Dissertation Proposal: A Computational Model of Spatial Navigation During Simulated Motion in Rats

Steven J Doubleday

Institute for Mathematical Behavioral Sciences

University of California, Irvine

Contents

Aims

The principal objective of this dissertation is to provide a neuroscientifically-plausible account of spatial navigation during simulated motion in rats, using a relatively novel computational modeling approach. There is evidence that during quiet awake periods a rat generates sequences of activation ("trajectories") in the hippocampus that depict ("preplay") its future motion in two-dimensional space (Pfeiffer & Foster, 2013). The mechanism by which these preplay trajectories are generated is not known; the theory to be explored here is that they are part of a process of the animal simulating its future motion, a process that will be proposed to begin in cortical areas outside the hippocampus. An alternate hypothesis is that the trajectories originate in the hippocampus. An origin within the hippocampus has been proposed for similar trajectories that appear to "replay" the animal's past actions. These replay trajectories arise in somewhat different behavioral circumstances and are presumed to have different functions (Wilson & McNaughton, 1994; Lee & Wilson, 2002; Diba & Buzsáki, 2007), but the default hypothesis that all these trajectories are generated by processes local to the hippocampus would be parsimonious. The burden of this work is to explore the theory that in some circumstances trajectories are a byproduct of simulated motion, and are driven from outside the hippocampus.

Simulations can be defined as "imaginative constructions of hypothetical events or scenarios" (Schacter et al., 2012). There is substantial evidence for simulation in humans, for example during language understanding (B. K. Bergen, 2012). I suggest that a subset of simulation mechanisms may also be employed by rats. In rats, the focus will be narrowed to simulated motion in two-dimensional space. Simulated motion is the execution of a motor routine without activation of the motor neurons that would cause actual physical motion. I suggest that the simulated motion updates a path integration system, which has been hypothesized as existing mainly in the medial entorhinal cortex (McNaughton, Battaglia, Jensen, Moser, & Moser, 2006). In turn, this activates the place fields in the hippocampus that constitute the observed trajectory.

2

I suggest that the adaptive value of simulated motion is as a fast way of exploring the likely rewards from moving in various possible directions. This proposal raises the question of what additional information is provided by simulated motion towards a goal location, if the animal has previously found the target location rewarding. One hypothesis to be explored here is that the path towards the goal location encodes additional information about possible rewards. An alternate hypothesis is that the target location itself provides the reward information, and the path towards the goal location provides no additional information.

The core of the work is development of a computational model of spatial navigation during simulated motion, termed the Simulated Motion System (SMS). If this work is successful, it will demonstrate that a theory and supporting computational model of spatial navigation during simulated motion is consistent with much of current understanding of spatial behavior in rats. The theory focuses on processing in areas that are important to spatial navigation, principally the entorhinal cortex and hippocampus. The theory will have much less to say about processing in other cortical areas, although it will include a simple model of the interaction between spatial navigation processes and the neocortex.

The computational model will make use of Petri nets, a formalism developed for the modeling of distributed systems. A Petri net defines states, and transitions between states. Directed arcs define the conditions under which transitions execute, to move the Petri net from one state to another. Petri nets execute without a centralized point of control, which makes them suitable for modeling asynchronous and concurrent processes, including neural processes. Extensions to Petri nets, developed in prior work by the author, enable larger Petri nets to be built from smaller Petri nets and from other software components. This makes it feasible to use Petri nets to build models of complex processes, such as those to be explored here. The Petri nets make visible important events and the connections between processes.

The specific aims of the dissertation depend on the development of a computational model of real and simulated motion in rats (SMS). SMS will implement and extend an existing set of models of components and processes that support spatial navigation in rats. A major underlying concept in these models is that of an attractor network, structured in some cases as an autoassociative network model of memory, and in other cases as a representation of aspects of the animal's current spatial position. Although not universally adopted in the literature, these attractor models do

account for a wide range of experimental findings (Navratilova & McNaughton, 2014). Recent revisions to these models (Colgin et al., 2010; J. K. Leutgeb, Leutgeb, Moser, & Moser, 2007) provide an account of the functioning of the entorhinal cortex and the hippocampus in support of spatial navigation that will serve as the basis for the computational model. Once the computational model has been developed, it will be validated by demonstrating its ability to simulate important navigational behaviors and some of their neural correlates. Once the model is validated, its support for various hypotheses will be assessed. The specific aims are as follows:

1. **Assess the support provided by the computational model for each of two hypotheses regarding the mechanism that generates novel trajectories in the hippocampus during quiet awake periods.**

   A primary function of the hippocampus is to represent locations that the animal has previously experienced. This is seen most clearly by the activation of hippocampal "place" cells upon repeated visits to a given location, but recent experiments (Dragoi & Tonegawa, 2011, 2013; Ólafsdóttir, Barry, Saleem, Hassabis, & Spiers, 2015) have shown that place cells can be allocated to represent locations that the animal has never physically visited. These findings provide a clear area within which to distinguish two hypotheses. What I will refer to as the **hippocampal** hypothesis, proposed by Dragoi and Tonegawa (2011), suggests that a local process within the hippocampus makes populations of cells available for use in representing future locations that have not yet been visited. This addresses the mechanism by which novel trajectories can be generated (although it does not address the function of such trajectories.) In contrast, the **simulated motion** hypothesis proposed here is that processes outside the hippocampus drive the generation of trajectories, including novel trajectories. A target location or direction is selected, and simulated motion towards the target causes updates in the path integration system, thereby activating a new population of place cells in the hippocampus to represent the novel trajectory. The hippocampal and simulated motion hypotheses make different predictions, but only the simulated motion hypothesis is consistent with the other findings explained by the attractor models. If the behavior of the computational model matches the predictions, this will constitute a challenge for the hippocampal hypothesis, and support for the simulated motion hypothesis.

2. **Assess the support provided by the the computational model for each of two hypotheses regarding the information provided by simulated motion towards a goal location.**

In the experiment of Pfeiffer and Foster (2013), the trajectories during quiet awake periods represented paths towards a home location more frequently than towards any other location in the arena. Under the experimental design, the home location is intermittently rewarding, at a rate greater than for any other location in the arena. The greater number of home-directed trajectories suggests that the animal is aware that the home location may be rewarding. If that is the case, what is the additional value of simulating motion towards home? The **target location** hypothesis is that the population of place cells that represent the home location are associated with a higher probability of reward, and there is no additional value in simulating motion towards the goal. The **target trajectory** hypothesis, on the other hand, suggests that the sequence of place fields that constitute a trajectory might encode additional information about the probability of reward once the target is reached. In spatial alternation tasks, where the rewarding location differs from trial to trial, the rate of place cell firing varies on portions of the physical path that are the same, and these differences correlate with the reward location. Perhaps an analogous mechanism exists for trajectories during quiet awake periods. The target trajectory hypothesis suggests that the trajectory can act as a context within which the animal evaluates the probability that the target location will be rewarding. Distinguishing between these hypotheses will require enhancements to the computational model, which will make testable predictions.

