On the use of cognitive maps

adam johnson

University of Minnesota

This is to certify that I have examined this copy of a doctoral dissertation by

Adam Conrad Johnson

and have found that it is complete and satisfactory in all respects, and that any and all revisions required by the final examining committee have been made.

Name of Faculty Adviser

Signature of Faculty Adviser

Date

Graduate school

On the use of cognitive maps

A dissertation submitted to the faculty of the graduate school of the
University of Minnesota by

Adam Conrad Johnson

in partial fulfillment of the requirements for the degree of
doctor of philosophy

Adviser: A. David Redish

June 2008

Acknowledgements

I am profoundly grateful to many people for their contributions to my graduate studies and this thesis.

Dave Redish fundamentally shaped this work. Even though he was surprised by many of these experimental findings and theoretical conclusions, he provided an environment of support and challenge within which these theoretical perspectives developed and these observations were made. His willingness to chase a ideas and to challenge dogma made science exciting and interesting where it might have otherwise been a simple collection of facts.

The members of the Redish lab - Mallika Arudi, Deborah Bang, Dan Bernal, Chris Boldt, Alex Colvin, Giuseppe Cortese, John Ferguson, Maniezheh Firouzi, Mandy Huber, Jadin Jackson, Steve Jensen, Sarah Jutila, Monica Kumar, Zeb Kurth-Nelson, Morgan Little, Beth Masimore, Matt van der Meer, Susan Nwoke, Neil Schmitzer-Torbert, Kelsey Seeland - all provided fundamental contributions to this work, through technical support, experimental assistance or discussions. I would like to highlight Matt van der Meer's questions, discussions and all-around resourcefulness which prompted and refined many of these ideas.

Paul Schrater provided important theoretical discussions and feedback along with much statistical assistance. J. Bruce Overmier and the "Mental time travel in animals" seminar provided much of the conceptual fodder that connected a variety of disparate ideas and observations. Bruce also provided much encouragement to go back to many original sources including Tolman's observations on vicarious trial and error. My interactions with Paul Schrater and Bruce Overmier developed through the Center for Cognitive Sciences and has been a source of inspiration and answers throughout my education. The Cognitive Critique reading group, particularly Celia Gershonson, also provided valuable feedback on early ideas and drafts of the paper on cognitive maps.

Edvard and May-Britt Moser and the Center for the Biology of Memory deserve special mention for their time and support. Although no experiments from their lab are presented

in the thesis, many of the ideas were refined in their lab. Gerit Pfuhl and Robert Beigler provided much support and many discussions on cognitive maps.

David Crowe, Alexandra Basford, Thomas Naselaris and Matt Chafee provided important points of perspective along the way. From the start, David Crowe challenged me to think deeply about the difference between an experimental result and an experimental insight.

Finally, I have been amazingly blessed and continually surprised by the grace of my family - Holly, Isabelle and Elise Johnson - as I wrestled with these experiments and ideas.

Funding sources

This work was supported by NIH (MH68029), NSF (IGERT9870633), a training fellowship from the Center for Cognitive Sciences (NIH-T32HD007151), fellowships from 3M and the graduate school (a first year graduate student fellowship and a doctoral dissertation fellowship) and a Fulbright scholarship.

To I and E who teach me how to see.

Abstract

Cognitive maps were proposed as an alternative to stimulus-response explanations of animal behavior. Although the concept of cognitive maps advanced treatments of complex animal behavior, it has remained resistant to theoretical definition. A simplified perspective on cognitive maps that focused on spatial behavior and the construction of spatial maps has provided an important approach to understanding the role of the hippocampus in spatial behavior and spatially modulated neural activity, particularly within the hippocampus. However, this perspective leaves open many questions on how spatial maps and neural activities within the hippocampus are used and how they contribute to selection of adaptive actions.

A reinforcement learning approach to animal behavior was used to develop a theory of cognitive map function. Reinforcement learning provides a theoretical framework within which the components of cognitive map function can be readily defined and explored. This approach addresses long-standing criticisms of cognitive map theory by explicit mapping of stimuli to action via specific, albeit behaviorally unobservable, computations. Flexible behavior associated with cognitive maps implies the use of transition models in reinforcement learning algorithms. In contrast to model-free algorithms that depend on current experience only, model-based reinforcement algorithms represent sensory or state information beyond the modeled animal's current sensory experience. As a result, model-based reinforcement learning provides a principled approach to analysis of neural representations and the dynamic processes that support cognition.

Neurophysiological recordings in the hippocampus showed that apparent noise present in spatially modulated place cell activity could be explained as coherent spatial representations that deviated from the animal's position on the maze. These non-local representations were associated with fast spatial representation dynamics and were typically found when the animal was at feeder locations or choice points. Non-local representations at choice points shifted forward of the animal to potential future spatial positions and were associated with theta and gamma local field potential activity. Forward-shifted spatial representations were associated with vicarious-trial-and-error behaviors and were task and experience dependent. In sum, these results suggest how cognitive maps in the hippocampus can contribute to selection of adaptive actions through the construction of past events and potential future experiences.

Contents

List of Figures

List of Tables

Chapter 1

Cognition and cognitive maps

1.1 Problem statement

The basic question of this thesis is how animals construct, evaluate and implement plans. To make this question more tractable it has been approached within the context of navigation, cognitive maps and hippocampal place cells. Animal behavior, within and beyond the laboratory, suggests that animals form complex representations that allow them to construct near-optimal patterns of behavior (Stephens and Krebs, 1987; Gallistel, 1990). The approaches of Tolman (1948), O'Keefe and Nadel (1978), and Redish (1999) have provided a foundation for the following discussion that asks how cognitive maps contribute to the organization of behavior. By asking this question the trajectory of cognitive maps begun by Tolman (1948) seems to have come full circle; while previous work has focused on the construction of cognitive maps (O'Keefe and Nadel, 1978; Gallistel, 1990; Redish, 1999) at the expense of explanations of the specific use of cognitive maps, this thesis focuses on understanding how cognitive maps organize behavior.

The following discussion of cognitive maps is organized in three parts. The first part provides a simple explanation of cognitive maps. The first chapter provides a brief review of cognitive maps, cognitive map-based behavior and current perspectives on the neural substrates of cognitive maps. The second part outlines a simple theoretical framework for analysis of cognitive map function. Chapter two introduces and briefly reviews reinforcement learning and several prominent models that have been used to model flexible behavior. Chapter three outlines a simple reinforcement learning model that supports flexible behavior, connects reinforcement learning signals to experimentally observed neural signals and makes several experimentally testable predictions. Part three outlines several experiment findings that indicate how cognitive maps in the hippocampus are potentially used. Chapter four outlines an analytical approach to identification of covert, cognitive signals within neural activity. Chapter five presents several basic experimental findings that support the use of cognitive maps within the hippocampus during decision-making. Chapter six summarizes the basic theoretical perspective, analytical approach, and experimental findings and provides a short discussion of these perspectives and findings in light of Tolman's original ideas on cognitive maps.

1.2 Early observations and the formation of cognitive maps

There is a basic question of why we should return to old papers and ideas about cognition. This question is all the more important when those papers and ideas predate so many conceptual and methodological breakthroughs. The simple answer to our question is in reframing our own perspectives on cognition. The descriptions, analogies, reasoning and insights of these original papers and their ideas are unfettered by our modern conceptual and methodological perspectives. Where these original papers lack modern precision, they offer a candid view of cognition that does not divide cognition into the neat little boxes and areas that we might today. And it is for exactly this reason that we should return to the writings of these astute observers of behavior and puzzle over the connections they make, particularly those that seem most foreign to us.

One of the basic questions of cognition is whether cognition can exist in animals other than humans and, if so, what is it and how might it be manifested through behavior? This was the central question within Edward Tolman's work and remains a central issue in subsequent interpretation of Tolman's ideas and our modern discussions of cognition.

The majority of Tolman's work was conducted as psychology struggled to assert itself as a scientific discipline. The struggle to define psychological science resulted in heated discussion on the division between topics that were scientifically tenable and those that were too airy or imprecise for solid scientific treatment, no matter how interesting they might be. And animal cognition was and remains close in proximity to this divide. Tolman's preferred perspective on cognition was based on animal rather than human behavior. He wrote:

> *There are many ... findings which suggest the operation of something like a pure cognitive or curiosity need in animals and also findings which indicate the nature of the dependence or the independence of this pure cognitive drive upon such more practical wants as thirst, hunger, sex, fear. Furthermore, we, or at any rate I, see these facts and relationships about*

cognitive needs more clearly when they have been observed in rats or apes than when they have been merely noted in a common-sense way in human beings. – Tolman, 1954

In these comments, Tolman distinguishes himself from many of his behaviorist contemporaries who would not treat cognition in either humans or animals. Tolman's perspectives can also be directly contrasted to later views of cognition that emphasized human (and quite often very 'book-ish') aspects of cognition.

For reasons that are not entirely clear, the battle between these two schools of thought [S-R learning versus cognitive learning] *has generally been waged at the level of animal behaivor. Edward Tolman, for example, has based his defense of cognitive organization almost entirely on his studies of the behavior of rats — surely one of the least promising areas in which to investigate intellectual accomplishments.* – Miller et al., 1960 (p.8)

Consequently Tolman's writings, though influential, have often lacked a comfortable place in the history of psychology. A variety of recent investigations of animal behavior have once again emphasized the role of cognition and more recent investigations of human cognition have wondered about correspondences with animal behavior. And so it is that we revisit Tolman's perspectives on cognition.

The following sections revisit Tolman's development of cognitive maps, particularly the underlying features of his earlier work that gave rise to these ideas. Tolman (1948) reviewed a series of five experiments as the basis for cognitive maps. Exploring each of these experiments in turn, the following discussion examines Tolman's original definitions and their evolution through Tolman's later work with links to more recent investigation.

1.2.1 Tolman and *cognitive maps*

Tolman developed the idea of a cognitive map as an alternative to the then-common metaphor of a central office switchboard for learning and memory, typical of stimulus-response formulations. He wrote,

We assert that the central office itself is far more like a map control room than it is like an old-fashioned telephone exchange. The stimuli, which are allowed in, are not connected by just simple one-to-one switches to the outgoing response. Rather, the incoming impulses are usually worked over and elaborated in the central control room into a tentative, cognitive-like map of the environment. And it is this tentative map, indicating routes and paths and en-

vironmental relationships, which finally determines what responses, if any, the animal will finally release. – Tolman, 1948

The origins of the cognitive map are evident even in Tolman's early writings (Tolman, 1932; Tolman and Krechevsky, 1933; Tolman and Brunswik, 1935). Tolman's anti-reductionist emphasis on purpose and macroscopic or *molar* behavior stood in stark contrast to many of the learning theories of his contemporaries. Though Tolman was greatly influenced by Watson's behaviorism and considered himself a behaviorist throughout the majority of his career, his framing of behavioral questions, particularly in relationship to expectancies, signalled a key difference from his contemporaries. This initial position was called *purposive behaviorism* (Tolman, 1932) and provided the foundation for Tolman's later perspectives on animal cognition. Purposive behaviorism can be most succinctly summarized as the idea that animals develop expectancies of their dynamic world and through these expectancies they organize their behavior.

What is expectancy?[1] Tolman struggled with this question throughout his work. His early writings described an expectancy as an animal's multifaceted interaction or *commerce* with its environment. For instance, Tinklepaugh (1928) observed that monkeys would engage in 'surprise hunting behaviors' when one food type was substituted for another expected food type. When the animal found a reward that differed from its previous experience – even though the reward the animal found was supposedly 'just as rewarding' – it would continue to search for a stimulus that matched its previous experience. Such behaviors signaled to Tolman that animals, even those as simple as rats, maintained a set of complex, integrated expectancies of their world (Tolman, 1932). Tolman would later call these complex integrated expectancies cognitive maps (Tolman, 1948). His later writings emphasized the use of multiple expectancies to inform behavioral performance, particularly when an animal is faced with a choice (Tolman, 1954).

Because the expectancies an animal holds and its use of them develop with experience, different cues or sets of cues that underlie behavior change as a function of experience (Tolman and Brunswik, 1935; Tolman, 1949). That is, both the content and the use of cognitive maps change with experience. From our modern perspective, this statement and its ramifications may be readily apparent; rats, for instance, tend to navigate using map-like place-learning given certain sets of environmental conditions and experiences whereas they

---

[1] It should be noted that this question is tightly coupled with the question of representation (Schneidman et al., 2003): a question that was not adequately addressed by Tolman or his contemporaries and continues to cause debate.

tend to navigate using response-learning given other sets of environmental conditions and experiences (O'Keefe and Nadel, 1978; Packard and McGaugh, 1996). These ramifications were less evident during Tolman's era and led to one of the more contentious debates on the basis of rat navigation (Hull, 1943; Tolman, 1948). Tulving and Madigan (1970) describe the debate between place-learning, a position advocated by Tolman, and response-learning, a position advocated by Hull, by stating: 'place-learning organisms, guided by cognitive maps in their head, successfully negotiated obstacle courses to food at Berkeley, while their response-learning counterparts, propelled by habits and drives, performed similar feats at Yale' (p. 440). The debate's resolution has been that animal behavior is dependent on rich sets of cue information during early performance of novel tasks and with overtraining this behavior becomes dependent on an increasingly smaller and more specific set cue information. As such, it is worth noting that Tolman was primarily interested in the case of early, relatively novel learning while many of his contemporaries (such as Hull) were investigating the case of later, overtrained learning (Restle, 1957).

In his original formulation of *cognitive maps*, Tolman (1948) discussed five basic experiments to develop his perspectives on cognition:

***Latent learning*:** Learning can occur without observable changes in behavioral performance. This form of learning can be accomplished with a completely random (e.g. passive) method of exploration.

***Vicarious trial and error*:** Learning occurs through active investigation, either in terms of investigating the signaling stimulus definition or its signaling contingencies. This form of learning depends on coincident signalling stimulus and contingency learning.

***Searching for the stimulus*:** Learning occurs through active investigation by the animal – highly salient outcomes yield a search for a cause. This search, however, does not indicate how credit assignment is made in terms of consistency and valence.

***Hypotheses*:** Development of expectancies requires testing and outcome stability. Hypothesis behavior is based on a previous set of behavioral expectancies – that some given change in behavior should produce a change in environmental outcome.

***Spatial orientation or the short-cut*:** The efficient use of hypothesis behavior in a novel circumstance.

The majority of these ideas have gained a multitude of meanings since Tolman originally presented his formulation of the cognitive map. In order to understand more thoroughly

Tolman's conceptualization of cognitive maps, we revisit his earlier writings and investigate why he draws on these experiments.

Latent learning

Latent learning can be defined as learning that occurs in the absence of generally observable changes in behavioral performance. Given appropriate experimental conditions, this learning can be subsequently uncovered. Tolman highlighted latent learning as the first experimental elaboration of cognitive maps to demonstrate that learning occurs across multiple modalities and that learning generally occurs even when that learning is not manifest in behavior. This integrated learning across multiple modalities provides the basis for an animal to perform different sets of behaviors as different needs arise (e.g. thirst, hunger, etc.).

In a traditional example of latent learning (Spence and Lippitt, 1946), two sets of fully fed and fully watered animals are allowed to navigate a Y-maze. Food is available at the end of one arm of the Y and water is available at the end of the other arm. The first set of rats is then water deprived while the second set of rats is food deprived. When the two sets of rats are then placed on the Y-stem, each set runs to the appropriate arm at levels much greater than chance (rats deprived of water run to water and rats deprived of food run to food).

Latent learning provides the fundamental basis for cognitive maps by allowing for organisms to learn covertly. Animals need not display all that they have learned at a given moment. By beginning with latent learning, Tolman explicitly argues against the assumption that a lack of performance is synonymous with a failure to learn. Performance, Tolman maintains, is not a simple mixture of rewarded responses, but an interaction of need and expectancy. The cognitive map is useful only insofar as it allows an organism to learn and develop a set of expectancies that anticipate potential future needs (in the Y-maze example for instance, the rats did not know whether they would be water or food deprived).

Vicarious trial and error

Vicarious trial and error is a set of experimentally observable behaviors where an animal attends to and sometimes approaches a specific choice option but does not commit to it (Muenzinger, 1938; Tolman, 1938, 1939). At choice points within T mazes or radial arm mazes, vicarious trial and error appears as a vacillation between potential options; the rat orients toward one maze arm then re-orients toward another until it finally makes its choice.

Although the term *vicarious trial and error* (VTE) carries with it a variety of meanings, Tolman appears to have used this description to emphasize the changing interaction an animal has with its environment. Within the context of cognitive maps, Tolman specifically identifies the temporal relationship between VTE behavior and task performance as a significant point of interest. In tasks where VTE occurs, rats usually display an increase in VTE behaviors immediately before dramatic improvements in task performance. Once performance reaches ceiling, vicariously trial and error behaviors diminish. If the task is suddenly changed or the discrimination is made more difficult, rats will again display increased VTE behavior as they learn the new set of contingencies. Tolman's hypothesis is that VTE signals a state of cognitive map re-organization that provides the substrates for changing behavioral performance.

Tolman's earlier writings suggest a distinction between learning the causal (or coincident) couplings in the environment and the effects of the organism's own behavior (Tolman and Brunswik, 1935; Tolman, 1938). As developed by Tolman, vicarious trial and error behaviors are hypothesized to display the tension between two driving forces: (1) the learning of the set of signaling cues within a task (e.g. a grey cue card located at one arm) and (2) the task contingency (e.g. food is located at the end of the alley with the grey cue card). Failed behavioral performance can occur due to either an inability to distinguish the signifier, to learn the signalled contingency, or to link the two. For example, a grey cue card that signals the rewarded arm on a Y-maze likely holds no distinction in a rat's perception until the rat learns that there is some correlation between its environment and reward availability. Now the rat must discriminate the boundaries of the correlation between environment and reward availability, namely the grey cue card. As a result, the animal orients toward one maze arm in a search for some potential incoming signal that stably correlates with the rat's previous experience (Tolman, 1938, 1939). The interactive quality of identification of cues (signifiers) and their contingencies provide a central feature of cognitive maps: beyond a form of random exploration that might be used in latent learning, an organism actively attends to a given stimulus set in order to determine its boundaries and identify its signalling contingencies.

Searching for the stimulus

An organism's search for a causal stimulus continues Tolman's development of active interaction between an organism and its environment. The set of experiments highlighted by Tolman (1948) describe searching for a stimulus as the post-hoc attribution of an outcome (in most cases a shock) to a stimulus. When the paired stimulus was removed immediately following the outcome, the stimulus attribution was incorrectly made to another stimulus if it was even made at all. This attribution is what has come to be known as the *credit assignment problem*. Tolman's focus on the active assignment of credit and the temporal order of this assignment (outcome leads stimulus) contrasted with much previous research that focused on the passive transfer of the value or valence of an unconditioned reinforcer to a previously activated stimulus (stimulus leads outcome).

Tolman's perspective is predicated on a set of pre-existing expectancies that an organism has about its environment. It is important to note that the searching for the stimulus described by Tolman (1948) is *novel, single-trial* learning; that is, an experimental manipulation greatly deviates from the organisms set of pre-existing expectancies. When the organism's expectancies are violated, the organism investigates and assigns credit. While searching for stimulus behaviors are indicative of active credit assignment in single trial learning, an organism will develop another set of active credit assignment behaviors when faced with multiple different problems of the same class. Tolman called these behaviors *hypotheses*.

Hypotheses

In their paper on *The organism and the causal texture of the environment*, Tolman and Brunswik (1935) provide an explanation for objective hypotheses: the appearance of systematic rather than chance distributions of behavior. The term 'objective' denotes the raw statistical nature of the behavior rather than a subjective hypothetical experience within the organism. Such a statement can be interpreted as either non-sensically trivial or excruciatingly complex. The behavioral performance of an animal that displays hypothesis behavior may not appear to be different than random at a macroscopic level; however, on closer inspection these animals switch from one strategy to another in a rather discrete manner. Citing Krechesvsky (1932), Tolman illustrates this point with observations from a sequential Y decision task: for instance, a rat might first attempt all left turns, then all right turns, then a progressive mixture of the two by alternating between right and left turns.

A basic problem with observations of this type is how they should be reported – a point that is particularly salient because animals rarely follow the same trajectory through a given hypothesis space. At a macroscopic level, animals might select one option at chance probabilities; at a microscopic level, the patterns of behavior might not be predictable. Standard statistical treatments of learning typically describe behavioral patterns in terms of averages. However, such averages across multiple animals often mask abrupt transitions from chance performance on a task to nearly perfect performance within a single animal (Gallistel et al., 2004). Because modern change point analysis were not available (see Smith et al., 2004 as an example), Tolman resorts to reporting multiple single cases in an attempt to describe the behavior of rats at spatial choice points (Tolman, 1938, 1939).[2]

The basis for hypothesis behavior is *functional chunking* of the environment. Tolman and Krechevsky (1933) emphasize that animals do not attempt nonsense intermediate strategies (e.g. attempting to go straight on a T-maze and hitting the wall) but rather shift from one coherent *molar* pattern of behavior to another. In essence what Tolman has done is to have allowed the organism to be a good scientist: it maintains a consistent behavior to determine the predictability of the environmental (task) outcome. Tolman's theoretical development suggests that hypotheses are composed of *functional chunks*, or to use his term, expectancies. A rather simplistic version of hypothesis behavior might be used to explain behavior on the Y-maze during the test portion of latent learning: an *approach* functional chunk (here a set of motor outputs) combined with the previous learning experience produces a hypothesis – *approach deprived outcome*. The hypothesis is utilized as a complete chunk and is rewarded. This simple hypothesis behavior is typical of the Y-maze but is easily elaborated to produce much more interesting behaviors within other more complex tasks as we will see.

Spatial orientation

Much has been made of the short-cut behavior highlighted by Tolman in the sunburst maze (Lashley, 1929; Tolman et al., 1946). In this task, rats are pre-trained to navigate through an indirect single sequence of alleys to a food reward site. During the test phase, the alley sequence from pre-training is blocked and instead multiple radial paths are presented. One of these radial paths leads directly to the food site. The majority of rats select the short-cut – the radial path that leads directly to food (Tolman et al., 1946).

While a number of methodological issues have been identified in this experiment (e.g. a light was placed above the food site in the original experiment; O'Keefe and Nadel, 1978), Tolman argues that *'the rats had, it would seem, acquired not merely a strip-map to the effect that the original specifically trained on path led to food but rather, a wider comprehensive map...'* (Tolman, 1948). The short-cut behavior arises from the capacity to combine pre-existing behavioral expectancies with environmental cues which have not been explicitly rewarded. The appeal of the short-cut has, at its basis, an efficiency argument; the short-cut is the most efficient task solution. The resulting question is over what cue

---

[2] Abrupt transitions in behavioral performance appear to be hallmarks of certain choice tasks, particularly those which elicit vicarious trial and error.

sets and manipulations can an animal continue to produce the most efficient method of task solution. Tolman's use of the term *width* to describe cognitive maps seems to denote a measure of efficiency or level of energy-savings over a broad variety of environmental cues and manipulations.

The width of a map might be considered to be a question of specificity and generalization. In contrast to others' subsequent interpretation of cognitive maps (e.g. O'Keefe and Nadel, 1978), Tolman (1948) does not indicate how wide or narrow a cognitive map should be or what specific set behaviors are indicative of *the* cognitive map. Beyond the basic set of illustrative behaviors listed above, Tolman simply states that certain conditions such as brain damage, inadequate environmental cues, over-training, very high motivational or emotional states lead to narrower, strip-like maps behavior. Given appropriately impoverished conditions, cognitive maps can be reduced to very narrow *stimulus-response* maps; however, the crux of Tolman's formulation of cognitive maps suggests that the deductive approach will not yield the complex interactions made possible through broad cognitive maps and observed in less-impoverished environments.

1.2.2 Subsequent readings and perspectives

An early reading of Tolman by MacCorquodale and Meehl (1954) emphasized the similarity of his learning theory with other learning theories. The observation that Tolman's use of the term *expectancy* contained multiple different meanings and implications led MacCorquodale and Meehl (1954) to describe Tolman's formulations as unnecessarily imprecise and, perhaps, even shoddy. Their "precise" reformulation considers Tolman's learning theory in terms of *learning to* rather than *learning that* (MacCorquodale and Meehl, 1954). The end result of their action-based reformulation was a set of functional and mathematical relations that were virtually indistinguishable from other contemporary theories of learning (c.f. Hull, 1943). The similarity of his reduced theory to Hull's theory and others was not disputed by Tolman (1955) and to a certain extent this perspective was anticipated by Tolman (1948), but neither did Tolman fully agree with this position.

Two critical points identified by MacCorquodale and Meehl (1954) were to become a powerful central influence in subsequent readings of Tolman. These are first that "Tolman's complex cognition-statements are not themselves truth-functions of their components" and second that when a rat's expectancies are discussed rather than its habits, the questions of reference and intention become unavoidable. Each point highlights the increasingly prominent position that *representation* would come to hold within descriptions of animal behavior and particularly within the nascent field of cognitive psychology.[3] The first point highlights the subjective component of animal behavior and begins to move away from the perspective that the animal somehow extracts purely objective information about its task or more specifically that an animal might infer something about its task that was not simply derived from bottom-up processes. The second point highlights the required discussion of establishing representations relative to specific reference points or intentions. The place-learning/response-learning debate (Hull, 1943; Tolman, 1948; Restle, 1957) can be interpreted as an argument of the representational substrates of learning. And it was such conceptual developments of place representations within cognitive maps that aroused the interest of later cognitive and neurobiological research.

1.2.3 Current behavioral approaches to cognitive maps

Research on cognitive maps continues to be an area of current active interest. This interest is based on two fundamental sets of observations – place cells in the hippocampus (O'Keefe and Nadel, 1978; Redish, 1999) and comparative ethological approaches to animal foraging and navigation (see review by Gallistel, 1990).

In their influential book *The Hippocampus as a Cognitive Map*, O'Keefe and Nadel (1978) suggested that the recent discovery of *place cells* within the hippocampus (O'Keefe and Dostrovsky, 1971; O'Keefe, 1976) provided the neurobiological substrates of the cognitive map. O'Keefe and Nadel (1978) elaborated this hypothesis with an extensive review of the behavioral consequences of damage to the hippocampus and the mechanisms of hippocampal function that potentially supported learning spatial maps. This assertion revitalized discussion of cognitive maps which had waned within laboratory settings following the arguments by MacCorquodale and Meehl (1954) and Restle (1957) that Tolman's learning theory was, at its basis, equivalent to Hull's theory (Hull, 1943). O'Keefe and Nadel explicitly argued against treating the learning described by Tolman (1948) and the learning described by Hull (1943) as two phases of a single learning process. Instead they suggested a distinction between multiple memory systems: a *locale* system based on hippocampal function that supported spatial learning and a series of *taxon* systems that supported other simpler forms of learning. Based on the phenomenology of neuropsychology and neurophysiology of the hippocampus, O'Keefe and Nadel (1978) argued that the hippocampus

---

[3] It should be noted that while an emerging field of research that would become cognitive psychology was developing in the 1950s, these researchers were primarily physicists, mathematicians and engineers who had become interested in behavior (von Neumann and Morgenstern, 1944; Simon, 1955; Newell and Simon, 1972; Turing, 1992) and whose perspectives were relatively independent of Tolman's ideas (Hilgard and Bower, 1975).

provided the neural basis for cognitive maps.

The idea that animals use maps to navigate in their natural environment has its roots in folk psychology and its modern origins in ethology. The observed search patterns in many species suggest that animals combine various classes of sensory information to compute spatial location (see reviews by O'Keefe and Nadel, 1978; Gallistel, 1990; Redish, 1999). The complex spatial behaviors of animals as simple as arthropods suggest the use of metric spatial information and construction of cognitive maps. The ethological approach to spatial behavior is typified by the mathematical perspective of Gallistel (1990). Based partially on evolutionary considerations of animal behavior, this perspective suggests that the strategies animals use to navigate approach optimal patterns of behavior. In contrast to the multiple memory systems perspective within the treatment of O'Keefe and Nadel (1978), ethological perspectives typically hold that a single integrated (or unified) system informs spatial behavior.

While there are many differences between these modern perspectives on cognitive maps, they share a common emphasis on the pure spatial (or geometric) component of cognitive map function. Observations that animals can navigate even in the absence of spatial cues (by dead reckoning) suggests that spatial maps are based on metric information and that a large class of geometric relations can be used for spatial inference and wayfinding (Gallistel, 1990). In sum, these perspectives on cognitive maps have been devoted to explaining how animals construct and use spatial representations that underlie the complex patterns of spatial behavior observed both within the laboratory and more naturalistic settings. Although these perspectives retain certain aspects of the original treatment of cognitive maps (e.g. spatial short-cut behavior), they also underscore the great divergence of modern theories of cognitive maps from Tolman's ideas.

Place learning and cognitive maps

The original conceptualization of cognitive maps was based on five clear experimental observations and an opaque formal behavioral learning theory (Tolman, 1948). O'Keefe and Nadel (1978) presented a more formalized approach to cognitive map function that emphasized the spatial components of Tolman's original theory and specifically place learning. The basic psychological premise of the theory[4] forwarded by O'Keefe and Nadel (1978) was the distinction between routes and maps (see Table 1.1).

---

[4] This discussion is focused on the psychological contributions to cognitive maps made by O'Keefe and Nadel (1978) and leaves discussion of their neurobiological contributions to the next section.

	Route	Map
Motivation	The final stimulus is the goal; the route is built with this in mind	No object or place on the map is a goal; the map is usually built out of curiosity
Flexibility	Routes are rather rigid; the are rendered useless by any damage, or by the loss of a guidance, or direction	Maps are extremely flexible, and relatively invulnerable to noise and damage
Speed	Very fast	Relatively slow
Information content	Relatively little; each route contains only a small amount of data	Maps are one of the most efficient information storage devices known, with very large capacity
Access	No specialized knowledge is required for access; no coding strategies	Special knowledge of coding strategies required
Manipulation	None	Maps can be compared; places on maps can be compared

Table 1.1: Properties of routes and maps hypothesized by O'Keefe and Nadel (1978) in their treatment of cognitive maps (Table 1 from O'Keefe and Nadel, 1978, p.89).

Route-based navigation, argued O'Keefe and Nadel, was based on a sequence of stimulus-response pairings. Several important consequences emerge from this perspective on routes (Table 1.1). Routes are created based on a goal and established as a means for obtaining that goal. They contain a simple set of stimulus related information that allows for the development of subgoals; and this spatial information is organized relative to the current heading of the animal rather than a environment-based coordinate system. While these characteristics make route information easily accessed and fast to use, they also underscore the relative inflexibility of routes and the heightened susceptibility to becoming lost when using route information. As a final note, routes are not reversible; inverting a route in order to retrace one's steps produces a myriad of complications that are typically insoluble beyond the most trivial of circumstances.

O'Keefe and Nadel (1978) contrast route based navigation with map based navigation. They argued that map based representations allow for the use of multiple spatial transformations and that other non-spatial forms of information are embedded or symbolized within the map. This treatment of map based navigation leads to several important consequences. O'Keefe and Nadel argued that maps are constructed on the basis of curiosity rather than on the basis of obtaining a specific goal. They have high information content and confer great navigational flexibility on the animal using maps. Map availability suggests that even when an animal becomes lost, it can quickly recover its location and resume its journey. However, maps require special coding strategies and are relatively slow to access.

Given the basic phenomenology of wayfinding behaviors in human and non-human animals and the organization and function of the hippocampus (treated below), O'Keefe and Nadel (1978) hypothesized two types of memory systems: a *locale* system subserved by the hippocampus and a series of *taxon* systems mediated by extra-hippocampal areas. O'Keefe and Nadel (1978) argued that map-based behavior was based on locale system function while route-based behavior was based on the functions of the various taxon systems. As suggested by the title of their book, they primarily focused on treatment of the properties and functions of the locale system. Table 1.2 outlines the basic characteristics and functionality of the taxon and locale systems.

Several important differences between taxon and locale systems mark O'Keefe and Nadel's theoretical formulation of cognitive maps. In contrast to incremental learning within taxon systems, learning in the locale system is accomplished through an all-or-none mechanism. Because locale system learning progresses in following an all-or-none rule, abrupt shifts in behavioral performance or hypothesis behaviors emerge (Krechesvsky, 1932; Tolman and Krechevsky, 1933). Furthermore, each system uses cue (stimulus) information

	Taxon	Locale
Motivation for learning	Biological need: to obtain reward or avoid punishment	Cognitive curiosity: to construct and update maps of the environment
Learning change	Incremental or decremental	All-or-none
Persistence	High esp. orientation hypotheses	Low
Temporal changes after activation	Marked changes in threshold and strength as a function of time after activation: sensitive to intertrial interval	Minimal changes with time after activation; insensitive to intertrial interval
Interference between similar items	High	Low

Table 1.2: Properties of taxon and locale systems hypothesized by O'Keefe and Nadel (1978) in their treatment of cognitive maps (Table 2, from O'Keefe and Nadel, 1978 p.100).

differently. While specific sets of cue (stimulus) information are critical for appropriate taxon system function, no single cue is necessary for appropriate locale system function. Place inference within the locale system is mediated by a specific constellation of cue (stimulus) information that is robust to single cue perturbations and interference between similar items. Finally, taxon and locale systems can be contrasted in the interaction between sensory and motor function. The locale system combines sensory and motor information within spatial representations through a dead-reckoning system. In contrast, taxon systems are primarily organized and controlled by sensory stimulus patterns that release a specific motor program.[5]

O'Keefe and Nadel (1978) provide only a terse discussion on the arbitration of control between taxon and locale systems. The two primary factors considered in arbitration of control are reinforcement and uncertainty. Reinforcement, in O'Keefe and Nadel's view, serves to maintain the current balance of control between taxon and locale systems and non-reinforcement leads toward changing the balance of control. Similarly and perhaps consequently, uncertainty can produce changes in arbitration of control. O'Keefe and Nadel (1978) suggest that increased uncertainty associated with a single cue used by a taxon

---

[5] This distinction between the integration of stimulus and action information within taxon and locale systems results in a specific prediction regarding pavlovian and instrumental conditioning. O'Keefe and Nadel (1978) explicitly suggest that although taxon systems mediate all forms pavlovian conditioning and many forms of instrumental conditioning, certain types of instrumental conditioning are dependent on locale systems (p.316). However, it should be noted that O'Keefe and Nadel further argue that learning that generalizes across contexts is based on taxon rather than locale systems.

system will shift behavioral control to the locale system while increased variability in the cues used for place inference will shift behavioral control away from the locale system and toward taxon systems.

More recent treatments that follow from O'Keefe and Nadel (1978) have further elaborated the mechanisms and function of cognitive maps. Redish (1999), for instance, outlines five different forms of navigation: random, taxon, praxic, route and locale navigation. These different forms of navigation each require different sets of underlying computations. Supporting these distinctions, he reviewed the hippocampal and extra-hippocampal dependence of these patterns of navigation and specified a formal computational theory of cognitive map function based on attractor dynamics (McNaughton et al., 1996; Touretzky and Redish, 1996; Samsonovich and McNaughton, 1997; Redish, 1997; Redish and Touretzky, 1997, 1998a). This theory of cognitive maps based on attractor dynamics has been used to reproduce many of the behavioral phenomena hypothesized by cognitive map function (as defined by O'Keefe and Nadel, 1978 – particularly experimental results similar to experiments 1,3 and 5 from Tolman, 1948) and hippocampal place cell activity.

Although Redish (1999) moves toward the use of multiple maps related to goals, his treatment maintains a primary focus on the construction of cognitive maps rather than specifying their use, a perspective central to the original treatment of O'Keefe and Nadel (1978). This perspective shared by O'Keefe and Nadel (1978) and Redish (1999) rests on the tenet that once an animal locates itself on a given map, it will immediately be able to construct a goal location toward which the animal can move. Although this perspective represents an important contribution, it leaves open several questions related to the formation, evaluation and use of goals within behavior. Clearly these goal related functions depend on an interaction between route and map-based systems that are absent within the competitive control based formulations of O'Keefe and Nadel (1978). Part of this neglect is due to the hippocampus-centric view found within research on the neural substrates of cognitive maps. A more integrated view of cognitive maps and their use within spatial navigation comes from ethological observations.

Ethological approaches to cognitive maps

In contrast to psychological and neurobiological approaches to cognitive maps that emphasize differentiation of multiple memory systems, ethological approaches typically focus on whole, intact behavior and the interactions between the animal and its environment (Gallistel, 1990). The ethological approach to cognitive maps builds on development of metric information and subsequent inclusion of specific stimulus information within a constructed metric space.[6] This unified approach begins with a simplified view of goal behavior in terms of path integration and elaborates this view of goal behavior in terms piloting.

Animals make circuitous trajectories as they forage for food. However, even after foraging over long periods and great distances, they find their way back to the nest. Rather than retracing their trajectory, many animals select a nearly direct path home (Mittelstaedt and Mittelstaedt, 1980; Wehner and Srinivasan, 1981; Müller and Wehner, 1988). At its simplest, this form of navigation requires only the storage of metric information relative to a starting position. The cumulative displacement relative to a starting position that develops through the trajectory can be subsequently used to compute an approximate direction and radial distance. This computation is called path integration (Mittelstaedt and Mittelstaedt, 1980; Wehner and Srinivasan, 1981) or dead reckoning (Gallistel, 1990). Many animals, from arthropods to mammals, choose paths in the appropriate direction of their initial starting position and break off their search at the appropriate distance even when landmarks that might otherwise be used for navigation has been removed (Mittelstaedt and Mittelstaedt, 1980; Wehner and Srinivasan, 1981; Müller and Wehner, 1988; Gallistel, 1990; Etienne et al., 1996).

Animals' search strategies reflect both path integration and use of environmental feature information. Piloting represents a process that is described by Gallistel (1990, p.41) as moving "around unobserved obstacles toward unobserved goals by reference of observable features of the land and to a map that records the geometric relationship" between perception and the goal (whether it is to be found or avoided). In this description of piloting, Gallistel (1990) articulates a method for construction and evaluation of goal information. Goals are produced as a consequence of the content of a cognitive map and spatial behaviors represent the use of geometric- and feature-based inference in an attempt to realize these goals. The feature dependence of search strategies (Collett and Land, 1975; Collett, 1987; Arolfo et al., 1994) suggest that animals somehow represent the content of goal location (e.g. a snapshot or a sketch of landmark information) such that goal locations can be recognized and approached (Gallistel, 1990, p.121). How these goal acquisition behaviors are implemented remains an open question.

Within many theoretical treatments of an objective geometric spatial representation within cognitive maps, routes are defined as direct trajectories toward a location that end

---

[6] This discussion is heavily based on the discussion of navigation by Gallistel (1990) and will be generally focused on the development of goal representations. It should be noted that although Gallistel (1990) takes a modular approach to describing the processes that underlie behavior, these are part of a complete and unified framework. The organization of learning hypothesizes that a set of fully integrated modules produce near optimal behavior given natural environmental conditions.

with recognition of the predicted target. In simple theories, animal behavior at that target location is little more than random search. More complex theories have considered the organization of search processes. For instance, search patterns observed in ants following dead reckoning to a fictive target emanate in increasingly large concentric circles centered on the fictive target location (Wehner and Srinivasan, 1981). Beyond the use of uncertainty information within search, animals also actively perform information gathering behaviors similar to those described by Tolman (1948) as animals searched out a stimulus and performed vicarious trial and error behaviors.

In an interesting study on the active use of dynamic cue information in estimating distance, Ellard et al. (1984) describe information gathering behaviors of gerbils faced with the task of jumping across a gap. Gerbils estimated gap distance according to two type of parallax information (as reviewed by Gallistel, 1990). The *loom* or expansion of the visual image of the target was used when the runway preceding the gap was long enough to allow the animal to dart toward the target. Head bobbing was used when the track was sufficiently shortened to make darting impossible. These behaviors are a direct consequence of spatial problem solving. In order for animals to produce these behaviors they necessarily identified the problem to be solved (crossing the gap) and inferred a general solution strategy (a jump) for a given goal (reaching the target). These dynamic behaviors (darting toward the target on long runways and head bobbing on short runways) represent the explicit search for information related to (parameterization of the) solution of the problem rather than a simple headlong attempt toward goal acquisition.

In sum, ethological approaches to cognitive maps suggest a specific interaction between objective map information and the construction of subjective routes. Although the construction of goals and the spatial inference that provides the engine for realizing these goals remains poorly specified compared to the rigorous formulation of the spatial components of these theories, the movement toward explicit on-line integration of multiple computational processes (modules) represents a critical step forward.

1.2.4 Conclusions

The central question motivated by the previous discussion is how animals use the *content* of cognitive maps to inform behavior. How, exactly, does an animal formulate a trajectory through a map to a goal location? How, exactly, does an animal define and create goals? How does the search for a goal progress? The perspectives of O'Keefe and Nadel (1978) and those of ethologists differ in their treatment of goals within cognitive map function.

O'Keefe and Nadel (1978) remove the influence of goals and motivations from the *locale* system in order to simplify their discussion of cognitive maps. Instead they argue that

> *One thing is clear: routes imply goals which imply motivations. Why do you want to go to Upshire? The second stimulus in each S-R-S link can be looked on as a subgoal in the overall goal-directed chain. Routes direct attention to particular objects or specify turns within egocentric space. They are inflexible, must be use in the correct sequence, and only rarely allow freedom of choice to the traveler.* – O'Keefe and Nadel, 1978

O'Keefe and Nadel (1978) argue that route-based navigation has the great potential for becoming lost and contrast this with the much more robust form of map-based navigation. One problem in the argument forwarded by O'Keefe and Nadel (1978) is how cognitive maps are used for motivated behavior. The statement shown above suggests that all motivated or goal-based behavior is left to route-based taxon systems. Perhaps this explains why map-based navigation has such small potential for becoming lost because *how can an animal become lost if it has no goal?* And perhaps more importantly, we might ask, like Guthrie (1935) about how cognitive maps inform action.

O'Keefe and Nadel (1978) suggest that cognitive maps within the locale system are free of distortion potentially produced by motivation but leave open the interaction between the locale system and the various taxon systems which might use the contents of cognitive maps, form goals and inform action. Reinforcement within the locale system is included within the content of a map just as any other stimulus information would be. In contrast, reinforcement acts as a specialized channel for taxon systems. The suggestion that reinforcement information is included within cognitive map content implies some interaction between locale and taxon systems in the construction of cognitive map based routes. But how such routes are constructed remains unclear.[7]

Ethological perspectives have further developed the interaction between cognitive maps and spatial behavior strategies, particularly those related to search. However, these perspective leave open a number of important questions. Are goal locations identified and created based on previous experience – that is, were they tagged as potential goal locations at the time they were originally (or subsequently) visited. If so, how is the tagging of goal locations accomplished and is this subject to specific motivational influences? If, alterna-

---

[7] It is interesting to note that in contrast with taxon systems, O'Keefe and Nadel (1978) argue that the locale system and the hippocampus in particular "is specifically designed to produce variability in behaviour" (p.318). While the variability produced by curiosity might be considered as random, it could also be used to inform search processes – although it is not clear how this would be accomplished.

tively, goal locations are not tagged and are constructed after they were originally visited, how is this process organized?

Modern theoretical development of cognitive maps has primarily focused on the construction of objective spatial representations and the computational transformations required to support observed spatial behavior (O'Keefe and Nadel, 1978; Gallistel, 1990; Redish, 1999). This development has only peripherally examined the construction, evaluation and use of goals (O'Keefe and Speakman, 1987; Redish, 1999). As current research on cognitive map function and its neurobiological substrates has begun to consider these questions, it has returned to a more integrated view of cognitive map function that is consistent with Tolman's original formulation of cognitive maps. This research has focused on the relatively underdeveloped aspects of Tolman's theory (1948)[8], specifically hypothesis and vicarious trial and error behaviors (Hu and Amsel, 1995). The following section further elaborates the cognitive map function and its neurobiological instantiation.

1.3 Cognitive maps and the hippocampus

Damage to the hippocampus produces broad deficits in spatial learning and memory (O'Keefe and Nadel, 1978; Redish, 1999). The profound behavioral deficits in hippocampal lesioned animals combined with the discovery of place cells led O'Keefe and Nadel (1978) to hypothesize that the hippocampus formed the neural substrate for cognitive maps. Indeed O'Keefe and Nadel's original development of cognitive maps flows from the proposed equivalence of *spatial cognitive map function*[9] and hippocampal function: the hippocampus does what cognitive maps do and cognitive maps are what the hippocampus does. While this proposition has provided a rich framework for studying the functional properties of hippocampal place cells and interpreting the consequences of hippocampal lesions, it has also created an oddly moving definition of cognitive maps based on emerging evidence of hippocampal mechanisms and function. The conclusion of such a proposition is that any hippocampal- or cognitive-map- based function must be related to the place qualities of maps (O'Keefe, 1999); as a result, behaviors such as acquisition of trace conditioning (Solomon et al., 1986; Beylin et al., 2001) and retrieval of episodic memories which are also dependent on the hippocampus are considered epiphenomena of the mechanisms of cognitive map function.

---

[8] Although the authors of many of these studies tacitly claim their data disprove the existence or use of cognitive maps, their arguments seem directed toward the strong spatial cognitive map hypothesis forwarded by O'Keefe and Nadel (1978).

[9] I use the term *spatial cognitive map function* to denote the strong spatial position taken by O'Keefe (1999) wherein all cognitive map function is tied to place information.

The relationship between cognitive map function and the hippocampus proposed in the following discussion is less stringent than the strong spatial cognitive map proposal by O'Keefe (1999). Instead, it is focused on the interaction between processes which locate an animal on a cognitive map and those that inform animal behavior and through the evaluation and production of routes. So within the nomenclature developed by O'Keefe and Nadel (1978), this proposal of cognitive map function is based on the interaction between locale and taxon systems. Consequently, the neurobiological approach to cognitive map function within the present proposal extends beyond the hippocampus (and closely related hippocampal areas) and identifies the hippocampus as an important component within a system that mediates cognitive map function.

The basic organization of the hippocampal archicortex has been described as a trisynaptic circuit (see Figure 1.1). Entorhinal cortex (layer II) sends projections to the dentate gyrus via the perforant path; dentate gyrus sends projections to CA3 via mossy fibers; CA3 sends projections to CA1 via Schaffer collaterals; and CA1 sends its projections out of the hippocampus to either subiculum or entorhinal cortex. The hippocampus further contains recurrent projections within the CA3 region which have generated much speculation on the potential computational functions conferred by this architecture (Marr, 1971; Levy, 1996; Lisman, 1999; Treves, 2004; Koene et al., 2003; Koene and Hasselmo, 2008). The hippocampus and related parahippocampal areas receive projections from and project to many cortical and subcortical areas (Johnston and Amaral, 1998). The organization of these projections and the connectivity differences between the rodent dorsal and ventral hippocampus[10] suggest that the specificity of spatial information is much higher within dorsal hippocampus than in ventral hippocampus (Jung et al., 1994; Johnston and Amaral, 1998; Moser and Moser, 1998).

Early hippocampal recordings distinguished between local field potential activity associated with volitional movement and exploration characterized by strong 7-12Hz (theta) oscillations and activity associated with awake or asleep immobility characterized by large irregular activity (LIA) fluctuations and 2-4Hz low frequency (delta) oscillations (Vanderwolf, 1969, 1971). Hippocampal local field potentials also display brief periods of coordinated sharp wave activity that contain high frequency $\sim$ 200Hz ripples during LIA epochs (O'Keefe and Nadel, 1978, more recent investigations have also found ripples during theta epochs as well; see O'Neill et al., 2006). The relationship between single unit spiking activity

---

[10] The dorsal or septal hippocampus in the rodent roughly corresponds with the posterior hippocampus in the primate and the ventral or temporal hippocampus in the rodent roughly corresponds with the anterior hippocampus in the human.

