland, 1994),
- that random, strong connections between DG and CA3 drive place cell activity (McNaughton and Morris, 1987; McNaughton, 1989; Rolls, 1989; O'Reilly and McClelland, 1994),
- that connections between local view input and CA3, between CA3 and CA1, and between CA1 and the path integrator consist of synapses that show correlational Hebbian LTP.
- Finally, this model includes three different hippocampal modes: *storage*, *recall*, and *replay*. I will make the case that storage occurs during theta, recall occurs during awake LIA, and replay during sleep-state LIA.[7]

#### 3.2.4-D Self-localization (recall)

Touretzky and Redish (1996) argued that when the animal returns to a familiar environment the role of the place code is to reset the path integrator representation so that the animal could use the same coordinate system from experience to experience in an environment. In a sense, the place code *recalls* the previous representation given the local view inputs.

This concept has much in common with a number of other hippocampal navigation models that hypothesize that the hippocampus served as an associative memory (e.g. McNaughton, 1989; Leonard and McNaughton, 1990; Recce and Harris, 1996, see also McNaughton *et al.*, 1996; Samsonovich and McNaughton, 1997; Samsonovich, 1997), but these other models are all *online* models in that they assume the hippocampus is constantly using the local view to check and reset the hippocampal representation. The differences

[7] A number of authors have suggested that hippocampus shows two modes: *storage* and *replay* (Marr, 1970; Marr, 1971; Buzsáki, 1989; Pavlides and Winson, 1989; Hasselmo and Schnell, 1994; Wilson and McNaughton, 1994; Hennevin *et al.*, 1995; McClelland *et al.*, 1995; Smith, 1995; Kudrimoti *et al.*, 1996; McClelland and Goddard, 1996; Skaggs and McNaughton, 1996; Shen and McNaughton, 1996), and that storage occurs during theta, while replay occurs during LIA (Buzsáki, 1989; Wilson and McNaughton, 1994; Smith, 1995; Kudrimoti *et al.*, 1996; Skaggs and McNaughton, 1996; Shen and McNaughton, 1996, see also comments by McNaughton *et al.* in Seifert, 1983, p. 610).

between these models and the model being argued for here will be discussed in depth in Chapter 7.

The idea that the hippocampus is only necessary to re-instantiate context (i.e. that the recall process occurs on a *context-switch*) has much in common with ideas of *contextual-retrieval* (Hirsh, 1974; Nadel and Willner, 1980; Rawlins, 1985; Nadel, 1994; Nadel, 1995, see also Rotenberg *et al.*, 1996). None of the these authors, however, have suggested explicit mechanisms for the recall mechanism. The similarities and differences between these theories will be discussed in Chapter 4, see, in particular, Section 4.1.6.

Touretzky and Redish (1996) proposed a complex algorithm to explain the self-localization process, however a local-excitation, global-inhibition structure would suffice (Redish and Touretzky, 1997a; Samsonovich, 1997; Samsonovich and McNaughton, 1997). In the latter half of this thesis, I will discuss detailed aspects of this self-localization process and report simulations demonstrating its viability and relation to explicit experiments (see Part II, Chapters 7 and 8, and Part III, Chapters 10, 13, 14).

The key point is that local excitation in location space forms from random navigation (i.e. cells that represent nearby locations in the same environment[8] are more strongly connected than cells that represent distant position, Muller *et al.*, 1991b; Muller *et al.*, 1996, see also Wilson and McNaughton, 1994). A local-excitation/global-inhibition structure has stable states that represent coherent locations in space (Kohonen, 1982; Kohonen, 1984). In Chapters 7 and 8, I will argue that candidate locations provided by the local view are enough to force a noisy hippocampus to settle to a coherent representation of location within the time course of a single sharp wave.

### 3.2.4-E Route-learning (replay)

In Section 3.4, below, I will discuss the replay aspect of the place code by which asymmetries in the recurrent connections within the hippocampus form. The basic idea that the hippocampus stores temporary memories can be traced back to Marr (1970; 1971, but see also Pavlides and Winson, 1989; Wilson and McNaughton, 1994; Shen and McNaughton, 1996; Skaggs and McNaughton, 1996). The key point is that asymmetries in the recurrent connections in CA3 form because of specific properties of LTP (LTP is correlational, LTP is asymmetric) and specific neurophysiological properties (place fields have non-zero extent, phase precession in the place code). See Levy (1989); Abbott and Blum (1996); Blum and Abbott (1996); Levy (1996); Shen and McNaughton (1996). See also Section 3.4 below for an in-depth discussion of this route-learning and replay process.

These two modes (*recall* and *replay*) would seem to require incompatible connection matrices: *recall* requires symmetric connections, while *replay* requires asymmetric connections. I will argue in Chapter 9 that these requirements are not incompatible and will show simulations supporting the dual-role hypothesis.

[8]Actually in the same reference frame, see Section 3.3, below.

#### 3.2.4-F Aligning local view and path integrator representations in the place code

During normal navigation, both the path integrator and local view subsystems will provide excitatory input onto the same place cells. When this happens, we say that the representations are *consistent* with each other. However, when these representations become misaligned, it is the role of the place code to make them consistent again. This can be done either by resetting the path integrator or by changing the $\mathcal{LV} \times \mathcal{PI}$ association stored in the place code. The former corresponds to the *self-localization* process (*recall*), and the latter corresponds to deciding that the environment is novel which requires exploration (*storage*).

When an animal is placed into a familiar environment by an experimenter, the path integrator will presumably be tracking (if it is still tracking) the animal's position relative to the previous environment. This means that the $\mathcal{LV} \times \mathcal{PI}$ association will be inconsistent. If we assume that the environment is familiar and the animal is unfamiliar with the spatial relationship between the two environments, then the animal must self-localize. (I address how the animal decides whether the environment is novel or familiar later in this section.) In many experimental paradigms, animals are disoriented before being placed into an environment; in other paradigms, they are placed into a random location in the environment. If either of these cases is true then the animal has no information except local view with which to reset the path integration system. On the other hand if the animal is always put into the environment at the same location, then the animal could use additional information by priming the path integrator before the self-localization process. In a sense, all this does is provide additional candidate biases to the self-localization process.

What happens when the animal doesn't change environments, but still finds an inconsistency between local view and path integrator? Gothard *et al.* (1996a) trained rats to run back and forth on a linear track. Then they shortened the track (see Figure 2.19). Because they recorded from multiple place cells simultaneously, they could observe how the place code handles inconsistencies in the local view and path integrator. They found two effects: either the place code precessed faster along the track than the animal ran, so that the represented position eventually caught up with the animal; or the place code jumped to match the actual position. As pointed out by Gothard *et al.* (1996a, see also Samsonovich and McNaughton, 1997; Samsonovich, 1997), these two effects correspond to the two possible modes shown by the attractor networks in Appendix A: either the system precesses when the two inputs represent similar (but not identical) positions (e.g. Figure A.2) or it jumps when the two representations are very different (e.g. Figure A.4).

**Novel vs. familiar environments.** In the Gothard *et al.* (1996a) experiment, the animal has not re-encoded the $\mathcal{LV} \times \mathcal{PI}$ association, i.e. it has not decided the environment is novel. (We know this because the many of the same place cells appear in the same topological relation to each other. If the environment had been treated as novel, the topology of the place fields would have changed dramatically.) But there are experiments in which inconsistency is resolved by treating the new environment as novel (e.g. Kubie and Ranck, 1983; Muller and Kubie, 1987; Thompson and Best, 1989; Bostock *et al.*, 1991). When does an animal resolve the inconsistency by changing the $\mathcal{PI}$ representation and when does it resolve the inconsistency by treating the change as a novel environment?

An animal that treated an environment as novel whenever a single cue appeared out of place would never see a familiar environment. On the other hand, if enough has changed that the environment really is novel, the animal will be best served by treating all of the cues it sees as novel as well. This issue can be phrased in terms of *completion* and *separation* and is detailed in Figure 3.9. Sensory information from each environment is assumed to be represented as a distributed pattern of activity over a population of cells. Above some threshold of overlap of the current representation with a remembered representation, the animal will want to complete the new representation based on the old, while below that threshold, the animal will want to separate the two representations as much as possible to prevent memory interference.

This is discussed by McClelland *et al.* (O'Reilly and McClelland, 1994; McClelland *et al.*, 1995; McClelland and Goddard, 1996), who suggest the dentate gyrus is well suited for separation and CA3 for completion. McNaughton and Morris (1987, see also Marr, 1969; McNaughton, 1989; Rolls, 1989; 1996) made similar suggestions that DG separates inputs from EC into orthogonal representations, while the recurrent collaterals in CA3 form an associative memory to complete input representations. It has been shown that the activity of DG cells is much lower than the CA3 and CA1 representations (Jung and McNaughton, 1993, EX.90), and that DG is required for learning spatial tasks, but not for their performance (McNaughton *et al.*, 1989a; Sutherland and Hoesing, 1993, EX.15).

**Neuropharmacology: Acetylcholine (ACh).** When learning a new environment, the recurrent connections within CA3 will still drive the new representation towards an already stored one. This will cause interference between the two representations (Hasselmo and Bower, 1993; Hasselmo, 1993).

As discussed in Section 3.3 and Chapter 7, place cells in CA3 are most strongly connected to other cells that represent similar locations. This means that CA3 can be understood as an associative memory and the self-localization process can be understood as a recall of place and reference frame from an associative memory.

An important problem with auto-associative memories is that if the synaptic transmission continues to occur across recurrent connections when a memory is being stored in the system then incorrect correlations will occur between two memories that share neurons. This is called *interference* (Hasselmo and Bower, 1993). In order to fix the problem of interference, when LTP occurs in the system (i.e. when the system is storing a memory), the recurrent connections should be ineffective.

Hasselmo and Schnell (1994, EX.114) have shown that carbachol (a cholinergic agonist) infused into hippocampal slices reduces the size of the EPSP in CA1 produced by stimulating Schaffer collaterals by 90%, while only reducing that produced by stimulating the perforant path by 40%. Since the Schaffer collaterals are the same axons which form the

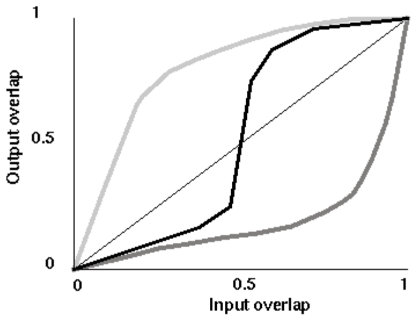



Figure 3.9: Correlation between two output patterns as a function of correlation between two input patterns. Thin line indicates 1:1 mapping. Completion (light grey line): output patterns are more similar than input; separation (dark grey line): output patterns are less similar than input. An optimal balance between separation and completion requires that dissimilar patterns be made more dissimilar while similar patterns be made more similar (heavy black line). This figure does not show simulation or experiment, it is meant to be explanatory only. After O'Reilly and McClelland (1994).

CA3 excitatory feedback pathway, this suggests that ACh may shut off these recurrent connections. Returning to the conceptual framework above, we can say that ACh turns off pattern completion while not affecting pattern separation. ACh also has an effect on learning: although it suppresses synaptic transmission, ACh enhances LTP in DG, CA1, and other structures (see Hasselmo, 1995, for a review). This means that ACh should enhance the learning and separation of place codes.

### 3.2.5 Goal Memory ($\mathcal{GM}$)

In this theory, the role of the place code is thus to determine the location of the animal, given the (possibly) conflicting internal and external cues. But knowing where it is in an environment will not help an animal find food unless it also also remembers where the food is in that same environment. Many of the demonstrations that animals can do navigation in natural settings come from cache behavior (Vander Wall, 1990; Sherry and Duff, 1996), and controlled experiments suggest that they use locale navigation to find their caches (Jacobs and Liman, 1991; Jacobs, 1992). However, if the locale navigation system determines the location of the animal relative to some reference point,[9] this information will not help the

[9] As discussed above, the head direction and path integrator representations do not have predefined correlates with the external environment, but once a direction has been identified with the preferred direction of a head direction cell, the preferred directions of all of the other cells are defined. This means that the head direction representations can be understood as representing the orientation of the animal relative to an

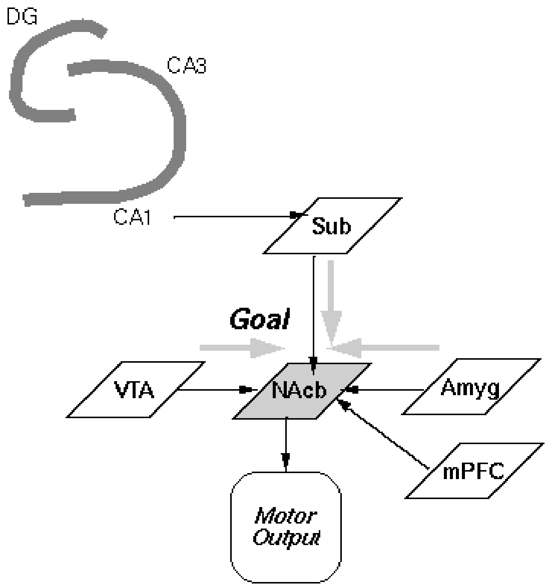



Figure 3.10: Anatomical connections of nucleus accumbens (NAcb). DG: dentate gyrus; CA3, CA1: hippocampus proper; Sub: subiculum; VTA: ventral tegmental area; Amyg: amygdala; mPFC: medial prefrontal cortex.

animal find its cache without a representation of the location of the cache relative to that same reference point. This is the role played by the *goal memory*.

### 3.2.5-A Anatomy

The role of the goal memory is to plan a route to a goal, given information about the position of the animal and the current needs and desires. The structure subserving the goal memory can thus be expected (1) to receive input from the place code or the path integrator, probably both, (2) to receive information about the animals current needs and desires, (3) to be a "motor area," that is, manipulations to it should affect locomotion.

Mogenson (1984) first suggested that the nucleus accumbens (NAcb) combines information from amygdala, VTA, CA1, and subiculum to produce locomotor actions via the subpallidal regions. Injections of dopaminergic agonists such as amphetamine produce excessive locomotion (Isaacson, 1974; Whishaw and Mittleman, 1991), while injections of glutamate antagonists such as l-glutamic acid diethyl ester HCl (GDEE) reduces locomotor

arbitrary reference direction. An analogous case can be made for the path integrator representing position relative to an arbitrary reference point. See discussion of $\mathcal{HD}$ (Section 3.2.2) and $\mathcal{PI}$ (Section 3.2.3) above.

activity produced by injections of carbachol (an ACh agonist) into hippocampus (Mogenson and Nielsen, 1984).

The nucleus accumbens is one of the major targets of the fornix (Witter *et al.*, 1990), which carries place information from subiculum and CA1. The amygdala and hypothalamus have been implicated in emotion and goal information by a number of studies (see Aggleton, 1993, and Davis *et al.*, 1994, for reviews), and NAcb receives afferent fibers from amygdala (Aggleton, 1993; Davis *et al.*, 1994; Finch, 1996), the ventral tegmental area (Wolske *et al.*, 1993), and the hypothalamus (Domesick, 1981). NAcb also receives afferents from the medial prefrontal cortex (Sesack *et al.*, 1989; Finch, 1996), which may supply contextual information (Kolb, 1990b). Finch (1996) has shown that many of these inputs (amygdala, hippocampal formation, medial prefrontal) all converge on single neurons in the accumbens. Additional studies have implicated NAcb as a locomotor structure (Jones and Mogenson, 1980; Mogenson, 1984; Mogenson and Nielsen, 1984), and particularly in locale navigation tasks (Annett *et al.*, 1989; Sutherland and Rodriguez, 1990; Seamans and Phillips, 1994; Floresco *et al.*, 1997, see EX.20, EX.47, and EX.48)

As reviewed in the Experimental Review, NAcb cells are correlated to reward and movement parameters (Lavoie and Mizumori, 1994, EX.111). Hippocampal lesions combined with caudate lesions produced a drastic increase in locomotor behavior, while hippocampal lesions combined with accumbens lesions produce a drastic decrease in locomotor behavior (Whishaw and Mittleman, 1991, EX.34). This led Whishaw and Mittleman to suggest that caudate and accumbens are balanced to produce stereotypy and locomotion, respectively, and that hippocampus affects locomotion through accumbens. This is a particularly interesting view when held in comparison to the evidence that caudate is involved in taxon and praxic navigation strategies (Section 3.1, above).

Animals with fimbria-fornix lesions and septal grafts can still learn the hidden platform water maze (Nilsson *et al.*, 1987, EX.26). This suggests that there must be alternate pathways from hippocampus to the motor structures. One candidate may be the posterior cingulate cortex (Sutherland and Hoesing, 1993), see discussion of *routes*, Section 3.4, below. It is not clear what the role of the nucleus accumbens is relative to this alternative pathway, but even the grafted animals are impaired relative to normals (although they are much better than lesioned, non-grafted animals). Further work needs to be done to completely ferret out the role played by the nucleus accumbens in the navigation system, but recent work on dopamine and the striatum in general has opened another door.

#### 3.2.5-B The role of dopamine

A tremendous amount of work has recently been done in the non-human primate examining the role of dopamine and the striatum (caudate nucleus and nucleus accumbens). Because there is not an equivalent literature of work done on the rodent, I will quickly review the primate data here, remembering that it is not necessarily true that the role of the striatum in primate and rodent are identical. For further reading, I refer the reader to Houk *et al.* (1995). The anatomical structure of the basal ganglia is rather complex and will not be reviewed here.

The role of dopamine has recently fallen into a coherent picture, based most strongly on results from Schultz *et al.* (Schultz *et al.*, 1995; Schultz, 1997, EX.117). They've shown that dopamine neurons in VTA respond most vigorously to an unexpected reward, and, conversely, that they decrease their firing rates when an expected reward is not given. As suggested by Schultz *et al.*, this means that we can interpret the information represented by these neurons as an error signal, a representation of $R(t) - \hat{R}(t)$, where $R(t)$ is the actual reward received at time $t$, and $\hat{R}(t)$, the expected reward at time $t$.

Many learning theories ranging from *Pavlovian conditioning* (Rescorla and Wagner, 1972) to *backpropogation of error* (Rumelhart *et al.*, 1986b; Hertz *et al.*, 1991) and *reinforcement learning* (Sutton, 1992) require an error signal in order to update an associative strength. Many of these theories work on a variant of the following: weights are updated by a product of three terms: activity (firing rate) of the presynaptic neuron, activity (firing rate) of the postsynaptic neuron, and an error signal, indicating whether the system should encourage or discourage the pair to fire simultaneously (Schultz *et al.*, 1995). For this to work, the signals have to all meet together, i.e. dopaminergic synapses and glutamatergic synapses have to meet in a three-way pair with dendritic spines. Sesack and Pickel (1990) have shown by looking at double-stainings of anterograde stains (horseradish peroxidase) injected into subiculum and dopamine-sensitive stains (tyrosine hydroxylase) injected into accumbens that they do.

#### 3.2.5-C Path planning and obstacles

An important role of the goal memory is to plan a trajectory from the animal's current position to a goal. If there are obstacles in the way or if the animal must take a complex path to reach the goal, the goal memory should be able to plan these kinds of complex trajectories. Rodents can learn extremely complex trajectories such as the mazes in Figure 2.17.

It is not currently known how obstacles are represented in the rodent brain or how the planning is done to avoid them. A number of robotics algorithms have been developed to avoid obstacles or plan complex paths (such as *potential field navigation*, Khatib, 1986; Connolly *et al.*, 1990; Tarassenko and Blake, 1991; Connolly and Grupen, 1992; *occupancy grids*, Moravec, 1988; *sinusoidal transforms*, Pratt, 1991a; 1991b; and *graph search*, Muller *et al.*, 1991b; 1996; see also Trullier *et al.*, 1997, for a review of some additional proposals), but they do not lend themselves easily to neural implementations, and so will not be reviewed here.

It is not even clear what brain structures subserve these roles. Because of the lack of data I will not speculate on how or where obstacle avoidance is handled by the rodent brain. Throughout this thesis, I will stick to navigation tasks in simple arenas in which a path can be planned from start to goal by a simple vector subtraction. There is still plenty of data to choose from, including tasks done in open fields, small enclosures, water mazes, etc. Even the radial maze can be handled this way if the simulation always follows the wall in the best approximation to the direct bath to the goal.

## 3.3 Reference Frames ($\mathcal{RF}$)

Place cells allow the system to self-localize and thus to resolve ambiguous local views; this allows the animal to plan a path to a goal even when some cues have been moved or even removed entirely from the environment. This suggests that anytime the $\mathcal{LV} \times \mathcal{PI}$ association changes significantly, the place code will have to change entirely. In other words, place cells have to be sensitive to environment. They are (O'Keefe and Conway, 1978; Kubie and Ranck, 1983; Muller and Kubie, 1987; Thompson and Best, 1989, EX.79).

But place cells are also sensitive to a host of other factors, including direction along repeated paths (McNaughton *et al.*, 1983a; Markus *et al.*, 1995), task (Markus *et al.*, 1995), and subtask (Eichenbaum *et al.*, 1987; Otto and Eichenbaum, 1992; Cohen and Eichenbaum, 1993; Hampson *et al.*, 1993; Gothard *et al.*, 1996a; Gothard *et al.*, 1996b). In this section, I will argue that place cell sensitivity to all of these non-spatial factors (environment, task, etc.) can be understood as realizations of a single effect: that of changing *reference frame*.

When we say cells are "sensitive to a non-spatial aspect" (such as task), what we mean is that if a cell has a place field under one condition, it may or may not show a place field under the other, and if two cells both show place fields under both conditions, then the spatial relationships between them may change drastically from one condition to the other. Essentially, a cell's place field (in fact whether it has a place field at all) is independent of its field in other reference frames.

Note how this is different from the head direction (Chapter 5) and path integration (Chapter 6) representations. Although the *reference orientation* (for the HD system) or the *reference point* (for the PI system) may change from environment to environment, the topology of the representation does not change. Whenever two PoS or ATN HD cells have been recorded simultaneously, the difference between their preferred directions never changes, even though the actual preferred directions may rotate when distal cues are moved (Taube *et al.*, 1990b; Taube, 1995; Taube *et al.*, 1996, EX.97). Similarly, the place fields of ECs cells do not change between two similar environments (Quirk *et al.*, 1992, EX.91), nor do the place fields of subiculum cells (Sharp, 1996b, EX.92), unlike hippocampal CA3 and CA1 place fields (O'Keefe and Conway, 1978; Kubie and Ranck, 1983; Muller and Kubie, 1987; Thompson and Best, 1989, EX.79).

### 3.3.1 Definition

Together a *reference point* and a *reference orientation* define a coordinate system.[10] With the addition of a *mental set*, derived from the goals or activities associated with this coordinate system, these components form a *reference frame*. Each location in each reference frame is represented in the hippocampus by a different set of place cells; the place code in

[10] One should probably include a *distance metric* as well (B. McNaughton, personal communication, see Samsonovich, 1997).

hippocampus represents, not just location, but *location within a reference frame.*[11]

From the earliest theoretical descriptions of the *cognitive map hypothesis* (O'Keefe and Nadel, 1978), there was the idea that different environments would be encoded by different *maps.* O'Keefe and Nadel pointed out that the cognitive map had to be internally consistent: if some of the cues moved while others did not, the map should rigidly translate and/or rotate to follow some consistent subset of cues. If too many of the cues changed, a new map would be needed.

One way to think of this is that in each reference frame there are a different set of potential place fields in the hippocampus. For example, Muller and Kubie (1987) have suggested that different environments have different *active subsets* of place cells: this active subset forms a set of potential place fields from which the specific cells active at each location in the environment are drawn.

McNaughton *et al.* (McNaughton *et al.*, 1996; Samsonovich and McNaughton, 1997; Samsonovich, 1997) have suggested that there are a set of *charts* in the hippocampus, which are used for different environments. In fact, McNaughton *et al.* suggest that the charts are pre-wired in the hippocampus and that external cues are associated with representations of location on a chart. Thus, in their model, the hippocampus represents not location per se, but *location within a chart.*

I prefer, instead, to think of the issue of *reference frame* as a property of the whole system and not just the hippocampus. For example, if the path integrator *reference point* changes, then the $\mathcal{PI} \times \mathcal{LV}$ association will change and the place code will have to change to accommodate it.[12] In Chapter 10, I will show that one explanation of the recent results of Barnes *et al.* (1997, EX.83) is that the reference point is reset to the incorrect value in old animals which changes the $\mathcal{PI} \times \mathcal{LV}$ association. Thus the animal must use a new map in the hippocampus.

A number of authors (McNaughton *et al.*, 1991; Muller *et al.*, 1991a; McNaughton *et al.*, 1994a; Markus *et al.*, 1995; Recce and Harris, 1996) have suggested that some of the sensitivity of place cells to cues other than visible landmarks can be attributed to changes in attention to *virtual landmarks.* These are locations in space from which distance and bearing information can be derived, just as for real landmarks. However, since virtual landmarks are not tied to perceivable objects, instead of being tracked by the perceptual system, they must be tracked by path integration. Although a reference frame could be said to employ a virtual landmark as the reference point, reference frames also include a canonical orientation, so the two proposals are not equivalent. Questions of how many virtual landmarks can be tracked

[11] Although this does mean that place cells represent more than place, I will continue to refer to them as "place cells" for historical reasons.

[12] What does it mean to say that the "path integrator reference point changes"? As we saw in the path integrator discussion above (see Section 3.2.3), there is no "center" of the $\mathcal{PI}$ representation. But mathematically, the path integrator represents the position of the animal relative to some point in space. Thus what it means neurophysiologically when we say the "reference point has changed" is that the path integrator place fields translate or rotate relative to the external world. (Remember that, in this theory, the path integrator place field topology never changes.)

by the system are dependent solely on processing power issues, analogous to questions of how many normal landmarks can comprise a local view. In contrast, because the animal has only one path integrator and one head direction code, it should only be able to represent a single reference frame.

### 3.3.2 Non-spatial aspects of place cells

The *multi-map* hypothesis (*multiple maps in the hippocampus* (O'Keefe and Nadel, 1978), *active subsets* (Muller and Kubie, 1987), *reference frames* (Wan *et al.*, 1994b; Wan *et al.*, 1994c; Touretzky and Redish, 1996; Redish and Touretzky, 1997a), *charts* (McNaughton *et al.*, 1996; Samsonovich and McNaughton, 1997; Samsonovich, 1997)) can explain most of the non-spatial aspects seen in place cell recordings.

**Environmental manipulations.** Since different environments will require different $\mathcal{PI} \times \mathcal{LV}$ associations, they will require different place codes. As reviewed in the Experimental Review (EX.79), the topology of place fields changes between two environments.

An important point to make is that all of these experiments show that when the animal is returned to an environment, the place fields return to the representation encoding that environment. In other words, place fields are often stable from session to session (Muller *et al.*, 1987), particularly for normal animals that are very familiar with the environment. For example, Thompson and Best (1990) report recording a stable place field for months.

However, older animals show unstable representations of the environment (Barnes *et al.*, 1997, EX.83). For young animals, the ensemble correlation between the place fields seen during two subsequent 25 minute experiences show a unimodal distribution (around 0.7, indicative of a similar representation between experiences); but for senescent animals, the ensemble correlation was bimodal (around 0, indicative of a complete remapping, i.e. a new reference frame, and around 0.7, indicative of a similar representation between experiences). Within a single 25 minute run, the ensemble correlation (taken between two halves of the run) was always high (around 0.8). This is consistent with the hypothesis that animals must self-localize on returning to the environment, but not while traveling through the environment and suggests that old rats were occasionally self-localizing to a different reference frame when they were returned to the track.

This same stability-within-session/instability-across-sessions has been seen in animals with genetic deficits to NMDA receptors (Rotenberg *et al.*, 1996, EX.84). Other examples of this last point (that reference frames don't change within a single session barring task-related or other changes, see below), were reviewed in the Experimental Review: Bostock *et al.* (1991, EX.80) recorded from place cells in a cylindrical arena (Figure 2.14), first with a white cue card, and then with a black cue card. Sometimes the place fields were similar and sometimes they were unrelated (as if the two situations were encoded as different environments). However, once a place field changed when the cue card was changed then all other place fields recorded subsequently from the same animal changed with the cue cards. When the white cue card was returned, the place field returned to its original configuration. This implies that after the two cards are represented differently, they are always represented by two different reference frames. Importantly, the animal was removed from the environment between each session.

On the other hand, reference frame transitions have been shown even within a single session. For example, Sharp *et al.* (1995, EX.81) measured place cells in a circular arena with a four-way symmetrical local view. Halfway through the session, Sharp *et al.* rotated the arena by 90° (which left the local view unchanged). Like the Bostock *et al.* (1991) results, they found first that the place fields followed the rotation, but they also found that after a number of sessions, the fields changed suddenly and dramatically after the rotation. Like the Bostock *et al.* (1991) results, once the fields began remapping, they always remapped after the experience.

**Task manipulations.** Markus *et al.* (1995) also found that within-session task manipulations could produce reference frame transitions. When rats were trained to search for food on a large elevated platform either randomly or at the corners of a diamond, different subsets of place cells were active for each task, and some cells that were active for both tasks had different place fields, as if the animals were encoding the tasks as different environments (Markus *et al.*, 1995). When the animal switched between these two tasks, the change between representations was rapid, suggesting a shift in a property encompassing the entire system, such as reference frame.

Markus *et al.* (1995) also found some cells with similar fields in both tasks, suggesting that there may be two levels of representation here: the physical environment and the task within the environment. This does not imply, however, that there are two simultaneously active reference frames. Since the two tasks are occurring in the same environment, some cells may be tuned to the same cues in both tasks. The place fields of these cells would not change from task to task. On the other hand, the topology of the overall place code would still change dramatically from one task to the other. An alternative possibility is that some of the animals may have been using a new reference frame while others used the same reference frame. This would predict that all cells would showed similar fields or none of them would (as was seen by Barnes *et al.*, 1997, EX.83); Markus *et al.* did not specifically examine this question.

In more complex tasks than simply finding food scattered on the floor of the arena, place cells do not always fire when the animal is in the place field. Eichenbaum *et al.* (Eichenbaum *et al.*, 1987; Eichenbaum and Cohen, 1988; Cohen and Eichenbaum, 1993, EX.86 and EX.87) tested rats in an odor-detection task and found that some place cells were dependent on whether the rat was going to the reward location or not. In a similar task, Eichenbaum *et al.* (Otto and Eichenbaum, 1992; Cohen and Eichenbaum, 1993) found that cells responded when two odors matched in a delayed match-to-sample task, but not when they didn't. Animals in the Eichenbaum *et al.* (1987) task had been trained to go to a reward location given one set of odors ($S^+$) but not given another ($S^-$). This meant that the animals were learning to take two different paths to reward depending on the whether the odor was in the

$S^+$ or $S^-$ set. Eichenbaum *et al.* (1987) report that hippocampal pyramidal cells ("place" cells) are odor sensitive. However, they are not sensitive to different odors, they are really sensitive to different reward conditions, indicated in this experiment by different *odor sets*.

This means that the Eichenbaum *et al.* tasks can be explained by reference frame transitions (Touretzky and Redish, 1996; Redish and Touretzky, 1997a). Somehow, the animals associate the $S^+$ or $S^-$ odor sets with different reference frames (presumably because different odor sets imply the animals should take different paths to reward). Equivalently, the Otto and Eichenbaum (1992) experiment showed place cells sensitive to $S^+$ and $S^-$ odor-pairs.

In addition to sensitivity to reward availability, Eichenbaum *et al.* report that some cells are correlated with location during stage of a task. For example, one cell might show a place field when the animal is approaching a sniff port to sample the odor (to determine whether it is an $S^+$ or an $S^-$ odor), but not when the animal leaves the sniff port to either go to the reward location or back to the starting point. Again, this is explicable by reference frame transitions that occur between *stages* of the task (Touretzky and Redish, 1996; Redish and Touretzky, 1997a).

Hampson *et al.* (1993, EX.87) trained rats to do a multiple lever-pressing delayed-match-to-sample task and found place cells dependent on whether a lever had already been pressed or not. They also found cells sensitive to which lever had been pressed (and therefore which lever had to be pressed to receive reward). Again, these stages of the task require different path integration reference points and so might require different reference frames. And would then be expected to be encoded by different hippocampal "maps" (Touretzky and Redish, 1996; Redish and Touretzky, 1997a).

All of these task manipulations share one important trait in common: the different reference frames all occur when the animal is learning to take different routes to reward. In the Markus *et al.* (1995) task, the food was either distributed randomly (and approximately uniformly) across the environment, or at the four corners of a square. In the Eichenbaum *et al.* task, when the rat detected an $S^+$ odor (or an $S^+$ pair) it could get reward at the reward location; when it detected an $S^-$ odor (or an $S^-$ pair), it had to return to the starting point.

**Directional place cells.** An early debate in the place cell literature was whether place cells were directional or not (McNaughton *et al.*, 1983a; Muller *et al.*, 1987; Muller *et al.*, 1994, EX.64). This debate was partially resolved by data showing that place cells were directional when the animal took repeated, restricted movement paths, but not when the animal had complete freedom and criss-crossed its path from every direction (Markus *et al.*, 1995). In other words, when the animal only entered a place field from a few discrete directions, place fields were highly directional, but when it entered the field from a continuum of directions, the fields were not directional at all.

One possible explanation for directionality of place cells on linear tracks is that animals use two reference frames to encode the track (Wan *et al.*, 1994c; Gothard *et al.*, 1996a; McNaughton *et al.*, 1996; Touretzky and Redish, 1996; Redish and Touretzky, 1997a;

Samsonovich and McNaughton, 1997; Samsonovich, 1997). When the animal runs back and forth along a linear track, it defines a reference point for each end of the path. When the animal travels in one direction, one reference frame is active, but when the animal travels in the other direction, the other frame is active. Because location within each reference frame is encoded by a different representation, place fields appear directional. In contrast, when wandering around open arenas like those used by Muller *et al.* (1994), place cells would be non-directional (according to this theory) because animals use a single map to encode the environment (Wan *et al.*, 1994c; McNaughton *et al.*, 1996; Touretzky and Redish, 1996; Redish and Touretzky, 1997a; Samsonovich and McNaughton, 1997; Samsonovich, 1997).[13]

### 3.3.3 Anatomy

Reference frames can be differentiated based on local view (such as which landmarks are present), on routes travelled, on other sensory cues (such as olfactory information) or memory cues (such as whether the environment was recently rotated or not). These cues enter the hippocampus, are mixed with location information (spatial aspects of landmarks represented in the local view) and path integrator representations in the dentate gyrus (DG) and then passed into CA3 which represents a specific position in a specific reference frame.

Let us begin by examining where each of these aspects is represented and follow the anatomical pathways from those representations into the hippocampal formation. As was discussed in Section 3.2.1, spatial aspects of landmarks are represented in the posterior parietal cortex and passed to the superficial layers of the entorhinal cortex (ECs), and from there into DG and CA3 via the perforant path.

When the local view was first discussed (back in Section 3.2.1), data was reviewed suggesting that visual information was divided into a dorsal and a ventral stream in the primate and probably also in the rodent. Spatial aspects of the local view are likely to be represented in the dorsal stream, which includes parietal cortex, while non-spatial aspects are represented in the ventral stream, which includes the inferotemporal cortex in the monkey and area Te2 in the rodent. Area Te2 also sends efferents to the superficial entorhinal cortex (ECs).

Recent memories are probably represented in the prefrontal cortex. Strong evidence exists in the primate that during short delays within a single context, recurrent networks in prefrontal cortex can store temporary values (Goldman-Rakic *et al.*, 1990). Zipser (1991) has shown artificial neural network simulations to demonstrate the viability of this hypothesis. Floresco *et al.* (1997, EX.47) report that lesions disconnecting prefrontal cortex from the ventral hippocampus produces difficulties in delayed radial maze foraging tasks, but not in similar tasks with no delay imposed. If we assume that this within-context working memory

[13] A number of early explanations of place cell directionality included "shifting attention" between directions (e.g. McNaughton *et al.*, 1994a; Markus *et al.*, 1995) but these theories explained the non-directionality of place cells in tasks like those used by Muller *et al.* (1994) by "quickly shifting attention" among landmarks (McNaughton *et al.*, 1994a, p. 593), which is not equivalent to the single reference frame hypothesis.

is represented in frontal cortex in the rodent as well (Kolb, 1990b), then we should expect a pathway from prefrontal cortex to the EC and thence into hippocampus.

Olfactory information enters the system from the piriform cortex via the lateral olfactory tract. It seems to be more strongly represented in the lateral entorhinal cortex (LEC) than the medial (MEC). In an odor-to-place matching task, LEC lesions produce incorrect odor-to-place matches, but when the LEC lesions encroached on MEC, the rats also made spatial errors (Otto *et al.*, 1996, EX.33).

An interesting possibility is that while the location information may enter the system through MEC, the reference frame selection information (such as object identity, olfactory information, etc.) may enter via LEC. When Quirk *et al.* (1992) recorded from MEC cells, they found broad place-like behavior, but no sensitivity to environment. No one has (to my knowledge) recorded from LEC cells.

### 3.3.4 Detecting RF transitions

A transition between reference frames can be observed by a decorrelation between the place fields at the transition.[14] Because place fields cover continuous areas with a finite spatial extent, the firing rates of the population of cells at position $x$ and those at position $x + \Delta x$ will be highly correlated as long as $\Delta x$ is less than the average size of a place field. But the place fields between two reference frames are independent. Therefore, after a reference frame transition, the population of place fields will be uncorrelated. This idea has been used earlier in this section (Environmental manipulations, above) to argue that the data from Barnes *et al.* (1997) showed that naive animals sometimes returned to the wrong reference frame.

The reference frame transition can be seen by plotting the ensemble correlation of the population activity of cells at time $t$ with the same population at time $s$ for all times $t$ and $s$ throughout a journey. If the journey is represented by a single reference frame, the plot will show high correlation along the diagonal. The decorrelation can be seen as a "pinch point," shown in Figure 3.11 about a quarter the way along the diagonal.

### 3.3.5 Reference Frame transitions

There are two important cases in which a reference frame transition can occur: when the animal has just entered an environment (e.g. when it is placed into the water in the water maze task) and when something has changed in the local view or possible-route representations (e.g. when it reaches the end of a linear track and turns around).

These two cases differ in the amount of knowledge the animal has of the relationship between the reference frames. In the former case, the animal has no information about the spatial relationship between its position before and after the transition. Thus it must determine its location entirely from external cues (local view). In the latter case, there is

[14] This analysis method is due to B. McNaughton, personal communication.

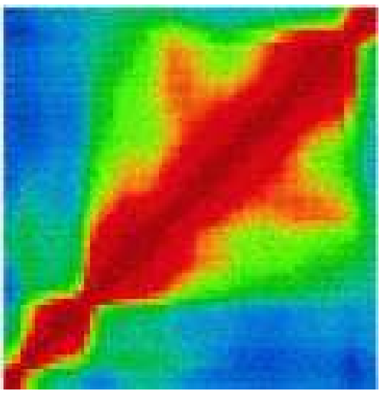


Figure 3.11: Pinch-point from Gothard *et al.* (1996a). Each pixel shows the ensemble cross-correlation between the population of recorded place cells at each point in the environment with each other point in the environment. Thus the diagonal consists of correlations of 1.0. A *pinch-point* (possibly indicative of a reference frame transition) is visible approximately one quarter the way up the diagonal. See text for details. Original figure in color. Reprinted by permission of author and publisher.

a constant relationship between the coordinates in the pre- and post-transition reference frames. Each of these cases can occur in novel or familiar environments.

**Reference frame transitions on returning to a familiar environment.** I have already discussed how an animal *self-localizes* on returning to a familiar environment (Section 3.2.4, see also Chapter 8). This process is dependent on a connection structure within CA3 such that each cell has the strongest synaptic ties to cells that represent similar locations. Because place cells are only co-active when they have overlapping place fields within a reference frame, correlational LTP will produce a connection matrix such that each cell actually has the strongest synaptic ties to cells that represent similar locations *within a reference frame*. This means that the self-localization process described in Section 3.2.4-D (see also Chapter 8) will settle not only to a representation of location, but also to reference frame. Simulations demonstrating the viability of this proposal will be shown in Chapter 8.

**Reference frame transitions on entering a novel environment.** When an animal enters a novel environment, there will be no stored $\mathcal{LV} \times \mathcal{PI}$ association stored, so a new one will need to be learned. A new reference frame needs to be created (by creating a new place code separated from previously learned ones) and a new association to and from that place code need to be stored. There are two possible mechanisms by which this association can be learned: either the reference frames can be separated by an attractor process in CA3 (as originally suggested by Samsonovich and McNaughton, 1997) or they can be separated by an orthogonalization process in DG (as originally suggested by McNaughton and Morris (1987, see also Marr, 1969; McNaughton, 1989; Rolls, 1989; Rolls, 1996).

We can best understand each of these hypotheses by examining how they explain the recent result from Barnes *et al.* (1997) that old animals show unstable place field representations (EX.83). Barnes *et al.* allowed an animal to explore an environment, removed it from the environment for an hour, and then returned the animal to the same environment. They found that old animals sometimes return to a different reference frame. After enough experience, normal young animals always return to the same reference frame.

The key idea of the **CA3-hypothesis** is that the connection structure described as a means of dealing with entering familiar environments above is pre-wired in the hippocampus. This means that when an animal enters a novel environment, the settling process will coalesce on a representation of one of the previously available maps (Samsonovich and McNaughton, 1997). Because old animals have deficient LTP, the local view does not force them into the same map as before.

The key idea of the **DG-hypothesis** is that a DG cell fires given a specific PI representation and a specific LV representation (due to random pre-wired connections). When entering a novel environment, the path integrator settles to a random position, and due to the orthogonalizing nature of DG, a subset of DG cells becomes active, which drives a subset of CA3 cells. Local view information is associated with the CA3 representation which is associated with the path integrator representation. Because old animals have deficient LTP, the system does not reset the path integrator to the same representation on returning to the environment. This means that there will be a different $\mathcal{PI} \times \mathcal{LV}$ association in DG and a different set of CA3 cells active. This demonstrates the necessity of understanding reference frame in terms of the entire system. According to this second hypothesis, the old animals use different maps in their hippocampi because they are (sometimes) using an incorrect path integrator reference point.

This experiment and a comparison between the two hypotheses will be described in detail in Chapter 10.

**Reference frame transitions within a familiar environment.** Let us assume that an animal has been representing an environment by multiple reference frames (such as the two directions on a linear track). How does it change from one reference frame to another when it switches directions? There are two possible mechanisms which cannot be distinguished with the current data, but which make predictions by which we can distinguish them.

The first is an explicit transition mechanism, by which some system recognizes an inconsistency between local view and path integrator coordinates (presumably by an incoherency in the place code Touretzky and Redish, 1996) and forces a new self-localization process. Because the animal has not been disoriented between this transition (it presumably happens without experimenter interference), the current path integrator representation can be used to prime the system. The combination of the new local view and path integrator representation will force the system to settle to the new reference frame.

The alternative is an implicit transition mechanism, based on the internal dynamics of the attractor described in Appendix A (see also Samsonovich and McNaughton, 1997; Samsonovich, 1997). The change in local view will produce a new "candidate location" represented on a new "candidate reference frame." If this new representation is strong enough it will produce a jump in the representation (as shown in Figure A.4).

An intriguing question is whether the dentate gyrus is necessary for this kind of reference frame transition or whether the jump can be driven from learned connections. Data from McNaughton *et al.* (1989a) that even after DG lesions (by colchicine) place cells are directional implies that DG is not necessary for this kind of reference frame transition.

If the first (explicit) hypothesis is true then there should be a self-localization process that occurs at reference frame transitions. We cannot say whether a sharp wave will be necessary or whether some more subtle change can occur, but some specific change would have to occur. The alternative (implicit) hypothesis would not require this. Gothard *et al.* (1996a) may have seen a reference frame transition (see Figure 3.11) but do not report any abnormal processes occurring at the transition. This supports the second (implicit) hypothesis, but Gothard *et al.* were not explicitly looking for such a process.

**Reference frame transitions within a novel environment.** How quickly do different intra-environment reference frames get separated? How do they get separated? The data from Sharp *et al.* (1995, EX.81) imply that the difference is sometimes learned over time. Sharp's (1991) model would suggest that directionality would be learned over time. However, directionality on the linear track is observable from the first experience a rodent has on it (B. McNaughton, personal communication). This issue is reminiscent of the question of how quickly place cells show their place fields in a novel environment (EX.72): some authors report that they appear very quickly (Hill, 1978; Wilson and McNaughton, 1994), while others report that they can take hours to fill out (Austin *et al.*, 1993; Tanila *et al.*, 1997b). The mechanism by which intra-environmental references frames are separated is not known at this time, but can be explained by combining the hypotheses from the previous explanations of transitions within a familiar environment and exploration of a novel environment. Either of the two novel-environment hypotheses can be combined with either of the two within-environment hypotheses. This is an open question and experiments examining the onset of directionality (particularly at the transition points such as the ends of a linear track) would be well worth examining.

### 3.3.6 Discussion

**Reference frames and context.** Hippocampal place cells are sensitive to a host of non-spatial factors such as environment, direction, task, and subtask; in this section, we have seen how sensitivity to each of these factors can be explained by the concept of *reference frames*. These factors form the *context* in which the behavior occurs. An important question is whether place cells should still be primarily considered as "place" cells instead of more general “context” cells. According to the multi-map hypothesis, there is a fundamental difference between the representation of spatial coordinates and the representation of reference frame. When a small change occurs in spatial location, a correspondingly small change occurs in the activity of the place cell population; as the change increases, the place cell representation changes continuously. On the other hand, when a small change occurs in reference frame, no change occurs in the place cells; as the change increases, there is a sharp and sudden change in the representation. If the hippocampus was merely representing the context, then changes in non-spatial variables should be encoded identically with spatial. They are not. Therefore, I believe that it is still useful to consider place cells as representing location, albeit location within a reference frame.

**Predictions of the cholinergic interference hypothesis.** An intriguing prediction comes from an interaction between the interference and reference frame theories: in animals with cholinergic deficits, cells with fields in multiple reference frames should pull their neighbors from one frame into the other. The key point made by the interference hypothesis is that representations should blend together along connections made by neurons that take part in multiple representations. The reference frame hypothesis requires cells with overlapping place fields in one reference frame to have strong connections. Therefore we can predict which cells will be drawn into a reference frame. Simultaneous recordings of multiple cells made under low ACh conditions should show this effect very clearly.

## 3.4 Routes

The place code in the hippocampus allows an animal to self-localize to a specific location in a specific reference frame when it returns to a familiar environment. Once there is a coherent representation in the place code, the goal memory can plan a trajectory to a goal. However the hippocampus has also been implicated in the storage and replay of sequences and recent memories (Scoville and Milner, 1957; Marr, 1970; Marr, 1971; Squire and Zola-Morgan, 1988; Pavlides and Winson, 1989; Levy, 1989; Squire and Zola-Morgan, 1991; Squire, 1992; Cohen and Eichenbaum, 1993; Zola-Morgan and Squire, 1993; Wilson and McNaughton, 1994; Hennevin *et al.*, 1995; McClelland *et al.*, 1995; Smith, 1995; Squire and Alvarez, 1995; Kudrimoti *et al.*, 1996; Levy, 1996; Shen and McNaughton, 1996; Rempel-Clower *et al.*, 1996; Redish and Touretzky, 1997c); in the context of rodent navigation this sequence storage and replay appears as storage and replay of *routes* travelled (Levy, 1989; Abbott and Blum, 1996; Blum and Abbott, 1996; Kudrimoti *et al.*, 1996; Levy, 1996; Skaggs and McNaughton, 1996; Shen and McNaughton, 1996; Gerstner and Abbott, 1997; Redish and Touretzky, 1997c).

### 3.4.1 Storing routes

There are four neurophysiological effects that allow the hippocampus to store routes as the animal travels:

1. **LTP is preferentially asymmetric** (see Abbott and Blum, 1996, and Levy, 1996, for reviews). If neuron $a$ fires shortly before neuron $b$ then it is the $a \to b$ connection that is potentiated, not the $b \to a$ connection.

2. **The actual timing of spikes fired by hippocampal place cells precesses along the theta cycle** (O'Keefe and Recce, 1993; Skaggs *et al.*, 1996, EX.74). Effectively, the position represented in the place code sweeps across the actual position from back to front with each theta cycle (Tsodyks *et al.*, 1996).

3. **LTP is correlational** (see McNaughton and Morris, 1987; Bliss and Lynch, 1988; Brown *et al.*, 1991; McNaughton, 1993; Malenka, 1995 for reviews). The connection between two neurons is potentiated only when spikes in the presynaptic neuron are combined with a depolarization of the postsynaptic neuron. Thus the increase in connection strength between two neurons $a$ and $b$ is proportional to the product of their firing rates $F_a$ and $F_b$.

4. **Cells near the route will have higher-firing neighbors closer to the route.** A cell with a place field not centered at the location of the animal will be more strongly connected to cells that have place fields centered at the location of the animal than to cells with distant place fields.

Effects 1 and 2 combine to produce asymmetries along the routes traveled; effects 3 and 4 combine to produce asymmetries leading towards those routes.

**Asymmetries along the route traveled.** Recent work by O'Keefe and Recce (1993) and Skaggs *et al.* (1996) has shown that the timing of action potentials fired by place cells has an interesting interaction with the theta rhythm. As the animal then moves through the place field, the cell generally fires its spikes earlier and earlier in the theta cycle (EX.74). Because cells with place fields centered in front of the animal fire later in the theta cycle than cells with place fields centered behind the animal, the position represented by the population of active cells sweeps across the animal from back to front with each cycle. When combined with the biophysical time course of LTP (effect 1), this will favor connections pointing along recently traveled routes (Skaggs, 1995; Skaggs *et al.*, 1996).

A number of proposals have been put forward to explain the phase precession effect, including that it is an intrinsic process formed by an interaction of two rhythms with similar frequencies that are just different enough to produce beats (O'Keefe and Recce, 1993) and that it can be generated by asymmetric connections (Tsodyks *et al.*, 1996). Although both of these hypotheses could play a role in generating phase precession, neither of these explanations can fully explain the data. The O'Keefe and Recce hypothesis cannot explain why place cells always start firing at a phase of 90°–120° after the theta peak. The Tsodyks *et al.* hypothesis requires the asymmetric connections to precede the phase precession. However, phase precession occurs from the first traversal through an environment, before the asymmetric connections have had time to be trained (McNaughton, personal communication).

In addition, Skaggs *et al.* (1996) report that DG cells also show phase precession. With data from Bragin *et al.* (1995), suggesting that theta in CA3 is a consequence of theta-related activity in EC, this suggests that the processes that cause phase precession may be extra-hippocampal.

Burgess *et al.* (Burgess *et al.*, 1994; O'Keefe and Burgess, 1996; Burgess and O'Keefe, 1996) suggest a model by which phase precession is a consequence of attention sweeping from landmarks behind the animal to landmarks forward of the animal. However, phase precession also occurs in the dark (Weaver *et al.*, 1996), so this model cannot be the whole story either.

Samsonovich and McNaughton (1996) suggest that phase precession is a consequence of an interaction between the mechanism that drives path integration and internal neuronal dynamics. However their hypothesis that the hippocampus is necessary for path integration has been recently disproven (Alyan *et al.*, 1997, EX.59, see also Chapter 6). Therefore this model cannot be the whole story either. However, this doesn't disprove their phase precession mechanism, but there isn't any data supporting it, either.

What mechanism produces phase precession is still an open question.

**Off-route asymmetries.** Imagine the animal at an instant along the route taken. Because place fields cover an extended area, the animal will be inside the place fields of cells whose place field centers are slightly off the route. But, because the place field is centered slightly off the route, the cell will not be firing at its maximum rate. Cells whose place fields are closer to the route will have higher firing rates than cells with fields farther from the route. Because LTP is correlational (i.e. it is dependent on the firing rate of both the presynaptic and postsynaptic cells) the synaptic strengths of the outputs of these off-route cells will be biased asymmetrically toward the path traveled.

Figure 3.12 shows the asymmetries in a hippocampal simulation after the simulated animal traversed four routes to a goal. (See Chapter 9 for simulation details.)

Although no one has been able to show asymmetries in an actual hippocampus (measuring synaptic weight between two cells *in vivo* is not possible with current technologies), two experiments have been done that strongly suggest that hippocampal recurrent connections do show the hypothesized asymmetries.

Mehta *et al.* (1996) showed that place fields shift backwards along well-traversed paths (EX.75). And Skaggs and McNaughton (1996) have shown data that suggests that routes are replayed during sleep (EX.88). These are both suggestive of asymmetries in the weight matrix.

### 3.4.2 Recalling routes

In order to say that routes have been stored in the hippocampus, there must be a mechanism for those routes to be recalled.

When there is no sensory input into the hippocampal system, the hippocampus will still settle to a coherent activity pattern that represents a valid location due to the symmetric

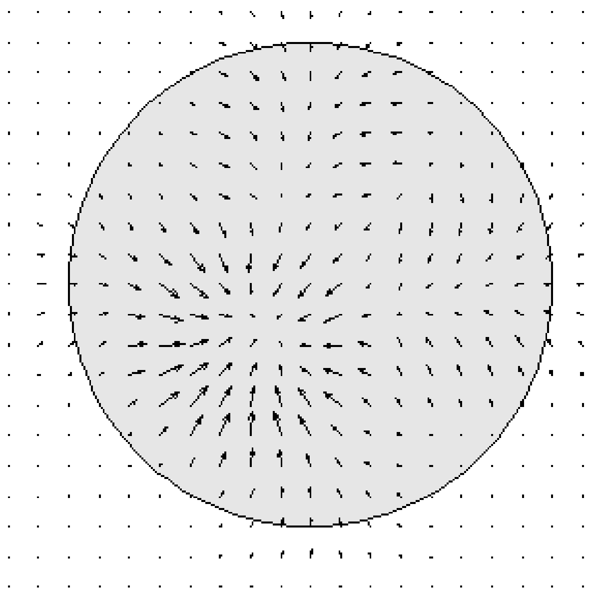


Figure 3.12: Routes stored in CA3. For each cell $j$ in CA3, we calculated the center of mass of the output connection weights, and plotted an arrow from the place field center toward the center of mass. Length of arrow is linearly proportional to the distance between the center of cell $j$'s place field and the center of mass of cell $j$'s output connection weights. From Redish and Touretzky (1997c). Similar simulations have been reported by Abbott *et al.* (Blum and Abbott, 1996; Gerstner and Abbott, 1997).

component of the recurrent connections in CA3. However, due to the asymmetric component, this representation will then drift along the remembered route (Redish and Touretzky, 1997c, see also Tsodyks *et al.*, 1996). Figure 9.5 shows the place code settling to a coherent representation of location which then drifts along a recently traveled route. See Chapter 9 for simulations and citations relating to this replay process.

### 3.4.3 Memory consolidation

The suggestion that memories are written into hippocampus during theta and recalled during sharp waves was first discussed in detail by Buzsáki (1989). Chrobak and Buzsáki (1994) have recently shown that during theta, cells in superficial layers of EC fire in a pattern correlated to the theta rhythm, while cells in deep layers of EC do not. In contrast, during LIA, cells in deep layers of EC fire in a pattern correlated to the sharp waves, while cells in superficial EC do not. This suggests that the deep layers of EC form the output pathway from hippocampus for the consolidation of memories. Compare this to the evidence that the fornix is the output pathway from the hippocampus for navigation (see Section 3.2.5).

Data supporting a replay of recent experience in hippocampus during sleep had been reported by Pavlides and Winson (1989), Wilson and McNaughton (1994), and Skaggs and McNaughton (1996), see EX.88. Simulations of this have been reported by Shen and McNaughton (1996).

REM sleep is characterized by a number of distinctive physiological processes, such as bursts of rapid-eye-movements (REM), a total relaxation of the muscles, and an EEG trace that looks remarkably like theta. (Thus REM sleep is sometimes called paradoxical sleep, because the brain looks like it should be awake, but it isn't.) See Vanderwolf (1990) and Hennevin *et al.* (1995) for reviews.

The original indication that REM sleep affected learning came from REM sleep deprivation studies (Hennevin *et al.*, 1995). Because REM sleep is so distinctive it is easy to wake a subject up whenever it goes into REM sleep. There is a window of a few hours during which animals must be able to get REM sleep in order to learn a task such as the water maze (Smith, 1995).

During sleep, rats first enter slow-wave sleep. LIA occurs throughout the hippocampus, and irregular slow activity occurs throughout the cortex. With the onset of REM sleep, all of the animal's muscles go limp and theta appears in hippocampus, while low voltage fast activity (LVFA) occurs in the cortex. This bout of REM sleep lasts for about 2 minutes (Vanderwolf, 1990). Plenty of time to retrace a route.

One possibility is that during the final sharp wave of the LIA block, the system settles on a coherent location and then throughout the following bout of REM sleep, the system retraces the route.

Another possibility is suggested by data from Smith (1995) and Kudrimoti *et al.* (1996). Smith (1995) reviews data that there is a post-REM NMDA-window in which NMDA blockers such as MK-801 will also disrupt learning. For example, animals with REM deprivation windows of 5-8 hrs post-training had NMDA-windows of 9-12 hrs post-training, and animals with REM deprivation windows of 1-4 hrs post-training had NMDA-windows of 5-16 hrs post-training.