These are not the only two plausible hypotheses regarding the observed trajectories. For example, it is plausible to suggest that some mechanism internal to the hippocampus, such as gradient descent towards a remembered goal location, is responsible for generating the novel trajectories. The objective here is not to eliminate such a hypothesis, but rather to examine the evidence for and consequences of the simulated motion hypothesis. The burden of this aim is to suggest a mechanism that might account for the observed evidence by explaining the circumstances under which simulated motion might be advantageous.

3. **Assess the usefulness of Petri nets as an approach for modeling complex neural processes.**

   Although the Petri net formalism is Turing complete, i.e., it can solve any problem that a Turing machine can solve, it does not seem well-adapted to high-fidelity modeling of the details of the processing of individual neurons. The conjecture to be tested by this work is that Petri nets, with extensions to enable the testing and integration of multiple components, will be useful for modeling complex neural processes at some middle level of abstraction. The goal is to model simulated motion in rats with enough abstraction to make the problem tractable, but with sufficient detail to test the specific hypotheses above.

Significance

There is growing evidence that simulating action is a principal mechanism by which humans understand events, including language (B. K. Bergen, 2012), (B. Bergen, Chang, & Narayan, 2004). The results reported in Pfeiffer and Foster (2013) are consistent with a hypothesis that rats also use some form of simulation. This work provides a theory about the role that spatial navigation processes in the entorhinal cortex and hippocampus might play in simulated motion. The theory makes testable neuroscientific predictions about simulated motion in rats.

Implementing this theory in a complete computational model of simulated motion in rats is a very challenging objective, as numerous cortical processes would be involved. Reducing the scope to focus primarily on the spatial navigation aspects of simulated motion still entails a computational model of significant complexity. The computational model allows testing of hypotheses about specific mechanisms at work in simulated motion. The computational model is open source, and its components could be reused or updated to address other questions. Finally, the work demonstrates the usefulness of the Petri net formalism in creating neuroscientifically plausible computational models.

Background

Neuroscience of spatial navigation

The account of simulated motion in rats will focus in detail on spatial navigation, and will briefly address how physical motion is motivated and accomplished.

6

Understanding spatial navigation in rats begins with the discovery that some neurons in the hippocampus of freely moving rats fire reliably when the animal is in a given location of the current environment, and are largely silent elsewhere (O'Keefe & Dostrovsky, 1971; O'Keefe, 1976). The activity of these "place cells" varies with two major modes of rat behavior, characterized by patterns in the local field potentials (LFP) measured in the hippocampus (O'Keefe & Nadel, 1978). LFP oscillations in the hippocampal and retrohippocampal formations in the frequency range of 6-10 Hz are known as "theta", and are associated with exploration of novel environments or novel objects in known environments, and movement that involves spatial displacement. "Large-amplitude irregular activity" (LIA) is a second pattern found in the hippocampus, with a major frequency slower than theta; it is observed during quiet periods without spatial displacement, e.g., while sitting quietly, grooming, eating, drinking, or during slow-wave sleep. At times during LIA there are "sharp wave ripples" (SWR), consisting of a large amplitude change in LFPs accompanied by a high-frequency burst of activity, originating in CA3 and propagating to CA1, the subiculum and deep layers of the entorhinal cortex (Buzsáki, S., & Vanderwolf, 1983; Buzsáki, 1986; Buzsáki, Horvath, Urioste, Hetke, & Wise, 1992).

As a rat moves from place to place during active exploration, sequences of place cells, "trajectories", are activated, whose place fields depict the animal's physical path. O'Keefe and Recce (1993) found that the firing of these place cells in both CA1 and CA3 happens at a slightly higher frequency than the theta oscillation that accompanies exploration. This results in "precession" – upon entry to a given place field, place cells fire late in the theta cycle, and fire earlier in each successive phase cycle. The functional significance of precession may be to associate successive place fields with each other within a window of synaptic plasticity, such that trajectories are strengthened (O'Keefe & Recce, 1993; Skaggs, McNaughton, Wilson, & Barnes, 1996; Mehta, Barnes, & McNaughton, 1997).

Trajectories are observed under circumstances other than physical movement. Some appear to "replay" the animal's past actions, during REM sleep (Louie & Wilson, 2001), or during SWRs in slow-wave sleep (Wilson & McNaughton, 1994; Lee & Wilson, 2002), or during SWRs in quiet awake periods (Foster & Wilson, 2006; Diba & Buzsáki, 2007; Karlsson & Frank, 2009). In the context of a general theory of the relationship between the hippocampus and the neocortex (McClelland, McNaughton, & O'Reilly, 1995), replay has been suggested as a candidate for a mechanism to re-

peatedly update the neocortex with recent experiences, so that the neocortex can gradually learn the structure of those experiences. Under this hypothesis of their function, it is plausible to think of replay trajectories as originating in the hippocampus through some local process. Within an autoassociative model of the CA3 area of the hippocampus, this process could be as simple as re-activation of sequences through random changes in activation levels (Wilson & McNaughton, 1994).

Other trajectories appear to "preplay" the animal's actions, i.e., to reflect the future activity of the animal. This has been observed during active exploration, particularly during vicarious trial and error (VTE) behavior at a choice point (Johnson & Redish, 2007). During VTE, animals are observed to orient briefly toward each possible path in turn. Simultaneously, hippocampal trajectories are observed, sweeping forward down the available paths in turn. Trajectories that reflect future activity are also observed during quiet awake SWRs (Pfeiffer & Foster, 2013). Although they seem to have different functions, as a practical matter, replay and preplay trajectories are often found together and may be difficult to distinguish (Diba & Buzsáki, 2007; Davidson, Kloosterman, & Wilson, 2009; Singer & Frank, 2009).

Given the similarities and co-occurrence of replay and preplay trajectories, it is parsimonious to suggest that they all originate in the hippocampus. An alternative to this hippocampal hypothesis is the simulated motion hypothesis, which suggests that preplay trajectories are part of a process of simulated motion, where the animal simulates exploration of the environment in a given direction to ascertain whether that direction seems likely to be rewarding. In this hypothesis, the origin of a preplay trajectory is external to the hippocampus.

These are far from the only plausible hypotheses regarding the generation of trajectories. Trajectories occur under very different behavioral circumstances, and they proceed at varying time scales that may differ greatly from that of physical motion. It is conceivable that there are multiple mechanisms that generate trajectories, to reflect these differences. For simplification, however, this work will focus only on the two hypotheses above.