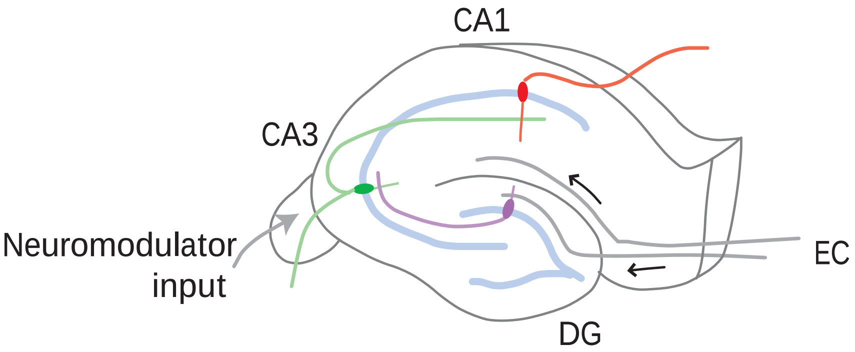

Figure 1.1: Basic circuitry of the hippocampus. Superficial layers of the entorhinal cortex project to hippocampus; layer II projects to dentate gyrus and CA3 while layer III directly projects to CA1. Dentate gyrus projects to CA3. CA3 contains a dense set of recurrent connections (not shown), projects to the contralateral CA3 and CA1 via the anterior commissure, and projects to CA1. CA1 projects to subiculum and deep layers of entorhinal cortex (after Neves et al., 2008).

and the phase of one or more of these oscillations has been used to characterize the various types of projection neurons and interneurons found in the hippocampus (Klausberger et al., 2003; Somogyi and Klausberger, 2005). Hippocampal projection neurons (pyramidal neurons in CA1 and CA3 and granule cells in dentate gyrus) display sparse spiking activity that is modulated by theta oscillations during waking behavior and by sharp wave ripple activity during sleep/awake immobility (O'Keefe and Nadel, 1978; Klausberger et al., 2003; Somogyi and Klausberger, 2005; Buzsáki, 2006). Because projection neurons ostensibly provide the primary pathway for neural information transmission, the relationship between spatial behavior, single unit spiking activity and local field potential activity has been viewed as a critical component in understanding the contribution of the hippocampus to cognitive map function and spatial behavior. These relationships are discussed below.

1.3.1 Spatial behavior and the hippocampus

A standard characterization of the hippocampus dependent map-based spatial behavior is provided by a series of experiments using the water maze (Morris et al., 1982, 1986). Within the water maze an animal learns to find a platform hidden beneath the surface of the water by using location information derived from distal landmarks (Morris, 1981). Animal behavior on the water maze is characterized by increasingly direct paths to the platform location from any initial position over the course of training and, in the absence of the platform, selective search of the location of the missing platform. In agreement with other forms of hippocampus dependent spatial behavior (O'Keefe and Nadel, 1978), behavior on the water maze is dependent on distal cues and shifts of these distal cues produce corresponding shifts in the selectively searched quadrant (Morris, 1981). Further evidence suggests that water maze behavior is dependent on a constellation of cues rather than individual cues: removal of a subset of spatial cues, or the use of partial distal cue information, causes negligible behavioral deficits (Nakazawa et al., 2002). Hippocampal lesions or inactivation produce massive deficits in acquisition and performance within the water maze as measured by either path length or search selectivity (Morris et al., 1982, 1986). In the partial cue version of the water maze, CA3 specific genetic ablation of NMDA receptor compromises quadrant selective search patterns (Nakazawa et al., 2002). Animals with lesions to the hippocampus are capable of solving the water maze if initially placed in an identical location and orientation within the water maze apparatus across training (Eichenbaum et al., 1990). Similarly, water maze behavior becomes independent of the hippocampus when animals are trained with a constant initial position and orientation (Whishaw and Mittleman, 1986; Eichenbaum et al., 1990).

Behavior on the water maze is characterized by the interaction of spatial information (map-relations) and goal information (map-content) that leads to the production of an efficient route to the platform. Protocols that compromise access to visual information interfere with normal water maze acquisition and performance and suggest that behavior on the water maze is at least partially dependent on a continuous stream of visual information (Arolfo et al., 1994). This evidence supports a tight coupling between spatial location and the expected sensory content for a given location as would be suggested within piloting-based navigation. *But how and when are these trajectories constructed?* While it might be argued that an efficient trajectory to the platform is constructed immediately following identification of the animal's position in the water maze (Redish and Touretzky, 1998a), the continued use of sensory input suggests that trajectory computations integrate sensory information throughout the trajectory (Arolfo et al., 1994). Moreover, the search trajectories are also dependent on previously experienced patterns of reward (Olton, 1979). The active integration of environmental cue information and patterns of previous reward become critical considerations within more naturalistic foraging situations where trajectories must be constructed and coordinated with respect to multiple potential reward sites.

The organization of multiple goal trajectories and its relationship to spatial working memory has been studied in a series of experiments by Olton and colleagues. The radial arm maze is characterized by a set of alleys emanating from a central platform (Olton and

Samuelson, 1976; Olton et al., 1979, 1980). In the (non-delayed) spatial working memory version of the task, each arm is baited once within a session. In the delayed version of the spatial working memory task animals are removed from the track after several arm entries and returned after a delay. Errors are classified as entries into unbaited (previously selected) arms. Spatial working memory on the radial arm maze is hippocampus dependent in the delayed version of the task (Rawlins, 1985; Jarrard, 1993; Redish, 1999; but see Barnes, 1988 for a contrasting view). It is also dependent on the the connections between ventral CA1/subiculum and prelimbic mPFC in the delayed version of the task and connections between ventral CA1/subiculum and ventral striatum in the non-delayed version of the task (Floresco et al., 1997). These studies suggest that interactions between the hippocampus and other brain areas, specifically the ventral striatum and prelimbic mPFC, organize goal-directed spatial behavior.

Rats within the radial arm maze display organized spatial trajectories and rarely visit previously visited arms even across long delays (Olton and Samuelson, 1976; Olton et al., 1979). Moreover, the decision to select a given arm is primarily related to whether it was previously visited; most evidence suggests that each decision is independent of the order of the previously visited arms (Olton and Samuelson, 1976; Brown, 1992; an upper bound of approximately 15 items has been calculated for the capacity of working; as noted by Olton, 1979). This suggests that rats do not utilize a single generalized problem solving strategy for solution of the radial arm maze but determine trajectories according to spatial working memory as a sequential decision process. Consistent with a sequential decision process, maze arm entry is often associated with vicarious trial and error or microchoice behavior (Brown, 1992). Vicarious trial and error behaviors are dependent on the hippocampus (Hu and Amsel, 1995) and correlated with increased hippocampal metabolism (Hu et al., 2006). The macroscopic and microscopic organization of goal-directed behavior within the radial arm maze and its relationship to hippocampal function suggests that the hippocampus contributes to a sequential decision process based on spatial working memory stores, but does not address what specific contribution the hippocampus makes (e.g. what is the content of spatial working memory) to the creation of different goal trajectories.

A recent study by Tse et al. (2007) on the construction and use of spatial schemata provides an important insight on the contribution of the hippocampus to the organization of behavior. Rats in this study were initially trained on odor-place paired associates. The presentation of a specific odor indicated that a reward was available at a specific position (food-well) within the task arena. Animals were initially trained with six paired associates and slowly acquired the task over multiple sessions. Hippocampal lesions prior to training impaired basic task acquisition. Intact rats were trained on two new paired associates following initial training of the six paired associates. These animals displayed nearly immediate acquisition of the new paired associates. Subsequent lesions of the hippocampus produced no behavioral deficits for either the initial six paired associates or the subsequently learned two paired associates but did block further odor-place paired associate learning. In contrast, hippocampus intact animals quickly learned subsequently presented paired associates. The explanation of these data by Tse et al. (2007) suggests that the hippocampus mediates the construction of integrated spatial representations or spatial schemata that allow fast integration of new information. Hippocampal lesions suggest that the hippocampus is required for construction of spatial schemata and integration of new information into existing schemata. Although the hippocampus is not required for production of previously learned cue-specific spatial trajectories, these results suggest that previously learned spatial information directly contributes to the construction of trajectories relative to novel cues within the paired-associate task.

In order to understand how the intrinsic organization of information within spatial schemata contributes to learning, Tse et al. (2007) explored the the effects of consistent versus inconsistent paired-associate training. In this experiment, rats were trained with six consistent odor-place pairings in one room and trained with six odor-place pairings that varied throughout training in a second room. These animals were then trained on two novel odor-place pairings (and these novel pairs were identical between the experimental rooms – identical odors and relative positions). The rats learned the two new odor-place pairings in the consistent condition room while the rats did not learn in the inconsistent condition room.[11] Tse et al. (2007) suggest that the rats were able to form a spatial schema that contributed to later learning and behavioral performance in the consistent condition while they were unable to do so in the inconsistent condition.

The study by Tse et al. (2007) suggests that a search trajectory constructed on the basis of a cognitive map is much more efficient than one formed without access to the hippocampal component of a cognitive map because it includes information about what locations were *not* suggested by the cue as well as what locations were suggested by the cue. Furthermore, the studies described above suggest an intriguing interaction between the hippocampus and other brain areas (e.g ventral striatum and prelimbic mPFC) in the organization of flexible, goal-directed spatial behavior, particularly during early learning regimes. These results

---

[11] It would be interesting to see whether rats with hippocampus lesions in the inconsistent condition learn the novel paired associates more quickly than hippocampus intact animals due to hippocampus-mediated interference.

underscore the organization of cognitive map-based behavior and the specific contribution of the hippocampus. These observations further suggest the computations performed by the hippocampus and a set of potential coding strategies and dynamics within hippocampal neural activity.

1.3.2 Place cells

Observed patterns of single unit spiking activity within the hippocampus are highly correlated with an animal's spatial location (O'Keefe and Dostrovsky, 1971; Ono et al., 1991a,b; Ekstrom et al., 2003; Ulanovsky and Moss, 2007). Hippocampal pyramidal neurons or *place cells* show increased spiking activity at a specific location called a *place field* (O'Keefe and Dostrovsky, 1971; O'Keefe and Nadel, 1978). Individual place fields for a single cell are stable within a single experimental session and can be stable between experimental sessions separated by periods up to several months (Thompson and Best, 1990). Like many of the spatial behaviors discussed above (Morris, 1981), place cell activity is closely coupled with distal visual cues and place field shifts are predicted by experimental cues shifts (Muller and Kubie, 1987). But place cells are not wholly dependent on distal visual cues; they are stable following cue removal (O'Keefe and Conway, 1978; Muller et al., 1987; Markus et al., 1994) and experientially decouple from highly variable cues (Knierim et al., 1995). These basic results suggest that place fields form a consistent map of the animal's environment by integrating distal multisensory cue information.

Place fields appear to be randomly and equivalently distributed across a given environment over a wide variety of tasks (O'Keefe and Nadel, 1978; Redish et al., 2001). Although several studies (Hollup et al., 2001b,a; Hok et al., 2007) have found slightly increased density of place fields associated with goal locations, place cell activity and place field distributions are generally independent of reinforcement. Further studies have examined the distribution of place fields relative to anatomical position within the hippocampus. In contrast with other brain areas which display a topographic anatomical organization of a represented information, analysis of the place field distributions suggests that no such anatomical topography is found within the hippocampus (O'Keefe and Nadel, 1978; Redish et al., 2001).

Map stability and spatial performance. Place cells display different place fields in different environments (O'Keefe and Conway, 1978; Muller and Kubie, 1987). The distribution of place fields in one environment does not predict the distribution of place fields in a second environment and suggests that the hippocampus maintains independent maps of each environment (Muller and Kubie, 1987; Guzowski et al., 1999; Redish, 1999). Place cells display pattern separation/pattern completion characteristics in animals that have been previously trained in two different environments (e.g. circle versus square) and were subsequently tested in cue-conflict or intermediate environments (Lee et al., 2004b; Wills et al., 2005; Leutgeb et al., 2005a). The coherent, coordinated shifts in place field distributions correspond to several predicted computations related to attractor dynamics and suggest that the hippocampus competitively selects a single map for a given environment over long timescales (Redish and Touretzky, 1997; Samsonovich and McNaughton, 1997; Redish and Touretzky, 1998a; Tsodyks, 1999; Doboli et al., 2000; Tsodyks, 2005).

A variety of studies have examined the stability of spatial maps within a single environment. Manipulation of behavioral tasks (or task contingencies) influence map stability and the distribution of place fields within a single environment (Markus et al., 1995; Moita et al., 2004; Kentros et al., 2004). Aged rats displayed increased place field instability across repeated sessions of the same task compared to controls (Barnes et al., 1997). Moita et al. (2004) showed that contextual fear conditioning produced place field remapping in rats. And Kentros et al. (2004) showed that whereas place fields were relatively unstable for a standard open-field foraging task in mice, they were much more stable in an attention-demanding version of the task. Kentros and colleagues further demonstrated that NMDA receptor blockade compromised the development of long-term place field stability in novel environments (Kentros et al., 1998) and a D1/D5 dopamine agonist (SKF 38393) facilitated place field stability even without a behavioral task while a D1/D5 dopamine antagonist (SCH 23390) compromised place field stability within a foraging task. A number of studies have also shown a link between the observed hippocampal map and behavior and suggest the active use of hippocampal spatial representations underlies spatial navigation (O'Keefe and Speakman, 1987; Lenck-Santini et al., 2001, 2002; Rosenzweig et al., 2003). These observations suggest that place field maps support spatial behavior and map stability is related to the cognitive demands of the behavioral task.

Trajectory coding. In contrast with place fields observed on random foraging tasks, place fields on alley-based tasks are sometimes dependent on the rat's running trajectory. Place fields on the linear track are dependent on the rat's running direction (McNaughton et al., 1983) and place fields on the common stem of spatial alternation tasks are dependent on the ($R \rightarrow L$, $L \rightarrow R$) trajectory (Wood et al., 2000; Frank et al., 2000). Ferbinteanu and Shapiro (2003) extended these initial observations to show that separate trajectory dependent place cell populations predict the direction the rat will turn and the direction from which the rat came on a plus maze task while a third population displayed no trajectory dependence. Although trajectory dependent place fields have been examined on tasks more complicated than linear track and spatial alternation tasks, the trajectory dependence of place fields is task dependent (Eichenbaum et al., 1987; O'Keefe and Speakman, 1987; Ainge et al., 2007a,b) and also dependent on the specific training regimen used (Bower et al., 2005).[12] These observations suggest that the hippocampus parses behavioral tasks into spatial components. Further evidence also suggests a similar parsing of non-spatial task components (Eichenbaum et al., 1987; Wood et al., 1999; Griffin et al., 2007) and supports the general hypothesis that hippocampal pyramidal neurons code task relevant information in both hippocampus dependent and hippocampus independent tasks.

1.3.3 Cell assemblies

In an attempt to describe the neural substrates of learning, Hebb (1949) formulated the concept of the cell assembly. Hebb (1949) hypothesized that cell assemblies were formed from multiple correlated cellular activities based on the rule that coincident activation of two neurons with shared connections strengthened these connections and subsequently allowed the activity of one cell to propagate to another. Hebb (1949) identified two primary consequences of this formulation of learning. First, a cell assembly provided the basis for construction of an integrated representation across multiple stimuli (similar to the proposal of Gestalt psychologists). This is a forerunner of modern pattern recognition and pattern completion algorithms (Hopfield, 1982; Hertz et al., 1991). Second, the temporal dynamics of cell assemblies suggested that the coordinated, sequential propagation of cell assemblies (what Hebb called a "phase-sequence") allowed access to information not present in immediate sensory experience. Evidence for the basic plasticity rule Hebb proposed was first discovered in the hippocampus (Lomo, 1971, but has now been found in many other brain areas Bear, 1996) and further evidence suggests that hippocampal pyramidal neurons form cell assemblies (Harris et al., 2003; Jackson and Redish, 2007).

Cell assemblies within the hippocampus Place cell activity is much more variable than would be expected given the set of model parameters typically used to describe single unit place field activity (e.g. position dependent spike emission as a Poisson process; Fenton and Muller, 1998). Fenton and colleagues hypothesized that this unexplained variability could result from fast-switching between multiple maps (Olypher et al., 2002). Consistent

---

[12] It should be noted, however, that the functionality of these observed codes remains in question. Ainge et al. (2007a) showed that place fields on a hippocampus dependent task displayed reduced trajectory dependence relative to place fields on a hippocampus independent task.

with this hypothesis, place cell activity on a linear track is better predicted by a combination of position, theta phase, and the set of simultaneously active cells than by a combination of position and theta phase alone (Harris et al., 2002). Harris et al. (2002, 2003) suggested the set of simultaneously active place cells represented cell assemblies. Jackson and Redish (2007) decomposed hippocampal place cell activity into two maps and showed that these maps dynamically shift during task performance and that dynamical map-switching could account for much of the previously unexplained variability observed within place cell activity. These data support the hypothesized formation of cell assemblies as coordinated simultaneous activity within hippocampal place cells. Further evidence for cell assemblies in the hippocampus comes from observations on the dynamics of place cell activity and hippocampal replay.

Place cell activity replays previously observed task activity during sleep episodes. Early observations indicated that place cell activity during task performance was predictive of subsequent cellular activity during sleep (Pavlides and Winson, 1989; Wilson and McNaughton, 1994). Skaggs and McNaughton (1996) showed that the temporal ordering of place cell spiking activity (as identified by cross correlations) during subsequent sleep episodes was similar to activity observed during task performance. And evidence from linear track tasks showed reactivation of full temporal sequences within sharp wave ripple activity during slow wave sleep (Nádasdy et al., 1999; Kudrimoti et al., 1999; Lee and Wilson, 2002) and during REM sleep (Louie and Wilson, 2001). Based on these observations, the term *forward route replay* has been used to describe the time compressed reactivation observed during sleep because of the similarity between the temporal ordering of place cell spiking activity observed during sharp wave ripple activity and the temporal ordering of place cell spiking activity observed during task performance. Finally, replay observed during sleep following task performance is much more frequent and temporally ordered than replay observed during sleep prior to task performance (Wilson and McNaughton, 1994; Skaggs and McNaughton, 1996; Kudrimoti et al., 1999; Lee and Wilson, 2002) and suggests that replay develops as a function of experience.

More recent observations suggest that route replay occurs during both periods of sleep and wakefulness. Patterns of local field potential patterns within the hippocampus during awake behavior shift between epochs of high theta power associated with task performance and LIA epochs associated with periods of awake immobility (Vanderwolf, 1969, 1971). Much like LIA observed during sleep, sharp wave ripple activity observed in periods of awake immobility (O'Keefe and Nadel, 1978) are associated with replay (Jackson et al., 2006; O'Neill et al., 2006; Diba and Buzsàki, 2007; O'Neill et al., 2008). Sharp wave associated replay during awake states displays more complicated dynamics than those observed during sleep and include backward route replay (Foster and Wilson, 2006; Csicsvari et al., 2007; Diba and Buzsàki, 2007), forward route replay (Diba and Buzsàki, 2007) and mixed replay dynamics (personal observation and observations from Matt Wilson). Foster and Wilson (2006) showed that backward route replay occurs on the first lap of a linear track task and decays with experience. In contrast, sharp wave ripple-associated reactivation develops as a function of experience and behavioral repetition (Jackson et al., 2006; O'Neill et al., 2008) and suggests that forward and backward route replay are supported by fundamentally different neural mechanisms. These observations are consistent with cell assembly dynamics proposed by Hebb (1949) and suggest the formation of cell assemblies within hippocampal place cell activity.

Cell assemblies beyond the hippocampus. Hebb's cell assembly proposal suggested that a cell assembly was "... a diffuse structure comprising cells in the cortex and diencephalon, capable of acting briefly as a closed system, delivering facilitation to other such systems" (Hebb, 1949). This suggests that cell assembly dynamics observed within the hippocampus likely extends beyond the hippocampus to other structures. Current evidence supports this hypothesis (Qin et al., 1997). Hoffman and McNaughton (2002) found coordinated reactivation across multiple neocortical sites. Tatsuno et al. (2006) showed reactivation associated with a novel experience persisted in both the prefrontal cortices and hippocampus; while hippocampal reactivation lasted much longer than prefrontal reactivation after a single novel experience, Euston and McNaughton (2007) showed that prefrontal reactivation persists across much longer periods following practice. Ji and Wilson (2007) showed that sharp wave associated hippocampal replay observed during sleep coincided with replayed activity in primary visual cortex in a visually complex spatial alternation task and, like hippocampal replay, coordinated replay in the hippocampus and visual cortex developed experientially. And evidence from Pennartz et al. (2004) suggests that hippocampal sharp wave ripple activity propagates to ventral striatum during sleep. The observed correspondence between sharp waves and neocortical up-states (Battaglia et al., 2004) and that hippocampal replay slightly precedes cortical replay (by approximately 50ms; Ji and Wilson, 2007) suggests replay is centrally organized by the hippocampus during sleep.

Although these observations support the theory of cell assembly formation across multiple brain areas, there remains little evidence on the functional contribution of cell assemblies to online behavior. McClelland et al. (1995) suggested that memories are quickly stored as cell assemblies in the hippocampus and that these memories are subsequently consolidated to neocortex through slow organization of neocortical cell assemblies. The memory consolidation proposal is consistent with the broad connectivity of the hippocampus and the temporal gradient of hippocampal memory dependence (Zola-Morgan and Squire, 1990; Squire and Zola-Morgan, 1991; Squire and Alvarez, 1995; Teng and Squire, 1999) but suggests that the hippocampus makes no contribution to online behavior – it is simply a fast-learning observer and its contribution occurs through offline training of the neocortex.

The proposal by McClelland et al. (1995) for the development of cortical cell assemblies is consistent with Hebb's description of cell assemblies that could be used to support perceptual learning (Hebb, 1949). However, it markedly contrasts with the functional description originally presented by Hebb (1949) in which cell assemblies further provided the neural mechanisms for sequence prediction. While this online functional perspective on cell assembly function seems well aligned with the anatomy of the hippocampus, its patterns of place cell activity, and its necessity for behaviors such as vicarious trial and error, it is currently unknown whether the hippocampus organizes cortical and subcortical activity during waking behavior as it appears to do during sleep.

1.3.4 Mental time travel, spatial imagery and episodic memory

Discussions of human hippocampal function and the deficits caused by its damage are traditionally centered on episodic memory (Tulving, 1983, 1984, 2002). Episodic memory is characterized by *mental time travel* or the ability to re-experience a previous sequence of events based on memory. While damage to the hippocampus in humans produces profound deficits in episodic memory (Squire, 1992; Corkin, 2002; Tulving, 2002; Squire et al., 2004), debate on a non-human analog of episodic memory has been contentious (Tulving, 2002; Clayton and Dickinson, 1998; Roberts, 2002; Clayton et al., 2003; Emery and Clayton, 2004). Early descriptions of episodic memory were described as memories that united *what-where-when* information (Tulving, 1983, 1984). Tulving's more recent descriptions of episodic memory are focused on the autonoetic experiential recollection of one's own past (Tulving, 2002). This experiential definition has precluded most if not all non-human investigations of Tulving's version of episodic memory.

The term *episodic-like* memory has been used to describe recent observations that animals are capable of organizing behavior with respect to *what-where-when* information (Clayton and Dickinson, 1998; Eacott and Norman, 2004) and in response to the changing definition of episodic memory (Tulving, 2001, 2002). Because scrub-jays were able to locate a preferred food item based on how much time had passed since it was stored, Clayton and Dickinson (1998) argued for the presence of episodic-like what-where-when memory in scrub-jays. However, the specific nature of the temporal component of memory continues to be debated. Roberts et al. (2008) showed that memory performance in rats was better predicted by "how long ago" than by "when" while Eacott and Norman, 2004 have suggested that the temporal components of episodic-like memory are related to and embedded contextual information (*what-where-which*) that is dependent on the hippocampus and parahippocampal cortices.

More recent considerations of episodic memory have focused on mental time travel as the construction of previous experience and its potential for construction of future experience. Studies of long-term memory highlight the constructive aspects of human memory and its susceptibility to errors via post-hoc suggestion (Loftus and Palmer, 1974; Schacter, 1995). Within these studies, experimental manipulations that provide information following the original to-be-remembered-experience interfere with the original memory. Subjects within these studies recollect something they did not witness and do so with high confidence (Loftus and Palmer, 1974). In contrast to Tulving's approach to episodic memory that suggests mental time travel retrieves only previous experience (Tulving, 2002), a construction-based perspective suggests that mental time travel allows for false (unexperienced) memories to be constructed as well. Functional (Atance and O'Neill, 2001, *episodic future thinking*) and evolutionary arguments (Suddendorf and Corballis, 1997; Suddendorf and Busby, 2003) suggest that mental time travel allows an animal to experientially navigate both the future and the past. Indeed, Suddendorf has argued that the ability to re-experience the past is an epiphenomenon related to the evolution of the ability to think-ahead and imagine potential futures (Suddendorf and Corballis, 1997; Suddendorf and Busby, 2003; Suddendorf and Corballis, 2007). These function-based arguments (Atance and O'Neill, 2001; Suddendorf and Corballis, 1997; Suddendorf and Busby, 2003; Suddendorf and Corballis, 2007) on episodic memory are centered on the ability of an organism to predict the future and, consequently, organize its behavior appropriately.

Investigations of neurobiological substrates of mental time travel, spatial imagery and episodic memory have focused on a fronto-temporal network (Buckner and Carroll, 2007; Buckner et al., 2008). A recent study by Hassabis et al. (2007) demonstrates that individuals with damage to the hippocampus have profound deficits in spatial imagination. The hallmark of these deficits is not a failure to visualize any particular element of a scene, but an ability to integrate these elements into a coherent whole. These ideas parallel those of Buckner and Schacter (Buckner and Carroll, 2007; Schacter and Addis, 2007a,b; Schacter et al., 2007) who suggest that a critical aspect of cognition is the ability to richly imagine potential future circumstances. Functional imaging during spatial and episodic memory function is punctuated by coordinated activation of the hippocampus and frontal lobes (Buckner and Carroll, 2007).

1.3.5 Conclusions

Previous research on the neural substrates of cognitive maps has centered on the hippocampus. O'Keefe and Nadel (1978) hypothesized that cognitive maps were constructed within the hippocampus based on the observed deficits in many flexible spatial behaviors following damage to the hippocampus and the observation of hippocampal place cells. Many research findings on the hippocampal dependence of spatial behavior and place cell function have supported and elaborated this hypothesis. While O'Keefe and Nadel's hypothesis continues to provide insight into hippocampal function, some critics have argued that the strong spatial requirements and hippocampal focus have limited the utility of cognitive maps as a hypothetical construct (Cheng et al., 2007).

More recent research on place cell activity has focused on the formation and dynamics of cell assemblies. Place cells form stable maps and display pattern completion/pattern separation characteristics at long timescales (Leutgeb et al., 2005a; Wills et al., 2005). At fine timescales, place cell activity is organized into multiple cell assemblies (Harris et al., 2002, 2003; Jackson and Redish, 2007). Further evidence for the Hebbian cell assemblies comes from observations of hippocampal route replay and associated activity across multiple cortical (Qin et al., 1997; Hoffman and McNaughton, 2002; Euston and McNaughton, 2007; Ji and Wilson, 2007) and subcortical (Pennartz et al., 2004) areas during sleep. Although these studies on the formation and dynamics of cell assemblies have primarily focused on cell assembly phenomenology rather than function, they parallel recent research on the functionality conferred by episodic memory.

Recent studies of episodic memory have examined its role in planning and decision making. In contrast with previous purely phenomenological treatments of episodic memory (Tulving, 2002), recent efforts have focused on the utility of representational dynamics that underlie episodic (or episodic-like) memory (Clayton and Dickinson, 1998; Atance and O'Neill, 2001; Suddendorf and Corballis, 1997; Suddendorf and Busby, 2003; Eacott and Norman, 2004; Suddendorf and Corballis, 2007). These conceptualizations of episodic memory suggest that episodic memory[13] provides the basis for an organism to plan behavior according to a predicted or imagined experiential sequence of events. One particularly salient aspect of these descriptions of episodic memory is the coordinated integration of

---

[13] Or episodic-like memory for non-human animals.

multiple sensory modalities or contextualization of imagery. Integrated spatial imagery in humans is dependent on the hippocampus (Hassabis et al., 2007) and the contextualization of memory is dependent on the hippocampus and parahippocampal cortices in the rat (Eacott and Norman, 2004). And recent imaging studies have identified a fronto-temporal network that includes the hippocampus in the consideration of future events (Buckner and Carroll, 2007).

How such cognition is manifest in action continues to be a relevant, open question in research on spatial decision-making, planning and imagery. Although some critics have suggested that cognitive maps provide little utility for understanding the organization and neural substrates of behavior, this argument (Cheng et al., 2007) appears directed toward O'Keefe's hippocampus-centric proposal of spatial cognitive maps (O'Keefe and Nadel, 1978; O'Keefe, 1999). Indeed many recent of perspectives on hippocampal and episodic memory function are consistent with a more generalized conceptualization of cognitive maps (Tolman, 1948; Redish, 1999). These perspectives continue to emphasize the importance of spatial cognition, particulary as a vehicle for understanding human and animal cognition, but have disregarded many of the constraints suggested by purely spatial perspectives of cognitive map function. Furthermore, these perspectives have also emphasized active interactions between the hippocampus and other brain areas highlighted by research on cell assemblies (Harris et al., 2002, 2003; Jackson and Redish, 2007; Lee and Wilson, 2002; Jackson et al., 2006; Diba and Buzsàki, 2007; Foster and Wilson, 2006; Ji and Wilson, 2007) and the frontotemporal network activated during construction and consideration of possible future events (Buckner and Carroll, 2007). These perspectives suggest that cognitive map function occurs through interactions with other learning and memory systems that extend beyond the hippocampus.

1.4 Multiple memory systems

One critical development within general theoretical formulations of learning was the identification of multiple memory systems (O'Keefe and Nadel, 1978). That some mnemonic functions are devastated following brain area specific damage while other functions remain intact suggests that memory was not the single unified system previously hypothesized (Scoville and Milner, 1957; Cohen and Squire, 1980). Within this context, O'Keefe and Nadel (1978) suggested that cognitive maps mediated by the hippocampus comprise one memory system while another set of extra-hippocampal brain areas subserve a second set of memory function in which learning occurs through processes much like those proposed by Hull (1943).[14]

Analysis of behavior on navigation tasks (Restle, 1957; Barnes, 1979; Packard and McGaugh, 1996) demonstrates that the information used to solve a task and the set of brain areas that underlie behavior are experience and task dependent. On a simple plus maze task, rats initially use a hippocampus dependent place-strategy during early learning phases, but switch to a dorsal striatum dependent response-strategy with increasing experience in later learning phases (Restle, 1957; Packard and McGaugh, 1996; Barnes, 1979). This type of switch between multiple memory systems is typical of simple tasks within a sufficiently stationary environment (Redish, 1999). The switch from a place-strategy to a response strategy is highlighted by a developing insensitivity to environmental information. For instance, an animal that uses a response strategy and is released at an atypical location within an environment will produce a typical series of responses even though this response pattern leads to a previously non-rewarded location (Packard and McGaugh, 1996) or leads the rat directly into a wall at full speed (Carr and Watson, 1908). The reinforcement consequences of such task failures suggest that animals should quickly shift back toward an increased environmental sensitivity. And in many navigation tasks, rats quickly modify their behavior based on relevant environmental stimuli following failure to discover an expected reward or some more adverse consequence such as running into a wall at full tilt.

In sum, these examples from animal navigation suggest a dynamic interaction of multiple memory systems. This division produces a series of distinctions between systems based on the relative insensitivity to specific environmental changes and the hypothesized predictions or expectations used by an animal in task solution. The following discussion provides a brief overview of multiple memory systems from several non-spatial conditioning perspectives in order to further examine the specific predictive computations or expectation components that influence spatial behavior and cognitive map function.

1.4.1 Outcome valuation within instrumental tasks

Choice behavior can be characterized as the selection of actions that lead to reinforcement. The organization of action has been broadly studied using instrumental conditioning tasks in which reward delivery is contingent on a specific action (see Figure 1.2A). Multiple instrumental tasks display developing insensitivity to devaluation and support a distinction between multiple memory systems (Balleine and Dickinson, 1998; Adams and Dickinson, 1981). In a typical instrumental experiment, rats initially learn to press a lever for a food

---

[14] Given only behavioral observations, Tolman (1949) anticipated many features of the multiple memory systems debate and even provided a rudimentary outline of what he called multiple types of learning.

reward. When the lever-press response is subsequently devalued by changing the reward contingency, pre-feeding the animal to satiety prior to test, or pairing the food reward with illness (usually induced with lithium chloride), the lever-press response patterns are dependent on the extent of prior training. Animals within early learning phases quickly cease lever pressing following devaluation protocols, while animals within later, overtrained learning phases continue to perform the same pattern of lever-press response and appear oblivious to devaluation protocols. Sensitivity to devaluation is dependent on multiple brain areas including the ventral striatum (Balleine and Dickinson, 1998), prelimbic medial prefrontal cortex (prelimbic mPFC Killcross and Coutureau, 2003; Corbit and Balleine, 2003b), and posterior dorsomedial striatum (Yin et al., 2005). Developing insensitivity to devaluation protocols is dependent on anterior dorsolateral striatum (Yin et al., 2004) and infralimbic medial prefrontal cortex (Coutureau and Killcross, 2003) and compromised function within these areas leads to prolonged sensitivity to devaluation protocols.

Animal behavior within both instrumental learning and spatial learning tasks displays progressive formation of action habits that are insensitive to changes in outcome. Progressive development of behavioral habits in both instrumental and spatial learning is dependent on the dorsolateral striatum (Yin et al., 2005; Packard and McGaugh, 1996). The similarities between instrumental learning and spatial learning observed during habit regimes, both in terms of behavior and neural substrates, does not appear to extend to flexible and dynamic learning regimes. For instance, while spatial learning is typically dependent on the hippocampus (O'Keefe and Nadel, 1978; Redish, 1999), projection sparing lesions of the hippocampus produce no changes in the sensitivity to devaluation within simple instrumental conditioning (Corbit et al., 2002). These differences are likely attributable to the differential use of stimulus information; simple instrumental learning tasks are described in terms of action and outcome only while spatial learning tasks are described in terms of stimuli, actions and outcomes (see Figure 1.2). Dickinson and Balleine (2000) have argued that goal-directed action and causal cognition requires only (1) representation of current value of an outcome and (2) representation of the causal relationship between an action and its outcome. As such, experiments that test sensitivity to outcome devaluation provide only partial support for goal directed action and require further experimental tests for sensitivity to action-outcome degradation.

1.4.2 Instrumental contingencies

The observation that rats are sensitive to contingency degradation within instrumental tasks (Balleine and Dickinson, 1998) suggests that rats are capable of goal-directed action (Dickinson and Balleine, 2000). Analysis of the neural substrates for contingency representations has been performed in a series of studies by Balleine and colleagues (Balleine and Dickinson, 1998; Corbit and Balleine, 2000; Corbit et al., 2002; Corbit and Balleine, 2003a).

Corbit and Balleine made an interesting methodological variation to examining instrumental contingencies within their studies. Typical approaches to contingency degradation fully compromise a given contingency by modifying all action outcomes. Within lever-pressing tasks, contingency degradation protocols usually compromise the outcomes of lever pressing *and* non-lever pressing responses (Bolles, 1972). In contrast, contingency degradation in the studies by Corbit and Balleine was accomplished by modifying the outcome for only the no-lever press condition. Following typical instrumental conditioning protocols, Corbit and Balleine trained rats to lever-press for specific rewards (e.g. left lever for sucrose solution and right lever for food pellets). Following acquisition of instrumental responding, non-contingent reward of a single type was provided probabilistically at each lever (non-contingent sucrose solution was provided at each lever). For one lever, the non-contingent reward matched the contingent reward; for the other lever, the non-contingent reward did not match. Contingency degradation was hypothesized to be evidenced by reduced responding at the lever on which the contingent and non-contingent rewards matched.[15]

Corbit and Balleine (2000) initially examined the role of the hippocampus in encoding the causal relationship between actions and their consequences. This hypothesis followed from earlier work by (Devenport and Holloway, 1980; Devenport et al., 1981) that suggested that animals without a hippocampus were dependent on simple response relationships rather than richer action-outcome expectancies. In order to examine the type of expectancy information mediated by the hippocampus, Corbit and Balleine (2000) examined instrumental behavior following electrolytic lesions of the hippocampus and specifically sought to determine the dependence on reward devaluation and contingency degradation on hippocampal function.

Lesions of the dorsal hippocampus produced no deficits in devaluation but produced clear deficits in contingency degradation. Similar to sham-controls, hippocampus-lesioned rats responded much less for (satiety) devalued reward. In contrast, hippocampus-lesioned rats displayed reduced responding following contingency degradation protocols but did not distinguish between degraded and non-degraded contingency outcomes as sham-controls did. These results demonstrated that rats with electrolytic hippocampal lesions were unable to integrate multiple contextual action-outcome contingencies but were able to integrate new

---

[15] It should be noted that all tests of contingency degradation were performed during extinction; that is, no rewards were provided in these sessions.

value information into action selection.

A complication of electrolytic lesions is the potential for damaging fibers that project to downstream areas. To better assess the specific contribution of the hippocampus and related areas to instrumental learning, Corbit et al. (2002) examined contingency degradation following NMDA lesions to the dorsal hippocampus, subiculum, and entorhinal cortex. Replicating their earlier findings (Corbit and Balleine, 2000), electrolytic lesions of the dorsal hippocampus produced deficits in contingency degradation. In contrast, NMDA lesions of the hippocampus did not produce deficits in contingency degradation. Devaluation remained intact in rats with NMDA lesions of the hippocampus. Instead NMDA lesions of the retrohippocampal cortex caused deficits in contingency degradation but did not affect devaluation and can account for earlier observations (Corbit et al., 2002). Furthermore disconnecting the hippocampus and entorhinal cortex produced clear deficits in contingency degradation while no deficits were found for NMDA disconnection of the subiculum and hippocampus. These results indicate that the entorhinal cortex and input projections that pass through the dorsal hippocampus are required for acquisition and use of instrumental contingency information within behavior.

Theory

The contingency-based inference examined by Corbit and Balleine can be described using Bayes' rule. The goal of instrumental contingency inference is to determine whether a given action a_i will provide a given desired outcome $\hat{o}$ any more than a set of alternative actions a_j for all $j \neq i$.

$$p(a_i|\hat{o}) = \frac{p(\hat{o}|a_i)p(a_i)}{\sum_j p(\hat{o}|a_j)p(a_j)} \tag{1.1}$$

A typical stimulus-response treatment hypothesizes that acquisition of a response occurs independently of other actions and outcomes. This is equivalent to the argument that animals learn about a single action a_i given a desired outcome $p(a_i|\hat{o})$ rather than $p(a|\hat{o})$ for all actions. In contrast, the A-O approach used by Corbit and Balleine suggests that non-contingent actions (a_j where $j \neq i$) also contributes to selection of action a_i. In the protocol used by Corbit and Balleine, $p(o|\text{lever press})\,p(\text{lever press})$ and $p(o|\text{no lever press})\,p(\text{no lever press})$ were very different during initial conditioning. Because $p(o|\text{lever press})\,p(\text{lever press}) >> p(o|\text{no lever press})\,p(\text{no lever press})$, the probability of the desired outcome was much greater for pressing the lever than for not pressing the lever. However, $p(o|\text{lever press})\,p(\text{lever press})$ and $p(o|\text{no lever press})\,p(\text{no lever press})$ were approximately equal for the degraded contingency lever following contingency degradation. Consequently, neither pressing the lever nor not pressing the lever provided any greater access to the given desired outcome.

Contingency degradation protocols can be contrasted with reward devaluation protocols by considering the outcome of an action in the context of other potential actions. Reward devaluation protocols produce a change in the outcome of the single learned contingent action given a (supposedly) desired outcome $p(a_i|o)$ rather than in other non-contingent actions implied by $p(o|a_j)$ (where $j \neq i$). As a result, reward devaluation is a specific manipulation of the desired outcome o or motivation rather than any manipulation of contingencies $p(o|a_i)$ that are used for the development of causal or contextual expectancies.

1.4.3 Goals and outcomes

Tolman and Brunswik (1935) suggested that an animal learns the "causal texture" of its environment and, furthermore, that such learning provided the basis for the development of expectations and goal directed action (Tolman and Brunswik, 1935; Tolman, 1948, 1949). This approach to causal cognition fundamentally differs from Dickinson and Balleine's approach to causal cognition as goal-directed action (Balleine and Dickinson, 1998; Dickinson and Balleine, 2000) in its integration of stimulus information into action selection. Within this combined stimulus-action context, goals are composed of stimulus information that includes outcome (reward) information.

The influence of stimulus information on action has been most thoroughly explored within Pavlovian conditioning (Pavlov, 1927; Mackintosh, 1974). The action within a Pavlovian task is an unconditioned response (e.g. startle) initiated by the presentation of an unconditioned stimulus (e.g. shock). Both the unconditioned response and the unconditioned stimulus are species specific, inflexible and, presumably, a result of evolution. Pavlovian conditioning describes a process wherein a conditioned stimulus predictive of the unconditioned stimulus comes to illicit a conditioned response. The conditioned response is identical to the unconditioned response (and most often inflexible, evolutionarily defined). And though action is typically used to assess learning, Pavlovian conditioning is most often conceptualized as learning a stimulus-outcome association (Figure 1.2D). Pavlovian conditioning, like instrumental conditioning, is sensitive to post-training manipulations of outcome value (Holland and Straub, 1979). However, because it is not clear whether 'true' Pavlovian conditioning is sensitive to contingency degradation, Dickinson and Balleine (2000) argue that Pavlovian conditioning does not meet the criteria for goal-directed behavior but can contribute to instrumental responding (Kruse et al., 1983, Pavlovian instrumental transfer).

Investigations on the neural substrates of Pavlovian conditioning suggest a variety of brain areas and neural signals support this type of learning. One particularly salient observation is the correspondence between activity of midbrain dopaminergic neurons (Ljungberg et al., 1992; Schultz et al., 1993) and theoretical predictions of Pavlovian conditioning (Rescorla and Wagner, 1972; Sutton and Barto, 1981; Schultz et al., 1997). Midbrain dopamine neurons display increased activity following reward receipt during initial learning in Pavlovian conditioning tasks (Ljungberg et al., 1992; Schultz et al., 1993, 1997; Pan et al., 2005; Roesch et al., 2007). With further training, this heightened dopamine response associated with reward shifted to the conditioned stimulus. Consistent with the hypothesis that increased midbrain dopamine activity is associated with reward prediction, reward omission after acquisition of the Pavlovian association produced transient reductions in spiking activity relative to baseline spiking. Although midbrain dopaminergic activity clearly corresponds to predictive learning signals within simple Pavlovian conditioning tasks, the specific contribution of dopaminergic activity to decisions and action remains a matter of debate (Montague et al., 1995, 1996; Schultz et al., 1997; Schultz, 1998; Berridge and Robinson, 1998; Schultz, 2002; Berridge, 2006).

Midbrain dopamine neurons project to a variety of areas implicated in Pavlovian conditioning including the ventral striatum, amygdala and prefrontal cortex. Dopaminergic activity in the ventral striatum (Dalley et al., 2005) and medial prefrontal cortex (Grace and Rosenkranz, 2002) is required for acquisition of simple appetitive Pavlovian tasks. Lesion and inactivations studies have shown that Pavlovian decision tasks are dependent on the basolateral amygdala and orbitofrontal cortex as well (Schoenbaum et al., 1998; Pickens et al., 2003; Ostlund and Balleine, 2007). And neural signals associated with Pavlovian outcomes have been found in the ventral striatal (nucleus accumbens) (Nicola et al., 2004a,b; Yun et al., 2004) orbitofrontal cortex (Schoenbaum et al., 1998, 1999, 2000, 2005; Tremblay and Schultz, 1999; Padoa-Schioppa and Assad, 2006), and amygdala (Schoenbaum et al., 1998; Saddoris et al., 2005; Schoenbaum et al., 1999, 2000).

While these findings have important implications for more complicated decision processes, relatively little is known about the signals that support complex goal construction and evaluation. These complex goal processes are likely supported by coordinated neural activity across multiple brain areas. And though a portion of the signals and brain areas implicated in simple goal-directed behavior will play a role in more complex behaviors, how extensible these signals are for more complex goal-related behaviors remains an open question. Clearly, other signals and brain areas will likely be recruited for these more complex decision processes. A recent study by Ramus et al. (2007) showed that anticipatory activity observed within the OFC during a delayed matching task is dependent on the hippocampus and suggests that goal related activity observed in more complex tasks may be supported by a more complex network than required for simple tasks. This also suggests that some simple goal-related signals (and perhaps brain areas) may no longer be critical for complex goal-directed behavior. This raises the intriguing question of whether simple goal expectancy signals like dopamine activity continue to show reward correlates for covert cognitive processes that remain behaviorally unrealized.

1.4.4 Conclusions

Discussions of multiple memory systems often present each memory system as independent (O'Keefe and Nadel, 1978 – locale map versus taxon systems; Cohen and Squire, 1980 – what versus how systems; etc.). Although there are clear distinctions to be made between these systems in very simple behaviors, there are also clear interactions between these multiple memory systems in more complex, goal behaviors (Tolman, 1949; Kruse et al., 1983; Squire, 1987; Zola-Morgan and Squire, 1993). Cognitive maps were originally hypothesized to examine the interaction of these different memory systems (Tolman, 1948, 1949). While much progress has been made in delineating different memory systems and their neural substrates, less is known about their potential interactions, particularly as these interactions support cognitive function.

1.5 A modern theoretical/statistical/Bayesian approach to cognitive maps

A systematic formulation of Tolman's learning theory was first approached by MacCorquodale and Meehl (1954). Its reductionist perspective clearly frustrated Tolman (Tolman, 1955), but Tolman and the entire field lacked much of the appropriate vocabulary and mathematical and algorithmic sophistication to clearly articulate their ideas. Tolman's treatment of learning seems to fall much in line with recent hierarchical Bayesian approaches to behavior, reasoning and inference (Tenenbaum et al., 2006; Kemp et al., 2007; Griffiths and Tenenbaum, 2007). While Tolman did not explicitly formulate his theories in equations, he applauded early attempts to examine probabilistic learning that were the forerunners of modern probability-based (Bayesian) approaches to learning (Tolman and Brunswik, 1935). Hierarchical Bayesian approaches to learning and inference provide a systematic framework that appears to match much of Tolman's conceptualization of cognitive maps.