This is particularly intriguing given recent data from Kudrimoti *et al.* (1996, EX.88), who measured the increase in correlation of cells whose place fields overlapped during a recent exploration of an environment during three blocks of LIA in SWS. They found that the correlation decreased through subsequent SWS blocks but increased between SWS blocks when a bout of REM sleep intervened. An intriguing possibility is that replay occurs during both SWS and REM. During SWS, information may be written out to cortex and partially erased from the hippocampus, while during REM, the hippocampal information may be strengthened so that it can be read out again during SWS.

More work is clearly needed looking at the representation of location in hippocampus during sleep to explore whether the replay occurs during slow-wave or REM sleep, and what the corresponding roles of slow-wave and REM sleep are.

### 3.4.4 Where are the routes stored?

An important question that we have not yet addressed is where these routes are stored in cortex. More work needs to be done to determine exactly where the routes are stored, but we can define some requirements the cortical region must meet:

- It must receive input from local view (hypothesized to be in parietal cortex).
- It must receive input from hippocampus.
- It must receive input from head direction areas such as the anterior thalamic nuclei or the postsubiculum.
- It must send efferent projections to motor structures such as nucleus accumbens or anterior motor cortices.
- It must be able to represent intended actions and directions of motion.
- Combined lesions of the area with hippocampus should cause impairments in navigation tasks. Although lesions of the area alone might not cause severe impairments, combined lesions should produce devastating results.

One possibility suggested by Sutherland and Hoesing (1993) (see also McNaughton *et al.*, 1991; Chen *et al.*, 1994b; Chen *et al.*, 1994a) is that posterior cingulate cortex (also known as retrosplenial cortex or area 29) is a good candidate for the anatomical locus of route-storage in long term memory. It is bidirectionally connected with parietal cortex (Vogt, 1985), and with postsubiculum (Wyss and van Groen, 1992; Finch, 1993), receives input from the anterior thalamic nuclei (Sripanidkulchai and Wyss, 1986; van Groen *et al.*, 1993) and subiculum (Vogt, 1985; Wyss and van Groen, 1992), and sends efferents projections to motor cortex (Finch, 1993). Single cell recordings from posterior cingulate cortex show (rare) correlations to head direction, and (more common) correlations to behavior (Chen *et al.*, 1994a).

Supporting this hypothesis is data from Sutherland and Hoesing (1993), who report that both cingulate and parietal lesions continue to have devastating effects, even if they occur 12 weeks after training. They suggest that posterior parietal cortex represents stimulus information, and posterior cingulate cortex then transforms this representation back into motor commands. Posterior parietal cortex may play a significant role in the local view system (Section 3.2.1) and posterior cingulate may play a role in the goal memory system (Section 3.2.5). Unilateral lesions of hippocampus combined with contralateral (but unilateral) lesions of posterior cingulate cortex are absolutely devastating to the ability to perform the hidden platform water maze (Sutherland and Hoesing, 1993). They are as devastating as bilateral posterior cingulate lesions or bilateral hippocampal lesions.

Figure 3.13 shows the areas that may be involved in route replay and storage.

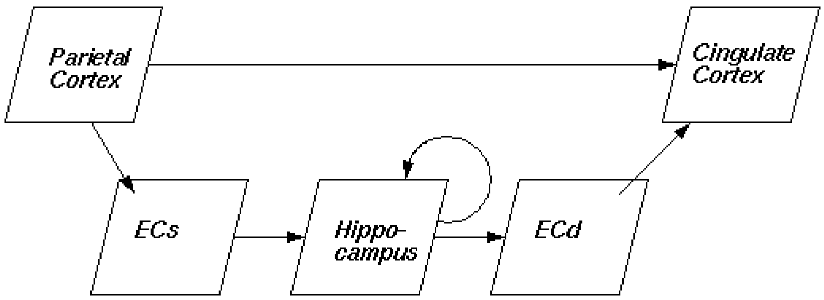



Figure 3.13: Areas that may be involved in route storage and replay. ECs: superficial entorhinal cortex. ECd: deep entorhinal cortex. Routes are stored in hippocampus during theta and recalled during sleep-state LIA.

## 3.5 Functional decomposition

The complete model consisting of the union of these subsystems is shown in Figure 3.14. Figure 3.15 shows their anatomical realization. Note that there is not a one-to-one mapping between brain structure and behavioral function.

Having now synthesized a comprehensive theory of rodent navigation, I will review the theoretical literature on the role of the hippocampus and then move on to specific contributions. In Part II, I examine a number of specific components of the theory (and include simulations): the head direction system (Chapter 5), the path integration system (Chapter 6), and the role of the hippocampus (Chapters 7–9); in Part III, I show simulations of specific experiments, including the bimodality in the representations of an environment by old animals but not young (Barnes *et al.*, 1997, Chapter 10), the interaction between entry

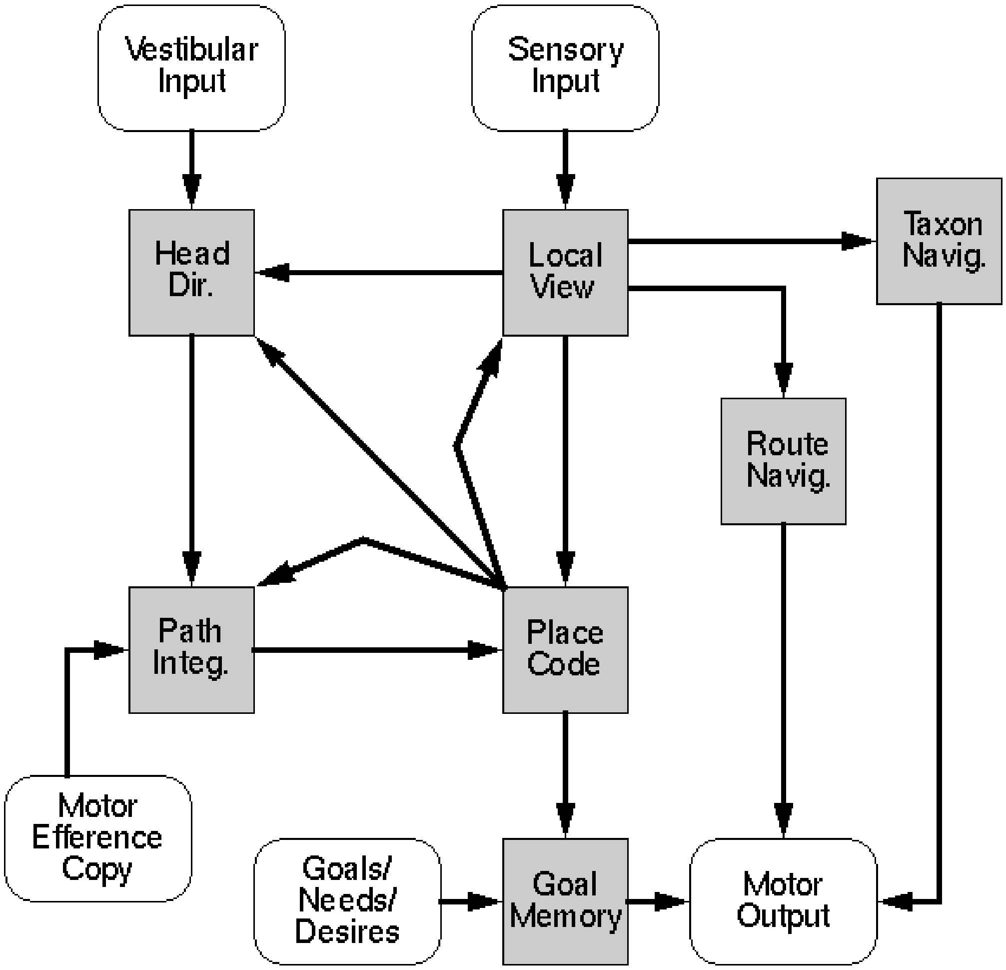



Figure 3.14: Functional decomposition of navigation.

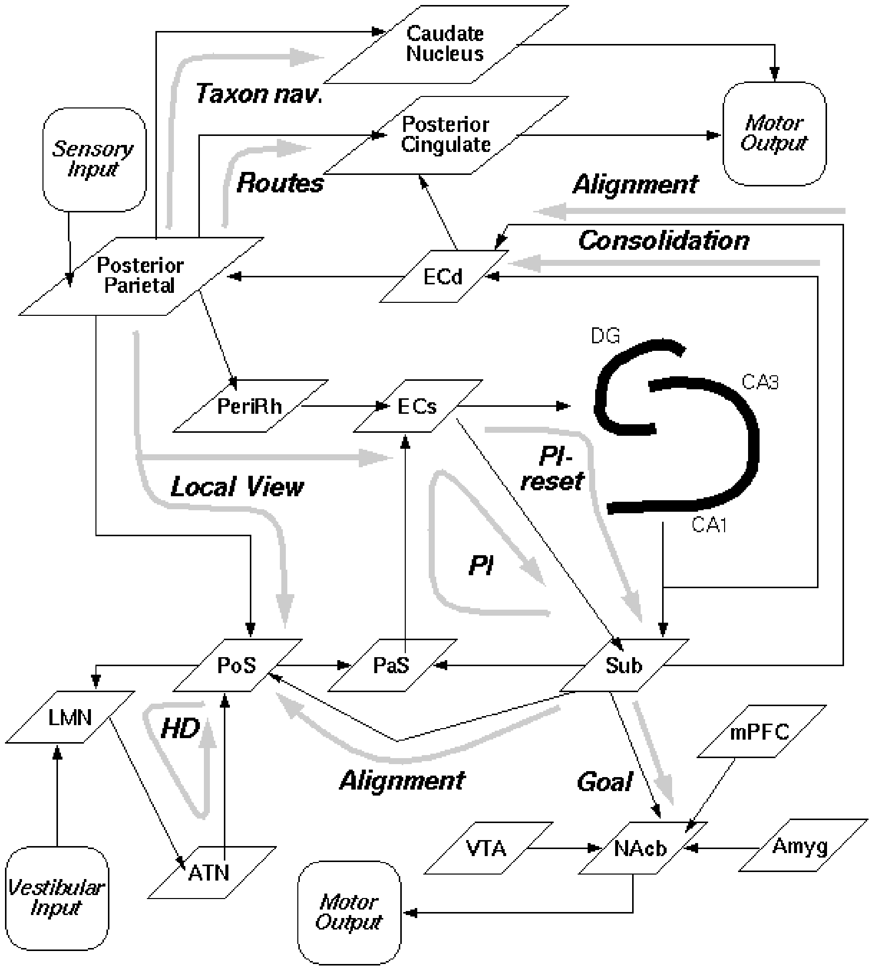



Figure 3.15: Anatomical realization: PeriRh: Perirhinal cortex; ECs: superficial entorhinal cortex; ECd: deep entorhinal cortex; DG: dentate gyrus; CA3,CA1: hippocampus proper; Sub: subiculum; PaS: parasubiculum; PoS: postsubiculum; LMN: lateral mammillary nuclei; ATN: anterior thalamic nuclei; NAcb: nucleus accumbens; VTA: ventral tegmental area; Amyg: amygdala; mPFC: medial prefrontal cortex. Not all anatomical structures or connections are shown. Functional pathways are meant to be indicative only; structures not directly on a labeled pathway may also be involved in that subsystem.

point and cue cards (Sharp *et al.*, 1990, Chapter 11), the differential effects of disorientation or non-disorientation (Cheng, 1986, and Margules and Gallistel, 1988, Chapter 12), self-localization in the Morris water maze (Morris, 1981, Chapter 13), and effects of manipulating the landmark array on search by gerbils (Collett *et al.*, 1986, see also experiments in Section 2.5, simulations in Chapter 14).

An important aspect of this theory presented here that it treats an entire domain, and thus uses a huge corpus of data with which to constrain the hypotheses. By looking at a large corpus of experiments across a variety of paradigms, I have been able to bring the theoretical literature into a consistent picture (rejecting some theories, accepting some, and modifying others). By not assuming a one-to-one mapping between structure and function, I have been able to pull together a theory that fits data ranging from the anatomical, through the neuropharmacological and neurophysiological levels, all the way to behavior. No anatomical structure works alone. It is only through interactions between structures that behavior is accomplished.

# Chapter 4

# Theories of hippocampal function

We have seen how navigation abilities can be divided into two categories: *taxon/praxic navigation* and *locale navigation*, and how locale navigation can be understood as an interaction between five subsystems: *local view*, *head direction*, *path integration*, *place code*, and *goal memory*. We have seen how routes are stored and recalled in the place code and how the place code is realized by hippocampal place cells. We have seen how this implies a need for multiple reference frames, realized by different place field mappings. And we have seen how the interaction of these systems forms a comprehensive, computational theory of navigation.

What does this imply for the role of the hippocampus? The hippocampus is one of the most studied structures in the brain, and a number of theories have been proposed to explain its role. It is generally accepted that the hippocampus is involved in spatial navigation, but not that the hippocampus is solely spatial. Although this thesis is not a "hippocampal model," it addresses a number of issues about the hippocampus, so before proceeding to my specific contributions, I will review the major theories of hippocampal function and address how the understanding brought out in the previous chapter relates to this question of the role of the hippocampus.

Although hippocampal studies have been performed on rodents, rabbits, and both human- and non-human primates, I have discussed only rodents. I will continue to do this for two reasons: first of all for simplicity, but also because it is not necessarily true that the hippocampus plays the same role across diverse species. I will note analogies to non-rodent data when it illuminates an intriguing hypothesis, but do not want to make any strong claims about the role of the hippocampus in species other than rodents.

## 4.1 The theories

### 4.1.1 The hippocampus as emotional center

Early theories, taking their cue from the fact that the hippocampus forms strong connections with other parts of the limbic system, suggested that it mediated the emotional state of the animal (Papez, 1937; Isaacson, 1974). The limbic system is popularly known as the "reptilian" brain, and was believed to control the emotions and feelings that we can not control consciously. This theory traces its roots to Freud (1923) and the concept of the *triune brain* (see Isaacson, 1974). It comes mostly from early inabilities to separate structures that control emotion (such as amygdala and anterior cingulate cortex) from nearby structures (such as hippocampus and posterior cingulate).

Most of the data used to inform this theory comes from early lesion data. Much of the early lesion data examined effects in tasks that were either under or overcontrolled. Lesion data (particularly the early lesion data) is generally problematic for a number of reasons including (1) the fact that it is nearly impossible for a lesion not to encroach on adjacent structures and early lesions often affected fibers of passage and other structures that may have been on the way to the target, (2) if an animal can no longer perform a task, the reason why not may not be clear, and (3) conversely if an animal can still perform a task, it may not be using the same strategy after the lesion as it used before.

The main problem with this theory is its inadequacy to explain the extensive neurophysiological data that has energized the field over the last 20 years (see Experimental Review, *passim*).

### 4.1.2 Hippocampus as a predictive comparator

Gray (1982b, see also Gray, 1982a) attempted to bring the disparate fields of the study of anxiety (as evidenced by the effects of antianxiety drugs and specific lesions) and the understanding of the hippocampus (in particular its connections to the septum via the fimbria-fornix) into a comprehensive theory. In this theory, Gray suggested that a *behavioral inhibition system* (BIS) senses signals of punishments, nonreward, and novel stimuli, sets up conditions of behavioral inhibition, and increases arousal and attention. The effect of antianxiety drugs is hypothesized to reduce anxiety by impairing this BIS.

Gray then takes the questionable step of equating the anatomical instantiation of the BIS with the septo-hippocampal system. The main function of the hippocampus in Gray's theory is as a predictor, sending information to the subiculum which acts as a comparator: By predicting the future and comparing it to subsequent sensory cues, the animal can decide whether to continue on its current path or to stop and re-evaluate the situation.

The main problem with this theory is that it is unable to explain much of the data presented in the Experimental Review. For example, it is completely unable to address the reasons that animals cannot solve the hidden platform water maze after hippocampal lesions (EX.14). A more likely anatomical instantiation for the behavioral inhibition system may be the dopamine neurons recorded by Schultz *et al.* (1995) which are sensitive to changes in reward and novelty.

### 4.1.3 Configural association theory

Sutherland and Rudy (1989) suggested that the right way to think about the hippocampus is by differentiating between *simple associations* and *configural associations*: a *simple association* relates two elementary stimuli, while a *configural association* combines the stimuli into a unique representation in order to differentiate similar representations that can then be associated with other stimuli.

Sutherland and Rudy originally suggested this as a means of solving the *negative patterning discrimination problem*, which is equivalent to the *XOR* problem: respond in one way to (A) or (B) but in another way to (A and B) or (not A and not B). A linear network cannot solve this problem (Minsky and Papert, 1969), however a multi-layer network with simple non-linearities can (Rumelhart *et al.*, 1986b, see Hertz *et al.*, 1991). In a sense, Sutherland and Rudy are suggesting that the hippocampus is the middle layer of a three-layer network and that the basic learning rule for the rest of cortex is essentially linear. Gluck and Myers (1993; 1996) have made this explicit: they simulate the "hippocampus" as a three-layer autoassociative network trained by backpropagation of error and use the hidden layer representation as an output to inform the cortex.

The main problem with this theory is that the "configural associations" that rodents have trouble with all involve an environmental component (i.e. the environment is one of the elements of the configuration). In fact, in a configural association that did not involve changing environments, Jarrard (1993) found that animals with ibotenic lesions acquired the association normally (where A was a tone and B was a light). As suggested by Nadel (1994; 1995, see also O'Keefe and Nadel, 1978; Nadel and Willner, 1980; Jarrard, 1993, and Section 4.1.6), it may well be that environmental context is handled specially.

The configural association theory also suggests that there should be no difference between spatial tuning and non-spatial tuning of hippocampal pyramidal cells. As discussed in the Navigation Overview (see Section 3.3), although hippocampal pyramidal cells show correlations to non-spatial aspects, these must be seen as second order correlations in that they determine how place field topologies change depending on the non-spatial aspects.

### 4.1.4 Working memory

Based on data from a radial maze (Figure 2.16, see Experimental Review, Section 2.1.6) showing that fimbria-fornix lesioned rats couldn't remember which arms they had already visited but could remember which arms were never baited, Olton *et al.* (1979) suggested a difference between memory for within-task information (*working memory*) and between-task information (*reference memory*). Olton *et al.* suggested that the hippocampus mediates working memory but not reference memory. However, the hidden platform water maze (which is strongly hippocampus-dependent, EX.14) requires reference memory, implying that the hippocampus is also involved in reference memory.

Jarrard (1993) explicitly tested reference memory and working memory on spatial and cued versions of the radial maze (EX.45) and found that rats with ibotenic hippocampal lesions showed errors of all three kinds (entries into never-baited arms, repeated entries into once-baited arms, and repeated entries into never-baited arms) in the place version, implying that the rats were totally lost. In the cued version, however, rats avoided cued arms that were never baited (reference memory) and did eventually learn to not repeat cued arms (working memory) but took much longer to train. Jarrard (1993) takes this as a disproof of the working memory hypothesis. Along with the involvement of the hippocampus in the hidden platform water maze, the hypothesis that the hippocampus mediates working memory and not reference memory must be abandoned.

As pointed out in the commentary associated with Olton *et al.* (1979), the working memory distinction is a good operationalization of the radial maze and certain other tasks, even if it is not necessarily a good description of the impact of hippocampal lesions.

The concept of working versus reference memory may still be useful for understanding how the hippocampus is involved in these tasks. As noted by Olton *et al.* (1979), there is a strong relation between the working memory concept and the general concept of *interference* (see Hasselmo and Bower, 1993, and Hasselmo, 1993, for a good description of interference).

As pointed out in the commentary associated with Olton *et al.* (1979), another problem with the working memory hippocampal hypothesis is that Olton *et al.* (1979) lumped together a variety of hippocampal lesions. More recent experiments have brought to light differences between the various components of the hippocampal formation (see Navigation Overview, Chapter 3). Many of the lesion experiments reviewed by Olton *et al.* (1979) involved fimbria-fornix lesions. Cholinergic antagonists have been implicated in impairments of working memory more than reference memory, even on spatial tasks (Eckerman *et al.*, 1980; Wirsching *et al.*, 1984; Buresova *et al.*, 1986, EX.54). One possibility is that acetylcholine affects interference (Hasselmo and Bower, 1993; Hasselmo, 1993) and that reference memory is less suceptable to interference than is working memory on the radial maze. This is supported by data from Wirsching *et al.* (1984) showing that low doses of scopolomine (a cholinergic antagonist) produce working memory errors but reference memory errors start appearing at higher doses (EX.54).

Other important issues include whether the animals have been pre-trained before the lesion or not. As reviewed by Barnes (1988), most post-lesion training methodologies have found strong place dependencies but little or no cue dependence in both reference and working memory aspects. However, the pre-lesion data reviewed by Barnes (1988) is less clear. Specifics of the pre-training methodology may be as important as the working memory vs. reference memory dichotomy proposed by Olton *et al.*

An interesting question is the possible relationship between mislocalization and the working memory vs. reference memory distinction. As suggested by the *multi-map hippocampal hypothesis* (see Section 3.3), the $n$-arm radial maze may be encoded by as many as $2n + 1$ maps. Reference memory errors might then consist of between-environment mislocalizations while working memory errors might consist of within-environment mislocalizations. As suggested by Redish and Touretzky (1997a), there might be a particular role for cholinergic inputs in the separation of reference frames particularly in within-environment transitions. The data reported by Mizumori *et al.* (1989, EX.77) would seem to suggest that there is no connection between CA1 place fields and working memory errors, however, Mizumori *et al.* did not make specific correlations and they do report that CA3 place fields were disrupted during the time period working memory errors were being made. Whether or not the radial maze is encoded by multiple maps and whether or not working and reference memory errors correspond to mislocalizations remains to be tested but is an intriguing possibility.

### 4.1.5 Cognitive maps

In contrast to the four disproven theories described above, there are three theories which have survived to the present day and which I will argue have some strong connection to the actual role of the hippocampus. The first of these theories is the *cognitive map theory* of O'Keefe and Nadel (1978). (The other two are *contextual retrieval*, see section 4.1.6, and *episodic memory*, see section 4.1.7.)

The cognitive map theory proposes that the hippocampus stores a cognitive map (Tolman, 1948) of the environment (O'Keefe and Nadel, 1978, see also O'Keefe and Nadel, 1979). The strongest support for the theory comes from data showing that the first-order correlate of the firing rate of hippocampal pyramidal cells (the primary projection cells in hippocampus) is the location of the animal (O'Keefe and Dostrovsky, 1971, see Section 2.2.2-A). As reviewed in the Experimental Review (e.g. EX.14 and EX.45), hippocampal lesions have also been shown to have strong effects on spatial navigation abilities of rodents. This theory has clearly caught on, judging by the tremendous number of experiments that have been performed based on its hypotheses. In addition, extensive modeling work has been done, trying to understand what drives place cells to fire the way they do and what role they might play in navigation; see Chapter 7.

The cognitive map theory was so successful that it spawned a large number of more specific theories that attempted to address the specifics of the role played by the hippocampus in navigation. Many of the theories address how the hippocampus represents the location of the animal and then how the system can use that representation (e.g. Burgess *et al.*, 1994; Wan *et al.*, 1994b; Wan *et al.*, 1994c; Brown and Sharp, 1995; Burgess and O'Keefe, 1996; Recce and Harris, 1996; Sharp *et al.*, 1996; Touretzky and Redish, 1996; see also Chapter 7), but some theories hypothesize other roles for the hippocampus.

#### 4.1.5-A Associative memory for reinstantiating local views

An early theory of why place cells show place fields was that they were sensitive to combinations of landmarks (see, for example, Zipser, 1985; Leonard and McNaughton, 1990; Sharp, 1991; Hetherington and Shapiro, 1993; Shapiro and Hetherington, 1993, and Section 7.2.1), however since place cells continued to show place fields in the dark (O'Keefe, 1976; McNaughton *et al.*, 1989b; Quirk *et al.*, 1990; Markus *et al.*, 1994, EX.71), it was recognized that these models couldn't be the whole story (see Chapter 7 for a discussion of this issue). Early suggestions that the hippocampus might show associative memory properties (Marr, 1971) suggested one possible explanation: some cells were sensitive to remaining landmarks and the rest of the cells were "completed" by the associative memory properties of the hippocampus (McNaughton and Morris, 1987; McNaughton, 1989; Rolls, 1989; Leonard and McNaughton, 1990; Poucet, 1993; Schölkopf and Mallot, 1993; Benhamou *et al.*, 1995; Recce and Harris, 1996).

In the total absence of cues, this theory suggested that animals could also associate local view representations with self-motion information to predict the next expected local view representation (McNaughton, 1989; Leonard and McNaughton, 1990). This would allow the animal to use a sort of dead-reckoning to keep track of its position in the absence of cues. However, this hypothesis is not equivalent to the path integration abilities discussed in Section 3.2.3 (see also Chapter 6) because this mechanism cannot path integrate across an unexplored area — an animal must first experience a transition in order for it to be associated with the subsequent local view. This also means that rodents would not be able to plan shortcuts or unexperienced paths through an environment. As reviewed in the Experimental Review, rodents can plan solve the water maze, even if they must pass through unexplored areas (Keith and McVety, 1988; Matthews *et al.*, 1995, EX.29).

Recce and Harris (1996) have suggested that the associative memory property should not be seen as performing the path integration function, but instead as just filling out a partial local view. They suggest that the path integrator is extrinsic to the hippocampus and that path integration is used to update the local representation outside hippocampus. This has the problem that errors that accumulate will distort the map. A more robust alternative is for the path integration representation (external to the hippocampus) to be input to the hippocampus directly along with the local view (as suggested by O'Keefe, 1976; Wan *et al.*, 1994c; Touretzky and Redish, 1996; Redish and Touretzky, 1997a, see Chapter 7).

#### 4.1.5-B Fractured maps

Worden (1992) hypothesized a navigation ability based on the concept of *fragment filters*, each of which represents the animal's location relative to a small number of landmarks (Worden used three). If an animal needs to find a goal in one fragment, but only has access to landmarks in a different fragment, then there must be a mechanism for fitting two fragments together. It may even be necessary to link more than two fragments in a chain to reach from the currently available landmarks to a distant goal. Worden refers to this as *fragment fitting*. Although Worden associates certain mathematical operations necessary for his algorithm (such as calculating displacements and rotations necessary for fragment fitting) with anatomical structures (DG and CA3 respectively), he supplies little support for this theory beyond the basic data that hippocampal cells are involved in spatial mapping (and the consequent properties of that, e.g. that there are place cells in CA3). The computations required by the processing modules in the Worden (1992) theory do not seem feasible given the detailed hippocampal anatomy reviewed in Section 3.2.4-A.

#### 4.1.5-C The cognitive graph

Muller *et al.* (1991b; 1996) point out that correlational (Hebbian) LTP combined with random exploration of an environment will produce a synaptic weight function such that the weight between two place fields is inversely proportional to the overlap between their place fields. This means that the synaptic weight can represent the distance between the place field centers and the matrix represents the topology of the space. Muller *et al.* suggest the term *cognitive graph* for this structure. Evidence that this structure does exist in the hippocampus after exploration has been shown by Wilson and McNaughton (1994).[1]

Having made this observation, Muller *et al.* then suggest that a *graph-search algorithm* could plan paths using this structure. They propose Dijkstra's shortest-path algorithm (Dijkstra, 1959, discovered independently by Whiting and Hillier, 1960, see Bondy and Murty, 1976, pp. 15–21, and Cormen *et al.*, 1990, chapter 25) a good candidate, but also note the Bellman-Ford algorithm (Bellman, 1958; Ford and Fulkerson, 1962, see also Cormen *et al.*, 1990, chapter 25) as another possibility. This main problem with this theory is that neither of these algorithms submit to an easy or obvious neural implementation. Nor have Muller *et al.* proposed one.

#### 4.1.5-D Hippocampus as a path integrator

McNaughton *et al.* (1996, see also Samsonovich and McNaughton, 1997; Samsonovich, 1997) proposed a mechanism that could do path integration (see Section 3.2.3 and Chapter 6) and proposed that the hippocampus plays a crucial role, notably as the *attractor network* which enforces a coherent representation of position. They suggest that the subicular complex is the source of the *offset connections* which allow the system to track self-motion.

In particular, they require that the cognitive graph be pre-wired into the system before an animal explores an environment. They also note that when one takes into account the fact that place cells show different place fields in different environments, the actual synaptic weight between two cells should be a function of the minimum distance between their place fields over all environments. They term this extension of the cognitive graph the *multi-chart* model of hippocampus and each map a *chart*. I have already reviewed the relationship between the multi-chart theory and the other multi-map hypotheses (*multiple maps in the hippocampus* (O'Keefe and Nadel, 1978), *active subsets* (Muller and Kubie, 1987), and *reference frames* (Wan *et al.*, 1994c; Wan *et al.*, 1994b; Touretzky and Redish, 1996; Redish and Touretzky, 1997a)).

This theory fits a lot of the data, particularly the place cell data (see Section 2.2.2-A), but it requires a highly complex connection structure because of the multiple charts in

[1] Wilson and McNaughton (1994) only showed that the cognitive graph exists after exploration. Whether it also exists prior to exploration (as suggested by McNaughton *et al.*, 1996) or is learned through exploration (as suggested by Muller *et al.*, 1991b) is still an open question.

the hippocampus (see Chapter 6 for a discussion of the complexities involved). Although the complexity of the connection structure does not doom this hypothesis, it did drive proposals that the mechanism may work on a single map outside of hippocampus (Redish and Touretzky, 1995; Sharp, 1996b; Redish and Touretzky, 1996a; Redish and Touretzky, 1997a).

As proposed by Redish and Touretzky (1997a), an experiment in which rodents were tested in the Mittelstaedt and Mittelstaedt (1980) pup retrieval task after ibotenic hippocampal lesions would differentiate these hypotheses. The recent experiment of Alyan *et al.* (1997, EX.59) has shown that rats can show path integrative abilities without a hippocampus which can be taken as disproving the *hippocampus as behavioral path integrator* hypothesis.

#### 4.1.5-E Route-learning as navigation

As described in Section 3.4, the combination of asymmetric LTP and phase precession produces an asymmetric connection matrix in the CA3 recurrent connections that can represent recently traveled routes (Skaggs, 1995; Blum and Abbott, 1996; Skaggs *et al.*, 1996). One effect of these asymmetric connections predicted by Blum and Abbott (that place fields would stretch backwards along well-traveled routes) has been confirmed experimentally (Mehta and McNaughton, 1996; Mehta *et al.*, 1996). Blum and Abbott (1996, see also Gerstner and Abbott, 1997) suggest that these asymmetric connections can be used to guide navigation.

One problem with this hypothesis lies in its inability to address transfer tasks. It suggests that animals plan shortcuts by interpolating vector fields. This means that shortcuts to a goal can be planned from a location only if that location lies between two previously experienced locations. The transfer experiments of Matthews *et al.* (1995, EX.29) belie this.

Another problem with this hypothesis is that the hippocampus is not necessary for online navigation (Alyan *et al.*, 1997): animals can navigate back to a starting point without a hippocampus. I have argued that the effect of the asymmetric connections is better understood as a means of *replay* rather than *recall*, Sections 3.2.4 and 3.4, see also Chapters 7–9.

### 4.1.6 Contextual retrieval

Hirsh (1974) suggested that the hippocampus plays the role of supplying context to normal stimulus-response mechanisms. This *contextual retrieval* theory explained the reversal deficit seen in hippocampally ablated animals: they can acquire a single discrimination at a normal rate, but tend to perseverate which element is chosen when the reward is reversed. It also explained the negative patterning discrimination problem (respond to stimulus A or to stimulus B, but not to the stimulus pair AB) and latent learning (animals learn tasks faster in familiar than unfamiliar environments).

The contextual retrieval theory is similar to both the configural association (Section 4.1.3) and cognitive map (Section 4.1.5) theories. It is similar to the configural association theory in that it suggests the role of the hippocampus is to provide a way of augmenting a basic stimulus-response learning ability, however, like the cognitive map theory, it suggests that the augmenting is purely environmental. In a sense, the dichotomy drawn by this theory can be seen as the same dichotomy as between taxon and locale navigation drawn by O'Keefe and Nadel (1978, see Chapter 3), and as the same dichotomy that originally led Tolman *et al.* (see Tolman *et al.*, 1946a; 1946b; Tolman, 1948) to hypothesize the existence of a *cognitive map* in contrast to the stimulus-response theories of the Behaviorists (e.g. Hull, 1943). Similar proposals have been made by Nadel *et al.* (Nadel and Willner, 1980; Nadel, 1994; Nadel, 1995) and Jarrard (1993).

The hippocampus has been implicated in *contextual conditioning*. Honey and Good (Good and Honey, 1991; Honey and Good, 1993) found that the association itself was unaffected by hippocampal lesions, but that hippocampectomized animals carried the association between environments unlike normals. As pointed out by Jarrard (1993) and Nadel (1994), this suggests that the hippocampus mediates a contextual component in the association. Conversely, animals with hippocampal lesions cannot learn to make one association in one environment and a different association in another (Selden *et al.*, 1991; Phillips and Le Doux, 1992).

#### 4.1.6-A Temporal discontiguity

A similar proposal has been put forward by Rawlins (1985) in which he suggested that the hippocampal lesion data could be summarized as a problem with spanning delays. However, all of the data cited in support of his theory include contextual transitions within the delay. Rawlins emphasizes the temporal rather than the discontiguous aspects of his theory, but his ideas share much with the contextual retrieval points of Hirsh (1974), Nadel *et al.* (Nadel and Willner, 1980; Nadel, 1994; Nadel, 1995), and Jarrard (1993).

Rawlins's theory also shares much with the mislocalization of reference frames (see Section 3.3): a temporal discontiguity implies a necessary *recall* process (see Chapters 7 and 8). Errors in the recall process would produce hippocampally-mediated effects dependent on temporal discontiguities.

An interesting question becomes what exactly produces a *context-switch*. As pointed out by Rawlins, delay learning in normals is facilitated by removal from the environment during the delay, but it is also facilitated by a salient event occurring soon before the choice (Rawlins, 1985). As seen in a number of cases (Eichenbaum *et al.*, 1987; Eichenbaum and Cohen, 1988; Quirk *et al.*, 1990; Sharp, 1991; Cohen and Eichenbaum, 1993; Hampson *et al.*, 1993; Markus *et al.*, 1995; Gothard *et al.*, 1996b, see EX.64, EX.81, EX.85, and EX.87), reference frame switches can occur within a single environment in a complex task. Mislocalizations that occur at these internal transition points would also produce context-based errors that should probably also be thought of as discontiguities.

#### 4.1.6-B Recognition memory

An interesting question that this analysis opens up is the relation between these *contextual retrieval* theories and the *recognition memory* theory (Gaffan, 1972; 1974). Essentially, Gaffan suggested that the hippocampus is involved in the detection of familiarity.

Gaffan seems to have abandoned this theory when Gaffan *et al.* (1984) found that fornix lesions had no effect on stimulus recognition experiments in monkeys. However, Gaffan's early experiments on rats (Gaffan, 1972; 1974) were based on environmental recognition while Gaffan *et al.* (1984) tested stimulus recognition. As discussed above, environmental cues may be handled separately from other aspects.

A mislocalization to an incorrect reference frame might be considered an incorrect recognition of a familiar environment. Barnes *et al.* (1997, EX.83) show an example of where old animals (presumably with deficient LTP and deficient hippocampi) do not use the same representation between multiple experiences in an environment. Conversely, Tanila *et al.* (1997b, EX.82) show an example where elder animals (again with presumably deficient hippocampi) do not recognize a novel environment and perseverate the use of inappropriate representations.

### 4.1.7 Episodic memory

In contrast to the rodent data, early studies on humans with temporal lobe lesions, such as that of H.M. (Scoville and Milner, 1957), suggested that the hippocampal formation may be involved in memory formation. Hippocampal lesions in humans produce devastating impairments in memory (Scoville and Milner, 1957; Squire and Zola-Morgan, 1988; Squire and Zola-Morgan, 1991; Squire, 1992; Cohen and Eichenbaum, 1993; Zola-Morgan and Squire, 1993; Squire and Alvarez, 1995; Rempel-Clower *et al.*, 1996).

Although these patients perform immediate recall tasks normally, they are strongly impaired at times greater than a few minutes. These patients show a total inability to form new long-term memories, thus demonstrating anterograde amnesia. In addition, they also show a retrograde amnesia (an inability to recall things that happened prior to the occurrence of the lesion), but the retrograde amnesia only reaches back a limited time. Thus the earliest memories are still intact.

Further studies showed that the hippocampal formation was not involved in all kinds of memory, but only *declarative memory*. The distinction between *declarative* and *procedural memory*, first made by Cohen and Squire (1980) and extended by Squire (1992), is that *declarative memory* stores facts, names, events, and episodes (thus declarative memory is sometimes called *episodic memory*) while *procedural memory* stores skills acquired through practice. Declarative memory is fast (things can be learned in one trial) and usually consists of "what," while procedural memory usually consists of "how" and requires many practice trials. For example, if you notice that this sentence occurs on page 142, you know it immediately. You can now shut your eyes and if asked, tell someone what page you were just reading. On the other hand, even if I told you exactly how to throw a baseball, until you'd actually tried to throw a baseball a hundred or a thousand times, you wouldn't be very good at it. The first example is declarative, the second procedural.

People with hippocampal lesions can still learn to perform procedural tasks (Milner *et al.*, 1968; Corkin, 1968; Cohen and Squire, 1980; Squire, 1992; Cohen and Eichenbaum, 1993). For example, Corkin (1968) reports that H.M. shows marked improvements on motor tasks (such as the rotary-pursuit task). Even though he reports that he has never seen the task before, and refuses to believe that he can do it, he performs it very well.

The episodic memory theory proposes that the hippocampus serves as a temporary store for declarative memory (Marr, 1970; Marr, 1971; Buzsáki, 1989; Zola-Morgan and Squire, 1990; Squire and Zola-Morgan, 1991; Cohen and Eichenbaum, 1993; Squire, 1992; McClelland *et al.*, 1995; Squire and Alvarez, 1995), but that procedural memory does not require the hippocampus.

#### 4.1.7-A Storage and replay of sequences

Levy *et al.* (see Levy, 1989; 1996) have suggested that the role of the hippocampus is to store and replay sequences. They show that the recurrent connections in CA3 are well-suited for the storage of sequences. The route-storage and replay discussed in Section 3.4 also stores and replays sequences using similar mechanisms. In order to store sequences with overlapping codes, they introduce the concept of "context units" which disambiguate the overlapping codes (Levy, 1996; Levy and Wu, 1996; Wu *et al.*, 1996). (Essentially, these are hidden units in an autoassociative network, see Kohonen, 1977; Kohonen, 1980; Rumelhart *et al.*, 1986a.) The concept of reference frames (Section 3.3) serves a similar purpose to disambiguate crossing routes (Section 3.4).

#### 4.1.7-B Replay of memories during sleep

In his theory of neocortex, Marr (1970) suggested that the role of the neocortex was classification and long term memory and that in order for the neocortex to classify two similar pieces of information it would be useful for the system to be trained from another system during sleep. In his theory of archicortex (hippocampus), Marr (1971) suggested that the hippocampus might play this role. In Marr's theory, the critical question was *where* to store the information, but more recent examinations of this question have addressed other possible reasons for the consolidation of memories through hippocampal function.

McClelland *et al.* (1995) suggested that one possible reason for the temporary storage of memories in hippocampus is that connectionist or PDP networks (in which memories are stored by changes in weight matrices) are subject to *catastrophic interference*. The effect of catastrophic interference is that when an animal learns a second item that *interferes* with an item learned earlier, it forgets the first item. The experiment by Hirsh *et al.* (1978, EX.39) is an example of this: when Hirsh *et al.* tried to train an animal with fimbria-fornix lesions to solve two tasks in a Y-maze alternatively, they found that the rats unlearned one task as they learned the other.

The key to all of these episodic memory theories (Squire, 1992; Cohen and Eichenbaum, 1993; McClelland *et al.*, 1995) is the idea that the hippocampus is specially structured to store memories quickly but that neocortex stores memories slowly. McClelland *et al.* (1995) suggested that this allows memories stored in neocortex to be interleaved which would counteract the catastrophic interference problem.

As reviewed in the Experimental Review (EX.88), there is quite a lot of data supporting the theory that recent memories are replayed in hippocampus during sleep (Pavlides and Winson, 1989; Wilson and McNaughton, 1994; Skaggs and McNaughton, 1996; Kudrimoti *et al.*, 1996). See Navigation Overview, Section 3.4, for a discussion of this phenomenon and the issue of whether the replay occurs during slow-wave sleep or REM sleep.

## 4.2 How do the theories compare?

All three of the surviving theories have their problems: The cognitive map theory cannot explain the fact that hippocampal cells are also correlated to non-spatial aspects. Nor can the contextual retrieval theory explain the difference between locational representations and reference frame representations. On the other hand, it is not yet understood how episodes (declarative memories) are represented in the brain. We understand the mathematics of space, we know the difference between two- and three-dimensions, what it takes to represent position in space. We've been working on those kinds of problems since Euclid, more than 2000 years ago. But we have no comparable mathematics for the representation of episodes. This makes it very difficult to build explicit models of episodic memory in real tasks. And, in fact, no models of hippocampal function in specific episodic memory tasks exist. All published models of declarative memory that include simulations or computational analyses only address storage and retrieval of random binary vectors (Marr, 1971; McNaughton and Morris, 1987; McNaughton, 1989; Rolls, 1989; Alvarez and Squire, 1994; Hasselmo and Schnell, 1994; O'Reilly and McClelland, 1994; Levy, 1996; McClelland and Goddard, 1996; Rolls, 1996). Although these models can address general principles involved in memory, they cannot address the role of the hippocampus in specific tasks. This makes it difficult to compare their results with real experiments or to generate testable predictions. The one exception to this is Shen and McNaughton (1996) which shows storage and retrieval of recently experienced environments and locations.

## 4.3 So what does the hippocampus really do?

Chapter 3 has argued that the role of the hippocampus in the navigation domain is two-fold:

- to allow the animal to self-localize upon re-entering a familiar environment. This allows the animal to reset its path integrator representation from the local view.
- to replay recently travelled routes during sleep LIA states.

However, hippocampal lesions affect more than just spatial tasks. Eichenbaum (1996) reviews a collection of non-spatial effects of hippocampal lesions, including social transmission of food preferences, see also Winocur (1990). When one normal rat is exposed to another that has just eaten a food, the first animal prefers to eat that food instead of a different food when given a choice. It turns out that rodents are sensitive to the combination of carbon disulfide ($CS_2$, a product of digestion) and the food odor. Essentially, if the other rat ate the food and didn't die, the food must be edible. Bunsey and Eichenbaum (1996) showed that the hippocampal lesions only have an effect if a delay is imposed: even with hippocampal lesions, animals showed normal preferences at zero-delay.

As another example, Wood *et al.* (1996) examined rats with ischemic lesions.[2] They found that ischemic rats were impaired in a delayed non-match to sample (DNMS) task. In this task, the animals first displaced an object in a central area to get food, then the object was removed, while the animal waited for a delay (4 sec). After the delay, doors were opened to allow the animal to go to one of two objects (one novel, the other familiar). If the animal displaced the novel object, it received a second food reward.

These examples show that hippocampal lesions affect *offline* or *delay* tasks, in other words, tasks that involve a *context-switch*. In these tasks, what the animal has to do is to reinstate the context after the delay. This is exactly what the hippocampus is doing in the navigation domain: upon re-entry into an environment (read "context"), the animal must self-localize to find its position in that environment (read "determine its current contextual state").[3]

Although we have tried to concentrate on rodents throughout, we should take a moment to examine two primate results, one human and one non-human. People with hippocampal lesions, such as H.M., can hold normal conversations, but they can't remember facts across different conversations (Scoville, 1968; Milner *et al.*, 1968; Cohen and Eichenbaum, 1993). For example, H.M. can solve difficult crossword puzzles but then forgets them soon afterwards (Scoville, 1968). The classic example (told of Jimmy G. a patient with Korsakoff's syndrome[4]) is that of a patient who can talk to the someone, addresses that person by name, etc., but if the questioner leaves the room and returns, Jimmy G. will claim he never met that person before in his life (Sacks, 1985). This effect can be explained because when you leave the room and return, these amnesic patients have to reinstantiate the context.

In a recent paper, Alvarez and Squire (1994) tested rhesus macaques with hippocampal lesions in a standard DNMS task. They found much worse deficits after 10 and 40 minute delays than after delays of 1 minute or less. They attributed the effect to the length of

[2] In both rodents and primates, including both human and non-human primates, a short ischemia obliterates CA1, while leaving most of the rest of the hippocampal structure intact (Squire, 1992). It has devastating effects on declarative memory (Wood *et al.*, 1996; Rempel-Clower *et al.*, 1996). Recent work, however, has called the limited effect of ischemia into question by suggesting that some extra-hippocampal structures may be covertly damaged with ischemia (Wood, personal communication).

[3] Note the similarity between this hippocampal role and the *contextual retrieval theory*, Section 4.1.6.

[4] Patients with Korsakoff's syndrome show similar but not identical anterograde and retrograde amnesias to patients with ischemia-caused lesions and other hippocampally lesioned patients (Squire, 1992).

the delay, but during the 10 and 40 minute delays, they removed the animals from the experimental situation, which they didn't do for the short delays. Although this does not affect normals, it might have a severe effect on hippocampals if the hippocampus is involved in the retrieval of the context in which the task was done (Nadel, 1995, see Section 4.1.6). In support of this hypothesis, Murray and Mishkin (1996) showed that hippocampal-lesioned rhesus macaques are not impaired on a CNMS task at 10 or 40 minutes. The continuous-non-match-to-sample (CNMS) task is similar to the DNMS except that animals are shown a sequence of example objects and then shown novel pairs in reverse order. This means that although there is a delay between the time the animal sees the first object and when it sees the corresponding last pair, the animal never leaves the experimental situation. If the animals do not leave the context, then they can perform the task well even without a hippocampus.

This theory makes some interesting predictions about the effect of hippocampal lesions on rodents. It suggests that hippocampal lesions will have strong non-navigational effects (in contrast to suggestions by O'Keefe and Nadel, 1978), and that they will affect both working and reference memory (in contrast to suggestions by Olton *et al.*, 1980). It suggests that what matters is not the absolute time but whether a context-switch is involved or not (in contrast to Rawlins, 1985), and it suggests that intra-maze context shifts will be hippocampally dependent (in contrast to Hirsh, 1974, and Nadel, 1994; 1995). It also suggests that animals with hippocampal lesions will be able to perform path integration as long as the animal does not leave the environment (in contrast to suggestions by Samsonovich and McNaughton, 1997), and that they will be able to handle decisions based on multiple cues, again, as long as the animal does not leave the environment (in contrast to suggestions by Sutherland and Rudy, 1989).

Examples of some of these cases have already been discussed. The numerous experiments showing that rats cannot solve the hidden platform water maze are examples of hippocampal influences on reference memory. The Wood *et al.* (1996) and Bunsey and Eichenbaum (1996) experiments are both examples of non-spatial hippocampal effects. The Bunsey and Eichenbaum effect is an example in the configural association domain. The fact that there is no effect at zero delay implies that rodents can perform configural associations as long as they don't have to re-instate the context.

Finally, the question of path integration with hippocampal lesions has just recently been tested (Alyan *et al.*, 1997). As reviewed in the Experimental Review (EX.59), Alyan *et al.* confirmed the prediction that animals should be able to show path integrative abilities even with ibotenic lesions of the hippocampus (originally made in Redish and Touretzky, 1997a).

We can also predict that rodents with hippocampal lesions should be able to perform the hidden platform water maze task if it is simplified in the following way: once the animal reaches the goal location, carry it smoothly and directly back to the release point. Presumably, an animal in such a situation does not have to relocalize at the wall's edge. Normally, rodents in the water maze task are carried back to a home cage to rest or are disoriented by being carried around the room before being placed at the release point again.

We know of no experiments examining this task, but if the prediction is correct, the result would invalidate a number of hippocampal navigation models. Notably, the models which say that the hippocampus is necessary for online navigation (e.g. Dayan, 1991; Blum and Abbott, 1996; Dayan and Singh, 1996; Muller *et al.*, 1996; Samsonovich and McNaughton, 1997). These models have to be seriously questioned due to the recent data from Alyan *et al.* (1997).

## 4.4 Non-spatial aspects of hippocampal function

Hippocampal pyramidal cells are sometimes correlated to other aspects of the environment beyond the location of the animal, including local view (Muller *et al.*, 1987; Muller and Kubie, 1987; Bostock *et al.*, 1991), speed, direction, turning angle (McNaughton *et al.*, 1983a; Markus *et al.*, 1994; Wiener *et al.*, 1989), texture underfoot (Young *et al.*, 1994), odor (Eichenbaum *et al.*, 1987), task (Markus *et al.*, 1995), and stage of task (Eichenbaum *et al.*, 1987), as well as recent environmental effects such as a wall being lifted (Wilson and McNaughton, 1993), the lights turning on or off (Quirk *et al.*, 1990), or the environment rotating (Sharp *et al.*, 1995). We have argued (Section 3.3) that many of these cues are necessary for the selection of reference frame, but it is also possible that some of these cues are uncorrelated with reference frame or location. They could then be treated as noise for the determination of reference frame and location.

If these non-locational cells fire in sequence then they will form interconnections, just as the location cells do. In effect, they will store routes in state space. Dreams will replay these routes in state space, just as they do in real space. This means that the same consolidation mechanism discussed in Section 3.4 can occur with non-spatial tasks as well.

In tasks that do not require the animal to learn sequences, we can still ask what representations will be reinstantiated during dreams. In the Winocur (1990) task, the neurons in the representation of getting the food associated with the carbon disulfide will have stronger interconnections than other neurons. This means that, when the hippocampus settles to a stable representation in the absence of sensory input (during dreams), the representation of that food is more likely to be reinstantiated than other representations with weaker interconnections. The representation can be said to have a greater basin of attraction in the state space of hippocampal associations. In other words, the animal is more likely to dream about the food it associated with the smell of carbon disulfide. After enough dreams, the food-$CS_2$ association will be stored elsewhere than hippocampus and the animal will still choose that food over other foods even after hippocampal lesions. Shen and McNaughton (1996) have presented a model of this, showing that recently primed experiences will be preferentially replayed during sleep.

So, to summarize, what the hippocampus does (according to this theory) is re-instatiate context. In the navigation domain, this means that the hippocampus allows an animal to self-localize on re-entry into a familiar environment. This allows it to act as a fast memory for situations (or episodes). In other words to act as a *declarative memory.* Then routes and often revisted memories are stored in the recurrent connnections within hippocampus. During dreams these routes and often revisited memories are replayed and used to train more procedural memories in the rest of the cortical system.

I noted at the start of Section 4.2 that concrete models of episodic memory were too difficult to actually implement. But by understanding the navigation domain in depth and in detail, we can begin to address episodic memory in a concrete way, with models of specific tasks that make explicit predictions. That is, however, beyond the scope of this thesis, and I will have to leave it for future work.

# Part II

# Specific contributions on selected subsystems

Having reviewed the experimental literature that forms the domain of rodent navigation (Chapter 2), the theoretical structure of rodent navigation itself (Chapter 3), and the theoretical work relating to the hippocampus, (Chapter 4), I will now discuss specific contributions I have made to the field of rodent navigation. I have already discussed two specific experimental contributions (Sections 2.5 and 2.6), and in Part III, I will present simulations of specific experiments. In the following chapters, I present:

- (Chapter 5) the first simulation of the head direction system that can track multiple head direction speeds accurately. From these simulations, I made an observation about the tuning curves of cells in ATN. This observation has been confirmed experimentally by Blair *et al.* (1997).
- (Chapter 6) a novel proposal for the anatomical locus of the path integrator. By examining the computational requirements and the anatomical data, I find fault with the previous hypotheses and provide a novel one. I also describe experiments that can differentiate between my proposal and previous hypotheses, one of which has recently been confirmed by Alyan *et al.* (1997).
- (Chapter 7) a novel understanding of the role of the hippocampus in navigation: specifically that it has three modes: *storage*, *recall*, and *replay*.
- (Chapter 8) simulations demonstrating the viability of my hippocampal proposal, including storage and recall of locations within single environments, with ambiguous inputs, and in multiple environments. I will also present a novel explanation of the reason that rodents return to their starting points as they explore a novel environment.
- (Chapter 9) simulations demonstrating the viability of the dual-role hippocampus, showing that the recall and replay modes can coexist within the hippocampus. (The two roles would seem to require incompatible connection matrices.)

# Chapter 5

# Head Direction

## 5.1 Anatomy

Head direction cells in the postsubiculum (PoS, also known as dorsal presubiculum) were first described by Ranck (1984). As reviewed in the Experimental Review (EX.94), these cells can be characterized as having triangular tuning curves: the firing rate drops off linearly from a peak at the preferred direction until it reaches a baseline value (Taube *et al.*, 1990a). However, these curves can also be modeled very closely by Gaussians with an average standard deviation of approximately 66° (Blair and Sharp, 1995; Zhang, 1996a). Similar cells have been found in the lateral mammillary nucleus (LMN: Leonhard *et al.*, 1996, EX.96) and in the anterior thalamic nuclei (ATN), particularly in the anterior dorsal (AD) nucleus (Blair and Sharp, 1995; Knierim *et al.*, 1995; Taube, 1995, EX.95).

These areas are anatomically linked: The AD nucleus of the ATN and PoS are directly interconnected (van Groen and Wyss, 1990); PoS also sends a strong projection to LMN (van Groen and Wyss, 1990), which in turn sends a strong projection to the AD nucleus (Bentivoglio *et al.*, 1993).

As reviewed in the Experimental Review, PoS, LMN, and ATN head direction cells are sensitive to rotation of distal cues (Ranck, 1984; Taube *et al.*, 1990b; Goodridge and Taube, 1995; Knierim *et al.*, 1995; Taube, 1995; Taube and Burton, 1995; Leonhard *et al.*, 1996, see EX.98). However, they do not require visual input to show a strong directional signal (Ranck, 1984; Goodridge and Taube, 1995; Taube *et al.*, 1996, EX.99).

When multiple head direction cells have been recorded from ATN, PoS, or LMN the difference between their preferred directions is a constant across all environments (Taube *et al.*, 1990b; Goodridge and Taube, 1995; Taube and Burton, 1995, Taube, personal communication, EX.97). As discussed in the Navigation Review (see Section 3.2.2), one way of interpreting the activity of these cells is as a distributed representation of the rat's current head direction. A population of HD cells with preferred directions $\tilde{\phi}_i$ evenly distributed through 360° represents the direction of the weighted vector sum $\sum_i F_i \cdot \vec{v}_i$, where $F_i$ is the normalized firing rate of cell $i$ and $\vec{v}_i$ is a unit vector pointing in direction $\tilde{\phi}_i$. This is the weighted circular mean (Mardia, 1972), and is also known as a *population vector* encoding (Georgopoulos *et al.*, 1983). Of course, real head direction cells might not be evenly distributed, but as long as the cells are approximately evenly distributed, this interpretation is still valid (Georgopoulos *et al.*, 1988; Redish and Touretzky, 1994).

Recently, important distinctions have been made between each of the three key areas (ATN, PoS, and LMN): ATN cell activity is best correlated not with current head direction, but with future head direction, while PoS head direction cells are best correlated with the animal's recent head direction (Blair and Sharp, 1995; Taube and Muller, 1995, EX.101). Also, LMN cell activity is strongly correlated with angular velocity as well as direction (Leonhard *et al.*, 1996, EX.96).

## 5.2 Models of the head direction system

### 5.2.1 McNaughton *et al.* (1991)

McNaughton *et al.* (1991) proposed an associative mapping model for updating head direction based on angular velocity. In this model, a head direction cell population $\mathcal{H}$ and an angular velocity population $\mathcal{H}'$ jointly produce activity in a direction velocity population $\mathcal{H}\mathcal{H}'$. Each $\mathcal{H}\mathcal{H}'$ cell is tuned to both a specific direction and a specific angular velocity. Thus, the representation of head direction in the $\mathcal{H}$ population can be updated by a linear associative memory given the representation in the $\mathcal{H}\mathcal{H}'$ population. Various anatomical areas, such as parietal and retrosplenial cortex, are discussed as sources of the $\mathcal{H}$ and $\mathcal{H}'$ signals, and PoS is suggested as a possible site for the $\mathcal{H}\mathcal{H}'$ population. This model is a state shift model in which a new representation of head direction is generated at each time step from the representation of the head direction $\times$ angular velocity state. Although it presented the first model that could explain how head direction could be maintained from vestibular and reset by extrinsic cues, no simulations were reported.

### 5.2.2 Blair (1996)

Blair (1996) proposed a somewhat different shift register model in which clockwise and counterclockwise angular velocity modulated head direction cells (AVHD cells) in the reticular thalamic nucleus selectively inhibit cells in ATN that are offset from the corresponding AVHD cells. Blair (1996) suggested that the activity of ATN cells is also governed by a modulatory input from angular speed cells hypothesized to exist in the mammillary bodies; this varies the rate of shifting because the ATN cells provide input that helps drive the AVHD cells. In this model, ATN also drives HD cells in PoS and retrosplenial cortex. Blair (1996) was the first to propose a model in which the ATN representation leads the PoS representation. However, there is no evidence for the necessary cell types in the reticular thalamus or the mammillary bodies. Also, the model did not produce realistic tuning curves, and Blair did not report tracking performance on realistic data sets. He did show that his simulations could track a single turn, but did not examine rotations at multiple angular velocities.