The simulation hypothesis is situated here within the context of several theories regarding the function of the hippocampal and retrohippocampal areas in support of spatial navigation (reviewed in Navratilova and McNaughton (2014)). A basic concept is that of an attractor network, a dynamic system consisting of an N-dimensional space whose state at a given moment is defined by a vector, and where attractor states can be

identified, toward which states in a close neighborhood of the attractor tend to evolve over time. When configured as an "autoassociative network", where linkages exist among all the elements that constitute a given system state, these models have been used to represent memory (Hopfield, 1982). Attractor models have also been used to represent aspects of the spatial position of an animal (McNaughton & Nadel, 1990). An example of this followed the discovery that the allocentric angular orientation of a rat with respect to a fixed landmark is represented by "head direction" cells (Ranck Jr, 1985), which appear to behave as a system (Taube, Muller, & Ranck, 1990a, 1990b). This means that if one head direction cell changes its preferred orientation of firing, other head direction cells will be observed to change their orientations in the same direction and with the same angular displacement. This head direction system can be modeled through a one-dimensional attractor model conceptualized as a ring (McNaughton, Leonard, & Chen, 1989; McNaughton, Chen, & Markus, 1991; Skaggs, Knierim, Kudrimoti, & McNaughton, 1995; Zhang, 1996; Redish, Elga, & Touretzky, 1996), where the animal's current orientation is represented by a "bump", a population of cells whose activation is self-sustaining. Self-motion cues and visual cues will cause the bump to move, representing the updated orientation.

Many species are capable of path integration, tracking movements through an environment using self-motion cues, possibly corrected by sensory cues from landmarks, such that direct routes to arbitrary points elsewhere in the environment can be calculated as needed (Mittelstaedt & Mittelstaedt, 1980). Coupling the head direction model with another attractor model of the animal's position in two-dimensional space, McNaughton et al. (1996) and Samsonovich and McNaughton (1997) generated a model that accounts for path integration behavior and many findings regarding experimentally-induced discrepancies between self-motion and sensory cues (Gothard, Skaggs, & McNaughton, 1996; O'Keefe & Burgess, 1996; Knierim, Kudrimoti, & McNaughton, 1995; Sharp, Kubie, & Muller, 1990; Wilson & McNaughton, 1993). The model introduced the concept of a "chart", an abstract plane where place cells are arranged such that the pattern of neural activation is a localized activity packet centered on the animal's head, when mapped to a given environment. Each different environment has its own chart.

This model pre-dated the discovery of "grid cells" in the medial entorhinal cortex (MEC), whose firing pattern resembles a pattern of tessellated equilateral triangles overlaid on the rat's environment, providing a stable metric for space (Hafting, Fyhn,

Molden, Moser, & Moser, 2005). The model required relatively minor changes to accommodate grid cells, and made some predictions that were consistent with later grid cell discoveries. Grid cells appear to be organized in discrete modules at increasing spatial scales along the dorsal-ventral axis of the medial entorhinal cortex (Stensola et al., 2012). This discrete spatial structure was predicted by the spatial attractor model, and can account for an increase in spatial scale in the hippocampus along its dorsal-ventral axis, and several findings related to place field size (Jung, Wiener, & McNaughton, 1994; Terrazas et al., 2005; Maurer, VanRhoads, Sutherland, Lipa, & McNaughton, 2005). This accumulation of evidence, in combination with the discovery of other cell types in MEC predicted by the spatial attractor model (Sargolini et al., 2006; Couey et al., 2013), provides significant support for the hypothesis that the path integration function and spatial attractor dynamics reside in the MEC, with collaboration from the hippocampus and lateral entorhinal cortex (LEC) (Colgin et al., 2010; J. K. Leutgeb et al., 2007).

In contrasting the roles of the medial and lateral entorhinal cortices, Deshmukh (2014) suggests that the MEC processes self-motion cues in support of path integration, while LEC generates both spatial and non-spatial representations from external sensory input. Among other sources, the LEC receives input from the amygdala, so the non-spatial representations may also include information as to whether the current inputs are rewarding. The introduction of the animal to a novel environment causes "global remapping" – allocation of a new chart in MEC. In turn, the population of grid cells in that new chart activates a new population of place cells in the hippocampus to represent place fields at the various locations on that chart. Changes in the inputs from LEC, on the other hand, cause "rate remapping" – modification of the firing rates of the place cells allocated to the current chart, to represent information about landmarks, the current task, and rewards (S. Leutgeb et al., 2005; Navratilova, Hoang, Schwindel, Tatsuno, & McNaughton, 2012).

A primary point for this work is that these models suggest that the spatial attractor dynamics of where the animal is on the current chart is driven out of the medial entorhinal cortex. The dynamics will cause one population of grid cells to be activated on one chart, representing the animal's current location, and the bump will move across the chart in response to self-motion cues. The grid cells project directly to place cells in the hippocampus, and the activation of place cells is an immediate consequence of the activation of grid cells one synapse upstream. Global remapping

is interpreted as occurring when the attractor dynamics in MEC cause a different, possibly novel, chart to be activated, causing a different population of grid cells to be activated, and consequently, a different population of place cells. This theory has different consequences for the hippocampal and simulated motion hypotheses for the origin of trajectories, consequences to be explored in the computational model.

Although the evidence supporting the spatial attractor model is substantial, for challenges to various aspects of these models, see Van Cauter, Poucet, and Save (2008); O'Reilly, Alarcon, and Ferbinteanu (2014); Ferbinteanu, Shirvalkar, and Shapiro (2011). For conflicting reports on the relation between entorhinal cortex and hippocampus under theta oscillations, see Brandon, Koenig, Leutgeb, and Leutgeb (2014) and Wang, Romani, Lustig, Leonardo, and Pastalkova (2015), with an attempted reconciliation by Ormond and McNaughton (2015)).

The structure of the CA3 area of the hippocampus, with support from the dentate gyrus, is compatible with the hypothesis that CA3 functions as an autoassociative network (McNaughton & Morris, 1987; McNaughton & Nadel, 1990; Treves & Rolls, 1992). Colgin et al. (2010) suggested that CA3 could support the encoding of sequences of place fields as trajectories. In addition, the place field sequences could function as index-code sequences to enable the retrieval of episodic memories whose details are stored in various areas of neocortex (McClelland et al., 1995; Teyler & DiScenna, 1986; Paller, 1997). These functions do not conflict with each other, and for purposes of the computational model to be developed here, I will assume CA3 fulfills both functions.