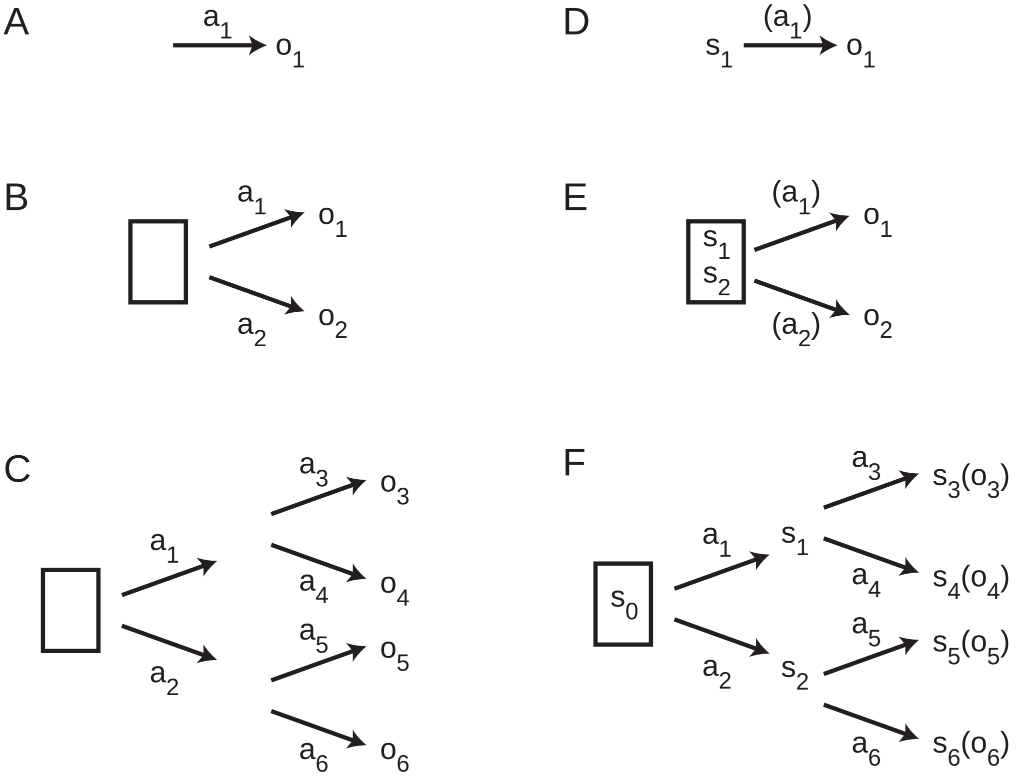

Figure 1.2: A graphical comparison of instrumental and Pavlovian conditioning and spatial learning. Balleine and colleagues have described goal directed behavior as sensitive to both outcome devaluation and contingency degradation. A Basic instrumental conditioning can be described as an action-outcome association. Animals display sensitivity to outcome devaluation only during early training (Adams and Dickinson, 1981; Balleine and Dickinson, 1998) and developing devaluation insensitivity is dependent on the dorsolateral striatum (Yin et al., 2004, 2006). B Multiple action-outcome paradigms have been used to assess the dependence of action-outcome contingencies for behavior. Action-outcome associations are dependent on prelimbic mPFC (Corbit and Balleine, 2003a) and the entorhinal cortex (Corbit et al., 2002). These tasks are dependent on context (shown as an open box). C More recently Balleine has shown that multi-action-outcome associations are dependent on medial agranular premotor cortex (unpublished observations). Note that Balleine's goal directed behavior is independent of stimulus information. D Pavlovian conditioning is described as stimulus dependent behavior wherein the unconditioned response to an outcome is associated with a conditioned stimulus s. E Pavlovian decision tasks are dependent on the basolateral amygdala and orbitofrontal cortex (Schoenbaum et al., 1998; Pickens et al., 2003; Ostlund and Balleine, 2007). F Spatial learning involves multi-step interaction of both stimulus and action -based processing. This is similar to causal texture of the environment proposed by Tolman and Brunswik (1935).

Three parallels between Tolman's cognitive maps and hierarchical Bayesian treatments of learning and inference are particularly important to consider. These are (1) the global computations required for both cognitive maps and Bayesian inference, (2) the explicit development of hypotheses for interpreting sparse data, and (3) the changing forms of information utilized to make decisions and how this organizes behavior. Note that the frequent discussion of *prior distributions* can be viewed as a component of development of hypotheses for interpreting sparse data. Let us provide a brief overview of Bayes' rule before turning to its use in learning and inference.

At its simplest, Bayes' rule describes the relationship between probabilities $p(a|b)$ and $p(b|a)$ where $p(a|b)$ is read as *the probability of a given b.* Bayes' rule can be written as

$$p(a|b) = \frac{p(b|a)p(a)}{\sum_{a' \in A} p(b|a')p(a')} = \frac{p(b|a)p(a)}{p(b)} \tag{1.2}$$

where the terms $p(a)$ and $p(b)$ are prior distributions describing the probability of observing a or b. In this form, Bayes' rule forms the basis for statistically appropriate expectations related to variable a based on the coincidence of another variable b. An important computational aspect of Bayes' rule is that the inference $a|b$ requires consideration of every alternative to a. In other words, given the observation of b, we must consider *all* possible a (that is all $a' \in A$). This global computation is a hallmark of any Bayesian treatment.

More generally, Bayes' rule has been used to examine multiple hypotheses given sparse data. The conditional form of Bayes' rule (equation 1.3) has been used to describe the probability of a particular hypothesis h given an observation o and a theoretical framework T.

$$p(h|o,T) = \frac{p(o|h,T)p(h|T)}{\sum_{h' \in H_T} p(o|h',T)p(h'|T)} \tag{1.3}$$

This conditional form of Bayes' rule suggests a hierarchical framework for inference: competing or overlapping hypotheses can be tested in light of observation data o and a greater, overarching theoretical framework T (see Figure 1.3). Tenenbaum et al. (2006) have used this framework to show how abstract rules and principles can be used and derived given only sparse language data. While hierarchical Bayesian perspectives have been used primarily within the arena of human inference (particularly within language), these perspectives can be more broadly applied to behavioral inference of animals, particularly the sort of behaviors and experiential regimes Tolman used to describe cognitive maps.

In terms of animal behavior, a hierarchical Bayesian treatment suggests that animals form and test hypotheses. These hypotheses structure behavior so that inferences can be made even with relatively little experience (sparse data). In order to develop several

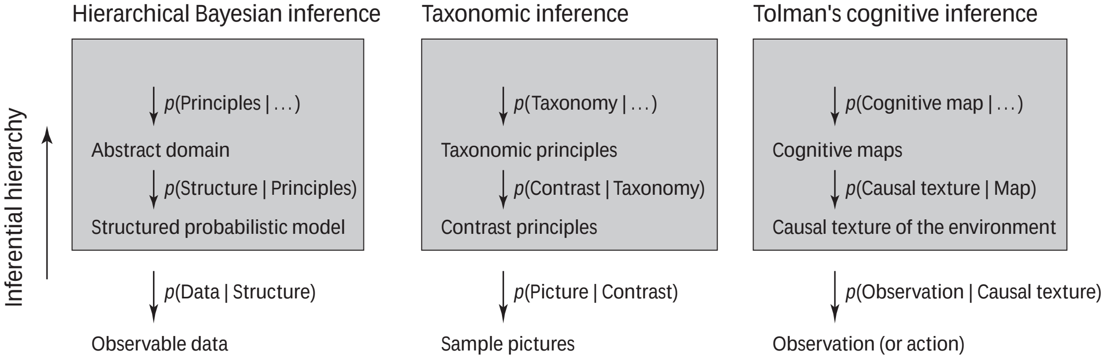

Figure 1.3: Hierarchical approaches to Bayes' rule for cognitive research (after Tenenbaum et al., 2006). The left column shows the basic structure for hierarchical Bayesian analysis. Observations are interpreted based on an inferential hierarchy. At the lowest level of this hierarchy are structured probabilistic models that are explicit hypotheses on the distribution of observations. Higher levels allow comparison of multiple probabilistic models relative to data and abstract domain principles. And these hierarchies can be further extended to include higher order theoretical principles. The central column shows how hierarchical Bayesian analysis has been used for taxonomic inference for pictures by Tenenbaum and Xu (2000). Within this example, low hierarchical levels are used for analyzing picture contrast and higher hierarchical levels are used for category and word selection (taxonomy). The right column shows an interpretation of Tolman's ideas on cognitive inference using a hierarchical Bayesian approach. Tolman argued that animals learn the causal texture of the environment and led to the formation of cognitive maps and higher order cognitive structure (Tolman and Brunswik, 1935; Tolman, 1948, 1949). Hierarchical Bayesian approaches explicitly suggest how cognitive maps fundamentally alter an animal's perception of its environment, its remembrance of prior experience and, consequently, its inference (Tolman, 1949).

important conceptual points related to this treatment, let us examine two sets of behavioral experiments on the partial reinforcement extinction effect (Capaldi, 1957) and the use of spatial schema by rats (Tse et al., 2007).

1.5.1 Experimental evidence

Extinction and the partial reinforcement extinction effect. The partial reinforcement extinction effect (PREE) highlights the importance of prior experience and expectations. The basic observation can be described as follows. While animals initially trained with deterministic reinforcement schedules (e.g. FR1 lever press schedule) abruptly extinguish responding when the reward is made unavailable, animals that are trained with probabilistic reinforcement schedules display an increased resistance to extinction (Capaldi, 1957). Hilgard and Bower (1975) cite Tolman and Brunswik (1935) and their development of an organism's probabilistic interactions with its environment as one of the first theoretical predictions of PREE.

Statistical treatments of PREE (Courville et al., 2006; Redish et al., 2007) suggests how the level of uncertainty within the initial reinforcement schedule influences extinction. The initial probabilistic training schedule produces an expectation that a lever press might not yield reward. Such an expectation means that when a lever press does not yield reward, it could be caused by either (1) a continuation of the probabilistic training schedule or (2) a new training schedule. As a result, the expectation or prior developed through the initial training profoundly affects the interpretation of a lever press without reward and how an animal responds. Animals initially trained with a deterministic reward schedule have no basis for interpreting the failed lever press as continuation of the original schedule and can, consequently, interpret the failed lever press as the signal for a new training schedule.

A number of interpretations identify extinction as new learning and not unlearning. These treatments suggest that during extinction a new cognitive state or context is learned wherein a given action fails to deliver on a previously learned expectation (Bouton, 2002; Redish et al., 2007). The formation of separate and separable states in which previous learning is valid or is invalid creates an inferential hierarchy based on the construction of new cognitive or inferential structures. Tolman (1949) predicted the formation of such new cognitive states and their influence in inference. He wrote, "in the course of the usual learning experiment there may be acquired not only a specific [*cognitive map*] but also new modes or ways of perceiving, remembering and inferring...which may be the utilized by the given organism in still other later environmental set-ups."

Spatial schemas. Spatial schemas provide a useful starting point for considering how new cognitive states contribute to and organize behavior. The recent study by Tse et al. (2007) on spatial schemas in rats highlights several important aspects of cognitive maps and hierarchical Bayesian inference. As previous described, rats were initially trained on odor-place paired associates. The presentation of a specific odor indicated that a reward was available at a specific position within the task arena. Rats were trained with six consistent odor-place pairings in one room and with six odor-place pairings that varied throughout training in a second room. Animals were then trained on two novel odor-place pairings (and these novel pairs were identical between the experiment rooms – identical odors and places). The rats trained in the consistent condition learned the two new odor-place pairings while they did not in the inconsistent condition. Tse et al. (2007) suggest that the rats in the consistent condition were able to form a spatial schema that contributed to later learning while the rats in the inconsistent condition were unable to form this spatial schema.

But what is a spatial schema and how does it organize behavior? The critical distinction between the two conditions in this experiment is the probability that a single odor would signal a single location. This is what Tolman and Brunswik (1935) called the causal texture of the environment. In the consistent condition, this probability was one ($p = 1$) while, in the inconsistent condition, the probability was less than one ($p < 1$). The argument for schema-use follows that when animals were able to form a spatial schema for the odor-place pairings (consistent condition), they were able to learn the set of novel pairs because they also knew where the locations were not (because they were already taken by other odors). This corresponds to the global computation with Bayes' rule that one must integrate all possible alternatives (all $a' \in A$ and that includes *not* a) in order to correctly make an inference and precisely matches cognitive map function.[16]

Furthermore, the experiment by Tse et al. (2007) shows that animals in the consistent condition were able to learn novel odor-place pairs much more quickly with experience than when they had little experience. That is, animals with increasing experience were able to interpret and structure their behavior appropriately given many fewer training trials. The authors' argument that the development of a spatial schema allows animals to more easily learn fits well within Bayesian treatments of learning from sparse data.

---

[16] Tolman argued that both the expectation of food availability at the end of one path and the expectation of food unavailability at the end of other paths contribute to an animal's choice (p.180). While this point was derisively noted by MacCorquodale and Meehl (1954), it is well-suited to a Bayesian framework.

1.5.2 Discussion.

The vast majority of Tolman's experiments on animal cognition were focused on understanding inference or, in Tolman's words, the animal's sign-gestalt expectancy. In contrast to the experiments of Hull and his contemporaries, Tolman's experiments examined animal behavior during early learning regimes when animals had relatively little (or sparse) experience. Krechevsky and Tolman's observations that animals appeared to organize and structure their behavior even when many theories of their contemporaries suggested otherwise led them to wonder whether these animals were in fact testing hypotheses and what cognitive factors formed the basis for these hypothesis behaviors (Tolman and Krechevsky, 1933). Tolman and Brunswik (1935) suggested that animals maintained some representation of *the causal texture of the environment* and that this served as the basis for inference. *Cognitive maps* represented Tolman's best efforts to define how animals learn this causal structure of the environment and, consequently, how this structure organizes behavior.

The difficulties Tolman encountered in examining early learning regimes continue to be a challenge for current cognition research. And while many researchers have sought to avoid this challenge by concerning themselves primarily with asymptotic learning regimes, inference in the face of sparse data remains a central question for cognitive scientists. That Tolman's conceptual development of cognitive maps so well aligns with Bayesian treatments of learning is no doubt a result of continued interest in this question.

There exist three primary points of intersection between Tolman's cognitive maps and Bayesian treatments of learning. The simplest point of intersection is inference based on sparse data. Tolman uses the cognitive map as an explicit attempt to explain the apparent reasoning and structure found within animal behavior, particularly within novel environments and novel experimental perturbations that set one set of cues against another. Shortcuts, hypothesis testing and latent learning highlight this type of inference. This parallels Bayesian treatments of sparse data given a set of competing hypotheses, particularly within studies on cue competition and integration (Cheng et al., 2007; Battaglia and Schrater, 2007).

The second point of intersection follows from Tolman's expectancies and 'commerce' with environmental stimuli (Tolman, 1932). Tolman suggests that an animal will examine its environment and reduce uncertainty through vicarious trial error and searching for the stimulus. Each of these types of behavior depend on some representation of the causal texture of the environment and lead to information-seeking behaviors when observations and predictions diverge. These integrated representations well match the computations of Bayes' rule: new information propagates throughout the entire network – not simply within a single dimension or information level. And the prediction-observation comparisons highlighted by searching for the stimulus strongly resemble Bayesian filter approaches used in machine learning and robotics (Thrun et al., 2005).

The third point of intersection is the hierarchical approach to learning. Tolman (1949) explicitly identifies multiple interactive types of learning and suggests a hierarchy of learning processes. He argues that these different learning systems have different constraints and their interaction results in "new modes or ways of perceiving, remembering and inferring" over the usual course of learning. Hypothesis testing and vicarious trial and error highlight the development of hierarchical inference – Tolman maintains these behaviors represent an active search for higher environmental principles rather than simply a random search for a successful stimulus-response pattern. This hierarchical aspect of learning directly corresponds to Bayesian treatments of language that examine the hierarchical development of rules, principles and theories (Tenenbaum et al., 2006).

1.6 Conclusions

The proposal that animals, even as simple as rats, might possess internal models of the world was not novel within 1930's behaviorist literature (Hull, 1930). However, Tolman's treatment of cognitive function in animals left some critics to wonder whether animals would be left '*buried in thought*'. Guthrie (1935) wrote:

> *In his concern with what goes on in the rat's mind, Tolman has neglected to predict what the rat will do. So far as the theory is concerned the rat is left buried in thought; if he gets to the food-box at the end that is [the rat's] concern, not the concern of the theory.*

Tolman (1955) saw this type of getting lost in thought as a natural consequence of learning and cognitive function, particularly in humans and even more so those in academics (Tolman, 1954) and he left the problem of how cognition, particularly cognition related to deeper inferential processes, is manifested in behavior only loosely specified. This problem increasingly drove many behaviorists away from cognitive map based formulations of learning and inference and toward reduced behavioral paradigms in which the organization of action can be simplified.

The following chapter develops the theoretical basis of cognitive maps and how it influences action.

Chapter 2

Reinforcement learning

The theoretical arguments and perspectives developed in this thesis are based on considerations of reinforcement learning. Reinforcement learning has been approached from a number of directions and this thesis emphasizes animal and machine learning perspectives. The theoretical considerations and complex applications of reinforcement learning have produced a rich framework for understanding information dynamics and the processes that contribute to learning (Sutton and Barto, 1998). The goal of this chapter is to develop a series of qualitative and quantitative predictions about memory, its organization and its use based on reinforcement learning.

The machine learning approach to reinforcement learning is grounded in optimization. *What behaviors should an agent*[1] *perform in order to maximize reward receipt and minimize cost?* Within traditional machine learning approaches to reinforcement a robot learns a series of actions (a *policy*) that will maximize the reward and minimize cost over a state-space defined by sensor information. And though there are clearly situations in which sensor-information is unwieldy, there remains a fundamental, and unanswered question of how state is optimally defined. For simplicity and tractability, state is often defined *a priori* according to a static set of rules and it is over this state-space that the agent learns. While this simplification is generally acceptable for certain types of problems, it becomes less tenable for problems of increasing complexity and interest. We will return to state definition issues later in this chapter, particularly the implication of acquisition of state information and some intriguing issues associated implementation of state-estimation in uncertain environments. For the present, let us turn to a formal development of policy evaluation with a known state-space.

---

[1] In models of animal behavior, the agent refers to the modeled animal.

2.1 Basic formalisms

The value of a state given a policy π is defined as the total expected future reward that will result from occupying state s at time t (Sutton and Barto, 1998). This is written as

$$V^{\pi}(s_t) = \sum_{\tau \geq t} E\left[r_\tau | s_t\right] \tag{2.1}$$

where $V^{\pi}(s_t)$ is the value of state s_t, r_τ is the reward at time τ and $E[\cdot]$ is the expectation. By defining a reward function $\mathbf{R}_{s,s',r} = p(r_t = r|s_t = s, s_{t+1} = s')$ and a Markov transition model $\mathbf{T}_{s,s'} = p(s_{t+1} = s'|s_t = s)$, the value function can be defined recursively.

$$\begin{aligned} V(s_t) &= E\left[R_{s_t}\right] + E\left[V(s')\right] && (2.2)\\ &= E\left[R_{s_t}\right] + \sum_{s' \in S} \mathbf{T}_{s,s'} V(s') && (2.3)\\ \mathbf{V} &= E\left[\mathbf{R}\right] + \mathbf{TV} && (2.4) \end{aligned}$$

For a Markov decision process, the value function $V^*(s)$ for an optimum policy $\pi^*(s,a)$ can be found according to

$$V^*(s) = \max_a \sum_{s'} \mathbf{T}_{s,a,s'} \left(r\mathbf{R}_{s,a,r} + V^*(s')\right) \tag{2.5}$$

where $\mathbf{T}_{s,a,s'} = p(s_{t+1} = s'|s_t = s, a_t = a)$. However, the recursion can be extended for multiple steps given $\mathbf{T}$ and provides a basic program for estimating $\mathbf{V}$ and optimizing agent control (Bellman, 1957; Daw, 2003).

Within the previous derivation, action is absorbed into the state definition and the transition model. This is not a requirement of the Bellman equation. Q-learning describes a common method for computing the value function over paired state-action pairs (that is $\langle s, a\rangle$ instead of simply s) for a given policy (Watkins and Dayan, 1992). While the subtle change to the definition of state does not greatly influence the validity (or convergence) of the Bellman equation, the optimal policy $Q^*(s,a)$ can be found according to

$$\begin{aligned} V^*(s) &= \max_a Q^*(s,a) && (2.6)\\ Q^*(s,a) &= \sum_r r\mathbf{R}_{s,a,r} + \sum_{s'} \mathbf{T}_{s,s'} V^*(s'). && (2.7) \end{aligned}$$

Note that the term $\mathbf{R}_{s,s',r}$ has been modified to $\mathbf{R}_{s,a,r}$. Many of the following examples will use $Q(s,a)$ implementations.

It should be noted that equation 2.4 requires knowledge of the transition function $\mathbf{T}_{s,s'}$. The transition function or world-model can be implicitly estimated through sampling (experience) or explicitly modeled (and fitted to experience). Although the Bellman equation proves that the agent will uncover the optimal policy (and value function) under standard infinite time and sampling conditions, the implementation of transition models fundamentally alters the trajectory for uncovering it. We will return to model-based reinforcement learning after briefly covering model-free reinforcement learning.

2.2 Model-free reinforcement learning

The Bellman equation (equation 2.2) can be used to derive a model-free[2] algorithm for estimating the value of a given policy. This algorithm is called $TD(0)$ (Sutton and Barto, 1998). By making the assumption that the quantity r_t from equation 2.2 represents a sample from $E\left[R_{s_t}\right]$ and that $\hat{V}(s_{t+1})$ represents a sample from $E\left[V(s_{t+1})\right]$, it follows $\hat{V}^{\pi}(s_t)$ should be equivalent to the quantity $\left[r_t + \hat{V}(s_{t+1})\right]$ if the estimated value of the policy is correct. The difference or prediction error δ between $\hat{V}(s_t)$ and $r_t + \hat{V}(s_{t+1})$ is defined as

$$\delta_t = r_t + \hat{V}(s_{t+1}) - \hat{V}(s_t). \tag{2.8}$$

If the prediction error δ_t is non-zero, updating the current estimated value function $\hat{V}(s_t)$ by δ allows estimated value function to approach the optimal value function.[3] This leads to learning rules reminiscent of those proposed by Rescorla and Wagner (1972) for animal conditioning and by Widrow and Hoff (1960) for adaptive linear systems.

$$\hat{V}(s_t) \leftarrow \hat{V}(s_t) + \alpha\delta_t \tag{2.11}$$

---

[2] It should be noted that the term *model-free* is somewhat disingenuous. As shown in the later sections of this chapter, the definition of the state-space over which the value function is computed also represents *model* information. The usage within this chapter reflects current usage for the use of transition model information within reinforcement learning algorithms.

[3] Most implementations of this algorithm include a discounting term to avoid the insolubility of the value calculations with an infinite horizon. The discounting term γ is used to modify future reward expectancy according to

$$V^{\pi}(s_t) = \sum_{\tau \geq t} \gamma^{\tau} r_{\tau} \tag{2.9}$$

The consequence of this formulation of the value function is that it weights rewards within near future more greatly than those in the distant future. This redefinition of the value function requires a change in the prediction error (equation 2.8). That is

$$\delta_t = r_t + \gamma\hat{V}(s_{t+1}) - \hat{V}(s_t) \tag{2.10}$$

where γ is the discounting term.

Here α is a learning rate ($0 < \alpha \leq 1$) and is usually set to a number less than one in order to produce greater long term stability. The difference between Rescorla and Wagner's theory and temporal difference learning is the temporal dynamics that transfers value backward in time to the earliest state or stimulus that predicts reward (Sutton and Barto, 1981). As Daw (2003) notes, this simple rule provides a rough explanation for many behavioral observations including acquisition (Pavlov, 1927), cessation of responding (Pavlov, 1927), overshadowing (Pavlov, 1927), blocking (Kamin, 1969), and conditioned inhibition (Rescorla, 1969). Although many other behaviors cannot be predicted by this algorithm, the distinct correspondence between this basic algorithm and behavior suggests that the brain performs a similar set of computations and that signatures of these computations should be evident within neural signals.

The activity of dopaminergic neurons in the midbrain roughly corresponds to the prediction error (δ) within reinforcement learning algorithms (Montague et al., 1995, 1996; Schultz et al., 1997; Waelti et al., 2001; Bayer and Glimcher, 2005). Single unit activity recorded in the macaque midbrain (substantia nigra pars compacta and the ventral tegmental area) on simple appetitive learning tasks suggest that neurons within these areas compute a prediction error signal (Schultz et al., 1997; Schultz, 1998; Roesch et al., 2007). Further support for the hypothesis that these dopaminergic populations code for prediction-error learning signals comes from the broad anatomical projections these neurons send to cortex and striatum (Schultz, 1998; Wickens et al., 2003; Calabresi et al., 2007). While various algorithm implementation and neural instantiation related complications remain (Pan et al., 2005; Roesch et al., 2007), much of the activity of dopamine neurons within these simple tasks can be explained by temporal difference reinforcement learning models (but see Gurney et al., 2004).

How the brain implements other parts of a reinforcement learning algorithm that either compute or use the prediction error signal remains an open question. Neural correlates of the signals that presumably generate the prediction error – the value function and specific reward receipt – have yet to be uncovered. The likelihood of identifying the neural correlates of these signals is heavily dependent on how evaluation computations are performed and whether other (specifically model-based reinforcement learning) systems also contribute to or utilize prediction error signals (c.f. Houk et al., 1995; Dayan and Balleine, 2002; Daw et al., 2005; Niv et al., 2006a). Given the broad projections of dopamine neurons, it seems likely that the prediction error signals are utilized by multiple systems (Redish et al., 2007). How prediction error information carried by the activity of dopamine neurons might contribute more widely to reinforcement learning is considered in the sections below.

In summary, the model-free approach to reinforcement learning provides a simple, elegant and important introduction to reinforcement learning. Temporal difference reinforcement learning provides a solution for the Bellman equation without explicit definition and use of a transition model. Temporal difference learning computations require only the representation of the current state (or *state-action* pair) and its value, the value of the predicted state, and reward. These representations constrain decision-making to the present; the agent or modeled animal cannot *think ahead*, remember previous experiences or represent potential outcomes within a model-free framework. While such constraints are clearly at odds with the phenomenology of memory and decision-making, they also impose severe limitations on behavior within uncertain and non-stationary environments with non-constant or probabilistic patterns of reward.

2.3 Model-based reinforcement learning

Reinforcement learning models that operate within uncertain and non-stationary environments or in probabilistic or non-stationary reward tasks appear to require additional knowledge of environmental dynamics. Tolman and Brunswik (1935) called acquisition of this knowledge *learning the causal texture of the environment.* Within machine learning approaches to reinforcement learning, this knowledge is included within a model of task dependent state transitions called a *transition model.* The inclusion of transition models within machine learning approaches to reward-based behavior have directly contributed to dynamics- and uncertainty-related issues in navigation and reward acquisition within probabilistic environments (Thrun et al., 2001, 2005). The inclusion of transition models within reinforcement learning models of animal behavior produce behavioral dynamics that are much more consistent with observed animal behaviors than those predicted by standard model-free approaches to reinforcement learning (Daw et al., 2005; Johnson and Redish, 2005a; Courville et al., 2006; Niv et al., 2006b; Zilli and Hasselmo, 2008).

2.3.1 Multiple memory systems

The inclusion of transition models within reinforcement learning algorithms suggests several important questions. The first, and perhaps, most obvious question is what to do with the model-free learning rules. Though efficient, model-free reinforcement learning rules are generally inflexible algorithms and require many sample experiences for substantial changes in behavior. In contrast, model-based reinforcement learning algorithms are quite flexible but typically require significant computational expense. The expense of these algorithms is apparent within stable environments for well practiced tasks where these algorithms are most inefficient. While many machine learning approaches have sought to identify a single optimal algorithm, neurobiological approaches to learning has emphasized the use of multiple learning and memory systems (O'Keefe and Nadel, 1978; Squire, 1987; Cohen and Eichenbaum, 1993; Nadel, 1994; Redish, 1999). Consonant with neurobiological approaches to learning and memory, recent theoretical considerations of the algorithmic basis for animal behavior have included both model-free and model-based reinforcement learning (Daw et al., 2005).

The central question in developing a reinforcement learning theory on the simultaneous function of multiple memory systems is to identify what learning and memory system underlies behavior and how systemic control of behavior is accomplished. As described in the previous chapter, animals display flexible behavior that becomes more rigid across repeated experience within both navigation and instrumental tasks. These observations suggest that behavior is mediated by a flexible transition-model-based learning system during early learning phases and, with increasing experience, behavioral control is transferred to an inflexible model-free learning system.

2.3.2 Arbitration of behavioral control – Daw et al. (2005)

The critical contribution of the model by Daw et al. (2005) is treatment of multiple memory systems within a reinforcement learning framework and specifically analysis of the mechanisms for arbitrating behavioral control between these systems. Daw et al. (2005) propose two memory systems based on a Bayesian Q-learning algorithm initially proposed by Dearden et al. (1998). The first memory system is a simple model-free or cache-based system in which the value function is updated according to the immediate experience of the agent. The second memory system is a model-based *tree* system in which the value function is updated through both immediate experience and inferred potential experience based on a transition model. Because Bayesian Q-learning maintains a distribution of values over each state rather than a single scalar value, it provides access to the expected value of a given state *and* the uncertainty of the value for that state. In contrast to standard decision systems that use only expected value information (Sutton and Barto, 1998), systems that maintain value distributions allow decisions to be made according to both the expected value of a future state and the uncertainty of its value (Dearden et al., 1998). Daw et al. (2005) show that within models of simple instrumental tasks, the expected value and value uncertainty estimates differ between the cache and tree learning algorithms. By selecting the system that minimizes the value uncertainty information, they reproduce two basic behavioral results within instrumental learning.

Model

As in the previous sections, the transition model $\mathbf{T}_{s,a,s'} = p(s_{t+1} = s'|s_t = s, a_t = a)$ specified the probability that state $s' \in \mathcal{S}$ would result from action $a \in \mathcal{A}$ in state $s \in \mathcal{S}$. The reward function $\mathbf{R}_{s,s',r} = p(r = 1|s_t = s, s_{t+1} = s')$ defined the probability of reward receipt. The rewarded state was also a terminal state in all modeled tasks. The state-action value function $Q(s, a)$ is the expected probability of reward receipt given action a in state s.

$$Q(s,a) = \begin{cases} R(s) & s \text{ is terminal } (a = \emptyset) \\ \sum_{s'} T(s,a,s') \cdot \max_{a'} \left[Q(s',a')\right] & \text{otherwise} \end{cases} \tag{2.12}$$

While typical reinforcement learning approaches only track a single scalar value for $Q(s, a)$, Daw et al. (2005) used a Bayesian variation that estimated a posterior distribution $\mathbf{Q}_{s,a}(q) = p\left(Q(s,a) = q|\text{data}\right)$ that indicated the optimal probability of future reward q given the experiential evidence 'data' about transitions and outcomes.

Value distributions for the cache system $\mathbf{Q}^{cache}_{s,a}$ were found by bootstrapping. In contrast, value distributions for the tree system $\mathbf{Q}^{tree}_{s,a}$ were found by policy iteration. It should be noted that policy iteration was accomplished through Bayesian methods based on a current estimate of the transition model. While the transition model was estimated based on previous experience (although it was unclear how $\mathbf{T}_{s,a,s'}$ was estimated), it should be noted that the list of states were previously defined. The system with the least variance in the $\mathbf{Q}_{s,a}$ distribution was then used to set the expected value $Q_{s,a} = \langle \mathbf{Q}_{s,a} \rangle$ and actions were selected according to the Boltzmann distribution.

Dynamics and behavior

The model by Daw et al. (2005) was explicitly developed to model devaluation within instrumental paradigms (Killcross and Coutureau, 2003; Holland, 2004). More specifically, the model was used to explore two related theoretical issues: general behavioral patterns associated with instrumental responding and reward approach behaviors. Basic observations of animal behavior suggest that sensitivity to devaluation for actions that are distant from reward receipt like lever-pressing are dependent on the extent of previous experience (Killcross and Coutureau, 2003). In contrast, actions that are in close proximity to reward receipt like magazine entry behavior remain sensitive to devaluation even after extensive training (Killcross and Coutureau, 2003). Interestingly, animals given multiple actions and outcomes remain sensitive to devaluation even after extensive training (Holland, 2004). The multiple memory system model proposed by Daw et al. (2005) reproduces each of these behavioral observations.

Conclusion and commentary

The model forwarded by Daw et al. (2005) represents a simple and elegant proposal for the interaction and control of multiple memory systems within a general reinforcement learning framework. Its formulation of a model-free cache system is consistent with previous hypotheses on the correspondence of dopamine activity and the prediction error (Montague et al., 1996; Schultz et al., 1997). Although the model-based tree system is only loosely specified, both in computational terms and associated neural substrates, this reflects a general lack of specificity present within current psychological and neurobiological approaches to higher cognitive function.

Based on experimental findings that show inactivations or lesions of the infralimbic prefrontal cortex following overtraining reinstates sensitivity to devaluation protocols (Coutureau and Killcross, 2003) while lesions of the prelimbic medial prefrontal cortex impairs sensitivity to devaluation protocols (Killcross and Coutureau, 2003), Daw et al. (2005) hypothesized that behavioral control between the cache (model-free) and tree (model-based) systems is arbitrated by the medial prefrontal cortex. Daw et al. (2005) further proposed that the cache system is situated within the dorsolateral striatum while the tree system is situated within the prefrontal cortex. Although the cache-system/dorsolateral striatum proposal is consistent with the relative simplicity of cache system computations and other treatments of model-free reinforcement learning (Houk et al., 1995), the tree system/prefrontal cortex proposal appears less tenable. It is more likely that the greater computational complexity of the the tree system requires interaction of multiple brain areas beyond the prefrontal cortex *and* that the prefrontal cortex forms an integral part of this larger network. This larger network perspective suggests that considerations of the dynamic interaction of the prefrontal cortices and other brain areas may provide critical insights into the organization and computations of the tree system.

A particularly interesting question opened by the use of tree systems within reinforcement learning is how the transition model contributes to construction of a given policy (value function). The tree system proposed by Daw et al. (2005) depends on policy evaluation using policy iteration.[4] Although policy evaluation can be theoretically accomplished

---

[4] Solving for the optimal value function using policy iteration requires integration of value information over the full state space (and over all potential values if some variant of Bayesian Q-learning is used). These

simultaneously for all states at each time step (Sutton and Barto, 1998), more practical approaches to policy evaluation have used value updating based on hypothetical state trajectories derived from the transition model in addition to experience-based state trajectories (e.g. Dyna-Q algorithms Sutton, 1990; Sutton and Barto, 1998). The construction of hypothetical state trajectories from previous experience closely parallels recent conceptualizations of fictive experience based on episodic memory (Suddendorf and Corballis, 1997; Buckner and Carroll, 2007, see discussion of episodic memory as construction from chapter 1). Furthermore, several recent studies suggest that imagery and fictive experience are supported by an interaction between the prefrontal and temporal lobes (Addis et al., 2007; Schacter and Addis, 2007a; Schacter et al., 2007; Buckner and Carroll, 2007) and suggest that tree system dynamics may be mediated by similar network interactions.

In summary, the model by Daw et al. (2005) suggests that dynamical state representations form a critical component of model-based reinforcement learning. In contrast to value function construction in the cache system, dynamical state representations in the tree system allow fast propagation of value information over a state-space. The transfer of control from the tree system to a cache system provides a simple mechanism for modeling animals' changing sensitivity to devaluation protocols. Although little is known about these dynamic state representations, they can be described as memory dynamics and appear to correspond to recent treatments of episodic memory. These considerations suggest that further understanding of tree-like systems within reinforcement learning may come from including memory dynamics within reinforcement learning models.

2.3.3 Memory dynamics in reinforcement learning

A variety of memory processes have been explored within human and animal psychology literatures. Working memory, procedural memory, semantic and episodic memory, recognition, and recall each represent distinct memory processes with partially overlapping mechanisms (O'Keefe and Nadel, 1978; Squire, 1987; Cohen and Eichenbaum, 1993; Baddeley, 1997; Redish, 1999, present several well known taxonomies). While model-free reinforcement learning has been hypothesized to correspond to procedural memory, there is, as yet, no hypothesized correspondence between model-based reinforcement learning and other forms of

---

computations quickly become intractable as the number of states and potential transitions become large. Standard approaches to tree-based reinforcement learning over large state-spaces typically use pruning to minimize the portion of the tree to be searched/updated or redefine state information in order to reduce the the size of the state-space (Sutton and Barto, 1998). Given a reduced state-space (menu), near instantaneous policy evaluation becomes much more feasible. Neural coding consistent with this policy evaluation in a reduced state-space has been found in orbitofrontal cortex (Padoa-Schioppa and Assad, 2006).

(declarative?) memory. A basic question resulting from this dichotomy is whether current perspectives on memory, its constraints and its dynamics can contribute to understanding model-based reinforcement learning.

Recent approaches to reinforcement learning have begun to examine the contribution of memory-related processes and dynamics to models of animal behavior (Johnson and Redish, 2005a; Zilli and Hasselmo, 2008). These proposed models have been developed to better understand different memory processes, how they potentially interact, and how they contribute to observed behavior. Within the reinforcement learning framework, memory processes must be operationally and algorithmically defined. And though current definitions are underdeveloped, they provide a starting point for further understanding the computational substrates of different forms of memory. Finally, it should be noted that these reinforcement learning models markedly contrast to normative decision making models that utilize infinite capacity memory (Daw et al., 2005). Instead, the intent of these models is to understand the behavioral constraints that result from use of a particular theory of memory or memory system.

2.3.4 Analysis of working and episodic memory – Zilli and Hasselmo (2008)

State information is usually defined in terms of the modeled animal's current sensory experience. A cue within an instrumental task or a position within a maze is often used to signify state information. A recent model by Zilli and Hasselmo (2008) suggested that standard state information can be supplemented by state information derived from working and episodic memory. The result of modifying the definition of state in this way is that decision processes are not entirely dependent on current environmentally-based state information. Instead, current environmentally-based state information can be combined with memory-based state information to inform decision processes. Zilli and Hasselmo (2008) provide operational definitions of working memory and episodic memory such that these mnemonic processes can be defined as state information and show how this redefinition of state contributes to behavior.

Model

The model proposed by Zilli and Hasselmo (2008) used a standard actor-critic reinforcement learning algorithm with eligibility traces (Sutton and Barto, 1998). The critical difference between their model and previous models is its definition of state. Typical approaches to modeling animal behavior in navigation tasks have defined state information based on environmental position or location information (e.g. Foster et al., 2000; Johnson and Redish, 2005a). Zilli and Hasselmo (2008) defined state as a composite of location information S_L, working memory information S_{WM} and episodic memory information S_{EP}. The composite state-space $S = S_L \times S_{WM} \times S_{EP}$ situates animal behavior in terms of both external, experimentally observable motor actions and internal, experimentally covert mnemonic actions.

Motor actions produced a change in task state while mnemonic actions produced only changes in modeled working or episodic memory state. *Motor actions* were defined to produce a change in location state S_L. Working memory was defined a simple supplementation of state location information S_L and *working memory actions* were defined as either writing the current location state into working memory ($S_{WM} \leftarrow S_L$) or clearing working memory state information ($S_{WM} \leftarrow \emptyset$). Because the model maintained a temporally ordered list of the n previously visited location states $[S_L(t-1), \ldots, S_L(t-n)]$, episodic memory state S_{EP} was defined to contain either a single state from this list or be left empty.[5] *Episodic memory actions* allowed the modeled animal to retrieve an element from the episodic memory list that matched the current state location information S_L, advance retrieved episodic memory state information according to the temporally ordered list, or clear episodic memory state information.

Dynamics and behavior

Within the context of this basic actor-critic architecture Zilli and Hasselmo (2008) compared the performance of standard algorithms using only state location information and algorithms that included working memory or both working memory and episodic memory across a variety of behavioral tasks. These tasks include continuous and delayed spatial sequence disambiguation, continuous and delayed spatial alternation (Wood et al., 2000; Lee et al., 2006; Ainge et al., 2007b), continuous and delayed non-match to position (Griffin et al., 2007), continuous and delayed non-match to sample (Hampson and Deadwyler, 1996), continuous and delayed odor sequence disambiguation (Agster et al., 2002), and continuous and delayed tone-cued spatial alternation (Jog et al., 1999; Johnson and Redish, 2007).

The standard algorithm performs at chance levels for all of modeled tasks. The addition of working memory information to basic location state information allowed reinforcement learning algorithms to achieve high levels of performance on all non-delayed tasks. And

---

[5] Zilli and Hasselmo (2008) define episodic memory as the only content addressable memory within their set of memory systems. Although episodic memory almost certainly requires content addressable memory, this is likely a wider characteristic of state.

only the combined working and episodic memory algorithm[6] was able to solve delayed tasks (with the exception of delayed odor sequence disambiguation).

Conclusion and commentary

The model proposed by Zilli and Hasselmo (2008) underscores the importance of an appropriate description of state. Zilli and Hasselmo (2008) definitively show that a basic sensory-based (location) state-space is insufficient for modeling animal behavior on a variety of navigation and sequence tasks and that a composite (location×memory) state-space can be used to reproduce a basic set of behavioral results from these tasks. Given multiple memory systems (and associated state-spaces), this observation implies that the brain somehow selects or constructs a task appropriate state-space or memory system over which reinforcement learning computations operate. An intriguing consequence of Zilli and Hasselmo's treatment is that the (location) and (location×working memory) state-spaces are subspaces of the composite (location×working memory×episodic memory) state-space. Coupled with the cost/reward contingency structure for each task,[7] the state-space structure suggests that the set of state trajectories defined by an optimal policy are increasingly biased against working memory and episodic memory use (see figure 2.1). Implicit within this formulation is a mechanism for arbitration of control between multiple memory systems and selection of an effective state-space.

The composite state-space $S_L \times S_{WM} \times S_{EP}$ suggests that state is not encapsulated by a single brain area and instead results from a combination of signals across multiple brain areas. Working memory is dependent on the prefrontal cortex (Braver et al., 2001; Wang et al., 2004); spatial working memory is dependent on the hippocampus (Floresco et al., 1997; Wilkerson and Levin, 1999; Lee and Kesner, 2002); and episodic memory is dependent on the hippocampus (Tulving, 2002; Squire et al., 2004). Furthermore, several behavioral and theoretical accounts of prefrontal cortical function have implicated it in construction of an appropriate state-space (Bouton, 2002; Milad and Quirk, 2002; Lebron et al., 2004;

---

[6] No episodic memory only algorithm was used in this paper.

[7] A subtle and important aspect of the implementation used for the model by Zilli and Hasselmo (2008) is the structure of reward contingencies used within each modeled task. For most of the tasks modeled in their paper, the agent received a reward of $+9.5$ for a correct response, -6.0 for an incorrect response, -1.0 for impossible actions (such as selecting an action that moved the agent outside the task boundaries), -1.0 for re-entering a location the agent had just visited, and -0.05 for all other actions. The importance of this final penalty should not be underestimated. The authors suggest that this penalty encourages the agent to find the shortest path that leads to task solution (which is a combination of motor and mnemonic actions) rather than selecting only mnemonic actions. The bias toward motor action rather than cognitive or memory actions is a signature of motivation and stress (see Niv et al., 2006b for a cogent treatment of motivation in reinforcement learning).

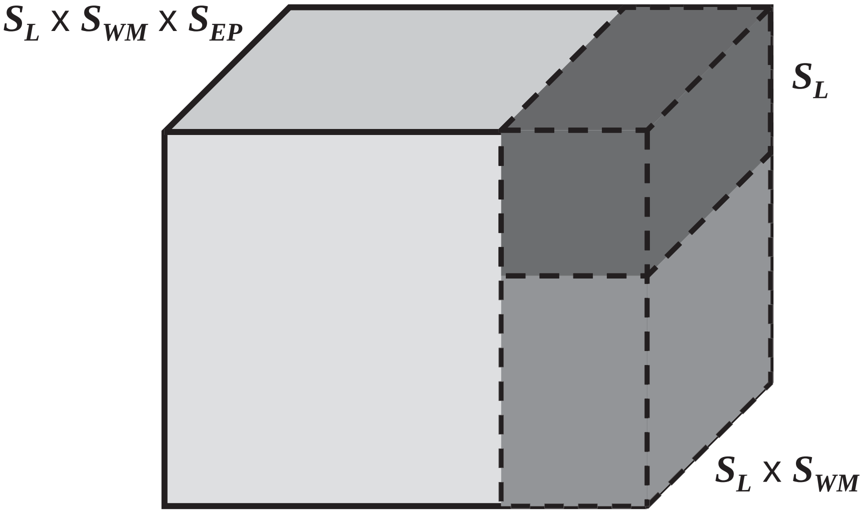

Figure 2.1: A sketch of the state-space defined by Zilli and Hasselmo (2008) for location, working memory and episodic memory. Zilli and Hasselmo's definition identifies $S_L \in [S_L \times S_{WM}] \in [S_L \times S_{WM} \times S_{EP}]$. Because only actions within the S_L space are rewarded, the set of state trajectories defined by an optimal policy are increasingly biased against working memory and episodic memory use.

Milad et al., 2004; Santini et al., 2004; Bouton et al., 2006; Quirk et al., 2006; Redish et al., 2007) and state-space or memory system selection (Coutureau and Killcross, 2003; Killcross and Coutureau, 2003; Rhodes and Killcross, 2004; Daw et al., 2005; Haddon and Killcross, 2005).

Although working and episodic memory processes are controlled by a simple reinforcement learning algorithm within Zilli and Hasslemo's model, the timing and extent of their use is task dependent. For example, working memory actions occur only at the initial ambiguous state within the spatial sequence disambiguation tasks. Such non-uniform or differential use of memory may provide an explanation for directional place fields found within the hippocampus (Frank et al., 2000; Wood et al., 2000).[8] The contribution of the episodic memory algorithm to decision-making within Zilli and Hasselmo's model is consistent with episodic memory as mental time travel for past events (Tulving, 2002) but is not consistent with episodic memory as episodic future thinking (Atance and O'Neill, 2001) or

---

[8] One particularly interesting point is that the episodic memory system within Zilli and Hasselmo's model is that it must be primed according to the modeled animal's current location. This requirement leads to microscopic exploration of the state-space for use of episodic memory. Such microscopic choice behavior and the associated memory dynamics closely parallel vicarious trial and error (Muenzinger, 1938; Tolman, 1939; Brown, 1992).

foresight (Suddendorf and Corballis, 2007) because these require construction of potential future events.

In summary, Zilli and Hasselmo's proposed interaction between multiple memory systems provides a deep insight into the organization and use of memory within reinforcement learning tasks. Memory function is not a ubiquitous phenomenon, but is selectively distributed according to task demands. This contrasts with the model proposed by Daw et al. (2005) that tree-based memory function occurs non-selectively within a task and across task performance; instead, Zilli and Hasselmo (2008) suggest that learning within a task entails identification of specific points within a task that represent important sites of mnemonic processing. While these aspects of Zilli and Hasselmo's theory remain unexplored, they provide an intriguing area of future experimental and theoretical work.

2.4 Discussion

The basic reinforcement learning formalisms outlined in this chapter form a substantial part of current reinforcement learning theory as it applies to modeling animal behavior its neural substrates. The previous discussion examined reinforcement learning in terms of model-free and model-based algorithms. This distinction is a natural theoretical result following the derivation of the Bellman equation (2.4). Although the distinction between model-free and model-based reinforcement learning provides a useful system for classification of reinforcement learning models (Sutton and Barto, 1998; Daw et al., 2005), some reinforcement learning models are poorly fit by these categories (Foster et al., 2000; Johnson and Redish, 2005a; Zilli and Hasselmo, 2008, note that these models are not normative). For example, Zilli and Hasselmo's model used a basic actor-critic model-free TD(λ) system but also included a variant of a transition model for episodic memory system. The correspondence between transition model use and different forms of (declarative?) memory represents an intriguing open question that is partially treated in subsequent chapters.