### 5.2.3 Attractor networks (Skaggs *et al.*, 1995; Redish *et al.*, 1996; Zhang, 1996a)

Skaggs *et al.* (1995) sketched a model of the head direction system which accounted for both the shape of the head direction tuning curves and a means of updating the representation as the animal moved through space. In their proposal, a population of head direction cells forms an *attractor network* in which *coherent representations of head direction* are stable states of the network. A coherent representation is one in which the firing rates of all component cells are consistent with current representation.

In their proposal, head direction cells also project to corresponding left and right *rotation cells*, whose activity is controlled by *vestibular cells* that fire when the animal is making a left or right turn. The rotation cells play a role similar to the $\mathcal{HH}'$ cells of McNaughton *et al.* (1991) and project back to either the left or right neighbors of the head direction cell that drives them. Thus, during a turn, the hill of activity over the head direction cell population gradually shifts. This model was an abstract proposal and did not include simulation results, however, it formed the basis for the subsequent simulations by a number of authors (Elga *et al.*, 1996; Zhang, 1996a; Redish *et al.*, 1996), including those presented here.

Zhang (1996a) presented the first simulation results of a pure attractor model of the head direction system. He defined a Gaussian-shaped hill of activation and derived weights and an activation function that produce self-sustaining activity patterns of this form. He then derived another weight function to dynamically translate the activity pattern to the left or right. But Zhang did not separate the roles of ATN and PoS and did not measure tracking ability on realistic data sets.

The models of Skaggs *et al.* and Zhang can be understood as a combination of two different weight matrices: (1) during periods of non-rotation, the weight matrix among the head direction cells has a local-excitation and global-inhibition structure, and (2) during rotations, the offset cells become active and input extra activity into the system which effectively makes the weight matrix asymmetric in the direction of rotation. This kind of local-excitation/global-inhibition network has been well studied, both with symmetric weight-matrices (showing that when the weight matrix is symmetric a "hill" of activation is a stable state in the network, Wilson and Cowan, 1973; Amari, 1977; Ermentrout and Cowan, 1979; Kishimoto and Amari, 1979; Kohonen, 1982; Kohonen, 1984; Murray, 1989), as well as asymmetric (in which travelling waves can form, Wilson and Cowan, 1973; Ermentrout and Cowan, 1979; Murray, 1989). Fundamentally similar models (moving a hill of activation along a one-dimensional or two-dimensional attractor) have also been proposed as explanations of saccade-generation by the superior colliculus (Droulez and

Berthoz, 1991; Munoz *et al.*, 1991; Lefévre and Galiana, 1992; van Opstal and Kappen, 1993; Arai *et al.*, 1994).

## 5.3 Contributions

We[1] began by simulating a modification of the networks of Skaggs *et al.* (1995) and Zhang (1996a). We separated out the maintenance and update functions of the network by having two populations, both of which formed attractor networks, which we identified with PoS and ATN.

In this model, there are two types of connections between the excitatory neurons of the two populations: *matching* and *offset*. *Matching connections* connect cells in PoS with cells in ATN and cells in ATN with cells in PoS that have similar preferred directions. *Offset connections* connect cells in PoS with cells in ATN that have slightly offset preferred directions.

This network structure is similar to that of Skaggs *et al.* (1995) and Zhang (1996a) in that the mechanism for maintenance of the stability of the representation occurs by an attractor network and the mechanism for updating the representation from vestibular cues occurs by offset connections. It is different in that Skaggs *et al.* and Zhang suggested that the head direction $\times$ angular velocity component was reintegrated together with the attractor network in a single population. Zhang simulated both mechanisms within the same population, but unlike in our simulations, Zhang did not show tracking at multiple angular velocities. Skaggs *et al.* did not report simulations or address the complexities of tracking at multiple angular velocities.

This work demonstrated three things:

1. This model can accurately track a real behavioral sequence of turns with realistic neurophysiological head direction tuning curves.

2. The representation of head direction in the offset population (which we identified with ATN) led the representation of head direction in the maintenance population (which we identified with PoS).

3. We found that without attractor dynamics in the ATN population, the head direction tuning curves deformed (they "stretched" to one side during periods of rotation).

Because the previous data had suggested that ATN tuning curves do not deform during rotations (in contrast to point 3, above), we included attractor dynamics in our ATN component. Since this work was published, Blair *et al.* (1997) have reported that ATN tuning curves do deform during rotations (EX.102). Since the simulations were done before Blair *et al.*'s experimental work, I will first present the simulations and the conclusions

[1] This work was done in collaboration with Adam Elga and David S. Touretzky. This work has been previously reported in conference (Elga *et al.*, 1996), and journal (Redish *et al.*, 1996) form.

from it, but after I have presented them, I will discuss these recent experimental results as well as some more recent theoretical work which has extended this theory (Zhang, 1996b; Goodridge *et al.*, 1997).

## 5.4 Simulation details

The neuron model used in our simulations consisted of three equations. For an extensive discussion of the derivation, see Wilson and Cowan (1972) and Pinto *et al.* (1996) (Eqs. 1.1–1.3).

### 5.4.1 Maintaining A Stable Representation

To create a population with a triangular attractor state (actually a Gaussian with standard deviation of roughly 66°), the network needs to have local excitation and global inhibition (Wilson and Cowan, 1973; Amari, 1977; Ermentrout and Cowan, 1979; Kohonen, 1982; Kohonen, 1984; Murray, 1989; Skaggs *et al.*, 1995; Zhang, 1996a). The simulations included an excitatory pool $E$ and an inhibitory pool $I$, each composed of units governed by equations (1.1)–(1.3).

For simplicity, we assumed that the units in each pool have evenly distributed preferred directions.[2] A unit in the excitatory pool strongly excites those units in both pools whose preferred directions are close to it. A unit in the inhibitory pool weakly inhibits practically all units in both pools, but units close in preferred direction are inhibited slightly more. See Figure 5.1.

Thus the connection weight from neuron $j$ to neuron $i$, $w_{ij}$, is a product of two terms: a weight $\tilde{w}_{P(i)P(j)}$, dependent on the pools in which $i$ and $j$ reside (excitatory pool E or inhibitory pool I), and a Gaussian $G_{P(j)}(\Delta\phi_{ij})$ dependent on the pool of the input neuron $j$, $P(j)$ and the difference in preferred direction between $i$ and $j$, $\Delta\phi_{ij}$. See Table 5.1 for specific parameter instantiations used.

With appropriately chosen parameters (see Table 5.1 for parameter values used), the excitatory and inhibitory pools will have single regions of excitation in each pool as stable attractor states.

### 5.4.2 Updating the Representation

This model of PoS-ATN interactions consists of two coupled attractor networks and an additional inhibitory gain control population synapsing on ATN, perhaps coming from the

[2] In the simulations presented here, both pools included the same number of units, evenly distributed through 360°. It is not necessary that the inhibitory neurons show directional tuning or that they preferentially inhibit excitatory cells with similar preferred directions, but the simulations presented here included both effects. Simulations with a single inhibitory interneuron also produces similar (but not identical) results (with appropriate parameters).

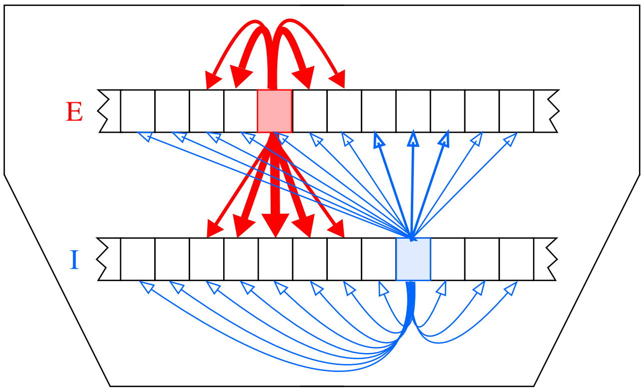



Figure 5.1: Two-pool attractor network. A neuron in the E pool strongly excites those units in both pools with similar preferred directions, but a neuron in the I pool weakly inhibits practically all units in both pools, although units close in preferred direction are inhibited more. From Redish *et al.* (1996).

reticular thalamus or another unknown structure.[3] All connections between the two attractor networks are between their excitatory pools.

In addition to the intrinsic connections within each pool and between the two pools making up each attractor, which are necessary for the maintenance of the representation, there are two types of connections between the excitatory PoS:E and ATN:E pools: *matching* and *offset*. Matching connections are reciprocal connections between units with corresponding preferred directions. In the absence of any head rotation, matching connections dominate and serve to synchronize the locations of the peaks in the PoS and ATN pools.

Offset connections are responsible for changing the represented head direction. They come in two forms: left and right. Each element in the excitatory pool of the PoS structure with preferred direction $\tilde{\phi}$ has a left offset connection to the unit in the excitatory pool of the ATN structure with preferred direction $\tilde{\phi} - \delta$ and a right offset connection to the unit with preferred direction $\tilde{\phi} + \delta$, where $\delta$ is the amount of the offset and is the same for all units.[4] The attractor nature of the ATN structure will move the stable state towards any off-peak excitation. In practice, we used $\delta = 10°$ because this does not deform the tuning curves very much and produces a smoother progression of represented head direction values.

All offset connections have strengths modulated by angular head velocity in the following way: While the head is rotating to the right, right offset connections have strength proportional to the speed of rotation and left offset connections have strength zero. The

[3] The inhibitory gain control is not necessary if the ATN tuning curves are allowed to deform with rotations. See Section 5.6, below.

[4] $\delta$ does not have to be the same for all units, but this approximation simplifies the simulations.

opposite holds true for rotations to the left. During periods when the animal is not turning, both sets of offset connections have zero strength; only the matching connections remain effective, synchronizing the PoS and ATN representations.

We used an empirically determined function, $\xi(\dot{\psi})$, for the angular velocity modulation of the offset connections to ATN. As shown in Figure 5.2, the relationship is nonlinear but monotonic. The simulations also included a compensatory gain control signal which we implemented by setting the tonic inhibition on all ATN:E cells to $\gamma_E = -\frac{1}{2}\xi(\dot{\psi})$. This nonspecific modulation was applied to all units in the ATN:E population, independent of preferred direction.

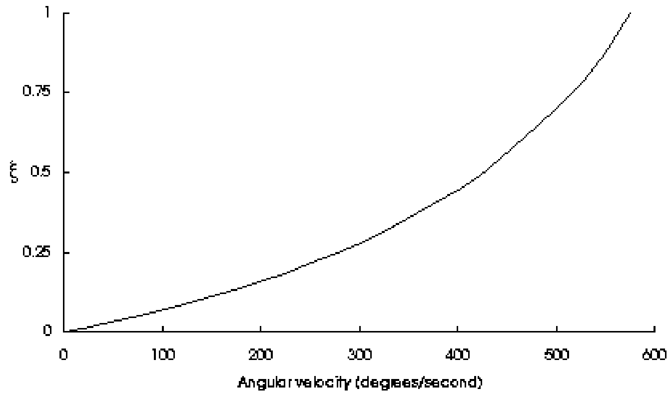



Figure 5.2: Connection strength of the offset connection between PoS:E and AD as a function of angular velocity. This function $\xi(\dot{\psi})$ was experimentally determined. From Redish *et al.* (1996).

This matching-plus-offset connections architecture entails that the locations of Gaussians in the two attractor modules will be synchronized during periods of no rotation, but during rotations ATN will lead PoS. Furthermore, the amount of lead will depend on the angular velocity of the rotation, but due to the gain control mechanism, the shape of the hill in ATN will remain largely unchanged.

## 5.5 Simulation results

We begin by comparing the model to data recorded by Blair and Sharp from freely moving rats (for details on recording methods, see Blair and Sharp, 1995). This data included the rat's head direction $\psi_t$, sampled at 60 Hz. Missing data points were linearly interpolated. To counteract quantization error, head direction at each time step was calculated as the average direction over a 133 ms window centered at that time. This filtering smoothed out fine fluctuations without removing important detail from the angular velocity trace. The rat's angular head velocity was then estimated over the 16 ms time period between two

| | | |
|---|---|---|
| Time step | dt | 0.1 ms |
| Number of excitatory units in a pool | $N_E$ | 100 |
| Excitatory unit time constant | $\tau_E$ | 1 ms |
| Excitatory unit tonic inhibition | $\gamma_E$ | -1.5 |
| Number of excitatory units in a pool | $N_I$ | 100 |
| Inhibitory unit time constant | $\tau_I$ | 0.2 ms |
| Inhibitory unit tonic inhibition | $\gamma_I$ | -7.5 |
| Standard deviation of $G_E$ | $\sigma_E$ | 30° |
| E $\rightarrow$ E connection weight | $\tilde{w}_{EE}$ | 5.0 |
| E $\rightarrow$ I connection weight | $\tilde{w}_{IE}$ | 16.0 |
| Standard deviation of $G_I$ | $\sigma_I$ | 360° |
| I $\rightarrow$ I connection weight | $\tilde{w}_{II}$ | -8.0 |
| I $\rightarrow$ E connection weight | $\tilde{w}_{EI}$ | -12.0 |
| PoS $\rightarrow$ ATN matching connection strength | $W_{\text{PoS} \rightarrow \text{ATN}}$ | 1.0 |
| ATN $\rightarrow$ PoS matching connection strength | $W_{\text{ATN} \rightarrow \text{PoS}}$ | 0.6 |
| Offset | $\delta$ | 10° |
| PoS $\rightarrow$ ATN offset connection strength | $\xi(\dot{\psi})$ | see Figure 5.2 |
| $\gamma$ correction factor in ATN | $\gamma_{\text{ATN}}$ | $-\frac{1}{2} \cdot \xi(\dot{\psi})$ |

Table 5.1: Simulation Parameters.

samples $\psi_t$ and $\psi_{t+1}$ as the change in head direction divided by 16 ms. These head velocity estimates served as the vestibular input $\dot{\psi}_t$ for our simulations.

Simulations thus consisted of (1) initializing the units to random states, (2) allowing the two components to settle to stable attractor states (approx 20–30 simulated ms), (3) identifying the direction represented in PoS:E with the initial head direction sample, and (4) allowing the system to run using the $\dot{\psi}_t$ sequence calculated as per above, and at each step comparing the direction represented in the E pool of PoS with the measured head direction of the animal.

Figure 5.3 shows the model's ability to integrate actual rodent head movements. Cumulative HD tracking error fluctuated, but typically did not exceed 20° for simulations shorter than 3 minutes.

Tracking accuracy was largely dependent on four parameters: how strongly vestibular input modulated the left- and right-offset cells, the offset amount $\delta$, and the time constants $\tau_E$ and $\tau_I$. The offset modulation controlled the angular velocity of the representation for a given input head velocity, while the amount of offset $\delta$ controlled the lead of the ATN population. The $\tau$ parameters controlled the resistance of units to changing activation, determining the inertia of each pool.

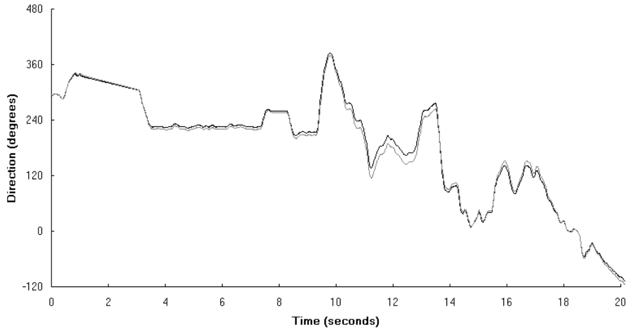
In agreement with Ermentrout and Cowan (1979), we observed that large values of $\tau_I$ can cause oscillatory behavior. $\tau_E$ controlled the qualitative behavior of the system: small values (< 2ms) allowed accurate tracking of turns at up to 600°/sec with little distortion of the hill shapes and small differences between PoS and ATN hill locations. Larger values of $\tau_E$ produced distorted hill shapes when tracking fast turns (even with the attractor dynamics in model ATN), but yielded larger differences between PoS and ATN hill locations.

Tuning curves of model excitatory PoS and ATN cells (say cell $i$) were determined by recording $F_i(t)$ at each time $t$ and storing $F_i(t)$ and the actual head direction $\psi_t$ (not the represented direction) at that point in the simulation. A histogram of $F_i$ was then generated, binned by head direction in 10° bins. Tuning curves of real PoS and ATN cells were generated from spike timing data supplied by Blair and Sharp. As can be seen in Figure 5.4, the model shows an excellent fit to the data.

Given the parameters used by Redish *et al.* (1996) (see Table 5.1), the direction represented by the ATN population leads the direction represented by the PoS population by approximately 10 ms over a wide range of turning speeds. Although this is smaller than the 20–40 ms discrepancy reported in neurophysiological experiments (Blair and Sharp, 1995; Taube and Muller, 1995; Taube *et al.*, 1996, EX.101), this model at least replicates the qualitative observation that ATN leads PoS.

## 5.6 Deformations of the tuning curve

We found that when the ATN population did not include attractor dynamics, the tuning curves of ATN component neurons deformed during turns by becoming asymmetric. Because early reports implied that ATN cell tuning curves did not deform during rotations (Blair and



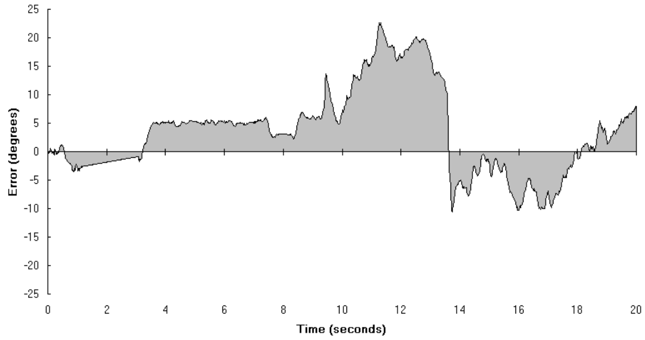




Figure 5.3: Simulated tracking ability. (Top) Black line: sequence of head orientations recorded by T. Blair and P. Sharp. Grey line: head direction represented by excitatory pool of simulated post-subiculum. Direction has been unrolled for clarity. (Bottom) Cumulative tracking error: angular difference between actual and simulated head directions. From Redish *et al.* (1996).

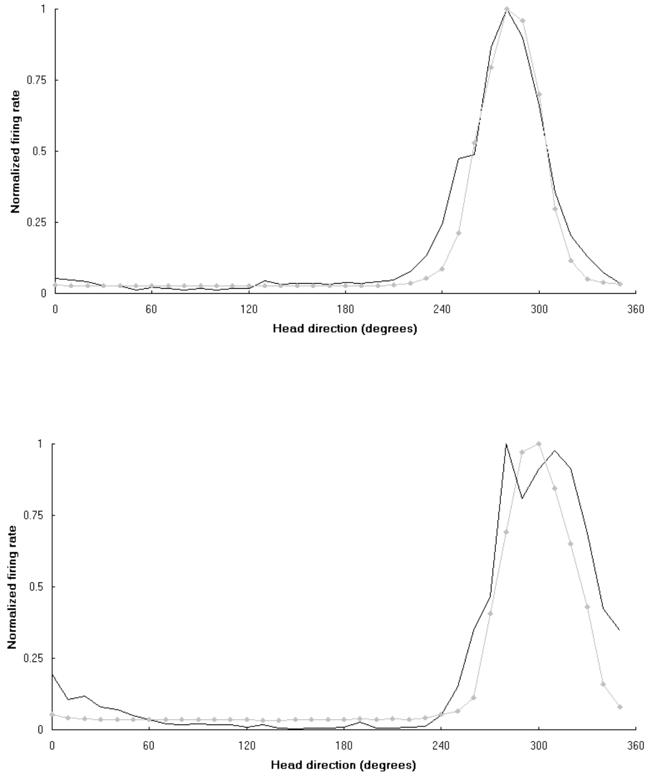



Figure 5.4: Simulated HD tuning curves. Black line: data from T. Blair and P. Sharp; grey line with diamonds: simulations. (top) Postsubiculum. (bottom) ATN. Note that these simulations included ATN as an attractor network, and thus the ATN tuning curve does not show the "two-peak" effect (EX.102, see Figure 2.24) From Redish *et al.* (1996).

Sharp, 1995; Taube, 1995), we simulated the model with attractor dynamics in the model ATN, but noted that this was incompatible with the anatomy of the anterior thalamic nuclei.

Attractor dynamics in the ATN require that there be (1) intrinsic excitatory connections and (2) inhibitory interneurons in the ATN. Most head direction neurons recorded from ATN have been recorded from the anterior dorsal (AD) nucleus (Taube, 1995; Blair and Sharp, 1995), but there is little evidence for either of these requirements in the AD nucleus (Bentivoglio *et al.*, 1993). Anatomical evidence indicates few if any interneurons in the rat ATN (although some interneurons have been found in cat and monkey) (Bentivoglio *et al.*, 1993). There do seem to be GAD-positive fibers in the anteroventral (AV) and anterodorsal (AD) nuclei, but not in the anteromedial (AM) nucleus (Bentivoglio *et al.*, 1993). GAD staining is indicative of the presence of GABA-mediated inhibition, but it is not clear where these GAD-positive fibers come from.

Recent data from Blair *et al.* (1997) show that ATN tuning curves do deform during rotations and that they deform in exactly the way we observed when ATN was not an attractor (EX.102, see Figure 2.24). Blair *et al.* (1997) report that when the animal is not turning, head direction cells in the ATN show a bimodality in the tuning curves. The separation of the two peaks is proportional to the *anticipatory time interval* (ATI) of the cell (i.e. the time in the future which produces the tightest tuning curve). Blair *et al.* also report that when the animal is turning in one direction or the other, one peak grows more than the other. This is exactly what one would predict given stronger offset than matching connections and no (or weak) attractor dynamics in the ATN.

Figure 5.5 shows how the two peaks can be accounted for by a model that includes strong offset connections and no attractor dynamics in ATN.[5] Work is ongoing to examine how well this new model accounts for the tuning curves of ATN cells with varied ATIs.

# 5.7 Other predictions and future work

## 5.7.1 The role of the postsubiculum

The anatomical instantiation of the attractor network theory suggests that the role of the postsubiculum is as the attractor network itself, i.e. as the $\mathcal{H}$ population of McNaughton *et al.* (1991) and the ring of *head direction cells* of Skaggs *et al.* (1995). This functional understanding of the role of the postsubiculum predicts that (1) the postsubiculum should be an attractor, and (2) lesions to the anterior thalamic nuclei or the postsubiculum should disrupt the head direction system as well as navigational strategies dependent on the system.

An aspect of first prediction can be tested using current technologies, but to my knowledge has not been. The attractor nature of the system (as a whole) could be directly demonstrated using multi-unit recording. If one records from multiple head direction cells simultaneously (say 20 cells), then the population encoding (Figure 3.5) can be seen directly.

[5] This work was done with Jeremy Goodridge and David S. Touretzky. Preliminary results will be presented at the 1997 Society for Neuroscience Conference, see Goodridge *et al.* (1997).

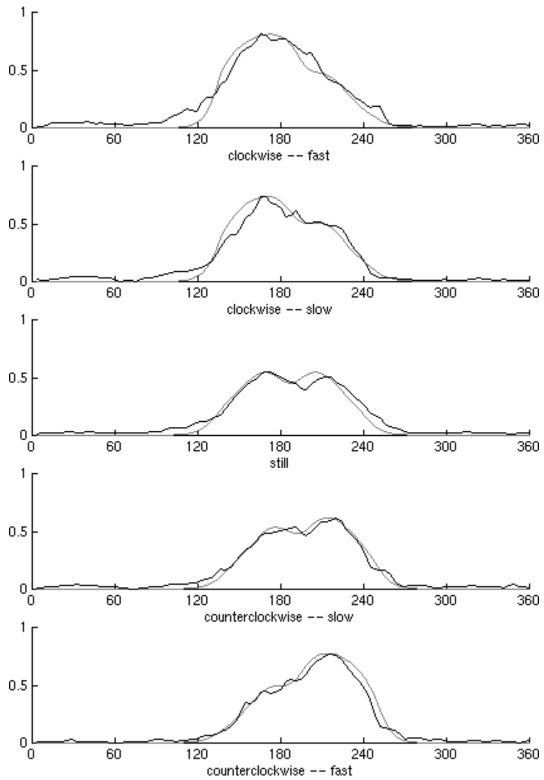



Figure 5.5: An ATN cell showing the "two-peak effect." (Black line) Data courtesy T. Blair and P. Sharp. Data has been interpolated and normalized. (Grey line) Computer simulation, courtesy of J. Goodridge.

If noise is injected into the system by microstimulation, then the representation should be transiently disrupted. Since nearby postsubicular cells encode different directions (Sharp, 1996a), microstimulation should effectively add noise into the system by firing a set of cells with unrelated preferred directions. If the system includes an attractor network, then, even in darkness, without external or self-motion cues, the PoS population's firing rates should return to a well-formed representation of head direction.

The second prediction, however, has been tested. Although anterior thalamic and postsubicular lesions do disrupt many spatial tasks (Sutherland and Rodriguez, 1990; Taube *et al.*, 1992, EX.22), and anterior thalamic lesions disrupt postsubicular cell head direction sensitivity (Goodridge and Taube, 1994; Taube *et al.*, 1996, EX.106), postsubiculum lesions do not seem to disrupt the directional selectivity of AD cells (Goodridge and Taube, 1994; Taube *et al.*, 1996, EX.108).

This last reviewed result is a problem for this theory. There are three possible explanations that can make the theory and data compatible again:

- The postsubiculum may be only part of a larger structure, such as the presubiculum (Taube *et al.*, 1990a), at which point the lesions made by Goodridge and Taube were incomplete. However, the anatomical connections of the ventral presubiculum are slightly different from that of the postsubiculum (van Groen and Wyss, 1990) suggesting that the two structures may really play different roles. No one has yet recorded from the presubiculum, so we don't know if there are head direction cells there.

- Another brain structure can play the role normally played by the postsubiculum. The cingulate cortex also sends projections to the ATN (Bentivoglio *et al.*, 1993; van Groen *et al.*, 1993), and Chen *et al.* (1994b) found head direction cells in cingulate cortex, so perhaps the cingulate cortex can take over the role normally played by postsubiculum.

- The theory may be wrong. For example, the head direction signal may be generated by deeper subcortical structures projecting through LMN to ATN and thence to PoS (Taube *et al.*, 1996). This is always a possibility, but this theory is compatible the all of the available data except for this one lesion study and I hesitate to suggest abandoning the theory without at least exploring the two other possibilities.

### 5.7.2 The role of the lateral mammillary nuclei

The simulations presented here (Sections 5.4 and 5.5) included three connection matrices from PoS:E to ATN:E: a set of *matching* connections, and two sets of *offset* connections. The "weight" of each offset connection was a function of the angular velocity of the animal ($\xi(\dot{\psi})$). The naive implementation of this would require multiplicative synapses in which both PoS cells and neurons with firing rates representing the angular velocity of the animal (actually $\xi(\dot{\psi})$) synapse jointly onto each ATN cell. A much more reasonable implementation would include a population of cells with firing rates correlated to head direction $\times$ angular velocity (again, actually, $\xi(\dot{\psi})$). Zhang (1996b) has shown simulations of a two-population model with each population sensitive to HD $\times$ AV. The simulations of Goodridge *et al.* (1997) (see Figure 5.5) include simulations of LMN which show HD $\times$ AV tuning curves.

McNaughton *et al.* (1991) first suggested that these HD $\times$ AV cells might play an important role in the system. McNaughton *et al.* called them $\mathcal{HH}'$ cells, and Skaggs *et al.* called them *rotation* cells. Both Taube *et al.* (1990a) and Sharp (1996a) report that some cells in postsubiculum are sensitive to both head direction and angular velocity. McNaughton *et al.* (1991) has also reported cells in parietal cortex with similar sensitivities to both head direction and angular velocity.

However, recent data from Taube *et al.* has suggested that the HD $\times$ AV role may be played by the lateral mammillary nuclei (LMN). LMN is situated in a loop with PoS and ATN (van Groen and Wyss, 1990; Bentivoglio *et al.*, 1993; Taube *et al.*, 1996) and receives vestibular input via the dorsal tegmental nucleus (Taube *et al.*, 1996). LMN cell activity is strongly correlated with angular velocity as well as direction (Leonhard *et al.*, 1996, EX.96).

This three-stage interpretation of the head direction system thus has three components: an attractor network which maintains a coherent representation of head direction (in PoS), a multiplicative network where head direction representations are combined with angular velocity information coming from the vestibular system (combined in LMN), and a representation of the head direction at time $t + \Delta t$ (in ATN). These three structures form a loop which may maintain a representation of the head direction system and update it from vestibular input.

# Chapter 6

# Path integration

As discussed in the Navigation Overview (see Section 3.2.3), path integration is the ability to return directly to a starting point (sometimes called a *home base* or *reference point*), even after a circuitous journey, even in the absence of external cues (Barlow, 1964; Gallistel, 1990; Maurer and Seguinot, 1995). A path integration ability has been demonstrated in gerbils (Mittelstaedt and Mittelstaedt, 1980; Mittelstaedt and Glasauer, 1991), hamsters (Etienne *et al.*, 1986; Etienne, 1987; Etienne *et al.*, 1988; Etienne, 1992; Chapuis and Scardigli, 1993), house mice (Alyan and Jander, 1994), rats (Tolman, 1948; Alyan *et al.*, 1997), birds (Mittelstaedt and Mittelstaedt, 1982; von Saint Paul, 1982), insects (Wehner and Srinivasan, 1981), and arthropods (Mittelstaedt, 1983), as well as dogs, cats, and humans (Beritashvili, 1965). See EX.57.

The basic idea of path integration is shown in Figure 3.6: position at time $t + \Delta t$ can be calculated from information about position, speed, and direction at time $t$. The main problem with path integration is that if your measurement of speed and direction are wrong, your representation of position will be increasingly inaccurate with ongoing movement.

Note that because errors in the path integrator can be corrected from local view information via the association stored in the place code, there may be some systematic error in the path integrator which is corrected by external cues. Maurer and Seguinot (1995) review data showing that systematic errors in path integration do exist in a variety of species, including rodents (see in particular Seguinot *et al.*, 1993). However, a path integrator that drifts too much will be useless.

## 6.1 Models of path integration

### 6.1.1 Kybernetik models

Early models of path integration addressed issues of the computational components of path integration, attempting to account for the animals' behavior, but generally not addressing the anatomy or neurophysiology (e.g. Jander, 1957; Mittelstaedt, 1962; Mittelstaedt, 1983; Mittlestaedt, 1985; Müller and Wehner, 1988; see Gallistel, 1990, and Maurer and Seguinot, 1995, for reviews). Many of the Kybernetik models address path integration in insects. I do not mean to suggest that path integration in insects and rodents necessarily occur by the same mechanism, but I will review the major Kybernetik models as they are possible explanations of path integration, even in rodents.

#### 6.1.1-A Jander (1957)

Jander (1957) suggested that animals maintained a representation of the *angle of return* $\alpha_r$ (the angle home) by calculating the integral

$$\alpha_r = \frac{1}{t_Z - t_N} \int_{t_N}^{t_Z} \alpha_t dt \tag{6.1}$$

where the journey lasts from $t_N$ to $t_Z$, and $\alpha_t$ is the angle turned at time $t$. Note that Jander's model does not include a representation of the distance necessary to return home, only the direction. As pointed out by Maurer and Seguinot (1995), although this equation works for specific cases, it does not match either perfect path integration or the actual abilities of insects.[1]

#### 6.1.1-B Mittelstaedt (1962).

Mittelstaedt (1962) suggested that animals explicitly represent the Cartesian coordinates of the path home, so that after a step of distance $d$ at angle $\phi$, the animal changes its representation of position by

$$\begin{aligned} x(t+\Delta t) &= x(t) + d(t)\cdot\cos(\phi(t)) \\ y(t+\Delta t) &= y(t) + d(t)\cdot\sin(\phi(t)) \end{aligned} \tag{6.2}$$

This corresponds to the mathematically ideal path integration. Noise in the system produces scatter in the position returned to (Benhamou *et al.*, 1990), however, as pointed out by Maurer and Seguinot (1995), this model cannot explain systematic errors in path integration.

Fujita *et al.* (1990) examined a linearization of the mathematically accurate path integrator model but they found it a very poor fit to the path integration abilities of ants. (It made errors in the opposite directions from the actual ants.)

[1] This description of Jander (1957) [original in German] is a paraphrasing of the review of Maurer and Seguinot (1995).

#### 6.1.1-C Müller and Wehner (1988).

Müller and Wehner (1988) presented a mean-direction hypothesis in which the animal keeps track of the mean direction home by weighted averages of distance traveled along each direction. Although this model does not address neurophysiology, it addresses inaccuracies seen in ants.

Müller and Wehner (1988) suggested that the representation of the direction home is updated at each step by the following equation:

$$\phi_{n+1} = \phi_n + k\frac{(180° + \delta)(180° - \delta)}{l_n} \tag{6.3}$$

where $\delta$ is the angle turned in step $n$, $k$ is a constant (empirically determined to be $4.009 \times 10^{-5}\,\mathrm{deg}^{-2}$ for *cataglyphis fortis* ants), and $l_n$ is a measure of the distance from the home base and is updated by:

$$l_{n+1} = l_n + 1 - \frac{\delta}{90°} \tag{6.4}$$

This model is not a true path integration model in that it does not perfectly track the vector home, but Müller and Wehner show that it does fit the path integration abilities of ants remarkably well. Seguinot *et al.* (1993) show that it is also a reasonable approximation to the path integration abilities of hamsters taken on short journeys consisting of 1–5 straight segments.

### 6.1.2 Connectionist models

The Kybernetik models do not attempt to address the neurophysiology of path integration. There have been three major models of path integration in rodents which address the neurophysiology.

#### 6.1.2-A Early McNaughton *et al.* models (McNaughton, 1989; McNaughton and Nadel, 1990; McNaughton *et al.*, 1991; McNaughton *et al.*, 1994a).

McNaughton *et al.* (McNaughton, 1989; McNaughton and Nadel, 1990; McNaughton *et al.*, 1991; McNaughton *et al.*, 1994a) suggested that path integration could be accomplished by an associative memory in which a representation of position was associated with a representation of spatial displacement. The association produced a representation of the new position. The discrete form of this is a table lookup model and the continuous version can be described by a linear associator. No simulations of these proposals were reported.

These models do not show true path integration in that they require exploration of the environment before an animal can show path integration in that environment. This is because they use self-motion information to regenerate local views. Because animals can path integrate from a first experience in an environment (Beritashvili, 1965; Mittelstaedt and Mittelstaedt, 1980; Wehner and Srinivasan, 1981; von Saint Paul, 1982; Alyan and Jander, 1994), these models are insufficient to fully describe path integration.

#### 6.1.2-B Touretzky *et al.* (1993).

Touretzky *et al.* (1993) proposed a representation of 2D vectors that accommodated several vector arithmetic operations including addition, in which the error grew linearly. Because the error grows linearly, this model can be seen as an implementation of the vector addition theory of Mittelstaedt (1962) and the effect of noise will be similar to that seen by Benhamou *et al.* (1990).

The representation consists of a population of cells in which each cell fires at a rate

$$F_i = b_i + k_i \cdot r \cos(\phi - \phi_i) \tag{6.5}$$

where $b_i$ is a baseline firing rate, $k_i$ is a gain parameter, and $\phi_i$ is a preferred direction, analogous to that of head direction cells. This representation is one possible extension of the head direction representation shown in Chapter 5. It is also an extension of a represention found in primary motor cortex and first described by Georgopoulos *et al.* (1983). We have shown that some motor control results can be explained by it (Redish and Touretzky, 1994).

Wittmann and Schwegler (1995) showed how this encoding could be used to model path integration in ants, however, no such cells have yet been found in the rodent.

#### 6.1.2-C Two-dimensional attractor networks.

Another extension of the head direction representation was suggested as a model of path integration by McNaughton *et al.* (McNaughton *et al.*, 1996; Samsonovich and McNaughton, 1997; Samsonovich, 1997) and Zhang (1996a). In this extension each cell has a preferred spatial coordinate (in two dimensions) and fires at a rate proportional to a Gaussian of the distance between the represented coordinates and its preferred coordinates. Preferred coordinates are analogous to preferred directions in the head direction representation.

An appropriately formulated connection matrix will produce a two-dimensional attractor analogous to the one-dimensional attractor described in Chapter 5 (Kohonen, 1982; Kohonen, 1984; Droulez and Berthoz, 1991; Munoz *et al.*, 1991; Arai *et al.*, 1994; McNaughton *et al.*, 1996; Zhang, 1996a; Samsonovich and McNaughton, 1997; Samsonovich, 1997); offset-connections can move the two-dimensional representation around just as they can the one-dimensional representation in Chapter 5 (Droulez and Berthoz, 1991; Munoz *et al.*, 1991; Arai *et al.*, 1994; Zhang, 1996a; McNaughton *et al.*, 1996; Samsonovich and McNaughton, 1997; Samsonovich, 1997).

The basic theory is a straight-forward extension of the head direction system (see Section 3.2.2 and Chapter 5). Samsonovich and McNaughton (1997) suggest that the system works in two stages: a $\mathcal{P}$-stage which represents the position (analogous to the representation of head direction role assigned to PoS in Chapter 5) and an $\mathcal{I}$-stage which represents position $\times$ velocity (analogous to the representation of head direction $\times$ angular velocity role assigned to LMN in Chapter 5). One can also imagine a third stage (which I will label $\mathcal{R}$ for *reintegration*, with a role analogous to the role assigned to ATN in Chapter 5). This third stage was not separated from the $\mathcal{P}$-stage in Samsonovich and McNaughton (1997). All three stages were incorporated into a single population in Zhang (1996a). There are two possible connection structures that can accommodate this process. I will review both here, but further work needs to be done to differentiate between them.

**A**. One possibility (the one suggested by Samsonovich and McNaughton, 1997) is that the $\mathcal{I}$ cells are sensitive to place, direction, and speed, i.e. they have a preferred orientation $\tilde{\phi}_i$ and a preferred speed $\tilde{\delta}_i$ as well as preferred coordinates $\langle \tilde{x}_i, \tilde{y}_i \rangle$. These cells would then synapse on $\mathcal{P}$ cells with preferred coordinates $\langle \tilde{x}_i + \tilde{\delta}_i \cdot \cos(\tilde{\phi}_i), \tilde{y}_i + \tilde{\delta}_i \cdot \sin(\tilde{\phi}_i) \rangle$. These cells are 2D analogues of the AV $\times$ HD cells in the head direction model.

**B**. Another possibility not explicitly addressed by Samsonovich and McNaughton is that each $\mathcal{P}$ cell connects to a population of $\mathcal{R}$ cells in a ring around the corresponding preferred coordinates. Each point in the ring receives multiplicative input from the head direction representation in the corresponding direction. So the effective connection strength $w_{ijk}$ between three cells, $\mathcal{I}$ place cell $i$ with preferred coordinates $\langle \tilde{x}_i, \tilde{y}_i \rangle$, head direction cell $j$ with preferred direction $\tilde{\phi}_j$, and $\mathcal{R}$ place cell $k$ with preferred coordinates $\langle \tilde{x}_k, \tilde{y}_k \rangle$ is proportional to

$$w_{ijk} \propto \| \langle \tilde{x}_k, \tilde{y}_k \rangle - \langle \tilde{x}_i + \delta \cdot \cos(\tilde{\phi}_j),\ \tilde{y}_i + \delta \cdot \sin(\tilde{\phi}_j) \rangle \| \tag{6.6}$$

where $\delta$ is a radial offset. In the animal, $\delta$ is more likely to be a distribution than a constant. Effective connection strength can occur as a function of probability of connection or synaptic weight or both. The effect of this synapse on cell $k$ will be $w_{ijk} \cdot F_i F_j$. That is, it will be a three-way, multiplicative synapse. This is analogous to the weighting of inputs from PoS to ATN discussed in the original Redish *et al.* (1996) HD model and reviewed in Chapter 5.

## 6.2 Anatomy

### 6.2.1 Computational requirements

Having reviewed the previous models of path integration, I now address the question of what brain structures are directly involved in path integration. The brain structures that allow the rodent to accomplish path integration must satisfy the following requirements:

1. **They must collectively be able to represent the position of the animal.** That is, the cells must show activity patterns correlated with the position of the animal.
2. **They must receive input from the head direction system.** In order to update the representation of position, the path integration system must know the direction the animal is moving.[2]

[2] Although head direction is generally correlated with direction of motion, animals can obviously travel in directions they are not facing. It is not clear how the path integration system handles motion in a

3. **They must receive information about self-motion from the motor and vestibular systems.** In addition to direction of motion, path integration requires information about the speed of the animal.

4. **They must update the representation as the animal moves around the environment.** In other words, the path integration system must be able to perform the equivalent of the vector addition shown in Figure 3.6, at least for small vectors along the direction of motion.

5. **They must send output to the area associated with the place code.** If hippocampal place cell activity in the dark is a consequence of path integrator input (O'Keefe, 1976; McNaughton *et al.*, 1989b; Leonard and McNaughton, 1990; Markus *et al.*, 1994; Wan *et al.*, 1994b; Wan *et al.*, 1994c; McNaughton *et al.*, 1996; Touretzky and Redish, 1996; Redish and Touretzky, 1997a; Samsonovich and McNaughton, 1997, see also EX.71), then the hippocampus must either be directly involved in path integration or the cells must receive input from the path integrator.

### 6.2.2 Previous hypotheses

#### 6.2.2-A Hippocampus

McNaughton *et al.* (McNaughton and Nadel, 1990; McNaughton *et al.*, 1996; Samsonovich and McNaughton, 1997; Samsonovich, 1997) suggested that the key to a rodent's ability to path integrate lies in the hippocampus. The place cell representation can represent the position of the animal, so it meets criterion 1.

Each place cell can be modeled as a Gaussian representation of position. Because the set of Gaussians covering all positions forms a basis set representation of position, the population of place cells can be viewed as a basis set.[3] Basis sets have the important property that any function of their inputs can be approximated by a linear combination of the activities of the elements of the basis set. This means that any function of position can be calculated by a linear transformation applied to the representation of position in the place code. This includes the position update function necessary for path integration.

Because animals can return to their starting point even in novel environments (Eilam and Golani, 1989; Leonard and McNaughton, 1990; Golani *et al.*, 1993; Touretzky *et al.*, 1996, see

direction different from the animal's heading (McNaughton *et al.*, 1996; Touretzky and Redish, 1996; Redish and Touretzky, 1997a; Samsonovich and McNaughton, 1997). One possibility is that there is an override mechanism which updates the path integrator correctly; alternatively, the path integrator might produce incorrect results. A third possibility (B. McNaughton, personal communication) is that some head direction cells are actually tuned to direction of motion and that when the animal moves in a direction it is not facing, these cells will fire relative to the direction of motion, not the direction the animal is facing. Experiments have not yet been done to distinguish between these three possibilities.

[3] Gaussian representations covering a space are sometimes called *Radial Basis Function* representations and are used in many neural networks (Hertz *et al.*, 1991, §9.7).

Experiment Results EX.60 and EX.130), the path integrator is probably not dependent on exploration of the environment. This means that the mechanism subserving path integration must be already wired up before the animal enters an environment. However, the topology of place fields changes from environment to environment (O'Keefe and Conway, 1978; Kubie and Ranck, 1983; Thompson and Best, 1989, EX.79, see also Section 3.3), which means that the update function allowing the vector arithmetic (point 4) would have to be be pre-wired in the hippocampus for each separate reference frame. Samsonovich and McNaughton (1997) have suggested exactly this, calling each pre-wired reference frame a *chart*.

The mechanism suggested by McNaughton *et al.* and Zhang (Section 6.1.2-C) requires a very complex interaction between the head direction system and the representation of location in the path integrator: in one case (Theory A) there must be an interpretation function which transforms environment-sensitive representation of location (in the $\mathcal{P}$-stage, identified with hippocampus) into environment-independent representation of location $\times$ velocity. In the other version (Theory B) there must be multiplicative connections between the $\mathcal{P}$ stage and the $\mathcal{I}$ stage which are correctly matched with head direction input. Samsonovich and McNaughton (1997, see also Samsonovich, 1997) have argued that this complex connection matrix may be trained up early in life and that a limited number of environment-dependent maps (called *charts* by Samsonovich and McNaughton) make the theory plausible, but they have not given any method (nor shown any simulations) by which this connection structure could be trained. I find the identification of the $\mathcal{P}$-stage with the hippocampus implausible and unnecessarily complex. I will show (see below) an alternative hypothesis that does not require the complex pre-wiring required by this instantiation.

A strong prediction of this hypothesis is that hippocampal lesions should devastate path integration in rodents. The recent experiments by Alyan *et al.* (1997, EX.59) showing that rodents can perform path integration tasks even with cytotoxic lesions of hippocampus are a serious problem for the hippocampus as path integrator theory.

#### 6.2.2-B Subiculum

Sharp (1996b) suggests the subiculum as the locus of the path integrator because subicular cells show similar place fields across different environments (EX.92). This result suggests that the subiculum does not have the problem just discussed for hippocampus. Because each environment is represented by the same subset of place cells (Sharp, 1996b), only one update function needs to be learned. This means that it can be learned once early in life, or even pre-wired genetically. Because there are place cells in subiculum (Sharp and Green, 1994; Sharp, 1996b), it can represent the current position of the animal (meeting criterion 1), and subicular cells do show a (weak) directional signal (thus meeting criterion 2) (Sharp and Green, 1994). However, the subiculum does not send output directly to the place code; it sends output to the postsubiculum, to the parasubiculum, to the deep layers of the entorhinal cortex, and to a variety of other structures via the fornix but not to the hippocampus (Kohler, 1986; Kohler, 1988; Witter *et al.*, 1990). Thus the subiculum acting alone does not meet criterion 5.

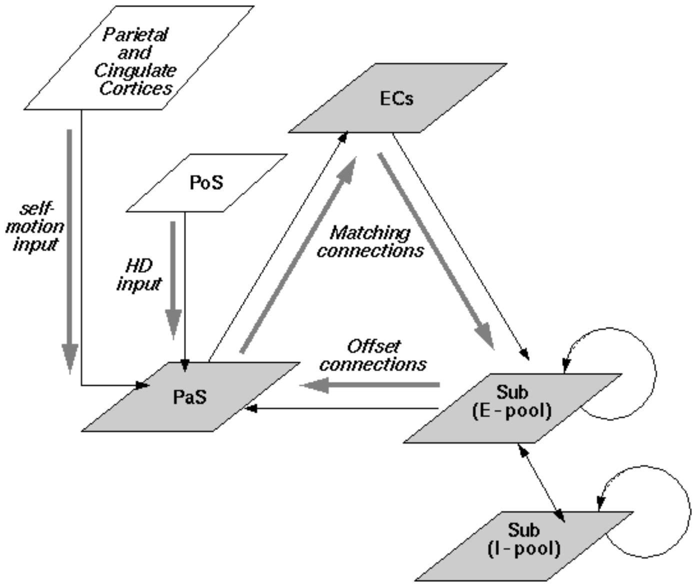



Figure 6.1: The path integrator loop. ECs: superficial layers of entorhinal cortex; Sub: subiculum; PaS: parasubiculum; PoS: postsubiculum.

### 6.2.3 A proposal

We[4] propose that the path integrator is a two-dimensional attractor as reviewed above (Theory 6.1.2-C), but that path integration occurs via a loop composed of the subiculum (Sub), the parasubiculum (PaS), and the superficial layers of entorhinal cortex (ECs). They are connected in a loop as shown in Figure 6.1 (Kohler, 1986; Kohler, 1988; Witter *et al.*, 1990; van Groen and Wyss, 1990; Wyss and van Groen, 1992). PaS receives input from the postsubiculum (van Groen and Wyss, 1990) which is a key component of the head direction system (see Chapter 5). And PaS is interconnected with posterior cingulate cortex (Wyss and van Groen, 1992), which includes directional and behavioral representations (Chen *et al.*, 1994a) that could supply self-motion information.

The proposed loop of Sub, PaS, and ECs meet the five criteria listed above:

1. Because each of these areas show place cells, each of the three can represent the position of the animal (Quirk *et al.*, 1992; Sharp and Green, 1994; Taube, 1996).

2. The parasubiculum receives input from the postsubiculum (van Groen and Wyss, 1990), which represents the head direction of the animal (see Chapter 5).

[4] This work was done with D. S. Touretzky and has been published in abstract (Redish and Touretzky, 1996a) and journal (Redish and Touretzky, 1997a) form.

3. The parasubiculum is interconnected with parietal and cingulate cortex (Wyss and van Groen, 1992), which includes directional and behavioral representations (Chen *et al.*, 1994a).

4. There is no data showing that the update mechanism described in Section 6.1.2-C occurs in the Sub-PaS-ECs loop, but it is plausible because only one complex connection structure would have to be learned. Thus a mechanism that produced such a structure early in life (for example from exploration in an early-experienced environment) could be used in any subsequent environment. I will take this mechanism as a hypothesis and note some predictions that come from it below (see Section 6.4).

5. As required, this path integrator proposal includes direct projections into the hippocampus (from ECs, a proposed component of the path integrator.)

## 6.3 Issues: edge effects

There are three possible mechanisms that might allow the system to handle problems at the edges of its representations:

- **The map might be toroidal**, which would mean that there are no edges. Representations that fall off the top appear at the bottom while representations that fall off the left edge appear on the right and vice versa. If this were true, this would make a very strange and intriguing prediction: when exposed to a large enough environment, the place fields recorded from the path integrator components (Sub, PaS, ECs) would repeat after a certain distance travelled.

- **The map might be scalable**, so that when faced with a large environment, the place fields of the PI components would scale appropriately. This is easily testable, but there is a limit to how large an environment scaling can accommodate. As fields get larger, the accuracy decreases, because as the fields enlarge, a distance $\Delta x$ is distinguished by smaller and smaller changes in firing rates. In addition, if an animal misjudged the size of an environment, the scale of the place fields would be incorrect and some recovery mechanism would have to exist.

- **The system can shift centers when it approaches the edge of the representation.** This would ensure that it can always have a valid representation of its position in some coordinate frame, but if the animal is to be able to path integrate back across that *reference frame shift*, then there must be a representation of how the two reference frames are associated. Evidence supporting this third possibility was discussed in Section 3.3.

## 6.4 Predictions

### 6.4.1 Prediction of position

Whichever mechanism produces the offset connections, the post-offset area should show prediction of position (much like ATN does in the head direction system, Blair and Sharp, 1995; Taube and Muller, 1995; Blair *et al.*, 1997, EX.101). Taube (1996) reports that parasubicular place cells are optimally tuned to a slight prediction of current position. This is similar to the demonstration that ATN cells predict future orientation (Blair and Sharp, 1995), but more recent work by Blair *et al.* (1997) suggests that the orientation prediction in ATN is occurring as a consequence of offset connections impinging on ATN cells (see Figure 2.24). No such demonstration has yet been made for the parasubicular place cells.

### 6.4.2 Multiplicative connections

According to the second possible instantiation of the mechanism hypothesized to underlie path integration (see Section 6.1.2-C, paragraph B), neurons in the pre-offset area should connect to neurons in the post-offset area with a ring structure and the connections should be multiplicative with head direction input. Because none of the areas (Sub, PaS, ECs) have cells laid out topographically with respect to their preferred coordinates, this ring structure will not be anatomically evident, but the multiplicative synapses would be visible by double-staining methods at the electron microscope level. The connection structure might also be discernible from correlational measurements of simultaneous recordings from multiple structures.

### 6.4.3 Activity in the dark

Cells in all three structures should show normal activity in the dark. I know of no recordings from any of these structures in the absence of sensory cues, although Quirk *et al.* (1992) does report that ECs cells continued to show normal place fields when a cue card was removed.

### 6.4.4 Topology of place fields

The topology of the place fields in these structures should never change, although the overall layout may translate or rotate. That is: if one measures the place field of a cell (say $a$) in one environment, nothing can be predicted about the location of its place field in another environment (except that it must exist). On the other hand, if one measures the place fields of two cells (say $a$ and $b$) in one environment, the distance between their place fields in any other environment must be identical. If one records the place fields of three cells (say $a$, $b$ and $c$) in one environment, one cannot predict where the center of the triangle formed will be, but the triangle their place fields form must be identical (modulo translation, rotation and scaling) across environments.

### 6.4.5 Lesions

Lesions to any of these structures (Sub, PaS, ECs), should impair the ability of the animal to return to its starting point after a circuitous path in the absence of sensory cues (i.e. to path integrate). Lesions to the hippocampus should not impair this ability.

In order to explicitly test this last prediction, Alyan *et al.* (1997) tested rats in the classic Mittelstaedt and Mittelstaedt (1980) and Etienne (1987) paradigm (EX.59). Alyan *et al.* (1997) found that rats with ibotenic hippocampal lesions showed normal path integration abilities in this task. This result disproves the hypothesis that the hippocampus is a critical component of behavioral path integration.

# Chapter 7

# Place cells

Spikes fired by dentate granule cells, and CA3 and CA1 pyramidal cells are strongly correlated with the location of the rat: each cell fires when the animal is in a specific place (called the *place field* of the cell). These *place cells* are some of the most-studied neurons in the rodent brain; they have been examined in a wide variety of environmental manipulations (see Experimental review, Section 2.2.2-A). The hippocampus is also probably the most extensively modeled system in the rodent brain. I will begin this chapter with a review of the critical experimental results and then review previous models of place cells. I will then put these ideas together into an amalgam model of place cells which I will use throughout the rest of the thesis.

## 7.1 Place cell properties

Because place cells show such a clear correlate between firing rate and a spatial variable, many experiments have been done to explore how they react to environmental manipulations. As shown in Figure 2.18, a typical place field covers an area a few times the size of the rodent. But place cells show a number of other properties. When I review the place cell models, I will compare the models to these properties:

- When distal landmarks are moved, place fields also move proportionately (Muller and Kubie, 1987; O'Keefe and Speakman, 1987; McNaughton *et al.*, 1994a; Knierim *et al.*, 1995; Cressant *et al.*, 1997, EX.67).
- Place cells continue to show clean place fields when landmarks are removed (O'Keefe and Conway, 1978; Muller and Kubie, 1987; O'Keefe and Speakman, 1987; Pico *et al.*, 1985, EX.70).
- Place cells continue to show compact fields in the dark (O'Keefe, 1976; McNaughton *et al.*, 1989b; Quirk *et al.*, 1990; Markus *et al.*, 1994, EX.71).

- Firing rates of hippocampal granule and pyramidal cells are correlated to more than the location of the animal, (EX.78), including odor set (Eichenbaum *et al.*, 1987; Eichenbaum and Cohen, 1988; Otto and Eichenbaum, 1992; Cohen and Eichenbaum, 1993, EX.86), task (Markus *et al.*, 1995, EX.85), and stage of task (Eichenbaum *et al.*, 1987; Otto and Eichenbaum, 1992; Hampson *et al.*, 1993, EX.87).
- Place cells show different place fields in different environments (O'Keefe and Conway, 1978; Kubie and Ranck, 1983; Muller and Kubie, 1987; Sharp *et al.*, 1990; Bostock *et al.*, 1991; Sharp *et al.*, 1995; Sharp, 1996b, EX.79).
- Place cells are directional when an animal takes limited paths, but non-directional when wandering randomly on open fields. (McNaughton *et al.*, 1983a; Markus *et al.*, 1994; Muller *et al.*, 1994; Gothard *et al.*, 1996a, EX.64).
- Not all place fields are convex, some fields are crescent-shaped (Muller *et al.*, 1987); they hug the arena walls (see Figure 2.20, EX.65).
- Not all place fields are compact. Some cells show multiple subfields, even within a single environment (see Figure 2.21, EX.66).

## 7.2 Models of place cells

From the first discovery of place cells in the 1970s, people have questioned what drives hippocampal pyramidal cells to show place fields. There is no direct cue which tells the animal when it is at a certain location; it must use a combination of cues to determine its location. This was the original hypothesis proposed by O'Keefe and Nadel (1978): locale navigation depends on a combination of cues, in contrast to taxon or praxic navigation, which depends on a single cue.

As reviewed in the Chapter 4, many theories of hippocampal function have been proposed. There are also a number of specific models of place cells, many including simulations and comparisons to data (Zipser, 1985; Zipser, 1986; O'Keefe, 1989; McNaughton and Morris, 1987; McNaughton *et al.*, 1989b; McNaughton, 1989; Leonard and McNaughton, 1990; Sharp, 1991; Treves *et al.*, 1992; Hetherington and Shapiro, 1993; Schmajuk and Blair, 1993; Shapiro and Hetherington, 1993; Burgess *et al.*, 1994; Wan *et al.*, 1994b; Wan *et al.*, 1994c; McNaughton *et al.*, 1994a; Burgess and O'Keefe, 1996; O'Keefe and Burgess, 1996; Recce and Harris, 1996; Touretzky and Redish, 1996; Fuhs *et al.*, 1997; Redish and Touretzky, 1997a; Samsonovich and McNaughton, 1997). These models fall into three major classes, *local view models*, which depend solely on the local view to explain place cell firing, *path integration models*, which depend on a combination of local view and path integration to form the place field, and *associative memory models*, which depend on internal dynamics of the hippocampal network to produce the key place field properties.

Before discussing these models in depth, let us take note of some mathematical properties of space. Local view is a high-dimensional continuous space, consisting presumably of representations of spatial aspects of local and distal landmarks, such as distance and egocentric or allocentric bearing. Although local view is a high dimensional space, in any single environment, the position of the animal at any point in time can be described by two variables (its coordinates on the plane). This means that the animal only experiences a two-dimensional manifold of the high-dimensional space. Both the local view representations and this manifold share an important property: they are continuous.[1] This means that a cell tuned to a compact section of the high-dimensional space will also be tuned to a compact section of the two-dimensional manifold, i.e. it will be a *place cell* and will show a *place field*. Any learning mechanism that allows cells to distribute themselves around this two-dimensional manifold will produce place cells.