Although physical proximity of place cells in the hippocampus does not imply physical proximity of their place fields (O'Keefe, 1976; O'Keefe & Speakman, 1987; Gothard et al., 1996), the establishment of sequences of place cells through the traversal of physical trajectories does establish a mapping of relative distance between physical locations in the hippocampus for a given environment. This could enable novel trajectories to be generated in the hippocampus, through some mechanism such as gradient descent toward locations that have been previously identified as rewarding. This is consistent with the hippocampal hypothesis of the origin of preplay trajectories.

There is some evidence, however, that SWRs containing preplay trajectories may occur prior to an animal's first physical exploration of an environment (Dragoi & Tonegawa, 2011, 2013; Ólafsdóttir et al., 2015), (but see Silva, Feng, and Fos-

ter (2015) for a suggestion that these results have been incorrectly interpreted, and therefore that physical experience is required for expression of a place field). Under the spatial attractor model, these reports are not consistent with the hippocampal hypothesis of a hippocampal origin for preplay trajectories. Entry into a new environment should cause global remapping, driven by allocation of a new chart in MEC. The new place cells corresponding to this chart cannot be predicted endogenously within the hippocampus – they are allocated by a process external to the hippocampus, solely by virtue of their connectivity with the grid cells of the new chart. These results are consistent, on the other hand, with the simulated motion hypothesis. Simulated motion in the motor cortex might cause the active location on the new chart to move to chart coordinates that correspond to areas of the environment that have not yet been physically explored. This would cause activation of the connected place cells for the first time, generating a trajectory prior to physical exploration.

The hypothesis of simulated motion implies that at least some active SWRs do not originate in the hippocampus, but are triggered by processes in motor areas, in turn updating medial entorhinal cortex and then hippocampus. This presupposes a communication mechanism between these areas, i.e., activity at the same frequency and phase across areas (Fries, 2005). Specifically, we should look for assemblies oscillating at the same frequency and phase at the first step upstream from the hippocampus, in medial entorhinal cortex. The research on SWRs largely contradicts this supposition, instead suggesting that SWRs originate in the hippocampus (Buzsáki et al., 1983; Buzsáki, 1986; Buzsáki et al., 1992; Chrobak & Buzsáki, 1994, 1996) (but with some small counterevidence in Chrobak and Buzsáki (1996), and see Norimoto, Matsumoto, Miyawaki, Matsuki, and Ikegaya (2013)). Although the evidence is equivocal that there is communication from EC to CA3 in the high frequency range of SWRs, perhaps the communication is possible at another frequency range. Colgin et al. (2009) found two distinct frequency ranges that appear important in communication between the entorhinal cortex and the hippocampus, and within the hippocampus. "Fast gamma", roughly 65-140 Hz, has a significant level of synchronization between MEC layer 3 and CA1. There is also some synchronization in these frequency ranges between CA3 and CA1, which the authors speculate may be due to synchronization between layers 2 and 3 in MEC, as well as outputs from layer 2 to CA3. "Slow gamma", roughly 25-50 Hz, shows a significant level of synchronization between CA3 and CA1. Pfeiffer and Foster (2015) studied the timing of simulated

movement within quiet awake SWRs recorded in CA1, and found that movement occurred at cycles of roughly 30 to 40 Hz, within the slow gamma frequency range. This suggests a testable prediction, that recording from MEC layer 2 and CA3 simultaneously during awake SWRs should show signs of coherent communication between the areas, possibly at the SWR frequencies, or perhaps more likely, at low gamma frequencies. Similar communication might be expected during active exploration, especially during periods of vicarious trial and error.

Pfeiffer and Foster experiment

In the experiment that is the subject of Pfeiffer and Foster (2013), rats learned to leave a home location to forage randomly for a reward, and upon finding the reward, to then return to the home location to receive a second reward. The rats' performance in finding the second reward at the home location was significantly faster than their performance in finding the first reward at the random location. Monitoring was done in dorsal hippocampal CA1 using 40 tetrodes and recording 100-250 units. During quiet periods, while the rat was consuming a reward, the recordings showed the generation of trajectory events, "temporally compressed, two-dimensional trajectories across the environment". Trajectory events coincided with sharp-wave ripple patterns in local field potentials. Trajectory events averaged 104 ms in length, and corresponded to physical path lengths of 40-200 cm. There were from 144-373 trajectory events per session. These trajectories significantly predicted the rats subsequent movements, both when rats were at the home location and when they were away from the home location. The over-representation of the home goal while the rat was at an away location is suggestive of some planning mechanism. To quantify this effect, the angular displacement was calculated between the trajectory and the animal's immediate future path and immediate previous path. The angular displacement between the trajectory and both the future path and the previous path was clustered around zero, but the distribution was much more tightly clustered for the future path. Dividing the away events between those that preceded paths to home and those that preceded paths to non-home locations elsewhere in the arena, in both cases the trajectory was significantly more predictive of the future path than of the previous path. This suggests that the trajectory has broad planning significance. The authors looked for, but did not find, significant firing of the place cells representing the home location prior to the trajectory events (Pfeiffer, 2015). An interpretation of this might be that the

trajectory was defined only by a particular angular orientation, rather than a vector consisting of an angular orientation and a defined displacement.

A basic assumption of simulation is that the simulated action results in the animal having more information about some pending action or decision. The results of Pfeiffer and Foster (2013) show that the animals generated on the order of twice as many preplay trajectories ending at the home location as at any other location. This suggests that the representation of the home location was associated with possible reward. If the home location is already associated with a reward, what might be the adaptive value of simulating motion toward home? The experimental design is a type of spatial alternation task. In spatial alternation tasks, a given location is not uniformly rewarding. For example, a T-maze might have rewards alternating between the right and left arms. Frank, Brown, and Wilson (2000); Wood, Dudchenko, Robitsek, and Eichenbaum (2000) tested the activity of place cells representing the fixed portion of a path in spatial alternation tasks (the stem of a T-maze, in this example). They found that the firing rates in the common portions of the animal's path were correlated with the subsequent action of the animal. This suggests that under some circumstances, portions of an animal's physical trajectory may represent aspects of its future action. This has been termed "prospective coding", and is a form of rate remapping, where firing rates for place cells are modified by information other than physical location. Perhaps a mechanism analogous to prospective coding is also associated with simulated motion.

The "target trajectory" hypothesis suggests that the process of mentally traversing a sequence of place fields provides the animal with additional information over and above the information associated with the target location itself. The alternative is the "target location" hypothesis, which suggests that the reward information is associated only with the target location, and the simulated motion provides no additional information.