The discussion of learning presented in this chapter emphasizes model-based reinforcement learning. In contrast to model-free reinforcement learning theories that describe learning in terms of specific cellular mechanisms (Houk et al., 1995), little is known about the cellular substrates of model-based reinforcement learning or even the neuronal dynamics that would support transition model use. As a consequence, model-based reinforcement learning has been primarily discussed in terms of high level memory processes supported by entire brain areas rather that specific neural circuits. This suggests that characterization of the specific neural substrates that support transition model use represents an important avenue for understanding model-based reinforcement learning. The next chapter outlines a simple model that defines the transition model and its dynamics in terms of cellular processes within the hippocampus (Johnson and Redish, 2005a).

Chapter 3

State dynamics and learning

3.1 Introduction

Standard temporal difference learning algorithms learn very slowly over many episodes (requiring thousands or tens of thousands of episodes, Sutton and Barto, 1998). Various methods have been developed to accelerate learning. Some of these methods use *practice* signals which allow the agent to use previously recorded experience or a model to simulate experience following the episode and update value estimates and policy. In the Dyna-Q family of algorithms (Sutton, 1990; Sutton and Barto, 1998), a state and action are selected from stored previous experience and used to update the value function between episodes. Dyna-Q thus requires a replay of the agent's recent experience during periods of rest between episodes.

These technical considerations within reinforcement learning are well matched to sequential activity found within the rodent hippocampus. Replay of previous experience in the hippocampus during periods of rest and sleep (Pavlides and Winson, 1989; Wilson and McNaughton, 1994; Skaggs and McNaughton, 1996; Nádasdy et al., 1999; Louie and Wilson, 2001; Lee and Wilson, 2002) may provide a mechanism implementing a practice component of reinforcement learning algorithms. Early experimental observations of hippocampal pyramidal cell activity showed that cells activated during a task remained active during sleep following the task performance (Pavlides and Winson, 1989) and that the activity of place cells with overlapping place fields was highly correlated during sleep in comparison to cells without overlapping place fields (Wilson and McNaughton, 1994). Later studies indicated that place cell activity is temporally ordered; during sleep episodes place cells fire in the order they were encountered by the rat during task performance (Skaggs et al., 1996; Nádasdy et al., 1999; Louie and Wilson, 2001; Lee and Wilson, 2002). This is called *route replay* (See Redish, 1999, for review). Single unit activity in sleep or rest periods shows increased spike probability associated with sharp wave ripple events (O'Keefe and Nadel, 1978; Buzsáki, 1989; Kudrimoti et al., 1999). Further evidence showed that during slow wave sleep, replay occurs during these sharp wave ripple events (Kudrimoti et al., 1999; Lee and Wilson, 2002). Because place cell activity shows increased temporal structure following task performance, hippocampal activity during replay has been suggested to be associated with memory consolidation processes (Marr, 1971; McNaughton, 1983; Buzsáki, 1989; Pavlides and Winson, 1989; Wilson and McNaughton, 1994; Skaggs et al., 1996; Redish and Touretzky, 1998a; Redish, 1999; Louie and Wilson, 2001; Lee and Wilson, 2002).

Hippocampal replay has been examined on simple runway tasks (Wilson and McNaughton, 1994; Skaggs et al., 1996; Nádasdy et al., 1999; Louie and Wilson, 2001; Lee and Wilson, 2002; Foster and Wilson, 2006) but not choice-based tasks. To gauge the potential effect of replay on behavior in an explicit choice-task, we examined a model of reinforcement learning a multiple-T task (described by Schmitzer-Torbert and Redish, 2002, 2004; Johnson and Redish, 2007; Figure 3.1A). In order to receive the reward, the animal had to successfully navigate a sequence of T choices. Although the task used by Schmitzer-Torbert and Redish (2002, 2004) formed a loop and animals were not removed from the Multiple-T maze between laps, they tended to run the maze episodically, pausing for a long time (mean 27 seconds) at the second feeder before running another lap quickly (mean 16 seconds) (Masimore et al., 2005; Schmitzer-Torbert, 2005). Rats running the multiple-T task showed two differentiable learning rates: a fast decrease in the number of errors (the number of incorrect choices per lap) and a slow increase in the regularity of the path on each lap (path stereotypy) (Schmitzer-Torbert and Redish, 2002).

3.2 Model

3.2.1 Reinforcement learning.

The reinforcement learning component of the model was based on a standard SARSA Q-learning algorithm with a continuous state-space and discrete actions (Sutton and Barto, 1998). Spatial state information was approximated by a set of n_s randomly distributed radial basis functions. A state-action value $Q(s, a)$ function was found by associating a set of *place* neurons (2-D radial basis functions) distributed over the spatial state-space and n_a actions to with a scalar value, $Q(s_i, a_j)$. Action directions were distributed evenly over the interval $[0, 2\pi)$. State-action value indicated the "quality" of taking action a_j in state s_i. The value of any state-action in the continuous state-space was thus given by the sum of values weighted by their similarity to the current state s:

$$Q(s, a_j) = \sum_{i=1}^{n_s} Q(s_i, a_j) G(s - s_i) \tag{3.1}$$

where s_i is the *static* position of the ith neuron in state-space and $G(s - s_i)$ is a Gaussian kernel with standard deviation σ. Note that G is effectively a *place code* with place field centers located at s_i.

Following standard discrete time temporal difference approaches, the prediction error signal δ was found by

$$\delta = r + \gamma Q(s(t+1), a(t+1)) - Q(s(t), a(t)). \tag{3.2}$$

and was used to update $Q(s_i, a_j)$ by

$$Q(s_i, a_j) \leftarrow Q(s_i, a_j) + \eta \delta G(s(t) - s_i) \lfloor \langle a(t).a_j \rangle \rfloor_+ \tag{3.3}$$

where η is a learning rate parameter and $\lfloor \langle a.a_j \rangle \rfloor_+$ is the positive component of the dot product of the last action a and each possible action a_j. Consequently, $G(s - s_i)$ indicates the activity of any place cell centered at s_i and its contribution to the representation of actual position. The quantity $\lfloor \langle a.a_j \rangle \rfloor_+$ provides an indication of the contribution of action a_j to $Q(s_i, a_j)$ based on the difference between a_j and the actual action.

Actions were selected according to the softmax rule,

$$P(a(t) = a_j) = \frac{e^{\beta Q(s(t), a_j)}}{\sum_j e^{\beta Q(s(t), a_j)}} \tag{3.4}$$

where β is a temperature parameter and $Q(s, a_j)$ is the value of each action a_j at the current position s.

3.2.2 The practice signal.

Replay was implemented probabilistically. Random cellular activity was propagated via a transition matrix $W_{transition}$ that developed with experience. The transition matrix was initialized to 0 to begin simulation of each new session (i.e. each day) and maintained across laps (i.e. within each session). Within the maze, the transition matrix was updated at each timestep according to

$$W_{transition} \leftarrow W_{transition} + \zeta \arctan\left(\bar{P}^T(t) \left[\bar{P}(t) - \bar{P}(t-1) \right] \right) \tag{3.5}$$

Parameter		
n_a	8	number of actions
n_s	430	number of place cells
σ	2.5	place field width
β	1.0	softmax temperature
stepsize	0.8	movement length
γ	0.99	discounting factor
η	0.6	value function learning rate
n_r	8	number of replays attempted per lap
$r_{threshold}$	$n_s/50$	replay activity threshold
ζ	1.0	transition matrix learning rate
tracksize	$(38, 38)$	
trackwidth	2	

Table 3.1: Parameters used in the model by (Johnson and Redish, 2005a).

where $\bar{P}(t)$ is the vector of current place cell activity ($\bar{P}(t) = G(s(t) - s_i)\forall$ cells $i, 1 \leq i \leq n_s$), and ζ is a learning rate. This is similar to a discretized time version of the weight matrix used by Blum and Abbott (1996) or an approximation of the learning that occurs through the combination of spike-time-dependent-plasticity and phase-precession (Redish and Touretzky, 1998a).

n_r replays were attempted at the completion of each lap. Each replay attempt began by activating one randomly selected place cell. The activity of that selected cell was set to 1. This initial random activity propagated to other cells following learned transition matrix weights. Subsequent place cell activity was found found by application of winner-take-all rule where the cell with the largest activity produced by the outputs of the previous active cells was selected as the subsequently active cell. If the activity of the next active cell was greater than threshold, $r_{threshold}$, it was counted as part of the replay, its activity was set to 1, and its activity was allowed to propagate throughout the network. This process was repeated until either the activity did not reach threshold or the replay reached the end of the maze. At each step of each replay, the state-action value function $Q(s, a)$ was updated according to the equations given above except that the action was estimated as the one most similar to the change in state.

The values for the parameters used in the model are given in Table 3.1.

3.2.3 Analysis

Twenty-four sessions were simulated with replay and without replay. Each session consisted of eighty laps. If the agent did not complete the maze (i.e. “failure” or “reward” in Figure 3.1A) within 2000 steps, the agent was removed from the lap and no reward was given, that is, the lap ended in “failure”. The minimum number of steps an agent would require to get to reward depended on the specific configuration of the maze, but was approximately 200 steps.

Errors were defined as entry into an incorrect arm at each T choice point. Because repeated (or extended) entries into an incorrect arm were only counted once, a maximum of four errors could occur on a maze with four T choices.

Path stereotypy was defined as the correlation of spatial path through the maze for each lap following the methods outlined by Schmitzer-Torbert and Redish (2002). Briefly, this included calculation of a path correlation matrix that summarized the correlation between the pixelated path for each lap with each other lap. The linearized path correlation at lap i, identified as *path stereotypy* hereafter, was defined as the mean correlation of lap i with all other laps.

Replay was quantified by counting the length of each cascade as number of place cells activated following each lap. This number was summed over each of the n_r replays for each lap.

3.3 Results

The basic results of the model are summarized in Figures 3.2 and 3.3. The simulation provided four primary results: the temporal difference model learned the correct path, errors decreased with experience, path stereotypy increased with experience and route replay developed with experience.

Simulated rats successfully learned to navigate the multiple-T maze. Simulations with and without replay learned the multiple-T and showed a decrease the number of steps required to complete each lap and an increase the average reward per lap over a training session (figure 3.2). While the model with replay shows a slightly faster initial decrease in the number of steps required for each lap, the model without replay reaches a slightly lower steady state (beyond 50 laps). The model with replay acquired more reward on average than the model without replay.

Similar to animal behavior, the number of errors decreased early in the simulated training session. However, it should be noted that the decrease in errors in the simulation was slower than what is observed in animal behavior (compare Schmitzer-Torbert and Redish, 2002). The path stereotypy increased over the simulated training session. In combination with the decrease in errors, this indicates a trend toward a more efficient path through the maze. However, the increase in path stereotypy was still slower than what was observed in animal behavior (compare Schmitzer-Torbert and Redish, 2002, 2004).

Simulations showed slow development of route replay (Figure 3.4). Because replay develops probabilistically in this model, increasing the number of replays at the completion of each lap, n_r, will magnify differences between the models.

3.4 Discussion

3.4.1 Assumptions

Route replay occurs during awake behavior.

The model requires replay during task performance. While replay was situated in the pauses taken between laps in this model, replay can occur during any rest interval. Extensive evidence has shown that route replay occurs during sleep states in between experimental sessions (Skaggs and McNaughton, 1996; Nádasdy et al., 1999; Louie and Wilson, 2001; Lee and Wilson, 2002). Following early observations of sharp wave ripple activity during non-sleep periods of rest and immobility (O'Keefe and Nadel, 1978; Buzsáki et al., 1983), several recent experiments show evidence for replay during awake behavior (O'Neill et al., 2006; Jackson et al., 2006; Diba and Buzsàki, 2007; O'Neill et al., 2008).

Route replay is dependent on asymmetric long term potentiation

The development of route replay (practice signals) was predicated on asymmetric long term potentiation. Without this assumption, standard symmetric long term potentiation-based replay leads to directionally non-specific activation. Extensive evidence for asymmetric plasticity has been found within the hippocampus (Levy and Steward, 1983; Bi and Poo, 2001). Furthermore, the spiking characteristics of phase precession (O'Keefe and Recce, 1993; Skaggs et al., 1996) match the temporal specificity required for spike-time dependent (asymmetric) long term potentiation and have been used to model place field expansion (Blum and Abbott, 1996; Redish and Touretzky, 1998a).

3.4.2 Predictions

Route replay develops within a session.

A basic prediction of the model is experience-dependent development of route replay (see Figure 3.4). Moreover, the model predicts that route replay-containing elements of particularly well sampled trajectories will develop faster than poorly sampled trajectories. Early observations of route replay showed increased replay activity following task performance relative to pre-task levels of replay and suggested experience-dependent development of route replay (Wilson and McNaughton, 1994; Skaggs and McNaughton, 1996; Kudrimoti et al., 1999; Lee and Wilson, 2002). More recent experimental observations suggest that replay activity develops within a task (Jackson et al., 2006; O'Neill et al., 2006).

Current theories suggest that sharp waves are indicative of activity cascades derived originally from recurrent connections in CA3 (Shen and McNaughton, 1996; Csicsvari et al., 2000; Ylinen et al., 1995) and that the replay component follows asymmetries that arise through learning (Buzsáki, 1989; Skaggs and McNaughton, 1996; Redish and Touretzky, 1998a; Nádasdy et al., 1999; Redish, 1999). Consistent with the predictions of the model and current theories of sharp wave generation, Jackson et al. (2006) showed that the frequency of sharp ripple emission increased as a function of experience in the multipleT task and that increased sharp-wave emission in CA3 leads CA1 increases (see Figure 3.5).

Another prediction of this model is a consequences of the synaptic plasticity mechanisms hypothesized to support development of route replay. The model suggests that route replay develops through the interaction of asymmetric plasticity (Levy and Steward, 1983; Bi and Poo, 2001) and phase precession (O'Keefe and Recce, 1993; Skaggs et al., 1996). Several theories (Blum and Abbott, 1996; Redish and Touretzky, 1998a) have proposed these synaptic plasticity mechanisms as the basis of a backward expansion of place fields (Mehta et al., 1997, 2000; Lee et al., 2004a). These hypotheses predict that sharp wave ripple activity should develop with a time-course similar to that of place-field expansion. Furthermore, place field expansion is dependent on NMDA integrity (Shen et al., 1997; Ekstrom et al., 2001). If experience dependent development of route replay and place field expansion share the same plasticity mechanisms, blocking NMDA receptor function in CA3 should reduce development of route replay during sharp waves.

Hippocampal inactivation should increase errors but should facilitate the development of path stereotypy.

Hippocampal replay provides an indirect practice signal which enables faster learning. As such, it accelerates learning and reduces errors (see Figure 3.3A). However, because the replay overlearns actual experience, it slows the development of optimized fine-motor performance paths (see Figure 3.3B). These predictions can be tested by removing the hippocampal contribution through direct inactivation studies. Hippocampal inactivation should increase errors during learning (i.e. slow the correction of errors), but should speed up the development of path stereotypy.

3.4.3 Open questions

What behavior is replayed?

The model presented here neglects the potential influences of cognitive or top-down processes to structure replay and instead uses experimentally observable cellular mechanisms to structure replay. This leaves open a critical question: if relatively simple cellular mechanisms mediate replay, at what rate does synaptic modification occur? The rate of synaptic change dictates the content of route replay and the practice signal; very fast synaptic modifications will result in the replay of recent episodes, while slow synaptic modifications will result in a more general averaged replay. Theories of hippocampal function suggest that synaptic modifications within the hippocampus, particularly within the CA3 recurrent collaterals, are very fast (Marr, 1971; Levy, 1996). Alternatively, slow modifications may result in a split in replay at T choice points, unless the network acts to maintain a consistent representation (as in an attractor network, Samsonovich and McNaughton, 1997; Redish and Touretzky, 1997; Tsodyks, 1999; Doboli et al., 2000). Experimental examination of route replay have typically employed highly practiced behaviors (Wilson and McNaughton, 1994; Skaggs and McNaughton, 1996; Kudrimoti et al., 1999; Louie and Wilson, 2001; Lee and Wilson, 2002). Evidence from Jackson et al. (2006) suggests that coherent reactivation develops with experience but did not distinguish whether replay reflect recent experience or averaged experience.

Clearly, the content of the replayed practice signal directly influences the development of the value function and behavioral plasticity. The synaptic modification rates used in the current model were quite slow, providing a generalized average replay and relatively small changes in the value function. Faster synaptic modification rates lead to greater temporal specificity in replay and larger changes in the value function. In sum, the degree of plasticity within the value function directly affects the level of behavioral flexibility of the modeled animal.

Is replay prioritized?

Current models of replay (including this one) have generally hypothesized the initiation of replay as random (e.g. Redish and Touretzky, 1998a), or related to the ratio of experience of the agent (e.g. Shen and McNaughton, 1996). However, various reinforcement learning models have shown that a prioritized-sweep provides for a significant improvement in learning speed (Sutton and Barto, 1998; Zilli and Hasselmo, 2008). Experimental studies of hippocampal replay have only focused on simple, hippocampal-independent, runway tasks (i.e. not requiring choices). It is not yet known whether the replayed patterns are selected at random or whether they are prioritized in some way.

The reinforcement learning literature suggests that, given a large state-space, replay should be prioritized with a preference for state-transitions in which there was a large change in the value function, that is experiences in which the value-prediction error signal δ was large (Sutton and Barto, 1998). If replay were prioritized in this way, replay should begin nearest to the source of reward and follow the greatest change in the δ signal – this might provide some explanation for backward route replay (Foster and Wilson, 2006; Diba and Buzsàki, 2007). Given the hypothesis that phasic dopamine carries the δ signal (Montague et al., 1996; Schultz et al., 1997; Schultz, 1998), this would also predict that dopamine should have an effect on prioritizing states for replay. Past research has shown that dopamine enhances early long term potentiation in CA1 (Otmakhova and Lisman, 1996, 1999) and that dopamine agonists enhance the stability of hippocampal pyramidal cell place fields while dopamine antagonists destabilize them (Kentros et al., 2004). While it remains unclear whether phasic dopamine, corresponding to a δ signal, is the basis for these observed changes place cell stability, one interesting possibility is that this dopamine signal is responsible for place field modulation by behaviorally relevant learning signals such as those seen by Moita et al. (2003).

3.5 Conclusion

Model-based reinforcement learning algorithms display much faster learning than model-free learning algorithms. Practice, or off-line replay of recently experienced states, has been used to accelerate temporal-difference learning algorithms (Sutton and Barto, 1998).

The purpose of the model presented in this chapter was to link a simple model-based reinforcement algorithm with known neuronal dynamics and synaptic plasticity.

Flexible spatial learning is dependent on the hippocampus. The observation of route replay parallels proposed mechanisms for accelerating reinforcement learning algorithms. Although passive development of replay facilitated performance in a model of behavior on a multiple-T maze, behavioral performance within simulations continued lag behind observed animal behavior. Development of prioritized replay (Sutton and Barto, 1998) or active search (Daw et al., 2005) could be used to further facilitate model performance. While there is currently little evidence for prioritized replay or search during sharp wave ripple activity, a number of representation dynamics observed within hippocampal theta regimes might support this function (O'Keefe and Recce, 1993; Jensen and Lisman, 2005; Johnson and Redish, 2007). These observations suggest that active memory mechanisms mediated by the hippocampus may play an important role in reinforcement learning.

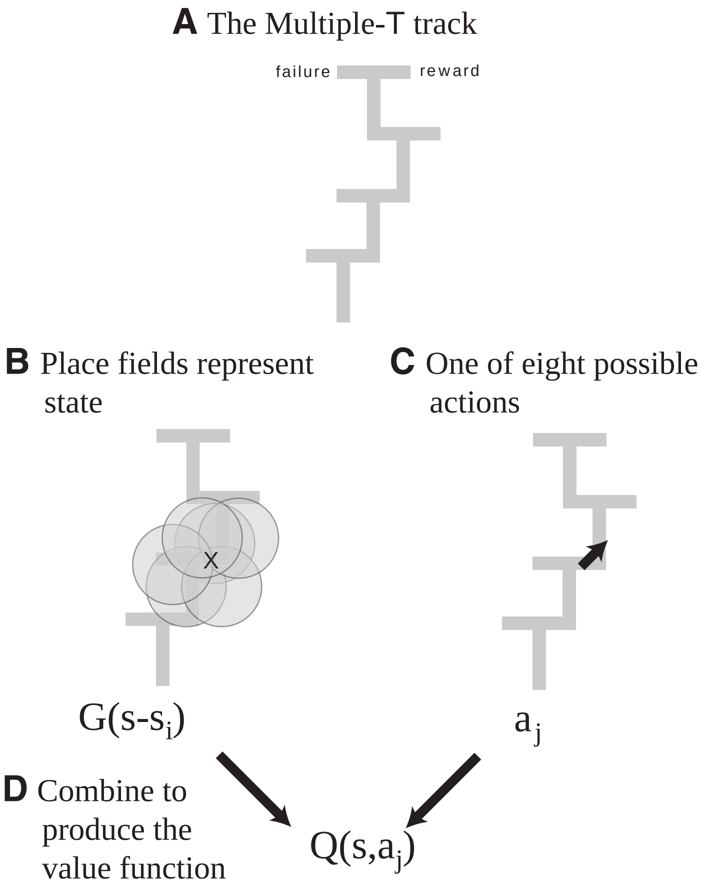

Figure 3.1: Summary of the multiple-T task and model. The task consisted of a sequence of four T choices with reward available on one arm of the final T. The model used radial basis functions to compute the state-action value $Q(s, a)$ over a continuous state, discrete action space. Each action was selected using *softmax*. **A**: One example track as used in the model. **B**: Place fields represent state through a distributed encoding. x indicates the position of the animal, active units are shown. **D**: State (**B**) and one of the eight possible actions (**C**) were each associated with a value function $Q(s, a_j)$. Figure used with permission (Johnson and Redish, 2005a).

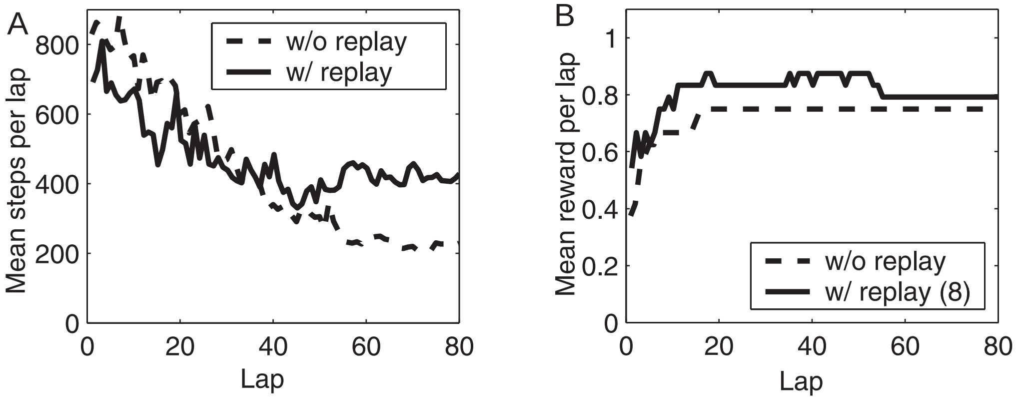

Figure 3.2: Comparison of learning in TDRL models with and without developing replay-based practice. **A**: Both models show a decrease in the number of steps to complete each lap. The number of steps required to achieve the optimal path depended on the configuration of the maze, but was approximately 200 steps. **B**: The model with replay typically acquires more reward than the standard model. For agents that completed a lap, chance would be 50%, however, because agents were removed after 2000 steps, chance is actually much lower than 50%. Figure used with permission (Johnson and Redish, 2005a).

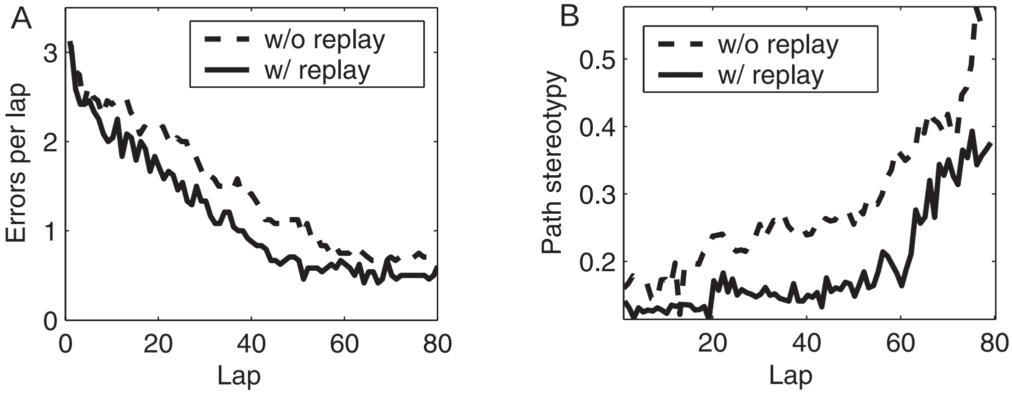

Figure 3.3: Comparison of TDRL models with and without developing replay-based practice over twenty-four sessions. **A**: The model with replay shows a faster decrease in the number of errors per lap than the standard model. **B**: The model with replay shows a slower onset of path stereotypy than the standard model. Figure used with permission (Johnson and Redish, 2005a).

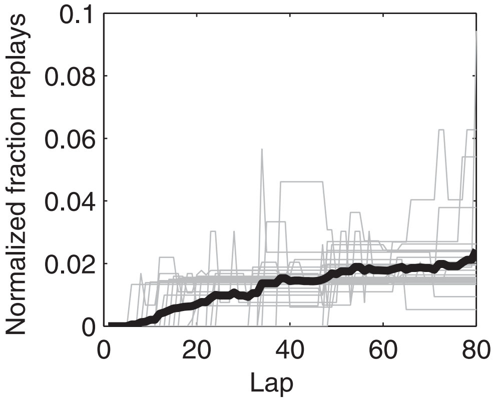

Figure 3.4: Replay developed over the course of the session. Replay strength was measured as the length of the replayed sequence. Plotted are the proportion of replay strength occurring at each lap (length of replay on each lap divided by total length of replay over entire session). On early laps, the replay tended to fail early, thus producing short replays, while on later laps, replay tended to proceed robustly, thus producing longer replays. Figure used with permission (Johnson and Redish, 2005a).

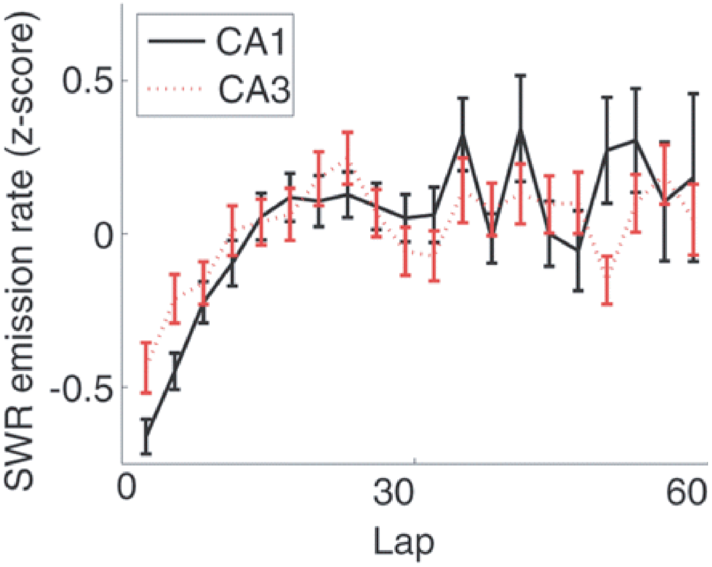

Figure 3.5: Sharp wave emission develops across experience in a multipleT task. Sharp waves were detected using standard methods (see Jackson et al., 2006 for details). Note that increased sharp wave emission in CA3 leads increased sharp wave emission in CA1. Figure used with permission (Jackson et al., 2006).

Chapter 4

Statistical methods

4.1 Neural representations

As an animal interacts with the world, it encounters various problems for which it must find a solution. The description of the world and the problems encountered within it play a fundamental role in how an animal behaves and finds a solution. Sensory and memory processes within the brain provide a description of the world and within that description the brain's decision making processes must select some course of action or behavior.

How is information about the world represented and organized in the brain? The use of information from the world in behavior involves two critical processes. The first process is appropriate transformation of information about the world into a representation that is relevant and useful for behavior. The second process is the projection of that representation onto a behavior that allows the animal to select appropriate actions for interacting with its world. The following discussion uses the term *neural representation* to describe any representation of the world within the brain or any transformation of that representation toward behavior (even if the behavior is not executed). This definition is intentionally broad such that the operations underlying directly observable behavior *and* covert mental activities can be considered.

Neural representations are distributed across populations of neurons. Although traditional approaches to investigating neural representations have typically focused on single unit activity, more recent approaches have examined the simultaneous activity of multiple units. Current recording technology allows simultaneous recording of large neural ensembles of more than 100 cells from awake behaving animals. Neural ensemble recordings provide a distinct advantage over single unit recordings because they make coordinated dynamics across multiple cells accessible. More specifically, patterns of neural activity that deviate from hypothesized encoding models (experimentally observable or directly inferrable variables) might be the result of neural noise or indicate internal cognitive processes. Differentiating between noise and internal cognitive processes has been difficult within single unit recording studies and largely constrained these studies to investigation of the correspondence between neural activity and a directly observable stimulus or pattern of motor behavior. In contrast, differentiating between neural noise and internal cognitive processes is much more feasible within neural ensemble recording studies. This technological advance coupled with theoretical developments in the analysis of neural ensembles has allowed new insights into the organization and dynamics of neural representations.

The following discussion develops a strategy for decoding dynamical representations in multiple information processing modes. This chapter is divided into two sections. The first section outlines a basic approach to neural decoding strategies based on neural correlates of observed behavior and extends previous methods for use neural representations reflecting covert information processing with no *a priori* behavioral correlate at fast timescales (ms, tens of ms). The second section outlines a method for examining the comparative validity of decoded representations and how these can be used to identify multiple different information processing modes.

4.2 Analysis of neural representations

4.2.1 Encoding and tuning curves

What makes a neuron fire? The question can be asked with respect to the neuron's immediate environment — its afferents and ion channels — and with respect to the world beyond. Answering the latter question requires knowing what information is *encoded* by the neuron. An *encoding model* describes an experimental hypothesis relating the information represented by a single neuron (sensory, perceptual, motivational, motor, etc.) to its observed activity. The hypothesized relationship between the encoded information x and the neural activity, typically considered in terms of spikes, s can be written as the function

$$p(s_t) = T(x_t) \tag{4.1}$$

where $p(s_t)$ probability of a spike at time s_t.[1] This definition is easily extended to include both preceding experience and planned future behaviors in the encoded information x. For

---

[1] Of course, the actual activity is also a function of the history of spiking of the neuron (e.g. neurons show refractory periods, bursting, and other history-dependent processes).

simplicity, the present discussion neglects identification of the precise temporal offset in describing the relationship between s_t and x_t.

These encoding models have classically been found by standard *tuning curves*. More recently, these encoding models have been stated in terms of Shannon information theory, identifying the mutual information between behavioral variables and spike timing (Rieke et al., 1997; Dayan and Abbott, 2001). Other encoding definitions have been based on *linear filter kernels*, which reflect the recent history of variable x in the temporal organization of spiking (Bialek et al., 1991) or bursting (Kepecs and Lisman, 2003) activity.[2] These encoding models can be measured relative to any available behavioral variable, whether it be immediate sensory input, such as the frequency of an auditory tone, an immediate motor output, such as the target of a saccade or the direction of a reach, or a cognitive variable, such as the location of an animal in the environment.

4.2.2 Population decoding

Because the variability of a single cell is usually insufficient to fully describe the entire space of encoded information, information is generally encoded across a population of neurons that differ in the parameterization of their tuning curves (often described by a *family of tuning curves* such as retinotopic Gaussians or place fields). If information is consistently represented across a population of neurons, then it should be possible to infer the expectation of the variable x by examining the neural activity across the population s. This inference can be made using Bayes' rule

$$p(x, s) = p(x|s)p(s) = p(s|x)p(x) \tag{4.2}$$

where $p(s|x)$ is the probability of observing some set of neural activities given the variable of interest and $p(x|s)$ is the probability of the variable of interest given some set of neural activity. This means that the variable x can be decoded from the neural activity across the population s by

$$p(x|s) = \frac{p(s|x)\,p(x)}{p(s)} \tag{4.3}$$

The probability $p(x|s)$ describes how information can be read out or *decoded* from the network. What should be clear from this simple account is that decoding critically depends on the *encoding model*, $p(s|x)$.

---

[2] Kernel based methods explain the observed neural activity in terms of both the represented information and the neuronal dynamics of the cell itself. The generative method below effectively extends this perspective to include unobserved variables that can only be determined by examining ensembles with decoding.

The term s in equation 4.3 reflects the pattern of neural activity across the entire population of cells and typically refers to population activity at a given time t. This analysis thus requires sufficient data to infer the probability density function across an n-dimensional space (where n is the number of cells in the ensemble). Because of combinatorial expansion, appropriate sampling of s often requires an inordinate amount of data for even moderately-sized ensembles. In many situations, it is convenient to assume that the activity of each cell is conditionally independent, relative to the represented variable x (Zhang et al., 1998; Brown et al., 1998), so that

$$p(x|s) = \prod_{i \in \text{cells}} \frac{p(s_i|x)\, p(x)}{p(s_i)} \tag{4.4}$$

This makes appropriate sampling much more viable. However, the validity of this assumption is still controversial (Nirenberg and Latham, 2003; Schneidman et al., 2003; Averbeck et al., 2006).

Although Bayes' rule (Eq. 4.3) provides an optimal solution for decoding, even the simplified version (Eq. 4.4) is often not computationally tractable. As a result, several other non-probability based methods have been developed for decoding (e.g. template matching, Wilson and McNaughton, 1993; Averbeck et al., 2003a; Zhang et al., 1998, linearly weighted averaging, Georgopoulos et al., 1983; Salinas and Abbott, 1994; Zhang et al., 1998). These decoding algorithms are often much more computationally efficient than fully Bayesian computations and can be considered as reduced forms of Bayes' rule (Dayan and Abbott, 2001).

4.2.3 Decoding at fast timescales

A critical open question on the application of neural decoding methods to spiking data is the temporal interval over which spiking data should be grouped. This question is closely related to the spiking versus firing rate code debate and has produced very different answers for similar data sets (e.g. hippocampal place cell data). For instance, Zhang et al. (1998) showed that the error distance between observed spatial trajectories and trajectories decoded from hippocampal place cells was minimized using long temporal intervals (on the order of one second) while Brown et al. (1998) showed similarly high accuracy neural decoding at much faster timescales (on the order of milliseconds) using predictive filter based methods. In contrast to standard Bayesian decoding methods that typically treat decoded neural representations at each time step as independent, predictive filter methods treat decoded neural representations as temporally dependent and constrain the temporal evolution of the posterior distribution.[3]

The term predictive filter refers to an explicit predictive model that is used to constrain the temporal evolution of neural representations. Predictive filter algorithms are characterized by the recursive application of a *prediction step* which predicts the temporal evolution of the neural representation given the previous prediction and the proposed dynamical model and a *correction step* which corrects the prediction based on the spikes observed at time t. The prediction can be written as

$$p(x_t|s_{t-1}) = \int p(x_t|x_{t-1})p(x_{t-1}|s_{t-1})\,dx_{t-1} \tag{4.5}$$

where $p(x_t|x_{t-1})$ describes the hypothesized representation dynamics, or what Brown et al. (1998) called the path model, and the term $p(x_{t-1}|s_{t-1})$ represents the predicted neural representation at the previous timestep. The correction step can be written as

$$p(x_t|s_t) = \frac{p(s_t|x_t)p(x_t|s_{t-1})}{p(s_t|s_{t-1})} \tag{4.6}$$

where $p(s_t|s_{t-1})$ is the probability of the neural activity set s_t given the previous set of neural activity s_{t-1}, $p(s_t|x_t)$ supplies current evidence, and the term $p(x_t|s_{t-1})$ represents the output of the predictive step (equation 4.5).

Predictive filter decoding strategies have been successfully used for decoding neural activity within a variety of brain areas at very fast timescales (e.g. Brown et al., 1998; Brockwell et al., 2004; Wu et al., 2006). In contrast to decoding methods applied at long timescales, predictive filter methods allow (require) specification of temporal processes that reflect single unit spiking dynamics such as the recent spiking history (Bialek et al., 1991) or bursting (Kepecs and Lisman, 2003) within the $p(s_t|s_{t-1})$ term and representation dynamics such as phase precession (O'Keefe and Recce, 1993) and route replay (Skaggs and McNaughton, 1996) within the $p(x_t|x_{t-1})$ term. Coupled with the fine timescale resolution of predictive filter decoding, the degree of explicit specification of temporal dynamics that underlie neural representations allowed by predictive filter methods suggests a distinct utility in investigation of memory and other dynamical cognitive processes.

4.2.4 Memory and cognition as non-local representations

Neural activity is typically analyzed with respect to an observable external variable x_t: $p(s_t) = T(x_t)$. However, a more accurate statement is that neural activity reflects an

---

[3] It should be noted that at sufficiently large temporal intervals, the contribution of a model of the temporal evolution neural representations becomes negligible.

internal representation of this variable. An important consequence of this subtle shift in perspective is that it suggests internal representations have the potential to coherently deviate from the external world. This point is particularly salient when investigating processes in which cognition potentially plays a role; one of the hallmarks of cognitive processing is the connection of the observable world with the animal's externally invisible goals or motivations (Tulving, 1983, 2001, 2002; Suddendorf and Busby, 2003; Gray, 2004; Ferbinteanu et al., 2006; Johnson and Redish, 2007). This perspective suggests that in memory tasks neural representations should align to observable stimulus or behavioral variables during encoding regimes and can potentially deviate from these variables during retrieval.

Multiple lines of evidence suggest that the hippocampus displays dynamics indicative of this distinction between encoding and retrieval (O'Keefe and Nadel, 1978; Buzsáki, 1989; Hasselmo and Bower, 1993; Hasselmo, 1993). As rats perform active behavioral tasks on an environment during normal navigation, the first order information within hippocampal pyramidal cells is the location of the animal (O'Keefe and Nadel, 1978; Redish, 1999). However, many recent experimental results related to phase precession and route replay have shown higher order fine temporal structure within hippocampal place cell activity that is not predicted by the animal's spatial position. Exactly how hippocampal place cell dynamics map onto the above distinction between encoding and retrieval remains unclear.

The *phase-precession* of place cell spiking activity during each theta cycle suggests a dynamic in which the spatial representation sweeps through positions recently occupied by the animal to positions that will likely be occupied by the animal (O'Keefe and Recce, 1993; Skaggs et al., 1996, see Fig. 4.1). Although place cell spiking is strongly modulated by the theta rhythm, this phase relationship is not constant and can be described more specifically than preferred phase (O'Keefe and Recce, 1993; Skaggs et al., 1996). O'Keefe and Recce (1993) showed that place cells tend to spike at later theta phases as an animal enters a cell's place field and precess toward earlier theta phases as the animal progresses through the field. The neural representation dynamics suggested by phase precession are time-compressed relative to animal behavior during task performance and have been calculated to move at speeds approximately 10-15 times faster than observed animal behavior (Skaggs et al., 1996). Because there is no clear observable behavioral variable onto which phase precession can be mapped, multiple groups have included preferred phase within hippocampal neural decoding in order to minimize the difference between predicted and observed spatial positions (Brown et al., 1998; Jensen and Lisman, 2000). However, several proposals by Hasselmo and colleagues (Hasselmo and Bower, 1993; Hasselmo, 1993; Koene et al., 2003; Koene and Hasselmo, 2008; Zilli and Hasselmo, 2008) suggest that differences between predicted and observed positions are not *errors*, but indicative of mnemonic processing.

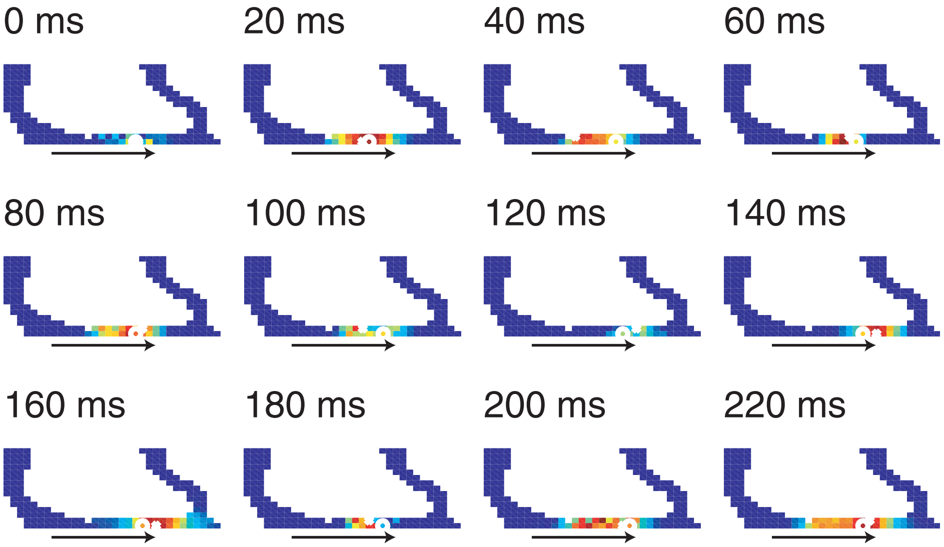

Figure 4.1: Phase precession during spatial navigation. As the rat runs from left to right the reconstructed location sweeps from behind the animal to in front of it. The reconstruction probability is indicated by color (red high probability, blue low probability). Panels arranged from left to right, top to bottom in 20 msec intervals.

The term *route replay* has been used to describe a dynamic that occurs during sharp wave ripple events. Sequences of neuronal activity observed during task performance are replayed within slow wave sleep following task performance (Kudrimoti et al., 1999; Nádasdy et al., 1999; Lee and Wilson, 2002). Sequential spiking activity in sharp-wave replay during slow wave sleep is time-compressed 40 times relative to animal behavior during the task (Nádasdy et al., 1999; Lee and Wilson, 2002). Similar sequential reactivation has been observed during awake states; however, unlike sleep, the ordering of spiking activity during periods of awake immobility can be forward, backward or mixed (Jensen and Lisman, 2000; Foster and Wilson, 2006; Jackson et al., 2006; O'Neill et al., 2006; Diba and Buzsàki, 2007, see Fig. 4.2). The interpretation of spiking activity during sharp wave ripple events as route replay, particularly during slow wave sleep, has received wide acceptance as memory process because there is little other explanation for these sequential activations and it is consistent with previous conceptualizations of online and offline *information processing modes* within the hippocampus as indicated by hippocampal local field potential activity (Vanderwolf, 1971; O'Keefe and Nadel, 1978; Buzsáki, 1989).

Application of neural decoding algorithms to both online and offline hippocampal spik-

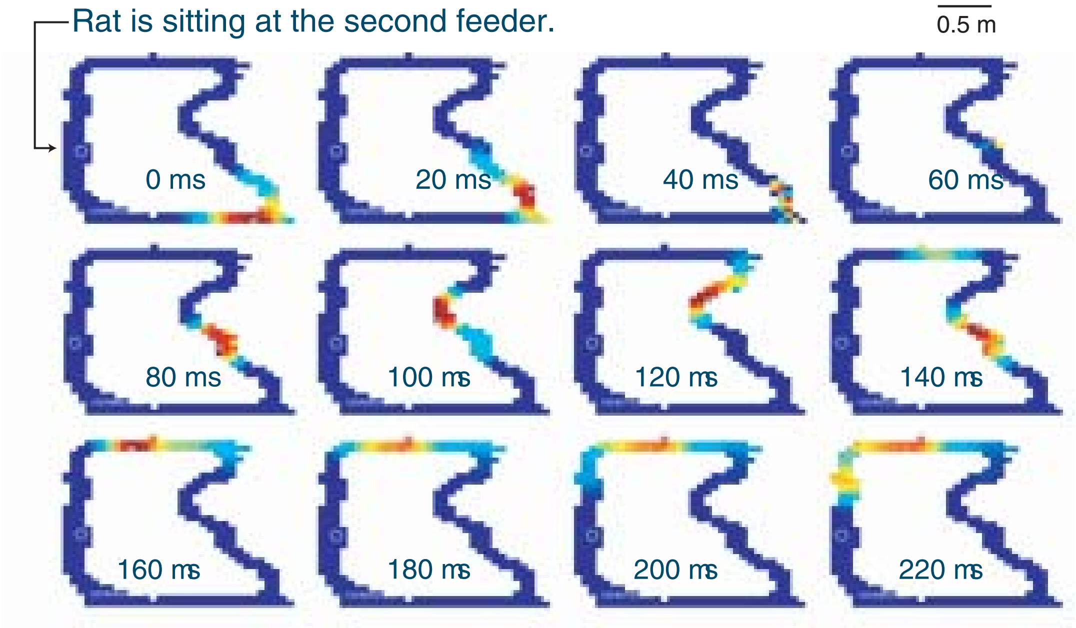

Figure 4.2: Replay of experience on the maze during an awake-sharp wave. The rat is sitting at the second feeder throughout the event. The distribution starts at the base of the first T and moves through the full maze in 220 msec (typical behavioral run times through this maze = 10–12 seconds). The reconstruction probability is indicated by color (red high probability, blue low probability). Panels arranged from left to right, top to bottom in 20 msec intervals. Note the coherent, but non-local reconstruction of the representation during the sharp wave.

ing data predict positions that deviate from the animal's observed positions. A variety of perspectives suggest that such differences should be considered non-local representations rather than simple decoding errors and that these spatial representation dynamics are, at least in part, indicative of multiple hippocampal information processing modes. Beyond differences in spatial representation dynamics, *information processing modes* within the hippocampus have been characterized by distinct local field potential activity, single unit activity patterns (in both pyramidal projection neurons and interneurons), and specific neuromodulation (Vanderwolf, 1971; O'Keefe and Nadel, 1978; Hasselmo and Bower, 1993; Somogyi and Klausberger, 2005). These modes are thought to be differentially involved in learning, storage, and recall (Buzsáki, 1989; Hasselmo and Bower, 1993; Redish, 1999). In sum, these observations provide strong support for the hypothesis that spatial representations in the hippocampus shift between primarily local representation dynamics and primarily non-local representation dynamics and suggest that identification of information processing models is critical for neural decoding.

4.3 Consistency analysis

The neural decoding methods have provided important insight into neural information processing. However, the methods outlined above run the risk of making an assumption that an animal's brain rigidly adheres to representing the present behavioral status of the animal and mistaking differences between predicted and observed behavioral variables as *errors*. A perspective that requires a match between observed and predicted status (e.g. spatial position) tacitly ignores the cognitive questions of memory and recall that are fundamental to brain function. However, departure from this traditional *prediction-matches-observation* perspective also removes the primary method used for testing the validity of the decoded representation.