This also means that any sufficient subset of spatial cues will force cells to show small, compact fields, where "sufficient" means that they are enough to specify a point in space. For example, distance to three landmarks, distance and allocentric orientation to a single landmark, allocentric orientation to two landmarks, etc. are all sufficient. It should also be noted that these properties do not have to be measured directly. Because distance is correlated with other aspects, such as retinal size, height relative to the horizon, etc., other aspects could also be used instead of distance. Differentiating between these input sets is extremely difficult; it would require very accurate measurements of place cell activity as one manipulates single cues. Due to the recurrent interactions in hippocampus and the consequent auto-associative properties (see below), it may even be impossible.

### 7.2.1 Local view models

The first neural model of place cells (including simulations) was presented by Zipser (1985; 1986). It was a radial basis function model in which each cell was assumed to be tuned to a set of cues; the cell fired with a rate proportional to a sum of Gaussians dependent on retinal area subtended by the distal landmarks.

O'Keefe (1989) suggested a local view model based on egocentric bearing to landmarks. In this model, the hippocampus determines the location of the animal by taking the centroid of the bearing to three or more distal landmarks, but no simulations were reported, nor were any direct comparisons to actual place fields made.

McNaughton *et al.* (1994a) proposed that each place cell could be understood as representing the vector from an animal to a specific landmark. This theory predicted that when two landmarks are separated by more during testing than during training, the place fields should dissociate, so that some fields follow one landmark, while other fields follow the other (McNaughton *et al.*, 1994a, p. 593). The data from Gothard *et al.* (1996a) that the population of place fields always represents a coherent location when the length of a track was shortened disproved this hypothesis.

Other local view models that did include simulations include those of Sharp (1991)

[1] Local view is not strictly continuous because of occlusions, but it is at least piece-wise continuous. It should be noted that none of the published models actually model the effect of occlusions.

which used inputs sensitive to distances to landmarks at the edge of the environment, Fuhs *et al.* (1997) which used retinal size as a distance correlate, and Burgess *et al.* (Burgess *et al.*, 1994; Burgess and O'Keefe, 1996; O'Keefe and Burgess, 1996) which incorporated a distance modifier: the closer the landmark, the more accurately distance was encoded. This has little effect on the basic observation that Gaussians of enough spatial parameters can encode a single location.

Neither the Zipser nor the Burgess *et al.* models incorporate learning. A number of other local view models did examine the effects of learning rules, including competitive learning (Sharp, 1991; Fuhs *et al.*, 1997), and genetic algorithms (Treves *et al.*, 1992). The learning mechanisms have minor effects on the place fields.

For example, Sharp (1991) used a limited field of 300° in her local view input. Combined with competitive learning, this makes place cells along repeated paths directional. If the rodent samples the local view at a place from a discrete number of directions, competitive learning separates views to represent the place differently at each discrete direction. Theoretically, at some level of sampling, the discrete directions will merge and become an omni-directional place field. Sharp shows that her model place fields are directional in radial mazes, but remain omni-directional in open fields. This theory predicts that cells in the center of the radial maze will not be directional, because all directions are sampled there. Markus *et al.* (1995) report that cells in the center of the radial maze are directional, which is inconsistent with the prediction.

Fuhs *et al.* (1997) showed that a product-of-Gaussians model based on noisy distance metrics such as retinal angle can produce excellent place cell models using real data, demonstrating place fields without explicit object recognition mechanisms. We took 25 black-and-white camera views of a complex scene of distal cues. These pictures were taken from one direction at each point on a $5 \times 5$ grid. Using standard visual preprocessing algorithms, we identified *blobs* (hyperellipses identified by color, total size, and squareness) in each image, and tracked the blobs from one image to the next. Using competitive learning, we trained cells on these blobs and showed place-like activity even from extremely noisy real data.

All of these models share the property that they are driven solely by local view. This means that they cannot explain place cell activity in the dark. One could hypothesize that place cell activity in the dark is driven by non-visual sensory cues, but that hypothesis is disproved by observations from Quirk *et al.* (1990) and Knierim *et al.* (1995) that place fields can drift.

Recce and Harris (1996) suggested what is essentially a local view model of hippocampal place cells based on storing "snapshots of cortical input," but they assumed that the local view representations could be updated by path integration in the dark. They assume that there is an egocentric representation of landmarks in cortex and that the hippocampus is an autoassociator storing egocentric maps. In order to explain the continued activity of place cells in the dark, Recce and Harris hypothesize an extrahippocampal path integrator which updates the egocentric map. Updating representations of landmarks by path integration is equivalent to the *virtual landmark* hypothesis (Muller *et al.*, 1991a) and tends to be computationally unstable. Because each of the landmarks is being tracked separately in the

Figure 7.1: One of the images used by Fuhs *et al.* (1997). Blobs found in the scene are indicated by white rectangles with small crosses denoting the center of the blob.

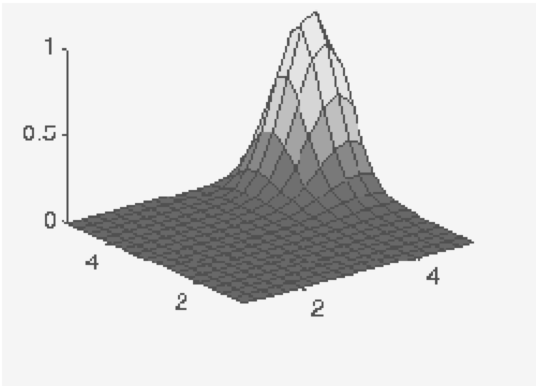



Figure 7.2: A sample simulated place field from Fuhs *et al.* (1997).

Recce and Harris model, errors that do not build up identically in each representation will distort the overall map.

### 7.2.2 Path integration models

In contrast to the local view models, a number of authors have suggested that place cell firing in the absence of cues is the direct result of path integration (O'Keefe, 1976; Muller *et al.*, 1991a; Wan *et al.*, 1994c; McNaughton *et al.*, 1996; Touretzky and Redish, 1996; Redish and Touretzky, 1997a).[2]

The first model to simulate place cells driven by both local view and path integrator information was that of Wan *et al.* (1994c, see Touretzky and Redish, 1996).[3] We modeled place cell activity by a product of six Gaussians including two Gaussians tuned to distance, two to allocentric bearing, one to retinal angle between a pair of landmarks, and one to path integrator coordinates.

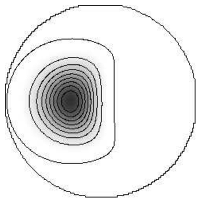


Figure 7.3: Simulated place field from Touretzky and Redish (1996) of a place cell partially tuned to a surface landmark.

We should note that in Touretzky and Redish (1996), we also included local surface orientation in the local view, so some of the landmarks were surface landmarks (such as the arena wall). Distance and bearing to a surface landmark were calculated as the normal to the surface (i.e. the vector of the shortest distance between the animal and the landmark). This allowed us to model crescent-shaped fields.

In Touretzky and Redish (1996), we suggested that when one of the cues was unavailable the gain on that Gaussian became negligible. Gain was defined as the inverse of the standard deviation of the Gaussian (i.e. $\sigma^{-2}$), so when gain became negligible, the Gaussian approached 1.0. Effectively, these inputs dropped out of the product equation.

[2] The early McNaughton *et al.* models (McNaughton *et al.*, 1989b; McNaughton, 1989; Leonard and McNaughton, 1990; McNaughton *et al.*, 1991; McNaughton *et al.*, 1994a) had some similarities to path integration models in that they included abilities to navigate in the dark by updating place representations from self-motion information. Because they accomplish this by associations between local views and self-motion information, they require exploration before showing path integration abilities. These models are better described as *associative memory models* and will be reviewed in Section 7.2.3.

[3] The papers presented prior to Wan *et al.* (1994c) hypothesized that place cells were partially driven by path integrator information, but none of them included explicit simulations.

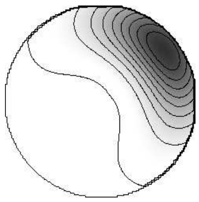


Figure 7.4: Simulated crescent-shaped place field from Touretzky and Redish (1996).

For example, when the animal navigated in the absence of cues, the first five Gaussians all dropped out, leaving only the path integrator input driving the cells. Conversely, when the animal was first placed into an environment, it could see the landmarks, but would not know its path integrator coordinates, so the last Gaussian would drop out and the place cell would be driven solely by local view information. We showed how, in this formulation, the place code could be used to reset the local view from path integrator information, and, conversely, the path integrator from local view information.

We showed that the model fits a large subset of the data, including the existence and shape of both normal convex place fields and crescent-shaped fields, as well as that place fields are tied to local landmarks, that they are unchanged when landmarks are removed, that place cells continue to show activation in the dark, and that place fields can be controlled by entry point. We simulated cue manipulations in radial maze tasks, disorientation in a rectangular arena, and navigation using arrays of local landmarks. We also hypothesized the existence of (but did not show simulations for) reference frames, which allowed us to explain directional place cells, place cells showing different place fields dependent on environment, task-dependent place fields, and goal-sensitive place fields using a single mechanism.

In an alternate model, McNaughton *et al.* (McNaughton *et al.*, 1996; Samsonovich and McNaughton, 1997) proposed that the hippocampus is, in fact, the path integrator proper. They took the hypothesis that synaptic weight between place cells is inversely proportional to the overlap of their place fields (Wilson and McNaughton, 1994, also known as the *cognitive graph*, Muller *et al.*, 1991b; 1996) as a basic starting point and proposed the existence of a loop between hippocampus and subiculum which performs path integration. Recent data from Alyan *et al.* (1997) has shown that animals can path integrate with hippocampal lesions, disproving this hypothesis.

### 7.2.3 Associative memory models

In addition to the local view and path integrator models, there are a number of *associative memory* models in which the place fields are essentially assumed to be driven by local view input, but in which autoassociative properties in the hippocampus change the properties of place fields significantly.

The idea that the hippocampus has autoassociative properties can be traced back to early writings by Marr (1971) and McNaughton and Morris (1987), (see also McNaughton, 1989; Rolls, 1989; Leonard and McNaughton, 1990; O'Reilly and McClelland, 1994; Hasselmo *et al.*, 1996; Rolls, 1996). Recurrent connections within the CA3 field are assumed to produce attractor states so that representations are "completed" given sufficient inputs.

Early models by McNaughton *et al.* (McNaughton and Morris, 1987; McNaughton, 1989; Leonard and McNaughton, 1990; McNaughton *et al.*, 1991; McNaughton *et al.*, 1994a) hypothesized that the hippocampus was an associative memory that associated local views with movements to predict future local views. This forms a sort of *transition table.* This theory can be seen as a sort of path integrator in that it updates the representation of place with each movement, but this theory requires exploration before a rodent can show path integration abilities in an environment.

The Recce and Harris (1996) and Samsonovich and McNaughton (1997) theories can be seen as associative memory theories in that they require attractor states which complete incomplete inputs. Recce and Harris (1996) assumes that local view inputs (corrected in the dark by path integration, external to the hippocampus) are input into the hippocampus and stored as stable states in a content-addressable memory. The pre-wired charts hypothesized by Samsonovich and McNaughton (1997) can be viewed as a continuous version of an autoassociative memory.

Another associative memory theory is that of Hetherington and Shapiro (Hetherington and Shapiro, 1993; Shapiro and Hetherington, 1993). In their model, place cells are identified with hidden units in a three-layer neural network with recurrent connections in the middle layer trained by the standard backpropagation algorithm (Rumelhart *et al.*, 1986b; Hertz *et al.*, 1991). Because of the recurrent connections, cells remain active in the dark. However, this theory requires that the animal see the environment in the light before the lights are turned out; if the animal is placed into the environment in the dark, the cells will not be active. Quirk *et al.* (1990) and Markus *et al.* (1994) report that place cells continue to show normal place fields, even when an animal first enters an environment in the dark, which is inconsistent with this prediction.

### 7.2.4 The three-mode place code model

The local view models share the idea that place fields are built from the combination of multiple spatial parameters. The path integration models share the idea that local view information is associated with path integrator information in the place code. The associative memory models suggest that the hippocampus can be understood as a content-addressable memory which stores associations among its input representations. These three aspects are not incompatible. Path integration can drive place cells originally; during exploration, local views can be associated with place cells; and then upon re-entry into an environment, local view can drive place cells when there is insufficient path integrator information to drive them directly. Local view information can also be used to correct for path integrator drift as the animal navigates around the environment. Associative memory properties allow the system to be insensitive to noise and to recall incomplete inputs.

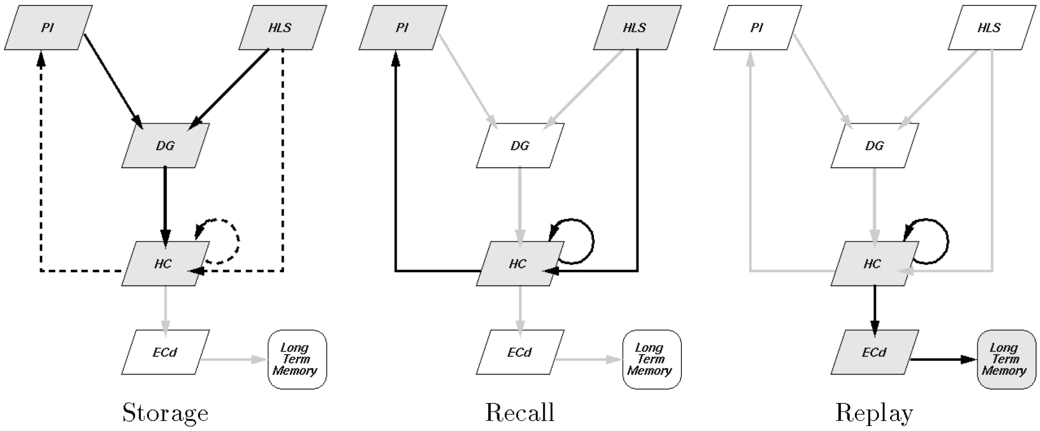



Figure 7.5: Major components of the three-mode place code model. In each mode, dark lines are hypothesized to transmit information, dashed lines to show LTP, and light lines are not involved.

Figure 7.5 shows the model of hippocampal place cells I will use throughout this thesis. It is an amalgam of ideas from previous place cell models as well as some of my own. Appropriate citations and attributions will be presented in the discussion immediately following the list of key points below.

This model does not differentiate between CA3 and CA1. This is not because I don't believe that there is a difference, but because the difference is not necessary for the simulations presented in this thesis. Because of this, I will refer to the model in terms of the hippocampus (HC), meaning the combined CA3 and CA1 fields, but not the dentate gyrus (DG), which is differentiated. I will therefore speak of the model HC as both having recurrent connections (as seen in CA3) and projecting to subiculum (as CA1 does).

The key points are:

- The path integrator (PI, outside the hippocampus) sends strong pre-wired but random connections into the dentate gyrus (DG). I will assume that these connections are unmodifiable.
- Areas representing processed sensory input (HLS[4], also outside the hippocampus) sends strong pre-wired but random connections into the dentate gyrus (DG). I will assume that these connections are unmodifiable.
- During normal navigation, the dentate gyrus performs an *and* function on these two

[4]HLS stands for *high level sensory* areas and is an abstraction of processed sensory inputs, such as might be expected to be included in the local view (see Section 3.2.1).

inputs so that a cell fires only if it gets sufficient activity from both its PI and HLS inputs.

- The dentate gyrus sends strong projections to HC. Each DG cell projects to only a few HC cells and each HC cell receives only a few DG projections. The effect of these strong connections, is that a single DG input can drive a HC cell to fire. I will assume that these connections are unmodifiable.
- HLS sends connections into the hippocampus. These connections are modifiable via a Hebbian mechanism such as LTP.
- Recurrent connections within the hippocampus are also modifiable via a Hebbian mechanism such as LTP.
- The hippocampus sends projections back to the path integrator. These connections are modifiable via a Hebbian mechanism such as LTP.
- The hippocampus has three different activation modes:

  **Storage.** During storage, LTP occurs in the hippocampal formation (between HLS and HC, within the recurrent connections in HC, between HC and PI), but these same connections show little or no synaptic transmission.

  **Recall.** During recall, LTP does not occur anywhere in the hippocampal formation, but the connections that once showed LTP now show synaptic transmission. In addition, I will assume that dentate gyrus is inactive, i.e. that it does not transmit information during recall.

  **Replay.** Replay is like recall, except that HLS and PI representations are assumed to be silent. They do not provide information into the hippocampus at all. The hippocampus acts from internal dynamics only.

The model described above has properties in common with all three of the preceding types. It is similar to the path integrator models discussed above (Wan *et al.*, 1994c; McNaughton *et al.*, 1996; Touretzky and Redish, 1996; Redish and Touretzky, 1997a; Samsonovich and McNaughton, 1997; Samsonovich, 1997) in that it requires both PI and HLS input to drive place cell activity. It differs from the McNaughton *et al.* models (McNaughton *et al.*, 1996; Samsonovich and McNaughton, 1997; Samsonovich, 1997) in that it assumes an extrinsic path integrator which works on a canonical map (referred to as "naive path integrator models" by Samsonovich and McNaughton, 1997).

This model is also similar to the local view models in that it assumes that extrinsic local view input into the hippocampus helps drive place cell activity. It differs from them in that during the normal *storage* mode the local view alone is insufficient to drive place cell activity; path integrator input is also required. (During *recall*, local view is sufficient to drive place cells via the HLS $\rightarrow$HC connections.)

This model is also similar to the associative memory models in that the role of the hippocampus is to associate local views with path integrator representations. As reviewed above, the idea that the hippocampus is completing inputs orthogonalized by the dentate gyrus was first proposed by McNaughton and Morris (1987), using a mechanism originally proposed by Marr (1969) for cerebellum. See also Rolls (1989; 1996) and McClelland *et al.* (O'Reilly and McClelland, 1994; McClelland and Goddard, 1996) for extensive discussions of this idea.

The role of the dentate gyrus in this model is to force the system to use a new place code when the $\mathcal{LV} \times \mathcal{PI}$ association changes. In Chapter 8, this allows the system to handle multiple environments that differ in their local view inputs. In Chapter 10, this forces the system to use different hippocampal representations when the path integrator changes reference points.

This model shares much with the Recce and Harris (1996) model in that in it the hippocampus is an associative memory that stores and recalls representations of location, but differs in that Recce and Harris assume the path integrator changes the local view external to the hippocampus, while in this model the path integrator representation is input directly into the place code.

The observation that the hippocampus shows multiple modes of activity was first made by Vanderwolf (1971). As reviewed in the Experimental Review (EX.62), the hippocampus does show at least two modes of activity (Vanderwolf, 1971; Vanderwolf and Leung, 1983; Buzsáki, 1989; Stewart and Fox, 1990; Vanderwolf, 1990) differentiated by distinctive EEG traces (LIA and theta).

A number of models have included suggestions that the hippocampus shows two modes: *storage* and *replay* (see, for example, Marr, 1970; Marr, 1971; Buzsáki, 1989; Hasselmo and Schnell, 1994; Wilson and McNaughton, 1994; McClelland *et al.*, 1995; McClelland and Goddard, 1996; Skaggs and McNaughton, 1996; Shen and McNaughton, 1996), and that storage occurs during theta, while replay occurs during LIA.

Some models have included *storage* and *recall*-like properties, but have not separated the two modes (Burgess *et al.*, 1993; Burgess *et al.*, 1994; Brown and Sharp, 1995; Blum and Abbott, 1996; Sharp *et al.*, 1996; McNaughton *et al.*, 1996; Gerstner and Abbott, 1997; Samsonovich and McNaughton, 1997; Samsonovich, 1997). I refer to these models as *online* models because they include the idea that recall occurs "online" during navigation (i.e. during theta).

Other researchers have previously hypothesized that the hippocampus does show explicit modes: *storage* and *recall* (Hasselmo and Bower, 1993; Hasselmo and Schnell, 1994; Nadel, 1995; Hasselmo *et al.*, 1996; Recce and Harris, 1996; Touretzky and Redish, 1996; Redish and Touretzky, 1997a; Rotenberg *et al.*, 1996). This issue is also related to the concept of *recognition memory* and the detection of novelty, particular novel environments (Gaffan, 1972; Gaffan, 1974; O'Keefe and Nadel, 1978; Nadel and Willner, 1980; Mishkin and Murray, 1994; Murray and Mishkin, 1996, see Section 4.1.6-B).

This theory differs from previous theories in that it requires the hippocampus to show three modes of activation: *storage*, *recall*, and *replay*. In this model, the storage mode occurs during theta as an animal explores around the environment, the recall mode occurs on a significant context-switch, such as when an animal is returned to an environment, and the replay mode occurs during sleep. Specifics of each of these three modes will be given in Chapter 8.

The main contribution of this work is to bring these ideas together. Although many of the ideas have been addressed separately, no one has yet looked at the effect of all of these ideas together, and all other models have incorporated aspects which I find do not fit the data as well as this model. In the following chapters, I will present simulations that address the complex issues involved in this three-mode model.

# Chapter 8

# Self-localization

As described in the Navigation Overview (Chapter 3), one role hypothesized for the place code is to associate the path integrator and local view so that when an animal enters a familiar environment, it can reset its path integrator to use the same coordinate system as during previous experiences in the environment. We call this process *path integrator reset* (Wan *et al.*, 1994a; Touretzky and Redish, 1996; Redish and Touretzky, 1997b) or *self-localization* (Redish and Touretzky, 1997a).

Many place cell models have included the idea that the local view influences the place code (see review, previous chapter, Section 7.2.1), but most of those models include the local view as the sole input. As reviewed previously, this must be incomplete.

Another model that addresses issues of associations between the path integrator and the local view is that of McNaughton *et al.* (McNaughton *et al.*, 1996; Samsonovich and McNaughton, 1997; Samsonovich, 1997). In the McNaughton *et al.* model, the hippocampus is the path integrator and local view input influences it. It differs from the model in this thesis in two major aspects: (1) the anatomical instantiation of the path integrator is extrahippocampal and (2) the local view influence on the path integrator is ongoing.

However, some researchers are now considering that the recall process may only occur on re-entry, because of recent data suggesting that place fields are only unstable across removal and re-entry (Bostock *et al.*, 1991; Rotenberg *et al.*, 1996; Barnes *et al.*, 1997; see also Scoville, 1968; Milner *et al.*, 1968; Sacks, 1985; Cohen and Eichenbaum, 1993; Murray and Mishkin, 1996 for similar effects in the primate lesion literature).

The concept of self-localization can be seen as an instance of *recalling a memory.* Other hippocampal models have included the concept of a general *recall* process, but none of these are models of navigation. Instead, they purport to be models of episodic memory and only deal with retrieval of binary vectors (e.g. Marr, 1971; McNaughton and Morris, 1987; Rolls, 1989; Hasselmo and Bower, 1993; Hasselmo and Schnell, 1994; Rolls, 1996). This issue is related to issues of *contextual retrieval* (Gaffan, 1972; Gaffan, 1974; Hirsh, 1974; O'Keefe and Nadel, 1978; Nadel and Willner, 1980; Nadel, 1994; Nadel, 1995, see Section 4.1.6).

## 8.1 The *self-localization* process

The mechanism we hypothesize as the implementation of this *recall* process is that local view provides candidate locations to the hippocampus, and a parallel relaxation process in the hippocampus resolves the incompatible constraints to provide a *coherent* representation of the animal's location within a reference frame.

In our early models (Wan *et al.*, 1994a; Wan *et al.*, 1994b; Wan *et al.*, 1994c; Touretzky and Redish, 1996; Redish and Touretzky, 1997b), we proposed an abstract implementation in which place cells provided candidate locations into the path integrator, which (because the path integrator could only represent a single location) limited the number of candidates. By slowly tuning the place cell's sensitivity to the path integrator, one candidate location eventually won out.

A more neural implementation of the self-localization process can occur as a consequence of local excitation and global inhibition in the hippocampus (McNaughton *et al.*, 1996; Zhang, 1996a; Redish and Touretzky, 1997a; Samsonovich, 1997; Samsonovich and McNaughton, 1997, see also Kohonen, 1982; Kohonen, 1984; Droulez and Berthoz, 1991; Munoz *et al.*, 1991; Arai *et al.*, 1994). Place cells that encode nearby locations on a single reference frame support each other. When combined with global inhibition, this produces a competitive dynamics similar to winner-take-all dynamics.

The final state of this system is an "activity bubble" (Kohonen's term) or "hill" of activation on the neural sheet represented by the cells. Cells with preferred locations near the represented location are very strongly active and cells with preferred locations farther from the represented location are progressively less and less active. An example of this is shown in the final panel of Figure 8.2.

Since the place cells are not laid out topologically[1] (McNaughton, 1989), the neural sheet has to be understood as occurring in the space of the represented position of the cells and not anatomically.

Because the final stable state of this system is not a single active cell, but a population of cells which encode a single location, I am going to refer to this as *pseudo winner-take-all* (pWTA) dynamics. It can be understood as winner-take-all dynamics in the context of location represented by the place code, but the final stable state is not one cell completely active (as would be true WTA dynamics), nor is it $k$ cells completely active (as would be true kWTA dynamics).

The fact that local excitation and global inhibition produces a single "hill" of activation on a one-dimensional sheet was first shown mathematically by Wilson and Cowan (1973; see also Amari, 1977; Ermentrout and Cowan, 1979; Kishimoto and Amari, 1979). The first simulations of this pWTA process were reported by Kohonen (1982; 1984) who showed simulations in both one and two dimensions.

[1]This is to be expected because place fields change from environment to environment (O'Keefe and Conway, 1978; Kubie and Ranck, 1983; Muller and Kubie, 1987; Thompson and Best, 1989, EX.79). This means that the mapping from anatomy to space changes from one environment to the next and thus place fields cannot be laid out topologically.

A recent model which also depends on pWTA dynamics is that of McNaughton *et al.* (McNaughton *et al.*, 1996; Samsonovich and McNaughton, 1997; Samsonovich, 1997), but this model incorporates additional unnecessary hypotheses about the hippocampus (i.e. that the hippocampus is the path integrator).

Another model which depends on pWTA dynamics is that of Shen and McNaughton (1996), which showed that LTP in intra-hippocampal connections would replay recent memories. (See Navigation Review, Section 3.4 and Chapter 9 for an extensive discussion of this and other similar models.) However, their model involves what I have called the *replay* mode and does not include an explicit *recall* mode.

## 8.2 Data support

The appropriate connection function needed for the self-localization process is simply that "nearby places support each other" combined with global non-specific inhibition.

The local excitation component can be learned by correlational Hebbian LTP and can be seen on a neural level as increased synaptic efficacy between place cells with overlapping place fields. It can be realized in the intra-CA3 connections if the synaptic efficacies between place cells are inversely related to the distance between the centers of their place fields (Muller *et al.*, 1991b; Wilson and McNaughton, 1994). I will follow Muller *et al.* (1991b; 1996) in calling this the *cognitive graph.* It can be learned by correlational LTP combined with random exploration of an environment. As the animal wanders around the environment, cells with overlapping place fields are more likely to be coactive than cells with well-separated fields. Combined with correlational LTP, in which the synaptic efficacy is increased when both cells are simultaneously active, the CA3 recurrent connections will be inversely related to the distance between the place field centers after a session of wandering an environment.[2] Figure 8.1 shows the cognitive graph learned by our simulations after a session of exploring a 100 cm diameter environment.

Hebb (1949) first suggested a correlational learning rule between neurons. Data suggesting that LTP is correlational (i.e. that the synaptic weight between two cells is increased only with presynaptic firing and postsynaptic depolarization) has been well established (see Landfield and Deadwyler, 1988; Brown *et al.*, 1991; McNaughton, 1993; Malenka, 1995 for reviews). LTP specifically has been shown in the recurrent connections in CA3, and in the Schaffer collaterals connecting CA3 to CA1 (see Landfield and Deadwyler, 1988; Brown *et al.*, 1991; McNaughton, 1993; Malenka, 1995 for reviews).

I have already reviewed data suggesting that place cells representing nearby locations are more strongly connected than cells that represent distal locations. Wilson and McNaughton (1994) recorded from more than a hundred hippocampal place cells simultaneously as the animal explored an environment and showed that the correlation between the

[2] Actually, this mechanism produces synaptic efficacy inversely related to the travel time between two place fields, but the mechanism will average over different speeds and should be an acceptable approximation for nearby place fields. Place fields that are far apart will have zero effective weight anyway.

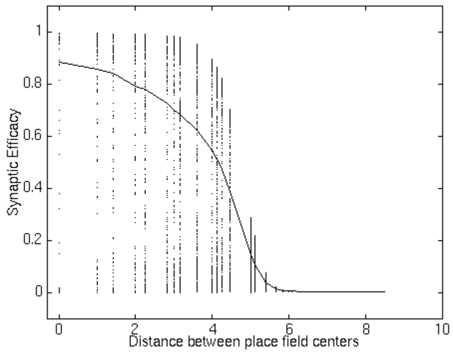



Figure 8.1: The cognitive graph: scatter plot of learned synaptic weights as a function of overlap between place fields of pairs of units. Line indicates mean. From Redish and Touretzky (1997c), see Redish and Touretzky (1997c) for simulation details.

specific timing of place cell spikes was stronger between cells with overlapping fields than between cells with widely separated fields, suggesting that the synaptic efficacies between them are stronger.

There are inhibitory neurons in the hippocampus with very broad arborizations (see Freund and Buzsáki, 1996, for a review). Inhibitory interneurons in the hippocampus also tend to fire over large portions of the environment (McNaughton *et al.*, 1983a; Christian and Deadwyler, 1986; Kubie *et al.*, 1990; Leonard and McNaughton, 1990; Mizumori *et al.*, 1990; Muller *et al.*, 1991a; Wilson and McNaughton, 1993), implying that they are very nonspecific.

These data provide strong support for the local-excitation/global-inhibition weight matrix required for pWTA dynamics.

An important aspect of this result is that it will also produce local excitation within a single map. That is, two place cells will be strongly connected if they have place fields near each other in one experienced map whether or not the have place fields near each other on another. This means that the mechanism will produce pWTA dynamics within each map. For a discussion of this type of network see also Samsonovich and McNaughton (Samsonovich, 1997; Samsonovich and McNaughton, 1997).

## 8.3 Hippocampal modes

As reviewed in the previous chapter, the hippocampus shows multiple modes of activity. In this model, the storage mode corresponds to hippocampal theta (in agreement with a number of models, see previous chapter), and recall occurs during sharp waves occurring during LIA in awake states.

There are two hypotheses being put forward in this chapter: (1) that there is an explicit recall process occurring on re-entry into an environment and (2) that the recall process occurs during sharp waves in LIA states. There are predictions one can make from each of these hypotheses but a disproof of the second hypothesis does not constitute a disproof of the first.

During LIA, hippocampal pyramidal cells tend to fire during sharp waves, but are then mostly silent (Buzsáki, 1989; Ylinen *et al.*, 1995). During theta, each cell fires only when the animal is in the corresponding place field (see Experimental Review, Section 2.2.2-A). Since each cell has a different place field, cells fire a few at a time.

During exploration (according to this theory), the rodent navigation system learns three things:

1. a mapping from local views to place codes (realized by LTP in the HLS $\rightarrow$ HC connections in the model presented in the previous chapter)

2. a connection function that enforces the place code to always consist of a coherent representation of location (realized by LTP within the recurrent connections of CA3 and between CA3 and CA1), and

3. a mapping from place codes to path integrator coordinates (realized by LTP between CA1 and subiculum).

In this theory, on entering an environment, the following sequence occurs:

1. Subiculum, parasubiculum, hippocampus, and entorhinal cortex are initially noisy.

2. Sensory cues in high level sensory areas such as parietal cortex (i.e. local views) are passed through superficial EC into the hippocampus proper, biasing the random firing rates with candidate locations.

3. The recurrent connections in CA3 allow one of these candidate locations to win out, forming a coherent code in hippocampus.

4. The connections between CA1 and subiculum reset the path integrator to the correct representation of the animal's location in path integrator coordinates.

We refer to this as the *self-localization* or *recall* process.[3]

We suggest that this happens within the course of a single sharp wave during LIA for three reasons. (1) In our simulations, the place code in CA3 is coherent within 50–100 ms. Figure 8.2 shows the first 70 ms of a simulated sharp wave. (2) During a sharp wave, place cells do not show normal place fields, many cells are simultaneously active (many more

[3]The *recall* process is similar to the *replay* process, which has been shown in simulation by a number of authors, such as Blum and Abbott (1996), Levy (1996), Tsodyks *et al.* (1996), and Shen and McNaughton (1996), see Navigation Overview (Chapter 3, Sections 3.2.4 and 3.4). However, the *recall* and *replay* processes differ in that the replay process does not include candidate locations in the input.

than during theta) (Buzsáki, 1989). (3) Data from Hasselmo and Schnell (1994, EX.114) suggests that storage occurs during theta and recall during LIA (Buzsáki, 1989; Hasselmo and Schnell, 1994; Wilson and McNaughton, 1994; Smith, 1995; Hasselmo *et al.*, 1996; Kudrimoti *et al.*, 1996; Skaggs and McNaughton, 1996; Shen and McNaughton, 1996, see also comments by McNaughton *et al.* in Seifert, 1983, p. 610).

This hypothesis also requires that ECs and ECd be active at different times: As the intra-hippocampal connections are learning (during theta), activity should not be transmitted through ECd. During the self-localization procedure, this theory requires the hippocampus to show LIA, but ECs cells should fire at a constant rate and would still be uncorrelated to LIA. Consistent with this hypothesis is data showing that while ECs cells are phase locked to the theta rhythm, they are uncorrelated to LIA, and conversely, while ECd cells are uncorrelated to theta, they are correlated to the LIA EEG signal (Chrobak and Buzsáki, 1994).

## 8.4 Predictions

The key prediction we can take from this theory about the role of the place code is that when the animal returns to a familiar environment, it should show the self-localization process.

### 8.4.1 LIA on entry

We have suggested that this process is realized by sharp waves during LIA. If this is true then we can predict that animals will show LIA on entry into a familiar environment. Note that we are not making any predictions about what happens when the animal enters a novel environment, only that on entering a familiar environment, it will show at least one sharp wave spike.

### 8.4.2 Coherency changes on entry

We can also predict that during the last sharp wave before the animal begins moving, the representation in CA3 will begin in an incoherent state and become coherent over the course of the sharp wave. At the end of the sharp wave, the place cells will encode the animal's current location accurately and these cells will be the initial ones active during the first theta cycles.

We can define the coherency of a representation mathematically as the inverse of the width of the confidence interval of the position represented by the population. It can be measured by the bootstrap algorithm (Efron, 1982) given enough cells. The simultaneous recording of a hundred cells or so described by Wilson and McNaughton (1993) should be sufficient.

Another way to view the coherency of a population is to plot the sum of the place fields of each cell, weighted by their current firing rates. A coherent representation of position

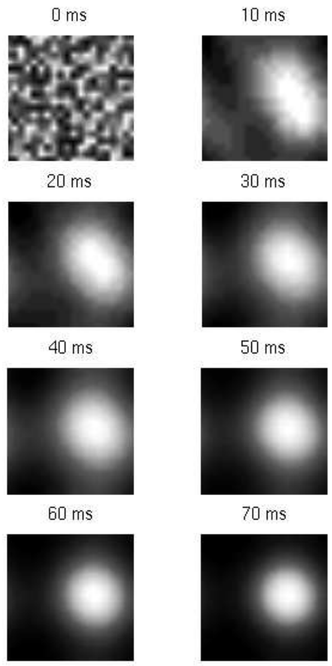



Figure 8.2: A simulated sharp wave. Each panel shows the firing rates of all of the simulated CA3 place cells at a single moment in time. The cells are laid out in a two-dimensional sheet with their locations in the sheet corresponding to the centers of their place fields in the environment. Intensity values have been interpolated for clarity. White indicates high firing rate, black low. At time 0, the simulation is primed with noise and small biases towards candidate locations at three points in the environment. The recurrent connections force it to settle to a single one of those candidate locations. From Redish and Touretzky (1997c), see Redish and Touretzky (1997c) for simulation details.

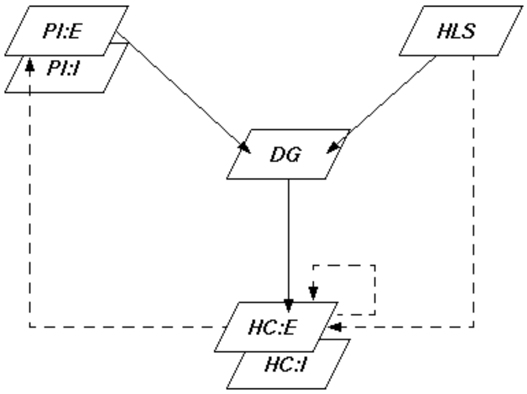



Figure 8.3: Areas used in the self-localization simulation. During storage, solid lines drive place cell activity and dashed lines show correlational learning; during recall, dashed lines show synaptic transmission and drive place cell activity and path integrator reset. See text for details.

will show up clearly as a single "hill" of activity.

### 8.4.3 The resolution issue

Another prediction we can make from this theory is that when the candidate locations are near each other (within the radius of a single place field), they will be averaged together, but when the candidate locations are far apart only one of them will be chosen and no interaction effect will be seen. This is because the hill of activity will precess towards nearby representations, but will suppress distal representations that are too weak to overrule it. See Appendix A for a discussion of this issue.

## 8.5 Simulations

In order to demonstrate the viability of this hypothesized hippocampal role, I will show simulations of three important situations. All three simulations used the same instantiation of the model described in the previous chapter. They were all trained on the same two environments and differ only in the environment in which the self-localization process occurs.

### 8.5.1 Simulation details

Simulations consisted of the model shown in Figure 8.3. All neurons used the three-equation neuron model based on the Pinto *et al.* (1996) interpretation of the Wilson and Cowan (1972) equations (Eqs. 1.1–1.3).

The path integrator (PI) simulation consisted of 400 excitatory neurons arranged in a grid ($20 \times 20$) and a single inhibitory interneuron. During exploration, the excitatory neurons of the path integrator (PI:E) were assumed to always represent the position of the animal accurately. This was done by representing the path integrator coordinates in the simulation separately and setting the firing rates of the 400 neurons to the appropriate values (see Chapter 6). This allowed us to examine questions of self-localization without addressing path integrator accuracy.

During the self-localization process, the path integrator simulation was initialized to a random value and allow to settle as described in Chapter 6, see Theory 6.1.2-C.

DG and HC neurons behaved normally based on the three-equation model described above. Each granule cell received input from one randomly chosen PI:E neuron and one randomly chosen HLS neuron and showed a high firing rate if both were active, but a low firing rate if either was inactive. Each hippocampal cell received inputs from 15 randomly chosen granule cells.

The parameters used are listed in Tables 8.1 and 8.2. Simulations used a time constant of 100 ms during exploration and 1 ms during self-localization.

| | | |
|---|---|---|
| Centers, $r$-type sensory neurons | $\tilde{r}$ | {0 cm, 10 cm, ..., 100 cm} |
| Gain, $r$-type sensory neurons | $\sigma_r$ | 25 cm |
| Centers, $\phi$-type sensory neurons | $\tilde{\phi}$ | {0°, 5°, ..., 355°} |
| Gain, $\phi$-type sensory neurons | $\sigma_\phi$ | 25° |
| Path integrator size (excitatory neurons) | $N_{PI:E}$ | 400 ($20 \times 20$) |
| Time constant, excitatory PI (recall) | $\tau_{PI:E}$ | 10 ms |
| Tonic inhibition, excitatory PI (recall) | $\gamma_{PI:E}$ | 0 |
| Path integrator size (inhibitory interneurons) | $N_{PI:I}$ | 1 |
| Time constant, inhibitory PI (recall) | $\tau_{PI:I}$ | 2 ms |
| Tonic inhibition, inhibitory PI (recall) | $\gamma_{PI:I}$ | -7.5 |
| Dentate gyrus | $N_{DG}$ | 10000 |
| Time constant, Dentate gyrus | $\tau_{DG}$ | 1 ms |
| Tonic inhibition, Dentate gyrus | $\gamma_{DG}$ | -6.0 |
| Hippocampus, excitatory neurons | $N_{HC:E}$ | 500 |
| Time constant, Hippocampus (recall) | $\tau_{HC:E}$ | 10 ms |
| Tonic inhibition, Hippocampus (exploration) | $\gamma_{HC:E}$ | -2.0 |
| Tonic inhibition, Hippocampus (recall) | $\gamma_{HC:E}$ | -20.0 |
| Hippocampus inhibitory neurons | $N_{HC:I}$ | 1 |
| Time constant, inhib Hippocampus (recall) | $\tau_{HC:I}$ | 2 ms |
| Tonic inhibition, inhib Hippocampus (recall) | $\gamma_{HC:I}$ | 0 |

Table 8.1: Parameters used in the self-localization simulations.

| | | |
|---|---|---|
| Weight, PI:E → DG | $W_{PI:E\rightarrow DG}$ | 4.5 |
| Weight, HLS → DG | $W_{HLS\rightarrow DG}$ | 4.5 |
| Weight, DG → HC | $W_{DG\rightarrow HC}$ | 3.0 |
| Gain, HLS → HC | $\tilde{w}_{HLS\rightarrow HC}$ | 1.5 |
| Learning rate, HLS → HC | $\eta_{HLS\rightarrow HC}$ | 1.0 |
| LTP sharpening constant, HLS → HC | $\mu_{HLS\rightarrow HC}$ | 5 |
| Gain, HC → HC | $\tilde{w}_{HC\rightarrow HC}$ | 1.0 |
| Learning rate, HC → HC | $\eta_{HC\rightarrow HC}$ | 1.0 |
| LTP sharpening constant, HC → HC | $\mu_{HC\rightarrow HC}$ | 5 |
| Gain, HC → PI:E | $\tilde{w}_{HC\rightarrow PI:E}$ | 2.0 |
| Learning rate, HC → PI:E | $\eta_{HC\rightarrow PI:E}$ | 1.0 |
| LTP sharpening constant, HC → PI:E | $\mu_{HC\rightarrow PI:E}$ | 5 |
| Weight, PI:E → PI:E | $W_{PI:E\rightarrow PI:E}$ | $6.0 \cdot \exp(-(\Delta_{ij})^2/(30\,\mathrm{cm})^2)$ |
| Weight, PI:I → PI:E | $W_{PI:I\rightarrow PI:E}$ | -10.0 |
| Weight, PI:E → PI:I | $W_{PI:E\rightarrow PI:I}$ | 0.075 |
| Weight, HC:I → HC:E | $W_{HC:I\rightarrow HC:E}$ | -20.0 |
| Weight, HC:E → HC:I | $W_{HC:E\rightarrow HC:I}$ | 0.002 |

Table 8.2: More parameters used in the self-localization simulations.

The two environments consisted of a pair of landmarks of distinct types at the center of a 50 cm radius cylinder. The first environment (hereafter, Environment AC), had two landmarks identified as "A" and "C"; the second environment (hereafter, Environment BC), had two landmarks identified as "B" and "C". Thus the environments shared some characteristics, but were differentiable. The simulation was allowed to explore randomly for 5 simulated minutes in each environment sequentially.

**LTP.** The learnable connections (HLS → HC, HC → HC, and HC → PI:E) were fully connected. We modeled LTP by the following equation

$$\eta \frac{dw_{ij}}{dt} = [-w_{ij} + \tilde{w} \cdot (S_j \cdot F_i)^{\mu}]^{+} \tag{8.1}$$

where $\eta$, $\tilde{w}$, and $\mu$ are constants. $\mu$ controls the sharpness of the LTP effect, i.e. the higher $\mu$ is, the larger the slope of LTP increase with increasing firing rates of the pre- and post-synaptic neurons. $[]^{+}$ signifies rectification at 0, so weights in the model can only increase. We do not model LTD.

**Local view representations.** Throughout this thesis, the local view representation, which will be referred to as *HLS* (*High Level Sensory* areas), serves as an abstraction of the complexities involved in perception.

As discussed in Section 3.2.1, there are two major categories of cues: local cues and distal cues. These will be represented by two different populations.

Distance measurements of distal cues are unavailable, and the allocentric bearing to distal cues does not change over the environment. Therefore, egocentric bearings to distal cues can be directly associated with head direction representations (McNaughton *et al.*, 1991; Skaggs *et al.*, 1995, see Section 3.2.2). To reflect this use, the representation of distal cues (when they are available) will consist of a population of neurons, each of which has an associated preferred direction $\tilde{\theta}_i$ and each of which show firing rates according to the following formula:

$$F_i(t) = \max_{(k \in \mathcal{D})} \mathrm{eq}(T_k, \tilde{T}_i) \cdot \exp(-(\theta_k - \tilde{\theta}_i)^2/\sigma^2) \tag{8.2}$$

where $k$ ranges over the set of distal landmarks $\mathcal{D}$. $\mathrm{eq}(T_k, \tilde{T}_i)$ is 1 if the type of landmark $k$, $T_k$, is the same as the expected type for neuron $i$, $\tilde{T}_i$, and 0 otherwise. $\theta_k$ is the egocentric bearing of landmark $k$, and $\tilde{\theta}_i$ is the preferred direction of neuron $i$. Angular differences were constrained to be in the range of $(-180°, 180°]$.

Both distance and egocentric bearings of local cues can be computed, however, the egocentric bearings of local cues are not useful to the system. I therefore included representations of distance and allocentric bearing in the representation of local cues. Allocentric bearing was computed from measured egocentric bearing and the current representation of head direction in the HD population.

The specific representation of local cues consisted of two populations of neurons, tuned to distance and allocentric bearing respectively. Each neuron showed firing rates determined by the following equation:

$$F_i = \max_{(k \in \mathcal{L})} \mathrm{eq}(T_k, \tilde{T}_i) \cdot \begin{cases} \exp(-(r_k - \tilde{r}_i)^2/\sigma_r^2) & r\text{-type neuron} \\ \exp(-(\phi_k - \tilde{\phi}_i)^2/\sigma_\phi^2) & \phi\text{-type neuron} \end{cases} \tag{8.3}$$

where $k$ ranges over the set of local landmarks $\mathcal{L}$. $r_k$ and $\phi_k$ are the distance and allocentric bearing to landmark $k$ respectively. $\tilde{r}_i$ is the preferred distance of $r$-type neuron $i$ and $\tilde{\phi}_i$ is the preferred bearing for $\phi$-type neuron $i$. $\sigma_r$ and $\sigma_\phi$ are gain parameters. Angular differences were constrained to be in the range of $(-180°, 180°]$. Both point landmarks and surface landmarks could serve as local landmarks. For surface landmarks, the distance was the shortest distance to the wall and the bearing used was that of the normal to the wall.

In some simulations, only local cues were assumed available, and in others, only distal cues were assumed available. Finally, in some simulations both were available. (In the ones in this chapter, all available landmarks are local.) Simulations only included a representation of the cues if the cues were available. I will identify in each simulation whether local or distal cues (or both) were available. I did not model noise in the sensory input.

### 8.5.2 Simulation results

#### 8.5.2-A Nonambiguous local view representations

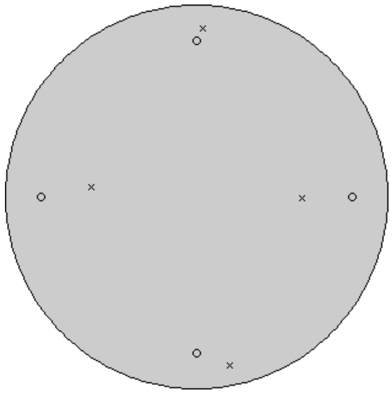


Figure 8.4: Self-localization at four non-ambiguous points in an environment. Each × marks the represented position in the path integrator simulation after self-localization attempts at each of the four locations in the environment (N, S, E, and W). The actual positions at which the sim-animal entered the environment are marked by small circles (∘). Landmarks "A" and "C" were located in the center of the environment.

The first case described is that of returning to a single environment with a nonambiguous local view. This actually only requires correlational LTP between HLS → HC and between HC → PI; the HC → HC connections do not change the results. The local view representation (in HLS) sets up a representation in the hippocampus which primes a representation in PI. Due to the psuedo-winner-take-all (pWTA) nature of the path integrator representation (see Chapter 6), the path integrator representation settles to a representation consistent with the local view. This means that when an animal returns to a location it has previously visited, the self-localization process will reset the PI representation to be identical to the previous experience. This means the PI × LV association will be the same and the same place code will be used and the place cells will seem to be stable from session to session.

Figure 8.4 shows the ability of the animal to self-localize at the four cardinal points in the AC environment.

#### 8.5.2-B Ambiguous local view representations

If the local view does not describe a unique location, then some process will need to choose one location over the other. When dealing with a single environment, the pWTA process in the path integrator will be sufficient, but the recurrent connections in hippocampus will also help. Figure 8.5 shows that when given an ambiguous local view, the simulations settle to one or the other candidate when the cues are sufficiently separated. The ambiguous local

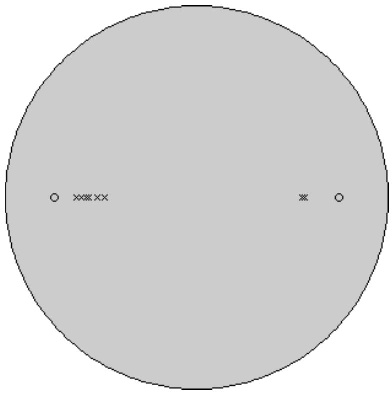


Figure 8.5: Self-localization with an ambiguous local view. Each $\times$ marks represented position in the path integrator simulation. The two $\circ$ marks indicate the two representations compatible with the local view. 10 self-localization attempts were made with an ambiguous local view. The represented locations cluster into two discrete cases both compatible with the local view.

view was generated by adding a second pair of landmarks offset from the original. This produced two possible candidate locations which provided an ambiguity for the simulation to resolve. This "resolution issue" is discussed in greater depth in Appendix A.

### 8.5.2-C Multiple environments

What is the role of the recurrent connections in hippocampus? When the ambiguity is not within the same environment, but between environments, then the non-linearity in the path integrator described in Chapter 6 is insufficient. We have hypothesized that the path integrator works on a single canonical map. This means that two different local view representations may suggest similar path integrator coordinates but on two different reference frames. They can only be separated because they are represented by different hippocampal populations.

These two populations are separated in simulation because of the orthogonalization properties in the model dentate gyrus. Although the two environments share half of their cues (the "A" landmark), half of the cues are also different ("B" vs. "C" landmarks). On examination, I found that the simulations also used two different path integrator reference points for the two environments. Because of this, the system must learn two different associations between path integrator and local view. Because of the *and* properties in the model dentate gyrus, two different hippocampal maps are thus used.

The average correlation between the two HC representations of the two environments is 0.1672. This means the simulation represents the two environments differently and allows the simulation to settle to the correct location in both environments. Although there was initially some overlap in the CA3 during the self-localization process, the overlap was de-

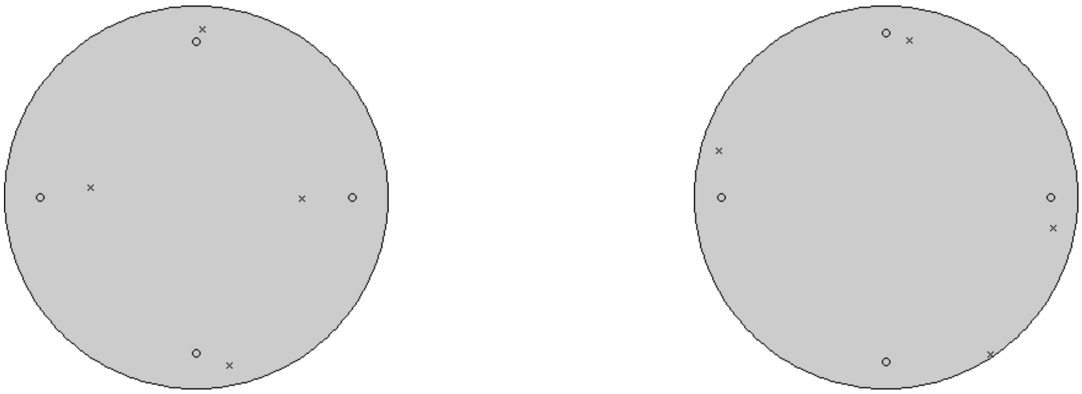


Figure 8.6: Self-localization at four non-ambiguous points in two environments. ×show represented positions in the path integrator simulation after self-localization attempts at each of the four locations in the environment (N, S, E, and W). The actual positions at which the sim-animal entered the environments are marked by small circles (∘). Landmarks "A" and "C" were located in the center of one environment and "A" and "B" at the center of the other. Note that both environments are reasonably accurate even with the limited simulations presented here.

creased as the recurrent connections drove the system to one of the two (separated) stable states. As can be seen in the figure, the simulations can self-localize reasonably well in both environments.

## 8.6 Discussion: Implications for exploration

As described in the Experimental Review (EX.60 and EX.130), rodents exploring a novel environment regularly return to *home bases* (Chance and Mead, 1955; Eilam and Golani, 1989; Leonard and McNaughton, 1990; Golani *et al.*, 1993; Touretzky *et al.*, 1996). These home bases can often be identified with the animal's initial entry points into the environment (Leonard and McNaughton, 1990; Touretzky *et al.*, 1996, see Section 2.6).

We have suggested that one possible explanation for this is that animals are making an association between local view and path integrator representations in the place code (Touretzky and Redish, 1996). If the path integrator drifts as an animal explores the environment then the wrong association will be made.

Samsonovich (1997) has suggested that because of path integrator drift, the association between local view representations and path integrator coordinates must be made directly and cannot be made using an intermediate representation unless the path integrator is very accurate. If the path integrator is not accurate, then after making a loop, the intermediate representation might not be the same and the wrong association might be made. As a consequence of this, they interpret that the hippocampus must update its information by path integration. However, the home base behavior of exploring animals can counteract drift even if there is an intermediate representation associating the local view and an extra-hippocampal path integrator.

At the home base, the animal has presumably learned an appropriate association between the local view and the path integrator. Then, as it makes a foray into the environment, the path integrator drifts. Near the home base, the drift won't be serious, so the association will be accurate, but as the animal moves farther into the environment, the drift will worsen and it will have to return to the base. By constantly returning to an area in which the association has been made, the animal is able to correct for any inaccuracies in its path integration ability. In a sense, we can say that he animal is *inductively* exploring the environment, constantly adding novel portions to its "known" area.

# Chapter 9

# The dual-role hippocampus

The place code in the hippocampus allows an animal to self-localize when it returns to a familiar environment. Once there is a coherent representation in the place code, the goal memory can plan a trajectory to a goal. When the animal traverses that trajectory, the synaptic efficacies of the recurrent connections within CA3 become asymmetric (see Navigation Overview, Section 3.4 for an extensive discussion and citations). This has the effect of storing recently traveled routes in the hippocampus.

In this chapter, I will show simulations demonstrating three important points:[1]

1. The routes replayed are paths to the goal.
2. The self-localization and route-storage roles are compatible, even though they seem to require different weight matrices. (Self-localization requires symmetric connections, route-storage requires asymmetric connections.)
3. The information contained in the replayed routes is sufficient to train a slowly-learning network to solve the water maze without the self-localization process (and thus perhaps without the hippocampus).

## 9.1 Relation to previous work

As discussed in the Navigation Overview (Section 3.4), the concept of consolidation of memories is an old one (Marr, 1970; Marr, 1971; Buzsáki, 1989; Pavlides and Winson, 1989; Zola-Morgan and Squire, 1990; Squire and Zola-Morgan, 1991; Squire, 1992; Cohen and Eichenbaum, 1993; Wilson and McNaughton, 1994; McClelland *et al.*, 1995; Squire and Alvarez, 1995; Skaggs and McNaughton, 1996), as is the idea that memories are stored

[1] This work was done with D. S. Touretzky. It has been published in abstract (Redish and Touretzky, 1996a), conference (Redish and Touretzky, 1997d), and journal (Redish and Touretzky, 1997c) form.

during theta mode and recalled during LIA (Buzsáki, 1989; Hasselmo and Bower, 1993; Hasselmo and Schnell, 1994; Chrobak and Buzsáki, 1994; Skaggs and McNaughton, 1996, see also comments by McNaughton *et al.* in Seifert, 1983, p. 610). On the other hand, all published models of declarative memory only include storage and retrieval of arbitrary binary vectors (Marr, 1970; Marr, 1971; McNaughton and Morris, 1987; McNaughton, 1989; Rolls, 1989; Alvarez and Squire, 1994; Hasselmo and Schnell, 1994; Levy, 1996; McClelland and Goddard, 1996; Rolls, 1996).[2] Although these models can address general principles involved in memory, they cannot address the role of the hippocampus in specific tasks. This makes it difficult to compare their results with real experiments or to generate testable predictions. By looking at the role of routes in an actual task, we have been able to show that (in simulation) the routes replayed are in fact the routes to the goal and that the route-replay process does not interfere with the self-localization process discussed in Chapter 8.

Blum and Abbott (1996) have shown that LTP combined with the normal navigation can store routes to a goal. However, they do not include a neural model in their simulations; they begin from the assumption of asymmetric LTP and analytically calculated the hypothesized shift in the position represented by the place cell population. We found that when we tried to simulate this result at a more neural level our simulations could not learn sequences (unpublished results). This occurred because the time course of asymmetric LTP is much to fast for normal navigation. As pointed out by Skaggs (1995), this can be solved by the effect of phase precession. We include phase precession in our simulations.