I suggest two hypothetical mechanisms that might support the target trajectory hypothesis. I am not aware of evidence in support of either mechanism. First, directly analogous to prospective coding during physical movement, perhaps there are differing firing rates within an SWR for place cells representing portions of a path that are fixed across multiple choices. A second, very different mechanism might suppose that an entire sequence of place fields is associated with information in other cortical areas. In the case of the Pfeiffer and Foster experiment, the relevant aspect of the
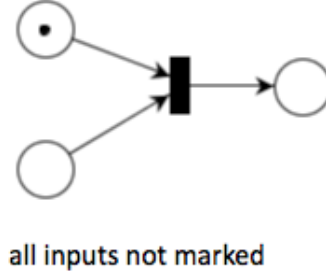
all inputs not marked

Figure 1. Disabled transition

sequence would be whether it begins at a non-home location that was rewarding. If so, a sequence that ended at the home location could be predicted to be rewarding. On the other hand, a home-directed trajectory that did not begin from a rewarding non-home location could be predicted not to be rewarding.

Petri nets

The Petri net formalism models distributed processing as a set of states (or "places"[1]) and transitions among them, connected by directed arcs. Transitions fire when enabled, that is, when the pre-conditions defined by their input arcs are met, generating updates to the places connected to their output arcs. (See Figure 1 for a disabled transition; Figure 2 shows an enabled transition.) The pre- and post-conditions take the form of one or more tokens in a place ("marking"). Input arcs are of two types. Normal arcs have integer-valued weights of one or more, and consume or produce tokens in the number of their weights. Inhibitory arcs prevent the firing of a transition, when their associated places are marked (see Figure 3). Transition firing is stochastic; if multiple transitions are enabled, one is selected at random to fire.

The basic Petri net formalism matches well to neural mechanisms. A Petri net place and its inputs and outputs can be interpreted as a single neuron, or as a population node, which is the interpretation in this model. Places can accumulate tokens, similar to a node accumulating input activations. The firing of a node is modeled explicitly as a transition between states. An arbitrary number of inputs to

---

[1]This should not be confused with places in the sense of physical locations, as generally used in this work.
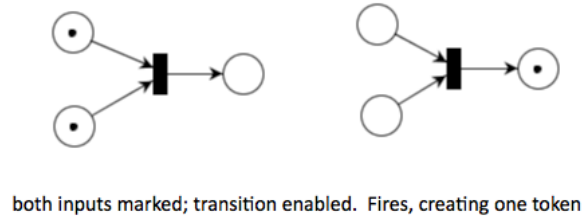
both inputs marked; transition enabled. Fires, creating one token

Figure 2. When all input places are marked, transition is enabled and fires, marking output place

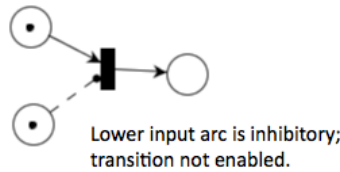Lower input arc is inhibitory; transition not enabled.

Figure 3. Transition inhibited if input place for inhibitory arc is marked.

and outputs from a node's firing can be represented by arcs from and to places. The strength of a connection between two nodes can be modeled as the weights of the arcs between them and an intervening transition; these weights govern the number of tokens to be consumed by firing.

There are areas in which Petri nets diverge from neural mechanisms. Unlike more realistic neural network models, inhibitory arcs are binary; their inputs are not summed. If the place driving the arc is populated, the target transition is prevented from firing, regardless of the status of the excitatory inputs. Also, in the implementation used here, Petri net places accumulate integer token counts; continuous token quantities are not supported. Neither of these problems is a major concern for SMS; the primary use of Petri nets is as control mechanisms, implementing simple logic functions. For components where what is required is modeling of a neural network whose outputs are defined by the interaction of many weighted connections, Matlab networks will be used. Another area where Petri nets do not have a direct mapping to neural processes is that time is not explicitly modeled in Petri nets. The complex temporal dynamics of neural systems, e,g., sub-threshold activations, are therefore not easy to model. Again, that is not a major concern here. Time is measured discretely

16

in SMS, and the interaction between components is not affected by sub-threshold dynamics.

Petri nets serve as a useful intermediate-level abstraction for the representation of actions. By including one Petri net within another, we can model complex actions as a composition of simpler actions. Because there is no central point of control for Petri net execution, we can model asynchronous interrupts and concurrent processing, both useful concepts in modeling cognitive processes.

Testable neuroscientific predictions

The hypothesis of simulated motion makes several neuroscientific predictions. Testing these predictions is outside the scope of this dissertation. The significance of each prediction is discussed in italics; in some cases, if subsequent experimental results do not support the predictions, it would cast serious doubt on the viability of the simulated motion hypothesis.

1. Recording from MEC and LEC layer 2, and CA3, simultaneously during quiet awake SWRs should show evidence of coherent communication between the areas, possibly at the SWR frequencies, or perhaps more likely, at low gamma frequencies. *The simulated motion theory suggests that the place field activations in CA3 and CA1 are due to activations of grid cell populations on the active chart in MEC, and activation of populations in LEC representing other dimensions, both of which necessitate communication between the entorhinal cortex and hippocampus during quiet awake SWRs. Failure to find evidence of such communication would constitute falsification of the simulated motion theory.*

2. In motor cortex, the populations that represent the motor routine for movement should be activated during the trajectory event in the hippocampus. *Under the simulated motion theory, communication passes from motor cortex to entorhinal cortex to hippocampus. Coherent communication between all these areas is a necessary component of the theory, and as above, failure to find evidence of such communication would constitute falsification of the simulated motion theory.*

3. The trajectory communicated between EC and CA3 during active SWRs and during vicarious trial and error during exploration is communicated at the same low gamma frequency (25-50 Hz). *It is parsimonious, but not absolutely essential, to suggest that simulated motion takes place at one rate within the motor*

system, and is accommodated within the hippocampal and retrohippocampal formations according to the current behavioral mode (theta or LIA).