How can we proceed without the prediction-matches-observation form of validity test? One potential answer comes from considering the differences between (1) precise, accurate recall, (2) precise, but inaccurate recall and (3) confusion. The decoded representations in the first two cases likely display a high degree of internal consistency; that is, for any given set of cells the similarity (or difference) in tuning functions predicts the similarity (or difference) in cellular activity at any moment in time. In contrast to cases (1) and (2) in which decoded neural representations display a high degree of internal consistency but differ in match between observed and predicted behavioral status, decoded neural representations in case (3) display low internal consistency. These considerations form the basis for the argument that consistency, rather than prediction-matches-observation, can be used as a powerful test of validity for examining neural representations as they relate to cognitive function.

4.3.1 Basic consistency analysis

A coherent or consistent representation is one in which the firing of a majority of neurons in a network conforms to an expected pattern. Ascertaining the consistency of a decoded neural representation can be accomplished in both probabilistic and non-probabilistic cases. In the probabilistic case, the maximum posterior probability (MAP) indicates the extent to which the reconstructed estimate is consistent across the population of neural activities. In the non-probabilistic case, the length of the population vector (Georgopoulos et al., 1983) can be used (Smyrnis et al., 1992, which is a measure of the variance of the circular distribution, Mardia, 1972, Batschelet, 1981). Georgopoulos et al. (1988) used this to measure development of a motor plan during mental rotation tasks. As the animal developed a motor plan, the length of the population vector increased.

A basic problem with each of these methods is how to appropriately define consistent representation. The upper bound for population vector length is unclear. And though the maximum posterior distribution is bounded, it is dependent on the specific characteristics of spiking activity and the shape of tuning curves within the population. A variety of comparative methods have been developed to circumvent this problem. For instance Averbeck and colleagues (Averbeck, 2001; Averbeck et al., 2002, 2003b) recorded from frontal and parietal neural ensembles in a shape-copying task and compared the neural activity patterns during the time monkeys were actually drawing the shapes with the neural activity patterns during the preparatory period. By describing population firing rates during each segment of the copying process as an n-dimensional tuple (e.g. $F_{\text{action 1}} = (f_1, f_2, \ldots, f_n)$), Averbeck et al. (2002) showed that the distance between ensembles firing rate patterns in the period leading up to the copy period and the ensemble firing rate patterns for each action component predicted component ordering for the to-be-copied shape. These basic observations suggest that the methods for assessing the consistency of neural representations form a integral part of understanding information processing in the brain.

4.3.2 Consistency and coherency

Place fields remap in response to environmental and task changes (Knierim et al., 1995; Markus et al., 1995). Given that place fields are sensitive to spatial reference frames, Redish et al. (2000) designed a task that set two spatial reference frames against each other. In this task the rat exited a box and ran down a linear track to an unmarked location where, if it paused, it would receive medial forebrain bundle stimulation. The unmarked location remained stationary in the room coordinate system while the linear track and the box were shifted relative to the room. In order to solve the task, rats were required to shift from use of a spatial reference frame aligned to the box to use of a spatial reference frame aligned to the room.

Redish et al. (2000) suggested that a mathematical comparison between observed and expected activity patterns could provide useful information about map switching dynamics. Redish et al. (2000) defined an activity packet as a weighted sum of the tuning curves

$$A(x,t) = \frac{\sum_k T_k(x) \cdot F_i(t)}{\sum_i T_i(x)} \tag{4.7}$$

where $T_i(x)$ is the tuning curve of cell i for a given variable x, and $F_i(t)$ is the firing rate of cell i at time t. Redish et al. (2000) also defined expected activity packets ($\hat{A}$) within box and room reference frames based on the expected firing activity given a relative position within each frame. This was accomplished by setting $F_i(t) = T_i(x_{\text{room}})$ or $F_i(t) = T_i(x_{\text{box}})$. *Coherency* was defined as the dot product of the actual and expected activity packets. A comparison of the box versus room coherency showed that the hippocampus transitioned from a spatial reference frame aligned to the rat's starting point to another reference frame aligned to the room.

In later treatments Jackson and Redish (2003, 2007) noted that a behaviorally observed variable (position relative to box or room coordinates) was not required for coherency calculations and that any hypothesized behavioral variable $\hat{x}$ could be used to define the expected activity packet $\hat{A}$. Based on this, Jackson and Redish (2003) defined the expected activity packet as the sum of the tuning curves weighted by the *expected* firing rate of the cells given a decoded position $\hat{x}(t)$

$$\hat{A}(x,t) = \frac{\sum_i T_i(x) \cdot E(F_i(t))}{\sum_i T_i(x)} \tag{4.8}$$

$$= \frac{\sum_i T_i(x) \cdot T_i(\hat{x}(t))}{\sum_i T_i(x)} \tag{4.9}$$

where $\hat{x}(t)$ is the hypothesized value of variable x at time t. Following construction of actual and expected activity packets, the consistency of the neural population relative to the hypothesized behavioral or decoded variable can be measured. Various consistency measures have been developed (Redish et al., 2000; Jackson and Redish, 2003; Johnson et al., 2005; Jackson and Redish, 2007 and see Jackson, 2006 for a comprehensive review).

Johnson et al. (2005) recorded neural ensembles of head direction cells from the post-subiculum of rats in a cylinder-foraging task and calculated the coherency of the head direction representation relative to the reconstructed head direction $\hat{\phi}$. Highly consistent patterns of activity were more likely to provide an accurate representation of the animal's current head direction than low consistency patterns of activity. Following the proposal of Jackson and Redish (2003), Johnson et al. (2005) suggested that if downstream structures used only self-consistent representations for making decisions, then the animal would be more likely to use accurate representations of the outside world.

The previous examples by Redish et al. (2000) and Johnson et al. (2005) suggest that consistency of an ensemble relative to hypothesized behavioral variable $\hat{x}$ remains consistent over long temporal intervals with some small switching probability. Statistically, this is equivalent to the probability of accepting the null hypothesis that the actual and expected activity packets are the same for some interval $t \in T$:

$$H_0 : \forall_x \forall_t A(x,t) = \hat{A}(x,t) \tag{4.10}$$

However, the validity of this hypothesis appears to vary over time and suggests that treating this hypothesis as as a function of time

$$H_0(t) : \forall_x A(x,t) = \hat{A}(x,t) \tag{4.11}$$

might provide greater insight into the dynamics of neural representations. Consonant with this hypothesis, Jackson and Redish (2007) showed that multiple hippocampal maps exist within the a single task environment and that these maps dynamically switch within task performance. Furthermore, Jackson and Redish (2007) showed that dynamic switching between multiple hippocampal maps better predicted place cell activity than a single averaged map.

4.3.3 A Bayesian approach to consistency

The previous non-probabilistic discussion developed consistency in terms of a comparison between actual and expected activity packets. The construction of both actual and expected activity packets rests on a number of assumptions. Some of these assumptions are merely incidental and can be proved negligible, but others remain and are difficult to treat in a statistically appropriate manner. For instance, coherency calculations using tuning curves project a distribution of firing rates for any given position onto a single number. This projection is usually accomplished in the form

$$F(x) = \int F \cdot p(F,x)\, dF \tag{4.12}$$

where $p(F,x)$ is the joint probability of firing rate and position. Similarly, coherency calculations use a single valued estimate for the decoded variable rather than a distribution. The marginalization of each of these distributions represents a massive loss of information. While such assumptions and associated information losses are required by many non-probabilistic decoding methods, they are not necessary within Bayesian formulations. Recent neural decoding strategies based on Bayesian and information measures (Rieke et al., 1997; Zhang et al., 1998; Brown et al., 1998; Zemel et al., 1998; Dayan and Abbott, 2001) suggest that question of consistency can be recast within a Bayesian framework.

In the non-probabilistic methods above, the activity packet (Eq. 4.7) measures the expected distribution of variable x given the firing rate at time t, $F(t)$. In Bayesian terms, the posterior distribution $p(x_t|s_t)$ provides an analogous term to the expected activity packet above (Eq. 4.9). Like the activity packet, the posterior distribution is a function over the variable x for each temporal interval. However, in contrast with coherency calculations which compare observed and expected neural activity within an activity packet space, Bayesian methods compare neural activity within a neural activity space. Examination of Bayes' rule in its simple (equations 4.3 and 4.4) or more complex predictive filter (equations 4.5 and 4.6) forms suggests how application of the hypothesized prior distributions for $p(x)$, $p(s)$, $p(s|x)$, $p(x_t|x_{t-1})$, and $p(s_t|s_{t-1})$ contribute to the posterior distribution. Together these terms specify a hypothesized model H_i within a generative model framework.

The generative model framework allows direct comparison of multiple models. In short, application of Bayes' rule to multiple models can be used to determine which generative model H_i is most probable given a hypothesized neural representation x and observed spiking data S.

$$p(H_i|x, S) = \frac{p(H_i|S)p(x|H_i, S)}{\sum_j p(H_j|S)p(x|H_j, S)} \tag{4.13}$$

Comparing the generative power of each hypothesized model can be found using standard methods (such as odds ratios and measured in decibans, Jaynes, 2003). Furthermore, generative models can be examined in either bulk trajectory form analogous to coherency (see equation 4.10) or in an instantaneous form as a function of time (see equation 4.11).

One particularly notable aspect of equation 4.13 is the term $p(x|H_i, S)$ which can be written as

$$p(x|H_i, S) = \frac{p(S|H_i, x)p(x|H_i)}{\sum_{x \in X} p(S|H_i, x)p(x|H_i)}. \tag{4.14}$$

Written in this form, a hypothesized model H_i that includes temporal dynamical constraints allows expansion of equation 4.14 as a predictive filter. Other hypothesized models allow probabilistic treatments of multiple hippocampal maps (Fuhs et al., 2005) that parallel non-Bayesian methods (Jackson and Redish, 2007). Comparison of multiple dynamical models applied to hippocampal data is explored below.

4.3.4 Multiple models in hippocampus

Spatial representations in the hippocampus have been explored using a variety of decoding methods (Wilson and McNaughton, 1994; Zhang et al., 1998; Brown et al., 1998; Jensen and Lisman, 2000). As described above, the neural activity of place cells and the decoded spatial representation generally predicts the animal's position within the environment; however, place cell activity can remain well-organized even when the decoded representation does not match the animal's position as in the case of phase precession (O'Keefe and Recce, 1993; Skaggs et al., 1996; Harris et al., 2002; Johnson and Redish, 2005b, see figure 4.1, above) and route replay (Skaggs et al., 1996; Lee and Wilson, 2002; Johnson and Redish, 2005b, 2007; Foster and Wilson, 2006, see figure 4.2, above). Phase precession and route replay differ in their spatiotemporal dynamics; phase precession is approximately $10-15\times$ compressed relative to animal behavior (Skaggs et al., 1996) whereas route replay is approximately $40\times$ compressed relative to animal behavior (Nádasdy et al., 1999; Lee and Wilson, 2002). Additionally, these dynamics correspond with distinct information processing states as indicated by characteristic local field potential activity within the hippocampus (Vanderwolf, 1971; O'Keefe and Nadel, 1978). Phase precession occurs during high theta power epochs (O'Keefe and Recce, 1993; Skaggs et al., 1996) while route replay occurs during high delta/high ripple power epochs (Kudrimoti et al., 1999; Nádasdy et al., 1999; Lee and Wilson, 2002; Jackson et al., 2006; Foster and Wilson, 2006; O'Neill et al., 2006; Diba and Buzsàki, 2007; O'Neill et al., 2008).

The predictive filter version of the generative model framework outlined above provides a method for examining spatial representations in the hippocampus with respect to multiple spatiotemporal dynamics. Explicit comparison of multiple models of hypothesized representation dynamics allows identification of the underlying dynamical state of the neural representation. A model of the dynamics of a neural representation can be most simply described as a Markov process $p(x_t|x_{t-1})$, that gives the transition probability for a representation at x_{t-1} to move to a new position x_t. The transition models used within predictive filters can be as simple as a Brownian walk or as complex as a rigidly specified directional flow.

Four generative models were used for neural decoding from a neural ensemble recorded from the CA1 and CA3 hippocampal subfields of an animal running on a 4T Multiple-T maze (Schmitzer-Torbert and Redish, 2002, 2004). These models were constructed by exploiting the Markov property which states

$$p(x_{t+n}|x_t) = p(x_{t+1}|x_t)^n \tag{4.15}$$

where n is a time compression factor. Each model specified the drift velocity of the probability distribution within the prediction step : $1\times \equiv p(x_t|x_{t-1})^1$, $15\times \equiv p(x_t|x_{t-1})^{15}$, $40\times \equiv p(x_t|x_{t-1})^{40}$, and $99\times \equiv p(x_t|x_{t-1})^{99}$, where $p(x_t|x_{t-1})$ was a Gaussian function with σ proportional to the average velocity of the rat. The $99\times$ model provided a nearly uniform distribution over the scale of the Multiple-T maze.

After decoding the neural representation for each of the proposed models, the generative

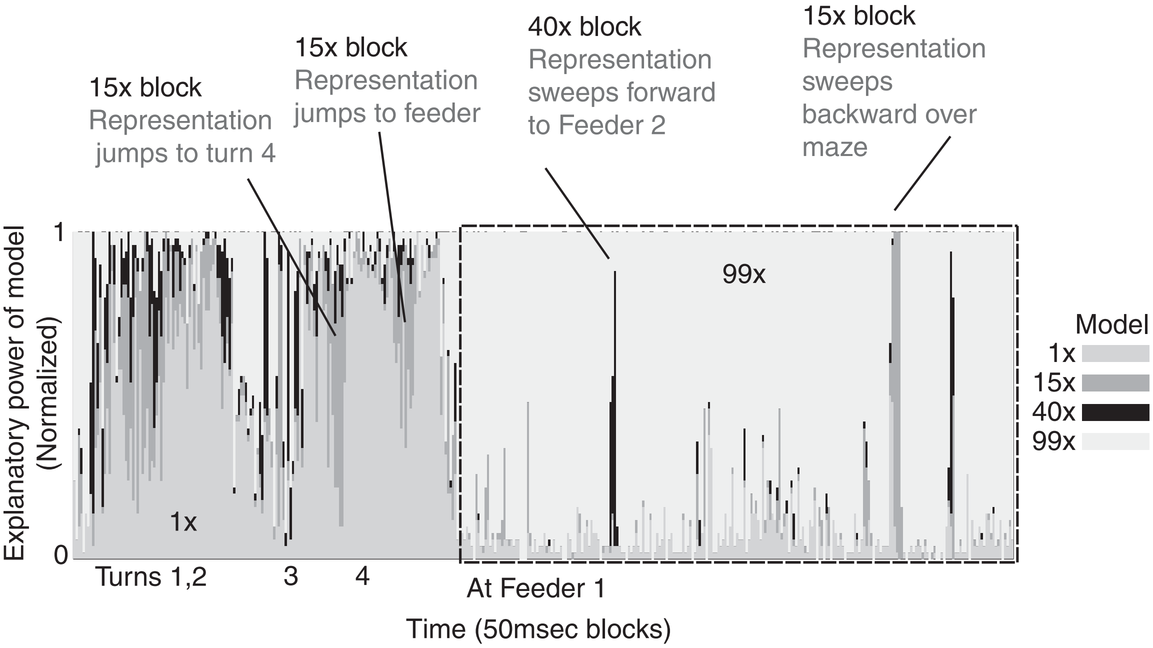

Figure 4.3: Multiple generative models in the hippocampus. Four generative models were examined 1×, 15×, 40×, and 99×. During the first portion (Turns 1–4), the animal was running through the maze. During the second portion, the animal paused at the first feeder to rest, groom, and eat. Colored columns display model selection distributions for 50ms blocks.

power for each of the hypothesized models can be determined by estimating how consistent the decoded representation is with the observed neural activity. Generative model selection was performed by taking the maximum posterior distribution. As shown in Figure 4.3, different spatiotemporal models were more consistent at different times, reflecting changes in the neural dynamics.

Each of the generative models was hypothesized to reflect different hippocampal information processing modes. The correspondence between these known information processing modes and their associated local field potential activity (O'Keefe and Nadel, 1978; Buzsáki, 1989, 2006; Hasselmo and Bower, 1993; Redish, 1999) suggests that the characteristic local field potential power spectrum for each hypothesized spatiotemporal model should show similar trends. Specifically, the 1× and 15× filters should show increased power within theta frequencies (7-10Hz) while the 40× filter should show increased power within slow wave delta (2-6Hz) and sharp wave ripple (170-200Hz) frequencies. The 99× model was used as a null (noise) hypothesis. Clear differences within the power spectra for the 1×,

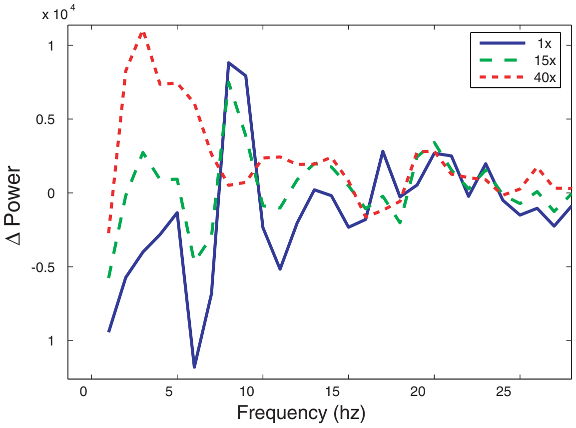

Figure 4.4: Average hippocampal local field potential power spectra for generative models in the hippocampus. Three generative models (1×, 15×, 40×) were compared to the noise model (99×).

15× and 40× models were found within slow wave delta and theta frequencies but no differences were found within sharp wave ripple frequencies (see Figure 4.4). Differences between the characteristic power spectra for each filter were similar between CA1 and CA3. Consistent with previous results (Lee et al., 2004b; Leutgeb et al., 2004), subfield analysis found that more distributed models (e.g. 99×) were more often selected in CA1 data sets, relative to the CA3 data sets (see Figure 4.5).

While generative models have been broadly used to explain and decode neural activity (e.g. Brown et al., 1998; Rao and Ballard, 1999; Lee and Mumford, 2003; Brockwell et al., 2004; Serruya et al., 2004; Wu et al., 2006), one notable distinction should be made between the typical generative model formulation and the present formulation. Because we are concerned with the dynamical regulation of neural representations by cognitive processes, particularly explicit memory retrieval, we suggest that multiple generative models are necessary to explain observed neural activity. A single model is generally not enough because cognition requires the interactive use of dynamical information based on sensory or motor processes *and* planning, motivation or, for lack of another word, cognitive processes. Within each of these types of representation, cognition modulates an ongoing process. This is precisely the type of modulation that is sought when examining learning and memory or any cognitive processes; mathematically it can be identified as a changes in the model prior $p(x)$. In terms of the generative model, this simply states that there exists a prior non-uniform distribution $p_1(x)$ which better describes the neural activity than a uniform distribution $p_2(x)$. The critical aspect of this formulation is that the goal is to completely generate the

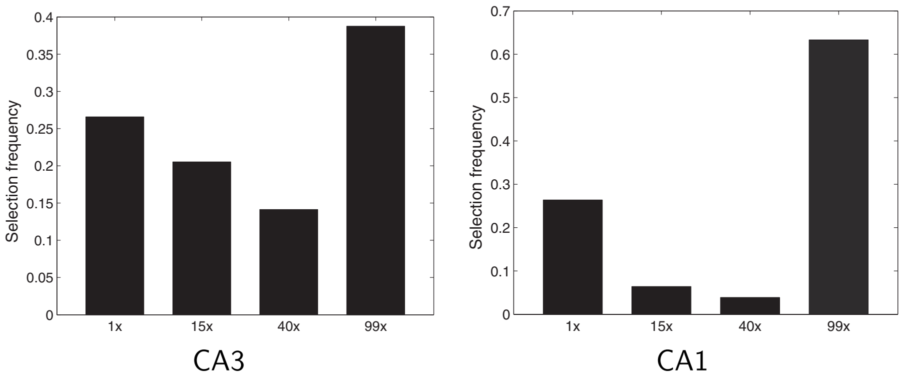

Figure 4.5: Percentage of samples in which each model was found to be the most consistent with the set of observed spiking data (equation. 4.13). The 99× model was often selected during jumps or intervals in which very few spikes were fired.

observed neural activity. Because of the probabilistic treatment, it becomes straightforward to integrate the contributions of both representation driven aspects of neural activity (e.g. above) and intrinsically driven neural dynamics such as refractory period (Frank et al., 2002).

4.3.5 Conclusions

A variety of experimental and theoretical results suggest the existence of cognitive processes that require active memory use in decision making. These processes are non-trivial to assess in human populations using such measures as self-report and are even more difficult to assess in non-human populations. Identifying such cognitive processes in non-human animals will require the development of measures to examine computations underlying these processes. Central to this approach is the development of statistical algorithms for decoding neural representations at multiple time scales and validation or error-assessment methods that allow characterization of cognitive processes related to, but not necessarily mirrored by, directly observable behavior. This chapter summarizes recent developments in methods for examining highly dynamic cognitive processes through observation of neural representations with multiple dynamics.

Neural decoding alone cannot be used to infer internal states of an animal's sensory and cognitive networks such as the difference between random firing and well-represented variables. This is particularly important when considering issues of memory and recall. One function of memory is to appropriately link a current experience to a past experience; in the case of the hippocampus, this may mean using the same spatial map as was previously used in an environment. However, a primary usefulness of a memory is in its ability to influence disconnected experiences through recall of past events or episodes. In this case of recall, one would expect that neuronal firing would, by definition, be disconnected from the current behavioral state of the animal. Recall may be detected by reconstruction methods identifying values very different from the current behavioral value. Usually, these values are considered noise to be removed from a reconstruction algorithm. Using consistency methods like those presented above allow discrimination between aberrant neural decoding and valid, non-local representational events.

Chapter 5

Experiments

A variety of observations support the existence of dynamic spatial representations in the rodent hippocampus (O'Keefe and Recce, 1993; Skaggs and McNaughton, 1996; Skaggs et al., 1996; Kudrimoti et al., 1999; Nádasdy et al., 1999; Louie and Wilson, 2001; Lee and Wilson, 2002; Foster and Wilson, 2006; O'Neill et al., 2006; Jackson et al., 2006). Although these dynamic representations have been used in theoretical explanations of spatial memory and navigation (Redish, 1999; Lisman, 1999; Jensen and Lisman, 2005), few experiments have examined dynamic spatial representations in the hippocampus during wayfinding behaviors. Based on recent methodological developments for decoding neural representations at fast timescales (Brown et al., 1998; Johnson et al., 2007, see chapter 4), the following experiments were performed to better describe the dynamic representational processes that support navigation.

5.1 Methods

5.1.1 Subjects

Male Fisher-Brown-Norway hybrid rats (Harlan, Indianapolis IN, age 7-14 mos at time of implantation) were maintained on a synchronous Day/Night cycle. Animals were handled for at least one week prior to beginning behavioral training. Rats were food-deprived to no less than 80% of their body weight during behavioral training; water was freely available in the home cage at all times. All procedures were in accordance with NIH guidelines for animal care and were approved by the IACUC at the University of Minnesota.

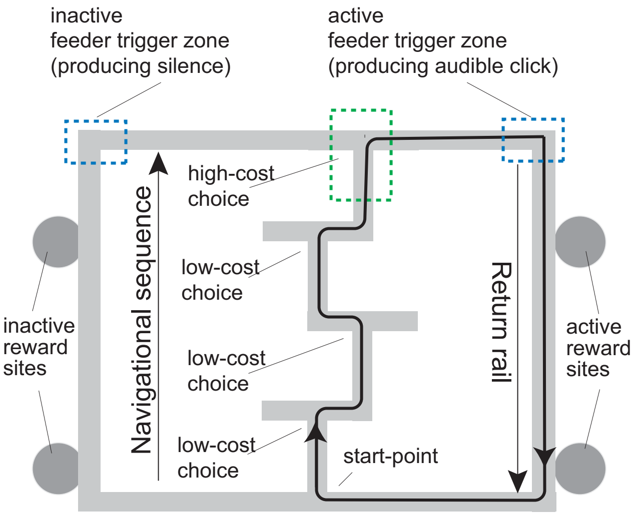

Figure 5.1: The Multiple-T maze. The task consists of four T choice points with food reward available at two sites on each return rail. Only feeders on one side of the track were rewarded in each session. Figure used with permission (Johnson and Redish, 2007).

5.1.2 Behavioral training

Tasks

The multiple-T task The task was identical to that used by Schmitzer-Torbert and Redish (2002, 2004, see Fig. 5.1). The task consisted of a sequence of 4 T turns with return rails after the final turn, making it a lap based task. Food was provided at two sites on each return rail. On any specific day, only one pair of sites (i.e. the right or left) provided food reward. However, both sites on that rewarded side provided food reward on each lap. Each feeder provided two 45 mg food pellets (Standard Formula P, Research Diets) on each lap. The sequence remained constant within a day, but changed between days. Errors were defined as entry into an incorrect maze arm. Rats ran one 40 minute session each day. Maze size is 165 cm × 165 cm. The trackwidth is 10 cm. There are $2^4 = 16$ available track configurations. On a 4T maze, chance of making a lap with no error is $(0.5)^4 = 0.0625$.

The cued-choice task The task consisted of a single T turn with return rails after the final turn, making it a lap based task. Food was provided at two sites on each return rail. On any specific lap, only one pair of sites (i.e. the right or left) provided food reward. A low frequency (3 kHz, for left) or high frequency (6 kHz, for right) cue tone played on the center stem and indicated which food sites would provide reward. A second tone that matched the cue tone sounded on the T arms if the rat made the correct decision (see Fig. 5.2). As a result, a rat often reversed its course if no tone was encountered on the T arms. Errors were defined as entrance into the incorrect tone area. Rats ran one 30-40 minute session each day. Maze size is 165 cm × 165 cm and the trackwidth is 10 cm.

Rats were trained following a pretraining protocol that first required the rats to run for laps to a single side of the track while the other side was blocked. The tone corresponding to the appropriate side was played at the primary and secondary cue sites on all laps and all laps were rewarded. Following single side pretraining, each rat was given two or three sessions of directed pretraining where they were blocked from making incorrect decisions. During this phase the rewarded side was determined according to pseudo-random order — reward was available pseudorandomly at the left or right feeders on each trial, but the number of left rewards and the number of right rewards was constrained to be equal within each ten- or twenty-trial block. Following directed pre-training the rats were provided undirected sessions with pseudo-random ordered blocks of ten or twenty laps. This ensured that the proportion of left and right choices remained approximately constant independent of the total number of laps run on a particular day, but also ensured that any decision-process that did not take the cue-tone into account would lead the rat to chance (50% correct). All neurophysiological analyses were taken from sessions in this final phase, which precluded the availability of any strategy that did not take the cue-tone into account.

Behavioral analysis

Pausing and variability of orientation of motion. Linear velocity was calculated using the adaptive windowing procedure proposed by Janabi-Sharifi et al. (2000). This algorithm provides for both accurate speed estimates and an accurate estimate of the time of speed change. Locations with a speed of less than 1 cm/sec were identified as *pausing locations*. Orientation of motion was calculated from the arctangent of the dx and dy terms of the linear velocity. Variability of orientation of motion was determined by applying the Janabi-Sharifi et al. (2000) algorithm to the orientation term. Locations with variability of motion $> \deg 120/\text{sec}$ were defined as high orientation variability locations.

Defining the "choice point". A small rectangle was identified around the final choice of the Multiple-T maze and the top of the central arm of the Cued-choice maze (see Figures 5.1 and 5.2). Anytime the animal entered this rectangle from the navigation sequence (Multiple-T) or the central arm (Cued-choice) was defined as a "choice". The time spent at the choice-point began when the animal entered the choice-point from the south and ended

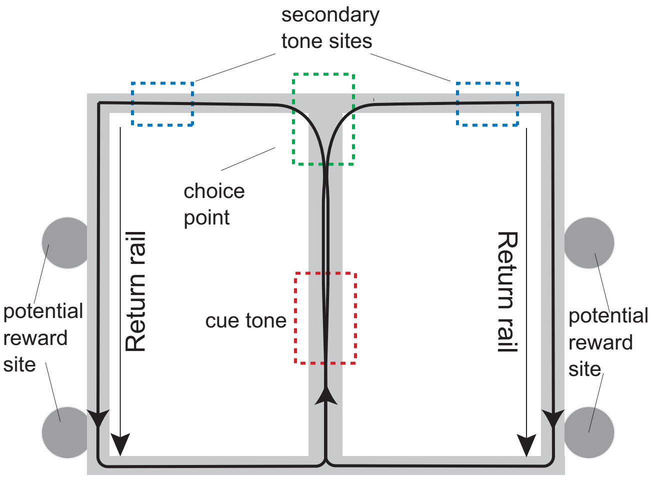

Figure 5.2: The Cued-choice maze. The task consists of a single T turn with food reward available at two sites on each return rail. On each lap a cue tone on the center stem (red dashed box) signalled whether the left or right arm would be rewarded. If and only if the rat made the correct turn, a matching tone would play at the T arm (blue dashed boxes). Figure used with permission (Johnson and Redish, 2007).

when the animal left the rectangle, turning either left or right. Times in which the animal returned across the choice-point rectangle after correcting an error (i.e. entering from the east or west) were not included in choice analyses.

5.1.3 Surgery and electrode implantation

After pre-training to proficiency, rats were implanted with a 14 tetrode microdrive (Kopf Neuro-Hyperdrive) directed toward the CA3 hippocampal subfield (Bregma -3.8mm A/P, 4.0mm M/L). Briefly, rats were anesthetized with Nembutal (sodium pentobarbital, 40-50 mg/kg, Abbott Laboratories, North Chicago, IL) and maintained using isoflurane (0.5–2% isoflurane vaporized in medical grade oxygen). The hyperdrive was secured in place with dental acrylic (Perm Reline and Repair Resin, The Hygenic Corp., Akron, OH) surrounding 8 anchor screws and a ground screw. Immediately following surgery, animals received Ketoprophen (5 mg/kg, subcutaneously). Following surgery, three doses of 5 mL sterile saline (0.9%) mixed with Baytril (in a dose of 25 mg/kg) were administered subcutaneously (one dose every 12 hours). Animals were allowed to recover in an incubator until they were ambulatory. Once animals were ambulatory, 0.8 mL Children's Tylenol was administered orally. For two days following surgery, rats had access to water containing Children's Ibuprophen (15 mg/kg, available in a dilution of 15 mg/kg/20-25 mL water). Rats were allowed two days to recover from surgery before resuming behavioral experiments. In the first few hours post-surgery, all tetrodes were advanced approximately 1mm. Twelve tetrodes were slowly advanced toward the pyramidal cell layer over the ten days following surgery, one tetrode was lowered to the hippocampal fissure, and one tetrode was left in corpus callosum or a quiet region of cortex to be used as a superficial reference.

Following task performance, rats were overdosed on nembutal and perfused with Formalin. After 24 hours in formalin, brains were transferred to a 30% sucrose-Formalin solution, sliced, and stained with cresyl violet using standard procedures. All recording locations were verified to lie in the CA3 region of the dorsal hippocampus.

5.1.4 Recordings

Neurophysiology

Spike activity was recorded using a 64 channel Neuralynx Cheetah system (Neuralynx, Tucson AZ), which sampled voltage at 32kHz for each channel, filtered between 600Hz and 6kHz. When the voltage on any of the four channels of a single tetrode reached threshold, a 1ms window of the spike waveform on each of the four channels on the tetrode was recorded and timestamped with microsecond resolution (Neuralynx). Following the experiment, spikes were manually clustered into putative cells on the basis of their waveform properties (MClust 3.4, Redish et al. http://www.umn.edu/ redish/mclust), based on automatically-derived clusters found via k-means and expectation-maximization algorithms (KlustaKwik 1.5, K.D. Harris http://klustakwik.sourceforge.net/). Local field potentials (LFP) were sampled at 2 kHz and bandpass filtered from 1–425 Hz. Cross-frequency correlation analysis was conducted as in Masimore et al. (2004). Data for theta-frequency analysis were taken from the fissure tetrode. Data for high-frequency ripple analysis were taken from pyramidal cell layer tetrodes. CA3 was initially distinguished by the presence of low frequency sharp wave ripples at the pyramidal cell layer (Ylinen et al., 1995; Csicsvari et al., 1999).

Behavior: tracking and control

5.1.5 LFP analysis

The frequency composition of local field potentials at the choice point was examined using frequency-detection methods capable of detecting transient oscillatory events in long data-series (Masimore et al., 2004). The method calculates the cross-frequency self-coherence within a signal. Because local field potential oscillations signals do not occur at a single, sharp frequency (i.e. sharp wave ripple complexes include components from the 100–220 Hz range O'Keefe and Nadel, 1978), they appear as identifiable pockets of high correlation in the cross-frequency correlation plot (Masimore et al., 2004).

Subsequent analysis examined specific frequency ranges or frequencies of interest. Sharp wave ripples were identified as epochs in which the local field potential power increased to greater than $7\sigma_{SW}$ within the 100-250Hz frequency range (Csicsvari et al., 1999). The standard deviation of the ripple power σ_{SW} was found from all data collected within the session. Similarly, gamma power was examined in the 30-80Hz frequency range and was also normalized by the standard deviation (σ_g) for comparison. The standard deviation of gamma power σ_g was found from all data collected within the session.

LFP Phase.

LFP data were bandpass-filtered at theta frequencies between 6-10Hz (128th order). Theta phase was estimated using basic Hilbert transform methods (Matlab). To determine whether sweeps times were aligned with theta, phase analysis compared the distribution of phases at sweep start and end times versus a random distribution using Watson's test for circular data (Zar, 1999).

5.2 Results

5.2.1 Behavior.

Rats learned to run both the Multiple-T and Cued-choice tasks successfully (rats made significantly more correct turns than chance; Multiple-T: binomial test, $p < 10^{-5}$; Cued-choice: binomial test, $p < 0.05$). On the Multiple-T task, rats began not different from chance and quickly reduced errors within a single session (Figure 5.3). On the Cued-choice task, rats chose correctly better than 70% of the time (see Figure 5.4). Because of the pretraining protocol in which each cue-tone was associated with a direction while the other side was blocked, it was possible for rats to start the open (decision-making) phase of the Cued-choice task above chance.

5.2.2 Electrophysiology.

Two hundred thirty-seven spike-trains were recorded from the dorsal CA3 region of three rats (four sessions with ensembles of 33 to 72 cells, see Table 5.1). Electrodes were confirmed

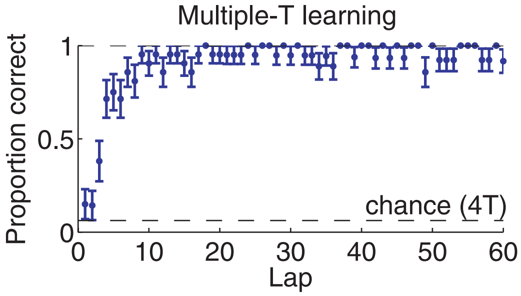

Figure 5.3: Learning on the Multiple-T task. Within each session, the probability of completing a lap without an error started not significantly different from chance and increased quickly over the initial laps. Mean and standard error shown for 21 novel sessions over six animals. Figure used with permission (Johnson and Redish, 2007).

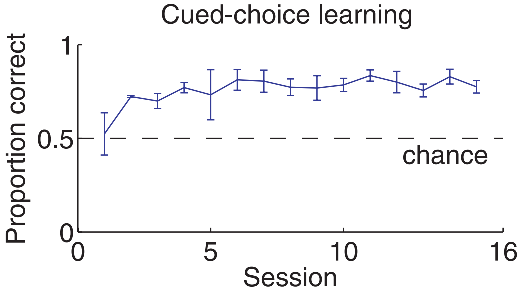

Figure 5.4: Learning on the Cued-choice task. Over 15 sessions, the probability of making the correct choice increased significantly. Mean and standard error shown for both left and right choices over two animals. Figure used with permission (Johnson and Redish, 2007).

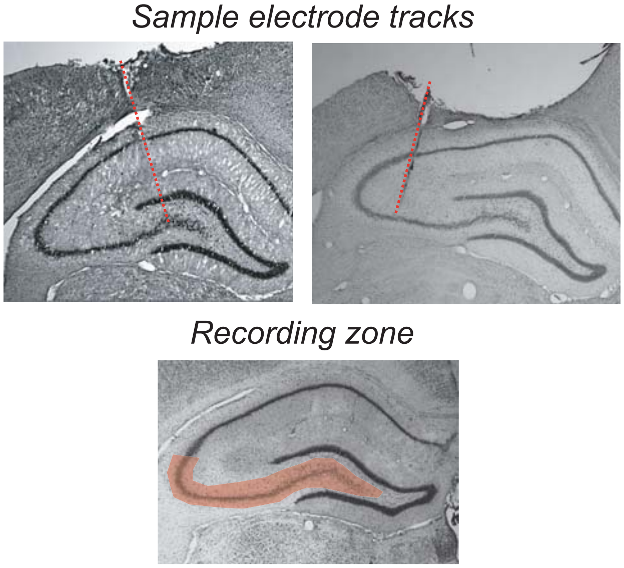

Figure 5.5: Recording sites came from the CA3 region of hippocampus. Figure used with permission (Johnson and Redish, 2007).

to lie in the CA3 region of the hippocampus (Fig. 5.5).

Rat	Session	Task	Number of cells	Laps	Performance
R051	2004-09-06	Multiple-T	62	36	92% ($p < 0.05$)
R076	2006-03-09	Multiple-T	70	85	78% ($p < 0.05$)
R076	2006-03-17	Cued-choice	72	49	63% ($p < 0.05$)
R085	2006-07-27	Cued-choice	33	61	93% ($p < 0.05$)

Table 5.1: Data summaries for each session. Probabilities calculated using a binomial test.

Individual cells showed well-defined place fields and sometimes multiple place fields similar to other track based tasks. The distribution of place fields covered the environment (place field coverage ranged from 2–25 cells at each pixel). However, these well-isolated cells also displayed scattered spiking activity outside their place fields during specific behaviors. Extra-field spiking occurred simultaneously across a subset of the neural ensemble at feeder sites during grooming, eating, and other pausing behaviors during sharp waves punctuating hippocampal non-theta, LIA states (consistent with previous observations of reactivation and replay at feeder sites Kudrimoti et al., 1999; Jensen and Lisman, 2000; Foster and Wilson, 2006; O'Neill et al., 2006; Jackson et al., 2006). However, extra-field firing also occurred during the theta local-field potential state, coincident with pausing behavior at the final choice point of the Multiple-T task, at the main choice point of the Cued-choice

Session	sweeps/lap	sweeps/second	length of sweeps	proportion of time in choice point spent in sweeps
R051-2004-09-06	1.10	1.2	120±9 ms	15%
R076-2006-03-09	0.36	0.30	90±5 ms	3%
R076-2006-03-17	1.10	0.46	190±26 ms	9%
R085-2006-07-27	0.21	0.50	190±16 ms	9%

Table 5.2: Sweep statistics.

task, and as the rat revised his choice after making a decision (error correction).

5.2.3 Extra-field firing at the high-cost decision-point.

When the rat reaches the end of the navigation sequence on the Multiple-T maze, he faces a high-cost choice — making the incorrect choice means a long (>3 meter) journey around the track before being given another chance. Rats often paused at that final choice. During those pauses, cells with primary place fields on the top arm and return rails fired spikes even though the rat was outside the cell's place field (See Figure 5.6.) Although the reconstructed distribution $p(x|S)$ tracked the animal well through the main path of the maze, when the rat paused at the final choice, the reconstructed distribution became non-local, sweeping ahead of the rat, first down one choice, then the other (Figure 5.7). This non-local neural activity occurred reliably when the animal paused at the final choice of the sequence. These events were rare, but reliably observable, occurring at a rate of 0.63 ± 0.21 sweeps/sec (mean ± SE, calculated over sessions) while animals were at the choice point. Sweeps lasted for 0.15 ± 0.025 seconds (mean ± SE, calculated over sessions). Animals spent an average of $9\% \pm 2\%$ (mean ± SE, calculated over sessions) of their time at the choice point in these non-local representations; during the remainder of the time, representations were primarily local. This meant that animals showed 0.7 ± 0.2 sweeps/lap (mean ± SE, calculated over sessions). See Table 5.2.

Non-local representations at the choice-point preferentially sweep forward of the animal.

To test whether the non-local reconstruction at choice points occurred preferentially ahead of the animal or behind the animal, we compared the proportion of the reconstructed probability distribution $p(x|S)$ reconstructing ahead of the animal to the proportion behind the animal. As can be seen in Fig. 5.8, the non-local reconstruction probability was concen-

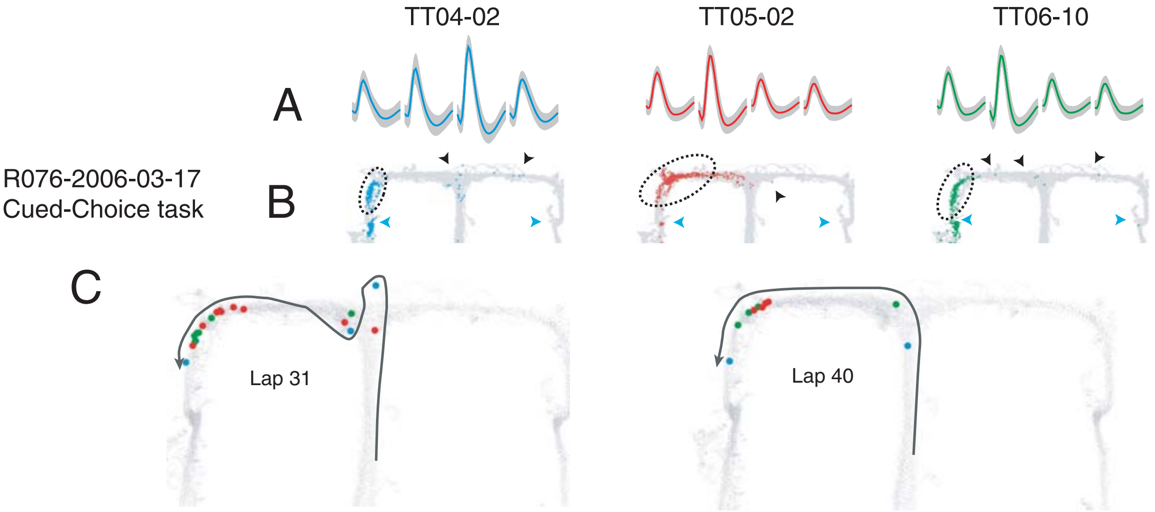

Figure 5.6: Extra field firing in hippocampal place cells at the choice-point on the Cued-choice task. A: Average waveforms (gray=SD) for three simultaneously recorded, well-isolated place cells. B: Spatial activity of each cell across the full session. The light gray lines indicate the rat's positions. Colored dots indicate the rat's position at the time of a spike. Blue triangles indicate feeder locations; black triangles indicate extra-field spiking. Dotted lines delineate place field boundaries. Place fields were defined as a contiguous region with average firing rate exceeding 20% peak field rate (Huxter et al., 2003; Leutgeb et al., 2004). Note the sparseness preceding the place field in the red cell (TT05-02) in panel B. It is possible that some of this sparse early activity is also extra-field firing. Classically-defined place fields are included for illustration only; all spikes were included in all analyses. C: Place cell activity on specific trajectories. Light gray lines again indicate animal positions across the full session. Colored dots show animal positions at the time of cell spiking. Black trajectory arrows indicate direction of motion along the trajectory. Note that extra field firing primarily occurs at the choice point. Figure used with permission (Johnson and Redish, 2007).

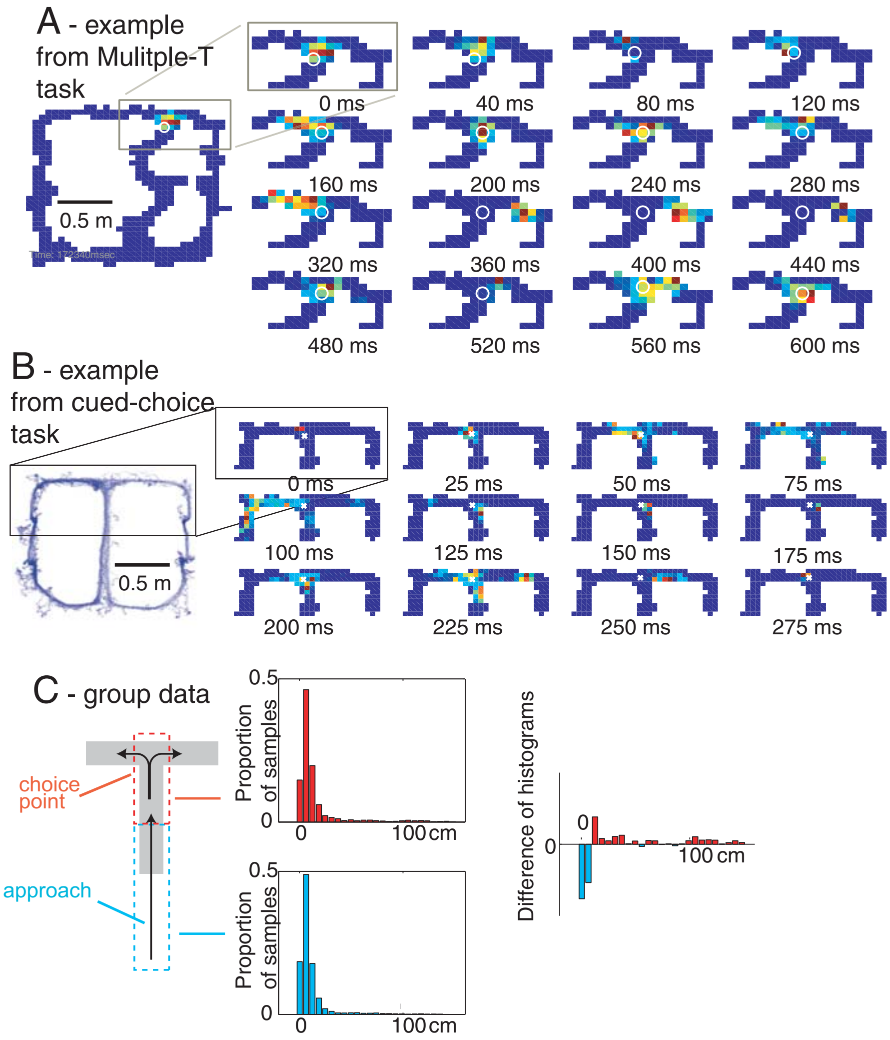

Figure 5.7: Forward-shifted neural representations at the choice point. The representation closely tracked the rat's position through the stem of the final T choice point for both the Multiple-T (A) and Cued-choice tasks (B). When the rat paused at the final choice point, the representation moved ahead of the animal and sampled each arm. The representation intensity is shown in pseudo-color (red high probability, blue low probability) and the animal's position shown as a white o. C: Distribution of distances between reconstructed location and actual location for the choice point (red,top) and for the immediately-preceding duration-matched approach to the choice point (cyan,bottom). The approach contains more local representations while the choice point contains more non-local representations. The medians of the two distributions are different (Wilcoxon rank-sum test, $p < 10^{-10}$). Figure used with permission (Johnson and Redish, 2007).

trated predominantly ahead of the animal. Forward locations were averaged together and the difference between forward locations and backward locations was calculated for each 125 ms time window that the animal was at the central choice point. The median difference (shown by the arrow) was significantly greater than 0 in all measured data sets. (SignRank test, $p < 10^{-10}$ for all sessions).