In addition, Blum and Abbott's theory requires the hippocampus for online navigation. As discussed in Chapter 7, this is not the same as the *replay* process being simulated here. Recent data from Alyan *et al.* (1997) suggests that the hippocampus is not required for online navigation.

Levy and his colleagues (Levy, 1996; Levy and Wu, 1996; Wu *et al.*, 1996) have explored sequence learning in recurrent hippocampus-like networks and have demonstrated that cells with temporally-extended effects can store and replay sequences of arbitrary binary vectors.

Shen and McNaughton (1996) have a model of attractor dynamics similar to that presented here in which cells with Gaussian place fields exhibit correlational (Hebbian) learning. When these two effects are combined with random exploration, a local-excitation weight matrix is formed (Muller *et al.*, 1991b; Muller *et al.*, 1996, see Section 8.2). Shen and McNaughton demonstrate that when presented with random input and allowed to settle to a stable state, cells with recently visited place fields are more active than other cells, corresponding to data from Pavlides and Winson (1989) and Wilson and McNaughton (1994). They did not, however, look at sequences, as we will here.

[2] The one exception to this is Shen and McNaughton (1996) which was done simultaneously with the work reported in this chapter. Shen and McNaughton (1996) have not looked at sequences, which are discussed in this chapter.

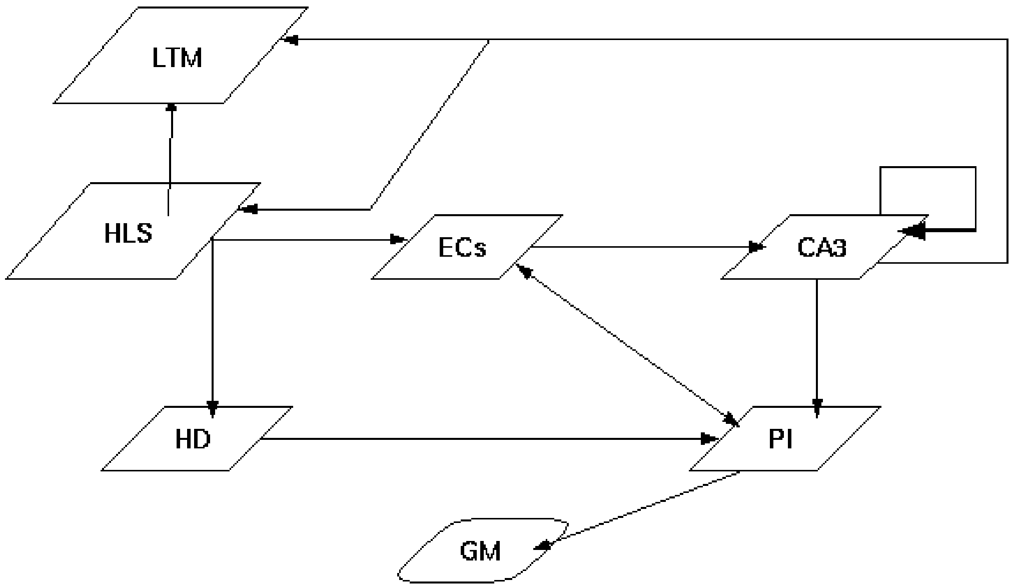



Figure 9.1: Components used in the simulation of route learning. HLS: high level sensory areas; LTM: long term route memory; ECs: superficial entorhinal cortex; PI: path integrator; HD: head direction system; GM: goal memory.

## 9.2 Simulation details

The simulations used a slightly earlier version of the model than that presented in other chapters of the thesis, but the key ideas are the same and the conclusions should not change.

As before, the neurons were simulated using the Pinto *et al.* (1996) interpretation of the Wilson and Cowan (1972) equations (Eqs. 1.1–1.3).

The network used in this chapter consisted of five 2D neural sheets (HLS, ECs, PI, HC:E, and HC:I); each consisted of 400 neurons arranged in a $20 \times 20$ grid. The local view simulation (HLS) was assumed to represent the position of the animal accurately at all times. The path integrator (PI) was assumed to track the animal accurately once it had been reset from the hippocampal representation.

We simulated the hippocampal population as two pools, one excitatory and one inhibitory (HC:E and HC:I, respectively). The excitatory neurons were interconnected within and between pools by a Gaussian with a standard deviation of 20 cm. Inhibitory neurons were broadly connected to both the excitatory and inhibitory pools. Essentially, this connection structure corresponds to local excitation and global inhibition.

The hippocampal excitatory neurons also showed phase precession (EX.74). We did not model phase precession as an emergent result of a complex process; instead we noted that phase precession exists in hippocampus and showed that, when combined with the asymmetric temporal nature of LTP, routes can be stored in the recurrent connections of the hippocampus (as suggested by Skaggs, 1995). In order to produce phase precession, I derived the *preferred phase* of each hippocampal neuron using the approximation in Figure 9.2. I then defined the firing rate of each neuron at time $t$ as

$$F_i(t) = \exp(-(\tilde{\Theta}_i(t) - \Theta(t))^2/\sigma^2) \cdot \hat{F}_i(t) \tag{9.1}$$

where $\tilde{\Theta}_i(t)$ is the preferred phase of neuron $i$, $\Theta(t)$ is the current phase of the theta rhythm, $\sigma$ is a constant, and $\hat{F}_i(t)$ is the peak firing rate at $\Theta(t) = \tilde{\Theta}(t)$. I assume a theta rhythm with a frequency of 7 Hz, so $\Theta(t) = (\frac{7 \cdot 360°}{sec} \cdot t) \bmod 360°$. $\hat{F}_i(t)$ is determined by Equation 1.2.

Thus the representation of position seems to sweep from behind the animal to in front of it with each theta cycle. I do not claim this as a model of how phase precession is actually generated in the rodent hippocampus, only that it produces a phase precession effect so that routes can be stored in the hippocampal recurrent connections.

We model the asymmetric nature of LTP by making the learning rule dependent on the product of the synaptic drive of the presynaptic neuron and the firing rate of the postsynaptic neuron:

$$\eta \frac{dw_{ij}}{dt} = [-w_{ij} + \tilde{w} \cdot (S_j \cdot F_i)^{\mu}]^{+} \tag{9.2}$$

This has the important dual effect of making LTP (1) correlational (because the weight increase is dependent on pre- and post-synaptic parameters) and (2) asymmetric (because it is dependent on firing rate $F_i$ of the post-synaptic neuron but synaptic drive $S_j$ of the pre-synaptic neuron. Synaptic drive can be understood as a decaying memory of recent firing rates shown by neuron $j$ with a decay time constant of $\tau_j$ (Pinto *et al.*, 1996).

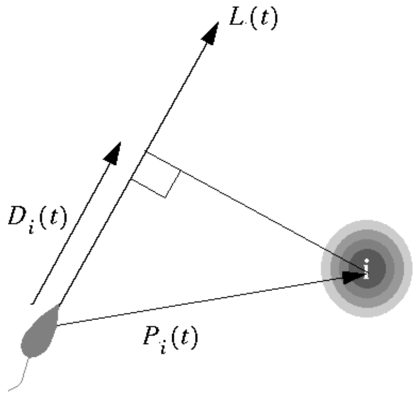



Figure 9.2: Approximation of phase precession. Let $L(t)$ be a ray originating at the simulated rodent's current position (as represented by the pyramidal cells in hippocampus), pointing in the direction of its current heading (as represented by the cells in postsubiculum). Let $P_i(t)$ be a vector from the represented position of the rodent to the center of the place field of place cell $i$, and $D_i(t)$ be the projection of $P_i(t)$ onto $L(t)$. Then the preferred phase of neuron $i$, $\tilde{\Theta}_i(t)$ is proportional to $D_i(t)$ using the relation $\tilde{\Theta}_i(t) = K \cdot D_i(t)$ where $K$ is a scale factor chosen to be small enough that the phase precession will not wrap around ($K = 1.2\,\mathrm{deg/cm}$ in our simulations). Thus, cells with place fields behind the represented position (in hippocampus) fire earlier in the theta cycle and cells ahead of the represented position fire later. We do not claim this as a model of how phase precession is actually generated in the rodent brain, only that it produces a phase precession effect so that routes can be stored in the hippocampal recurrent connections. From Redish and Touretzky (1997c).

The parameters used are listed in Tables 9.1 and 9.2. Simulations were done with a time step of 1 ms.

## 9.3 Simulation results

### 9.3.1 Storing routes

At this point, let us assume that an animal can, from local view information, self-localize and determine the path it must take to reach the hidden platform in the water maze, Figure 9.3 shows four paths that will be stored in the recurrent connections of the simulated hippocampus.

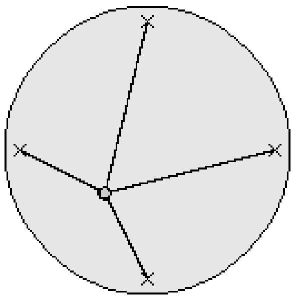


Figure 9.3: Four paths from the wall to the goal. Lines indicate trajectories taken by the simulated animal to reach the goal (indicated by small circle). An $\times$ has been drawn at the initial location of the animal in each position. Grey area denotes arena. From Redish and Touretzky (1997c).

The four effects detailed above combine to store routes in the recurrent connections of hippocampus. They produce a vector field pointing toward the path and then leading to the goal.

After the simulated animal traversed each route in Figure 9.3 once, we measured the vector field stored in the recurrent connections of hippocampus. For each cell $j$ in hippocampus, we calculated the center of mass of the output connection weights, and plotted an arrow from the place field center toward the center of mass. The length of the arrow in Figure 9.4 is linearly proportional to the distance between the center of cell $j$'s place field and the center of mass of cell $j$'s output connection weights.

### 9.3.2 Recalling routes

In order to say that routes have been stored in the hippocampus, there must be a mechanism for those routes to be recalled.

When there is sensory input into the hippocampus and the hippocampus is in LIA mode, sensory cues enter the system via superficial EC, and hippocampus place cells that

| | | |
|---|---|---|
| High-level Sensory (HLS) | $N_{HLS}$ | 400, $(20 \times 20)$ |
| Path integrator (PI) | $N_{PI}$ | 400, $20 \times 20$ |
| Superficial Entorhinal Cortex (ECs) | $N_{ECs}$ | 400 |
| Time constant, ECs | $\tau_{ECs}$ | 1 ms |
| Tonic inhibition, ECs | $\gamma_{ECs}$ | -2 |
| Hippocampal, Pyramidal pool HC:E | $N_{HC:E}$ | 400, $(20 \times 20)$ |
| Time constant, HC:E | $\tau_{HC:E}$ | 10 ms |
| Tonic inhibition, HC:E | $\gamma_{HC:E}$ | -1.5 |
| Phase precession parameter, HC:E | $\sigma$ | 25° |
| Hippocampal, inhibitory pool, HC:I | $N_{HC:I}$ | 400, $(20 \times 20)$ |
| Time constant, HC:I | $\tau_{HC:I}$ | 2 ms |
| Tonic inhibition, HC:I | $\gamma_{HC:I}$ | -7.5 |

Table 9.1: Parameters used in the dual-role hippocampal simulations.

| | | |
|---|---|---|
| Weight, HLS $\rightarrow$ ECs | $W_{HLS \rightarrow ECs}$ | 5.0 |
| Weight, ECs $\rightarrow$ HC:E | $W_{ECs \rightarrow HC:E}$ | 5.0 |
| Weight, ECs $\rightarrow$ PI | $W_{ECs \rightarrow PI}$ | 1.0 |
| Weight, PI $\rightarrow$ ECs | $W_{PI \rightarrow ECs}$ | 2.0 |
| Weight, HC:E $\rightarrow$ PI | $W_{HC:E \rightarrow PI}$ | 5.0 |
| Weight, HC:E $\rightarrow$ HC:I | $W_{HC:E \rightarrow HC:I}$ | 16.0 |
| Spread, HC:E $\rightarrow$ HC:I | $W_{HC:E \rightarrow HC:I}$ | 20 cm |
| Weight, HC:I $\rightarrow$ HC:E | $W_{HC:I \rightarrow HC:E}$ | -8.0 |
| Spread, HC:I $\rightarrow$ HC:E | $W_{HC:I \rightarrow HC:E}$ | 200 cm |
| Weight, HC:I $\rightarrow$ HC:I | $W_{HC:I \rightarrow HC:I}$ | -12.0 |
| Spread, HC:I $\rightarrow$ HC:I | $W_{HC:I \rightarrow HC:I}$ | 200 cm |
| Initial Weight, HC:E $\rightarrow$ HC:E | $W_{HC:E \rightarrow HC:E}$ | 5.0 |
| Spread, HC:E $\rightarrow$ HC:E | $W_{HC:E \rightarrow HC:E}$ | 20 cm |
| Learning rate, HC:E $\rightarrow$ HC:E | $\eta_{HC:E \rightarrow HC:E}$ | 0.01 |
| Power constant, HC:E $\rightarrow$ HC:E | $\eta_{HC:E \rightarrow HC:E}$ | 5.0 |
| Gain , HC:E $\rightarrow$ HC:E | $\tilde{w}_{HC:E \rightarrow HC:E}$ | 1.0 |

Table 9.2: More parameters used in the dual-role hippocampal simulations. Connections listed above the double line were instantiated as one-to-one connections (i.e. cell $j$ connects to cell $i$ iff they both correspond to the same location in the $20 \times 20$ grid). This was done for computational speed and should not affect the results. Other connections were instantiated as full connections in which the connection weight from input cell $j$ to output cell $i$, $w_{ij}$, falls off as a Gaussian function of distance between the place field centers of the two cells $i$ and $j$.

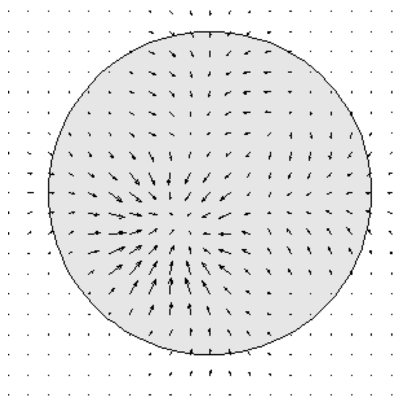


Figure 9.4: Asymmetric connections in hippocampus. For each cell $j$ in hippocampus, we calculated the center of mass of its output connection weights, and plotted an arrow from the place field center toward that center of mass. Length of arrow is linearly proportional to the distance between the center of cell $j$'s place field and the center of mass of cell $j$'s output connection weights. Arrows outside the environment are for cells with simulated place field centers outside the environment, but whose simulated fields overlap the environment. From Redish and Touretzky (1997c).

are consistent with the local view will be more active than those that are not. These place cells can be understood as representing candidate locations. This biases the place code to settle to a coherent representation of position, consistent with the local view, as shown in Chapter 8.

On the other hand, when there is no sensory input, this bias will be absent, but due to the symmetric component of the recurrent connections in hippocampus, the hippocampus will still settle to a coherent activity pattern that represents a valid location. Due to the asymmetric component, this representation will then drift along the remembered route. Figure 9.5 shows the place code settling to a coherent representation of location which then drifts along a recently traveled route.

Because these drifting representations consist of a coherent representation of location, we can measure the position represented. Figure 9.6 shows the believed position of the simulated rat during 20 of these route-replay sequences. This simulation is replaying routes leading to the goal.

### 9.3.3 Route-storage and self-localization: two incompatible processes?

The self-localization process described in Chapter 8 requires that the synaptic efficacy between two place cells be symmetric with strength proportional to the overlap between their place fields. The route-storage process described in this chapter requires that the synaptic


Figure 9.5: In the absence of sensory input, the hippocampus still settles to a coherent representation of position, which then drifts along the remembered route. From Redish and Touretzky (1997c).

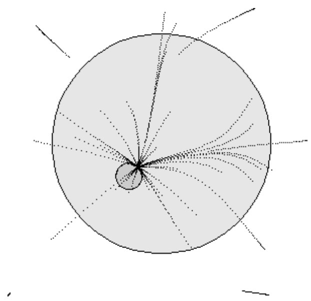


Figure 9.6: Replay of routes during LIA without sensory input. The drifting representations (see Figure 9.5) are replaying routes to the goal. Each dot indicates represented location at one time step during one replay sequence. Sequences were sampled at 10 Hz. 40 sequences are shown. From Redish and Touretzky (1997c).

efficacy be asymmetric, favoring connections projecting between cells with place fields along recently traveled routes. These two requirements would seem to be incompatible. They are not.

When the animal is self-localizing, there is sensory input from superficial EC providing candidate locations. The sensory input constrains the representation in hippocampus to one of the candidate locations and effectively locks the representation in place. Figure 9.7 shows the $x$-coordinate of the hippocampal representation after a simulated sharp wave occurred under three conditions: (light bars) when the simulated animal was at five different locations in an environment using only symmetric connection weights, (medium bars) the same five locations with both symmetric and asymmetric connection weights, and (dark bars) in the absence of sensory input, but when the sharp wave began as a representation near each of the five locations. There is no difference at all between the simulations shown in the light and medium bars, indicating that the local view input is sufficient to hold the representation in place. On the other hand, when there is no sensory input to hold the representation in place, it drifts, reaching the same final value in each case: the goal location. (See also Figure 9.6.) This shows that the recall and replay modes are in fact compatible, at least in simulation. There is also neurophysiological data suggesting that the two modes are not incompatible. (See Discussion, Section 9.4.2.)

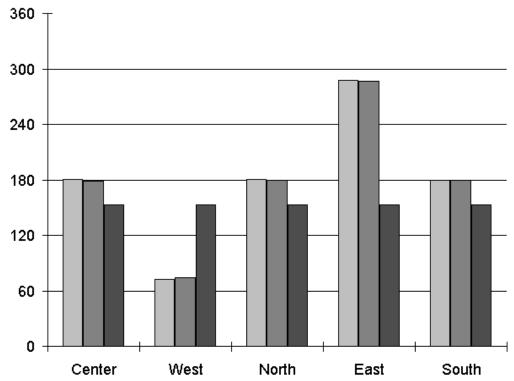



Figure 9.7: $X$-coordinate (on a toroidal map) of the final representation of the parallel relaxation process with (light bars) local view input and symmetric weights only, (medium bars) local view input and both symmetric and asymmetric weights, and (dark bars) both symmetric and asymmetric weights without local view input. $Y$-coordinates (not shown) are similar. From (Redish and Touretzky, 1997c).

| | | |
|---|---|---|
| HLS, size | $N_{\text{input}}$ | 400 |
| $\text{LTM}_H$, size | $N_{\text{hidden}}$ | 40 |
| $\text{LTM}_O$, size | $N_{\text{output}}$ | 2 |
| Learning Rate | $\eta$ | 0.0001 |
| Momentum | $\alpha$ | 0.1 |
| Derivative Increment | $\gamma$ | 0.2 |

Table 9.3: Parameters used in the consolidation simulation.

### 9.3.4 Training a network from the replayed routes

As reviewed in the Hippocampal Review (Chapter 4), the hippocampus may be involved in the temporary storage of memories, particularly of routes (Scoville and Milner, 1957; Squire and Zola-Morgan, 1988; Squire and Zola-Morgan, 1991; Squire, 1992; Cohen and Eichenbaum, 1993; Sutherland and Hoesing, 1993; Zola-Morgan and Squire, 1993; Squire and Alvarez, 1995; Rempel-Clower *et al.*, 1996, see also Navigation Overview, Section 3.4). However, recent results have called into question the limited nature of retrograde amnesia after hippocampal lesions in rats (Bolhuis *et al.*, 1994; Weisend *et al.*, 1996; Koerner *et al.*, 1996; Bohbot *et al.*, 1996).

In order to demonstrate the viability of the consolidation theory, we show that there is enough information in the sequences from Figure 9.6 to store the routes in Figure 9.3 in a slowly learning network.

We simulated long term memory as a three-layer feed-forward network (HLS, $\text{LTM}_H$, and $\text{LTM}_O$) and trained the weights (HLS $\rightarrow$ $\text{LTM}_H$ and $\text{LTM}_H \rightarrow \text{LTM}_O$) by the standard backpropagation algorithm (Rumelhart *et al.*, 1986b; Hertz *et al.*, 1991). Parameters are listed in Table 9.3. We do not want to claim that memory consolidation in cortex corresponds in any way to backpropagation of error. The important characteristics of this simulation are (1) that it learns slowly, (2) that there is enough information stored in the replayed hippocampal representations to generate the correct HLS $\rightarrow$ LTM function.

The training regimen for the LTM network consisted of 40 "dreams," each of which were 5 second sequences of hippocampal representations as the network was allowed to settle from noise and progress along the stored routes in replay mode as described in the previous section. The position represented by these sample inputs is shown in Figure 9.6. We used these representations to regenerate HLS representations and then generated directions of motion $\psi(t)$ by comparing each sample with its successor. We regenerated HLS($t$) by taking the represented position in hippocampus at time $t$ and setting the HLS representation to the representation of that position. In the animal, we expect this regeneration process to occur from feedback connections between ECd (deep entorhinal cortex) and parietal cortex. Because our hippocampal simulation is a 2D neural sheet, we can generate $\psi(t)$ by subtracting the center of mass of the hippocampus representation at time $t$ from the center of mass at time $t + \Delta t$. These input-output pairs formed the training set for the LTM network.

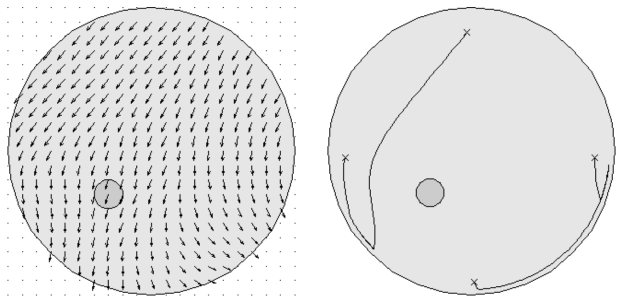
The three-layer LTM network had 400 inputs corresponding to each of the HLS neurons, 40 hidden units ($\mathrm{LTM}_H$), and 2 outputs ($\mathrm{LTM}_O$) representing $\cos(\psi)$ and $\sin(\psi)$. We do not claim that cortical representations look anything like this or that cortical learning occurs by backpropagation; we only want to demonstrate that enough information is contained in the sequences to train a cortical network on the desired input-output function. We trained the network on the 40 sequences for 1000 epochs (40,000 total presentations). Each sequence was sampled at 1 Hz, so five input-output pairs were selected from each sequence to train the network. This was done because subsequent pairs from a single sequence are very similar and this helped minimize computation time. In actual animals we expect the entire replayed route to be used and we expect each replayed route to be unique, rather than repeating the exact same one many times.

To test the simulated long term memory, we placed the sim-animal at evenly spaced positions in the environment (5 cm spacing) and measured the direction suggested at each location. Figure 9.8 (left) shows the output vector in $\mathrm{LTM}_O$ at each location.

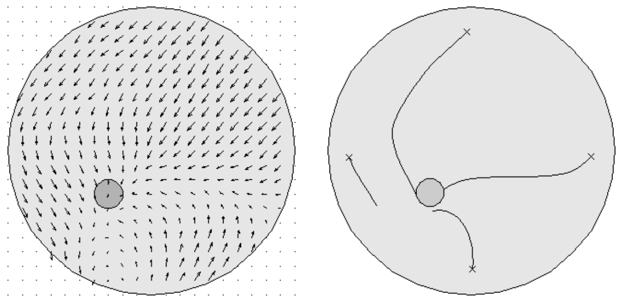
To demonstrate that the network successfully learned the task, we simulated placing the animal at each of the four cardinal locations in the water maze and moving it in the direction represented in $\mathrm{LTM}_O$. As the simulated animal moved through the environment, it was allowed to continue following the direction represented in $\mathrm{LTM}_O$ (which changed as the local view changed). If the animal hit the wall, it turned clockwise if the direction represented in $\mathrm{LTM}_O$ was to the right of the wall's normal vector and left otherwise. It turned in 45° increments until it could proceed. This has the effect of moving the animal along the wall. If the length of the $\mathrm{LTM}_O$ vector was zero, the animal proceeded along the shortest path to the wall until the $\mathrm{LTM}_O$ vector was non-zero or the animal hit the wall. This simulates the thigmotaxis seen in a variety of rodents. The simulation ended after 60 seconds or when the animal was within 10 cm of the goal. Figure 9.8 (right) shows the paths taken from the four starting locations.

Anterograde and retrograde amnesia can be clearly seen in Figure 9.8, as can the limited nature of the retrograde amnesia. If the hippocampus is lesioned before long term memory has been trained,[3] the animal cannot solve the task at all (anterograde amnesia). If the hippocampus is lesioned after long term memory has been partially trained, the animal does at least approach the goal, but still shows severe impairments (retrograde amnesia). Finally, if the long term memory is fully trained from the hippocampal dreams, then the animal can solve the task accurately even with hippocampal lesions (retrograde amnesia is limited).

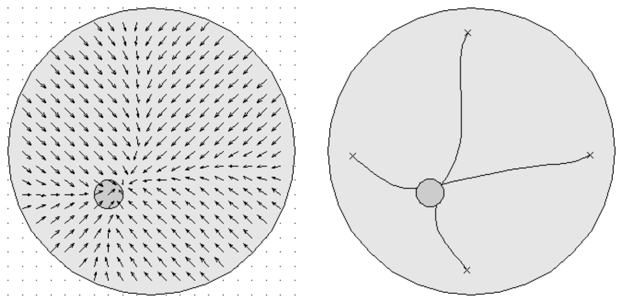
[3] The simulation had initial weights randomly and uniformly distributed in the range $[-0.5, +0.5]$. It also assumed that no learning took place in long term memory without training from replayed routes in hippocampus. Evidence that animals can learn to solve the hidden-platform water maze under certain conditions (belying this assumption) is discussed in Section 9.4.3, below.



(a) untrained



(b) partially trained



(c) completely trained

Figure 9.8: Memory consolidation in a network trained on replayed routes. (Left) The vector field of directions to go represented at each location by the long term memory network. (Right) Paths taken by a simulated animal with hippocampal lesions, navigating using just its cortical route-memory. Light grey area denotes arena; dark grey area denotes platform. From Redish and Touretzky (1997c).

## 9.4 Discussion

### 9.4.1 Relationship between the two modes

The two roles played by the hippocampus occur by variations of the same mechanism: both are simply examples of a pWTA network settling to a stable state but with different boundary conditions. The route-replay process occurs when the initial state is uniform noise and no extrinsic input is given; the self-localization process occurs when the initial state includes candidate locations and these candidate locations continue to be input throughout the self-localization process. The candidate locations can be seen as a signal input into the system mixed with noise. These two processes then are two extremes of a mathematical continuum: with no signal, the system performs route-replay, but with a strong signal, the system performs self-localization.

An interesting (and as yet unanswered) question is how a pWTA network behaves as the signal is increased from 0 (which produces replay) to a very strong signal (which produces recall). Is there partial replay with weak signals? Is there a point at which the system shows a phase change and suddenly shows strong replay instead of strong recall? The complex non-linearities of pWTA networks (see, for example, the *resolution issue*, Appendix A) make an analytic solution extremely difficult, however, simulations would be straight-forward and might show very interesting effects.

Although this is an interesting question that might be a fruitful avenue for further research, the claim being made here is independent of the dynamical behavior of the system at these intermediate parameterizations. The hypothesis here is that the rodent hippocampus plays two roles (recall and replay), which occur at different times (awake LIA and sleep LIA, respectively), and that these two roles can be accomplished by this same mechanism under these two boundary conditions. I refer to the system as having three modes (storage, recall, and replay) because the dual role hypothesis implies that the hippocampus does not utilize intermediate states.

### 9.4.2 Place fields shifting with experience

There is interesting neurophysiological evidence suggesting that the hippocampus may play two roles: increased experience in an environment increases the overlap along routes (improving route-storage) but does not change the centers of place fields (not impairing self-localization). Mehta *et al.* (Mehta *et al.*, 1996; Mehta and McNaughton, 1996) recently showed that the place fields of cells shift backwards along a much-repeated path as the animal runs along that path (EX.75). They trained a rat to run in a loop on a rectangular elevated maze and recorded place cells in CA1. They found that the place fields shifted backwards over the course of a single session. This would seem to be a problem for the self-localization process because the location represented by the place cell population changes with time. However, they also showed that while the area covered by the place field shifted by almost 50% (approx. 5–7 cm), the center of mass of the field actually shifted very little

(20%, approx. 2–3 cm) (Mehta and McNaughton, 1996). This is exactly what we need for the dual-role hippocampus (self-localization/route-storage). Because the number of cells overlapped by the expanded place field increases (due to the increase in area covered), longer routes can be stored, but because the center of mass of the field doesn't shift much, the self-localization process won't be affected very much.

### 9.4.3 Why can't the animal learn the task without a hippocampus?

Even with hippocampal lesions, animals can still learn simplified versions of the Morris water maze under specific conditions. If, for example, the platform is visible or the animal only begins from a single starting point or there is a landmark indicating the direction to swim (Eichenbaum *et al.*, 1990; McDonald and White, 1994), then the animal can perform at near-normal levels. These simplified versions of the task all allow routes to be learned based on single cues (the visible platform, a single motor program, a directional cue).

Brown and Sharp (1995) built a model of nucleus accumbens which uses reinforcement learning (Sutton, 1992) to back the direction-to-goal function from near the goal to locations distant from it. Their model learns to associate directions with representations of position. They use place cells as inputs, but there is no reason that the same process would not work with local view representations. However, such a mechanism (as envisioned by Brown and Sharp) must not exist separate from the hippocampus, because animals cannot learn to solve the Morris water maze without a hippocampus. As reviewed in the Navigation Overview (Chapter 3), there are three systems which allow animals to navigate within an environment:

- **Locale system.** Animals can use constellations of distal cues to localize themselves on a "cognitive map;" combined with a memory of the goal location on the same map, the animals can navigate to the goal. This system requires a hippocampus.
- **Taxon system.** Animals can use visual cues to drive navigation directly towards or away from the cue. Evidence exists that this system requires the caudate nucleus (McDonald and White, 1994; Packard and McGaugh, 1992; Packard and McGaugh, 1996). Because there is no visible cue in the hidden-platform water maze task, it would not help the animal find the platform.
- **Route system.** Routes stored in the hippocampus can be written out to cortex, so that directions necessary to reach a goal are associated with local views. This system requires training for each step the animal must take; it cannot learn to associate local views with directions to distant goals without hippocampal help (through route-replay).

If there was a way to show the animal the route to the goal, it might be possible to train the route system even without a hippocampus. Whishaw *et al.* (Whishaw *et al.*, 1995; Whishaw and Jarrard, 1996) and Schallert *et al.* (1996) both showed ways to train the route system directly and found that animals could learn to solve the water maze even with hippocampal lesions (EX.16).

Whishaw *et al.* trained animals with fimbria/fornix lesions (Whishaw *et al.*, 1995) and hippocampal lesions (Whishaw and Jarrard, 1996) to find a visible platform and then removed the visible platform. Whishaw *et al.*'s animals concentrated their search where the platform had been.

Schallert *et al.* (1996) used animals with kainate-colchicine hippocampal lesions (which destroys both the DG and CA3/CA1 fields). The animals were first trained with a large hidden platform that filled almost the entire maze. Once the animals could reach that platform reliably, it was shrunk trial by trial until it was the same size as a typical platform in a water maze task. Again, the animals could learn to solve the water maze without a hippocampus.

One way to explain these results is that non-hippocampal processes are being used to train the route system. In the Whishaw *et al.* experiments, the taxon system could be driving navigation to the visible platform and then the route system could learn from the actual traversals. In the Schallert *et al.* experiments, the route system may be being explicitly trained by their experimental methodology.

# Part III

# Simulations of specific experiments

Having presented reviews and overviews of the experimental, navigation, and hippocampal literature (Part I, Chapters 2–4), and having discussed specific contributions I have made to specific aspects of our understanding of rodent navigation (Part II, Chapters 5–9), I will now present simulations of specific experiments. In the following chapters, I present:

- (Chapter 10) a simulation of the recent result from Barnes *et al.* (1997), showing that the model produces a bimodality in the representation of location in animals with deficient LTP. Compare EX.83.
- (Chapter 11) a simulation of Sharp *et al.* (1990), showing the first simulations capable of replicating all the single place field conditions reported in their paper. Compare EX.73, Figure 2.22.
- (Chapter 12), simulations of Cheng (1986) and Margules and Gallistel (1988) showing the importance of disorientation in self-localization. Compare EX.35.
- (Chapter 13), simulations of Morris (1981), showing that the model can replicate navigation in the water maze.
- (Chapter 14), simulations of Collett *et al.* (1986) and our own gerbil navigation results (Section 2.5), showing that the model can replicate a number of reactions to different manipulated landmark arrays.

# Chapter 10

# Returning to the wrong map

Barnes *et al.* (1997) allowed an animal to walk around a Figure-Eight maze for 25 minutes. They then removed the animal for one hour, after which it was returned to the maze and allowed to walk around for another 25 minutes. During each 25 minute experience, Barnes *et al.* recorded about three dozen place cells simultaneously.

When young animals returned to the environment, they used the same set of place cells to encode location within the environment. But when old animals returned to the environment, they sometimes used a completely different set of cells. The correlation between the place fields in the two experiences was always high for young animals (approximately 0.7), but bimodal (sometimes near 0, othertimes near 0.7) in old animals. Within an experience, the place cells were very stable. Correlations between the first and second halves of a single experience were always high for both old and young animals. See EX.83.

Barnes *et al.* (1997, see also McNaughton *et al.*, 1996; Samsonovich and McNaughton, 1997; Samsonovich, 1997) interpret this as a problem in selecting the correct pre-configured cognitive map. As reviewed in the Hippocampal Review (Section 4.1.5-D), their theory includes a set of pre-wired *charts* in the hippocampus, such that the synaptic weight between two cells in hippocampus is inversely related to the minimum of their distances across all of the charts. This theory then presents the hippocampus as an attractor network which settles to a coherent representation of position on one of the charts (see Chapter 8 for a discussion of this type of network). As pointed out by Barnes *et al.*, this *mis-localization* effect can explain the bimodality seen in the ability of old animals' to solve the hidden platform water maze (EX.28).

According to the *multi-chart model* of hippocampus (McNaughton *et al.*, 1996; Samsonovich and McNaughton, 1997; Samsonovich, 1997), on entering a novel environment, one location on one *chart* will win the competition. As young animals explore the environment, a representation of the local view gets bound to the currently active chart. On returning to the environment, the local view representation biases the pseudo-winner-take-all dynamics in the hippocampus, and the same representation of location on the same chart is reinstantiated. Because old animals have LTP deficiencies (see Barnes, 1994; Barnes *et al.*, 1996; Barnes, 1996, for reviews), the local view would not become as tightly bound to the currently active chart. Thus (according to the multi-chart model), on returning to the environment, there would be a much weaker bias to select the same location on the same chart.

However, the complexity of the pre-wired intra-hippocampal connections required by the multi-chart model is unnecessary. The model discussed in Chapter 7 can explain this result while hypothesizing much simpler pre-wired connections.

## 10.1 Contributions

I propose that this result is not a consequence of pre-wired chart selection within the CA3 population, but rather an interaction between the non-linearity of the path integrator and the orthogonalization properties of dentate gyrus.

According to this theory, when young animals return to the environment, LTP has created associations between local view and hippocampus and between hippocampus and the path integrator. Thus during the self-localization process, the local view representation instantiates representations in hippocampus and representations in hippocampus force the path integrator to reset to the same representation of location as in the young animal's previous experience. In old animals, however, LTP is deficient (as reviewed by Barnes, 1994; Barnes *et al.*, 1996; Barnes, 1996) and thus there is little or no bias to reset the path integrator to the same location. Because each DG cell performs a logical "and" function of its path integrator and local view inputs, a change in path integrator representation will produce a dramatic change in DG representation which will be seen in hippocampus as a low overlap between each experience.

### 10.1.1 Similarities and differences

The interpretation I have proposed for this experiment and the interpretation described in the chart-model have some similarities and some crucial differences.

Because place cells are active on initial entry into the environment (Hill, 1978; Austin *et al.*, 1990; Wilson and McNaughton, 1994; Tanila *et al.*, 1997b, see EX.72), there must be some pre-wired connections producing place field activity. In the model discussed in Chapter 7, pre-wired connections between PI $\rightarrow$ DG, LV $\rightarrow$ DG, and DG $\rightarrow$ HC produce place cells with stable place fields on initial entry into the environment. The difference between these hypothesized pre-wired connections and the connections hypothesized by the McNaughton *et al.* chart-model is that the connections in the model presented here are totally random. There does not have to be any pre-defined correlation between them.

Because the place cell instability observed by Barnes *et al.* (1997) is bimodal in older animals, there must be some sort of nonlinear process. In the chart-model, this nonlinearity occurs by winner-take-all competitive dynamics between charts in CA3. In the model discussed in Chapter 7, it occurs because of nonlinearities in the settling of the path integrator.

This can be understood in terms of the resolution issue (Appendix A). If there is a very weak bias input into the path integrator representation and it has begun to settle to a nearby location, the bias will pull the representation to overlap strongly. If the path integrator representation has begun to settle to a distant location, the bias will simply be squashed and have no effect. (The bias is non-zero because old animals do have some weak LTP, it is not non-existent, see Barnes, 1994; Barnes *et al.*, 1996; Barnes, 1996, for reviews.)

This nonlinearity means that the path integrator will either strongly overlap with previous experience or not. If it does strongly overlap, the same set of DG cells will be active. (I assume that the local view has not changed from the previous experience.) If it does not strongly overlap, a different set of DG cells will be active.

Because the place cell instability observed by Barnes *et al.* (1997) only occurs on entry into the environment, there must be something special about entry into the environment. The self-localization process hypothesized to occur on entry into the environment would explain this. During normal navigation, the path integrator does not reset, it continues to be driven by internal dynamics more than external. But during self-localization, the path integrator is reset and external dynamics can have a strong influence.

## 10.2 Simulation details

Simulations consisted of the model shown in Figure 8.3 using the same basic structure and equations as in Chapter 8. In order to make the computations more tractable, the path integrator consisted of only 225 neurons ($15 \times 15$). This forced some minor parameter changes. In order to prevent any confusion, the complete parameter set used for the simulations in this chapter is listed in Tables 10.1 and 10.2.

Because we are only interested in the "returning" abilities of the simulation, the simulation was entered at a single location in the environment (which engendered a self-localization process) and allowed to learn at that location for a short time. It was then removed from the environment and returned to that same location with hippocampus initialized to random noise (but HLS accurately representing the position of the sim-animal). This cycle of removal and return was repeated 10 times each for the simulated young (normal LTP) and old (weak LTP) animals.

## 10.3 Simulation results

Because the connections from HLS to HC and HC to PI learned the correct mapping in the simulated "young" animals, the system always reset the path integrator to the same coordinates (Table 10.3, left). This then produced the same $\mathcal{PI} \times \mathcal{LV}$ association and thus the same DG and CA3 cells were active. All correlations seen are high, see left side of Figure 10.1.

However, in the simulated "old" animals, the path integrator settled to different coordinates most of the time (Table 10.3, right). Since the local view had not changed, this changed the $\mathcal{PI} \times \mathcal{LV}$ association. This engendered a change in the DG and thus the CA3 representations. Typical correlations were around 0. Occasionally, the path integrator would reset to a nearby location producing a rare high correlation (0.6–0.7), see right side of Figure 10.1.

## 10.4 Discussion

These simulations show that the dramatic nonlinearity seen by Barnes *et al.* (1997) does not necessarily imply that there must be pre-wired charts in the hippocampus. A model that includes random pre-wired connections from an external path integrator and an external local view into an orthogonalizing structure (such as dentate gyrus) as well as a single-map pre-wired path integrator is sufficient to produce the observed bimodality in animals with deficient LTP.

This demonstrates why it is important to consider *reference frame* as a property of the entire navigation system and not just the hippocampus (see Section 3.3). What I am suggesting here is that the old animals are on a different map (different chart, different reference frame) because they are using a different reference point for the path integrator. Because the path integrator representation is changed between experiences, the $\mathcal{PI} \times \mathcal{LV}$ association changes and the animal will have to use a different reference frame.

| | | |
|---|---|---|
| Centers, $r$-type sensory neurons | $\tilde{r}$ | $\{0 \text{ cm}, 10 \text{ cm}, \ldots, 100 \text{ cm}\}$ |
| Gain, $r$-type sensory neurons | $\sigma_r$ | 25 cm |
| Centers, $\phi$-type sensory neurons | $\tilde{\phi}$ | $\{0°, 5°, \ldots, 355°\}$ |
| Gain, $\phi$-type sensory neurons | $\sigma_\phi$ | 25° |
| Path integrator size (excitatory neurons) | $N_{PI:E}$ | 225 ($15 \times 15$) |
| Time constant, excitatory PI | $\tau_{PI:E}$ | 10 ms |
| Tonic inhibition, excitatory PI | $\gamma_{PI:E}$ | -1.5 |
| Path integrator size (inhibitory interneurons) | $N_{PI:I}$ | 1 |
| Time constant, inhibitory PI | $\tau_{PI:I}$ | 2 ms |
| Tonic inhibition, inhibitory PI | $\gamma_{PI:I}$ | -7.5 |
| Dentate gyrus | $N_{DG}$ | 10000 |
| Time constant, Dentate gyrus | $\tau_{DG}$ | 1 ms |
| Tonic inhibition, Dentate gyrus | $\gamma_{DG}$ | -6.0 |
| Hippocampus, excitatory neurons | $N_{HC:E}$ | 500 |
| Time constant, Hippocampus | $\tau_{HC:E}$ | 10 ms |
| Tonic inhibition, Hippocampus | $\gamma_{HC}$ | -2.0 |

Table 10.1: Parameters used in the simulation of Barnes *et al.* (1997).

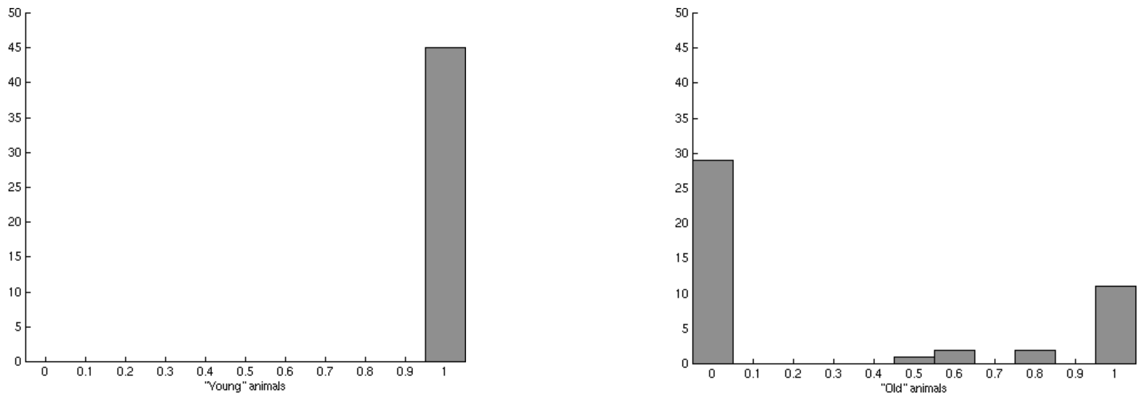



Figure 10.1: Histograms of correlations between representations of a specific location in the environment after multiple entries into that environment. Sim-animals were allowed to enter an environment 10 times and the hippocampal representation of a location in the environment was measured. Cross-correlations were made between each pair of experiences in the environment (45 pairs). Plotted is a histogram of the correlations found. (left) Simulations with strong LTP (i.e. "young" animals) always return to the same PI representation, so the correlation is always high. (right) Simulations with weak LTP (i.e. "old" animals) usually return to different PI representations and so usually have very low correlations, but occasionally return to similar PI representations and thus have rare highly correlated pairs.

| | | |
|---|---|---|
| Weight, PI:E → DG | $W_{PI:E \to DG}$ | 4.5 |
| Weight, HLS → DG | $W_{HLS \to DG}$ | 4.5 |
| Weight, DG → HC | $W_{DG \to HC}$ | 3.0 |
| Learning rate, HLS → HC ("young animals") | $\eta_{HLS \to HC}$ | 0.1 |
| Learning rate, HLS → HC ("old animals") | $\eta_{HLS \to HC}$ | 0.005 |
| Gain, HLS → HC | $\tilde{w}_{HLS \to HC}$ | 1.5 |
| LTP sharpening constant, HLS → HC | $\mu_{HLS \to HC}$ | 5 |
| Learning rate, HC → HC ("young animals") | $\eta_{HC \to HC}$ | 0.1 |
| Learning rate, HC → HC ("old animals") | $\eta_{HC \to HC}$ | 0.005 |
| Gain, HC → HC | $\tilde{w}_{HC \to HC}$ | 1.0 |
| LTP sharpening constant, HC → HC | $\mu_{HC \to HC}$ | 5 |
| Learning rate, HC → PI:E ("young animals") | $\eta_{HC \to PI:E}$ | 0.1 |
| Learning rate, HC → PI:E ("old animals") | $\eta_{HC \to PI:E}$ | 0.005 |
| Gain, HC → PI:E | $\tilde{w}_{HC \to PI:E}$ | 2.0 |
| LTP sharpening constant, HC → PI:E | $\mu_{HC \to PI:E}$ | 5 |
| Weight, PI:E → PI:E | $W_{PI:E \to PI:E}$ | $6.0 \cdot \exp(-(\Delta_{ij})^2/(12\,\mathrm{cm})^2)$ |
| Weight, PI:I → PI:E | $W_{PI:I \to PI:E}$ | -8.0 |
| Weight, PI:E → PI:I | $W_{PI:E \to PI:I}$ | 0.222 |

Table 10.2: More parameters used in the simulation of Barnes *et al.* (1997).

| Entry | "Young" animals | "Old" animals |
|---|---|---|
| 1 | $(120°, 12°)$ | $(120°, 12°)$ |
| 2 | $(120°, 12°)$ | $(35°, 131°)$ |
| 3 | $(120°, 12°)$ | $(36°, 113°)$ |
| 4 | $(120°, 12°)$ | $(91°, 63°)$ |
| 5 | $(120°, 12°)$ | $(62°, 109°)$ |
| 6 | $(120°, 12°)$ | $(178°, 326°)$ |
| 7 | $(120°, 12°)$ | $(128°, 11°)$ |
| 8 | $(120°, 12°)$ | $(126°, 12°)$ |
| 9 | $(120°, 12°)$ | $(126°, 12°)$ |
| 10 | $(120°, 12°)$ | $(125°, 12°)$ |

Table 10.3: Coordinates of the path integrator for the entry point after each of 10 entries into an environment. All entries were into the same location in same environment. Since the path integrator was simulated as a torus, PI coordinates are given as pairs of angles. These coordinates can be taken as indicative of the reference point that would be used for any subsequent navigation.

# Chapter 11

# Interactions between entry point and cue cards

Sharp *et al.* (1990) examined rats in a small cylindrical arena with a cue card subtending 90° (see EX.73). During training the animals were not disoriented before entering the arena, and they always entered at the same location (the northwest corner). Sharp *et al.* then tested the animals with (1) the original configuration, (2) the cue card on the opposite side from its original location, (3) both cue cards from configurations 1 and 2, and (4) two cue cards rotated by ±30°. They tested each of these cases in two conditions: with the animal entering the arena at the northwest corner and with the animal entering the arena at the southeast corner.

Sharp *et al.* found that cells usually continued to show a single compact place field, but that the location of the place field was a complex function of both the orientation of the cue cards and the entry point. See Figure 2.22.

The results found by Sharp *et al.* can be explained by interactions between *believed initial values* and *local view input* in the head direction system. Because the animal was always placed into the environment at the same location (the NW corner), I will assume that the animal expected the wall to be at a specific allocentric bearing, and the cards to be at another allocentric bearing. This can be modeled by assuming that there are three kinds of head direction input: an initial value (coming from the fact that the animal is not disoriented), local view information from the arena wall, and local view from distal landmarks (i.e. the head direction suggested by the cue card(s)).

## 11.1 Relation to prior work

In their discussion section, Sharp *et al.* (1990) pointed out that the data implied that the place fields had to be understood as a consequence of both *mnemonic* processes based on the entry point in the environment and *sensory* processes based on the location of the cue card(s). The mnemonic processes can be understood as equivalent to the hypothesis that the path integrator helps to drive place cell activity (see Section 3.2.4, as well as Chapter 7): once the path integrator has been reset based on the various external cues available, it drives the place fields and keeps them from doubling.

Muller *et al.* (1991a) suggested that the data could be explained if the orientation of the place field was found by a majority vote of three cues: (1) the cue card(s), (2) intial entry point (treated as a *virtual landmark*), and (3) the external (room) cues. Although this explanation works for specific cases (**a**–**f**, Figure 2.22), it cannot explain the interaction effect seen in cases **g** and **h** (SE entry point; two cue cards rotated by $\pm 30°$; place field rotates by $180° \pm 30°$).

Redish and Touretzky (1996b, see also Touretzky and Redish, 1996) presented simulations of this experiment. We suggested that the results could be a consequence of a complex re-entry process dependent on intial hypotheses for head direction and path integrator representations, as well as local views to the wall and cue cards. This process used a measurement of the *consistency* between the internal and external cues. We were able to explain and simulate all cases except case **c** (NW entry point; one cue card one opposite from training; place field does not change). We hypothesized that case **c** was a consequence of the cue card no longer influencing the place field. The explanation presented in this chapter is completely different from the model presented in Redish and Touretzky (1996b).

Samsonovich and McNaughton (1997) suggested that the reason one position is chosen over another in the two cue card case is because "the most clearly visible cue card" provides stronger input into the place code than the other, less visible cue card. This explanation is incompatible with case **c** (Figure 2.22) in which the place field is dependent on the entry point, but not the cue card.

## 11.2 Simulation details

The simulation of Sharp *et al.* (1990) consisted of three parts, which were simulated sequentially. First, the system determined a representation of orientation and then it determined a representation of position. Finally, the sim-animal was allowed to wander around the arena while a simulated place cell was recorded.

In this simulation, both distal cues (the cue cards) and local cues (the arena wall) were included. The distal cues informed the head direction system and the local cues informed the place code.

### 11.2.1 Determining orientation

As before, all neurons were simulated using the three equation model based on the Pinto *et al.* (1996) interpretation of the Wilson and Cowan (1972) model (Eqs. 1.1–1.3).

Simulations consisted of three neuronal populations, $HLS_{\mathcal{D}}$, HD:E, and HD:I. $HLS_{\mathcal{D}}$ represented the distal cues (the cue cards) using the representation described in Chapter 8.

The HD population consisted of 24 excitatory neurons evenly spaced in a ring around 360° and one inhibitory interneuron. The excitatory neurons were interconnected with a weight matrix with a Gaussian distribution. The HD population was initialized to the sum of a representation of a head direction of 0° (simulating non-disorientation) and a representation of a head direction prediction from the expected head direction given the nearest orientation of the arena wall (either 0° when the animal entered at the NW or 180° when the animal entered at the SE). This second aspect is an abstraction of the effect of the animal expecting the arena wall at a certain orientation, which can be assumed because it was always placed into the environment at the same location during training (the NW).

These simulations were allowed to settle to a stable state (100 ms at a time step of 1 ms).

| | | |
|---|---|---|
| $HLS_{\mathcal{D}}$ | $N_{HLS_{\mathcal{D}}}$ | 24 |
| $HLS_{\mathcal{D}}$, Gain | $\mu$ | 0.1 |
| $HLS_{\mathcal{D}}$, Spread | $\sigma$ | 10° |
| HD excitatory population | $N_{HD:E}$ | 24 |
| Time constant, HD:E pool | $\tau_{HD:E}$ | 10 ms |
| Tonic inhibition, HD:E pool | $\gamma_{HD:E}$ | -1.5 |
| HD inhibitory population | $N_{HD:I}$ | 1 |
| Time constant, HD:I pool | $\tau_{HD:I}$ | 2 ms |
| Tonic inhibition, HD:I pool | $\gamma_{HD:I}$ | -7.5 |
| Weight, $HLS_{\mathcal{D}} \rightarrow$ HD:E | $W_{HLS_{\mathcal{D}} \rightarrow HD:E}$ | 1.0 |
| Weight, HD:E $\rightarrow$ HD:E | $W_{HD:E \rightarrow HD:E}$ | $6 \cdot \exp(-(\Delta\phi_{ij})^2/(36°)^2)$ |
| Weight, HD:E $\rightarrow$ HD:I | $W_{HD:E \rightarrow HD:I}$ | 1.6 |
| Weight, HD:I $\rightarrow$ HD:E | $W_{HD:I \rightarrow HD:E}$ | -12 |

Table 11.1: Simulation parameters used in the directional component of the simulation of Sharp *et al.* (1990).

### 11.2.2 Determining location

After the head direction settling process is complete, the system performs a self-localization process (Chapter 8). The self-localization simulation consisted of the same five populations shown in Figure 8.3: $HLS_{\mathcal{L}}$, PI, DG, and HC. As before, $HLS_{\mathcal{L}}$ represented the local cues (the arena wall) using the representation described in Chapter 8. The modifiable connections were learned using the same model of LTP as in Chapter 8 and were trained by allowing the system to explore the environment for five simulated minutes (simulated at a time step of 10 ms).

The self-localization simulation was allowed to settle to a stable state (100 ms at a time step of 1 ms).

### 11.2.3 Mapping out the place field

Once the system had self-localized, the sim-animal was allowed to wander around the environment for 5 simulated minutes, and a place cell with a place field near the one shown by Sharp *et al.* was mapped out. This was done by measuring the firing rate $F_i$ of the cell and the position of the simulated animal at each time step. A two-dimensional histogram of average firing rate by position was then plotted. Simulations were done using a 10 ms time step.

Throughout this wandering the 1D ring of the head direction representation was assumed to always show a coherent representation of direction (i.e. it rotated as the animal rotated), but the initial representation of direction was that determined above. Similarly, the 2D neural sheet of the path integrator was assumed to correctly show a coherent representation of position (i.e. the representation moved as the simulated animal moved), but the initial representation of location was that determined above. This allowed me to examine the location of place fields without a complex simulation of accurate tracking of head direction or path integration (see Chapters 5 and 6 for discussions about accurate tracking of these systems).

## 11.3 Simulation results

The simulations qualitatively replicate all of the cases shown in Figure 2.22. The specific simulation results are shown in Figures 11.1 and 11.2.

An important result shown here is that when the cue card input is near the initial head direction representation (as suggested by either the entry point or the non-disoriented initial head direction), then it pulls the head direction representation to match it (see Figure A.2), but if the cue card input is dramatically different it does not reset the head direction representation (see Figure A.3). This is a demonstration of the resolution issue described in Appendix A.

A prediction can be made as a consequence of the resolution issue discussed in Appendix A. In Figure 11.1**c**, the cue card does not draw the place field from its normal orientation, but in Figure 11.2$\boldsymbol{\alpha}$ it does. This is because in 11.1**c**, the head direction suggestion made by the cue card is too different from the initial head direction representation but in 11.2$\boldsymbol{\alpha}$ it is close enough to draw it aside by precession (see Figure A.2). From this we can predict that if case **c** is repeated with the cue card near the original orientation (within 45°), then the place cell will follow the cue card.

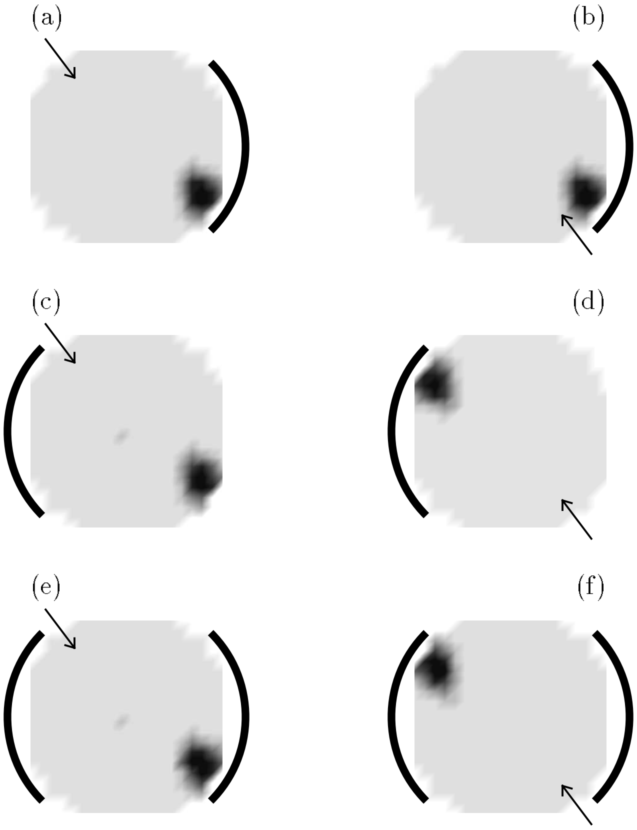



Figure 11.1: Simulations of results of Sharp *et al.* (1990), compare Figure 2.22. Letters indicate panels matching Figure 2.22; panel **a** shows the training case.