4. For quiet awake SWRs during spatial alternation tasks, it might be possible to predict the direction of the trajectory after the choice point, from the firing rates of the trajectory on the common path prior to the choice point. *Trajectories generated from physical movement encode retrospective and prospective information about the task (Frank et al., 2000; Wood et al., 2000); if such information is also encoded in quiet awake SWRs, it means that simulated motion can be more predictive of reward than just a representation of the target location itself. Note that this is not the mechanism used in the SMS model, but is a plausible alternative and is consistent with the overall simulated motion theory.*

5. Placement of a rest box in the center of a larger open arena without salient cues visible from the rest box, and with which the animal has had no previous experience, should lead to a set of trajectories that covers the arena in a roughly uniform distribution, and whose place fields are re-expressed on subsequent physical exploration. *Expanding on the design of Dragoi and Tonegawa (2013), this does not distinguish between the endogenous allocation of place fields that the authors hypothesize, from the simulated motion hypothesis, but the anticipated results would show the continued viability of the simulated motion theory as an alternative account.*

6. It is possible that significantly extended training of the experimental animals would show that they develop habitual responses that obviate the hypothesized advantage of simulated motion, as discussed below under the research plan for Aim #2. This would manifest itself as a lowered frequency of trajectories during quiet awake SWRs that predict the animal's next physical motions. Such a pattern might become more clear if the experimental design were modified to lower the opportunity for simulated motion, while maintaining the overall reward structure. An audible cue could be presented whenever the animal reaches the designated away location, indicating that returning home would now be rewarding. The away location might only be intermittently rewarding. When a reward is present, there would be an opportunity for quiet awake SWRs, as in the current design. When a reward is not present, the cue would be a signal that immediately returning to the home location would now be rewarding. An animal

that has reached habitual mastery of the task might then return to the home location without any quiet awake SWRs, and therefore, without the assistance of simulated motion. Note that absent quiet awake SWRs, it is possible that the animal might show virtual trial and error behavior prior to mastery at the level of habit. Under the interpretation of VTE adopted here, this would constitute support for the simulated motion hypothesis. *The context for this prediction is discussed under the research plan for Aim #2. Experimental results consistent with this prediction would constitute support for the interpretation of the target location hypothesis as a manifestation of what in Aim #2 is termed "phase 3" of the animal's learning. Failure to observe such results might mean that the nature of the experimental task is too complex for the animal to develop habitual mechanisms to master it; this would not mean that the mechanisms proposed under the Aim #2 research plan might not apply in other simpler experimental designs.*

Neuroscientific constraints on a computational model

From the chapter on spatial navigation, here are a set of constraints on the software model. The rationale for each is summarized in italics. In some cases, the software model is sufficiently simplified that the constraint is not applicable.

1. SWR trajectories during awake quiet periods most frequently predict the rat's next physical path. *The Pfeiffer and Foster (2013) results showed the trajectories most frequently predicting the next physical paths, with no systematic bias from an angular displacement of zero.*

2. At a lower level of frequency, trajectories during awake quiet periods may represent the rat's immediately previous path. *Replay is hypothesized as a mechanism to update neocortex with what has been learned.*

3. The mechanism(s) leading to temporally compressed trajectories observable in hippocampus should be compatible with different trajectory functions, which may differ as to origination point (inside vs. outside hippcampus) and overall behavioral state (SWS vs. REM vs. quiet awake vs. vicarious trial and error during active navigation). *The function performed should drive the mechanism that creates the trajectory.*

4. During periods of physical activity, especially vicarious trial and error, trajectories are generated that represent future movement. *Results from Johnson and Redish (2007) showed representation of alternate paths in T-maze.*

5. The parameter to simulated motion may be a vector, but may also just be an orientation, in which case the animal simulates movement in a particular direction. *Pfeiffer (2015) did not find representation of the goal location prior to trajectories during quiet awake SWRs.*

6. Simulated motion can cover an indefinite distance. *Davidson et al. (2009) found that a long trajectory was composed of multiple SWRs, and that total distance was linearly related to elapsed time.*

7. The direction of simulated motion should be consistent with simulated head direction. *Peyrache, Lacroix, Petersen, and Buzsaki (2015) found evidence of head direction changes during quiet awake SWRs. Head direction is a component of the spatial attractor model of path integration that I propose is used during simulated motion.*

Methods

Computational Model

The computational model to be developed in support of the proposed theory of spatial navigation during simulated motion consists of a system of software components modeling various neural subsystems, driven by a model of rat behavior implemented as Petri nets. This "Simulated Motion System" (SMS) is implemented in Matlab, Petri nets, and Java, and is maintained as open source: https://github.com/sjdayday/simulated-motion. The model simulates an animal's behavior in circumstances similar to those of the Pfeiffer & Foster experiment. The data flow of the system is depicted in Figure 4. The neural subsystems modeled in most detail are those in the hippocampal formation; the data flow in this area is depicted in Figure 5.

Components. The components of the system are modeled at a minimal level of detail, to keep the system tractable. The component name is meant to be suggestive of functions of one or a group of related neural subsystems. Components within the hippocampal formation are modeled in more detail, closer to their neural counterparts
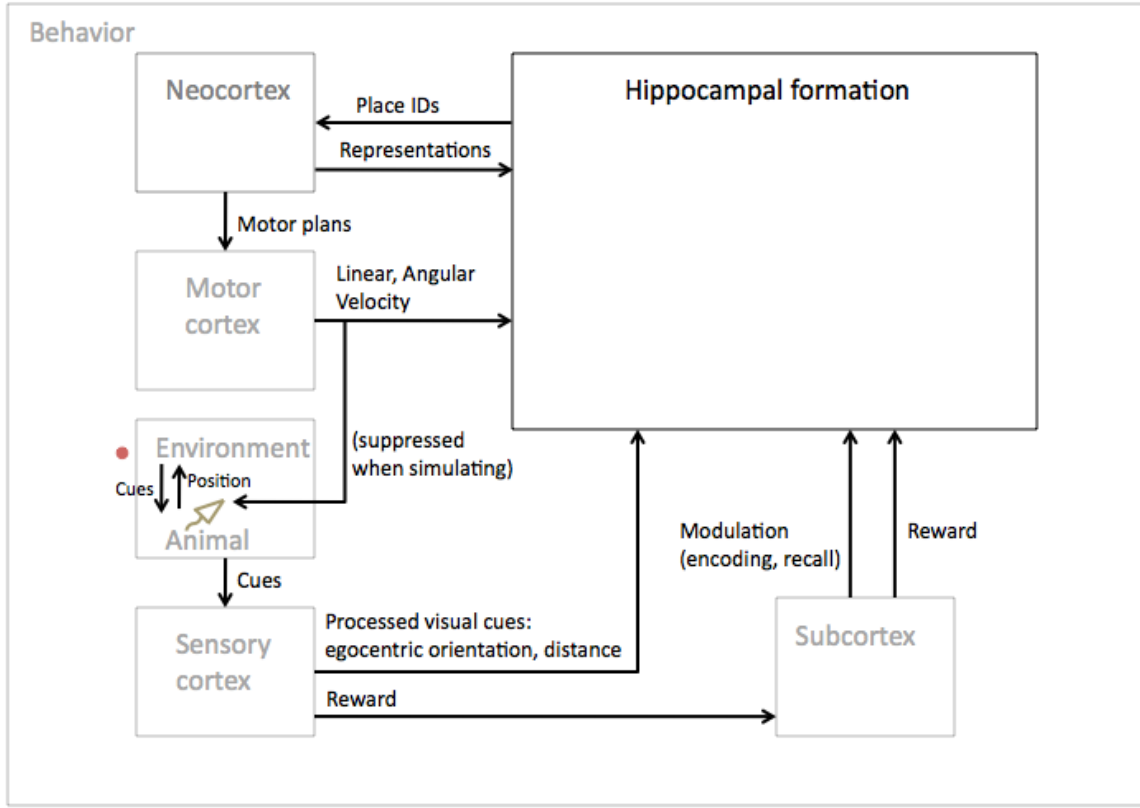
Figure 4. System data flow. (Areas in gray modeled in minimal detail.)

of the entorhinal cortex and hippocampus. The Environment and Animal components represent the physical arena and the physical rat behaving within that arena; the remaining components represent some aspect of neural functioning. The components are implemented in Matlab; the overall behavior of the system is driven by Petri nets, and Java code links the Petri nets with the Matlab components.