Non-local representations at the choice-point sweep down the two choices separately.

Given that the non-local reconstruction is concentrated forward of the animal rather than behind it, an important question is whether it is simply spreading forward of the animal in all available directions (as would be predicted by current models of phase precession, Skaggs et al., 1996; Yamaguchi, 2003; Jensen and Lisman, 2005), or whether it is concentrated on each choice separately. Examining the single time-window slices in Figure 5.7 suggests that the representation swept forward along only one choice at a time. In order to directly test this hypothesis, we measured the joint probability of the distribution of reconstruction contributing to the left and right boxes shown in Fig. 5.8. As can be seen in Fig. 5.9, the high probabilities were found primarily along the axes of the two-dimensional histogram and not near the diagonal, suggesting that decoded representations did not spread in both directions simultaneously.

5.2.4 Extra-field firing during error-correction.

The hypothesis that these extra-field firing and non-local representations occurring at key decision points (such as the high-cost choice point on the two tasks) predicts that extra-field firing and non-local representations should also occur at other decisions made within the tasks. On these two tasks, animals were allowed to turn around on the top arm (before the final turn onto the return rails). This means that rats were faced with a decision on error laps in which they made the wrong choice on the high-cost choice. On the Multiple-T task, this occurred when the rat did not hear the expected feeder click sound when it reached the feeder trigger zone (see Figure 5.1). On the Cued-choice task, this occurred when the rat did not hear the expected confirmation tone (secondary tone site, Figure 5.2). As predicted, place cells showed extra-field firing on error laps (Figure 5.10). Reconstruction analyses found that the $p(x|S)$ distribution swept down the two potential choices (forward to the feeder or back towards the choice point and the other choice, Figure 5.11).

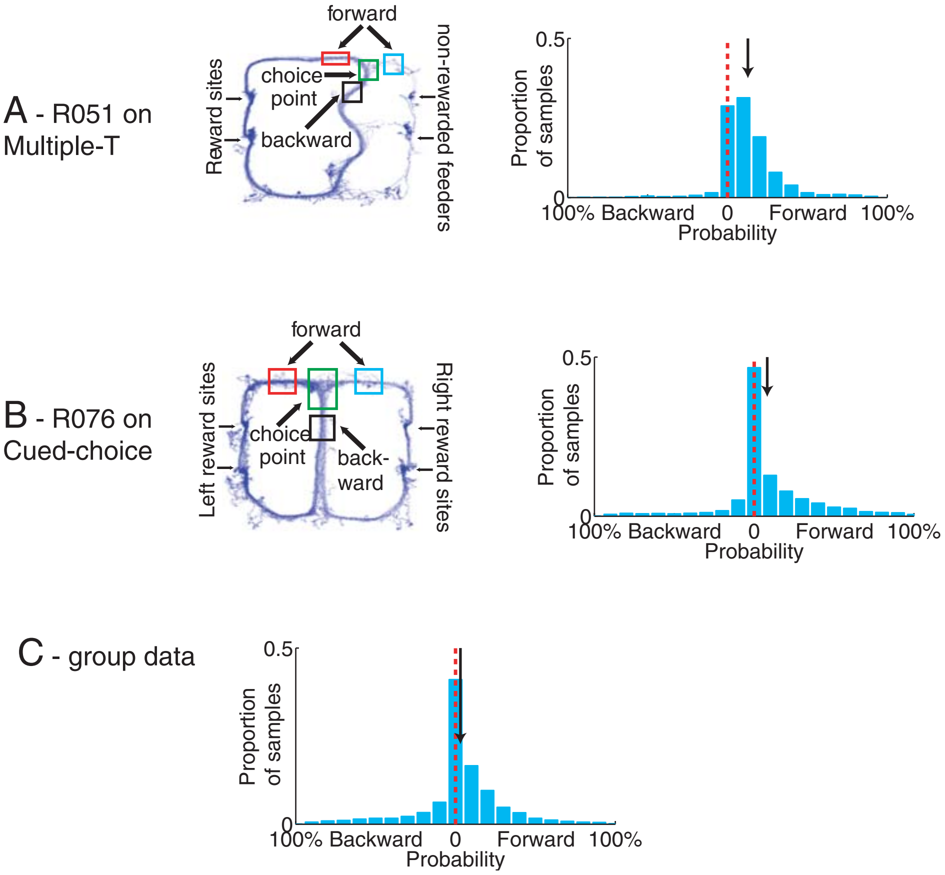

Figure 5.8: The non-local reconstruction events occurring at the choice point are forward of the animal. A,B: Specific examples from two data sets. Three regions of interest were defined (to the left of the animal, to the right of the animal, and behind the animal). Proportions of reconstructed probability $p(x|S)$ for each region were measured. Red line indicates 0, black arrow indices median. Median significantly greater than 0 in all measured data sets. A: Multiple-T. B: Cued-choice. C: Joint probability between forward probability intensity and backward probability intensity (log units). Note the strong preference for forward over backward. Figure used with permission (Johnson and Redish, 2007).

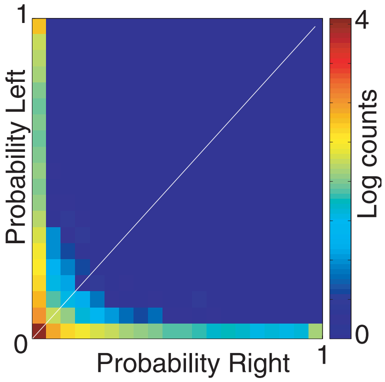

Figure 5.9: The non-local reconstruction events occurring at the choice point are concentrated in each arm separately. Joint probability of probability intensity concentrated on each arm (log units). Note the strong preference for one arm over the other with no joint probability in both directions simultaneously. Figure used with permission (Johnson and Redish, 2007).

Non-local representations are correlated but not defined by the orientation of motion.

As can be seen in Figure 5.9, the reconstruction probability forward of the animal predominantly concentrated at either the left or right directions. It is thus an important question to measure the extent to which these directions of forward representation relate to the orientation of the animal. Orientation of motion of the animal was measured as defined above (see Methods). While there was a significant correlation ($p(\text{slope}{=}0) < 10^{-10}$), there were also cases in which the probability was predominantly leftward while the rat moved rightward and vice versa (see Figure 5.12).

Non-local representations occur at locations of highly variable orientation of motion.

The observation that the choice point of the Cued-choice task, the final choice of the Multiple-T task, and at reversals on both tasks suggests that non-local representations may occur generally where rats pause and re-orient. In order to test this hypothesis, we calculated locations where animals paused (total linear velocity $<$ 1 cm/sec), locations where animals varied their orientation of motion (total angular velocity $>$ deg 120/ sec), and compared the reconstruction error distance during these conditions. As can be seen in Figure 5.13, the animals tended to show variable orientation of motion at the feeders, at

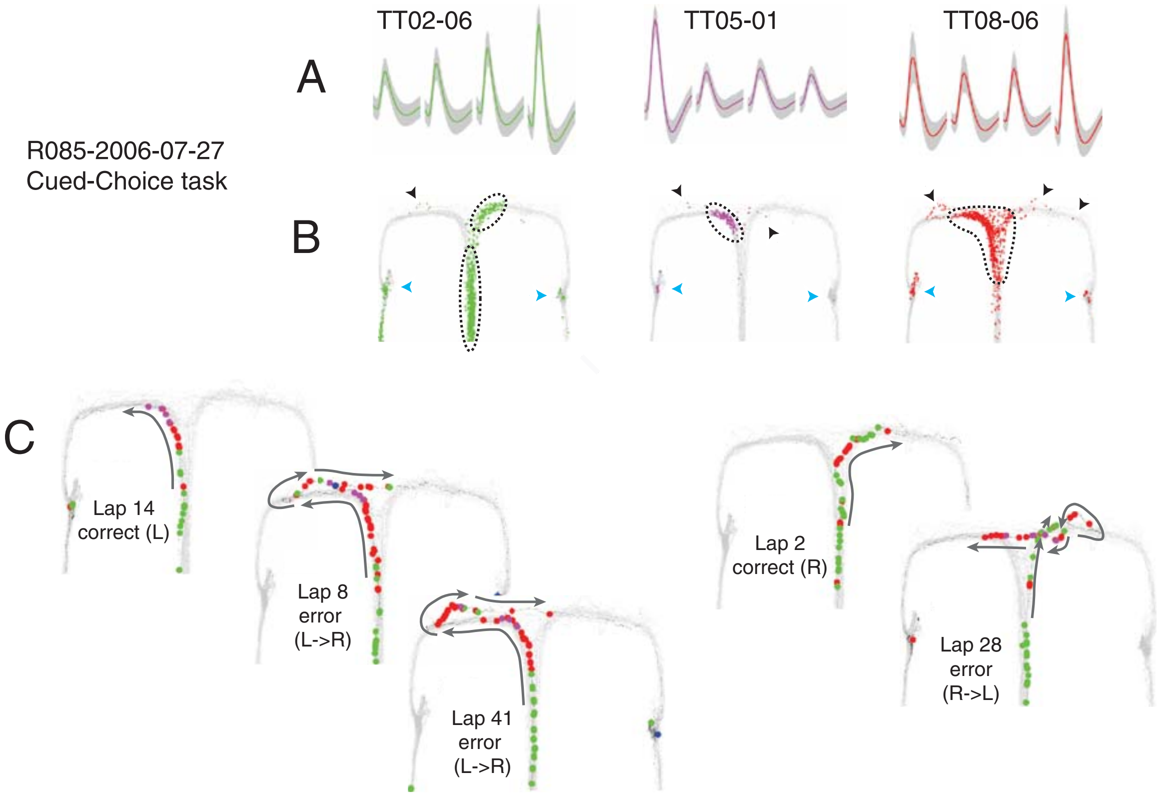

Figure 5.10: Extra field firing in hippocampal place cells during error-correction on the Cued-choice task. A: Average waveforms (gray=SD) for three simultaneously recorded, well-isolated place cells. B: Spatial activity of each cell across the full session. The light gray lines indicate the rat's positions. Colored dots indicate the rat's position at the time of a spike. Blue triangles indicate feeder locations; black triangles indicate extra-field spiking. Dotted lines delineate place field boundaries. Place fields were defined as a contiguous region with average firing rate exceeding 15% peak field rate (Huxter et al., 2003; Leutgeb et al., 2004). C: Place cell activity on specific trajectories. Light gray lines again indicate animal positions across the full session. Colored dots show animal positions at the time of cell spiking. Black trajectory arrows indicate direction of motion along the trajectory. Note that extra field firing sometimes occurs before the animal turns back toward the choice point on reversals. Figure used with permission (Johnson and Redish, 2007).

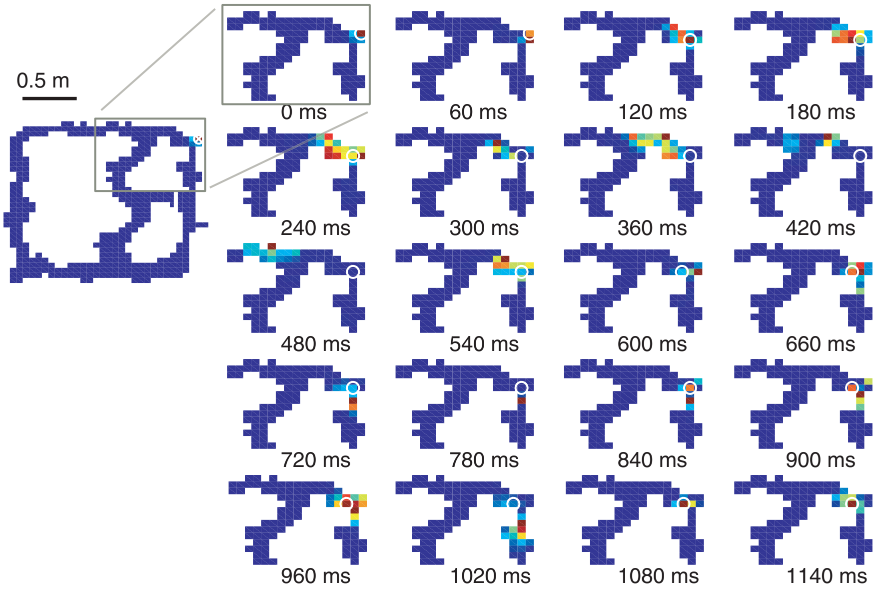

Figure 5.11: Error correction in the hippocampal neural ensemble. Again, the representation closely tracked the rat's position from the choicepoint to the feeder trigger zone. The rat turned back toward the choice point and the representation moved into the opposite maze arm. The representation intensity is indicated by color (red high probability, blue low probability), and the actual position of the animal is indicated by the o. Panels arranged from left to right, top to bottom in 60msec intervals. Figure used with permission (Johnson and Redish, 2007).

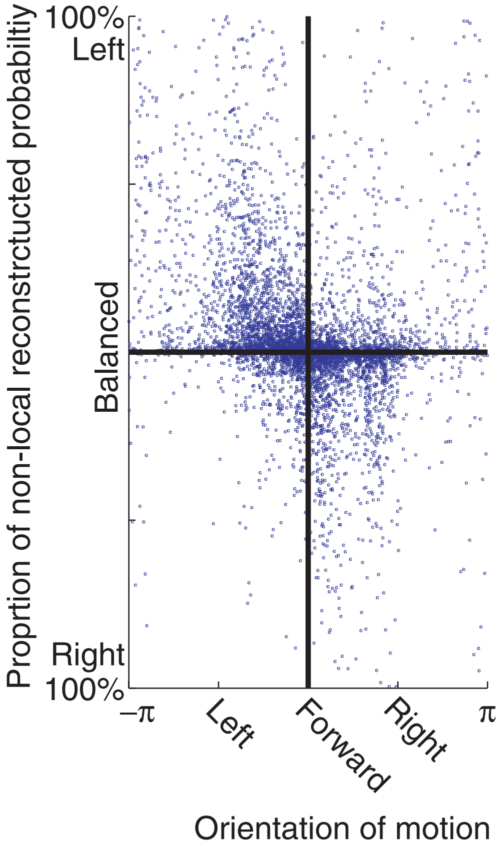

Figure 5.12: Direction of non-local reconstructed representation as a function of the orientation of motion of the animal. For all samples at the choice-point, direction of non-local reconstruction was measured as the proportion to the left of the animal minus the proportion to the right of the animal. While there was a significant correlation ($p(\text{slope}=0) < 10^{-10}$), there were also samples reaching leftward while the animal moved rightward and vice versa. Figure used with permission (Johnson and Redish, 2007).

the final choice-point, and at recovery from errors. When the distribution of reconstruction errors was compared between these variable orientation of motion positions and the rest of the task, variable orientation of motion positions included more non-local reconstructions (Wilcoxon rank-sum test, $p < 10^{-10}$).

Local field potentials.

Extra-field firing has been known to occur during sharp-wave ripple complexes within LIA episodes (O'Keefe and Nadel, 1978; Kudrimoti et al., 1999; Jensen and Lisman, 2000; Foster and Wilson, 2006; O'Neill et al., 2006; Jackson et al., 2006). Local field potential traces at the choice point displayed strong theta oscillations and no apparent sharp wave ripple activity (Fig. 5.14). To determine the hippocampal brain-state and identify whether sharp wave ripple complexes within theta might be occurring at the choice point, we examined the local field potential frequency composition using frequency-detection methods capable of detecting even very rare and transient oscillatory events in long data-series (Masimore et al., 2004). As shown in Fig. 5.15, the method clearly identifies both theta and sharp-wave frequency components across complete sessions (left panel). However, when only data in which the animal was at the choice point were considered, no sharp-wave components were seen. Instead, the analysis identified the presence of theta (7 Hz) and gamma (30-80 Hz) frequencies. This analysis method is sensitive enough to identify even transient sharp-wave ripple complexes occurring within theta episodes (eSWR, O'Neill et al., 2006). As can be seen in panels B and C, there was no power at sharp-wave ripple frequencies during pausing behavior at the stem or choice point. Comparisons between choice point local field potentials at the choice point, the maze stem and the feeders revealed slight differences between choice point and the stem local field potentials, but large differences from those at the feeders. Local field potentials at the feeders displayed diminished gamma power and increased sharp wave ripple activity. In summary, non-local representations were observed at both the feeders (reactivation, route replay, (Jackson et al., 2006; Johnson et al., 2007)) and at choice points (sweeps), but occurred in much different local field potential 'brain states.'

As suggested by Figure 5.14, these sweep events are occurring during hippocampal states including strong theta components. As shown in Table 5.2, the mean length of the sweep corresponded closely to that of a single theta cycle (mean sweep length = 153 ± 13 ms, not significantly different from 7 Hz [140 ms], z-test, $p = 0.86$). However, some sweeps did last much longer than a single theta cycle (length of longest sweep = 1.4 seconds). In order to analyze the phase-relationships between the sweeps and theta, sweep start and stop

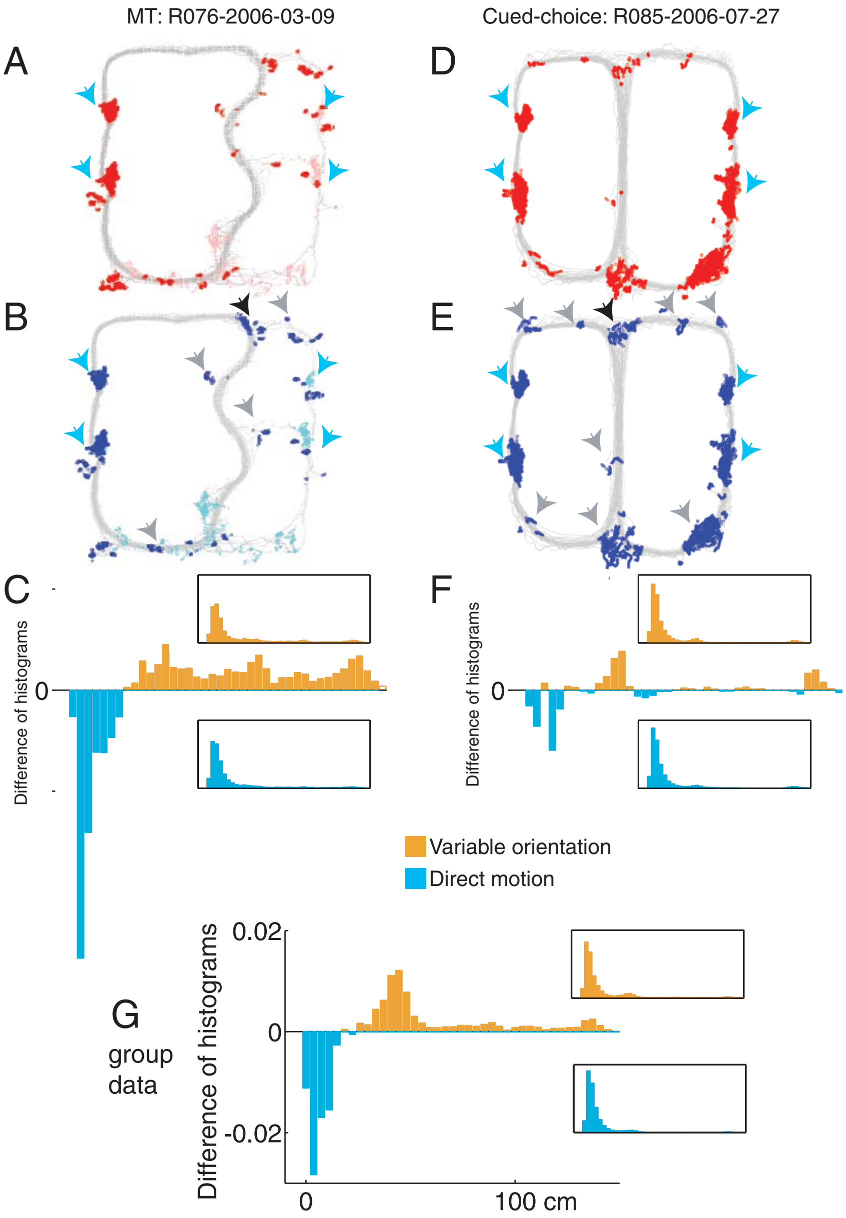
MT: R076-2006-03-09
Cued-choice: R085-2006-07-27
A
D
B
E
C
Difference of histograms
0
F
Difference of histograms
0
Variable orientation
Direct motion
G
group data
Difference of histograms
0.02
0
-0.02
0
100 cm

Figure 5.13: Reconstruction error is more non-local at locations where animals show highly variable orientation of motion (figure on previous page). (A,D) Locations where the animal paused. Gray dots indicate all locations sampled during the full session, red dots indicate locations where the animal paused. (Lighter color dots in panel A indicate pauses occurring in the last two laps, during which the animal no longer actively searched for food.) (B,E) Locations where the animal showed a high variability of orientation of motion. (Lighter color dots in panel B indicate variable orientations occurring in the last two laps, during which the animal no longer actively searched for food.) (C, F) Difference between histograms of reconstruction error distance during samples with variable orientation of motion [blue dots in panels B,E] and during the rest of the session. Reconstruction distance was greater during samples with variable orientation of motion (Wilcoxon rank-sum test, (C) $p < 10^{-10}$, (F) $p < 0.00005$). (A-C) Multiple-T. (D-F) Cued-choice. Blue triangles indicate feeder locations. Black triangles indicate choice point studied in Figure 5.7. Gray triangles indicate other pause locations during which animals spent time varying their orientation of motion, including the reversal locations studied in Figure 5.11. (G) Group data, showing that the tendency to show an increased non-locality of reconstruction during variable orientation motion is a general property of the entire data set. Figure used with permission (Johnson and Redish, 2007).

times were identified as described in the Methods. While sweep start times showed a non-significant trend to primarily occur at phases close to the peak of theta at the hippocampal fissure (Watson's circular $U^2 = 0.12, p = 0.1$), sweep stop times showed a statistically significant tendency to stop on the rising phase of theta (Watson's circular $U^2 = 0.52, p < 0.0001$). See Figure 5.16.

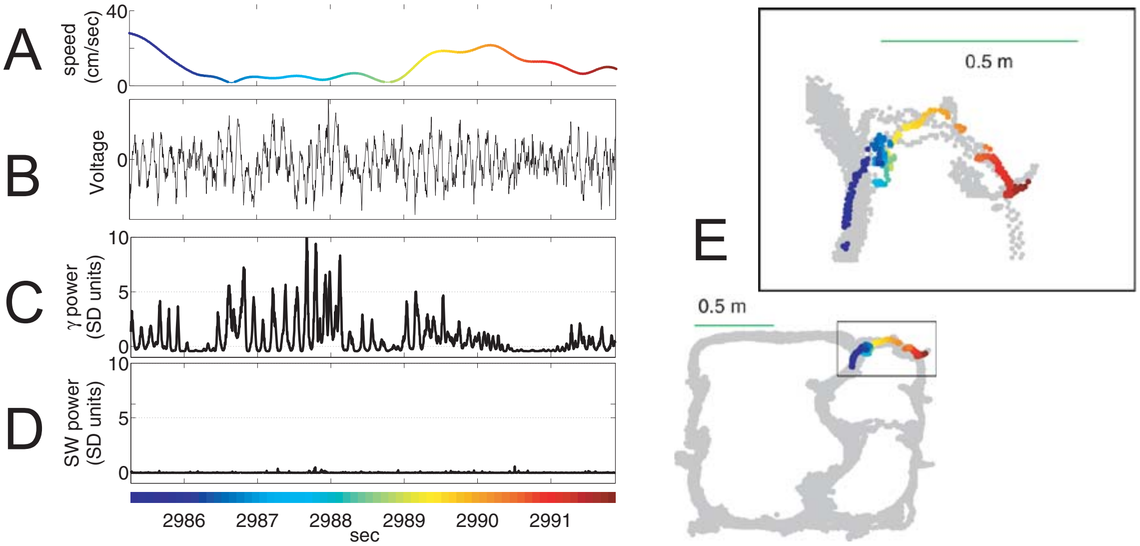

Figure 5.14: CA3 LFP during choice-point behavior. A-E: Sample local field potential during a pause at the choice point (see Figure 5.7). A: Absolute velocity. B: Raw local field potential trace. C Gamma power (filtered at 30–80 Hz) z-scores based on the distribution drawn from full session. D: Sharp wave ripple power (filtered at 100-250 Hz) z-scores based on the distribution drawn from full session. E: Gray dots indicate all positions sampled by the animal over the entire session. Colored dots indicate path of animal sampled during time corresponding to LFP data on the left. Colors indicate passage of time and are consistent between left and right panels. Figure used with permission (Johnson and Redish, 2007).

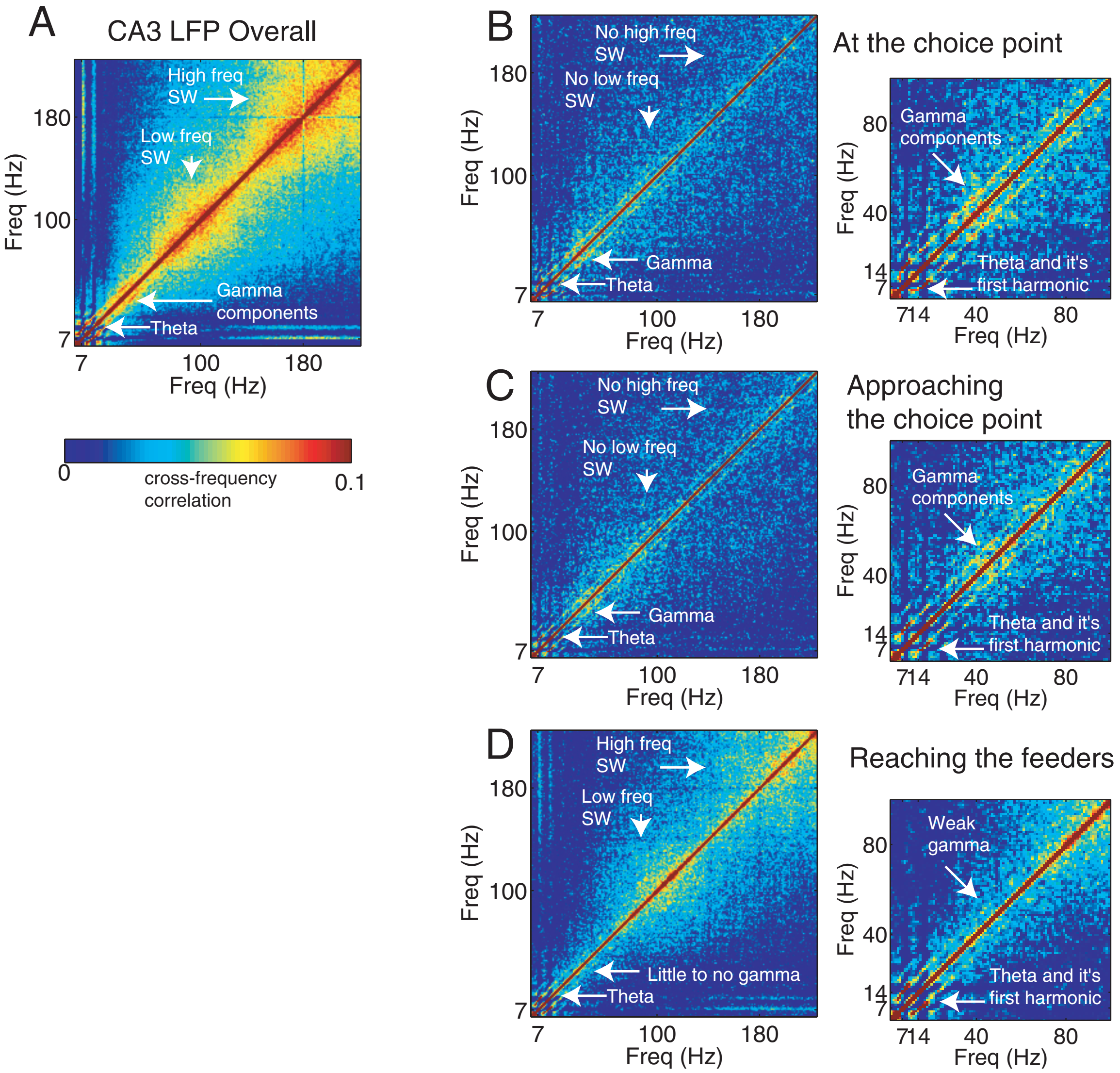
A
CA3 LFP Overall
High freq SW
Low freq SW
Gamma components
Theta
180
100
7
Freq (Hz)
7
100
180
Freq (Hz)
0
cross-frequency correlation
0.1
B
No high freq SW
No low freq SW
Gamma
Theta
At the choice point
Gamma components
Theta and it's first harmonic
C
No high freq SW
No low freq SW
Gamma
Theta
Approaching the choice point
Gamma components
Theta and it's first harmonic
D
High freq SW
Low freq SW
Little to no gamma
Theta
Reaching the feeders
Weak gamma
Theta and it's first harmonic
80
40
14
7
714
40
80

Figure 5.15: CA3 LFP during choice-point behavior and other maze areas (figure on previous page). A: Average cross-frequency correlations from all available 40 minute sessions. In these plots, important frequencies show up as high correlated blocks (Masimore et al., 2004). Data includes additional sessions from which spikes were not available. (Data averaged over 16 sessions from three animals). Note the clear indications of theta (at 7–10 Hz) and the low-frequency (100-150 Hz) and high-frequency (200 Hz) sharp waves in the overall plot (left). B: Average cross-frequency correlations at choice point times. C: Average cross-frequency correlations at stem times. D: Average cross-frequency correlations at feeder arrival. Note that local field potentials at the choice point and stem include gamma (30-80 Hz) but no low- or high-frequency sharp waves. This contrasts with feeder times which display decreased theta and gamma correlations, but strong correlations at low- and high-frequency sharp wave ripple frequencies. Figure used with permission (Johnson and Redish, 2007).

Development of non-local representations with experience.

If non-local representations within the hippocampus provide information to decision-making brain modules, then these representations may change as a function of task demands. To examine this possibility we compared within-session changes in the non-local spatial representations at the choice point on the Multiple-T and Cued-choice tasks. The primary difference between the two tasks is within-session learning, in that the spatial position of reward on the Multiple-T task is consistent within each session and changes from day to day (Schmitzer-Torbert and Redish, 2002). In contrast, the spatial position of reward on the Cued-choice task varies from lap to lap but the relationship between cues and reward is maintained from day to day.

Behaviorally, rats on the Cued-choice task paused for longer at the choice point than rats on the Multiple-T task (t(234)=2.49, p<0.05). Over the course of each session on the Multiple-T task, rats paused for less time on later laps than on earlier laps (significant negative slope, choice point time × lap linear regression; $F(1,91) = 21.59$, $p < 0.05$). In contrast, no such reduction was seen on the Cued-choice task (choice point time × lap linear regression; $F(1,81) = 0.03$, NS).

As can be seen in Fig. 5.17, during the first part of the Multiple-T task, the reconstructed distribution $p(x|S)$ first swept down both possibilities, but then, as the behavior automated, sweeps down the unrewarded choice (left on the session shown) decreased. In late laps, the total amount of non-local reconstructed probability decreased. In contrast, on the Cued-choice task the forward component of the sweep maintained non-locality sweeping down both directions throughout the task. There was a significant effect of both task, lap, and their interaction on the proportion of the probability sweeping to the left or the right at

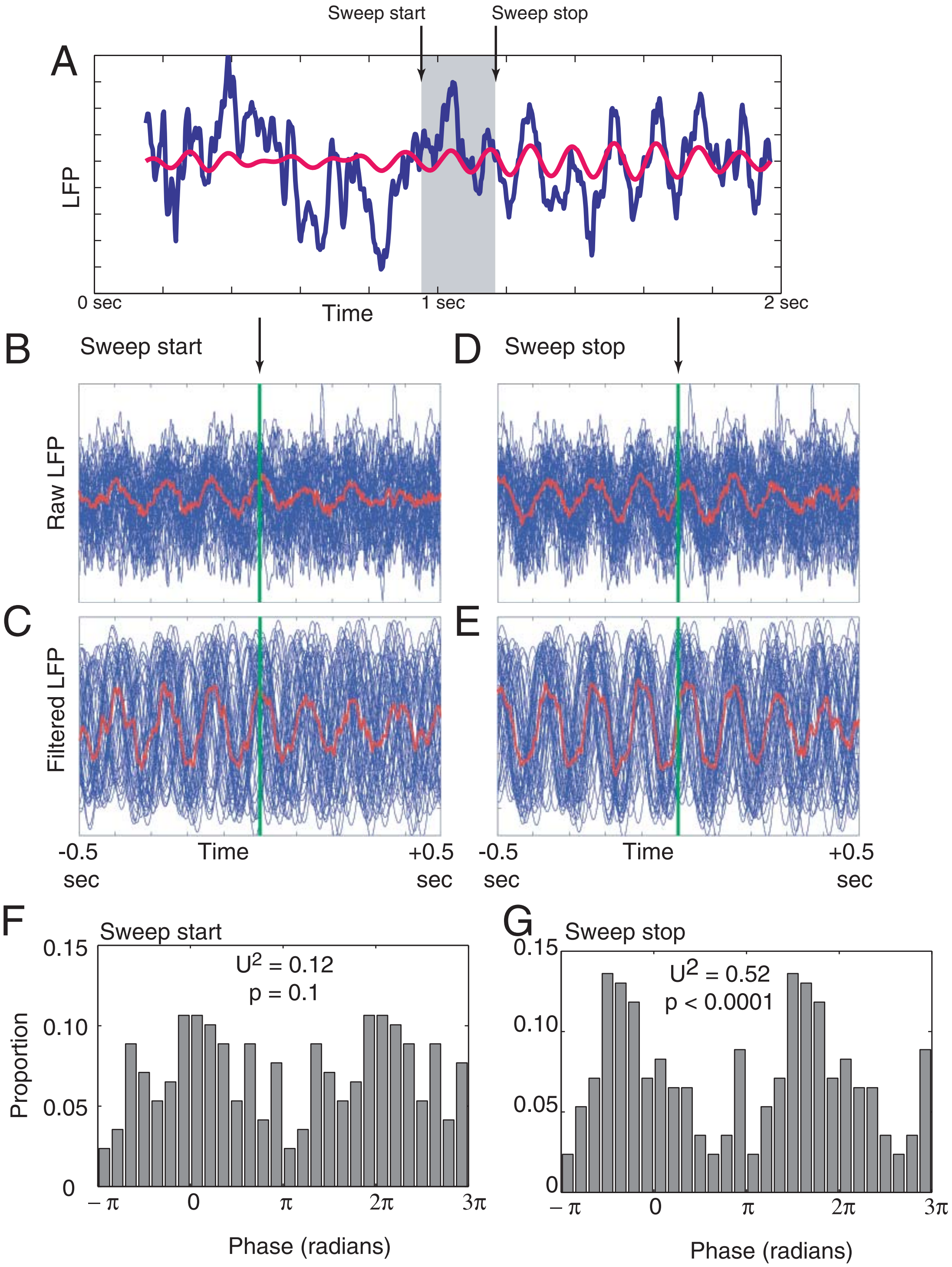
Sweep start
Sweep stop
A
LFP
0 sec
1 sec
2 sec
Time
B
Sweep start
D
Sweep stop
Raw LFP
C
E
Filtered LFP
-0.5 sec
Time
+0.5 sec
-0.5 sec
Time
+0.5 sec
F
Sweep start
$U^2 = 0.12$
$p = 0.1$
Proportion
0.15
0.10
0.05
0
$-\pi$
0
π
2π
3π
Phase (radians)
G
Sweep stop
$U^2 = 0.52$
$p < 0.0001$
0.15
0.10
0.05
0
$-\pi$
0
π
2π
3π
Phase (radians)

Figure 5.16: Phase-relationships between sweeps and theta (figure on previous page). A: Example local field potential showing the raw (blue) and theta-filtered (red) local field potential surrounding a 200 ms sweep. Note how the sweep spans several theta cycles. B: Distribution of LFP aligned to start times from all sweeps occurring within one session R076-2006-03-09. Start time corresponds to the green line. All samples are shown in blue, average is shown in red. C: Distribution of filtered LFP signals aligned to start times from all sweeps occurring within that same session. D,E: equivalent plots showing LFP alignment of sweep stop times. F: Group data for all sessions showing histogram of distribution of theta phase of sweep start times. G: Group data for all sessions showing histogram of distribution of theta phase of sweep stop times. Figure used with permission (Johnson and Redish, 2007).

any time (ANOVA, effect of Lap: df$= 38, F = 9.4, p < 10^{-10}$, effect of Task: df$= 1, F = 113.3, p < 10^{-10}$, interaction: df$= 38, F = 9.5, p < 10^{-10}$). In particular, the LR-balance of the forward sweeps (measured as the difference between left and right probabilities), was significantly larger than the mean difference only during the first five laps of the Multiple-T task. The proportion of the reconstruction probability at the location of the rat remained relatively constant, increasing slightly over laps on the Multiple-T task, decreasing slightly over laps on the Cued-choice task (ANOVA, effect of Lap: df$= 38, F = 7.1, p < 10^{-10}$, effect of Task: df$= 1, F = 91.6, p < 10^{-10}$, interaction: df$= 38, F = 8.9, p < 10^{-10}$).

5.3 Discussion

Neural ensembles recorded from the CA3 region of rats running on T-based decision tasks displayed transient activity at decision points indicative of positions different than the rat's current position (leading to non-local reconstruction). Projection of these activities onto spatial location showed that during these decision-making events the location reconstructed from the neural ensemble swept forward of the animal, first down one path and then the other. Reconstructed representations were coherent and preferentially swept ahead of the animal rather than behind the animal. Similar non-local representations were observed as rats recovered from an error. Local field potentials from these sites displayed strong power in theta and gamma frequencies and no sharp waves. Non-local spatial representations were influenced by task demands and experience.

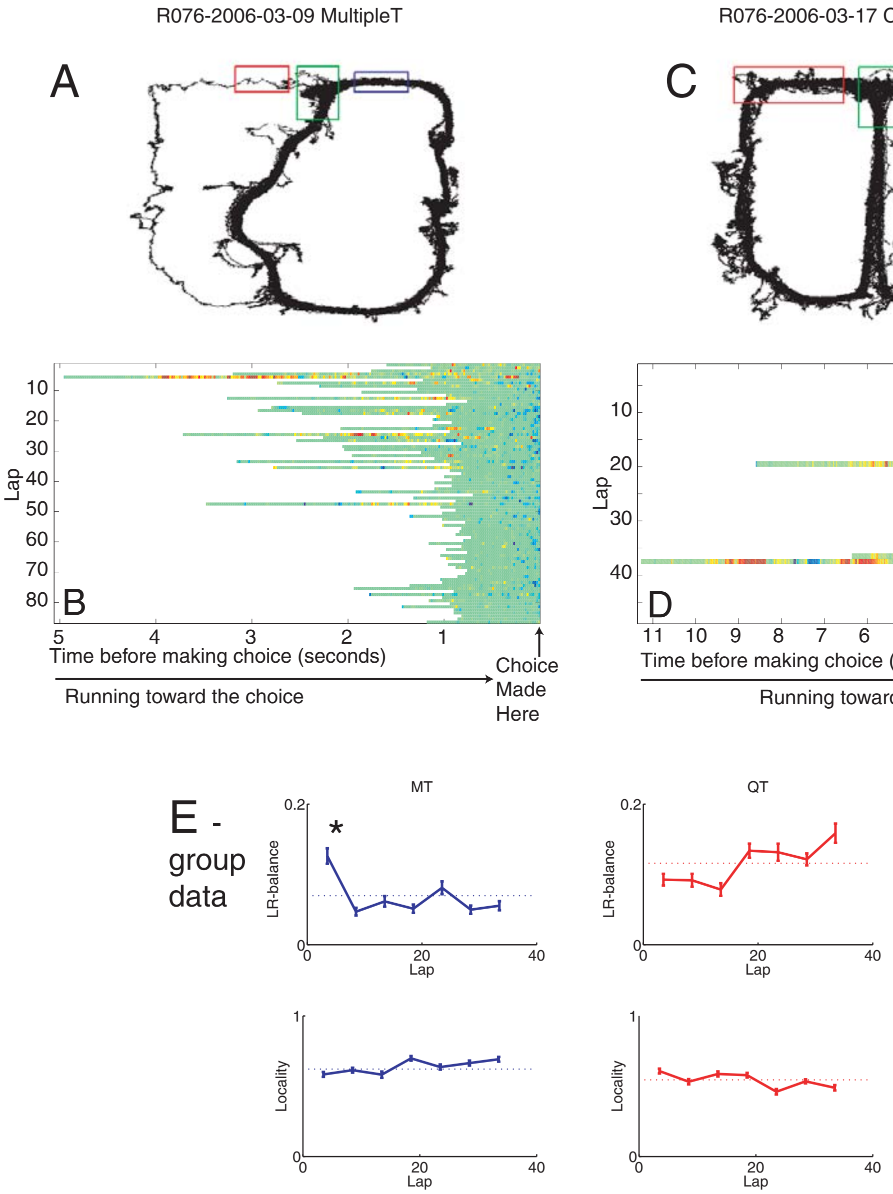
R076-2006-03-09 MultipleT
R076-2006-03-17 Cued-choice
A
C
Lap
10
20
30
40
50
60
70
80
B
5
4
3
2
1
Time before making choice (seconds)
Running toward the choice
Choice Made Here
Lap
10
20
30
40
D
11
10
9
8
7
6
5
4
3
2
1
Time before making choice (seconds)
Running toward the choice
Choice Made Here
E - group data
MT
QT
*
LR-balance
0.2
0
20
40
Lap
Locality
1
0
20
40
Lap

Figure 5.17: Choice point behavior and neural activity changes as a function of task and experience (figure on previous page). Both time in the center stem and non-local activity decreased with experience on the Multiple-T task; no such changes were observed on the Cued-choice task. Plots show spatial neural activity at the choice point (colors) localized to the choice point (green), the left arm (red), or the right arm (blue). Laps increase from top to bottom. Data are presented as time-series until exit from the choice point in 40msec samples. A,B: Multiple-T. C,D Cued-choice task. E: group data. LR-balance was defined as the mean absolute difference between the probabilities in the left and right arm areas for each lap. The LR-balance was significantly increased only on the first five laps of the Multiple-T task. The locality was the mean probability within the central choicepoint area for each lap. See text for details. Figure used with permission (Johnson and Redish, 2007).

5.3.1 Non-local representations.

That extra-field spiking maintained spatial organization across multiple simultaneously recorded cells during extra-field activity implies that the extra-field spiking may coincide with transient, non-local representations of events or locations. In these tasks, all essential information can be projected onto space (i.e. location of the animal). As a result, memory (sequences and episodes) and any decision-related processes present within hippocampal ensemble dynamics can be made observable by examining dynamic changes in the spatial representation. It is critical to note the reconstruction analysis makes no fundamental assumption regarding spatial representation within the hippocampus, only that memory and decision-making signals may be coherently projected onto space in these tasks.

Non-local decoding at the choice point displayed sequential activity associated with either of the T maze arms or the reward locations. Non-local representations and reconstructed locations were found predominantly forward of the animal's position at the choice point and were not temporally coincident with non-local representations behind the animal's position at the choice point or a general non-directional spreading of the representation (see Fig. 5.8). Reconstruction in front of but not behind the animal suggests that the information is related to representation of future paths rather than a replay of recent history. In order to determine whether this effect resulted from a non-specific forward shift or a directionally specific forward shift, we found the joint distribution of probabilities in the left and right arms (see Fig. 5.9). The non-local reconstructions observed here did not occur in both forward directions simultaneously: the representation encoded first one arm of the T and then the other.

5.3.2 Planning paths with cognitive maps.

The cognitive map was originally proposed to explain the flexible spatial navigation abilities observed in animals (O'Keefe and Nadel, 1978). Several researchers (O'Keefe and Speakman, 1987; Lenck-Santini et al., 2002; Rosenzweig et al., 2003) have shown that behavior aligns to the goal locations expected given an active place field distribution, while other researchers (Wood et al., 2000; Frank et al., 2000; Ferbinteanu and Shapiro, 2003; Ainge et al., 2007a; Hok et al., 2007) have shown goal-sensitivity during navigation processes. These previous results suggest that how the animal's current position is represented contributes to navigation, but the mechanism by which the hippocampal representation contributes to navigation remains an open question. The results presented here show that when animals pause during behavior, the hippocampal representation becomes transiently non-local. At high-cost choice points and at the correction of errors, the hippocampal representation sweeps forward along the potential paths available to the animal. These transient non-local signals could provide a potential substrate for the prediction of the consequences of decisions and the planning of paths to goals.

5.3.3 Does the representation have to reach the goal to be useful for planning?

Many of the sweeps proceed only partially down the potential forward paths and evaporate before they reach the goal (see chapter 2).[1] As has been established in the animal learning literature, it is possible for secondary reinforcers to train additional associations (Pavlov, 1927; Domjan, 1998). Both tasks include explicit secondary reinforcers (the click of the feeder [Multiple-T, Figure 5.1] and the secondary tones [Cued-choice, Figure 5.2]) and many of the sweeps do reach those secondary tones. However, it is also possible that the locations on the paths to the goal have themselves become secondary reinforcers. All that would be needed for evaluation of the potential choices would be that the non-local representation reach a point at which the expected value can be evaluated. Thus, a sweep that reaches to a point on the path leading to a goal could be used to appraise the expected reward on taking that path as long as there was value associated with that location (Sutton and Barto, 1998, see chapter 2). The critical test of this hypothesis will be to examine downstream structures during these non-local representations to determine how the non-local information is used.

---

[1] It is important, however, to note that the experiments reported here are based on ensembles of several dozen cells. It is possible that some of the observed sweeps proceed farther down the potential path toward the goal than can be observed with these recordings. It is possible that future recordings with larger ensembles may provide for more definitive statements about the dynamics of the forward sweeping process.

5.3.4 Vicarious trial and error.

Rats faced with difficult decisions vicariously sample the different choices before committing to a decision (Meunzinger, 1938; Hu and Amsel, 1995; Hu et al., 2006, see chapter 1). These "vicarious trial and error" (VTE) behaviors (Meunzinger, 1938; Tolman, 1939; Hu and Amsel, 1995; Hu et al., 2006) or "microchoices" (Brown, 1992; Hudon et al., 2002) appear as small head movements alternating between the potential choices. We observed VTE movements in both the Multiple-T and Cued-choice tasks. We identified VTE-like times quantitatively as times with highly variable orientation of motion. These VTE-like times tended to include more non-local representations than other navigation times (Figure 5.13). While VTE behaviors often appear as fast changes in movement orientation at timescales longer than 500 msec, it is possible that the multiple stages of information processing associated with VTE (representation of non-local information, evaluation and comparison to a criterion, re-orientation or choice) may occur at shorter timescales. VTE movements are diminished with hippocampal lesions (Hu and Amsel, 1995), and are related to hippocampal activity levels on hippocampus-dependent tasks (Hu et al., 2006). VTE has been proposed to provide a choice-consideration mechanism and to underlie complex decision planning (Tolman, 1939).