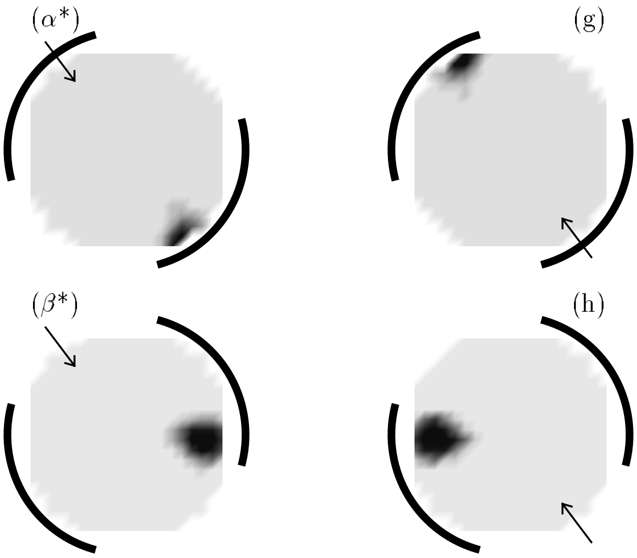



Figure 11.2: Simulations of results of Sharp *et al.* (1990), compare Figure 2.22. Letters indicate panels matching Figure 2.22. * Cases $\alpha$ and $\beta$ were not reported by Sharp *et al.*

# Chapter 12

# Disorientation in a rectangular arena

In Sharp *et al.* (1990), the animals were not disoriented; the theory in this thesis acknowledges an important distinction between whether animals have been disoriented or not. This distinction has been demonstrated behaviorally in rats by Gallistel *et al.* (Cheng, 1986; Margules and Gallistel, 1988; Gallistel, 1990, see EX.35).

Cheng (1986) tried to train rats to find food at one corner of a rectangular arena. In order to make the corners as distinct as possible, he placed a panel at each one, covered with a different type of material. In addition, the panels had different numbers of pinholes through which light was visible, and two of the panels had unique odorants behind them. Cheng disoriented the rats before placing them in the arena at a random location, and found that although the rats were able to distinguish one pair of diagonally opposed corners from the other, they could not distinguish the two corners in each pair. Cheng reports that the animals chose the correct corner in approximately 50% of the trials, and in the other 50%, they chose the corner opposite it. This suggests that the rats are sensitive to the geometric structure of the environment, and were ignoring other cues that could distinguish between the two corners. Margules and Gallistel (1988) replicated the experiment without disorientation and found that most animals had no difficulty selecting the correct corner over 75% of the time; some achieved better than 90% success rates.

We[1] will not present simulation results based on learning but can provide an explanation for the effect. Assume that each corner is actually represented by two landmarks — one based on geometric cues (e.g. lengths of adjacent walls) and one based on non-geometrical cues (odor, etc.). Then let us imagine the animal is disoriented and placed into the environment. During early trials, it may self-localize to the wrong orientation (off by 180°).

[1] This work was done with D. S. Touretzky (see Touretzky and Redish, 1996) and follows similar work done with D. S. Touretzky and H. S. Wan (see Touretzky *et al.*, 1994).

The association between the geometric components and the head direction representation will still be strengthened because it the environment is symmetrical around 180° rotation, but the association with the non-geometric components will not be. If we assume that this process has run its course and only geometric cues are associated with head direction representations, then we can show that the system will reorient correctly only 50% of the time (Figure 12.1, left). On the other hand, if there is an initial value that is close to the correct direction, the system will always orient correctly (Figure 12.1, right).

## 12.1 Relation to prior work

Gallistel (1990) hypothesized a *geometric module* involved in navigation that is primary over distinctive sensory-based cues. This can be seen as similar to the locale vs. taxon issue (O'Keefe and Nadel, 1978, see Navigation Overview, Chapter 3).

In Touretzky *et al.* (1994), we presented simulations of a similar experiment described by Cheng (1986) and Gallistel (1990). In these simulations, place cells were sensitive to Gaussians of allocentric bearing and distances to corners and the corners were differentiated based on the lengths of their adjacent walls (i.e. type A had a long left wall and a short right wall, while type B had a long right wall and a short left wall). Cells were also sensitive to path integrator input. Non-linearities in the self-localization process produced place fields that flipped when the arena was rotated beyond 180°.

In Touretzky and Redish (1996), we presented simulations of this experiment, but the effect occurred as a consequence of complex non-linearities in the head direction aspect of the self-localization algorithm. The basic concepts presented in this chapter are similar even though the particulars of the mechanism differ. (The mechanism here can be understood as a neural implementation of the algorithm presented in Touretzky and Redish, 1996.)

This simulation can be interpreted as following directly on the ideas of McNaughton *et al.* (1991), Skaggs *et al.* (1995), and Zhang (1996a). The key components are (1) that landmarks (the *geometric* cues) are associated with internal representations of direction (McNaughton *et al.*, 1991) and (2) that the the internal representation of direction is maintained by a ring attractor network (Skaggs *et al.*, 1995; Zhang, 1996a).

The idea that external cues are only associated with internal cues when the external cues are stable with respect to the internal cues was discussed extensively by Knierim *et al.* (1995), who also showed data supporting it. I am not going to show the learning sequence, but I will assume that in both cases only the geometric cues have been associated with the correct head direction.

What these simulations show is the difference between disoriented and non-disoriented rats. Disoriented rats have prior representations of head direction at a random orientation, and thus they will settle to one or the other representation 50% of the time (they settle to whichever orientation is closest to the initial one). Non-disoriented rats have prior representations and so will always settle to the correct orientation. These simulations can be interpreted as another example of the resolution issue and are direct consequences of the pWTA nature of the head direction attractor network (see Appendix A).

## 12.2 Simulation details

As before, all neurons were simulated using the three equation model based on the Pinto *et al.* (1996) interpretation of the Wilson and Cowan (1972) model (Eqs. 1.1–1.3).

In this simulation, only distal cues will be included. Thus, the simulations consisted of three neuronal populations, HLS$_{\mathcal{D}}$, E, and I. The HLS$_{\mathcal{D}}$ pool consisted of 36 neurons each associated with a preferred direction $\tilde{\theta}_i$, evenly spaced in a ring around 360°. Each HLS$_{\mathcal{D}}$ neuron was assumed to fire at a constant rate equal to

$$k \cdot \exp(-(\tilde{\phi}_i - 0°)^2/\sigma^2) + k \cdot \exp(-(\tilde{\phi}_i - 180°)^2/\sigma^2) \tag{12.1}$$

The HD population consisted of 36 excitatory neurons evenly spaced in a ring around 360° and one inhibitory interneuron. The excitatory neurons were interconnected with a weight matrix with a Gaussian distribution. Each head direction cell was initialized to

$$k \cdot \exp(-(\tilde{\phi}_i - \Phi_H)^2/\sigma^2) \tag{12.2}$$

In the "non-disoriented" simulations, $\Phi_H$ was always 0°, but in the "disoriented" simulations, $\Phi_H$ was a random value in the range $(0°, 360°]$, yielding a hill of activation at a random location in the head direction ring.

Each HLS$_{\mathcal{D}}$ cell was connected in a one-to-one fashion to its corresponding HD cell. The head direction system itself was thus a one-dimensional attractor, and compass points provided candidate orientations.

These simulations were allowed to settle to a stable state (100 ms at a time step of 1 ms).

## 12.3 Simulation results

Each simulation was allowed to settle from its initial state 50 times and the final orientation was categorized into one of four directions; a histogram of the final orientations is shown in Figure 12.1.

The results show that in the simulation of Cheng (1986) (left of Figure 12.1), the sim-animal orients to the correct direction or its 180° opposite. It very rarely ends up at an intermediate stage. On the other hand, in the simulation of Margules and Gallistel (1988) (right of Figure 12.1), the simulation always orients to the correct direction.

The simulation of Cheng (1986) does not end up at average the two candidate orientations because of the resolution issue discussed in Appendix A: one candidate or the other wins out. In the simulation of the Margules and Gallistel (1988) simulation, the second candidate never wins because the input is too weak to draw away the original representation (which is still supported by the first candidate direction).

| | | |
|---|---|---|
| $\mathrm{HLS}_{\mathcal{D}}$ population | $N_{HLS_{\mathcal{D}}}$ | 36 |
| Gain, $\mathrm{HLS}_{\mathcal{D}}$ | $k_{LV}$ | 0.1 |
| Spread, $\mathrm{HLS}_{\mathcal{D}}$ | $\sigma_{LV}$ | 30° |
| Gain, initial HD | $k_{HD}$ | 0.1 |
| Spread, initial HD | $\sigma_{HD}$ | 30° |
| HD excitatory population | $N_{HD:E}$ | 36 |
| Time constant, HD:E pool | $\tau_{HD:E}$ | 10 ms |
| Tonic inhibition, HD:E pool | $\gamma_{HD:E}$ | -1.5 |
| HD inhibitory population | $N_{HD:I}$ | 1 |
| Time constant, HD:I pool | $\tau_{HD:I}$ | 2 ms |
| Tonic inhibition, HD:I pool | $\gamma_{HD:I}$ | -7.5 |
| Weight, $\mathrm{HLS}_{\mathcal{D}} \rightarrow$ HD:E | $W_{HLS_{\mathcal{D}} \rightarrow HD:E}$ | 1.0 |
| Weight, HD:E $\rightarrow$ HD:E | $W_{HD:E \rightarrow HD:E}$ | $5 \cdot \exp(-(\Delta\phi_{ij})^2/(36°)^2)$ |
| Weight, HD:E $\rightarrow$ HD:I | $W_{HD:E \rightarrow HD:I}$ | 0.44 |
| Weight, HD:I $\rightarrow$ HD:E | $W_{HD:I \rightarrow HD:E}$ | -8 |

Table 12.1: Simulation parameters used in the simulation of Cheng (1986) and Margules and Gallistel (1988).

| | |
|---|---|
| 0.510 | 0.005 |
| 0.000 | 0.485 |

| | |
|---|---|
| 1.00 | 0.00 |
| 0.00 | 0.00 |

Figure 12.1: (left) Simulation of Cheng (1986). (right) Simulation of Margules and Gallistel (1988). Numbers indicate proportion of trials choosing that corner. See text for simulation details.

# Chapter 13

# The water maze

As reviewed in the Experimental Review (Section 2.1.2), the Morris water maze consists of a platform hidden in a pool of water made opaque by milk or chalk. Since the platform is never moved, once the animal knows its own location, then from any point in the environment, it can navigate to that platform.

In Chapter 9, we assumed that the animal was able to solve the water maze without using its route-learning mechanism. In this chapter, I show how it can solve the water maze using the self-localization process described in Chapter 8.

The simulations presented here to demonstrate learning in the water maze are simple: they demonstrate (1) the viability of the hypothesis that exploration learns an association between local view representations and internal representations and (2) that a pWTA process (*self-localization*, Chapter 8) can use this association to settle on a consistent representation of location.

The simulations presented here are from a slightly earlier version of the model and so do not correspond exactly to the simulations presented in Chapter 8. However, the key ideas are the same and the differences are minor.

The simulation used to model self-localization consisted of four pools: HD, HLS$_{\mathcal{L}}$, PC:E, and PC:I. HLS$_{\mathcal{L}}$ represented the local view as described in Chapter 8: distance and allocentric bearing to the wall. I assume that the head direction representation (HD) is not disoriented and represents an accurate (and correct) representation on entry. As before, the neurons were simulated using the Pinto *et al.* (1996) interpretation of the Wilson-Cowan (Wilson and Cowan, 1972) equations (Eqs. 1.1–1.3).

The PC population consisted of 900 excitatory neurons evenly spaced in a 2D neural sheet ($30 \times 30$) and one inhibitory interneuron. The excitatory neurons were interconnected with a weight matrix with a Gaussian distribution.

The place code was thus a two-dimensional attractor, and HLS representation provided candidate locations.

These simulations were allowed to settle to a stable state (100 ms at a time step of 1 ms).

| | | |
|---|---|---|
| Centers, $r$-type sensory neurons | $\tilde{r}$ | {0 cm, 5 cm, 10cm, ..., 100 cm} |
| Gain, $r$-type sensory neurons | $\sigma_r$ | 5 cm |
| Centers, $\phi$-type sensory neurons | $\tilde{\phi}$ | {0°, 10°, ..., 350°} |
| Gain, $\phi$-type sensory neurons | $\sigma_\phi$ | 10° |
| PC excitatory population | $N_{PC:E}$ | 900 (30 × 30) |
| Time constant, PC:E pool | $\tau_{PC:E}$ | 10 ms |
| Tonic inhibition, PC:E pool | $\gamma_{PC:E}$ | -1.5 |
| PC inhibitory population | $N_{PC:I}$ | 1 |
| Time constant, PC:I pool | $\tau_{PC:I}$ | 2 ms |
| Tonic inhibition, PC:I pool | $\gamma_{PC:I}$ | -7.5 |
| Weight, PC:E → PC:E | $W_{PC:E \rightarrow PC:E}$ | $6 \cdot \exp(-(\Delta_{ij})^2/(30\,\mathrm{cm})^2)$ |
| Weight, PC:E → PC:I | $W_{PC:E \rightarrow PC:I}$ | 0.055 |
| Weight, PC:I → PC:E | $W_{PC:I \rightarrow PC:E}$ | -12 |
| Learning rate | $\eta$ | 1.0 |
| Gain | $\tilde{w}$ | 1.0 |
| LTP sharpening constant | $\mu$ | 5 |

Table 13.1: Simulation parameters used in the simulation the Morris water maze.

HLS and PC:E were assumed to be fully connected (i.e. each HLS cell projected to every PC:E cell). The HLS → PC:E weights were trained by allowing a simulation to wander around the environment for 5 simulated minutes (simulated at a time step of 10 ms), during which time both the HLS and PC:E representations were assumed to be correct and the connections were modified by LTP as described by equation 8.1.

The simulations thus consisted of placing the simulated animal (sim-animal) in a random location in the environment, allowing the self-localization process to occur. This initialized the PC:E representation to a random (but coherent) representation of location. Then the sim-animal wandered randomly over the environment for 5 simulated minutes, using the PC:E representation as the initial PI representation. The area covered by the random walk is shown in Figure 13.1. As the sim-animal moved about the environment, simulated LTP occurred between the HLS and PC:E pools.

Then, to test the accuracy of the learned weights, the sim-animal was removed from the pool, placed back at the goal location (engendering a self-localization process) and the position represented by the PC:E pool was measured. The sim-animal was then removed and returned to each of the four cardinal locations. The position represented by the PC:E pool was measured. A direction was calculated by vector subtraction between the two positions represented by PC:E at the current location and at the goal (as determined previously). The sim-animal was then allowed to move in that direction until it reached a wall or was within 5 cm of the goal.

Figure 13.2 shows the result of our simulations.

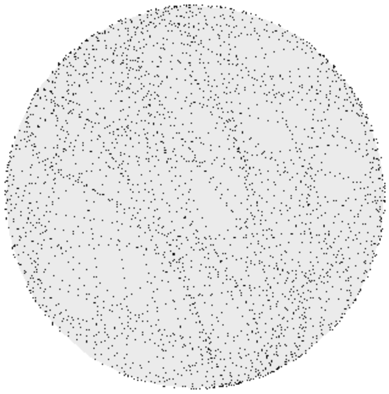


Figure 13.1: Route travelled by simulated rodent while exploring a 100 cm diameter circular environment for 5 minutes. Dots indicate position sampled every 10 seconds. Grey area denotes arena. From Redish and Touretzky (1997c).

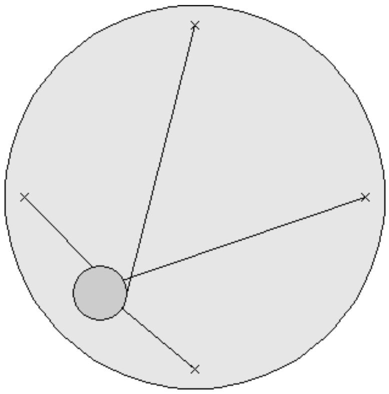


Figure 13.2: Simulated morris water maze. After being allowed to familiarize itself with the entire maze (with no platform), the simulated animal was placed at four locations within the environment. Lines indicate trajectories taken by the simulated animal to reach the goal (indicated by small circle). An $\times$ has been drawn at the four starting locations.

# Chapter 14

# Open field navigation

As reviewed in the Experimental Review (Section 2.1.1 and 2.5), rodents can be trained to find food relative to a set of local landmarks (EX.1–EX.12; EX.122–EX.129). By measuring distributions of time spent searching locations in the environment, one can determine where the animals seem to believe the food is hidden. By changing the landmark arrays, one can gain insights into the navigation system. A theory that purports to explain navigation should be able explain the different search distributions seen in response to manipulations of the landmark arrays. In this chapter, I present simulations that replicate these open-field navigation results.

## 14.1 Prior hypotheses

### 14.1.1 Vector voting (Collett *et al.*, 1986)

To account for the gerbils' behavior in their experiments, Collett *et al.* posited that the animals began planning their path to the goal by applying every learned landmark-to-goal vector (Figure 14.1) to all the landmarks they perceived. The animal tallied the votes, and the best-supported location was taken to be the goal. When there was a tie, multiple locations are searched.

### 14.1.2 Vector averaging (Cheng, 1989)

Cheng (1989) suggested that instead of using a majority vote scheme (as described by Collett *et al.*), animals average the vectors from each landmark. However, this hypothesis is incompatible with the data we collected in the stretched three-landmark task (Saksida *et al.*, 1995, EX.128): the "center of mass" solution (which is the equivalent to vector averaging) lies outside the 95% confidence interval of the search distribution. (The vector voting solution does lie within the 95% confidence interval.)


Figure 14.1: In the vector voting scheme, a vector from each landmark (solid circle) to the goal (small triangle) is learned during training trials (left). In later probe trials (right), these vectors are applied to all the perceived landmarks. The location with the most votes is searched. From Touretzky and Redish (1996).

### 14.1.3 Self-localization by parallel relaxation (Wan *et al.*, 1994a; Touretzky and Redish, 1996; Redish and Touretzky, 1997b)

In previous work, we suggested that animals do not attempt to localize the landmarks relative to their own position. Instead they attempt to determine their position relative to the landmarks. Mathematically, these two algorithms are equivalent, but they require different neural implementations.

When the animals are first placed in the environment, they can see a local view consisting of the landmarks and other cues, but they do not know their positions in this environment. In all three experiments, there was a strong *context shift* that occurred when the animal "entered the environment". In Collett *et al.* (1986), the animal was released from a "start box". In Gothard *et al.* (1996b), a door was lifted and the animal was free to roam the environment.[1] In Saksida *et al.* (1995), the animal was carried by hand from its home cage (in which it spent most of its life) and placed into the open environment.

According to the Touretzky and Redish (1996) theory, these context shifts are enough to trigger a self-localization process, which attempts to satisfy incompatible constraints from the (possibly manipulated) landmark array. In Touretzky and Redish (1996), we simulated this by using a population of "place units" which were tuned to products of Gaussians of local view parameters (such as distance and allocentric bearing to landmarks) and Gaussians of path integrator coordinates (represented mathematically by Cartesian coordinates). The parallel relaxation process described in Touretzky and Redish (1996) began with the place unit tuning to path integrator coordinates being very broad. This made the place units insensitive to path integrator coordinates. At each step the path integrator was set to represent the weighted average of all place unit coordinates. The place unit tuning to path integrator coordinates was then tightened slowly. Eventually, the place

[1] Place cell recordings from Gothard *et al.* (1996b) show place fields still tied to the location of the start box for a distance into the environment. Place fields do not become tied to the goal location until the animal is within 70 cm of the goal (Gothard *et al.*, 1996b). This may imply that the self-localization process does not occur as the animal leaves the start box, but instead occurs later as the animal approaches the goal.

unit representation settled on a coherent representation of location which was compatible with one of the candidate locations. In Touretzky and Redish (1996), we showed that the goal locations implied by these final stable states were compatible with the data reported by Collett *et al.* (1986, EX.1–EX.12).

## 14.2 Self-localization as a pWTA process

In this thesis, I have suggested that the self-localization process occurs as a pseudo-winner-take-all (pWTA) process, realized as one or more sharp waves (see Chapter 8 for details about this process and for comparisons between it and previous theories): The local view information is passed into the place code in hippocampus, which settles to a coherent representation of location. The path integrator is reset based on this coherent place code and then the goal memory plans a trajectory to reach the goal.

## 14.3 Simulation details.

The simulations presented here were identical to those presented in the previous chapter as simulations of the water maze (Chapter 13), with the exception that instead of being sensitive to the arena wall, the HLS cells were sensitive to the distance and allocentric bearing to the components of the landmark array (white cylinders in the experiments modeled here). In tasks which included more than one cylinder, the firing rate of each HLS cell was the maximum of the tuning to all landmarks.

### 14.3.1 Measuring accuracy

In order to measure the accuracy of the self-localization simulations, the simulation was allowed to explore the environment for 5 simulated minutes. Then the weight matrix learned from that exploration was used for 100 *entries*. Each entry consisted of placing the simulated animal at a random location in the environment, allowing the self-localization process to settle, and then measuring the represented location in the PC:E population. By vector subtracting the current representation from the previously measured representation of the goal location, an estimate of the position of the goal relative to the sim-animal's location (a *goal prediction*) was determined. A histogram of the 100 goal predictions was compared to the data from Collett *et al.* (1986, EX.1–EX.12) and our lab (EX.122–EX.129).

### 14.3.2 The resolution issue

An important difference between the vector voting and pWTA implementation of the self-localization processes is that, in the pWTA self-localization process, if two candidate locations are within the breadth of the hill of activity, they will be averaged together. The breadth of the hill is equal to the width of the typical place field. (See Appendix A for a discussion of this issue.)

Due to computational limitations, our three-landmark simulations (Figures 14.6–14.11) had to be done with just 900 cells (30 × 30) in our model hippocampus. Because of this, we observed an averaging effect in two cases in which the real animal does not show averaging. (See Figures 14.9 and 14.11.) To prevent this, I used a one-dimensional model with 90 cells for the two-landmark simulations (Figures 14.3–14.5).

## 14.4 Simulation results

### 14.4.1 One landmark experiments.

Collett *et al.* (1986) trained gerbils to find food placed near a single circularly-symmetric landmark (EX.1). The fact that the animals could learn to search at the correct bearing as well as distance implies that they have some independent means of determining bearing information. Collett *et al.* (1986) supposed that the animals were utilizing some external cue, despite their attempts to prevent this. The head direction subsystem (see Navigation Review, Section 3.2.2 and Chapter 5), once reset from the orientation of the start box (which did not change from trial to trial), would provide an internal cue that would allow the local view to represent allocentric bearing, which, combined with distance, is sufficient to localize the animal in the one-landmark case.

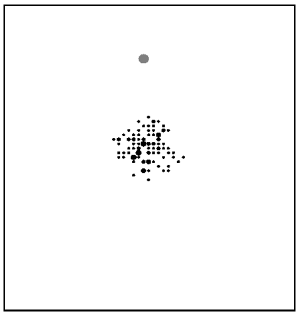


Figure 14.2: One landmark: Histogram of goal predictions by the simulation. Sold circle indicates landmark position. Size of blob indicates number of goal predictions at each location. Compare Figure 2.2.

### 14.4.2 Two landmark experiments.

Collett *et al.* also trained animals to find food at a location specified by a pair of landmarks. Well-trained animals searched at a single location, with the goal position well within the

95% confidence interval for the single Gaussian (EX.3; EX.122). Simulations also show a single peak, see Figure 14.3.

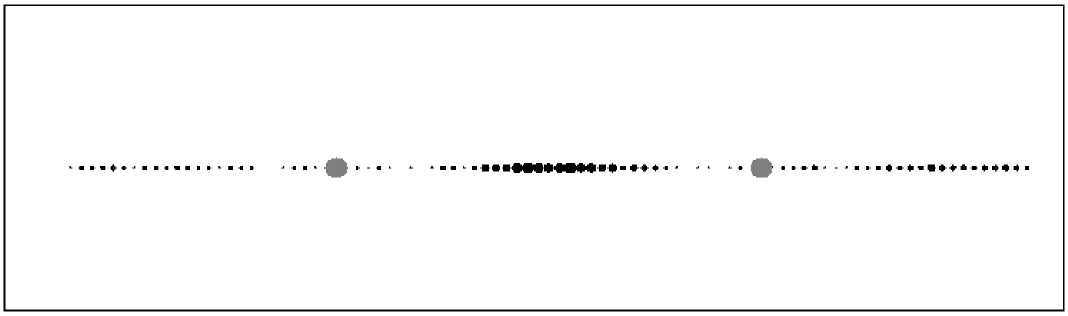


Figure 14.3: Two landmarks. Histogram of goal predictions by the simulation. Solid circles show landmark locations; size of blob indicates number of goal predictions at each location. Compare Figures 2.4 and 2.27.

With two landmarks, the array can be manipulated to provide ambiguous or inconsistent cues. For example one landmark can be removed. In this case, Collett *et al.* report that the gerbils searched alternately in two locations, each at the correct distance and bearing from one of the landmarks they had observed during training, EX.4. We found that the search distribution is best modeled by two Gaussians each accounting for 50% of the data, with the expected locations well within the 95% confidence intervals (EX.123). Our simulations also produce two peaks because the local view is insufficient to constrain the place code to a single location (see Figure 14.4). This produces an unstable state in the simulated place code. When the self-localization process ends, the system must end in one stable state or the other, it cannot remain in a superposition of states. In a sense, this forces the animal to consider the landmark as corresponding to one or the other of the remembered landmarks, not both.

The two-landmark array can also be stretched, see Figure 14.5. When trained with two landmarks and tested with the distance between them doubled, the gerbils again searched at two locations, each at the correct distance and bearing from the corresponding landmark (EX.5; EX.124). Our simulations also show this effect due to the same instability described for the two-minus-one case, above.

### 14.4.3 Three landmark experiments.

Both Collett *et al.* and we have also tested manipulations of a third configuration: reward placed at the center of an equilateral triangular array. As with the previous two configurations, well-trained animals searched for food at the center of the array (EX.7; EX.125). The analysis shows that this search distribution is best explained by a single Gaussian centered over the goal. Our simulations show a similar distribution (Figure 14.6).

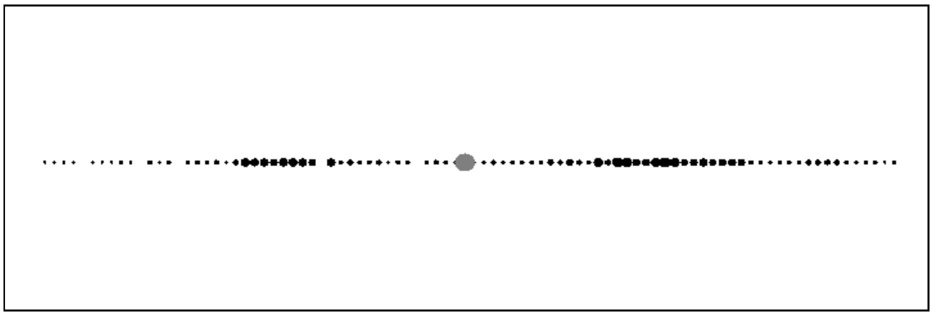


Figure 14.4: Tested with one landmark, after being trained with two. Solid circles show landmark locations; size of blob indicates number of goal predictions at each location. Compare Figures 2.4 and 2.27.

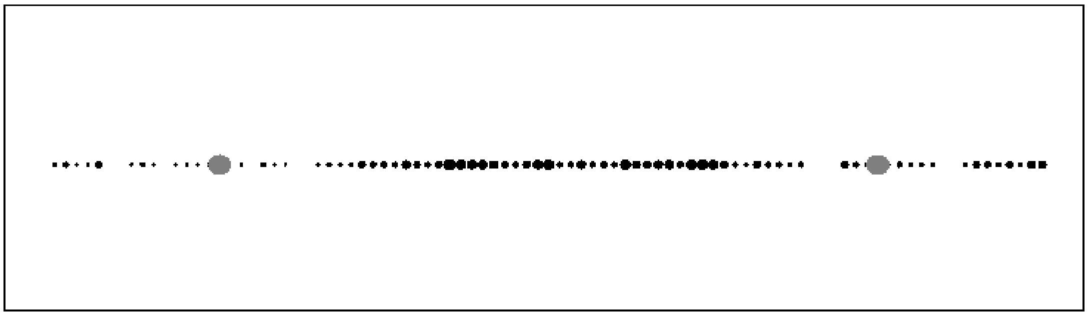


Figure 14.5: Tested with two landmarks more separated than during training. Histogram of goal predictions by the simulation. Solid circles show landmark locations; size of blob indicates number of goal predictions at each location. Compare Figures 2.3 and 2.26.

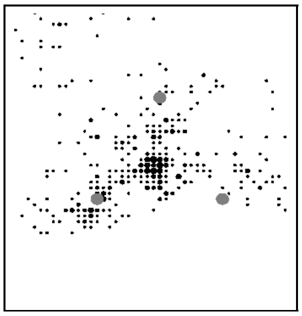


Figure 14.6: Three landmark experiment. Histogram of goal predictions by the simulation. Solid circles show landmark locations; size of blob indicates number of goal predictions at each location. Compare Figures 2.6 and 2.29.

As with the two-landmark array, manipulations of an array of three landmarks are also possible. We will describe five. First, one can add a landmark to form a second triangle with opposite orientation. Gerbils search at the center of the correct orientation (EX.8). Our simulations show a similar effect (see Figure 14.7).

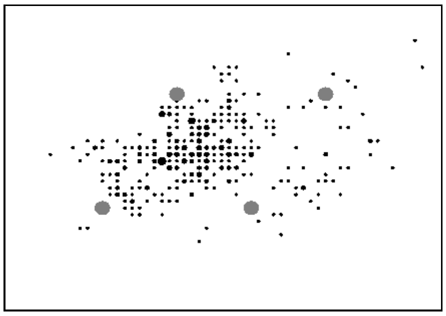


Figure 14.7: Tested with four landmarks, after being trained with three. Histogram of goal predictions by the simulation. Solid circles show landmark locations; size of blob indicates number of goal predictions at each location. Compare Figure 2.7.

Second, one can remove one or more landmarks. In our experiments, removing one landmark produced search on both sides of the line formed by the remaining landmarks (EX.126). Collett *et al.* reports that the animals search on only one side of the landmarks (EX.9), however, re-examination of their figure (reprinted as Figure 2.8) shows that there is some search on the opposite side of the line. Our simulations show a preferred search on the correctly-oriented side, but they also show a tendency towards the midline (Figure 14.8).

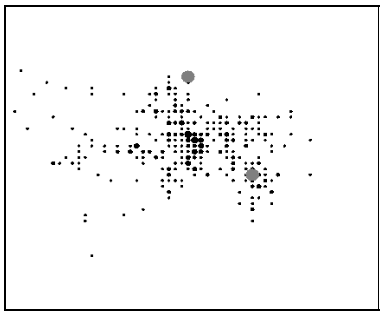


Figure 14.8: Tested with two landmarks after being trained with three. Histogram of goal predictions by the simulation. Solid circles show landmark locations; size of blob indicates number of goal predictions at each location. Compare Figures 2.8 and 2.30.

Collett *et al.* report that when trained with three landmarks and tested with one, gerbils searched at three locations corresponding to matching the landmark to each of the remembered landmarks in turn (EX.10). We were unable to replicate this result in our lab (EX.127), nor do our new simulations demonstrate this (Figure 14.9, left). The simulations using the neuronal model detailed above do not show the separated goal predictions as a consequence of the resolution issue also discussed above. In a more abstract simulation which does not suffer from these computational limitations (and thus included a higher resolution), we did observe the separation of goal predictions (Touretzky and Redish, 1996; Redish and Touretzky, 1997b), see Figure 14.9, right.

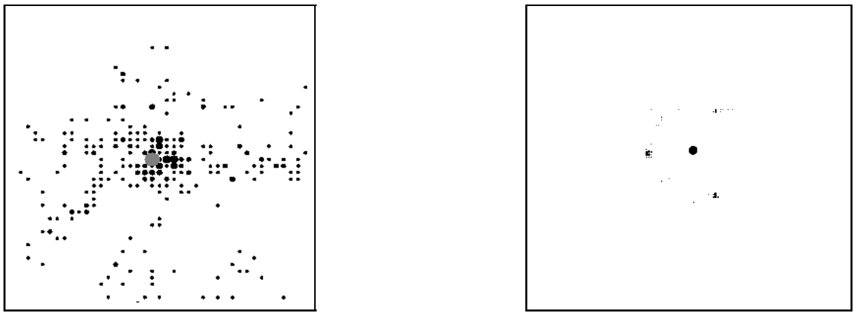


Figure 14.9: Tested with one landmarks after being trained with three. (left) Histogram of goal predictions by the simulation. (right) Histogram of goal predictions by earlier abstract simulation from Touretzky and Redish (1996). Solid circles show landmark locations; size of blob indicates number of goal predictions at each location. Compare Figures 2.9 and 2.31.

In addition to adding or removing landmarks, one can manipulate the positions of the three landmarks relative to each other. For example, one can stretch the triangle by moving one of the landmarks away from the other two, while leaving the distance between the other two unchanged. Gerbils searched at the correct distance from the two landmarks with the correct inter-landmark distance, but too far from the displaced landmark. We can use the analysis in Section 2.5.2-D to disprove the *vector averaging* hypothesis about the gerbils' navigation strategy: this point lies outside the 95% confidence interval of the search distribution. Thus we conclude that the animal *is* searching closer to the two "unchanged" landmarks than to the third "stretched" one. Our simulations show a similar effect, see Figure 14.10.

Finally, the triangle can be inverted (or equivalently, rotated by 60°). When this is done, Collett *et al.* report that the gerbils first search the center of the array and then proceed to search three exterior points (EX.12). Because we did not observe digging throughout the manipulated experiments, we cannot say that the gerbils first search the center and then the exterior points (EX.129).

Our current simulations show goal predictions at the center because of the resolution issue discussed above. The three external goal predictions are averaged together because

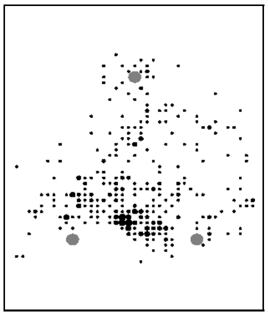


Figure 14.10: Training with three landmarks, testing with one displaced. Compare Figures 2.10 and 2.32.

the resolution of these simulations is too coarse. The wings shown in the left of Figure 14.11 are a consequence of the external points occasionally drawing the goal prediction toward itself.

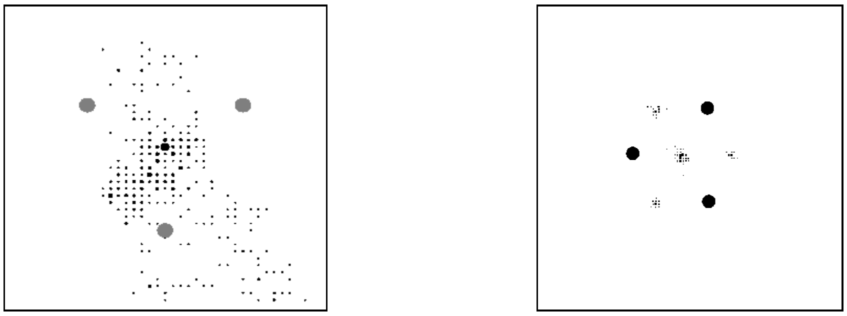


Figure 14.11: Trained with three landmarks, tested with triangle rotated by 180°. (left) Histogram of goal predictions by the simulation. (right) Histogram of goal predictions by earlier abstract simulation from Touretzky and Redish (1996). Solid circles show landmark locations; size of blob indicates number of goal predictions at each location. Compare Figures 2.11 and 2.33.

Our previous simulations (Figure 14.11, bottom right) produced search at the center and exterior points for different reasons. They searched at the center because the head direction representation occasionally flips by 180° in order to better match the rotated triangle. When this happens the task reverts to the original training configuration and the gerbil searches at the center of the triangle. On the other hand, when the head direction does not change (presumably due to inertia in the system or to the distal cues which have not moved), the simulated gerbils search at the exterior locations because they match two of the three landmarks. A prediction one can draw from these previous simulations is that when the animal searches in the center, the head direction representation should flip. This could be detected by recording from head direction cells as the animal performs this task.

# Part IV

# Conclusions and appendices

# Chapter 15

# Conclusions

## 15.1 Summary of the thesis

This thesis pulled together a comprehensive theory of rodent navigation from the extensive experimental and theoretical work done by researchers in the field over the past century, and examined specific components of that comprehensive theory (including testing aspects of them with computer simulations). By looking at a large corpus of experiments across a variety of paradigms (see Experimental Review, Chapter 2), I was able to bring the theoretical literature into a consistent picture (rejecting some theories, accepting some, and modifying others, see Navigation Overview (Chapter 3) and Hippocampal Theoretical Review (Chapter 4)).

However, a number of aspects of the theory must be left unresolved at this time. For example, there are two possible mechanisms which can separate reference frames in novel environments: orthogonalization in the dentate gyrus, and pre-wired maps in CA3 (see Section 3.3). Throughout this thesis, I have suggested that the DG hypothesis is more likely than the CA3 hypothesis. However, there is no data disproving the CA3 hypothesis, and the two hypotheses are not incompatible. Whether one, both, or neither mechanism occurs in the rodent brain will have to be left for future work.

Two other examples of unresolved issues include the mechanism which produces phase precession in the hippocampus and how the head direction system is reset from local cues. A number of proposals have been made about phase precession (O'Keefe and Recce, 1993; Burgess *et al.*, 1994; Samsonovich and McNaughton, 1996; Tsodyks *et al.*, 1996), but none of them fit all of the available data (see Section 3.4). In this thesis, I assumed that phase precession existed and simulated some consequences of it, but did not include a neurally plausible mechanism for generating it.

The second issue is the resetting the head direction representation from local cues. Resetting the head direction from external cues is straight-forward (McNaughton *et al.*, 1991;

Knierim *et al.*, 1995, see Section 3.2.2), but resetting head direction from local cues requires information about the location of the animal (presumably in the place code). Previous models of this have only included very abstract mechanisms (Wan *et al.*, 1994b; Touretzky and Redish, 1996; Redish and Touretzky, 1996b). As discussed in Appendix B, the evidence that rodents can actually reset head direction from local cues is not as robust as evidence that they can reset head direction from distal cues.

Even with some specifics unanswered, by looking at the big picture and examining interaction effects between systems, it is possible to draw novel conclusions. One important result of the work in this thesis is the hypothesis that although the hippocampus is strongly involved in navigation, it is not necessary for online navigation. It is only necessary for self-localization after significant context switches and in replaying routes during sleep to allow for memory consolidation.

This follows from the hypothesis that the path integrator is extra-hippocampal. We hypothesized that the path integrator was likely be extra-hippocampal because the topology of place fields changes between environments. This in turn suggests that the hippocampus should not be involved in online navigation tasks (tasks that do not require a context-switch). This prediction[1] has recently been confirmed by Alyan *et al.* (1997), who showed that rats continue to show path integration abilities during online navigation even after cytotoxic hippocampal lesions (EX.59).

In addition to synthesizing a comprehensive theory, this thesis examined (including simulations) a number of specific components of the theory: the head direction system (Chapter 5), the path integration system (Chapter 6), and the role of the hippocampus (Chapters 7–9). Finally, to demonstrate that the theory is truly comprehensive, I showed simulations of specific experimental results in the literature: the bimodality in the representations of an environment by old animals but not young (Barnes *et al.*, 1997, Chapter 10), the interaction between entry point and cue cards (Sharp *et al.*, 1990, Chapter 11), the differential effects of disorientation or non-disorientation (Cheng, 1986, and Margules and Gallistel, 1988, Chapter 12), self-localization in the Morris water maze (Morris, 1981, Chapter 13), and effects of manipulating the landmark array on search by gerbils (Collett *et al.*, 1986, additional experiments in Section 2.5, simulations in Chapter 14).

## 15.2 Modeling

Although it would be best to simulate the entire theory with a single computer program, it is not possible to do so at this time. Instead of a single unified simulation, I presented a family of simulations, all using the same neural model, with explicit assumptions identified for each component.

Each simulation examines one or more specific aspects of the theory. This allows each simulation to stand or fall on its own merits rather than being critically dependent on

[1] The prediction was originally published in Redish and Touretzky (1997a).

the other components. For example, the explanation for the Barnes *et al.* (1997) result (Chapter 10, see also EX.83) does not depend on a specific path integrator mechanism, only on the existence of an external path integrator that provides input into the dentate gyrus. This means that even if the mechanism or anatomy for path integration presented in Chapter 6 is wrong, the simulation in Chapter 10 may still be correct (if its assumptions are still met).

The other reason for presenting a family of simulations rather than a single unified program is computational. Many of these specialized simulations require large numbers of neurons and connections (thousands of neurons and perhaps a million connections), simulated at very small time steps (say 1 ms) for long time periods (e.g. 5–15 minutes of simulated time). Given the computational resources available today, building, debugging, and running a unified simulation able to replicate all of the results in Parts II and III is impossible. In addition, the simulations already require a large number of parameters, many of which do not have direct neural analogues that can be measured. Many of these parameters have highly non-linear effects (see, for example, the effect of external input discussed in Appendix A). A successful parameterization of a unified simulation would be much more difficult to find and interpret.

## 15.3 Major contributions

The major contributions of the thesis are the synthesis of the comprehensive theory presented in Part I plus a number of more specific contributions to the field of rodent navigation:

- By separating out the reintegration component of the head direction system (placing it anatomically in the ATN), we made a prediction that anterior thalamic head direction cells should deform their tuning curves during rotations. This prediction[2] has been recently confirmed by Blair *et al.* (1997).

- I presented a novel hypothesis for the anatomical location of the path integrator (subiculum-parasubiculum-superficial entorhinal cortex). This hypothesis predicted that the hippocampus would not be involved in path integration, which has been recently confirmed by Alyan *et al.* (1997).

- I presented a novel understanding of the role of the hippocampus in normal navigation: that it plays the role of *self-localization* during normal navigation. If this theory is true then the hippocampus should not be necessary for online navigation.

- I demonstrated through simulation that even if the hippocampus also plays a role in memory consolidation (route-replay), this is not incompatible with the self-localization role hypothesized. I also demonstrated that these two roles can occur by the same

[2] This prediction was originally published in Redish *et al.* (1996).

mechanism, depending only on whether candidate locations are available to the system or not.

- The concept of *reference frames*[3] unifies a diverse literature of non-spatial correlations of hippocampal place cells.

- I demonstrated the viability of the orthogonalizing-DG hypothesis of map separation. Although I cannot disprove the pre-wired-CA3 hypothesis, I showed that it is not necessary for certain experiments which had been previously taken as strong support for it.

## 15.4 Where to from here?

Although this thesis brought together a comprehensive theory of rodent navigation, it leaves a number of open questions. I have listed the major issues in Appendix B, but will note some important issues here.

There are both general aspects which require essentially exploratory experimental work before major theoretical work can be applied, and specific aspects which are ready for theoretical and simulation work. Two important examples of the former are:

- What is the representation of local view in the rodent? Important work has been done in primate (see Andersen *et al.*, 1993, and Colby *et al.*, 1995, for reviews). However, only preliminary work has been done in the rodent (McNaughton *et al.*, 1994b; Chen *et al.*, 1994a; Chen *et al.*, 1994b). An important consideration is that the primate parietal cortex consists of a host of subareas (Colby and Duhamel, 1991; Stein, 1992). Preliminary anatomical evidence suggests that this is also true of the rodent (Zilles, 1990).

- How do rodents plan complex routes and avoid obstacles? Animals are generally very successful at planning complex trajectories around obstacles, but it is not currently known how this planning occurs in the rodent brain. It is not even known how obstacles are represented in the rodent brain. Until the neurophysiology is done to find the neural correlates of obstacles, specific theoretical work understanding the representation is very difficult.

Two examples of open issues that are ready for theoretical or simulation work are:

- Can the path integration mechanism in Chapter 6 (a variation of that suggested by Samsonovich and McNaughton, 1997) be made to track actual trajectories of real animals? How accurate can it be? What is required to make it accurate? By simulating

[3] In Section 3.3, I discussed the relation between the multi-map concepts (*multiple maps in the hippocampus* (O'Keefe and Nadel, 1978), *active subsets* (Muller and Kubie, 1987), *reference frames* (Wan *et al.*, 1994b; Wan *et al.*, 1994c; Touretzky and Redish, 1996; Redish and Touretzky, 1997a), *charts* (McNaughton *et al.*, 1996; Samsonovich and McNaughton, 1997; Samsonovich, 1997))

the head direction system accurately, we found complexities (for example the deformation of the tuning curve) that were not obvious from the theoretical model. No one has yet reported an accurate simulation of two-dimensional path integration based on the moving hill hypothesis.

- What is the effect of changing the ratio of signal-to-noise in the boundary conditions for the pWTA mechanism? I have shown that the two extremes show qualitatively different modes and can solve qualitatively different navigational problems (self-localization and route-replay). However, it is still unknown how a pWTA network behaves as the sensory input signal is increased from 0 (which produces replay) to a strong signal (which produces recall). Is there partial replay with weak signals? Is there a point at which the system shows a phase change and suddenly shows strong recall? The complex non-linearities of pWTA networks (see, for example, Appendix A) make an analytic solution extremely difficult. But simulations would be straight-forward and might show very interesting effects.

In addition, there are a number of strong predictions one can take from the specific hypotheses presented in this thesis. Two examples are

- If it is true that the self-localization process occurs during awake LIA and is a pWTA process, then it should be possible to observe the self-localization process neurophysiologically. First, one needs to determine the place fields of each cell. This can be found by recording during an experience within an environment. Second, one needs to record the spikes of a population simultaneously during a self-localization process. The position represented by the population can be displayed at each moment by plotting the weighted sum of place fields (i.e. the sum of all fields of all cells that fired a spike at a moment in time). The representation should begin as a very broad (essentially uniform) distribution and then tighten up over the course of the self-localization process. In the end, it should be a coherent representation of the animal's position.

- The hypothesized self-localization process exhibits averaging of nearby candidate locations, but competition among distant candidate locations. This can be tested behaviorally by examining the search time of animals (such as gerbils) in a two-landmark task. If the animals are trained to find food at the center of a pair of landmarks and then tested with the landmarks at varied distances, the distance between candidate locations can be experimentally controlled. At large distances, the animal should alternate between two goal locations, but at small distances, the animal should search once in the center of the pair. Gerbils have been shown to search at two goal locations when faced with a stretched pair of landmarks (EX.5; EX.124), and at the center of the pair when faced with the same pair as training (EX.3; EX.122). But where this transition occurs has not been tested.

Finally, this thesis offers an example of one way that systems-level theoretical neuroscience might be done. It took a single, well-defined domain and built up a comprehensive theory based on experimental and theoretical work done by many other researchers. I believe that there are a number of other domains in systems-level neuroscience that are ripe for this kind of synthesis. These domains must be (1) well-defined, (2) have extensive experimental results from a variety of paradigms, and (3) have extensive theoretical work on components of the system. Two examples that fit these criteria are procedural learning in the basal ganglia (see Houk *et al.*, 1995, for a starting point) and the representation of egocentric space in primate parietal cortex (see Colby and Duhamel, 1991; Stein, 1992; Andersen *et al.*, 1993; Pouget and Sejnowski, 1995). I hope that this thesis will encourage and inspire others to attempt this kind of synthesis in other systems-level neuroscience domains.

# Appendix A

# Attractor networks

In this appendix, I review some important aspects of *attractor networks*. These networks play important roles in the various components described in this thesis. There are two important cases: one- and two-dimensions. In the one-dimensional case, I will assume that the cells are located along a ring (as head direction cells are, see Section 3.2.2 and Chapter 5). In the two-dimensional case, cells can be thought of as being located around a torus (see discussions of path integration, Section 3.2.3 and Chapter 6). In this appendix, I will only review the one-dimensional case. The two-dimensional analogy is straightforward.

If a population of head direction cells is interconnected such that cells with nearby preferred directions are more strongly interconnected (i.e. have a higher synaptic weight) than cells with distant preferred directions and there is global inhibition, then this system will have dynamics such that a stable state is a *hill* of activation. This local-excitation, global-inhibition connection structure has been studied by a number of researchers (in one-dimension: Wilson and Cowan, 1973; Amari, 1977; Ermentrout and Cowan, 1979; Kishimoto and Amari, 1979; Kohonen, 1982; Kohonen, 1984; Skaggs *et al.*, 1995; Elga *et al.*, 1996; Redish *et al.*, 1996; Zhang, 1996a; in two-dimensions: Kohonen, 1982; Kohonen, 1984; Droulez and Berthoz, 1991; Munoz *et al.*, 1991; Arai *et al.*, 1994; McNaughton *et al.*, 1996; Zhang, 1996a; Samsonovich and McNaughton, 1997; Samsonovich, 1997; Redish and Touretzky, 1997c).

Note that any direction is a stable attractor state and a coherent representation of direction. That is, the ring in Figure 3.5 can be rotated to any direction. There is no "global north." There is no correspondence to north or south at all. However, the system knows that when the population peak is at the top of the ring, the animal is facing 180° from when the population peak is at the bottom of the ring. The system can represent any direction and differentiate the different directions, but the correspondence between represented direction and environment is not hard-wired, it must be inferred by observation.

## A.1 Simulation details

The simulations for the attractor networks below consisted of 100 neurons connected with a local excitation function and a single inhibitory interneuron. All neurons (both excitatory and inhibitory) were based on the Pinto *et al.* (1996) interpretation of the Wilson and Cowan (1972) equations (Eqs. 1.1–1.3).

The 100 neurons (the E pool) were assigned preferred directions evenly distributed through 360°. The weight between them was proportional to a Gaussian function of the difference between their preferred directions:

$$w_{ij} = W_0 \cdot \exp(-(\Delta\phi_{ij})^2/\sigma^2) \tag{A.1}$$

The weight between the excitatory neurons and the single inhibitory neuron was a constant, as was the weight between the inhibitory neuron and the excitatory neurons.

All simulations were done with a time step of 1 ms. Parameters used are shown in Table A.1.

| | | |
|---|---|---|
| Number of excitatory neurons | $N_E$ | 100 |
| Excitatory time constant | $\tau_E$ | 10 ms |
| Excitatory tonic inhibition | $\gamma_E$ | -1.5 |
| Number of inhibitory neurons | $N_I$ | 1 |
| Inhibitory time constant | $\tau_I$ | 2 ms |
| Inhibitory tonic inhibition | $\gamma_I$ | -7.5 |
| Base weight, E →E | $W_0$ | 6 |
| Spread, E →E | $\sigma$ | 5 neurons |
| Weight, E →I | $W_{E\to I}$ | 0.5 |
| Weight, I →E | $W_{I\to E}$ | -8 |
| Weight, I →I | $W_{I\to I}$ | -4 |

Table A.1: Parameters used in the one-dimensional attractor network simulations.

## A.2 Simulation results

Figure A.1 shows an example of one of these attractor networks returning to a coherent representation of head direction after being initialized to a random state (which is not a coherent representation of head direction). Within 100 ms, the network has transformed the activity from an incoherent to a coherent state. The observation that a local-excitation/global-inhibition connection matrix produces a stable bump of activation can be traced to early work by Wilson and Cowan (1973), Amari (1977), and Kohonen (1982;

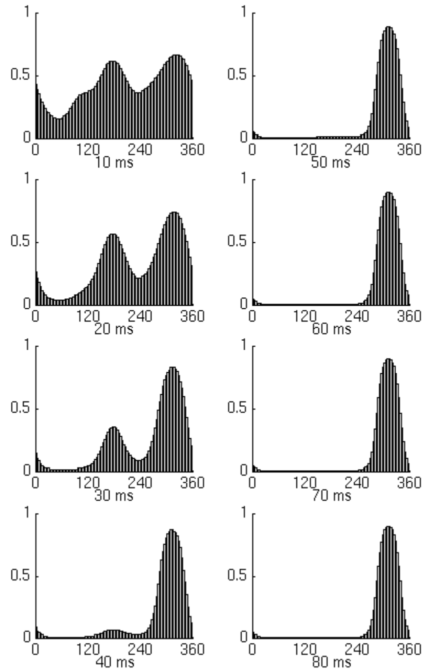



Figure A.1: A coherent representation of head direction forms from noise. Each figure shows the population activity of the E pool in a simulation of a one-dimensional attractor. The activity of each neuron at each moment in time is plotted at its preferred direction. Similar simulations have been reported by Kohonen (1982; 1984), Zhang (1996a), Elga *et al.* (1996), and Samsonovich (1997).

1984). The boundaryless ring condition was first suggested by Skaggs *et al.* (1995) and first simulated by Zhang (1996a).

## A.3 External excitatory input

What is the effect of excitatory input onto this attractor network? (I assume input only synapses on excitatory neurons.) There are four possible cases, depending on the location and magnitude of the extra-population input.

1. If an attractor network is in a stable state and receives input (synapsing on excitatory cells) that is peaked at the same direction as is currently being represented, then

nothing will change. The attractor network will still be in a stable state representing the same direction. The overall activity in the attractor network may increase slightly, but the represented direction will not change.

2. On the other hand, if the input is offset slightly, then the attractor network will precess until the new representation is centered at the input direction (Skaggs *et al.*, 1995; Elga *et al.*, 1996; Redish *et al.*, 1996; Zhang, 1996a; Samsonovich and McNaughton, 1997). Figure A.2 shows an attractor network with offset input. It quickly precesses from its current representation to a new representation compatible with the input.

3. If the input is offset by a large amount but is too small, it will not be strong enough to affect the current representation and the current representation will not change. Figure A.3 shows an attractor network with offset input that does not change the current representation.

4. Finally, if strong enough input is offset by a large amount, the representation of the current direction will disappear and the activity will reappear at the offset location (Zhang, 1996a; Samsonovich and McNaughton, 1997). See Figure A.4.

## A.4 The resolution issue

The two cases shown in Figures A.2 and A.3 used the same input strengths. When weak excitatory drive is input near the peak of a coherent representation, the peak precesses to align itself with the offset drive. But if weak excitatory drive is offset substantially from the peak, then the peak does not move.

This makes a prediction in cue conflict situations where external cues conflict with internal representations. If the external cues suggest a candidate location near the currently represented location, the representation will shift to match the cues. If all candidates are distant from the current representation, the representation will not change. Near and distant can be quantitatively defined as a function of the breadth of the tuning curve of the component cells.

Figure A.5 demonstrates this effect. Simulations were identical to the those described previously in this chapter. Note that the effect of the external input is linear out to approximately $\pm 60°$, but vanishes very quickly beyond that.

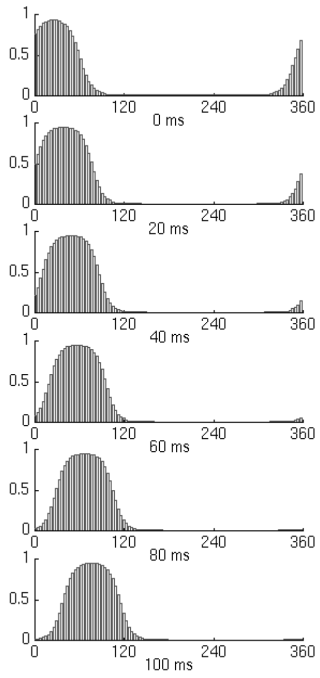



Figure A.2: An attractor network precessing from offset input. A small excitation (weight $= 2.0$, offset $= +20°$) is input for 100 ms slightly to the right of the center of the bump. Although this effect is small, continuously offset input is sufficient to track head direction accurately. Compare Figure A.4. Similar simulations have been reported by Zhang (1996a), Elga *et al.* (1996), and Samsonovich and McNaughton (1997).

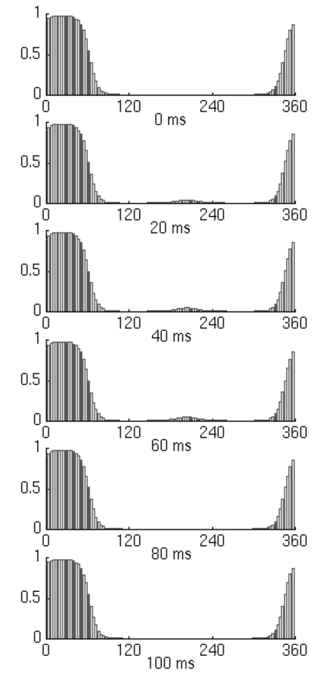



Figure A.3: An attractor network unaffected from offset input. A small excitation (weight $= 2.0$, offset $= +180°$) is input for 50 ms $180°$ off the center of the bump. Compare Figure A.4.

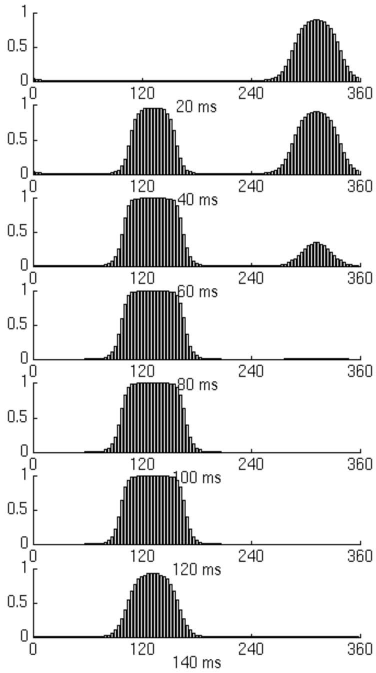



Figure A.4: Attractor network with tonic excitatory input. A sufficiently large excitation (weight $= 3.0$, offset $= +180°$) is input for 100 ms at another representation of head direction. The head direction jumps from the old representation to the new one. Compare Figure A.2. Similar simulations have been reported by Zhang (1996a) and Samsonovich and McNaughton (1997).