Each component has a number of parameters. A subset of these parameters are important in determining how the component behaves. No attempt is planned to find some optimal set of parameter values, beyond a set that enables the basic behaviors of the overall system to be validated. Nothing in principle prevents a machine learning process from adjusting the parameter values, but the theory of simulated motion being modeled does not define an objective function to optimize.

Environment: The Environment represents the physical arena within which the Animal moves. It consists of a set of walls, visual cues, and potentially rewarding
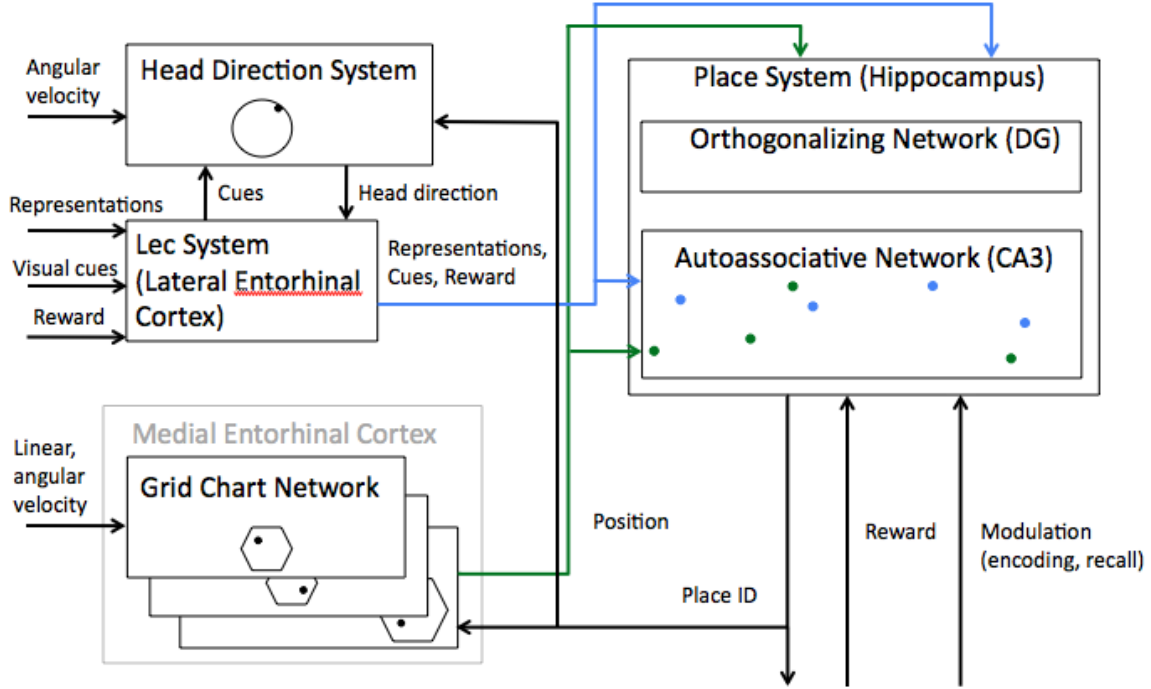
Figure 5. Data flow within the hippocampal formation

places. The Environment calculates the actual distance and angle from the animal to walls and cues, and passes this information to the Animal.

Animal: The Animal represents the physical rat, moving within the Environment. The Animal knows its position within the Environment, and updates that position as a result of the execution of motor plans. The Animal receives information from the Environment as to the distance and orientation to walls and cues.

Neocortex: The Neocortex generates motor plans to attain its goals of exploring the current environment or finding rewards. A motor plan is a combination of turns and straight runs in the Environment, away from the current place and orientation, for a distance or until arrival at a target place. Based on the experience accumulated through the animal's activity, the Neocortex updates representations of the current situation within a neural network, and uses these as input to the process of deciding on its next motor plan. The currently active representation is also sent to the HippocampalFormation. The Neocortex receives Place IDs from the Hippocampal Formation, which are associated within the Neocortex with the current representa-

tion, the current place, and the outcome of the current motor plan. Details of the representation mechanism are discussed under the Methods for Aim 2. Parameters include: number of nodes in the neural network, relative cost of simulated motion and physical motion, number of simulated trajectories per quiet awake period.

MotorCortex: The MotorCortex executes motor plans it receives from the Neocortex. In each time period, it executes a portion of a turn or a run. The MotorCortex has two modes of operation. In normal mode, the linear or angular velocity from the current time period is sent to the Animal, resulting in a change in the Animal's physical position. In simulated mode, the velocity signal is suppressed, and the Animal does not move. In both modes, however, the linear or angular velocity is sent to the Hippocampal Formation, analogous to "efferent copy". Parameters include: granularity of motion per time period (radians of rotation, units of straight distance).

SensoryCortex: The SensoryCortex receives updated visual cue orientation and distance information from the Animal, and converts these to a vector of egocentric orientation and distance cue tuples; the first cue in the vector is the most salient cue. One of the cues available from the Environment is initially designated as the most salient cue; this designation does not change. The vector of cues is output to the Hippocampal Formation. Parameters include: amount of random error to be introduced into the vector of cues, number of cues.

SubCortex: The SubCortex receives information from the SensoryCortex as to whether a reward is present at the current location. The SubCortex outputs the presence or absence of reward to the Hippocampal Formation. The SubCortex also modulates the operation of the Hippocampal Formation into two modes. "Recall" mode enables the attempted retrieval of pre-existing hippocampal memories of a place. "Encoding" mode enables hippocampal memory to be updated with new locations when memory place retrieval has failed. The modulation mode is driven by an interaction between the SubCortex and the behavioral Petri nets.