5.3.5 Phase precession

Both hippocampal interneurons and pyramidal cells fire in specific relationships to local field potentials (Klausberger et al., 2003; Somogyi and Klausberger, 2005). During the theta state, pyramidal cells fire at specific phases of the LFP theta rhythm which then precess through behavior (O'Keefe and Recce, 1993; Skaggs et al., 1996; Dragoi and Buzsaki, 2006). This phase-precession encodes a sweep of spatial representation from behind the animal to in front of the animal in each theta cycle. The sweep events reported here generally lasted approximately a single theta cycle, although some sweeps lasted much longer. Moreover, the distance encoded by our sweeps was often much larger than is typically seen during phase-precession. During sweeps that lasted long than the expected 140 ms (one cycle at 7 Hz), theta oscillations sometimes appeared to degrade transiently (see for example panel A of Figure 5.16), restarting only on the conclusion of the sweep (see panels D,E,G of Figure 5.16). This may suggest that the theta rhythm within the local field potential and the sweep process are both generated from similar processes.

However, the phase relationships predicted from single-cell recordings (Skaggs et al., 1996) suggest that phase-precession sweeps should start at the location of the rat at the peak of theta and end ahead of the rat at the subsequent rising phase of theta. This closely matches our observations (see Figure 5.16). Phase precession has been suggested as a means of planning future trajectories (Jensen and Lisman, 1998, 2005; Yamaguchi, 2003), but has never been studied in situations in which rats must select from multiple available paths. The exact relationship between the forward sweep phenomena reported here and phase precession remains to be fully elucidated, but studying phase precession using reconstructive techniques like those used here is likely to be fruitful.

5.3.6 Conclusion

Multiple learning and memory processes mediate decision-making behavior (Packard and McGaugh, 1996; Poldrack et al., 2001; Daw et al., 2005). The primary distinction between these learning and memory systems can be characterized by a trade-off between a flexible, fast-learning, computationally-costly planning system and an inflexible, slow-learning, computationally-efficient cache system (O'Keefe and Nadel, 1978; Redish, 1999; Daw et al., 2005; Redish and Johnson, 2007). A planning system allows an animal to flexibly compute outcomes resulting from variable task contingencies, a process which requires prospective memory (Daw et al., 2005; Redish and Johnson, 2007). Computational models of prospective memory have suggested that cortical-hippocampal interactions, particularly due to theta-gamma interactions, may allow hippocampal neurons to reactivate previously experienced temporal sequences (Jensen and Lisman, 1998, 2005; Koene et al., 2003; Eichenbaum, 2004). The similarities between these models and the observations of forward shifted neural representations and their coincidence with theta and gamma rhythms at the choice point suggest that the hippocampus may provide prospective memory signals to downstream evaluation and decision making modules.

Chapter 6

Discussion

The simple question presented at the beginning of this thesis was how cognitive maps contribute to the organization of behavior. Cognitive maps were originally proposed to explain five simple behavioral observations as representation based inference rather than simple stimulus-response associations (Tolman, 1948). The original proposal provided only a rough sketch of a theory and raised concerns related to its viability as an explanation for animal behavior (MacCorquodale and Meehl, 1948; Guthrie, 1952; MacCorquodale and Meehl, 1954). Many subsequent experimental findings support the utility of cognitive maps as hypothetical constructs (O'Keefe and Nadel, 1978; Gallistel, 1990; Redish, 1999; Cheng et al., 2007) and suggest that cognitive maps provide complex representation of an animal's (spatial and non-spatial) position within a task (O'Keefe and Nadel, 1978; Gallistel, 1990; Redish, 1999). Although these perspectives on cognitive maps avoid the problem of leaving an animal *buried in thought* (Guthrie, 1935), they provide little insight into the mechanisms that support flexible goal directed behavior and the contribution of cognitive maps to action selection.

Reinforcement learning (Sutton and Barto, 1998) provides a simple theoretical framework for analysis of cognitive maps as a mapping of states to actions. Flexible, goal-directed behavior is supported by the use of a transition model in reinforcement learning algorithms (Sutton and Barto, 1998; Daw et al., 2005; Niv et al., 2006b, see Figure 6.1). The transition model provides a computational definition of the causal texture of an environment (Tolman and Brunswik, 1935) that allows reinforcement learning algorithms to evaluate potential future outcomes and quickly generate patterns of action (Daw et al., 2005; Johnson and Redish, 2005a, see chapters 2 and 3) that reflect the modeled organism's current motivational state (Daw et al., 2005; Niv et al., 2006b). The covert state dynamics suggested by model-based reinforcement learning algorithms parallel recent considerations on the contribution of episodic memory to decision-making (Atance and O'Neill, 2001; Suddendorf and Busby, 2003; Suddendorf and Corballis, 2007; Buckner and Carroll, 2007; Schacter et al., 2007; Zilli and Hasselmo, 2008; Schacter et al., 2008; Buckner et al., 2008) and suggest that cognitive map use is likely supported by representation and evaluation of potential trajectories within a map. The specific content represented by these covert evaluative processes remains an open question (Niv et al., 2006b).

Together these theoretical considerations motivated the development of statistical methods that will allow observation and quantification of representational dynamics within neural activity (see chapter 4). And within this context, it becomes apparent that 'noise' within the activity of neural representations of gross observable experimental variables may support covert information processing related to cognition. This suggests that the basic correspondence between spiking activity and neural representations of an observable variable might continue to hold within 'noise' activity but demands new validation methods (Johnson et al., 2007, see chapter 4). In contrast to previous validations methods based on the correspondence between an observable sensorimotor variable and a prediction derived from neural activity, these new validation methods are based on the correspondence between observed spiking activity and predicted spiking activity derived (generated) from decoded neural representations. Harris et al. (2003), Harris (2005), Jackson and Redish (2007) and Johnson et al. (2007, see chapter 4) have shown that these generative approaches to analysis of spiking data provide greater explanatory power than simple sensorimotor decoding approaches.

Based on these theoretical and statistical considerations and much evidence implicating the hippocampus in cognitive map function, Johnson and Redish (2007, see chapter 5) examined spiking activity within the hippocampus and found evidence of covert information processing in the form of forward-shifted spatial representations at a highly salient choice point within multiple behavioral tasks. These observations provide continued support for the hippocampus as a cognitive map. However, these observations support a perspective on cognitive maps much more closely aligned with Tolman's original theory (Tolman, 1948) than with the minimal spatial form espoused by O'Keefe (O'Keefe and Nadel, 1978; O'Keefe, 1999). Furthermore, this broader perspective easily accommodates the role of the hippocampus in non-spatial memory, including episodic memory (Wallenstein et al., 1998; Eichenbaum et al., 1999; Redish, 1999; Squire et al., 2004), and several recent studies that suggest the hippocampus is required for learning cognitive map content while pure spatial learning occurs independently of the hippocampus (Gaskin et al., 2005; White and Gaskin,

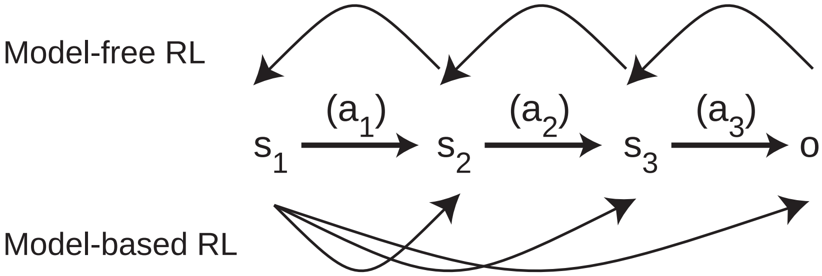

Figure 6.1: A graphical description of the differences in policy development for model-free and model-based reinforcement learning using a sequential learning task. States within the sequence task are denoted by s_i, actions by a_i and the outcome by o. *Top* Outcome information is slowly and progressively mapped backward across states and actions based on experience within model-free reinforcement learning. *Bottom* Outcome information can be covertly accessed and mapped onto future potential states and actions before action selection occurs without true behavioral experience via the transition model within model-based reinforcement learning.

2006; Gaskin and White, 2007; Stouffer and White, 2007).

6.1 Theory

The theoretical considerations of cognitive maps presented in this thesis were developed within the context of reinforcement learning. The reinforcement learning framework was specifically adopted to address the criticism that theoretical treatment of cognitive maps are concerned with what goes on in an animal's mind and neglects to predict what an animal will do (Guthrie, 1935, 1952). The simple premise of reinforcement learning is that animals will structure their (observable and covert cognitive) behavior in order to maximize reward receipt. The five original experiments used to outline cognitive maps can be used to describe a simple computational theory of cognitive map function within the reinforcement learning framework.

- *Short-cut behavior* – Short cut behavior naturally develops within transition model-based reinforcement learning algorithms operating in sufficiently complex (usually spatial) tasks (Sutton and Barto, 1998).
- *Latent learning* – Models of goal directed behavior that are sensitive to motivation-based changes in outcome utility (e.g. latent learning) require use of a transition model within reinforcement learning algorithms (Daw et al., 2005; Niv et al., 2006b).

- *Vicarious trial and error* – Vicarious trial and error suggests that policy evaluation is a sequential update process rather than a massive parallel update process that characterizes standard model-based reinforcement learning algorithms (Sutton and Barto, 1998; Doya, 2000; Daw et al., 2005; Niv et al., 2006b). Furthermore, vicarious trial and error behaviors suggest that this sequential evaluation process can be prompted (or initialized) by specific environmental stimuli (Zilli and Hasselmo, 2008) and should be distinguished from evaluation processes that operate offline (Sutton, 1990; Sutton and Barto, 1998; Johnson and Redish, 2005a).
- *Searching for the stimulus* – The search for a stimulus suggests that policy evaluation entails more than a simple representation of the expected value (reward) for a given state. These search patterns suggest that policies are generated relative to state-based representation of stimulus expectations. The extent and complexity of these expected stimulus representations remains an open theoretical and experimental question.
- *Hypothesis behavior* – The previous experiments explain behavior within a single (albeit motivation-dependent) cognitive map. Hypothesis behavior adds a hierarchical component to theoretical accounts of cognitive maps based on reinforcement learning. Cognitive maps are defined in terms of their content and transitions and fundamentally structure patterns of behavior (Redish et al., 2007). Switching between multiple maps produces important shifts in behavior (Redish, 1999).

In sum, this thesis outlines a reinforcement learning algorithm based approach to cognitive maps and provides specific definitions for much of the opaque terminology used by Tolman (1948) in describing cognitive maps. Although full construction of an algorithm that sufficiently models observed animal behavior on each of these tasks will require many implementation decisions, this approach provides a concrete conceptual framework for understanding cognitive map function.

6.2 Methodological considerations

Neural codes are generally defined as a proposed correspondence between neural activity and an observable variable (Rieke et al., 1997; Dayan and Abbott, 2001). Tuning curves are typically used to describe neural codes and quantify a cell's spiking activity as a function of given variable of interest. If a cell's spiking activity corresponds to the level of variable of interest, the cell is *tuned* to the variable of interest and its activity supports a neural representation of the variable. Tuning curves are usually constructed by averaging neural activity as an animal is presented with the same stimulus or as the animal performs the same behavior over many trials because averaging neural activity across many trials reduces the contribution of nuisance (noise) signals that are independent of the variable of interest. As a result, tuning curves can be used both as a description of observed neural activity and as a prediction of future neural activity given a description of the variable of interest. This suggests that multiple potential neural coding proposals can be distinguished by comparisons of explanatory power in predicting cellular activity (Fenton and Muller, 1998; Jackson and Redish, 2003; Johnson et al., 2007, see chapter 4).

Neural activity in hippocampal pyramidal cells is modulated by the spatial location of an animal within its environment (O'Keefe and Dostrovsky, 1971; O'Keefe, 1976). Place cells primarily spike when an animal is in the cell's place field which suggests that place cell activity can be explained as a spatial code. The spatial hypothesis explicitly predicts that place cell tuning curves (place fields) are stable in geometrically constant environments. However, although place fields can be stable for months in some environments (Thompson and Best, 1990), they can be unstable from day to day (Kentros et al., 2004), between multiple exposures to a single environment on a single day (Barnes et al., 1997), during long exposures to a single environment within a day (Ludvig, 1999), or even across subparts of a task (McNaughton et al., 1983; Eichenbaum et al., 1989; Cohen and Eichenbaum, 1993; Redish, 1999). Moreover, although an animal's position can be used to predict the activity of place cells, place cells fire in a more variable manner than would be predicted from the simple spatial position of the animal, both inside and outside of the place field (Fenton and Muller, 1998; McNaughton et al., 1983; Eichenbaum et al., 1987; Redish, 1999; Kentros et al., 2004; Jackson and Redish, 2007; Johnson and Redish, 2007). These observations suggest that while place cell activity can be described in terms of a spatial map (O'Keefe and Nadel, 1978, 'spatial hypothesis'), other (non-spatial) variables also contribute to place cell activity.

6.2.1 Place cell instability and multiple maps

Place cells can re-map in response to spatial and non-spatial task changes (Bostock et al., 1991; Cohen and Eichenbaum, 1993; Sharp et al., 1995; Redish, 1999). Remapping most often occurs as a result of spatial cue manipulations (Bostock et al., 1991; Tanila et al., 1997; Knierim et al., 1998; Hayman et al., 2003; Leutgeb et al., 2005b; Wills et al., 2005) and corresponds to information theoretic processing of cue likelihood given multiple contexts (Fuhs and Touretzky, 2007). Further experiments have shown that remapping can also occur across sessions without any task or environmental changes in senescent animals (Barnes et al., 1997), in animals made deficient in synaptic plasticity by genetic (Rotenberg et al., 2000; Cho et al., 1998) or pharmacological (Kentros et al., 1998) manipulations. Lesions of perirhinal cortex (Muir and Bilkey, 2001), post-subiculum, or anterodorsal thalamus (Calton et al., 2003) also produce remapping between sessions in a constant environment (but not within a single session).

Remapping can also happen in response to a change in behavioral strategy (Markus et al., 1995), coordinate systems (Gothard et al., 1996a,b; Redish et al., 2000; Rosenzweig et al., 2003) or context learning (Moita et al., 2004). Even in positions common to different trajectories, place cell activity can be strongly influenced by previous or future paths (McNaughton et al., 1983; O'Keefe and Recce, 1993; Gothard et al., 1996a; Wood et al., 2000; Frank et al., 2000; Ferbinteanu and Shapiro, 2003; Ainge et al., 2007a,b; Bower et al., 2005). Place cells also respond to the absence of expected stimuli (O'Keefe, 1976) and to the presence of unexpected but relevant stimuli (Fyhn et al., 2002; Moita et al., 2003). Spatial maps quickly form within novel environments but display significant levels of instability (Wilson and McNaughton, 1993; Frank et al., 2004). Kentros et al. (2004) showed that place fields in wild-type mice display higher levels of instability than place fields in rats tested under identical task conditions. Place cell stability increased in these mice during tasks that require increased attention.

These tuning curve-based observations suggest that hippocampal place fields are formed by experience, stored, and then retrieved (McNaughton and Nadel, 1990; Redish and Touretzky, 1998b; Redish, 1999). The reliable retrieval of a place cell map is dependent on its behavioral utility (Kentros et al., 2004) and is correlated with task performance (O'Keefe and Speakman, 1987; Barnes et al., 1997; Lenck-Santini et al., 2001, 2002; Rosenzweig et al., 2003). These data suggest that explanations of place cell activity that include covert variables related to attention, behavioral relevance, and goals better predict observed spiking activity than purely spatial explanations.

6.2.2 Place cell instability and fast representational dynamics

An interesting variant of the spatial hypothesis of hippocampal function suggests that evidence of map instability might result from coherent spatial representations that deviate from the animal's observed position. This modified hypothesis is consistent with transition model use within model-based reinforcement learning algorithms (see Figure 6.1) and essential for the theoretical treatment of cognitive maps outlined above. This hypothesis predicts that epochs of place cell instability or 'noise' might indicate covert cognitive processing modes that support cognitive map use. Several recent observations support this modified spatial hypothesis and suggest that variability within place cell activity can be explained in terms of fast spatial representation dynamics predicted from the activity of simultaneously recorded neural ensembles (Harris et al., 2003; Jackson and Redish, 2007; Johnson et al., 2007, see chapter 4).

Place cell activity within the place field

Even within simple foraging tasks in which stable place field and spatial maps are observed, place cells display unexpectedly high levels of spiking variability inside their respective place fields (Fenton and Muller, 1998; Olypher et al., 2002). In simple foraging tasks, a robust place cell may emit 20 or more action potentials on a single pass through a place field, but fail to emit even a single action potential seconds later on a pass that is behaviorally indistinguishable. The statistics of these deviations are incompatible with the hypothesis that place cell activity simply varies randomly about a mean described by a single spatial tuning curve (Fenton and Muller, 1998; Olypher et al., 2002) and instead support a hypothesis that place cell activity reflects a small number of spatial tuning curves that differ mainly in rate and are alternatively switched on and off with a period of about a second (Olypher et al., 2002). This proposal is analogous to the suggestion that the hippocampus maintains multiple spatial maps of the environment and somehow switches between those maps very quickly.

Support for this proposal comes from studies by Harris et al. (2003) and Jackson and Redish (2007). Harris et al. (2003) initially showed that predicting the spiking activity of hippocampal place cells using both position and the spiking activities of a set of simultaneously recorded place cells was significantly better than predicting hippocampal place cell activity using position information alone. Harris (2005) argued that the covariation of place cell spiking activity was evidence for use of multiple cell assemblies (Hebb, 1949) within the hippocampus. Jackson and Redish (2007) showed that coherent fast switching between multiple hippocampal cell assemblies could explain the excess variability observed within place cell spiking activity observed by Fenton and Muller (1998). Furthermore, Jackson and Redish (2007) showed that fast switching between cell assemblies was clearly aligned to specific behavioral phases in certain tasks and produced multiple and distinct spatial maps. The cell assemblies observed on the linear track, for instance, were generally aligned with the animal's running direction and their projection into spatial position was apparent as directional place fields. While directional place fields have been previously explained as indicative of multiple reference frames (maps) in linear track tasks (McNaughton et al., 1983; O'Keefe and Recce, 1993; Gothard et al., 1996a; Touretzky and Redish, 1996; Redish and

Touretzky, 1997), Jackson and Redish (2007) show these reference frames are not specific to linear track tasks and are generally present within hippocampal spiking activity (even though they may be less obvious).

The high levels of variability observed within single unit spiking activity of hippocampal place cells inside their place fields can be explained as coordinated activity across the hippocampus (Harris et al., 2003; Jackson and Redish, 2007). These observations further suggest that an unobservable-internal or *covert* variables mediate selection of a cell's active tuning curve and, consequently, determines the discharge of the cell. Although these data can be partially explained as non-cognitive cell assembly dynamics (Harris et al., 2003; Harris, 2005), observations that these cell assembly dynamics are modulated by cognitive demands (Olypher et al., 2002; Kentros et al., 2004) and aligned to specific task components (Jackson and Redish, 2007) suggest that cell assembly dynamics reflect covert cognitive processes.

Place cell activity outside the place field

Despite well-defined spatial tuning, place cells occasionally discharge outside of their firing fields. While this extra-field activity was easily explained as place cell noise during waking behavior, its presence was not so easily explained during sleep. Pavlides and Winson (1989) found that cells with place fields on the arms of an 8-arm maze and were reactivated during subsequent sleep states in the center of the maze. Subsequent studies showed this reactivation during sleep states to be both reliable and coherent. Cell pairs and ensembles that were active during awake behavior are reactivated during subsequent sleep states (Pavlides and Winson, 1989; Wilson and McNaughton, 1994; Kudrimoti et al., 1999; O'Neill et al., 2008) and have been found to maintain their temporal ordering as if experienced sequences of spatial locations were being replayed offline (Skaggs and McNaughton, 1996; Louie and Wilson, 2001; Nádasdy et al., 1999; Lee and Wilson, 2002). These replayed sequences are more likely to occur during sleep following task performance than during sleep preceding task performance (Kudrimoti et al., 1999; O'Neill et al., 2008; Tatsuno et al., 2006).

Further investigations suggest that replay is not limited to sleep states. O'Neill et al. (2006) showed that during pauses in exploratory activity in novel environments previously active place cells could reactivate and cause place cells to fire outside their firing fields. Jackson et al. (2006) showed that ensemble activity at feeder sites reflects the spatiotemporal structure observed within previous behavior and develops with experience. A number of further observations support the existence of forward and backward replay of place cell ensemble firing during periods of awake immobility (Foster and Wilson, 2006; Diba and

Buzsàki, 2007; Csicsvari et al., 2007; Johnson et al., 2007, see chapter 4). The spatiotemporal structure of these spiking sequences, coupled with their experience dependent development, suggests that spikes that occur outside of a place cell's place field might reflect replay of previous experiential episodes, rather than random noise.

Johnson and Redish (2007, chapter 6) studied place cell firing while animals showed vicarious trial-and-error (VTE) behaviors and, presumably, considered the available options at a decision point (Tolman, 1939). Place cells with place fields positioned forward of the animal sequentially fired outside their place fields such that decoding the activity revealed a sequential sweep of positions from the animal's current position to potential future positions on each arm (see chapter 5). The non-local forward representations contained sequential structure, were predominantly ahead of the animal, were related to the orientation of the animal during the VTE behavior and, much like replay activity, suggest that place cell activity that occurs outside of a cell's place field is not simply noise. Instead, this activation appears to signal a covert variable related to the potential choices of the animal, potentially related to the cognitive consideration of possibilities.

6.2.3 Statistical treatment

The idea of a *neural code* suggests a basic correspondence between neural activity and a specific behavioral, environmental or cognitive variable. In theory, a full understanding of a neural code implies that a variable of interest can be predicted from a given set of neural activity *or* that neural activity can be predicted from a given variable of interest. The prediction of a behavioral, environmental or cognitive variable from a given set of neural activity defines the *decoding* approach to neural codes (Georgopoulos et al., 1983, 1988, 1989; Wilson and McNaughton, 1993; Salinas and Abbott, 1994; Rieke et al., 1997; Zhang et al., 1998; Brown et al., 1998; Jensen and Lisman, 2000; Johnson and Redish, 2007). The prediction of neural activity from a given set of behavioral, environmental or cognitive variables defines the *generative* approach to neural codes (Rieke et al., 1997; Harris et al., 2002, 2003; Jackson and Redish, 2007; Johnson et al., 2007, see chapter 4). Using these approaches, different proposed neural codes can be compared in terms of their explanatory power or predictive validity (Johnson et al., 2007). In practice, several salient distinctions must be made between decoding and generative approaches to neural codes and produce fundamentally different strategies for testing the validity of a proposed neural code (Jackson and Redish, 2003; Johnson et al., 2007). These distinctions have important implications for the analysis of covert, cognitive processes.

Validity in decoding approaches to neural codes

Support for the validity of a proposed neural code within the decoding approach is related to how well it predicts an observed behavioral, environmental or cognitive variable from neural activity. The result of the decoding approach is that predictions that deviate from the observed state of the animal (e.g. the observed behavioral, environmental or cognitive variable) are interpreted as errors (Zhang et al., 1998; Brown et al., 1998; Jensen and Lisman, 2000). While average errors can be used to distinguish between multiple decoding algorithms (Salinas and Abbott, 1994; Zhang et al., 1998; Brown et al., 1998), individual error trajectories are thought to be indicative of noise.

Analysis of covert, cognitive processes present a distinct problem for the decoding approach to neural codes because the cognitive variables implied by these processes are not generally experimentally observable. These cognitive processes will likely appear as 'error' trajectories in decoding approaches. To avoid the problem of calling these trajectories 'errors', decoding approaches to covert, cognitive processes have either highlighted unexpected structure within 'error' trajectories or compared hypothesized time-series trajectories indicative of proposed cognitive processing with time-series trajectories of predicted (decoded) variables. The application of the decoding approach to cognitive processing has been successful in a number of studies on population vector rotation (Georgopoulos et al., 1989; Crowe et al., 2005) and route replay (Jensen and Lisman, 2000; Lee and Wilson, 2002; Foster and Wilson, 2006; Diba and Buzsàki, 2007; Csicsvari et al., 2007). However, the decoding approach to cognitive processing provides only a weak form of statistical validity because it depends on how hypothesized cognitive time-series trajectories are defined. As a result, decoding approaches to covert cognitive processes are usually reducible to an extended cross correlation analysis that is informed by each cell's tuning curve (Georgopoulos et al., 1989; Skaggs and McNaughton, 1996; Nádasdy et al., 1999; Lee and Wilson, 2002; Crowe et al., 2005; Jackson et al., 2006; Euston and McNaughton, 2007).

Validity in generative approaches to neural codes

Support for the validity of a proposed neural code within the generative approach is related to how well it predicts an observed set of neural activity from behavioral, environmental or cognitive variables (Rieke et al., 1997; Harris et al., 2002, 2003; Jackson and Redish, 2007; Johnson et al., 2007, see chapter 4). At a superficial level, the application of the generative approach to covert, cognitive processes presents the same problem encountered within decoding approaches (e.g. that a hypothetical time-series trajectory of cognitive variable must be proposed). However, several applications of the generative approach have circumvented this problem by inferring covert variable trajectories on the basis of ensemble activity (Harris et al., 2003; Jackson and Redish, 2007; Johnson et al., 2007, see chapter 4). These variants of the generative approach either implicitly (Harris et al., 2003) or explicitly (Jackson and Redish, 2007; Johnson et al., 2007) use decoding algorithms for inference of covert variable trajectories. Harris et al. (2003) predicted the spiking activity of a single place cell using the animal's position and the place cell activity of a small simultaneously recorded neural ensemble. Jackson and Redish (2007) predicted the spiking activity of a single place cell using the animal's position and a reference frame identified by the activity of simultaneously recorded place cells. Johnson et al. (2007, see chapter 4) showed that decoded spatial representations that sometimes deviated from an animal's observed spatial location better predicted the observed activity of a simultaneously recorded set of place cells than the animal's observed position and used these observations as the basis of subsequent analysis of non-local spatial representations (Johnson and Redish, 2007, see chapter 5).

6.2.4 Summary

The spatial information in hippocampal place fields provides an example of measurable tuning to a cognitive, yet directly observable variable. The observability of the place fields spatial tuning provides an opportunity to assess the active components of hippocampal processing. Place cell activity is much more variable than would be expected given a simple spatial hypothesis. Several observations suggest that higher than expected levels of spiking variability are organized with respect to covert, potentially cognitive, processes rather than noise (Pavlides and Winson, 1989; O'Keefe and Recce, 1993; Wilson and McNaughton, 1994; Olypher et al., 2002; Harris et al., 2003; Jackson and Redish, 2007; Johnson and Redish, 2007). Although decoding approaches to neural codes can be used in explanations of observed place cell activity, their use for analysis of covert, cognitive processes presents several statistical problems. In contrast, generative approaches to neural codes provide a powerful and statistically well founded explanation of hippocampal place cell activity (Johnson et al., 2007, see chapter 4). This generative approach has provided the basis for the experimental findings presented in this thesis. These and other findings reviewed above strongly suggest that hippocampal activity does not merely reflect the physical space an animal inhabits. Instead hippocampal activity is more parsimoniously described as an active construction of an internal, dynamic interpretation of subjective information extracted from the animals environment, dependent on cognitive variables such as attention and the behavioral relevance of the spatial cues.

6.3 Conclusions and open questions

The hippocampus forms a critical part of a neural system that supports spatial behavior, cognitive map function and episodic memory (O'Keefe and Nadel, 1978; Squire et al., 2004). Although hippocampal pyramidal cell activity can be simply explained in terms of the animal's environmental position (O'Keefe and Dostrovsky, 1971; O'Keefe, 1976; O'Keefe and Nadel, 1978), what elements of spatial experience are represented by this activity remains an open question. Purely spatial map-based behavior appears to be independent of the hippocampus and, instead, depends on a circuit that involves the entorhinal cortex and the fornix (Gaskin et al., 2005; Gaskin and White, 2007). In contrast, the hippocampus appears to be critical for encoding stimulus information within a spatial map supported by the entorhinal cortex and subsequent retrieval of this stimulus information (White and Gaskin, 2006; Stouffer and White, 2007). These findings suggest that place cell activity likely represents integrated stimulus information associated with place field position (though this stimulus information itself may be stored outside the hippocampus; Teyler and DiScenna, 1986).

Findings from several recent studies show that hippocampal neural activity is dynamically modulated by covert cognitive processes (Olypher et al., 2002; Jackson and Redish, 2007; Johnson and Redish, 2007, see chapters 4 and 5). Hippocampal spatial representations shift ahead of the animal's current position at choice points and suggest that an animal might be imagining potential future positions on the track (Johnson and Redish, 2007). This interpretation of hippocampal place cell activity is consistent with recent considerations of the contribution episodic memory makes toward decision-making processes (Suddendorf and Corballis, 1997; Atance and O'Neill, 2001; Suddendorf and Busby, 2003; Hassabis et al., 2007; Buckner and Carroll, 2007; Schacter et al., 2007; Suddendorf and Corballis, 2007; Buckner et al., 2008; Schacter et al., 2008). In humans, the hippocampus is critical for imagining coherent, integrated spatial scenes (Hassabis et al., 2007). Episodic memory based imagination activates a frontotemporal network that includes the hippocampus (Buckner and Carroll, 2007; Schacter et al., 2007; Suddendorf and Corballis, 2007; Buckner et al., 2008; Schacter et al., 2008). Although cell assemblies in the hippocampus are coordinated with activity beyond the hippocampus (McNaughton et al., 1996; Qin et al., 1997; Hoffman and McNaughton, 2002; Pennartz et al., 2004; Battaglia et al., 2004; Tatsuno et al., 2006; Euston and McNaughton, 2007; Fyhn et al., 2007; Ji and Wilson, 2007), understanding the contribution forward shifted spatial representations in the hippocampus make to behavior and their link to episodic memory and spatial imagination will require much future research.

A theoretical account of cognitive maps, spatial imagination and episodic memory will require development of new models and theory. While clear parallels between the state dynamics of model-based reinforcement learning algorithms and spatial imagination and episodic memory have been observed (Daw et al., 2005; Johnson and Redish, 2005a; Niv et al., 2006b; Zilli and Hasselmo, 2008), these parallels are only a loose approximation and provide a important trajectory for future work. One particularly interesting open question is whether transition model use in the model by Niv et al. (2006b) implies representation of a complete set of stimulus information at each state – something that would closely resemble episodic memory.

Finally, the complexity of episodic memory and spatial imagination suggests a parameter space too large to be explored through experiments alone. Future understanding of episodic memory and spatial imagination will come from coordinated interactions between theory and experiment. Theory supports identification of information rich portions of this cognition and memory parameter space and can be used for the development of deeply informative experiments. In turn, experimental results provide important theoretical constraints and provide increasing clarity in the boundaries of these complex hypothetical constructs (MacCorquodale and Meehl, 1948). This cycle of theory and experiment provides a clear path for understanding spatial imagination, episodic memory and cognitive maps.

Bibliography

Adams CD, Dickinson A (1981) Instrumental responding following reinforcer devaluation. Quarterly Journal of Experimental Psychology: Comparative and Physiological Psychology, 33(B):109–122.

Addis DR, Wong AT, Schacter DL (2007) Remembering the past and imagining the future: common and distinct neural substrates during event construction and elaboration. Neuropsychologia 45(7):1363–1377.

Agster KL, Fortin NJ, Eichenbaum H (2002) The hippocampus and disambiguation of overlapping sequences. Journal of Neuroscience 22(13):5760–5768.

Ainge JA, Tamosiunaite M, Woergoetter F, Dudchenko PA (2007a) Hippocampal CA1 place cells encode intended destination on a maze with multiple choice points. Journal of Neuroscience 27(36):9769–9779.

Ainge JA, van der Meer MAA, Langston RF, Wood ER (2007b) Exploring the role of context-dependent hippocampal activity in spatial alternation behavior. Hippocampus 17(10):988–1002.

Arolfo MP, Nerad L, Schenk F, Bures J (1994) Absence of snapshot memory of the target view interferes with place navigation learning by rats in the water maze. Behavioral Neuroscience 108(2):308–316.

Atance CM, O'Neill DK (2001) Episodic future thinking. Trends in Cognitive Sciences 5(12):533–539.

Averbeck BB (2001) Neural mechanisms of copying Geometrical Shapes. Ph.D. thesis, University of Minnesota.

Averbeck BB, Chafee MV, Crowe DA, Georgopoulos AP (2002) Parallel processing of serial movements in prefrontal cortex. Proc Natl Acad Sci U S A 99(20):13172–13177.

Averbeck BB, Chafee MV, Crowe DA, Georgopoulos AP (2003a) Neural activity in prefrontal cortex during copying geometrical shapes I. Single cells encode shape, sequence, and metric parameters. Experimental Brain Research 150:127–141.

Averbeck BB, Crowe DA, Chafee MV, Georgopoulos AP (2003b) Neural activity in prefrontal cortex during copying geometrical shapes II. Decoding shape segments from neural ensembles. Experimental Brain Research 150:142–153.

Averbeck BB, Latham PE, Pouget A (2006) Neural correlations, population coding and computation. Nature Reviews Neuroscience 7:358–366.

Baddeley AD (1997) Human memory: Theory and practice. Psychology Press.

Balleine BW, Dickinson A (1998) Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. Neuropharmacology 37(4-5):407–419.

Barnes CA (1979) Memory deficits associated with senscence: A neurophysiological and behavioral study in the rat. Journal of Comparative and Physiological Psychology 93:74–104.

Barnes CA (1988) Spatial learning and memory processes: The search for their neurobiological mechanisms in the rat. Trends in Neurosciences 11:163–169.

Barnes CA, Suster MS, Shen J, McNaughton BL (1997) Multistability of cognitive maps in the hippocampus of old rats. Nature 388(6639):272–275.

Batschelet E (1981) Circular statistics in biology. New York: Academic Press.

Battaglia FP, Sutherland GR, McNaughton BL (2004) Hippocampal sharp wave bursts coincide with neocortical "up-state" transitions. Learn Mem 11(6):697–704.

Battaglia PW, Schrater PR (2007) Humans trade off viewing time and movement duration to improve visuomotor accuracy in a fast reaching task. Journal of Neuroscience 27(26):6984–6994.

Bayer HM, Glimcher P (2005) Midbrain dopamine neurons encode a quantitative reward prediction error signal. Neuron 47:129–141.

Bear MF (1996) A synaptic basis for memory storage in the cerebral cortex. Proc Natl Acad Sci U S A 93(24):13453–13459.

Bellman R (1957) Dynamic Programming. Princeton, N.J.: Princton University Press.

Berridge KC (2006) The debate over dopamine's role in reward: the case for incentive salience. Psychopharmacology 191(3):391–431.

Berridge KC, Robinson TE (1998) What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? Brain Research Reviews 28:309–369.

Beylin AV, Gandhi CC, Wood GE, Talk AC, Matzel LD, Shors TJ (2001) The role of the hippocampus in trace conditioning: Temporal discontinuity or task difficulty? Neurobiology of Learning and Memory 76(3):447–461.

Bi G, Poo M (2001) Synaptic modification by correlated activity: Hebb's postulate revisited. Annual Review of Neuroscience 24(1):139–166.

Bialek W, Rieke F, de Ruyter van Steveninck RR, Warland D (1991) Reading a neural code. Science 252(5014):1854–1857.

Blum KI, Abbott LF (1996) A model of spatial map formation in the hippocampus of the rat. Neural Computation 8(1):85–93.

Bolles RC (1972) Reinforcement, expectancy, and learning. Psychological Review 79(5):394–409.

Bostock E, Muller RU, Kubie JL (1991) Experience-dependent modifications of hippocampal place cell firing. Hippocampus 1(2):193–206.

Bouton ME (2002) Context, ambiguity, and unlearning: sources of relapse after behavioral extinction. Biological Psychiatry 52:976–986.

Bouton ME, Westbrook RF, Corcoran KA, Maren S (2006) Contextual and temporal modulation of extinction: Behavioral and biological mechanisms. Biological Psychiatry 60(4):352–360.

Bower MR, Euston DR, McNaughton BL (2005) Sequential-context-dependent hippocampal activity is not necessary to learn sequences with repeated elements. Journal of Neuroscience 25(6):1313–1323.

Braver TS, Barch DM, Kelley WM, Buckner RL, Cohen NJ, Miezin FM, Snyder AZ, Ollinger JM, Akbudak E, Conturo TE, Petersen SE (2001) Direct comparison of prefrontal cortex regions engaged by working and long-term memory tasks. Neuroimage 14(1 Pt 1):48–59.

Brockwell AE, Rojas AL, Kass RE (2004) Recursive bayesian decoding of motor cortical signals by particle filtering. Journal of Neurophysiology 91(4):1899–1907.

Brown EN, Frank LM, Tang D, Quirk MC, Wilson MA (1998) A statistical paradigm for neural spike train decoding applied to position prediction from ensemble firing patterns of rat hippocampal place cells. Journal of Neuroscience 18(18):7411–7425.

Brown MF (1992) Does a cognitive map guide chioces in the radial-arm maze? Journal of Experimental Psychology 18(1):56–66.

Buckner RL, Andrews-Hanna JR, Schacter DL (2008) The brain's default network: Anatomy, function, and relevance to disease. Ann N Y Acad Sci 1124:1–38.

Buckner RL, Carroll DC (2007) Self-projection and the brain. Trends Cogn Sci 11(2):49–57.

Buzsáki G (1989) Two-stage model of memory trace formation: A role for "noisy" brain states. Neuroscience 31(3):551–570.

Buzsáki G (2006) Rhythms of the Brain. Oxford.

Buzsáki G, Leung LW, Vanderwolf CH (1983) Cellular bases of hippocampal EEG in the behaving rat. Brain Research 287(2):139–171.

Calabresi P, Picconi B, Tozzi A, Di Fillipo M (2007) Dopamine-mediated regulation of corticostriatal synaptic plasticity. Trends in Neurosciences 30:211–219.

Calton JL, Stackman RW, Goodridge JP, Archey WB, Dudchenko PA, Taube JS (2003) Hippocampal place cell instability after lesions of the head direction cell network. Journal of Neuroscience 23(30):9719–9731.

Capaldi EJ (1957) The effect of different amounts of alternating partial reinforcement on resistance to extinction. Am J Psychol 70(3):451–452.

Carr H, Watson JB (1908) Orientation in the white rat. Journal of Comparative Neurology and Psychology 18:27–44.

Cheng K, Shettleworth SJ, Huttenlocher J, Rieser JJ (2007) Bayesian integration of spatial information. Psychological Bulletin 133(4):625–637.

Cho YH, Giese KP, Tanila H, Silva AJ, Eichenbaum H (1998) Abnormal hippocampal spatial representations in alphacamkiit286a and crebalphadelta- mice. Science 279(5352):867–869.

Clayton NS, Bussey TJ, Emery NJ, Dickinson A (2003) Prometheus to proust: the case for behavioural criteria for 'mental time travel'. Trends Cogn Sci 7(10):436–7; author reply 437–8.

Clayton NS, Dickinson A (1998) Episodic-like memory during cache recovery by scrub jays. Nature 395(6699):272–274.

Cohen NJ, Eichenbaum H (1993) Memory, Amnesia, and the Hippocampal System. Cambridge, MA: MIT Press.

Cohen NJ, Squire LR (1980) Preserved learning and retention of pattern-analyzing skill in amnesia: Dissociation of knowing how and knowing that. Science 210:207–210.

Collett T, Land M (1975) Visual spatial memory in a hoverfly. Journal of Comparative Physiology 100:59–84.

Collett TS (1987) The use of visual landmarks by gerbils: Reaching a goal when landmarks are displaced. Journal of Comparative Physiology A 160:109–113.

Corbit LH, Balleine BW (2000) The role of the hippocampus in instrumental conditioning. Journal of Neuroscience 20(11):4233–4239.

Corbit LH, Balleine BW (2003a) Instrumental and pavlovian incentive processes have dissociable effects on components of a heterogeneous instrumental chain. Journal of Experimental Psychology: Animal Behavioral Processes 29(2):99–106.

Corbit LH, Balleine BW (2003b) The role of prelimbic cortex in instrumental conditioning. Behav Brain Res 146(1-2):145–157.

Corbit LH, Ostlund SB, Balleine BW (2002) Sensitivity to Instrumental Contingency Degradation Is Mediated by the Entorhinal Cortex and Its Efferents via the Dorsal Hippocampus. J Neurosci 22(24):10976–10984.

Corkin S (2002) What's new with the amnesiac patient h.m.? Nature Reviews Neuroscience 3(2):153–160.

Courville AC, Daw ND, Touretzky DS (2006) Bayesian theories of conditioning in a changing world. Trends in the Cognitive Sciences 10(7):294–300.

Coutureau E, Killcross S (2003) Inactivation of the infralimbic prefrontal cortex reinstates goal-directed responding in overtrained rats. Behavioural Brain Research 146:167–174.

Crowe D, Averbeck B, Chafee M, Georgopoulos A (2005) Dynamics of parietal neural activity during spatial cognitive processing. Neuron 47(6):1393–1413.

Csicsvari J, Hirase H, Czurk*ó* A, Mamiya A, Buzs*á*ki G (1999) Fast network oscillations in the hippocampal ca1 region of the behaving rat. Journal of Neuroscience 19(16):RC20.

Csicsvari J, Hirase H, Mamiya A, Buzsáki G (2000) Ensemble patterns of hippocampal CA3-CA1 neurons during sharp wave-associated population events. Neuron 28:585–594.

Csicsvari J, O'Neill J, Allen K, Senior T (2007) Place-selective firing contributes to the reverse-order reactivation of ca1 pyramidal cells during sharp waves in open-field exploration. European Journal of Neuroscience 26(3):704–716.

Dalley JW, L??ne K, Theobald DEH, Armstrong HC, Corlett PR, Chudasama Y, Robbins TW (2005) Time-limited modulation of appetitive pavlovian memory by d1 and nmda receptors in the nucleus accumbens. Proc Natl Acad Sci U S A 102(17):6189–6194.

Daw ND (2003) Reinforcement learning models of the dopamine system and their behavioral implications. Ph.D. thesis, Carnegie Mellon University.

Daw ND, Niv Y, Dayan P (2005) Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. Nature Neuroscience 8(12):1704–1711.

Dayan P, Abbott LF (2001) Theoretical Neuroscience. MIT Press.

Dayan P, Balleine BW (2002) Reward, motivation, and reinforcement learning. Neuron 36(2):285–298.

Dearden R, Friedman N, Russel S (1998) Bayesian q-learning. In: Proceedings of the 15th National Conference on Artificial Intelligence (AAAI), pp. 761–768.

Devenport LD, Devenport JA, Holloway FA (1981) Reward-induced stereotypy: modulation by the hippocampus. Science 212(4500):1288–1289.

Devenport LD, Holloway FA (1980) The rat's resistance to superstition: role of the hippocampus. J Comp Physiol Psychol 94(4):691–705.

Diba K, Buzsàki G (2007) Forward and reverse hippocampal place-cell sequences during ripples. Nature Neuroscience 10(10):1241–1242.

Dickinson A, Balleine BW (2000) The evolution of cognition, chap. Causal cognition and goal-directed action, pp. 185–204. Cambridge, MA: MIT Press.

Doboli S, Minai AA, Best PJ (2000) Latent attractors: a model for context-dependent place representations in the hippocampus. Neural Computation 12(5):1009–1043.

Domjan M (1998) The principles of learning and behavior. Brooks/Cole, fourth edition edn.

Doya K (2000) Reinforcement learning in continuous time and space. Neural computation 12:219–245.

Dragoi G, Buzsaki G (2006) Temporal encoding of place sequences by hippocampal cell assemblies. Neuron 50(1):145–157.

Eacott MJ, Norman G (2004) Integrated memory for object, place, and context in rats: A possible model of episodic-like memory? J Neurosci 24(8):1948–1953.

Eichenbaum H (2004) Hippocampus: Cognitive processes and neural representations that underlie declarative memory. Neuron 44:109–120.

Eichenbaum H, Dudchenko P, Wood E, Shapiro M, Tanila H (1999) The hippocampus, memory, and place cells: Is it spatial memory or a memory space? Neuron 23(2):209–226.

Eichenbaum H, Kuperstein M, Fagan A, 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, Morris RGM (1990) Hippocampal representation in place learning. Journal of Neuroscience 10(11):3531–3542.

Eichenbaum H, Wiener SI, Shapiro ML, Cohen NJ (1989) The organization of spatial coding in the hippocampus: A study of neural ensemble activity. Journal of Neuroscience 9(8):2764–2775.

Ekstrom AD, Kahana MJ, Caplan JB, Fields TA, Isham EA, Newman EL, Fried I (2003) Cellular networks underlying human spatial navigation. Nature 425(6954).

Ekstrom AD, Meltzer J, McNaughton BL, Barnes CA (2001) NMDA receptor antagonism blocks experience-dependent expansion of hippocampal “place fields”. Neuron 31:631–638.

Ellard C, Goodale M, Timney B (1984) Distance estimation in the mongolian gerbil: The role of dynamic depth cues. Behavioual Brain Research 14:29–39.

Emery NJ, Clayton NS (2004) The mentality of crows: Convergent evolution of intelligence in corvids and apes. Science 306(5703):1903–1907.

Etienne AS, Maurer R, Séguinot V (1996) Path integration in mammals and its interaction with visual landmarks. Journal of Experimental Biology 199:201–209.

Euston DR, McNaughton BL (2007) Apparent encoding of sequential context in rat medial prefrontal cortex is accounted for by behavioral variability. Journal of Neuroscience 26:13143–13155.

Fenton AA, Muller RU (1998) Place cell discharge is extremely variable during individual passes of the rat through the firing field. Proceedings of the National Academy of Sciences, USA 95:3182–3187.

Ferbinteanu J, Kennedy PJ, Shapiro ML (2006) Episodic memory — from brain to mind. Hippocampus 16(9):704–715.

Ferbinteanu J, Shapiro ML (2003) Prospective and retrospective memory coding in the hippocampus. Neuron 40(6):1227–1239.

Floresco SB, Seamans JK, Phillips AG (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.

Foster DJ, Morris RGM, Dayan P (2000) A model of hippocampally dependent navigation using the temporal difference learning rule. Hippocampus 10:1–6.

Foster DJ, Wilson MA (2006) Reverse replay of behavioural sequences in hippocampal place cells during the awake state. Nature 440(7084):680–683.

Frank LM, Brown EN, Wilson M (2000) Trajectory encoding in the hippocampus and entorhinal cortex. Neuron 27(1):169–178.

Frank LM, Eden UT, Solo V, Wilson MA, Brown EN (2002) Contrasting patterns of receptive field plasticity in the hippocampus and the entorhinal cortex: An adaptive filtering approach. Journal of Neuroscience 22(9):3817–3830.

Frank LM, Stanley GB, Brown EN (2004) Hippocampal plasticity across multiple days of exposure to novel environments. J Neurosci 24(35):7681–7689.

Fuhs MC, Touretzky DS (2007) Context learning in the rodent hippocampus. Neural Comput 19(12):3173–3215.

Fuhs MC, VanRhoads SR, Casale AE, McNaughton B, Touretzky DS (2005) Influence of Path Integration Versus Environmental Orientation on Place Cell Remapping Between Visually Identical Environments. J Neurophysiol 94(4):2603–2616.

Fyhn M, Hafting T, Treves A, Moser MB, Moser EI (2007) Hippocampal remapping and grid realignment in entorhinal cortex. Nature 446(7132):190–194.

Fyhn M, Molden S, Hollup S, Moser MB, Moser E (2002) Hippocampal neurons responding to first-time dislocation of a target object. Neuron 35(3):555–566.

Gallistel CR (1990) The Organization of Learning. Cambridge, MA: MIT Press.