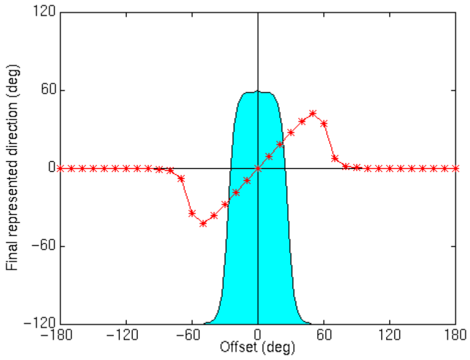



Figure A.5: The *'s indicate the final represented direction after excitatory drive was input offset to the peak by the listed angle. The bump shows the width of a typical "head direction" tuning curve in these simulations.

# Appendix B

# Open questions

In this appendix, I lay out what I consider to be some of the most important open questions in rodent navigation. Both sides of the debate were presented when each issue arose in the Navigation Overview (Chapter 3). Where possible, I suggest experiments that might be able to help resolve the debate.

Some of the open questions require extensive exploratory experimental and theoretical work. I refer to these as general issues, because answering them will take a major research effort. Other open questions relate more directly to specific components discussed in the thesis. Often these questions can be answered by a few well-designed experiments. In addition, there are predictions that can be made following hypotheses proposed in this thesis. In the following list, general issues are marked with ‡, specific issues with †, and predictions of hypotheses proposed here with $*$.

## B.1 Local View

‡ **How is the local view actually represented?** In Section 3.2.1, I identified some of the key variables that have to be represented in the local view (e.g. distance, egocentric and allocentric bearing) and some of the cortical areas that seem to be involved (e.g. LDN, posterior parietal cortex). Exactly how those variables are represented, however, is still an unresolved question. Tremendous work has been done in the primate (see Andersen *et al.*, 1993, and Colby *et al.*, 1995, for reviews), but I know of little equivalent work in the rodent. Additionally, in the primate, the parietal cortex consists of a host of subareas (Colby and Duhamel, 1991; Stein, 1992). Preliminary anatomical evidence suggests that this is also true in the rodent (Zilles, 1990), but no equivalent neurophysiological experiments have been done which could separate the areas functionally, as has been done in the primate.

‡ **Does the rodent visual system also separate into two streams?** There is extensive evidence suggesting that the primate visual system consists of two dissociable streams, one supporting spatial aspects (the dorsal stream, passing through parietal cortex), and one supporting identity aspects (the ventral stream, passing through inferotemporal cortex) (see Mishkin *et al.*, 1983; Ungerleider and Haxby, 1994; Goodale *et al.*, 1994 for reviews). There is some preliminary evidence that this is also true of the rodent (Kolb, 1990a; Kolb *et al.*, 1994), but more work could be done on this issue. An interesting corollary to this question is whether non-visual aspects also divide into these two streams.

## B.2 Head direction

∗ **Does the system include attractor dynamics?** A key component of the head direction system described in Section 3.2.2 and Chapter 5 is the attractor dynamics hypothesized to occur as a consequence of the pWTA network structure. This can be explicitly tested by multi-unit recording. If one records from multiple head direction cells simultaneously (say 20 cells), then the population encoding can be seen directly. If noise is injected into the system by microstimulation, then the representation should be transiently disrupted. Since nearby postsubicular cells encode different directions (Sharp, 1996a), microstimulation should effectively add noise into the system by firing a set of cells with unrelated preferred directions. If the system includes attractor dynamics, then even in darkness, without external or self-motion cues, the population's firing rates should return to a well-formed representation of head direction.

† **Is the head direction representation maintained in the absence of angular velocity input?** Stackman and Taube (1997) lesioned the vestibular system using intratympanic injections of sodium arsanilate (EX.107). After the surgery, Stackman and Taube could not find any head direction cells in the ATN. However, there are two possible explanations for this: (1) the population activity may have been a valid representation of head direction, but simply uncoupled from motor action, or (2) the representation itself may have been disrupted. This question is important because it bears on the attractor nature of the system: is the neural mechanism underlying head direction specifically constructed so as to produce a coherent representation of head direction, or is the coherent representation an epiphenomenon occurring as a consequence of the cells showing head direction tuning curves? The single-unit recording procedure used by Stackman and Taube could not differentiate these two possibilities, but multi-unit recording could.

† **How can the theory explain the data that postsubicular lesions do not affect ATN head direction tuning curves?** Postsubicular lesions do not disrupt ATN head direction cells (Goodridge and Taube, 1994, see also Taube *et al.*, 1996; EX.108). The theory in Section 3.2.2 and Chapter 5 predicts that postsubicular lesions should devastate ATN head direction correlations. As discussed in Section 5.7.1, there are three possible

explanations: (1) the postsubiculum may be part of a larger structure, (2) other areas may be able to play similar roles to that played by the postsubiculum, (3) there may be problems with the theory (or the experiment). Additional theoretical and experimental work is needed to bring the theory into alignment with this data.

‡ **Can the head direction system be reset from local cues? If so, how does this happen?** Head direction can be reset from distal cues directly because distal cues do not change their allocentric bearing over the environment (Gallistel, 1990; McNaughton *et al.*, 1991; Skaggs *et al.*, 1995). However, in order to reset head direction from local cues, the system requires a representation of the animal's position (e.g. the place code) (Wan *et al.*, 1994c; Redish and Touretzky, 1996b; Touretzky and Redish, 1996). Neural simulations of how this reset might occur have not been reported although we (Wan *et al.*, 1994c; Redish and Touretzky, 1996b; Touretzky and Redish, 1996) have previously simulated this reset process at an abstract level.

In this thesis, I have assumed that head direction reset occurs by distal cues only, including assuming that the cue card is a distal cue. Evidence that changing the width of the cue card does not change place fields (Muller and Kubie, 1987), that removing the cue card only changes the orientation of the place field relative to the center of the environment (Muller and Kubie, 1987; Muller *et al.*, 1991a; Quirk *et al.*, 1992), and that the head direction system can be reset by a cue card even after hippocampal lesions (Golob and Taube, 1997) all support this assumption. Collett *et al.* (1986) do report examples in which gerbils were trained to find food relative to a rotating pair of landmarks (which would require reset from local cues), but also report that training took a long time. In the one case discussed in this thesis where there may be an interaction between local cues and the head direction system (the Sharp *et al.* task, Chapter 11), I assumed that the head direction system was primed with the appropriate direction suggested by the expected orientation of the adjacent wall.

## B.3 Path integration

∗ **What is the anatomical locus of path integration?** The hypothesis proposed in Chapter 6 that path integration occurs by a loop between the subiculum, the parasubiculum, and the superficial layers of the entorhinal cortex implies that the hippocampus is not involved in path integration. The recent experiments of Alyan *et al.* (1997, EX.59) have confirmed this prediction. However, this does not prove that these other structures are involved in path integration. A number of experiments can be performed to test this hypothesis more directly. Lesions made to any of these three structures should impair the path integration ability of rodents. Neurophysiologically, place cells in all three structures should show normal activity in the dark. I know of no recordings from any of the three structures in the absence of sensory cues, although Quirk *et al.* (1992) report that ECs cells continued to show normal place fields after the cue card was removed.

If the anatomical hypothesis and mechanism discussed in Chapter 6 is correct, then two additional neurophysiological predictions can be made. First, one of the three structures should predict future position, the way that ATN cells predict future head direction (Blair *et al.*, 1997, see EX.102). Taube (1996) does report that parasubicular place cells are optimally tuned to a slight prediction (by $\sim$ 100 ms) of position (see EX.93), but more work needs to be done to show that this prediction is a consequence of the shifting-hill mechanism discussed in Chapter 6.

Second, the topology of fields in subiculum, parasubiculum, and superficial entorhinal cortex should not change from environment to environment. If one measures a Sub, PaS, or ECs place cell in one environment, nothing can be predicted about the location of its place field in another environment except that it must exist, but if one measures the place fields of two cells, the distance between their place fields must be identical in the two environments. Finally, if one records fields from three place cells, one cannot predict where the center of the triangle formed by their centers will be, but the triangle in the second environment must be *geometrically similar*[1] to the first.

† **Can the path integration mechanism in Chapter 6 be made to track actual trajectories of real animals?** How accurate can it be? What is required to make it accurate? By simulating the head direction system accurately, we found complexities (for example the deformation of the tuning curve) that were not obvious from the theoretical model. No one has yet reported an accurate simulation of two-dimensional path integration based on the moving hill hypothesis.

## B.4 Place code

‡ **What mechanism drives phase precession in the hippocampus?** Place cells in the hippocampus show *phase precession* (O'Keefe and Recce, 1993; Skaggs *et al.*, 1996, EX.74), that is the position represented by the place code sweeps across the actual position of the animal with each theta cycle (Tsodyks *et al.*, 1996). As reviewed in Section 3.4, there have been a number of proposals put forward to explain this effect, including that it is an intrinsic process formed by an interaction of two rhythms with similar frequencies that are just different enough to produce beats (O'Keefe and Recce, 1993), that it can be generated by asymmetric connections (Tsodyks *et al.*, 1996), that it is a consequence of attention sweeping from back to front with each theta cycle (Burgess *et al.*, 1994; O'Keefe and Burgess, 1996; Burgess and O'Keefe, 1996), and that it is a consequence of an interaction between the mechanism that drives path integration and internal neuronal dynamics (Samsonovich and McNaughton, 1996). Most of these hypotheses fail when compared to one data point or another. There is room here for theoretical work showing a simulation of phase precession

[1] Two triangles are geometrically similar if each pair of corresponding angles is the same, i.e. one triangle is a rescaled, rotated, and/or translated version of the other.

that fits all of the currently available data and quantitatively matches the recordings from Skaggs *et al.* (1996). One possibility that should be addressed is whether phase precession is generated extra-hippocampally. Bragin *et al.* (1995) have shown that the theta rhythm is partially driven from entorhinal cortex, but this possibility would also require that EC place cells show phase precession, which has not yet been tested.

$*$ **Observing the self-localization process.** If self-localization occurs as a consequence of pWTA dynamics in the hippocampus, then the self-localization process should be observable from multi-unit recording. If the place fields of a population of cells are recorded during an experience in an environment and then the spikes fired by all of the cells are recorded (simultaneously), the position represented by the population can be displayed at each moment in time by plotting the weighted sum of place fields (i.e. the sum of the place fields of all the cells that fired a spike at a moment in time). The representation should begin as a very broad (essentially uniform) distribution and then tighten up over the course of the self-localization process.

† **When does the self-localization process occur?** The self-localization process discussed in Chapters 7 and 8 is clearly necessary on returning to a familiar environment. We have proposed that self-localization occurs during a sharp wave during LIA. This predicts that one should see LIA on entry into a familiar environment, however, even if one doesn't see sharp waves, one should still see the tightening of the representation on entry.

It is not known whether the self-localization process occurs during motion within an environment. For example, does self-localization occur when an animal turns around at the end of an arm of a linear track or at the end of the arms of a radial maze? Given the two hypotheses that (1) different directions along a linear track are represented by different maps (Wan *et al.*, 1994c; McNaughton *et al.*, 1996; Touretzky and Redish, 1996; Redish and Touretzky, 1997a; Samsonovich and McNaughton, 1997; Samsonovich, 1997, see EX.64 and Section 3.2.4), (2) self-localization occurs during map transitions (Redish and Touretzky, 1997a, see also Section 3.3), then one would expect self-localization at the ends of the arms of a linear track. Or similarly, does self-localization occur along a well-travelled path? Gothard *et al.* (1996a) found realignment to occur along a journey at intermediate points along the journey, and one explanation for the data from Gothard *et al.* (1996b) is that the animals were realigning their representations to successive reference frames along their journey from the box to the room to the landmarks.

$*$ **Nearby candidate locations should be averaged but distant candidate locations should compete.** The hypothesized self-localization process averages nearby candidate locations, but selects between distant candidate locations. This can be tested behaviorally by examining search time of animals (such as gerbils) in a two-landmark task. If the animals are trained to find food at the center of a pair of landmarks and then tested with the landmarks at varied distances, the distance between candidate locations can be experimentally controlled. At large distances, the animal should alternate between two goal locations (as observed in EX.5 and EX.124), but at small distances, the animal should search at one location at the center of the pair.

† **What is the relationship between hippocampal mislocalization errors and behavioral errors?** If the radial maze is encoded by multiple maps, then there might be a relationship between reference memory errors, working memory errors and mislocalizations. Reference memory errors might consist of between-environment mislocalizations while working memory errors might consist of within-environment mislocalizations. There is data from Mizumori *et al.* (1989, see EX.77) suggesting that CA1 place fields did not change even though working memory errors were made, but they did find that CA3 fields were disrupted. More experiments explicitly designed to test this hypothesis might produce interesting results.

## B.5 Goal memory

‡ **What is the anatomy of the goal memory?** As discussed in Section 3.2.5, there is evidence that the nucleus accumbens is a key component of the goal memory (as first hypothesized by Mogenson, 1984), however, animals with fimbria-fornix lesions and septal grafts can still learn navigation tasks such as the water maze (Nilsson *et al.*, 1987, EX.26). This means that there must be alternate pathways from the hippocampus to motor structures. One candidate may be the posterior cingulate cortex (Sutherland and Hoesing, 1993). Further work needs to be done to completely ferret out the role played by the nucleus accumbens in the navigation system, as well as the role played by other structures such as the posterior cingulate cortex.

‡ **How do animals plan paths and avoid obstacles?** An important role of the goal memory is to plan a trajectory from the animal's current position to a goal. If there are obstacles in the way or if the animal must take a complex path to reach the goal, the goal memory should be able to plan these kinds of complex trajectories. Animals are generally very successful at planning complex trajectories and avoiding obstacles (for examples see Watson, 1907; Carr and Watson, 1908; Honzik, 1936; Chapuis and Scardigli, 1993). It is not currently known how obstacles are represented in the rodent brain or how the planning is done to avoid them. A number of robotics algorithms have been developed to avoid obstacles or plan complex paths, but they do not lend themselves to neural implementations: *potential field navigation* (Khatib, 1986; Connolly *et al.*, 1990; Tarassenko and Blake, 1991; Connolly and Grupen, 1992), *occupancy grids* (Moravec, 1988), *sinusoidal transforms* (Pratt, 1991a; 1991b), *graph search* (Muller *et al.*, 1991b; 1996), see also Trullier *et al.* (1997) for a discussion of this issue. More theoretical and experimental work is needed to understand how rodents plan paths and avoid obstacles.

## B.6 Reference frames

‡ **How are new reference frames separated in novel environments?** As described in Section 3.3, there are two possible mechanisms for separating reference frames between two similar environments: (1) CA3 could be pre-wired into a number of *charts* (Samsonovich and McNaughton, 1997) or (2) the dentate gyrus could orthogonalize the input (McNaughton and Morris, 1987, see also Marr, 1969; McNaughton, 1989; Rolls, 1989; O'Reilly and McClelland, 1994; Rolls, 1996 and Chapter 10). These two theories are not incompatible, and both may be correct. More experimental work is needed to understand which mechanisms occur in the rodent brain and more theoretical work is needed to understand the interaction between these two mechanisms.

There is important theoretical work to be done here also, examining the relationship between the two hypothesized mechanisms. Are both mechanisms viable? Are they compatible? The simulations used throughout the thesis assumed that the intra-hippocampal connections had very small random initial weights before experience in the simulated environment. This is obviously an oversimplification. How would it change the results if some of the connections had strong synaptic weights before experience?

‡ **Does the spatial and contextual input enter the hippocampus through separate structures?** There is some preliminary evidence that location information (such as landmark spatial aspects) enters the hippocampal formation through different pathways than reference frame selection information (such as object identity, olfactory information, etc.). Otto *et al.* (1996, EX.33) found a difference between LEC and MEC lesions. Quirk *et al.* (1992, EX.91) found broad sensitivities to location in MEC but no sensitivity to environment. No one has (to my knowledge) recorded from LEC.

One can also imagine theoretical work on this issue. Is it possible for both the spatial and contextual input to enter through the same pathway? How does this effect the path integrator input into the hippocampal formation? Is it necessary for the two inputs to be anatomically separated?

## B.7 Memory consolidation and the hippocampus

‡ **Can retrograde amnesia really be limited?** The role of the hippocampus in consolidation has been hotly debated (Nadel, 1991; Squire, 1992; Cohen and Eichenbaum, 1993), in particular, the question has been whether the hippocampus is only a temporary memory store or whether it is always necessary for navigation (Bolhuis *et al.*, 1994; Weisend *et al.*, 1996; Koerner *et al.*, 1996; Bohbot *et al.*, 1996). The Sutherland and Hoesing (1993) result (EX.15) would seem to conclusively lay the question to rest, but the lesion was made with colchicine (Sutherland, personal communication), which selectively targets dentate gyrus granule cells, and tends to spare the CA3 and CA1 fields. People have reported definite consolidation effects in rodents (Winocur, 1990; Cho *et al.*, 1993; Sutherland and Hoesing,

1993), but none of these lesions were purely or completely hippocampal. On the other hand, ischemic lesions produce clear limited retrograde amnesias in humans (Squire, 1992; Rempel-Clower *et al.*, 1996) and the only observable damage is to the CA1 hippocampal field. On the other hand, there may be covert damage beyond the hippocampus which produces the limited retrograde effect (E. Wood, personal communication). Further work is needed to resolve these issues.

‡ **What is the role of sleep states in memory consolidation?** As discussed in Section 3.4, the different roles of REM and SWS sleep states in consolidation is still an open question. Disruption experiments and multi-unit recording experiments have implicated both states in various aspects of consolidation, but the specifics roles played by the two is still an open question.

‡ **What structures subserve the consolidated routes?** An important question partially addressed in Section 3.4 is where are the routes stored if they are stored outside of hippocampus? One possibility is the posterior cingulate cortex (Sutherland and Hoesing, 1993, see also McNaughton *et al.*, 1991; Chen *et al.*, 1994a; Chen *et al.*, 1994b). However, the recordings presented by McNaughton *et al.* (McNaughton *et al.*, 1991; Chen *et al.*, 1994a; Chen *et al.*, 1994b) are preliminary; more extensive experimental work examining the role of posterior cingulate cortex is necessary before this hypothesis can be proven or disproven.

## B.8 Neuropharmacology

‡ **What are the roles of the neuromodulators in hippocampal function?** One of the most intriguing current issues in navigation is the question of the roles of neuromodulators such as acetylcholine, norepinephrine, serotonin, and dopamine (see Section 2.3, EX.114–EX.121). These neuromodulators may be involved in changing the state of the navigational system (e.g. from storage to recall or replay). Although there has been some remarkable progress in correlating dopaminergic and noradrenergic cell firing rates to behavioral variables (dopaminergic: Schultz *et al.*, 1995; Schultz, 1997, noradrenergic: Sara and Segal, 1991; Aston-Jones *et al.*, 1991; Rajkowski *et al.*, 1994), what is needed is an understanding of the influence of these neuromodulators on the representations in hippocampus and the other structures discussed in this thesis. There has been some progress in this issue for acetylcholine (Hasselmo and Bower, 1993; Hasselmo and Schnell, 1994; Fox *et al.*, 1997), but further work is still needed.

Navigation and the hippocampus is a particularly useful domain in which to study the role of neuromodulators, because we have a good understanding of the representation and general function of the structures. This means that if a representation or specific function is modified by a neuromodulator it may be possible to identify the informational change and make progress in understanding the role of the neuromodulator beyond the (important but limited) biophysical methods often used.

$*$ **A prediction from the interaction between the interference and mislocalization hypotheses.** An intriguing prediction comes from an interaction between the interference and reference frame theories: in animals with cholinergic deficits, cells with fields in multiple reference frames should pull their neighbors from one frame into the other. For example, imagine two cells $a$ and $b$ which have place fields near each other in one environment but only cell $b$ has a field in another environment. This prediction is that as the animal gains experience with the two environments, cell $a$ will begin to show a place field adjacent to cell $b$ in the second environment.

The key point made by the interference hypothesis is that representations should blend together along connections made by shared neurons. The reference frame hypothesis requires cells with overlapping place fields in one reference frame to have strong connections. Therefore we can predict which cells will be drawn into a reference frame. Simultaneous recordings of multiple cells made under low ACh conditions should show this effect.

# Appendix C

# Glossary

$\gamma_i$ **:** In the equations used throughout this thesis, $\gamma_i$ refers to a tonic input term (usually inhibitory) required for the dynamics of the system.

$\theta_i$ **:** Egocentric direction of a landmark.

$\tau_i$ **:** Decay of the synaptic drive of a neuron in the neural model used in this thesis.

$\Phi_H$ **:** Head direction of an animal.

$\phi_i$ **:** Allocentric bearing of a landmark.

**5-HT :** Serotonin, a neurotransmitter.

**Allocentric bearing :** Direction to a landmark, relative to an extrinsic direction (such as *north*).

**Acetylcholine (ACh) :** A neurotransmitter.

**ACh :** Acetylcholine, a neurotransmitter.

**ATN :** Anterior thalamic nuclei, a rodent brain structure, probably involved in the head direction system.

**CA1 :** A rodent brain structure, part of the hippocampus.

**CA3 :** A rodent brain structure, part of the hippocampus.

**Posterior cingulate cortex :** A rodent brain structure, also known as *retrosplenial cortex*.

**DA :** Dopamine, a neurotransmitter.

**Dentate gyrus (DG) :** A rodent brain structure, part of the hippocampus.

**DG :** Dentate gyrus, a rodent brain structure, part of the hippocampus.

**Dopamine (DA) :** A neurotransmitter.

**EC :** Entorhinal cortex, a rodent brain structure, divided into the superficial (ECs) and deep (ECd) layers.

**ECs :** Superficial layers of the entorhinal cortex, a rodent brain structure.

**ECd :** Deep layers of the entorhinal cortex, a rodent brain structure.

**EEG :** Electroencephelogram, a measurement of the ensemble neural activity in the brain, usually characterized by rhythmicities and dominant frequencies, see LIA and theta.

**Egocentric bearing :** Direction of a landmark, relative to the a property of the animal, such as the midline. Compare *allocentric bearing*.

**Fimbria :** A neural pathway connecting the septal nuclei with the hippocampus.

**Fornix :** A neural pathway connecting the hippocampus and subiculum with subcortical structures such as the septal nuclei, the nucleus accumbens, and others.

**GABA :** A neurotransmitter, generally inhibitory.

**Glutamate :** A neurotransmitter, generally excitatory.

**HD :** Head direction.

**Head direction (HD) :** A representation of the orientation of the animal with respect to the external world.

**Hippocampus :** A rodent brain structure, consisting of the dentate gyrus, CA3 and CA1 fields.

**Hippocampal formation :** The hippocampus and its adjacent structures including the entorhinal cortex, subiculum, postsubiculum, etc.

**HLS :** High level sensory areas: an abstraction of the complexities in sensory representations. In the text, representations of distal landmarks are referred to by $\mathrm{HLS}_{\mathcal{D}}$ and representations of local landmarks by $\mathrm{HLS}_{\mathcal{L}}$.

**LEC :** Lateral entorhinal cortex, a rodent brain structure.

**LDN :** Lateral dorsal nucleus of the thalamus, a rodent brain structure.

**LIA :** A very noisy EEG signal in the hippocampus characterized by sharp waves.

**Local view :** A representation of the spatial relationship between an animal and sensory landmarks. Usually refers only to simple spatial aspects directly computable from sensory cues. Sometimes includes non-spatial aspects as well.

**Locale navigation :** Navigation based on complex mapping abilities.

**LTP :** Long term potentiation, an increase in the synaptic efficacy between two neurons after certain conditions (e.g. the two neurons both fire action potentials). Thought to be the basis for learning in the brain.

**LVFA :** Low-voltage-fast-activity, an EEG signal occurring in cortex during REM sleep.

**MEC :** Medial entorhinal cortex, a rodent brain structure.

**Multiplicative synapse :** A three-way synapse where two inputs meet at the same point on the dendrite with a non-linear effect. Mathematically, the average synaptic depolarization is proportional to the product of the two input firing rates.

**NE :** Norepinephrine, a neurotransmitter, also known as noradrenaline.

**Noradrenaline (NE) :** A neurotransmitter, also known as norepinephrin.

**Norepinephrin (NE) :** A neurotransmitter, also known as noradrenaline.

**NMDA receptor :** A glutamatergic (excitatory) receptor thought to be strongly involved in LTP.

**Nucleus accumbens (NAcb) :** A rodent brain structure, probably involved in goal memory.

**Parasubiculum (PaS) :** A rodent brain structure, possibly involved in path integration.

**Parietal cortex :** A rodent brain structure, probably involved in local view.

**PaS :** Parasubiculum.

**Path integration (PI) :** The ability to return directly to a starting point from any location in an environment, using only internal cues. Implies a representation of a "vector home."

**PI :** Path integration.

**Place cell :** A cell that only fires strongly in a restricted portion of an environment (the *place field* of the cell).

**Place field :** The area of an environment in which a place cell fires strongly.

**Postsubiculum (PoS) :** A rodent brain structure, involved in the head direction system.

**PoS :** Postsubiculum.

**Praxic navigation :** Navigation based on a motor sequence dependent on simple cue responses or internally driven.

**Preferred direction :** The direction at which a head direction cell shows its maximal firing rate.

**Preferred location :** The location at which a place cell shows its maximal firing rate.

**Presubiculum :** A rodent brain structure. Some researchers argue that the postsubiculum is only a component of the presubiculum.

**pseudo-Winner-take-all network :** A pocket of activation is a stable state of the system. Any other state will progress to a compact pocket of activation. Because multiple cells are active but with a smooth fall-off as each cell's parameter varies from the currently represented value, this is not a true WTA network, nor is it a kWTA network.

**pWTA network :** pseudo-Winner-take-all network.

**REM :** A sleep state characterized by theta waves in the hippocampus and LVFA in the cortex.

**Septal nuclei :** A rodent brain structure, providing cholinergic and GABA-ergic input into the hippocampus and other structures.

**Serotonin (5-HT) :** A neurotransmitter.

**Sharp waves :** Short (~100–200 ms) bursts of activity in hippocampus occurring during LIA. Usually measured with EEG.

**Sim-animal :** Simulation of an animal in a simulated environment. Used in order to distinguish simulated from real data.

**Subiculum (Sub) :** A rodent brain structure, possibly involved in path integration.

**Slow wave sleep (SWS) :** A sleep state characterized by LIA in the hippocampus and sharp waves.

**Synaptic drive :** A non-measurable property of a neuron that can be understood as the effect of the neuron on all the neurons on which it synapses divided by the synaptic weight across each synapse.

**Taxon navigation :** Navigation based on approach towards or withdrawal away from a single landmark.

**Theta :** An EEG signal in hippocampus characterized by a strong 5–10 Hz signal.

# Part V

# Bibliography

# Bibliography

Abbott, L. F. and Blum, K. I. (1996). Functional significance of long-term potentiation for sequence learning and prediction. *Cerebral Cortex*, 6(3):406–416.

Aggleton, J. P. (1993). The contribution of the amygdala to normal and abnormal emotional states. *Trends in Neurosciences*, 1(8):328–333.

Alexander, G. E., DeLong, M. R., and Crutcher, M. D. (1992). Do cortical and basal ganglionic motor areas use "motor programs" to control movement? *Behavioral and Brain Sciences*, 15(4):656–665.

Alvarez, P. and Squire, L. R. (1994). Memory consolidation and the medial temporal lobe: A simple network model. *Proceedings of the National Academy of Sciences, USA*, 91:7041–7045.

Alyan, S. H. and Jander, R. (1994). Short-range homing in the house mouse *mus musculus:* stages in the learning of directions. *Animal Behavior*, 48(2):285–298.

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

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

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

Amari, S.-I. (1977). Dynamics of pattern formation in lateral-inhibition type neural fields. *Biological Cybernetics*, 27:77–87.

Andersen, R. (1988). The neurobiological basis of spatial cognition: Role of the parietal lobe. In Stiles-Davis, J., Kritchevsky, M., and Bellugi, U., editors, *Spatial Cognition: Brain Bases and Development.* Erlbaum, Hillsdale NJ.

Andersen, R. A., Snyder, L. H., Li, C. S., and Stricanne, B. (1993). Coordinate transformations in the representation of spatial information. *Current Opinion in Neurobiology*, 3:171–176.

Annett, L. E., McGregor, A., and Robbins, T. W. (1989). The effects of ibotenic acid lesions of the nucleus accumbens on spatial learning and extinction in the rat. *Behavioral Brain Research*, 31(3):231–242.

Arai, K., Keller, E. L., and Edelman, J. A. (1994). Two-dimensional neural network model of the primate saccadic system. *Neural Networks*, 7(6/7):1115–1135.

Aston-Jones, G., Chiang, C., and Alexinsky, T. (1991). Discharge of noradrenergic locus coeruleus neurons in behaving rats and monkeys suggests a role in vigilance. *Progress in Brain Research*, 88:501–520.

Austin, K. B., Fortin, W. F., and Shapiro, M. L. (1990). Place fields are altered by NMDA antagonist MK-801 during spatial learning. *Society for Neuroscience Abstracts*, 16:263.

Austin, K. B., White, L. H., and Shapiro, M. L. (1993). Short- and long-term effects of experience on hippocampal place fields. *Society for Neuroscience Abstracts*, 19:797.

Bach, M. E., Hawkins, R. D., Osman, M., Kandel, E. R., and Mayford, M. (1995). Impairment of spatial but not contextual memory in CaMKII mutant mice with a selective loss of hippocampal LTP in the range of the theta frequency. *Cell*, 81:905–915.

Barlow, J. S. (1964). Inertial navigation as a basis for animal navigation. *Journal of Theoretical Biology*, 6:76–117.

Barnes, C. A. (1979). Memory deficits associated with senscence: A neurophysiological and behavioral study in the rat. *Journal of Comparative and Physiological Psychology*, 93:74–104.

Barnes, C. A. (1988). Spatial learning and memory processes: The search for their neurobiological mechanisms in the rat. *Trends in Neurosciences*, 11:163–169.

Barnes, C. A. (1994). Normal aging: Regionally specific changes in hippocampal synaptic transmission. *Trends in Neurosciences*, 17(1):13–18.

Barnes, C. A. (1995). Involvement of LTP in memory: Are we "searching under the street light?". *Neuron*, 15:751–754.

Barnes, C. A. (1996). Memory changes during normal aging: Neurobiological correlates. In Martinez, Jr., J. L. and Kesner, R. P., editors, *Learning and Memory: A biological view*. Academic Press, New York. In Press.

Barnes, C. A., McNaughton, B. L., Mizumori, S. J. Y., Leonard, B. W., and Lin, L. H. (1990). 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*, pages 287–300. Elsevier Science Publishers B.V., New York.

Barnes, C. A., McNaughton, B. L., and O'Keefe, J. (1983). Loss of place specificity in hippocampal complex spike cells of senescent rat. *Neurobiology of Aging*, 4:113–119.

Barnes, C. A., Nadel, L., and Honig, W. K. (1980). Spatial memory deficit in senescent rats. *Canadian journal of psychology*, 34(1):29–39.

Barnes, C. A., Rao, G., and McNaughton, B. L. (1996). Functional integrity of NMDA-dependent LTP induction mechanisms across the lifespan of F-344 rats. *Learning and Memory*, 3:124–137.

Barnes, C. A., Suster, M. S., Shen, J., and McNaughton, B. L. (1997). Getting lost: Cognitive map multistability in aged rat hippocampus. Under editorial review.

Bellman, R. (1958). On a routing problem. *Quarterly Journal of Applied Mathematics*, 16(1):87–90.

Benhamou, S., Bovet, P., and Poucet, B. (1995). A model for place navigation in mammals. *Journal of Theoretical Biology*, 173:163–178.

Benhamou, S., Sauve, J. P., and Bovet, P. (1990). Spatial memory in large-scale movements: Efficiency and limitation of the egocentric coding process. *Journal of Theoretical Biology*, 145:1–12.

Bentivoglio, M., Kultas-Ilinsky, K., and Illinsky, I. (1993). Limbic thalamus: Structure, intrinsic organisation, and connections. In Vogt, B. A. and Gabriel, M., editors, *Neurobiology of Cingulate Cortex and Limbic Thalamus: A Comprehensive Handbook*, pages 71–122. Birkhauser, Boston.

Beritashvili, I. S. (1965). *Neural Mechanisms of Higher Vertebrate Behavior*. Little, Brown, Boston.

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

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

Bisiach, E. and Vallar, G. (1988). Hemineglect in humans. In Boller, F. and Grafman, J., editors, *Handbook of Neuropsychology*, volume 1, pages 195–222. Elsevier, New York.

Bjorklund, A., Nilsson, O. G., and Kalen, P. (1990). Reafferentation of the subcortically denervated hippocampus as a model for transplant-induced functional recovery in the CNS. *Progress in Brain Research*, 83:411–426.

Blair, H. T. (1996). A thalamocortical circuit for computing directional heading in the rat. In Touretzky, D. S., Mozer, M. C., and Hasselmo, M. E., editors, *Advances in Neural Information Processing Systems 8*, pages 152–158. MIT Press.

Blair, H. T., Lipscomb, B. W., and Sharp, P. E. (1997). Anticipatory time intervals of head-direction cells in the anterior thalamus of the rat, implications for path integration in the head-direction circuit. *Journal of Neurophysiology*. In press.

Blair, H. T. and Sharp, P. E. (1995). 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):6260–6270.

Bliss, T. V. P. and Lynch, M. A. (1988). Long-term potentiation of synaptic transmission in the hippocampus: Properties and mechanisms. In Landfield, P. W. and Deadwyler, S. A., editors, *Long-Term Potentiation: From Biophysics to Behavior*, pages 3–72. Liss, New York.

Blum, K. I. and Abbott, L. F. (1996). A model of spatial map formation in the hippocampus of the rat. *Neural Computation*, 8(1):85–93.

Bohbot, V., Liu, Z., Thurm, S. L., Nadel, L., and Bures, J. (1996). Spatial memory? Never without a hippocampus. *Society for Neuroscience Abstracts*, 22:1873.

Bolhuis, J. J., Stewart, C. A., and Forrest, E. M. (1994). Retrograde amnesia and memory reactivation in rats with ibotenate lesions to the hippocampus or subiculum. *Quarterly Journal of Experimental Psychology*, 47B(2):129–150.

Bondy, J. A. and Murty, U. S. R. (1976). *Graph theory with applications*. North-Holland, New York.

Bostock, E., Muller, R. U., and Kubie, J. L. (1991). Experience-dependent modifications of hippocampal place cell firing. *Hippocampus*, 1(2):193–206.

Bragin, A., Jando, G., Nadasdy, Z., Hetke, J., Wise, K., and Buzsáki, G. (1995). Gamma (40-100 hz) oscillation in the hippocampus of the behaving rat. *Journal of Neuroscience*, 15(1):47–60.

Brown, M. A. and Sharp, P. E. (1995). Simulation of spatial learning in the Morris water maze by a neural network model of the hippocampal formation and nucleus accumbens. *Hippocampus*, 5(3):171–188.

Brown, T. H., Zador, A. M., Mainen, Z. F., and Clairborne, B. J. (1991). Hebbian modifications in hippocampal neurons. In Boudry, M. and Davis, J. L., editors, *Long-Term Potentiation: A Debate of Current Issues*, pages 357–389. MIT Press, Cambridge MA.

Buhot, M.-C. (1997). Serotonin receptors in cognitive behaviors. *Current Opinion in Neurobiology*, 7:243–254.

Buhot, M.-C. and Naili, S. (1995). Changes in exploratory activity following stimulation of hippocampal 5-HT1A and 5-HT1B receptors in the rat. *Hippocampus*, 5(3):198–208.

Bunsey, M. and Eichenbaum, H. (1996). Conservation of hippocampal function in rats and humans. *Nature*, 379:255–257.

Buresova, O., Bolhuis, J. J., and Bures, J. (1986). Differential effects of cholinergic blockade on performance of rats in the water tank navigation task and in a radial water task. *Behavioral Neurology*, 100:476–482.

Burgess, N. and O'Keefe, J. (1996). Neuronal computations underlying the firing of place cells and their role in navigation. *Hippocampus*, 7:749–762.

Burgess, N., O'Keefe, J., and Recce, M. (1993). Using hippocampal 'place cells' for navigation, exploiting phase coding. In Hanson, S., Cowan, J., and Giles, L., editors, *Advances in Neural Information Processing Systems 5*. Morgan Kaufmann, San Mateo, CA.

Burgess, N., Recce, M., and O'Keefe, J. (1994). A model of hippocampal function. *Neural Networks*, 7(6/7):1065–1081.

Burwell, R. D., Witter, M. P., and Amaral, D. G. (1995). Perirhinal and postrhinal cortices of the rat: A review of the neuroanatomical literature and comparison with findings from the monkey brain. *Hippocampus*, 5(5):309–408.

Buzsáki, G. (1989). Two-stage model of memory trace formation: A role for "noisy" brain states. *Neuroscience*, 31(3):551–570.

Carr, H. (1917). Maze studies with the white rat. *Journal of Animal Behavior*, 7:259–306.

Carr, H. and Watson, J. B. (1908). Orientation in the white rat. *Journal of Comparative Neurology and Psychology*, 18:27–44.

Chance, M. R. A. and Mead, A. P. (1955). Competition between feeding and investigation in the rat. *Behavior*, 8:174–181.

Chandler, H. C., King, V., Corwin, J. V., and Reep, R. L. (1992). Thalamocortical connections of rat posterior parietal cortex. *Neuroscience Letters*, 143:237–242.

Chapuis, N. and Scardigli, P. (1993). Shortcut ability in hamsters (*mesocricetus auratus*): The role of environmental and kinesthetic information. *Animal Learning and Behavior*, 21(3):255–265.

Chen, L. L. (1989). Spatial representation in the rat posterior parietal cortex. Master's thesis, University of Colorado, Department of Psychology.

Chen, L. L. (1991). *Head-Directional Information Processing in the Rat Posterior Cortical Areas*. PhD thesis, University of Colorado.

Chen, L. L., Lin, L. H., Barnes, C. A., and McNaughton, B. L. (1994a). Head-direction cells in the rat posterior cortex: II. Contributions of visual and ideothetic information to the directional firing. *Experimental Brain Research*, 101:24–34.

Chen, L. L., Lin, L. H., Green, E. J., Barnes, C. A., and McNaughton, B. L. (1994b). Head-direction cells in the rat posterior cortex: I. Anatomical distribution and behavioral modulation. *Experimental Brain Research*, 101:8–23.

Cheng, K. (1986). A purely geometric module in the rat's spatial representation. *Cognition*, 23:149–178.

Cheng, K. (1989). The vector sum model of pigeon landmark use. *Journal of Experimental Psychology*, 15(4):366–375.

Cho, Y. H., Berracochea, D., and Jaffard, R. (1993). Extended temporal gradient for the retorgrade and anterograde amnesia produced by ibotenate entorhinal cortex lesions in mice. *Journal of Neuroscience*, 13(4):1759–1766.

Christian, E. P. and Deadwyler, S. A. (1986). Behavioral function and hippocampal cell types: Evidence for two non-overlapping populations in the rat. *Journal of Neurophysiology*, 55:331–348.

Chrobak, J. J. and Buzsáki, G. (1994). Selective activation of deep layer (V-VI) retrohippocampal neurons during hippocampal sharp waves in the behaving rat. *Journal of Neuroscience*, 14(10):6160–6170.

Cohen, N. J. and Eichenbaum, H. (1993). *Memory, Amnesia, and the Hippocampal System*. MIT Press, Cambridge, MA.

Cohen, N. J. and Squire, L. R. (1980). Preserved learning and retention of pattern-analyzing skill in amnesia: Dissociation of knowing how and knowing that. *Science*, 210:207–210.

Colby, C. L. and Duhamel, J.-R. (1991). Heterogeneity of extrastriate visual areas and multiple parietal areas in the macaque monkey. *Neuropsychologia*, 29(6):517–537.

Colby, C. L., Duhamel, J.-R., and Goldberg, M. E. (1995). Occulocentric spatial representation in parietal cortex. *Cerebral Cortex*, 5(5):470–481.

Collett, T., Cartwright, B. A., and Smith, B. A. (1986). Landmark learning and visuospatial memories in gerbils. *Journal of Comparative Physiology A*, 158:835–851.

Collinder, P. (1955). *A History of Marine Navigation.* St. Martin's Press, New York. Translated from the Swedish by M. Michael.

Connolly, C. I., Burns, J. B., and Weiss, R. (1990). Path planning using Laplace's equation. *Proceedings of the 1990 IEEE International Conference on Robotics and Automation*, pages 2102–2106.

Connolly, C. I. and Grupen, R. A. (1992). Applications of harmonic functions to robotics. Technical Report COINS-92-12, Computer and Information Science Department, University of Massachusetts.

Cooper, J. R., Bloom, F. E., and Roth, R. H. (1986). *The Biochemical Basis of Neuropharmacology.* Oxford University Press, New York, 5th edition.

Corkin, S. (1968). Acquisition of motor skill after bilateral medial temporal-lobe excision. *Neuropsychologia*, 6:255–265.

Cormen, T. H., Leiserson, C. E., and Rivest, R. L. (1990). *Introduction to Algorithms.* MIT Press and McGraw-Hill, Cambridge MA and New York NY.

Cotter, C. H. (1968). *A History of Nautical Astronomy.* American Elsevier, New York.

Cressant, A., Muller, R. U., and Poucet, B. (1997). Failure of centrally placed objects to control firing fields of hippocampal place cells. *Journal of Neuroscience*, 17(7):2531–2542.

Darwin, C. A. (1873a). Origin of certain instincts. *Nature*, 7:417–418.

Darwin, C. A. (1873b). Perception in the lower animals. *Nature*, 7:360.

Davis, M., Rainnie, D., and Cassell, M. (1994). Neurotransmission in the rat amygdala related to fear and anxiety. *Trends in Neurosciences*, 17(5):208–214.

Dayan, P. (1991). Navigating through temporal difference. In Lippmann, R. P., Moody, J. E., and Touretzky, D. S., editors, *Advances in Neural Information Processing Systems 3*, pages 464–470. Morgan Kauffman, San Mateo CA.

Dayan, P. and Singh, S. P. (1996). Long term potentiation, navigation, and dynamic programming. Presented at *CNS*96*, (the fifth annual *Computation in Neural Systems* meeting).

Dean, P. (1990). Sensory cortex: Visual perception functions. In Kolb, B. and Tees, R. C., editors, *The Cerebral Cortex of the Rat*, chapter 12, pages 275–308. MIT Press, Cambridge MA.

Decker, M. W. and McGaugh, J. L. (1991). The role of interactions between the cholinergic system and other neuromodulatory systems in learning and memory. *Synapse*, 7:151–168.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society B*, 39(1):1–38.

Dennis, W. (1932). Multiple visual discrimination in the block elevated maze. *Journal of Comparitive and Physiological Psychology*, 13:391–396.

Desimone, R., Schein, S. J., Moran, J., and Ungerleider, L. G. (1985). Contour, color, and shape analysis beyond the striate cortex. *Vision Research*, 25(3):444–452.

Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. *Numerische Mathematik*, 1:269–271.

DiMattia, B. V. D. and Kesner, R. P. (1988). Spatial cognitive maps: Differential role of parietal cortex and hippocampal formation. *Behavioral Neuroscience*, 102(4):471–480.

Domesick, V. B. (1981). Further observations on the anatomy of nucleus accumbens and caudatoputamen in the rat: Similarities and contrasts. In Chronister, R. B. and de France, J. F., editors, *The Neurobiology of the Nucleus Accumbens*. Haer Institute for Electrophysiological Research, Brunswick, Maine.

Droulez, J. and Berthoz, A. (1991). A neural network model of sensoritopic maps with predictive short-term memory properties. *Proceedings of the National Academy of Sciences, USA*, 88:9653–9657.

Dudchenko, P. A. and Taube, J. S. (1997). Correlation between head direction cell activity and spatial behavior on a radial arm maze. *Behavioral Neuroscience*, 111(1):3–19.

Eckerman, D. A., Gordon, W. A., Edwards, J. D., MacPhail, R. C., and Gage, M. I. (1980). Effects of scopolomine, pentobarbitol, and amphetamine on radial arm maze performance in the rat. *Pharmacology Biochemistry and Behavior*, 12:595–602.

Efron, B. (1982). *The Jackknife, the Bootstrap, and Other Resampling Plans*, volume 38 of *Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics, Philadelphia PA.

Eichenbaum, H. (1996). Is the rodent hippocampus just for "place?". *Current Opinion in Neurobiology*, 6:187–195.

Eichenbaum, H. and Cohen, N. J. (1988). Representation in the hippocampus: What do hippocampal neurons code? *Trends in Neurosciences*, 11(6):244–248.

Eichenbaum, H., Kuperstein, M., Fagan, A., and Nagode, J. (1987). Cue-sampling and goal-approach correlates of hippocampal unit activity in rats performing an odor-discrimination task. *Journal of Neuroscience*, 7(3):716–732.

Eichenbaum, H., Stewart, C., and Morris, R. G. M. (1990). Hippocampal representation in place learning. *Journal of Neuroscience*, 10(11):3531–3542.

Eilam, D. and Golani, I. (1989). Home base behavior of rats (*rattus norvegicus*) exploring a novel environment. *Behavioral Brian Research*, 34:199–211.

Elga, A. N., Redish, A. D., and Touretzky, D. S. (1996). A model of the rodent head direction system. *CNS*96 (Fifth Annual Computational Neuroscience Meeting).*

Ermentrout, B. and Cowan, J. (1979). A mathematical theory of visual hallucination patterns. *Biological Cybernetics*, 34:137–150.

Etienne, A. S. (1987). 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 I (*Experimental Animal Psychology and Ethology*), pages 233–251. Martinus Nijhoff Publishers, Boston.

Etienne, A. S. (1992). Navigation of a small mammal by dead reckoning and local cues. *Current Directions in Psychological Science*, 1(2):48–52.

Etienne, A. S., Maurer, R., and Saucy, F. (1988). Limitations in the assessment of path dependent information. *Behavior*, 106:81–111.

Etienne, A. S., Maurer, R., Saucy, F., and Teroni, E. (1986). Short-distance homing in the golden hamster after a passive outward journey. *Animal Behavior*, 34:696–715.

Farrall, L. (1984). *Unwritten Knowledge: Case Study of the Navigators of Micronesia.* Deakin University, Victoria 3217 Australia.

Finch, D. M. (1993). Hippocampal, subicular, and entorhinal afferents and synaptic integration in rodent cingulate cortex. In Vogt, B. A. and Gabriel, M., editors, *Neurobiology of Cingulate Cortex and Limbic Thalamus: A Comprehensive Handbook*, pages 224–248. Birkhauser, Boston.

Finch, D. M. (1996). Neurophysiology of converging synaptic inputs from rat prefrontal cortex, amygdala, midline thalamus, and hippocampal formation onto single neurons of the caudate/putamen and nucleus accumbens. *Hippocampus*, 6:495–512.

Flicker, C. and Geyer, M. A. (1982). Behavior during hippocampal microinfusions. I. Norepinephrine and diversive exploration. *Brain Research Reviews*, 4:79–103.

Floresco, S. B., Seamans, J. K., and Phillips, A. G. (1997). Selective roles for hippocampal, prefrontal cortical, and ventral striatal circuits in radial-arm maze tasks with or without a delay. *Journal of Neuroscience*, 17(5):1880–1890.

Ford, L. R. and Fulkerson, D. R. (1962). *Flows in Networks.* Princeton University Press, Princeton NJ.

Forde, H. (1873). Sense of direction. *Nature*, 7(463-464).

Fox, S. E., Ludvig, N., Kubie, J. L., Muller, R. U., Stead, M., and Fenton, A. (1997). The effects of scopolamine, delivered via intrahippocampal microdialysis, on the firing of local place cells. *Society for Neuroscience Abstracts*, 22:431.

Freud, S. (1923). *The Ego and the Id*, volume 19 of *The Standard Edition of the Complete Psychological Works of Sigmund Freud.* The Hogarth Press, London. (1961 edition.) Translated and Edited by J. Strachey.

Freund, T. F. and Buzsáki, G. (1996). Interneurons of the hippocampus. *Hippocampus*, 6(4):345–370.

Fuhs, M., Redish, A. D., and Touretzky, D. S. (1997). A visually driven hippocampal model. Presented at *CNS*97*, (the sixth annual *Computation in Neural Systems* meeting).

Fujita, N., Loomis, J. M., Klatzky, R. L., and Golledge, R. G. (1990). A minimal representation for dead-reckoning navigation: Updating the homing vector. *Geographical Analysis*, 22(4):326–335.

Gaffan, D. (1972). Loss of recognition memory in rats with lesions of the fornix. *Neuropsychologia*, 10:327–341.

Gaffan, D. (1974). Recognition impaired and association intact in the memory of monkeys after transection of the fornix. *Journal of Comparative and Physiological Psychology*, 86:1100–1109.

Gaffan, D., Sanders, R. C., Gaffan, E. A., Harrison, S., Shields, C., and Owen, M. J. (1984). Effects of fornix transection upon associative memory in monkeys: Role of the hippocampus in learned actions. *Quarterly Journal of Experimental Psychology*, 36B:173–221.

Gallagher, M., Burwell, R., and Burchinal, M. (1993). Severity of spatial learning impairment in aging: Development of a learning index for performance in the Morris water maze. *Behavioral Neuroscience*, 107(4):618–626.

Gallagher, M. and Burwell, R. D. (1989). Relationship of age-related decline across several behavioral domains. *Neurobiology of aging*, 10:691–708.

Gallagher, M. and Colombo, P. J. (1995). Ageing: The cholinergic hypothesis of cognitive decline. *Current Opinion in Neurobiology*, 5:161–168.

Gallagher, M., Nagahara, A. H., and Burwell, R. D. (1995). Cognition and hippocampal systems in aging: Animal models. In McGaugh, J. L., Weinberger, N. M., and Lynch, G., editors, *Brain and Memory: Modulation and Mediation of Neuroplasticity*, chapter 6, pages 103–126. Oxford, New York NY.

Gallagher, M. and Nicole, M. M. (1993). Animal models of normal aging: Relationship between cognitive decline and markers in hippocampal circuitry. *Behavioral Brain Research*, 57:155–162.

Gallistel, C. R. (1980). *The Organization of Action: A New Synthesis*. Lawrence Earlbaum Associates, Hillsdale NJ.

Gallistel, C. R. (1990). *The Organization of Learning*. MIT Press, Cambridge, MA.

Georgopoulos, A. P., Caminiti, R., Kalaska, J. F., and Massey, J. T. (1983). Spatial coding of movement: A hypothesis concerning the coding of movement direction by motor cortical populations. *Experimental Brain Research*, Supplement(7):327–336.

Georgopoulos, A. P., Kettner, R. E., and Schwartz, A. B. (1988). 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. *Journal of Neuroscience*, 8(8):2928–2937.

Gerstner, W. and Abbott, L. F. (1997). Learning navigational maps through potentiation and modulation of hippocampal place cells. *Journal of Computational Neuroscience*, 4:79–94.

Gladwin, T. (1970). *East is a Big Bird*. Harvard University Press, Cambridge MA.

Gluck, M. A. and Myers, C. E. (1993). Hippocampal mediation of stimulus representation: A computational theory. *Hippocampus*, 3(4):491–516.

Gluck, M. A. and Myers, C. E. (1996). Integrating behavioral and physiological models of hippocampal function. *Hippocampus*, 6(6):643–653.

Golani, I., Benjamini, Y., and Eilam, D. (1993). Stopping behavior: Constraints on exploration in rats (*rattus norvegicus*). *Behavioral Brain Research*, 53(1-2):21–33.

Goldman-Rakic, P. S., Funahashi, S., and Bruce, C. J. (1990). Neocortical memory circuits. *Cold Spring Harbor Symposia on Quantitative Biology*, LV:1025–1038.

Goldschmidt, R. B. and Steward, O. (1980). Preferential neurotoxicity of colchicine for granule cells of the dentate gyrus of the adult rat. *Proceedings of the National Academy of Sciences, USA*, 77(5):3047–3051.

Golob, E. J. and Taube, J. S. (1994). Head direction cells recorded from the postsubiculum in animals with lesions of the lateral dorsal thalamic nucleus. *Society for Neuroscience Abstracts*, 20:805.

Golob, E. J. and Taube, J. S. (1997). Head direction cells and episodic spatial information in rats without a hippocampus. *Proceedings of the National Academy of Sciences, USA*, 94:7645–7650.

Good, M. and Honey, R. C. (1991). Conditioning and contextual retrieval in hippocampal rats. *Behavioral Neuroscience*, 105:499–509.

Goodale, M. A. and Carey, D. P. (1990). The role of the cerebral cortex in visuomotor control. In Kolb, B. and Tees, R. C., editors, *The Cerebral Cortex of the Rat*, chapter 13, pages 309–340. MIT Press, Cambridge MA.

Goodale, M. A., Meenan, J. P., Bulthoff, H. H., Nicolle, D. A., Murphy, K. J., and Racicot, C. I. (1994). Separate neural pathways for the visual analysis of object shape in perception and prehension. *Current Biology*, 4(7):604–610.

Goodridge, J. P., Redish, A. D., Touretzky, D. S., Blair, H. T., and Sharp, P. E. (1997). Lateral mamillary input explains distortions in tuning curve shapes of anterior thalamic head direction cells. *Society for Neuroscience Abstracts*, 23. In press.

Goodridge, J. P. and Taube, J. S. (1994). The effect of lesions of the postsubiculum on head direction cell firing in the anterior thalamic nuclei. *Society for Neuroscience Abstracts*, 20(1):805.

Goodridge, J. P. and Taube, J. S. (1995). Preferential use of the landmark navigational system by head direction cells in rats. *Behavioral Neuroscience*, 109(1):49–61.

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

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

Gray, J. A. (1982a). Précis of *the neuropsychology of anxiety: an enquiry into the functions of the septo-hippocampal system*. *Behavioral and Brain Sciences*, 5:469–484. See also commentary and response, pages 484–534.

Gray, J. A. (1982b). *The Neuropsychology of Anxiety: An Enquiry into the Functions of the Septo-Hippocampal System*. Oxford University Press, New York.

Haas, H. (1983). Amine neurotransmitter actions in the hippocampus. In Seifert, W., editor, *Neurobiology of the hippocampus*, chapter 8, pages 139–156. Academic Press, London.

Hampson, R. E., Heyser, C. J., and Deadwyler, S. A. (1993). Hippocampal cell firing correlates of delayed-match-to-sample performance in the rat. *Behavioral Neuroscience*, 107(5):715–739.

Hasselmo, M. E. (1993). Acetylcholine and learning in a cortical associative memory. *Neural Computation*, 5:32–44.

Hasselmo, M. E. (1995). Neuromodulation and cortical function: Modeling the physiological basis of behavior. *Behavioral and Brain Research*, 67(1):1–27.

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

Hasselmo, M. E. and Schnell, E. (1994). Laminar selectivity of the cholinergic suppression of synaptic transmission in rat hippocampal region CA1: Computational modeling and brain slice physiology. *Journal of Neuroscience*, 14(6):3898–3914.

Hasselmo, M. E., Wyble, B. P., and Wallenstein, G. V. (1996). Encoding and retrieval of episodic memories: Role of cholinergic and GABAergic modulation in the hippocampus. *Hippocampus*, 6:693–708.

Hebb, D. O. (1949). *The Organization of Behavior*. Wiley, New York.

Hennevin, E., Hars, B., Maho, C., and Bloch, V. (1995). Processing of learned information in paradoxical sleep: Relevance for memory. *Behavioural Brain Research*, 69:125–135.

Hermer, L. and Spelke, E. S. (1994). A geometric process for spatial reorientation in young children. *Nature*, 370:57–9.

Hertz, J., Krogh, A., and Palmer, R. G. (1991). *Introduction to the Theory of Neural Computation*. Addison-Wesley, Reading, MA.

Hetherington, P. A. and Shapiro, M. L. (1993). A simple network model simulates hippocampal place fields: 2. Computing goal-directed trajectories and memory fields. *Behavioral Neuroscience*, 107(3):434–443.

Hill, A. J. (1978). First occurrence of hippocampal spatial firing in a new environment. *Experimental Neurology*, 62:282–297.

Hill, A. J. and Best, P. J. (1981). Effects of deafness and blindness on the spatial correlates of hippocampal unit activity in the rat. *Experimental neurology*, 74:204–217.

Hirsh, R. (1974). The hippocampus and contextual retrieval of information from memory: A theory. *Behavioral Biology*, 12:421–444.

Hirsh, R., Leber, B., and Gillman, K. (1978). Fornix fibers and motivational states as controllers of behavior: A study stimulated by the contextual retrieval theory. *Behavioral Biology*, 22:463–478.

Hodges, H., Allen, Y., Kershaw, T., Lantos, P. L., Gray, J. A., and Sinden, J. (1991a). Effects of cholinergic-rich neural grafts on radial maze performance of rats after excitotoxic lesions of the forebrain cholinergic system — I. Amelioration of cognitive deficits by transplants into cortex and hippocampus but not into basal forebrain. *Neuroscience*, 45(3):587–607.