Gallistel CR, Fairhurst S, Balsam P (2004) Inaugural article: The learning curve: Implications of a quantitative analysis. PNAS 101(36):13124–13131.

Gaskin S, Chai SC, White NM (2005) Inactivation of the dorsal hippocampus does not affect learning during exploration of a novel environment. Hippocampus 15(8):1085–1093.

Gaskin S, White NM (2007) Unreinforced spatial (latent) learning is mediated by a circuit that includes dorsal entorhinal cortex and fimbria fornix. Hippocampus 17(7):586–594.

Georgopoulos AP, Caminiti R, Kalaska JF, Massey JT (1983) Spatial coding of movement: A hypothesis concerning the coding of movement direction by motor cortical populations. Experimental Brain Research Suppl.(7):327–336.

Georgopoulos AP, Kettner RE, Schwartz AB (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.

Georgopoulos AP, Lurito JT, Petrides M, Schwartz AB, Massey JT (1989) Mental rotation of the neuronal population vector. Science 243:234–236.

Gothard KM, Skaggs WE, McNaughton BL (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 KM, Skaggs WE, Moore KM, McNaughton BL (1996b) Binding of hippocampal ca1 neural activity to multiple reference frames in a landmark-based navigation task. Journal of Neuroscience 16(2):823–835.

Grace AA, Rosenkranz JA (2002) Regulation of conditioned responses of basolateral amygdala neurons. Physiol Behav 77(4-5):489–493.

Gray J (2004) Consciousness: Creeping up on the Hard Problem. Oxford.

Griffin A, Eichenbaum H, Hasselmo M (2007) Spatial representations of hippocampal CA1 neurons are modulated by behavioral context in a hippocampus-dependent memory task. Journal of Neuroscience 27(9):2416–2423.

Griffiths TL, Tenenbaum JB (2007) From mere coincidences to meaningful discoveries. Cognition 103(2):180–226.

Gurney K, Prescott TJ, Wickens JR, Redgrave P (2004) Computational models of the basal ganglia: from robots to membranes. Trends in Neurosciences 27(8):453–459.

Guthrie ER (1935) The Psychology of Learning. New York: Harpers.

Guthrie ER (1952) The Psychology of Learning. New York: Harper.

Guzowski JF, McNaughton BL, Barnes CA, Worley PF (1999) Environment-specific expression of the immediate-early gene *ARC* in hippocampal neuronal ensembles. Nature Neuroscience 2(12):1120–1124.

Haddon JE, Killcross AS (2005) Medial prefrontal cortex lesions abolish contextual control of competing responses. J Exp Anal Behav 84(3):485–504.

Hampson RE, Deadwyler SA (1996) Ensemble codes involving hippocampal neurons are at risk during delayed performance tests. Proceedings of the National Academy of Sciences, USA 93:13487–13493.

Harris KD (2005) Neural signatures of cell assembly organization. Nature Reviews Neuroscience 6:399–407.

Harris KD, Csicsvari J, Hirase H, Dragoi G, Buzsaki G (2003) Organization of cell assemblies in the hippocampus. Nature 424(6948):552–556.

Harris KD, Henzi DA, Hirase H, Leinekugel X, Dragoi G, Czurko A, Buzsaki G (2002) Spike train dynamics predicts theta-related phase precession in hippocampal pyramdidal cells. Nature 417(6890):738–741.

Hassabis D, Kumaran D, Vann SD, Maguire EA (2007) Patients with hippocampal amnesia cannot imagine new experiences. Proc Natl Acad Sci U S A 104(5):1726–1731.

Hasselmo ME (1993) Acetylcholine and learning in a cortical associative memory. Neural Computation 5:32–44.

Hasselmo ME, Bower JM (1993) Acetylcholine and memory. Trends in Neurosciences 16(6):218–222.

Hayman RMA, Chakraborty S, Anderson MI, Jeffery KJ (2003) Context-specific acquisition of location discrimination by hippocampal place cells. Eur Journal of Neuroscience 18(10):2825–2834.

Hebb DO (1949) The Organization of Behavior. New York: Wiley. Reissued 2002 LEA.

Hertz J, Krogh A, Palmer RG (1991) Introduction to the Theory of Neural Computation. Reading MA: Addison-Wesley.

Hilgard ER, Bower GH (1975) Theories of Learning. Appleton-Century-Crofts.

Hoffman KL, McNaughton BL (2002) Coordinated reactivation of distributed memory traces in primate neocortex. Science 297(5589):2070–2073.

Hok V, Lenck-Santini PP, Roux S, Save E, Muller RU, Poucet B (2007) Goal-related activity in hippocampal place cells. Journal of Neuroscience 27(3):472–482.

Holland P (2004) Relations between pavlovian-instrumental transfer and reinfocer devaluation. Journal of Experimental Psychology: Animal Behavior Processes 30:104–117.

Holland P, Straub J (1979) Differential effects of two ways of devaluing the unconditioned stimulus after pavlovian appetitive conditioning. J Exp Psychol Anim Behav Process 5:65–78.

Hollup SA, Molden S, Donnett JG, Moser MB, Moser EI (2001a) Accumulation of hippocampal place fields at the goal location in an annular watermaze task. Journal of Neuroscience 21(5):1635–1644.

Hollup SA, Molden S, Donnett JG, Moser MB, Moser EI (2001b) Place fields of rat hippocampal pyramidal cells and spatial learning in the watermaze. European Journal of Neuroscience 13(6):1197–1197.

Hopfield JJ (1982) Neural networks and physical systems with emergent collective computational abilities. Proceedings of the National Academy of Sciences, USA 79:2554–2558.

Houk JC, Davis JL, Beiser DG, eds. (1995) Models of Information Processing in the Basal Ganglia. Cambridge MA: MIT Press.

Hu D, Amsel A (1995) A simple test of the vicarious trial-and-error hypothesis of hippocampal function. Proc Nat Acad Sci, USA 92(12):5506–5509.

Hu D, Xu X, Gonzalez-Lima F (2006) Vicarious trial-and-error behavior and hippocampal cytochrome oxidase activity during Y-maze discrimination learning in the rat. International Journal of Neuroscience 116(3):265–280.

Hudon C, Dor'e FY, Goulet S (2002) Spatial memory and choice behavior in the radial arm maze after fornix transection. Progress in Neuro-Psychopharmacology and Biological Psychiatry 26(6):1113–1123.

Hull CL (1930) Knowledge and purpose as habit mechanisms. Psychological Review 37(6):511–525.

Hull CL (1943) Principles of Behavior. New York: Appleton-Century-Crofts.

Huxter J, Burgess N, O'Keefe J (2003) Independent rate and temporal coding in hippocampal pyramidal cells. Nature 425(6960):828–832.

Jackson J (2006) Network Consistency and Hippocampal Dynamics: Using the properties of cell assemblies to probe the hippocampal representation of space. Ph.D. thesis, University of Minnesota.

Jackson J, Redish AD (2007) Network dynamics of hippocampal cell-assemblies resemble multiple spatial maps within single tasks. Hippocampus .

Jackson JC, Johnson A, Redish AD (2006) Hippocampal sharp waves and reactivation during awake states depend on repeated sequential experience. Journal of Neuroscience 26(48):12415–12426.

Jackson JC, Redish AD (2003) Detecting dynamical changes within a simulated neural ensemble using a measure of representational quality. Network: Computation in Neural Systems 14:629–645.

Janabi-Sharifi F, Hayward V, Chen CSJ (2000) Discrete-time adaptive windowing for velocity estimation. IEEE Transactions on Control Systems Technology 8(6):1003–1009.

Jarrard LE (1993) On the role of the hippocampus in learning and memory in the rat. Behavioral Neural Biology 60(1):9–26.

Jaynes ET (2003) Probability Theory. Cambridge.

Jensen O, Lisman JE (1998) An oscillatory short-term memory buffer model can account for data on the Sternberg task. Journal of Neuroscience 18(24):10688–10699.

Jensen O, Lisman JE (2000) Position reconstruction from an ensemble of hippocampal place cells: contribution of theta phase encoding. Journal of Neurophysiology 83(5):2602–2609.

Jensen O, Lisman JE (2005) Hippocampal sequence-encoding driven by a cortical multi-item working memory buffer. Trends in Neurosciences 28(2):67–72.

Ji D, Wilson MA (2007) Coordinated memory replay in the visual cortex and hippocampus during sleep. Nature Neuroscience 10(1):100–107.

Jog MS, Kubota Y, Connolly CI, Hillegaart V, Graybiel AM (1999) Building neural representations of habits. Science 286:1746–1749.

Johnson A, Jackson J, Redish A (2007) Measuring distributed properties of neural representation beyond the decoding of local variables - implications for cognition. In: Mechanisms of information processing in the Brain: Encoding of information in neural populations and networks (Holscher C, Munk, eds.), p. in press. Cambridge University Press.

Johnson A, Redish AD (2005a) Hippocampal replay contributes to within session learning in a temporal difference reinforcement learning model. Neural Networks 18(9):1163–1171.

Johnson A, Redish AD (2005b) Observation of transient neural dynamics in the rodent hippocampus during behavior of a sequential decision task using predictive filter methods. Acta Neurobiologiae Experimentalis 65(Supplement).

Johnson A, Redish AD (2007) Neural ensembles in ca3 transiently encode paths forward of the animal at a decision point. Journal of Neuroscience 27(45):12176–12189.

Johnson A, Seeland K, Redish AD (2005) Reconstruction of the postsubiculum head direction signal from neural ensembles. Hippocampus 15(1):86–96.

Johnston D, Amaral DG (1998) Hippocampus. In: The Synaptic Organization of the Brain (Shepherd GM, ed.), pp. 417–458. Oxford University Press, 4th edn.

Jung MW, Wiener SI, McNaughton BL (1994) Comparison of spatial firing characteristics of the dorsal and ventral hippocampus of the rat. Journal of Neuroscience 14(12):7347–7356.

Kamin L (1969) Predictability, surprise, attention and conditioning. In: Punishment and Aversive Behavior (Campbell B, Church R, eds.). New York: Appleton-Century-Crofts.

Kemp C, Perfors A, Tenenbaum JB (2007) Learning overhypotheses with hierarchical bayesian models. Dev Sci 10(3):307–321.

Kentros C, Hargreaves E, Hawkins RD, Kandel ER, Shapiro M, Muller RV (1998) Abolition of long-term stability of new hippocampal place cell maps by NMDA receptor blockade. Science 280(5372):2121–2126.

Kentros CG, Agnihotri NT, Streater S, Hawkins RD, Kandel ER (2004) Increased attention to spatial context increases both place field stability and spatial memory. Neuron 42(2):283–295.

Kepecs A, Lisman J (2003) Information encoding and computation with spikes and bursts. Network: Computation in Neural Systems 14(1):103–118.

Killcross S, Coutureau E (2003) Coordination of actions and habits in the medial prefrontral cortex of rats. Cerebral Cortex 13(8):400–408.

Klausberger T, Magill PJ, Márton LF, Roberts JDB, Cobden PM, Buzsáki G, Somogyi P (2003) Brain-state and cell-type-specific firing of hippocampal interneurons *in vivo*. Nature 421:844–848.

Knierim JJ, Kudrimoti HS, McNaughton BL (1995) Place cells, head direction cells, and the learning of landmark stability. Journal of Neuroscience 15:1648–1659.

Knierim JJ, Kudrimoti HS, McNaughton BL (1998) Interactions between idiothetic cues and external landmarks in the control of place cells and head direction cells. J Neurophysiol 80(1):425–446.

Koene RA, Gorchetchnikov A, Cannon RC, Hasselmo ME (2003) Modeling goal-directed spatial navigation in the rat based on physiological data from the hippocampal formation. Neural Networks 16(5-6):577–584.

Koene RA, Hasselmo ME (2008) Reversed and forward buffering of behavioral spike sequences enables retrospective and prospective retrieval in hippocampal regions ca3 and ca1. Neural Netw 21(2-3):276–288.

Krechesvsky I (1932) The genesis of ‘hypotheses’ in rats. Univ Calif Publ Psychol 6(4):46.

Kruse J, Overmier J, Konz W, Rokke E (1983) Pavlovian conditioned stimulus effects on instrumental choice behavior are reinforcer specific. Learning and Motivation 14(2):165–181.

Kudrimoti HS, Barnes CA, McNaughton BL (1999) Reactivation of hippocampal cell assemblies: Effects of behavioral state, experience, and EEG dynamics. Journal of Neuroscience 19(10):4090–4101.

Lashley K (1929) Brain mechanisms and intelligence: a quantitative study of injuries of the brain. Chicago: University of Chicago Press.

Lebron K, Milad MR, Quirk GJ (2004) Delayed recall of fear extinction in rats with lesions of ventral medial prefrontal cortex. Learning and Memory 11(5):544–548.

Lee AK, Wilson MA (2002) Memory of sequential experience in the hippocampus during slow wave sleep. Neuron 36(6):1183–1194.

Lee I, Griffin AL, Zilli EA, Eichenbaum H, Hasselmo ME (2006) Gradual translocation of spatial correlates of neuronal firing in the hippocampus toward prospective reward locations. Neuron 51:639–650.

Lee I, Kesner RP (2002) Differential contribution of NMDA receptors in hippocampal subregions to spatial working memory. Nature Neuroscience 5(2):162–168.

Lee I, Rao G, , Knierim JJ (2004a) A double dissociation between hippocampal subfields: differential time course of CA3 and CA1 place cells for processing changed environments. Neuron 42(5):803–815.

Lee I, Yoganarasimha D, Rao G, , Knierim JJ (2004b) Comparison of population coherence of place cells in hippocampal subfields of CA1 and CA3. Nature 430(6998):456–459.

Lee TS, Mumford D (2003) Hierarchical bayesian inference in the visual cortex. Journal of Optical Society of America, A 20(7):1434–1448.

Lenck-Santini PP, Muller RU, Save E, Poucet B (2002) Relationships between place cell firing fields and navigational decisions by rats. Journal of Neuroscience 22(20):9035–9047.

Lenck-Santini PP, Save E, Poucet B (2001) Evidence for a relationship between place cell firing and spatial memory performance. Hippocampus 11(4):377–390.

Leutgeb JK, Leutgeb S, Treves A, Meyer R, Barnes CA, Moser EI, McNaughton BL, Moser MB (2005a) Progressive transformation of hippocampal neuronal representations in "morphed" environments. Neuron 48(2):345–358.

Leutgeb S, Leutgeb JK, Barnes CA, Moser EI, McNaughton BL, Moser MB (2005b) Independent codes for spatial and episodic memory in hippocampal neuronal ensembles. Science 309(5734):619–623.

Leutgeb S, Leutgeb JK, Treves A, Moser MB, Moser EI (2004) Distinct ensemble codes in hippocampal areas CA3 and CA1. Science 305(5688):1295–1298.

Levy WB (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 WB, Steward O (1983) Temporal contiguity requirements for long-term associative potentiation/depression in the hippocampus. Neuroscience 8(4):791–7.

Lisman JE (1999) Relating hippocampal circuitry to function: Recall of memory sequences by reciprocal dentate-CA3 interactions. Neuron 22:233–242.

Ljungberg T, Apicella P, Schultz W (1992) Responses of monkey dopamine neurons during learning of behavioral reactions. Journal of Neurophysiology 67(1):145–163.

Loftus E, Palmer J (1974) Reconstruction of automobile destruction. Journal of Verbal Learning and Verbal Behavior 13:585–589.

Lomo T (1971) Patters of activation in a monosynaptic cortical pathway: the perforant path input to the dentate area of the hippocampal formation. Experimental Brain Research 12:18–45.

Louie K, Wilson MA (2001) Temporally structured replay of awake hippocampal ensemble activity during Rapid Eye Movement sleep. Neuron 29:145–156.

Ludvig N (1999) Place cells can flexibly terminate and develop their spatial firing. a new theory for their function. Physiol Behav 67(1):57–67.

MacCorquodale K, Meehl PE (1948) On a distinction between hypothetical constructs and intervening variables. Psychological Review 55:95–107.

MacCorquodale K, Meehl PE (1954) Edward c. tolman. In: Modern Learning Theory (Estes W, ed.), pp. 177–266. Appleton-Century-Crofts.

Mackintosh NJ (1974) The Psychology of Animal Learning. Academic Press.

Mardia KV (1972) Statistics of Directional Data. New York: Academic Press.

Markus EJ, Barnes CA, McNaughton BL, Gladden VL, Skaggs WE (1994) Spatial information content and reliability of hippocampal ca1 neurons: effects of visual input. Hippocampus 4(4):410–421.

Markus EJ, Qin YL, Leonard B, Skaggs WE, McNaughton BL, Barnes CA (1995) Interactions between location and task affect the spatial and directional firing of hippocampal neurons. Journal of Neuroscience 15(11):7079–7094.

Marr D (1971) Simple memory: A theory of archicortex. Philosophical Transactions of the Royal Society of London 262(841):23–81.

Masimore B, Kakalios J, Redish AD (2004) Measuring fundamental frequencies in local field potentials. Journal of Neuroscience Methods In Press.

Masimore B, Schmitzer-Torbert NC, Kakalios J, Redish AD (2005) Transient striatal [gamma] local field potentials signal movement initiation in rats. NeuroReport 16(18):2021–2024.

McClelland JL, McNaughton BL, O'Reilly RC (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.

McNaughton BL (1983) Associative properties of hippocampal long term potentiation. In: Neurobiology of the Hippocampus (Seifert W, ed.), pp. 433–447. New York: Academic Press.

McNaughton BL, Barnes CA, Gerrard JL, Gothard K, Jung MW, Knierim JJ, Kudrimoti H, Qin Y, Skaggs WE, Suster M, Weaver KL (1996) Deciphering the hippocampal polyglot: The hippocampus as a path integration system. Journal of Experimental Biology 199(1):173–186.

McNaughton BL, Barnes CA, O'Keefe J (1983) The contributions of position, direction, and velocity to single unit activity in the hippocampus of freely-moving rats. Experimental Brain Research 52:41–49.

McNaughton BL, Nadel L (1990) Hebb-Marr networks and the neurobiological representation of action in space. In: Neuroscience and Connectionist Theory (Gluck MA, Rumelhart DE, eds.), pp. 1–63. Hillsdale NJ: Erlbaum.

Mehta MR, Barnes CA, McNaughton BL (1997) Experience-dependent, asymmetric expansion of hippocampal place fields. Proceedings of the National Academy of Sciences USA 94:8918–8921.

Mehta MR, Quirk MC, Wilson MA (2000) Experience-dependent asymmetric shape of hippocampal receptive fields. Neuron 25:707–715.

Meunzinger KF (1938) Vicarious trial and error at a point of choice I. a general survey of its relation to learning efficiency. Journal of Genetic Psychology 53:75–86.

Milad MR, Quirk GJ (2002) Neurons in medial prefrontal cortex signal memory for fear extinction. Nature 420:70–74.

Milad MR, Vidal-Gonzalez I, Quirk GJ (2004) Electrical stimulation of medial prefrontal cortex reduces conditioned fear in a temporally specific manner. Behavioral Neuroscience 118(2):389–394.

Miller G, Galanter E, Pribram K (1960) Plans and the structure of behavior. New York: Adams Bannister Cox.

Mittelstaedt ML, Mittelstaedt H (1980) Homing by path integration in a mammal. Naturwissenschaften 67:566–567.

Moita MAP, Rosis S, Zhou Y, LeDoux JE, Blair HT (2003) Hippocampal place cells acquire location-specific responses to the conditioned stimulus during auditory fear conditioning. Neuron 37(3):485–497.

Moita MAP, Rosis S, Zhou Y, LeDoux JE, Blair HT (2004) Putting fear in its place: Remapping of hippocampal place cells during fear conditioning. J Neurosci 24(31):7015–7023.

Montague PR, Dayan P, Person C, Sejnowski TJ (1995) Bee foraging in uncertain environments using predictive hebbian learning. Nature 377(6551):725–728.

Montague PR, Dayan P, Sejnowski TJ (1996) A framework for mesencephalic dopamine systems based on predictive Hebbian learning. Journal of Neuroscience 16(5):1936–1947.

Morris RGM (1981) Spatial localization does not require the presence of local cues. Learning and Motivation 12:239–260.

Morris RGM, Anderson E, Lynch GS, Baudry M (1986) Selective impairment of learning and blockade of long-term potentiation by an N-methyl-D-aspartate receptor antagonist, AP5. Nature 319(6056):774–776.

Morris RGM, Garrud P, Rawlins JNP, O'Keefe J (1982) Place navigation impaired in rats with hippocampal lesions. Nature 297:681–683.

Moser M, Moser E (1998) Functional differentiation in the hippocampus. Hippocampus 8(6):608–619.

Muenzinger KF (1938) Vicarious trial and error at a point of choice: A general survey of its relation to learning efficiency. Journal of Genetic Psychology 53:75–86.

Muir GM, Bilkey DK (2001) Instability in the place field location of hippocampal place cells after lesions centered on the perirhinal cortex. Journal of Neuroscience 21(11):4016–4025.

Müller M, Wehner R (1988) Path integration in desert ants, *cataglyphis fortis*. Proceedings of the National Academy of Sciences USA 85:5287–5290.

Muller RU, Kubie JL (1987) The effects of changes in the environment on the spatial firing of hippocampal complex-spike cells. Journal of Neuroscience 7:1951–1968.

Muller RU, Kubie JL, Ranck Jr JB (1987) Spatial firing patterns of hippocampal complex-spike cells in a fixed environment. Journal of Neuroscience 7:1935–1950.

Nádasdy Z, Hirase H, Czurkó A, Csicsvari J, Buzsáki G (1999) Replay and time compression of recurring spike sequences in the hippocampus. Journal of Neuroscience 19(2):9497–9507.

Nadel L (1994) Multiple memory systems: What and Why, an update. In: Memory Systems 1994 (Schacter DL, Tulving E, eds.), pp. 39–64. Cambridge MA: MIT Press.

Nakazawa K, Quirk MC, Chitwood RA, Watanabe M, Yeckel MF, Sun LD, Wilson MA, Tonegawa S (2002) Requirement for hippocampal ca3 nmda receptors in associative memory recall. Science 279(5579):211–218.

Neves G, Cooke SF, Bliss TVP (2008) Synaptic plasticity, memory and the hippocampus: a neural network approach to causality. Nature Reviews Neuroscience 9(1):65–75.

Newell A, Simon HA (1972) Human problem solving. Englewood Cliffs, NJ: Prentice-Hall.

Nicola SM, Yun IA, Wakabayashi KT, Fields HL (2004a) Cue-evoked firing of nucleus accumbens neurons encodes motivational significance during a discriminative stimulus task. Journal of Neurophysiology 91(4):1840–1865.

Nicola SM, Yun IA, Wakabayashi KT, Fields HL (2004b) Cuing of nucleus accumbens neurons during the consummatory phase of a discriminative stimulus task depends on previous reward. Journal of Neurophysiology 91(4):1866–1882.

Nirenberg S, Latham PE (2003) Decoding neuronal spike trains: How important are correlations? Proceedings of the National Academy of Sciences, USA 100(12):7348–7353.

Niv Y, Daw ND, Dayan P (2006a) Choice values. Nature Neuroscience 9(8):987–988.

Niv Y, Joel D, Dayan P (2006b) A normative perspective on motivation. Trends Cogn Sci 10(8):375–381.

O'Keefe J (1976) Place units in the hippocampus of the freely moving rat. Experimental Neurology 51:78–109.

O'Keefe J (1999) Do hippocampal pyramidal cells signal non-spatial as well as spatial information? Hippocampus 9(4):352–365.

O'Keefe J, Conway DH (1978) Hippocampal place units in the freely moving rat: Why they fire where they fire. Experimental Brain Research 31:573–590.

O'Keefe J, Dostrovsky J (1971) The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely moving rat. Brain Research 34:171–175.

O'Keefe J, Nadel L (1978) The Hippocampus as a Cognitive Map. Oxford: Clarendon Press.

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

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

Olton D (1979) Mazes, maps, and memory. American Psychologist 34(7):583–596.

Olton DS, Becker JT, Handelmann GE (1979) Hippocampus, space, and memory. Behavioral and Brain Sciences 2:313–322. See also commentary and response, pages 323–366.

Olton DS, Becker JT, Handelmann GE (1980) Hippocampal function: Working memory or cognitive mapping? Physiological Psychology 8(2):239–246.

Olton DS, Samuelson RJ (1976) Remembrance of places passed: Spatial memory in rats. Journal of Experimental Psychology: Animal Behavior Processes 2(2):97–116.

Olypher AV, L?nsk? P, Fenton AA (2002) Properties of the extra-positional signal in hippocampal place cell discharge derived from the overdispersion in location-specific firing. Neuroscience 111(3):553–566.

O'Neill J, Senior T, Csicsvari J (2006) Place-selective firing of CA1 pyramidal cells during sharp wave/ripple network patterns in exploratory behavior. Neuron 49(1):143–155.

O'Neill J, Senior TJ, Allen K, Huxter JR, Csicsvari J (2008) Reactivation of experience-dependent cell assembly patterns in the hippocampus. Nature Neuroscience 11(2):209–215.

Ono T, Nakamura K, Fukuda M, Tamura R (1991a) Place recognition responses of neurons in monkey hippocampus. Neuroscience Letters 121:194–198.

Ono T, Tamura R, Nakamura K (1991b) The hippocampus and space: Are there "place neurons" in the monkey hippocampus. Hippocampus 1(3):253–257.

Ostlund SB, Balleine BW (2007) Orbitofrontal cortex mediates outcome encoding in pavlovian but not instrumental conditioning. Journal of Neuroscience 27(18):4819–4825.

Otmakhova NA, Lisman JE (1996) D1/D5 dopamine receptor activation increases the magnitude of early long-term potentiation at CA1 hippocampal synapses. Journal of Neuroscience 16(23):7478–7486.

Otmakhova NA, Lisman JE (1999) Dopamine selectively inhibits the direct cortical pathway to the ca1 hippocampal region. Journal of Neuroscience 19(4):1437–1445.

Packard MG, McGaugh JL (1996) Inactivation of hippocampus or caudate nucleus with lidocaine differentially affects expression of place and response learning. Neurobiology of Learning and Memory 65:65–72.

Padoa-Schioppa C, Assad JA (2006) Neurons in the orbitofrontal cortex encode economic value. Nature 441:223–226.

Pan WX, Schmidt R, Wickens JR, Hyland BI (2005) Dopamine cells respond to predicted events during classical conditioning: Evidence for eligibility traces in the reward-learning network. Journal of Neuroscience 25(26):6235–6242.

Pavlides C, 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.

Pavlov I (1927) Conditioned Reflexes. Oxford Univ Press.

Pennartz CMA, Lee E, Verheul J, Lipa P, Barnes CA, McNaughton BL (2004) The ventral striatum in off-line processing: Ensemble reactivation during sleep and modulation by hippocampal ripples. J Neurosci 24(29):6446–6456.

Pickens CL, Saddoris MP, Setlow B, Gallagher M, Holland PC, Schoenbaum G (2003) Different roles for orbitofrontal cortex and basolateral amygdala in a reinforcer devaluation task. Journal of Neuroscience 23(35):11078–11084.

Poldrack R, Clark J, Pare-Blagoev E, Shohamy D, Creso Moyano J, Myers C, Gluck M (2001) Interactive memory systems in the human brain. Nature 414(6863):546–550.

Qin YL, McNaughton BL, Skaggs WE, Barnes CA (1997) Memory reprocessing in corticocortical and hippocampocortical neuronal ensembles. Philosophical Transactions of the Royal Society, London B352(1360):1525–1533.

Quirk GJ, Garcia R, González-Lima F (2006) Prefrontal mechanisms in extinction of conditioned fear. Biological Psychiatry 60(4):337–343.

Ramus SJ, Davis JB, Donahue RJ, Discenza CB, Waite AA (2007) Interactions between the orbitofrontal cortex and the hippocampal memory system during the storage of long-term memory. Ann N Y Acad Sci 1121:216–231.

Rao RPN, Ballard DH (1999) Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. Nature Neuroscience 2(1):79–87.

Rawlins JNP (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.

Redish AD (1997) Beyond the cognitive map: Contributions to a computational neuroscience theory of rodent navigation. Ph.D. thesis, Carnegie Mellon University.

Redish AD (1999) Beyond the Cognitive Map: From Place Cells to Episodic Memory. Cambridge MA: MIT Press.

Redish AD, Battaglia FO, Chawla MK, Ekstrom AD, Gerard JL, Lipa P, Rosenzweig ES, Worley PF, Guzowski JF, McNaughton BL, Barnes CA (2001) Independence of firing correlates of anatomically proximate hippocampal pyramidal cells. Journal of Neuroscience 21(RC134):1–6.

Redish AD, Jensen S, Johnson A, Kurth-Nelson Z (2007) Reconciling reinforcement learning models with behavioral extinction and renewal: Implications for addiction, relapse, and problem gambling. Psychological Review 114(3):784–805.

Redish AD, Johnson A (2007) A computational model of craving and obsession. Annals of the New York Academy of Sciences 1104(1):324–339.

Redish AD, Rosenzweig ES, Bohanick JD, McNaughton BL, Barnes CA (2000) Dynamics of hippocampal ensemble realignment: Time vs. space. Journal of Neuroscience 20(24):9289–9309.

Redish AD, Touretzky DS (1997) Cognitive maps beyond the hippocampus. Hippocampus 7(1):15–35.

Redish AD, Touretzky DS (1998a) The role of the hippocampus in solving the Morris water maze. Neural Computation 10(1):73–111.

Redish AD, Touretzky DS (1998b) Separating hippocampal maps. In: Spatial Functions of the Hippocampal Formation and the Parietal Cortex (Burgess N, Jeffery K, O'Keefe J, eds.), pp. 203–219. New York: Oxford University Press.

Rescorla RA (1969) Pavlovian conditioned inhibition. Psychological Bulletin 72:77–94.

Rescorla RA, Wagner AR (1972) A theory of Pavlovian conditiong: Variations in the effectiveness of reinforcement and nonreinforcement. In: Classical Conditioning II: Current Research and Theory (Black AH, Prokesy WF, eds.), pp. 64–99. New York: Appleton Century Crofts.

Restle F (1957) Discrimination of cues in mazes: A resolution of the 'place-vs-response' question. Psychological Review 64:217–228.

Rhodes SEV, Killcross S (2004) Lesions of rat infralimbic cortex enhance recovery and reinstatement of an appetitive pavlovian response. Learn Mem 11(5):611–616.

Rieke F, Warland D, de Ruyter van Steveninck R, Bialek W (1997) Spikes. Cambridge MA: MIT Press.

Roberts WA (2002) Are animals stuck in time? Psychological Bulletin 128(3):473–489.

Roberts WA, Feeney MC, Macpherson K, Petter M, McMillan N, Musolino E (2008) Episodic-like memory in rats: is it based on when or how long ago? Science 320(5872):113–115.

Roesch MR, Calu DJ, Schoenbaum G (2007) Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. Nature Neuroscience 10(12):1615–1624.

Rosenzweig ES, Redish AD, McNaughton BL, Barnes CA (2003) Hippocampal map realignment and spatial learning. Nature Neuroscience 6(6):609–615.

Rotenberg A, Abel T, Hawkins RD, Kandel ER, Muller RU (2000) Parallel instabilities of long-term potentiation, place cells, and learning caused by decreased protein kinase a activity. Journal of Neuroscience 20(21):8096–8102.

Saddoris MP, Gallagher M, Schoenbaum G (2005) Rapid associative encoding in basolateral amygdala depends on connections with orbitofrontal cortex. Neuron 46(2):321–331.

Salinas E, Abbott L (1994) Vector reconstruction from firing rates. Journal of Computational Neuroscience 1(1-2):89–107.

Samsonovich AV, McNaughton BL (1997) Path integration and cognitive mapping in a continuous attractor neural network model. Journal of Neuroscience 17(15):5900–5920.

Santini E, Ge H, Ren K, de Ortiz SP, Quirk GJ (2004) Consolidation of fear extinction requires protein synthesis in the medial prefrontal cortex. Journal of Neuroscience 24(25):5704–5710.

Schacter D (1995) Memory Distortion: How Minds, Brains and Societies Reconstruct the Past, chap. Memory distortion: history and current status, pp. 1–43. Harvard Univ. Press.

Schacter DL, Addis DR (2007a) The cognitive neuroscience of constructive memory: remembering the past and imagining the future. Philos Trans R Soc Lond B Biol Sci 362(1481):773–786.

Schacter DL, Addis DR (2007b) Constructive memory: the ghosts of past and future. Nature 445(7123):27.

Schacter DL, Addis DR, Buckner RL (2007) Remembering the past to imagine the future: the prospective brain. Nature Reviews Neuroscience 8(9):657–661.

Schacter DL, Addis DR, Buckner RL (2008) Episodic simulation of future events: Concepts, data, and applications. Ann N Y Acad Sci 1124:39–60.

Schmitzer-Torbert NC (2005) Ph.D. thesis, University of Minnesota.

Schmitzer-Torbert NC, Redish AD (2002) Development of path stereotypy in a single day in rats on a multiple-t maze. Archives Italiennes de Biologie 140:295–301.

Schmitzer-Torbert NC, Redish AD (2004) Neuronal activity in the rodent dorsal striatum in sequential navigation: separation of spatial and reward responses on the multiple t task. Journal of Neurophysiology 91(5):2259–2272.

Schneidman E, Bialek W, Berry MJ (2003) Synergy, redundancy, and independence in population codes. Journal of Neuroscience 23(37):11539–11553.

Schoenbaum G, Chiba AA, Gallagher M (1998) Orbitofrontal cortex and basolateral amygdala encode expected outcomes during learning. Nature Neuroscience 1(2):155–159.

Schoenbaum G, Chiba AA, Gallagher M (1999) Neural encoding in orbitofrontal cortex and basolateral amygdala during olfactory discrimination learning. Journal of Neuroscience 19(5):1876–1884.

Schoenbaum G, Chiba AA, Gallagher M (2000) Changes in functional connectivity in orbitofrontal cortex and basolateral amygdala during learning and reversal training. Journal of Neuroscience 20(13):5179–5189.

Schoenbaum G, Setlow B, Saddoris MP, Gallagher M (2005) Encoding predicted outcome and acquired value in orbitofrontal cortex during cue sampling depends upon input from basolateral amygdala. Neuron 39(5):855–867.

Schultz W (1998) Predictive reward signal of dopamine neurons. Journal of Neurophysiology 80:1–27.

Schultz W (2002) Getting formal with dopamine and reward. Neuron 36(2):241–263.

Schultz W, Apicella P, Ljungberg T (1993) Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. Journal of Neuroscience 13:900–913.

Schultz W, Dayan P, Montague PR (1997) A neural substrate of prediction and reward. Science 275(5306):1593–1599.

Scoville WB, Milner B (1957) Loss of recent memory after bilateral hippocampal lesions. Journal of Neurology, Neurosurgery, and Psychiatry 20:11–21.

Serruya M, Hatsopoulos N, Fellows M, Paninski L, Donoghue J (2004) Robustness of neuroprosthetic decoding algorithms. Biological Cybernetics 88(3):219–228.

Sharp PE, Blair HT, Etkin D, Tzanetos DB (1995) Influences of vestibular and visual motion information on the spatial firing patterns of hippocampal place cells. Journal of Neuroscience 15(1):173–189.

Shen B, McNaughton BL (1996) Modeling the spontaneous reactivation of experience-specific hippocampal cell assembles during sleep. Hippocampus 6(6):685–693.

Shen J, Barnes CA, McNaughton BL, Skaggs WE, Weaver KL (1997) The effect of aging on experience-dependent plasticity of hippocampal place cells. Journal of Neuroscience 17(17):6769–6782.

Simon H (1955) A behavioral model of rational choice. The Quarterly Journal of Economics 69:99–118.

Skaggs WE, McNaughton BL (1996) Replay of neuronal firing sequences in rat hippocampus during sleep following spatial experience. Science 271:1870–1873.

Skaggs WE, McNaughton BL, Wilson MA, Barnes CA (1996) Theta phase precession in hippocampal neuronal populations and the compression of temporal sequences. Hippocampus 6(2):149–173.

Smith AC, Frank LM, Wirth S, Yanike M, Hu D, Kubota Y, Graybiel AM, Suzuki WA, Brown EN (2004) Dynaimic analysis of learning in behavioral experiments. Journal of Neuroscience 24(2):447–461.

Smyrnis M, Taira M, Ashe J, Georgopoulos AP (1992) Motor cortical activity in a memorized delay task. Experimental Brain Research 92:139–151.

Solomon PR, Schaaf ERV, Thompson RF, Weisz DJ (1986) Hippocampus and trace conditioning of the rabbit's classically conditioned nictitating membrane response. Behavioral Neuroscience 100(5):729–744.

Somogyi P, Klausberger T (2005) Defined types of cortical interneurone structure space and spike timing in the hippocampus. J Physiol (Lond) 562:9–26.

Spence KW, Lippitt R (1946) An experimental test of the sign-gestalt theory of trial and error learning. Journal of Experimental Psychology 36(6):491–502.

Squire LR (1987) Memory and Brain. New York: Oxford University Press.

Squire LR (1992) Memory and the hippocampus: A synthesis from findings with rats, monkeys, and humans. Psychology Review 99(2):195–231.

Squire LR, Alvarez P (1995) Retrograde amnesia and memory consolidation: A neurobiological perspective. Current Opinion in Neurobiology 5:169–177.

Squire LR, Stark CEL, Clark RE (2004) The medial temporal lobe. Annu Rev Neurosci 27:279–306.

Squire LR, Zola-Morgan SM (1991) The medial temporal lobe memory system. Science 253:1380–1386.

Stephens DW, Krebs JR (1987) Foraging Theory. Princeton.

Stouffer EM, White NM (2007) Roles of learning and motivation in preference behavior: mediation by entorhinal cortex, dorsal and ventral hippocampus. Hippocampus 17(2):147–160.

Suddendorf T, Busby J (2003) Mental time travel in animals? Trend in cognitive sciences 7(9):391–396.

Suddendorf T, Corballis MC (1997) Mental time travel and the evolution of the human mind. Genetic, Social & General Psychology Monographs 123(2):133–169.

Suddendorf T, Corballis MC (2007) The evolution of foresight: What is mental time travel, and is it unique to humans? Behav Brain Sci 30(3):299–313; discussion 313–51.

Sutton RS (1990) Integrated architectures for learning, planning, and reacting based on approximating dynamic programming. In: Proceedings of the Seventh International Conference on Machine Learning, pp. 216–224.

Sutton RS, Barto AG (1981) Toward a modern theory of adaptive networks: Expectation and prediction. Psychological Review 88(2):135–170.

Sutton RS, Barto AG (1998) Reinforcement learning: An introduction. Cambridge MA: MIT Press.

Tanila H, Shapiro ML, Eichenbaum H (1997) Discordance of spatial representation in ensembles of hippocampal place cells. Hippocampus 7(6):613–623.

Tatsuno M, Lipa P, McNaughton BL (2006) Methodological considerations on the use of template matching to study long-lasting memory trace replay. Journal of Neuroscience 26(42):10727–10742.

Tenenbaum J, Xu F (2000) Word learning as bayesian inference. In: 22nd Annu. Conf. Cogn. Sci. Soc. (Gleitman L, Joshi A, eds.), p. 517?522. Erlbaum.

Tenenbaum JB, Griffiths TL, Kemp C (2006) Theory-based bayesian models of inductive learning and reasoning. Trends Cogn Sci 10(7):309–318.

Teng E, Squire LR (1999) Memory for places learned long ago is intact after hippocampal damage. Nature 400:675–677.

Teyler TJ, DiScenna P (1986) The hippocampal memory indexing theory. Behavioral Neuroscience 100(2):147–154.

Thompson LT, Best PJ (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.

Thrun S, Burgard W, Fox D (2005) Probabilistic Robotics. MIT Press.

Thrun S, Langford J, Verma V (2001) Risk sensitive particle filters. In: Advances in Neural Information Processing Systems 14. Cambridge, MA: MIT Press.

Tinklepaugh O (1928) An experimental study of representative factors in monkeys. Journal of Comparative Psychology 8:197–236.

Tolman E, Brunswik E (1935) The organism and the causal texture of the environment. Psychological Review 42(1):43–77.

Tolman EC (1932) Purposive behavior in animals and men. New York: The Century Co.

Tolman EC (1938) The determiners of behavior at a choice point. Psychological Review 46:318–336.

Tolman EC (1939) Prediction of vicarious trial and error by means of the schematic sowbug. Psychological Review 46:318–336.

Tolman EC (1948) Cognitive maps in rats and men. Psychological Review 55:189–208.

Tolman EC (1949) There is more than one kind of learning. Psychological Review 56(3):144–155.

Tolman EC (1954) Freedom and the cognitive mind. American Psychologist 9(9):536–538.

Tolman EC (1955) Principles of performance. Psychological Review 62(5):315–326.

Tolman EC, Krechevsky I (1933) Means-end readiness and hypothesis. a contribution to comparative psychology. Psychological Review 40(1):60–70.

Tolman EC, Ritchie BF, Kalish D (1946) Studies in spatial learning. I. Orientation and the short-cut. Journal of Experimental Psychology 36:13–24.

Touretzky DS, Redish AD (1996) A theory of rodent navigation based on interacting representations of space. Hippocampus 6(3):247–270.

Tremblay L, Schultz W (1999) Relative reward preference in primate orbitofrontal cortex. Nature 398(6729):704–708.

Treves A (2004) Computational constraints between retrieving the past and predicting the future, and the ca3-ca1 differentiation. Hippocampus 14(5):539–556.

Tse D, Langston RF, Kakeyama M, Bethus I, Spooner PA, Wood ER, Witter MP, Morris RGM (2007) Schemas and memory consolidation. Science 316(5821):76–82.

Tsodyks M (1999) Attractor network models of spatial maps in hippocampus. Hippocampus 9(4):481–489.

Tsodyks M (2005) Attractor neural networks and spatial maps in hippocampus. Neuron 48(2):168–169.

Tulving E (1983) Elements of Episodic Memory. New York: Oxford University Press.

Tulving E (1984) Precis of elements of episodic memory. The Behavioral and Brain Sciences 7:223–238.

Tulving E (2001) Episodic memory and common sense: how far apart? Philosophical Transactions of the Royal Society B: Biological Sciences 356(1413):1505–1515.

Tulving E (2002) Episodic memory: From mind to brain. Annual Review of Psychology 53:1–25.

Tulving E, Madigan S (1970) Memory and verbal learning. Annual Reviews in Psychology 21:437–484.

Turing A (1992) Collected Works of A.M. Turing : Mechanical Intelligence. North Holland.

Ulanovsky N, Moss CF (2007) Hippocampal cellular and network activity in freely moving echolocating bats. Nature Neuroscience 10(2):224–233.

Vanderwolf CH (1969) Hippocampal electrical activity and voluntary movement in the rat. Clinical Neurophysiology 26:407–418.

Vanderwolf CH (1971) Limbic-diencephalic mechanisms of voluntary movement. Psychological Review 78(2):83–113.

von Neumann J, Morgenstern O (1944) Theory of games and economic behavior. Princeton NJ: Princeton University Press.

Waelti P, Dickinson A, Schultz W (2001) Dopamine responses comply with basic assumptions of formal learning theory. Nature 412:43–48.

Wallenstein GV, Eichenbaum H, Hasselmo ME (1998) The hippocampus as an associator of discontiguous events. Trends in Neurosciences 21(8):317–323.

Wang M, Vijayraghavan S, Goldman-Rakic PS (2004) Selective d2 receptor actions on the functional circuitry of working memory. Science 303(5660):853–856.

Watkins CJ, Dayan P (1992) Q-learning. Machine Learning 8(3):279–292.

Wehner R, Srinivasan MV (1981) Searching behavior of desert ants, genus *Cataglyphis* (Formicidae, Hymenera). Journal of Comparative Physiology A 142:315–338.

Whishaw IQ, Mittleman G (1986) Visits to starts, routes, and places by rats (*Rattus norvegicus*). Journal of Comparative Psychology 100(4):422–431.

White NM, Gaskin S (2006) Dorsal hippocampus function in learning and expressing a spatial discrimination. Learn Mem 13(2):119–122.

Wickens JR, Reynolds JNJ, Hyland BI (2003) Neural mechanisms of reward-related motor learning. Current Opinion in Neurobiology 13:685–690.

Widrow B, Hoff M (1960) Adaptive switching circuits. In: 1960 Institute of Radio Engineers, Western Electronic Show and Convention, pp. 96–104. Convention Record.

Wilkerson A, Levin ED (1999) Ventral hippocampal dopamine D1 and D2 systems and spatial working memory in rats. Neuroscience 89(3):743–749.

Wills TJ, Lever C, Cacucci F, Burgess N, O'Keefe J (2005) Attractor dynamics in the hippocampal representation of the local environment. Science 308(5723):873–876.

Wilson MA, McNaughton BL (1993) Dynamics of the hippocampal ensemble code for space. Science 261:1055–1058.

Wilson MA, McNaughton BL (1994) Reactivation of hippocampal ensemble memories during sleep. Science 265(5172):676–679.

Wood ER, Dudchenko PA, Eichenbaum H (1999) The global record of memory in hippocampal neuronal activity. Nature 397:613–616.

Wood ER, Dudchenko PA, Robitsek RJ, Eichenbaum H (2000) Hippocampal neurons encode information about different types of memory episodes occurring in the same location. Neuron 27(3):623–633.

Wu W, Gao Y, Bienenstock E, Donoghue JP, Black MJ (2006) Bayesian population decoding of motor cortical activity using a kalman filter. Neural Computation 18(1):80–118.

Yamaguchi Y (2003) A theory of hippocampal memory based on theta phase precession. Biological Cybernetics .

Yin HH, Knowlton BJ, Balleine BW (2004) Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. European Journal of Neuroscience 19:181–189.

Yin HH, Knowlton BJ, Balleine BW (2005) Blockade of nmda receptors in the dorsomedial striatum prevents action-outcome learning in instrumental conditioning. European Journal of Neuroscience 22(2):505–512.

Yin HH, Knowlton BJ, Balleine BW (2006) Inactivation of dorsolateral striatum enhances sensitivity to changes in the action-outcome contingency in instrumental conditioning. Behavioural Brain Research 166(2):189–196.

Ylinen A, Bragin A, Nadasdy Z, Jando G, Szabo I, Sik A, 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.

Yun IA, Wakabayashi KT, Fields HL, Nicola SM (2004) The ventral tegmental area is required for the behavioral and nucleus accumbens neuronal firing responses to incentive cues. Journal of Neuroscience 24(12):2923–2933.

Zar JH (1999) Biostatistical Analysis. Upper Saddle River, NJ: Prentice-Hall, fourth edn.

Zemel RS, Dayan P, Pouget A (1998) Probabilistic interpretation of population codes. Neural Computation 10(2):403–430.

Zhang K, Ginzburg I, McNaughton BL, Sejnowski TJ (1998) Interpreting neuronal population activity by reconstruction: Unified framework with application to hippocampal place cells. Journal of Neurophysiology 79:1017–1044.

Zilli EA, Hasselmo ME (2008) Modeling the role of working memory and episodic memory in behavioral tasks. Hippocampus 18(2):193–209.

Zola-Morgan S, Squire LR (1990) The primate hippocampal formation: Evidence for a time-limited role in memory storage. Science 250:288–290.

Zola-Morgan S, Squire LR (1993) Neuroanatomy of memory. Annual Review of Neuroscience 16:547–563.