Hodges, H., Allen, Y., Sinden, J., and Gray, J. A. (1991b). Effects of cholinergic-rich neural grafts on radial maze performance of rtas after excitotoxic lesions of the forebrain cholinergic system — II. Cholinergic drugs as probes to investigate lesion-induced deficits and transplant-induced functional recovery. *Neuroscience*, 45(3):609–623.

Honey, R. C. and Good, M. (1993). Selective hippocampal lesions abolish the contextual specificity of latent inhibition and of conditioning. *Behavioral Neuroscience*, 107:23–33.

Honzik, C. H. (1936). The sensory basis of maze learning in rats. *Comparative Psychology Monographs*, 13(4):1–113.

Houk, J. C., Davis, J. L., and Besier, D. G. (1995). *Models of Information Processing in the Basal Ganglia.* MIT Press, Cambridge MA.

Hull, C. L. (1943). *Principles of behavior.* Appelton-Century-Crofts, New York.

Hull, C. L. (1952). *A Behavior System: An Introduction to Behavior Theory Concerning the Individual Organism.* Yale University Press.

Hyvarinen, J. (1982). *The Parietal Cortex of Monkey and Man.* Springer-Verlag, New York.

Isaacson, R. L. (1974). *The Limbic System.* Plenum Press, New York. See also Second Edition, 1982.

Jacobs, L. F. (1992). Memory for cache locations in Merriam's kangaroo rats. *Animal Behaviour*, 43(4):585–593.

Jacobs, L. F. and Liman, E. R. (1991). Grey squirrels remember the locations of buried nuts. *Animal Behavior*, 41:103–110.

Jander, R. (1957). Die optische richtungsorientierung der roten waldameise (*formica rufa l.*). *Z. vergl. Physiol.*, 40:162–238. Cited in Maurer and Seguinot (1995).

Jarrard, L. E. (1993). On the role of the hippocampus in learning and memory in the rat. *Behavioral Neural Biology*, 60(1):9–26.

Jones, D. L. and Mogenson, G. J. (1980). Nucleus accumbens to globus pallidus GABA projection subserving ambulatory activity. *American Journal of Physiology*, 238:R65–R69.

Jones, R. S. G. (1993). Entorhinal-hippocampal connections: A speculative view of their function. *Trends in Neurosciences*, 16(2):58–64.

Jung, M. W. and McNaughton, B. L. (1993). Spatial selectivity of unit activity in the hippocampal granular layer. *Hippocampus*, 3(2):165–182.

Jung, M. W., Wiener, S. I., and McNaughton, B. L. (1994). Comparison of spatial firing characteristics of the dorsal and ventral hippocampus of the rat. *Journal of Neuroscience*, 14(12):7347–7356.

Keith, J. R. and McVety, K. M. (1988). Latent place learning in a novel environment and the influences of prior training in rats. *Psychobiology*, 16(2):146–151.

Kelly, J. B. (1990). Rat auditory cortex. In Kolb, B. and Tees, R. C., editors, *The Cerebral Cortex of the Rat*, chapter 15, pages 381–405. MIT Press, Cambridge MA.

Khatib, O. (1986). Real-time obstacle avoidance for manipulators and mobile robots. *International Journal of Robotics Research*, 5(1):90–98.

Kishimoto, K. and Amari, S. (1979). Existence and stability of local excitations in homogenous neural fields. *Journal of Mathematical Biology*, 7(4):303–318.

Kiyatkin, E. A. and Gratton, A. (1994). Electrochemical monitoring of extracellular dopamine in nucleus accumbens of rats lever-pressing for food. *Brain Research*, 652:225–234.

Knierim, J. J., Kudrimoti, H. S., and McNaughton, B. L. (1995). Place cells, head direction cells, and the learning of landmark stability. *Journal of Neuroscience*, 15:1648–1659.

Koerner, A., Thomas, M. J., Weisend, M. P., and Sutherland, R. J. (1996). Hippocampal-dependent memory consolidation: An evaluation of three hypotheses. *Society for Neuroscience Abstracts*, 22:1118.

Kohler, C. (1986). Intrinsic connections of the retrohippocampal region in the rat brain. II. The medial entorhinal area. *Journal of Comparative Neurology*, 246:149–169.

Kohler, C. (1988). Intrinsic connections of the retrohippocampal region in the rat brain. III. The lateral entorhinal area. *Journal of Comparative Neurology*, 271:208–228.

Kohonen, T. (1977). *Associative Memory: A System-Theoretical Approach*. Springer-Verlag, New York.

Kohonen, T. (1980). *Content-Addressable Memories.* Springer-Verlag, New York.

Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, 43:59–69.

Kohonen, T. (1984). *Self-Organization and Associative Memory.* Springer-Verlag, New York.

Kolb, B. (1990a). Posterior parietal and temporal association cortex. In Kolb, B. and Tees, R. C., editors, *The Cerebral Cortex of the Rat*, chapter 19, pages 459–471. MIT Press, Cambridge MA.

Kolb, B. (1990b). Prefrontal cortex. In Kolb, B. and Tees, R. C., editors, *The Cerebral Cortex of the Rat*, chapter 18, pages 437–458. MIT Press, Cambridge MA.

Kolb, B., Burhmann, K., McDonald, R., and Sutherland, R. J. (1994). Dissociation of the medial prefrontal, posterior parietal, and posterior temporal cortex for spatial navigation and recognition memory in the rat. *Cerebral Cortex*, 6:664–680.

Kolb, B. and Walkey, J. (1987). Behavioural and anatomical studies of the posterior parietal cortex in the rat. *Behavioral and Brain Research*, 23:127–145.

Krieg, W. J. S. (1946). Connections of the cerebral cortex. I. The albino rat: A. Topography of the cortical areas. *Journal of Comparative Neurology*, 84:221–275.

Kubie, J. L., Muller, R. U., and Bostock, E. (1990). Spatial firing properties of hippocampal theta cells. *Journal of Neuroscience*, 10(4):1110–1123.

Kubie, J. L. and Ranck, Jr., J. B. (1983). Sensory-behavioral correlates in individual hippocampus neurons in three situations: Space and context. In Seifert, W., editor, *Neurobiology of the Hippocampus*, chapter 22, pages 433–447. Academic Press Inc., New York.

Kudrimoti, H. S., Skaggs, W. E., Barnes, C. A., McNaughton, B. L., Gerrard, J. L., Suster, M. S., and Weaver, K. L. (1996). REM sleep and the reactivation of recent correlation patterns in hippocampal neuronal ensembles. *Society for Neuroscience Abstracts*, 22:1871.

Lanczos, C. (1970). *Space through the Ages: The Evolution of Geometrical Ideas from Pythagoras to Hilbert and Einstein.* Academic Press, New York.

Landfield, P. W. and Deadwyler, S. A., editors (1988). *Long-Term Pontentiation: From Biophysics to Behavior*, volume 35 of *Neurology and Neurobiology.* Alan R. Liss, Inc., New York.

Lavoie, A. M. and Mizumori, S. J. Y. (1994). Spatial-, movement- and reward-sensitive discharge by medial ventral striatum neurons in rats. *Brain Research*, 638:157–168.

Lefévre, P. and Galiana, H. L. (1992). Dynamic feedback to the superior colliculus in a neural network model of the gaze control system. *Neural Networks*, 5:871–890.

Leonard, B. and McNaughton, B. L. (1990). Spatial representation in the rat: Conceptual, behavioral, and neurophysiological perspectives. In Kesner, R. P. and Olton, D. S., editors, *Neurobiology of Comparative Cognition*, chapter 13, pages 363–422. Lawrence Earlbaum Associates, Hillsdale, NJ.

Leonhard, C. L., Stackman, R. W., and Taube, J. S. (1996). Head direction cells recorded from the lateral mammillary nuclei in rats. *Society for Neuroscience Abstracts*, 22:1873.

Levy, W. B. (1989). A computational approach to hippocampal function. In Hawkins, R. D. and Bower, G. H., editors, *Computational Models of Learning in Simple Neural Systems*, volume 23 of *The Psychology of Learning and Motivation*, pages 243–305. Academic Press, San Diego CA.

Levy, W. B. (1996). A sequence predicting CA3 is a flexible associator that learns and uses context to solve hippocampal-like tasks. *Hippocampus*, 6(6):579–591.

Levy, W. B. and Wu, X. (1996). The relationship of local context cues to sequence length memory capacity. *Network: Computation in Neural Systems*, 7:371–384.

Malenka, R. C. (1995). LTP and LTD: Dynamic and interactive processes of synaptic plasticity. *The Neuroscientist*, 1(1):35–42.

Mardia, K. V. (1972). *Statistics of Directional Data*. Academic Press, New York.

Margules, J. and Gallistel, C. R. (1988). Heading in the rat: Determination by environmental shape. *Animal Learning and Behavior*, 16(4):404–410.

Markus, E. J., Barnes, C. A., McNaughton, B. L., Gladden, V. L., and Skaggs, W. E. (1994). Spatial information content and reliability of hippocampal CA1 neurons: Effects of visual input. *Hippocampus*, 4:410–421.

Markus, E. J., Qin, Y., Leonard, B., Skaggs, W. E., McNaughton, B. L., and Barnes, C. A. (1995). Interactions between location and task affect the spatial and directional firing of hippocampal neurons. *Journal of Neuroscience*, 15:7079–7094.

Marr, D. (1969). A theory of cerebellar cortex. *Journal of Physiology*, 202:437–470. Reprinted in Marr (1991).

Marr, D. (1970). A theory for cerebral neocortex. *Proceedings of the Royal Society*, 176:161–234. Reprinted in Marr (1991).

Marr, D. (1971). Simple memory: A theory of archicortex. *Philosophical Transactions of the Royal Society of London*, 262(841):23–81. Reprinted in Marr (1991).

Marr, D. (1982). *Vision*. W. H. Freeman and Co., New York.

Marr, D. (1991). *From the Retina to the Neocortex: Selected Papers of David Marr*. Birkhäuser, Boston. Edited by L. M. Vaina.

Matthews, D. B., White, A. M., Brush, E. D., and Best, P. J. (1995). Construction of the spatial cognitive map does not require active exploration of the environment. *Society for Neuroscience Abstracts*, 21:2086.

Maurer, R. and Seguinot, V. (1995). What is modelling for? A critical review of the models of path integration. *Journal of Theoretical Biology*, 175:457–475.

McClelland, J. L. and Goddard, N. H. (1996). Considerations arising from a complementary learning systems perspective on hippocampus and neocortex. *Hippocampus*, 6(6):654–665.

McClelland, J. L., McNaughton, B. L., and O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, 102(3):419–457.

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

McHugh, T. J., Blum, K. I., Tsien, J. Z., Tonegawa, S., and Wilson, M. A. (1996). Impaired hippocampal represntation of space in CA1-specific NMDAR1 knockout mice. *Cell*, 87:1339–1349.

McNaughton, B. L. (1989). Neuronal mechanisms for spatial computation and information storage. In Nadel, L., Cooper, L., Culicover, P., and Harnish, R. M., editors, *Neural Connections, Mental Computation*, chapter 9, pages 285–350. MIT Press, Cambridge, MA.

McNaughton, B. L. (1993). The mechanism of expression of long-term enhancement of hippocampal synapses: Current issues and theoretical implications. *Annual Review of Physiology*, 55:375–396.

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. (1996). Deciphering the hippocampal polyglot: The hippocampus as a path integration system. *Journal of Experimental Biology*, 199(1):173–186.

McNaughton, B. L., Barnes, C. A., Meltzer, J., and Sutherland, R. J. (1989a). Hippocampal granule cells are necessary for normal spatial learning but not for spatially-selective pyramidal cell discharge. *Experimental Brain Research*, 76:485–496.

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

McNaughton, B. L., Chen, L. L., and Markus, E. J. (1991). "Dead reckoning," landmark learning, and the sense of direction: A neurophysiological and computational hypothesis. *Journal of Cognitive Neuroscience*, 3(2):190–202.

McNaughton, B. L., Knierim, J. J., and Wilson, M. A. (1994a). Vector encoding and the vestibular foundations of spatial cognition: Neurophysiological and computational mechanisms. In Gazzaniga, M., editor, *The Cognitive Neurosciences*, chapter 37, pages 585–595. MIT Press, Boston.

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

McNaughton, B. L., Mizumori, S. J. Y., Barnes, C. A., Leonard, B. J., Marquis, M., and Green, E. J. (1994b). Cortical representation of motion during unrestrained spatial navigation in the rat. *Cerebral Cortex*, 4(1):27–39.

McNaughton, B. L. and Morris, R. G. M. (1987). Hippocampal synaptic enhancement and information storage within a distributed memory system. *Trends in Neurosciences*, 10(10):408–415.

McNaughton, B. L. and Nadel, L. (1990). Hebb-Marr networks and the neurobiological representation of action in space. In Gluck, M. A. and Rumelhart, D. E., editors, *Neuroscience and Connectionist Theory*, chapter 1, pages 1–63. Lawrence Erlbaum Associates.

McNaughton, B. L., O'Keefe, J., and Barnes, C. A. (1983b). The stereotrode: A new technique for simultaneous isolation of several single units in the central nervous system from multiple unit records. *Journal of Neuroscience Methods*, 8:391–397.

Mehta, M. R. and McNaughton, B. L. (1996). Rapid changes in hippocampal population code during behavior: A case for Hebbian learning *in vivo*. Presented at *CNS*96*, (the fifth annual *Computation in Neural Systems* meeting).

Mehta, M. R., McNaughton, B. L., Barnes, C. A., Suster, M. S., Weaver, K. L., and Gerrard, J. L. (1996). Rapid changes in the hippocampal population code during behavior: A case for Hebbian learning *in vivo*. *Society for Neuroscience Abstracts*, 22:1872.

Miller, V. M. and Best, P. J. (1980). Spatial correlates of hippocampal unit activity are altered by lesions of the fornix and entorhinal cortex. *Brain Research*, 194:311–323.

Milner, B., Corkin, S., and Teuber, H. (1968). Further analysis of the hippocampal amnesia syndrome: 14-year follow-up study of H. M. *Neuropsychologia*, 6:215–234.

Minsky, M. and Papert, S. (1969). *Perceptrons*. MIT Press, Cambridge MA.

Mishkin, M. and Murray, E. A. (1994). Stimulus recognition. *Current Opinion in Neurobiology*, 4:200–206.

Mishkin, M., Ungerleider, L. G., and Macko, K. A. (1983). Object vision and spatial vision: Two cortical pathways. *Trends in Neurosciences*, 6:414–417.

Mittelstaedt, H. (1962). Control systems of orientation in insects. *Annual Review of Entemology*, 7:177–198. Cited in Maurer and Seguinot (1995).

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

Mittelstaedt, H. and Mittelstaedt, M.-L. (1982). Homing by path integration. In Papi, F. and Wallraff, H. G., editors, *Avian Navigation*, pages 290–297. Springer-Verlag, New York.

Mittelstaedt, M. and Glasauer, S. (1991). Idiothetic navigation in gerbils and humans. *Zool. Jb. Physiol.*, 95(427-435).

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

Mittlestaedt, H. (1985). Analytical cybernetics of spider navigation. In Barth, F., editor, *Neurobiology of Arachnids*. Springer-Verlag, Berlin. Cited in Maurer and Seguinot (1995).

Mizumori, S. J. Y., Barnes, C. A., and McNaughton, B. L. (1990). Behavioral correlates of theta-on and theta-off cells recorded from hippocampal formation of mature young and aged rats. *Experimental Brain Research*, 80:365–373.

Mizumori, S. J. Y. and Cooper, B. G. (1995). Spatial representations of dorsal caudate neurons of freely-behaving rats. *Society for Neuroscience Abstracts*, 21:1929.

Mizumori, S. J. Y., Lavoie, A. M., and Kalyani, A. (1996). Redistribution of spatial representation in the hippocampus of aged rats performing a spatial memory task. *Behavioral Neuroscience*, 110(5):1006–1016.

Mizumori, S. J. Y., McNaughton, B. L., Barnes, C. A., and Fox, K. B. (1989). Preserved spatial coding hippocampus CA1 pyramidal cells during reversible suppression in CA3c output: Evidence for pattern completion in hippocampus. *Journal of Neuroscience*, 9(11):3915–3928.

Mizumori, S. J. Y., Miya, D. Y., and Ward, K. E. (1994). Reversible inactivation of the lateral dorsal thalamus disrupts hippocampal place representation and impairs spatial learning. *Brain Research*, 644:168–174.

Mizumori, S. J. Y., Ward, K. E., and Lavoie, A. M. (1992). Medial septal modulation of entorhinal single unit activity in anesthetized and freely moving rats. *Brain Research*, 570:188–197.

Mizumori, S. J. Y. and Williams, J. D. (1993). Directionally selective mnemonic properties of neurons in the lateral dorsal nucleus of the thalamus of rats. *Journal of Neuroscience*, 13(9):4015–4028.

Mogenson, G. J. (1984). Limbic-motor integration – with emphasis on intiation of exploratory and goal-directed locamotion. In Bandler, R., editor, *Modulation of Sensorimotor Activity During Alterations in Behavioral States*, pages 121–138. Alan R. Liss Inc, New York.

Mogenson, G. J. and Nielsen, M. (1984). Neuropharmacological evidence to suggest that the nucleus accumbens and subpallidal region contribute to exploratory locomotion. *Behavioral and Neural Biology*, 42:52–60.

Moravec, H. P. (1988). Sensor fusion in certainty grids for mobile robots. *AI Magazine*, 9(2):61–74.

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

Morris, R. G. M., Garrud, P., Rawlins, J. N. P., and O'Keefe, J. (1982). Place navigation impaired in rats with hippocampal lesions. *Nature*, 297:681–683.

Morris, R. G. M., Schenk, F., Tweedie, F., and Jarrard, L. E. (1990). Ibotenate lesions of hippocampus and/or subiculum: Dissociating components of allocentric spatial learning. *European Journal of Neuroscience*, 2:1016–1028.

Müller, M. and Wehner, R. (1988). Path integration in desert ants, *cataglyphis fortis*. *Proceedings of the National Academy of Sciences USA*, 85:5287–5290.

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

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

Muller, R. U., Kubie, J. L., Bostock, E. M., Taube, J. S., and Quirk, G. J. (1991a). Spatial firing correlates of neurons in the hippocampal formation of freely moving rats. In Paillard, J., editor, *Brain and Space*, chapter 17, pages 296–333. Oxford University Press, New York.

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

Muller, R. U., Kubie, J. L., and Saypoff, R. (1991b). The hippocampus as a cognitive graph. *Hippocampus*, 1(3):243–246.

Muller, R. U., Stead, M., and Pach, J. (1996). The hippocampus as a cognitive graph. *Journal of General Physiology*, 107(6):663–694.

Munoz, D. P., Pélisson, D., and Guitton, D. (1991). Movement of neural activity on the superior colliculus motor map during gaze shifts. *Science*, 251:1358–1360.

Murphy, J. J. (1873). Instinct: A mechanical analogy. *Nature*, 7:483.

Murray, E. A. and Mishkin, M. (1996). 40-minute visual recognition memory in rhesus monkeys with hippocampal lesions. *Society for Neuroscience Abstracts*, 22:281.

Murray, J. D. (1989). *Mathematical Biology*, volume 19 of *Biomathematics*. Springer-Verlag, New York.

Myrher, T. (1991). Retroactive memory of a visual discrimination task in the rat: Role of temporal-entorhinal cortices and their connections. *Experimental Brain Research*, 84:517–524.

Nadel, L., editor (1991). *The hippocampus and space revisited, special issue*, volume 1(3) of *Hippocampus*. Churchill Livingstone, New York.

Nadel, L. (1994). Multiple memory systems: What and Why, an update. In Schacter, D. L. and Tulving, E., editors, *Memory Systems 1994*, chapter 2, pages 39–64. MIT Press, Cambridge MA.

Nadel, L. (1995). The role of the hippocampus in declarative memory: A commentary on Zola-Morgan, Squire, and Ramus 1994. *Hippocampus*, 5:232–234.

Nadel, L. and Willner, J. (1980). Context and conditioning: A place for space. *Physiological Psychology*, 8(2):218–228.

Nagahara, A. H., Otto, T., and Gallagher, M. (1995). Entorhinal-perirhinal lesions impair performance of rats on two versions of place learning in the Morris water maze. *Behavioral Neuroscience*, 109(1):3–9.

Nature (1873). Perception and instinct in the lower animals. *Nature* 7:377–378.

Navigation, Encyclopedia Britannica (1994). volume 8, page 566. Encyclopedia Britannica Inc., Chicago IL, 15th edition.

Nilsson, O. G., Shapiro, M. L., Gage, F. H., Olton, D. S., and Bjorklund, A. (1987). Spatial learning and memory following fimbria-fornix transection and grafting of fetal septal neurons to the hippocampus. *Experimental Brain Research*, 67:195–215.

Oatley, K. (1974). Mental maps for navigation. *New Scientist*, pages 863–866. Reprinted in Farrall (1984), pages 58–65.

Ohta, H., Matsumoto, K., and Watanabe, H. (1993). The interaction between central cholinergic and peripheral beta-adrenergic systems on radial maze performance in rats. *Brain Research*, 622(1–2):353–356.

O'Keefe, J. (1976). Place units in the hippocampus of the freely moving rat. *Experimental Neurology*, 51:78–109.

O'Keefe, J. (1989). Computations the hippocampus might perform. In Nadel, L., Cooper, L., Culicover, P., and Harnish, R. M., editors, *Neural Connections, Mental Computation*, chapter 8, pages 225–284. MIT Press, Cambridge, MA.

O'Keefe, J. and Burgess, N. (1996). Geometric determinants of the place fields of hippocampal neurons. *Nature*, 381:425–428.

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

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

O'Keefe, J. and Nadel, L. (1978). *The Hippocampus as a Cognitive Map.* Clarendon Press, Oxford.

O'Keefe, J. and Nadel, L. (1979). Précis of o'keefe and nadel's *the hippocampus as a cognitive map*. *Behavioral and Brain Sciences*, 2:487–494. See also commentary and response, pages 494–533.

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

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

Olton, D. S., Becker, J. T., and Handelmann, G. E. (1979). Hippocampus, space, and memory. *Behavioral and Brain Sciences*, 2:313–322. See also commentary and response, pages 323–366.

Olton, D. S., Becker, J. T., and Handelmann, G. E. (1980). Hippocampal function: Working memory or cognitive mapping? *Physiological psychology*, 8(2):239–246.

Olton, D. S., Branch, M., and Best, P. J. (1978). Spatial correlates of hippocampal unit activity. *Experimental Neurology*, 58:387–409.

Olton, D. S. and Samuelson, R. J. (1976). Remembrance of places passed: Spatial memory in rats. *Journal of Experimental Psychology: Animal Behavior Processes*, 2(2):97–116.

O'Reilly, R. C. and McClelland, J. L. (1994). Hippocampal conjunctive encoding, storage, and recall: Avoiding a trade-off. *Hippocampus*, 4(6):661–682.

Otto, T., Ding, C., Cousens, G., and Schiller, K. (1996). Effects of lateral vs. medial entorhinal cortex aspiration on the acquisition of odor-place associations. *Society for Neuroscience Abstracts*, 22:1120.

Otto, T. and Eichenbaum, H. (1992). Neuronal activity in the hippocampus during delayed non-match to sample performance in rats: Evidence for hippocampal processing in recognition memory. *Hippocampus*, 2(3):323–334.

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

Packard, M. G. and McGaugh, J. L. (1996). Inactivation of hippocampus or caudate nucleus with lidocaine differntialyl affects expression of place and response learning. *Neurobiology of Learning and Memory*, 65:65–72.

Papez, J. W. (1937). A proposed mechanism of emotion. *Archives of Neurology and Psychiatry*, 38:728–744.

Pavlides, C. and Winson, J. (1989). Influences of hippocampal place cell firing in the awake state on the activity of these cells during subsequent sleep episodes. *Journal of Neuroscience*, 9(8):2907–2918.

Phillips, R. G. and Le Doux, J. E. (1992). Differential contribution of amygdala and hippocampus to cued and contextual fear conditioning. *Behavioral Neuroscience*, 106(2):274–285.

Phillips, R. G., Tanila, H., and Eichenbaum, H. (1996). Differences in the spatial coding patterns of ventral subicular neurons in simple and complex environments. *Society for Neuroscience Abstracts*, 22:679.

Pico, R. M., Gerbrandt, L. K., Pondel, M., and Ivy, G. (1985). During stepwise cue deletion, rat place behaviors correlate with place unit responses. *Brain Research*, 330:369–372.

Pinto, D. J., Brumberg, J. C., Simons, D. J., and Ermentrout, G. B. (1996). A quantitative population model of whisker barrels: Re-examining the Wilson-Cowan equations. *Journal of Computational Neuroscience*, 3(3):247–264.

Plaznik, A., Danysz, W., and Kostowski, W. (1983). Some behavioral effects of microinjections of noradrenaline and serotonin into the hippocampus of the rat. *Physiological Behavior*, 31:625–631.

Potegal, M. (1982). Vestibular and neostriatal contributions to spatial orientation. In Potegal, M., editor, *Spatial Abilities: Development and Physiological Foundations*, chapter 15, pages 361–387. Academic Press, New York.

Poucet, B. (1993). Spatial cognitive maps in animals: New hypotheses on their structure and neural mechanisms. *Psychological Review*, 100(2):163–182.

Poucet, B., Chapuis, N., Durup, M., and Thinus-Blanc, C. (1986). A study of exploratory behavior as an index of spatial knowledge in hamsters. *Animal Learning and Behavior*, 14(1):93–100.

Poucet, B., Thinus-Blanc, C., and Muller, R. U. (1994). Place cells in the ventral hippocampus of rats. *NeuroReport*, 5(16):2045–2048.

Pouget, A. and Sejnowski, T. J. (1995). Spatial representations in parietal cortex may use basis functions. In Tesauro, G., Touretzky, D., and Leen, T., editors, *Advances in Neural Information Processing Systems 7*. MIT Press.

Pratt, I. (1991a). An algorithm for planning "sensible" routes. *Engineering Applications of Artificial Intelligence*, 4(2):97–108.

Pratt, I. (1991b). Path finding in free space using sinusoidal transforms. In Frank, A. and Mark, D., editors, *Cognitive and Linguistic Aspects of Geographic Space*, pages 219–233. Kluwer, Boston.

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

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

Rajkowski, J., Kubiak, P., and Aston-Jones, G. (1994). Locus coeruleus activity in monkey: Phasic and tonic changes are associated with altered vigilance. *Brain Research Bullettin*, 35(5-6):607–616.

Ranck, Jr., J. B. (1984). Head-direction cells in the deep cell layers of dorsal presubiculum in freely moving rats. *Society for Neuroscience Abstracts*, 10:599.

Rasmussen, M., Barnes, C. A., and McNaughton, B. L. (1989). A systematic test of cognitive mapping, working-memory, and temporal discontiguity theories of hippocampal function. *Psychobiology*, 17:335–348.

Rawlins, J. N. P. (1985). Associations across time: The hippocampus as a temporary memory store. *Behavioral and Brain Sciences*, 8:479–496. See also commentary and response, pages 497–528.

Recce, M. and Harris, K. D. (1996). Memory for places: A navigational model in support of Marr's theory of hippocampal function. *Hippocampus*, 6:735–748.

Redish, A. D., Elga, A. N., and Touretzky, D. S. (1996). A coupled attractor model of the rodent head direction system. *Network: Computation in Neural Systems*, 7(4):671–685.

Redish, A. D. and Touretzky, D. S. (1994). The reaching task: Evidence for vector arithmetic in the motor system? *Biological Cybernetics*, 71:307–317.

Redish, A. D. and Touretzky, D. S. (1995). Revisiting the Papez circuit: The role of hippocampus and its afferent and efferent structures in rodent navigation. *Society for Neuroscience Abstracts*, 21.

Redish, A. D. and Touretzky, D. S. (1996a). Details of a comprehensive theory of rodent navigation. *Society for Neuroscience Abstracts*, 22:678.

Redish, A. D. and Touretzky, D. S. (1996b). Modeling interactions of the rat's place and head direction systems. In Touretzky, D. S., Mozer, M. C., and Hasselmo, M. E., editors, *Advances in Neural Information Processing Systems 8*, pages 61–71.

Redish, A. D. and Touretzky, D. S. (1997a). Cognitive maps beyond the hippocampus. *Hippocampus*, 7(1):15–35.

Redish, A. D. and Touretzky, D. S. (1997b). Navigating with landmarks: Computing goal locations from place codes. In Ikeuchi, K. and Veloso, M., editors, *Symbolic Visual Learning*. Oxford University Press. In press.

Redish, A. D. and Touretzky, D. S. (1997c). The role of the hippocampus in solving the Morris water maze. *Neural Computation*. In press.

Redish, A. D. and Touretzky, D. S. (1997d). The role of the hippocampus in solving the morris water maze. Presented at *CNS*97*, (the sixth annual *Computation in Neural Systems* meeting).

Rempel-Clower, N. L., Zola, S. M., Squire, L. R., and Amaral, D. G. (1996). Three cases of enduring memory impairment after bilateral damage limited to the hippocampal formation. *Journal of Neuroscience*, 16:5233–5255.

Rescorla, R. A. and Wagner, A. R. (1972). A theory of Pavlovian conditiong: Variations in the effectiveness of reinforcement and nonreinforcement. In Black, A. H. and Prokesy, W. F., editors, *Classical Conditioning II: Current Research and Theory*, pages 64–99. Appelton Century Crofts, New York.

Rolls, E. T. (1989). The representation and storage of information in neuronal networks in the primate cerebral cortex and hippocampus. In Durbin, R., Miall, C., and Mitchison, G., editors, *The Computing Neuron*, chapter 8, pages 125–159. Addison-Wesley, Reading, MA.

Rolls, E. T. (1996). A theory of hippocampal function in memory. *Hippocampus*, 6:601–620.

Rotenberg, A., Mayford, M., Hawkins, R. D., Kandel, E. R., and Muller, R. U. (1996). Mice expressing activated CaMKII lack low frequency LTP and do not form stable place cells in the CA1 region of the hippocampus. *Cell*, 87:1351–1361.

Rumelhart, D. E., Hinton, G. E., and McClelland, J. L. (1986a). A general framework for parallel distributed processing. In Rumelhart, D. E. and McClelland, J. L., editors, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, volume 1, chapter 2, pages 45–76. MIT Press, Cambridge MA.

Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986b). Learning internal representations by error propogation. In Rumelhart, D. E. and McClelland, J. L., editors, *Parallel Distribted Processing: Explorations in the Microstructure of Cognition*, volume 1, chapter 8, pages 318–362. MIT Press.

Sacks, O. (1985). Lost mariner. In *The Man who Mistook his Wife for a Hat*, chapter 2, pages 22–41. Summitt Books, New York NY.

Saksida, L. M., Redish, A. D., Milberg, C. R., Gaulin, S. J., and Touretzky, D. S. (1995). Landmark-based navigation in gerbils supports vector voting. *Society for Neuroscience Abstracts*, 21.

Samsonovich, A. and McNaughton, B. L. (1996). Attractor-map-based path integrator model of the hippocampus reproduces the phase precession phenomenon. *Society for Neuroscience Abstracts*, 22:1872.

Samsonovich, A. and McNaughton, B. L. (1997). Path integration and cognitive mapping in a continuous attractor neural network model. *Journal of Neuroscience*, 17(15):5900–5920.

Samsonovich, A. V. (1997). *Attractor map theory of the hippocampal representation of space*. PhD thesis, Graduate Interdisciplinary Program in Applied Mathematics, University of Arizona.

Sara, S. J. and Segal, M. (1991). Plasticity of sensory responses of locus coeruleus neurons in the behaving rat: Implications for cognition. *Progress in Brain Research*, 88:571–585.

Save, E., Cressant, A., Thinus-Blanc, C., and Poucet, B. (1996). Early visual deprivation does not prevent hippocampal place cell firing in the rat. *Society for Neuroscience Abstracts*, 22:912.

Schallert, T., Day, L. B., Weisend, M., and Sutherland, R. J. (1996). Spatial learning by hippocampal rats in the Morris water task. *Society for Neuroscience Abstracts*, 22:678.

Schenk, F. and Morris, R. G. M. (1985). Dissociation between components of spatial memory in rats after recovery from the effects of retrohippocampal lesions. *Experimental Brain Research*, 58:11–28.

Schmajuk, N. A. and Blair, H. T. (1993). Place learning and the dynamics of spatial navigation: A neural network approach. *Adaptive Behavior*, 1:355–387.

Schölkopf, B. and Mallot, H. A. (1993). View-based cognitive mapping and path planning. Technical Report Technical Report 7, Max-Plank-Institut für Biologische Kybernetic.

Schöne, H. (1984). *Spatial Orientation.* Princeton, Princeton NJ. Translated by C. Strausfeld.

Schultz, W. (1997). Dopamine neurons and their role in reward mechanisms. *Current Opinion in Neurobiology*, 7:191–197.

Schultz, W., Apicella, P., Romo, R., and Scarnati, E. (1995). Context-dependent activity in primate striatum reflecting past and future behavioral events. In Houk, J. C., Davis, J. L., and Besier, D. G., editors, *Models of Information Processing in the Basal Ganglia*, chapter 2, pages 11–27. MIT Press, Cambridge MA.

Scoville, W. B. (1968). Amnesia after bilateral medial temporal-lobe excision: Introduction to case H. M. *Neuropsychologia*, 6:211–213.

Scoville, W. B. and Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery, and Psychiatry*, 20:11–21.

Seamans, J. K. and Phillips, A. G. (1994). Selective memory impairments produced by transient lidocaine-induced lesions of the nucleus accumbens in rats. *Behavioral Neuroscience*, 108(3):456–468.

Seguinot, V., Maurer, R., and Etienne, A. S. (1993). Dead reckoning in a small mammal: The evaluation of distance. *Journal of Comparative Physiology A*, 173:103–113.

Seifert, W., editor (1983). *Neurobiology of the Hippocampus.* Academic Press, London.

Selden, N. R., Everitt, B. J., Jarrard, L. E., and Robbins, T. W. (1991). Complementary roles for the amygdala and hippocampus in aversive conditioning to explicit and contextual cues. *Neuroscience*, 42(2):335–350.

Sesack, S. R., Deutch, A. Y., Roth, R. H., and Bunney, B. S. (1989). Topographical organization of the efferent projections of the medial prefrontal cortex in the rat: An anterograde tract-tracing study with *phaseolus vulgaris leucoagglutinin. Journal of Comparative Neurology*, 290(2):213–242.

Sesack, S. R. and Pickel, V. M. (1990). In the rat medial nucleus accumbens, hippocampal and catecholaminergic terminals converge on spiny neurons and are in apposition to each other. *Brain Research*, 527(2):266–279.

Shapiro, M. L. and Hetherington, P. A. (1993). A simple network model simulates hippocampal place fields: 1. Parametric analyses and physiological predictions. *Behavioral Neuroscience*, 107(1):34–50.

Shapiro, M. L., Simon, D. K., Olton, D. S., Gage III, F. H., Nilsson, O., and Bjorklund, A. (1989). Intrahippocampal grafts of fetal basal forebrain tissue alter place fields in the hippocampus of rats with fimbria-fornix lesions. *Neuroscience*, 32(1):1–18.

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

Sharp, P. E. (1996a). Multiple spatial/behavioral correlates for cells in the rat postsubiculum: Multiple regression analysis and comparison to other hippocampal areas. *Cerebral Cortex*, 6(2):238–259.

Sharp, P. E. (1996b). Subicular cells generate similar firing patterns in two geometrically and visually distinctive environments: Comparison with hippocampal place cells. *Behavioral Brain Research*. In press.

Sharp, P. E., Blair, H. T., and Brown, M. (1996). Neural network modeling of the hippocampal formation spatial signals and their possible role in navigation: A modular approach. *Hippocampus*, 6(6):735–748.

Sharp, P. E., Blair, H. T., Etkin, D., and Tzanetos, D. B. (1995). Influences of vestibular and visual motion information on the spatial firing patterns of hippocampal place cells. *Journal of Neuroscience*, 15(1):173–189.

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

Sharp, P. E., Kubie, J. L., and Muller, R. U. (1990). Firing properties of hippocampal neurons in a visually symmetrical environment: Contributions of multiple sensory cues and mnemonic processes. *Journal of Neuroscience*, 10(9):3093–3105.

Shen, B. and McNaughton, B. L. (1996). Modeling the spontaneous reactivation of experience-specific hippocampal cell assembles during sleep. *Hippocampus*, 6(6):685–693.

Shen, J., Barnes, C. A., Wenk, G. L., and McNauyghton, B. L. (1996). Differential effects of selective immunotoxic lesions of medial septal cholinergic cells on spatial working and reference memory. *Behavioral Neuroscience*, 110(5):1181–1186.

Sherry, D. F. and Duff, S. J. (1996). Behavioral and neural bases of orientation in food storing birds. *Journal of Experimental Biology*, 199(1):165–172.

Shorth, Encyclopedia of Statistical Sciences (1988). volume 8, page 448. John Wiley and Sons, Inc., New York.

Silva, A. J., Paylor, R., Wehner, J. M., and Tonegawa, S. (1992). Impaired spatial learning in $\alpha$-calcium-calmodulin kinase II mutant mice. *Science*, 257:206–211.

Skaggs, W. E. (1995). *Relations Between the Theta Rhythm and Activity Patterns of Hippocampal Neurons*. PhD thesis, University of Arizona.

Skaggs, W. E., Knierim, J. J., Kudrimoti, H. S., and McNaughton, B. L. (1995). 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*, pages 173–180. MIT Press.

Skaggs, W. E. and McNaughton, B. L. (1996). Replay of neuronal firing sequences in rat hippocampus during sleep following spatial experience. *Science*, 271:1870–1873.

Skaggs, W. E., McNaughton, B. L., Wilson, M. A., and Barnes, C. A. (1996). Theta phase precession in hippocampal neuronal populations and the compression of temporal sequences. *Hippocampus*, 6(2):149–173.

Smith, C. (1995). Sleep states and memory processes. *Behavioural Brain Research*, 69:137–145.

Squire, L. R. (1992). Memory and the hippocampus: A synthesis from findings with rats, monkeys, and humans. *Psychology Review*, 99(2):195–231.

Squire, L. R. and Alvarez, P. (1995). Retrograde amnesia and memory consolidation: A neurobiological perspective. *Current Opinion in Neurobiology*, 5:169–177.

Squire, L. R. and Zola-Morgan, S. (1988). Memory: Brain systems and behavior. *Trends in Neurosciences*, 11(4):170–175.

Squire, L. R. and Zola-Morgan, S. M. (1991). The medial temporal lobe memory system. *Science*, 253:1380–1386.

Sripanidkulchai, K. and Wyss, J. M. (1986). Thalamic projections to retrosplenial cortex in the rat. *Journal of Comparative Neurology*, 254:143–165.

Stackman, R. W. and Taube, J. S. (1997). Firing properties of head direction cells in the rat anterior thalamic nucleus: Dependence upon vestibular input. Under editorial review.

Staubli, U., Ivy, G., and Lynch, G. (1984). Hippocampal denervation causes rapid forgetting of olfactory information in rats. *Proceedings of the National Academy of Sciences, USA*, 81:5885–5887.

Stein, B. E. and Meredith, M. A. (1993). *The Merging of the Senses.* MIT Press, Cambridge MA.

Stein, J. F. (1991). Space and the parietal areas. In Paillard, J., editor, *Brain and Space*, chapter 11, pages 185–222. Oxford University Press, New York.

Stein, J. F. (1992). The representation of egocventric space in the posterior parietal cortex. *Brain and Behavioral Sciences*, 15:691–700.

Stewart, M. and Fox, S. E. (1990). Do septal neruons pace the hippocampal theta rhythm? *Trends in Neurosciences*, 13(5):163–168.

Stricanne, B., Andersen, R. A., and Mazzoni, P. (1996). Eye-centered, head-centered, and intermediate coding of remembered sound locations in area LIP. *Journal of Neurophysiology*, 76(3):2071–2076.

Sutherland, R. J., Chew, G. L., Baker, J. C., and Linggard, R. C. (1987). Some limitations on the use of distal cues in place navigation by rats. *Psychobiology*, 15(1):48–57.

Sutherland, R. J. and Hoesing, J. M. (1993). 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*, pages 461–477. Birkhauser, Boston.

Sutherland, R. J. and Rodriguez, A. J. (1990). The role of the fornix/fimbria and some related subcortical structures in place learning and memory. *Behavioral and Brain Research*, 32:265–277.

Sutherland, R. J. and Rudy, J. W. (1989). Configural association theory: The role of the hippocampal formation in learning, memory, and amnesia. *Psychobiology*, 17(2):129–144.

Sutherland, R. J., Whishaw, I. Q., and Kolb, B. (1983). A behavioral analysis of spatial localization following electrolytic, kainate-, or colchicine-induced damage to the hippocampal formation in the rat. *Behavioral Brain Research*, 7:133–153.

Sutherland, R. J., Whishaw, I. Q., and Kolb, B. (1988). Contributions of cingulate cortex to two forms of spatial learning and memory. *Journal of Neuroscience*, 8(6):1863–1872.

Sutherland, R. J., Whishaw, I. Q., and Regehr, J. C. (1982). Cholinergic receptor blockade impairs spatial localization by use of distal cues in the rat. *Journal of Comparative and Physiological Psychology*, 96(4):563–573.

Sutton, R. S., editor (1992). *Special issue on reinforcement learning*, volume 8(3/4) of *Machine Learning*. Kluwer Academic Publishers, Boston.

Suzuki, S., Augerinos, G., and Black, A. H. (1980). Stimulus control of spatial behavior on the eight-arm maze in rats. *Learning and Motivation*, 11:1–18.

Tanaka, K., Saito, H., Fukada, Y., and Moriya, M. (1991). Coding visual images of objects in the inferotemporal cortex of the macaque monkey. *Journal of neurophysiology*, 66(1):170–189.

Tanila, H., Shaprio, M., Gallagher, M., and Eichenbaum, H. (1997a). Brain aging: Changes in the nature of information coding by the hippocampus. *Journal of Neuroscience*, 17(13):5155–5166.

Tanila, H., Sipilä, P., Shapiro, M., and Eichenbaum, H. (1997b). Brain aging: Impaired coding of novel environmental cues. *Journal of Neuroscience*, 17(13):5167–5174.

Tarassenko, L. and Blake, A. (1991). Analogue computation of collision-free paths. *Proceedings of the 1991 IEEE International Conference on Robotics and Automation*, pages 540–545.

Taube, J. S. (1995). Head direction cells recorded in the anterior thalamic nuclei of freely moving rats. *Journal of Neuroscience*, 15(1):1953–1971.

Taube, J. S. (1996). Place cells recorded in the parasubiculum of freely moving rats. *Hippocampus*, 5(6):569–583.

Taube, J. S. and Burton, H. L. (1995). Head direction cell activity monitored in a novel environment and during a cue conflict situation. *Journal of Neurophysiology*, 74(5):1953–1971.

Taube, J. S., Goodridge, J. P., Golob, E. J., Dudchenko, P. A., and Stackman, R. W. (1996). Processing the head direction cell signal: A review and commentary. *Brain Research Bulletin*, 40(5/6):477–486.

Taube, J. S., Klesslak, J. P., and Cotman, C. W. (1992). Lesions of the rat postsubiculum impair performance on spatial tasks. *Behavioral and Neural Biology*, 5:131–143.

Taube, J. S., Muller, R. I., and Ranck, Jr., J. B. (1990a). Head direction cells recorded from the postsubiculum in freely moving rats. I. Description and quantitative analysis. *Journal of Neuroscience*, 10:420–435.

Taube, J. S., Muller, R. I., and Ranck, Jr., J. B. (1990b). Head direction cells recorded from the postsubiculum in freely moving rats. II. Effects of environmental manipulations. *Journal of Neuroscience*, 10:436–447.

Taube, J. S. and Muller, R. U. (1995). Head direction cell activity in the anterior thalamic nuclei, but not the postsubiculum, predicts the animal's future directional heading. *Society for Neuroscience Abtracts*, 21:946.

Thinus-Blanc, C., Bouzouba, L., Chaix, K., Chapuis, N., Durup, M., and Poucet, B. (1987). A study of spatial parameters encoded during exploration in hamsters. *Journal of Experimental Psychology*, 13(4):418–427.

Thinus-Blanc, C., Durup, M., and Poucet, B. (1992). The spatial parameters encoded by hamsters during exploration: A further study. *Behavioural Processes*, 26:43–57.

Thinus-Blanc, C., Save, E., Buhot, M.-C., and Poucet, B. (1991). The hippocampus, exploratory activity, and spatial memory. In Paillard, J., editor, *Brain and Space*, pages 334–352. Oxford University Press, New York.

Thompson, L. T. and Best, P. J. (1989). Place cells and silent cells in the hippocampus of freely-behaving rats. *Journal of Neuroscience*, 9(7):2382–2390.

Thompson, L. T. and Best, P. J. (1990). Long-term stability of the place-field activity of single units recorded from the dorsal hippocampus of freely behaving rats. *Brain Research*, 509(2):299–308.

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review*, 55:189–208.

Tolman, E. C., Ritchie, B. F., and Kalish, D. (1946a). Studies in spatial learning. I. Orientation and the short-cut. *Journal of Experimental Psychology*, 36:13–24.

Tolman, E. C., Ritchie, B. F., and Kalish, D. (1946b). Studies in spatial learning. II. Place learning versus response learning. *Journal of Experimental Psychology*, 36:221–229.

Touretzky, D. S., Gaulin, S. J. C., and Redish, A. D. (1996). Gerbils regularly return to their entry point when exploring a novel environment. *Society for Neuroscience Abstracts*, 22:449.

Touretzky, D. S. and Redish, A. D. (1996). A theory of rodent navigation based on interacting representations of space. *Hippocampus*, 6(3):247–270.

Touretzky, D. S., Redish, A. D., and Wan, H. S. (1993). Neural representation of space using sinusoidal arrays. *Neural Computation*, 5(6):869–884.

Touretzky, D. S., Wan, H. S., and Redish, A. D. (1994). Neural representation of space in rats and robots. In Zurada, J. M., II, R. J. M., and Robinson, C. J., editors, *Computational Intelligence: Imitating Life*, pages 57–68. IEEE Press.

Treves, A., Miglino, O., and Parisi, D. (1992). Rats, nets, maps, and the emergence of place cells. *Psychobiology*, 20(1):1–8.

Trullier, O., Wiener, S. I., Berthoz, A., and Meyer, J. A. (1997). Biologically based artificial navigation systems: Review and prospects. *Progress in Neurobiology*, 51(5):483–544.

Tsien, J. Z., Chen, D. F., Gerber, D., Tom, C., Mercer, E. H., Anderson, D. J., Mayford, M., and Kandel, E. R. (1996a). Subregion- and cell type-restricted gene knockout in mouse brain. *Cell*, 87:1317–1326.

Tsien, J. Z., Huerta, P. T., and Tonegawa, S. (1996b). The essential role of hippocampal CA1 NMDA receptor dependent synaptic plasticity in spatial memory. *Cell*, 87:1327–1338.

Tsodyks, M. V., Skaggs, W. E., Sejnowski, T. J., and McNaughton, B. L. (1996). Population dynamics and theta rhythm phase precession of hippocampal place cell firing: A spiking neuron model. *Hippocampus*, 6(3):271–280.

Ungerleider, L. G. and Haxby, J. V. (1994). 'What' and 'where' in the human brain. *Current Opinion in Neurobiology*, 4:157–165.

van Groen, T., Vogt, B. A., and Wyss, J. M. (1993). Interconnections between the thalamus and retrosplenial cortex in the rodent brain. In Vogt, B. A. and Gabriel, M., editors, *Neurobiology of Cingulate Cortex and Limbic Thalamus: A Comprehensive Handbook*, pages 123–150. Birkhauser, Boston.

van Groen, T. and Wyss, J. M. (1990). The postsubicular cortex in the rat: Characterization of the fourth region of the subicular cortex and its connections. *Brain Research*, 529:165–177.

van Opstal, A. J. and Kappen, H. (1993). A two-dimensional ensemble coding model for spatial-temporal transformation of saccades in monkey superior colliculus. *Network: Computation in Neural Systems*, 4:19–38.

Vander Wall, S. B. (1990). *Food Hoarding in Animals*. University of Chicago Press, Chicago.

Vanderwolf, C. H. (1971). Limbic-diencephalic mechanisms of voluntary movement. *Psychological Review*, 78(2):83–113.

Vanderwolf, C. H. (1990). An introduction to the electrical activity of the cerebral cortex: Relations to behavior and control by subcortical inputs. In Kolb, B. and Tees, R. C., editors, *The Cerebral Cortex of the Rat*, chapter 6, pages 151–189. MIT Press, Cambridge MA.

Vanderwolf, C. H. and Leung, L.-W. S. (1983). Hippocampal rhythmical slow activity: A brief history and the effects of entorhinal lesions and phencyclidine. In Seifert, W., editor, *Neurobiology of the Hippocampus*, pages 275–302. Academic Press, New York.

Vogt, B. A. (1985). Cingulate cortex. In Peters, A. and Jones, E., editors, *Association and Auditory Cortices*, volume 4 of *Cerebral Cortex*, pages 89–150. Plenum Press, New York.

von Saint Paul, U. (1982). Do geese use path integration for walking home? In Papi, F. and Wallraff, H. G., editors, *Avian Navigation*, pages 298–307. Springer-Verlag, New York.

Wallace, A. R. (1873a). Inherited feeling. *Nature*, 7:303.

Wallace, A. R. (1873b). Perception and instinct in the lower animals. *Nature*, 8:65–66.

Wan, H. S., Touretzky, D. S., and Redish, A. D. (1994a). Computing goal locations from place codes. In *Proceedings of the 16th Annual Conference of the Cognitive Science Society*, pages 922–927. Lawrence Earlbaum Associates, Hillsdale NJ.

Wan, H. S., Touretzky, D. S., and Redish, A. D. (1994b). A rodent navigation model that combines place code, head direction, and path integration information. *Society for Neuroscience Abstracts*, 20:1205.

Wan, H. S., Touretzky, D. S., and Redish, A. D. (1994c). Towards a computational theory of rat navigation. In Mozer, M., Smolensky, P., Touretzky, D., Elman, J., and Weigend, A., editors, *Proceedings of the 1993 Connectionist Models Summer School*, pages 11–19. Lawrence Earlbaum Associates, Hillsdale NJ.

Watson, J. B. (1907). Kinaesthetic and organic sensations: Their role in the reactions of the white rat to the maze. *Psychological Review*, 8(2):43–100.

Weaver, K. L., Shen, J., McNaughton, B. L., Barnes, C. A., Skaggs, W. E., Suster, M. S., and Gerrard, J. L. (1996). Theta phase precession in hippocampal pyramidal cells does not depend on visual input. *Society for Neuroscience Abstracts*, 22:1871.

Wehner, R. and Srinivasan, M. V. (1981). Searching behavior of desert ants, genus *cataglyphis* (Formicidae, Hymenera). *Journal of Comparative Physiology A*, 142:315–338.

Weisend, M. P., Astur, R. S., and Sutherland, R. J. (1996). The specificity and temporal characteristics of retrograde amnesia after hippocampal lesions. *Society for Neuroscience Abstracts*, 22:1118.

Whishaw, I. Q. (1985). Cholinergic receptor blockade in the rat impairs locale but not taxon strategies for place navigation in a swimming pool. *Behavioral Neuroscience*, 99(5):979–1005.

Whishaw, I. Q., Cassel, J.-C., and Jarrard, L. E. (1995). 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):5779–5788.

Whishaw, I. Q. and Jarrard, L. E. (1996). Evidence for extrahippocampal involvement in place learning and hippocampal involvement in path integration. *Hippocampus*, 6:513–524.

Whishaw, I. Q. and Mittleman, G. (1986). Visits to starts, routes, and places by rats (*rattus norvegicus*). *Journal of Comparative Psychology*, 100(4):422–431.

Whishaw, I. Q. and Mittleman, G. (1991). Hippocampal modulation of nucleus accumbens: Behavioral evidence from amphetamine-induced activity profiles. *Behavioral and Neural Biology*, 55:289–306.

White, N. M. (1997). Mnemonic functions of the basal ganglia. *Current Opinion in Neurobiology*, 7:164–169.

Whiting, P. D. and Hillier, J. A. (1960). A method for finding the shortest route through a road network. *Operational Research Quarterly*, 11:37–40.

Wiener, S. I. (1993). Spatial and behavioral correlates of striatal neurons in rats performing a self-initiated navigation task. *Journal of Neuroscience*, 13(9):3802–3817.

Wiener, S. I., Paul, C. A., and Eichenbaum, H. (1989). Spatial and behavioral correlates of hippocampal neuronal activity. *Journal of Neuroscience*, 9(8):2737–2783.

Wilson, H. R. and Cowan, J. D. (1972). Excitatory and inhibitory interactions in localized populations of model neurons. *Biophysical Journal*, 12(1):1–24.

Wilson, H. R. and Cowan, J. D. (1973). A mathematical theory of the functional dynamics of cortical and thalamic tissue. *Kybernetik*, 13:55–80.

Wilson, M. A. and McNaughton, B. L. (1993). Dynamics of the hippocampal ensemble code for space. *Science*, 261:1055–1058.

Wilson, M. A. and McNaughton, B. L. (1994). Reactivation of hippocampal ensemble memories during sleep. *Science*, 265:676–679.

Winocur, G. (1990). Anterograde and retrograde amnesia in rats with dorsal hippocampal or dorsomedial thalamic lesions. *Behavioural Brain Research*, 38:145–154.

Wirsching, B. A., Beninger, R. J., Jhamandas, K., Boegman, R. J., and El-Defrawy, S. R. (1984). Differential effects of scopolomine on working and reference memory of rats in the radial maze. *Pharmacology Biochemistry and Behavior*, 20:659–662.

Witter, M. P. (1993). Organization of the entorhinal-hippocampal system: A review of current anatomical data. *Hippocampus*, 3:33–44.

Witter, M. P., Ostendorf, R. H., and Groenwegen, H. J. (1990). Heterogeneity in the dorsal subiculum of the rat. Distinct neuronal zones project to different cortical and subcortical targets. *European Journal of Neuroscience*, 2:718–725.

Wittmann, T. and Schwegler, H. (1995). Path integration – a network model. *Biological Cybernetics*, 73(6):569–575.

Wolske, M., Rompre, P. P., Wise, R. A., and West, M. O. (1993). Activation of single neurons in the rat nucleus accumbens during self-stimulation of the ventral tegmental area. *Journal of Neuroscience*, 13(1):1–12.

Wood, E. R., Mumby, D. G., Pinel, J. P. J., and Phillips, A. G. (1996). Impaired object recognition memory in rats following ischemia-induced damage to the hippocampus. *Behavioral Neuroscience*, 107(1):51–62.

Worden, R. (1992). Navigation by fragment fitting: A theory of hippocampal function. *Hippocampus*, 2(2):165–87.

Wu, X., Baxter, R. A., and Levy, W. B. (1996). Context codes and the effect of noisy learning on a simplified hippocampal CA3 model. *Biological Cybernetics*, 74:159–165.

Wyss, J. M. and van Groen, T. (1992). Connections between the retrosplenial cortex and the hippocampal formation in the rat: A review. *Hippocampus*, 2(1):1–12.

Ylinen, A., Bragin, A., Nadasdy, Z., Jando, G., Szabo, I., Sik, A., and Buzsáki, G. (1995). Sharp wave-associated high-frequency oscillation (200 Hz) in the intact hippocampus: Network and intracellular mechanisms. *Journal of Neuroscience*, 15(1):30–46.

Young, B. J., Fox, G. D., and Eichenbaum, H. (1994). Correlates of hippocampal complex-spike cell activity in rats performing a nonspatial radial maze task. *Journal of Neuroscience*, 14(11):6553–6563.

Zhang, K. (1996a). Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: A theory. *Journal of Neuroscience*, 16(6):2112–2126.

Zhang, K. (1996b). Spatial orientation and the Papez circuit: Dynamics of the interacting cell groups. Presented at *CNS*96*, (the fifth annual *Computation in Neural Systems* meeting).

Zilles, K. (1990). Anatomy of the neocortex: Cytoarchitecture and myeloarchitecture. In Kolb, B. and Tees, R. C., editors, *The Cerebral Cortex of the Rat*, chapter 4, pages 77–112. MIT Press, Cambridge MA.

Zilles, K. J. (1985). *The cortex of the rat: A stereotaxis atlas.* Springer-Verlag, New York.

Zipser, D. (1985). A computational model of hippocampal place fields. *Behavioral Neuroscience*, 99(5):1006–1018.

Zipser, D. (1986). Biologically plausible models of place recognition and goal location. In *Parallel Distribted Processing: Explorations in the Microstructure of Cognition*, volume 2, chapter 23, pages 423–470. MIT Press, Cambridge MA.

Zipser, D. (1991). Recurrent network model of the neural mechanism of short-term active memory. *Neural Computation*, 3:179–193.

Zola-Morgan, S. and Squire, L. R. (1993). Neuroanatomy of memory. *Annual Review of Neuroscience*, 16:547–563.

Zola-Morgan, S., Squire, L. R., and Ramus, S. J. (1994). Severity of memory impairment in monkeys as a function of locus and extent of damage within the medial temporal lobe memory system. *Hippocampus*, 4:483–495.

Zola-Morgan, S. M. and Squire, L. R. (1990). The primate hippocampal formation: Evidence for a time-limited role in memory storage. *Science*, 250:288–290.

Zoladek, L. and Roberts, W. A. (1978). The sensory basis of spatial memory in the rat. *Animal Learning and Behavior*, 6(1):77–81.