HippocampalFormation: The Hippocampal Formation is an umbrella component that corresponds to the entorhinal cortex and hippocampus. It sends and receives information to and from the other system components and distributes it appropriately among the following components: HeadDirectionSystem, LecSystem, multiple GridChartNetworks, and the PlaceSystem. Parameters include: number of GridChartNetworks, number of different orientations represented across the GridChartNetworks, number of different spatial gains represented across the GridChartNetworks.

HeadDirectionSystem: The HeadDirectionSystem receives an angular velocity signal from the MotorCortex and maintains a representation of the Animal's orientation in relation to the most salient cue in the Environment. The system implements continuous attractor dynamics for a single dimension as in Trappenberg (2010), following the head direction model from Skaggs et al. (1995). The current head direction is output to the LecSystem. The current cues from the LecSystem and the Place ID from the PlaceSystem are associated to the current head direction. This supports retrieving a known orientation upon re-entry to a previously explored Environment. Parameters include: number of nodes representing possible directions.

GridChartNetwork: The GridChartNetwork receives linear and angular velocity signals from the MotorCortex and maintains a representation of the Animal's position on a "chart" of the environment, configured as a torus. The system implements continuous attractor dynamics for two dimensions as in Trappenberg (2010), roughly following the model from Samsonovich and McNaughton (1997). The activation pattern takes the form of a hexagonal grid, as found in the medial entorhinal cortex. The model of grid activation is adapted from Guanella, Kiper, and Verschure (2007), as implemented by Zilli (2012). The Guanella model makes a simplifying assumption that is not biologically plausible, in that the weights that drive the attractor dynamics are recalculated in each time step and directly include the velocity signal. As a consequence of this simplifying assumption, however, the model generates repeatable grid positions with low error rates, which makes it sufficient for the purposes of this system. The current grid position is output to the PlaceSystem. Similarly to the HeadDirectionSystem, the Place ID that is output from the PlaceSystem is associated with the current grid position. This supports the later retrieval of a known position upon re-entry to a previously explored Environment. In this implementation, there is no single component that corresponds to the medial entorhinal cortex (MEC); rather, there are multiple GridChartNetworks, and their combined positional outputs drive activation of an approximately unique place code in the PlaceSystem. Parameters include: number of cells comprising the network, Place ID learning rate, number of bits to represent a signal to the PlaceSystem, amount of random error to be introduced into the signals sent to the PlaceSystem.

LecSystem: The LecSystem models some aspects of the lateral entorhinal cortex (LEC), providing both spatial and non-spatial inputs to the PlaceSystem. The LecSystem receives the vector of cues from the SensoryCortex, the representation of

the current situation from the Neocortex, and the signal as to whether the current place is rewarding from SubCortex. After processing, this information is then passed on to the PlaceSystem. The vector of cues is also passed to the HeadDirectionSystem, which returns the current head direction. This is used by the LecSystem to create a "canonical" view of the cues. For purposes of creating a single representation of a place within the PlaceSystem, the animal's current head direction is rotated to orient directly towards the most salient cue, and all other cues are rotated accordingly. Creation of the canonical view is the only processing performed by the LecSystem. This processing supports the creation of a representation in the PlaceSystem that encodes the distance and orientation to a cue. Although such processing is neurologically plausible, it is likely that multiple areas are involved, and it may be that LEC only receives the output of processing that happens elsewhere. Parameters include: number of bits to represent a signal to the PlaceSystem, amount of random error to be introduced into the signals sent to the PlaceSystem.

PlaceSystem: The PlaceSystem is an umbrella component that models some functions of the hippocampus. It is comprised of two components, the OrthogonalizingNetwork and the AutoassociativeNetwork, corresponding roughly to the dentate gyrus and CA3. For purposes of this system, separate models of CA1 and subiculum are not needed. The PlaceSystem sends and receives information to and from other components in the Hippocampal Formation: HeadDirectionSystem, LecSystem, and the GridChartNetworks. It also receives information from the SubCortex. Its only other processing is to initially connect the OrthogonalizingNetwork and the AutoassociativeNetwork together. An important simplifying assumption is that the PlaceSystem only maintains binary representations; all information is encoded as either one or zero. The interpretation is that a given population node is either active or inactive, and this implies that the model cannot accommodate sub-threshold activation. Parameters include: size of the OrthogonalizingNetwork, size of the AutoassociativeNetwork.

OrthogonalizingNetwork: The OrthogonalizingNetwork implements one aspect of the processing of a model of the dentate gyrus, to orthogonalize inputs from the LecSystem and the GridChartNetworks. For example, when representing adjacent physical locations, lower error rates will result if the overlap in bits representing the locations is minimized. The OrthogonalizingNetwork establishes a random but stable connection between the input vectors received from the LecSystem and the Grid-

ChartNetworks, and the output vector to the AutoassociativeNetwork. As presently implemented, the OrthogonalizingNetwork performs no processing to make the output representation more sparse than the inputs, and does no learning. Parameters include: size of the input vectors, size of the output vectors.

AutoassociativeNetwork: The AutoassociativeNetwork models some aspects of CA3, interpreted as a Hebb-Marr autoassociative network, as defined in McNaughton and Nadel (1990). The inputs from the OrthogonalizingNetwork are treated analogously to "detonator" synapses, i.e., their firing causes the nodes to which they are connected in the AutoassociativeNetwork to fire with probability one. These inputs from the OrthogonalizingNetwork originated in the LecSystem and GridChartNetworks. Analogously to the perforant path connections from entorhinal cortex to hippocampus, the same inputs are also directly sent to the AutoassociativeNetwork (see Figure 5). On this second path the role of these inputs is equivalent to projections to the "modifiable" synapses of the Hebb-Marr model, where they undergo Hebbian learning when firing in conjunction with the firing of inputs from the OrthogonalizingNetwork pathway. As initially implemented, these synapses are given a weight of one upon first firing, and do not decay. The AutoassociativeNetwork operates in two modes, driven by modulation from the SubCortex. When encoding, both the detonator and modifiable pathways are active, strengthening synaptic connections. When recalling, the detonator pathway is suppressed, and only the signal along the modifiable pathway is active. At this point, the third synaptic type in the Hebb-Marr model becomes involved, that of inhibitory synapses, to control the output pattern that is retrieved from the AutoassociativeNetwork. The inhibitory synapses implement the equivalent of integer division, ensuring that the only output nodes in the AutoassociativeNetwork that fire are those where the number of active inputs is at least equal to the number of strengthened synapses of the original input pattern. This constitutes an implementation of point attractor dynamics. If no pattern is retrieved, the inhibition level is lowered on successive attempts, until the inputs match some subset of a stored pattern. If no pattern is ever retrieved, this serves as a signal of novelty, which causes the overall system to switch to encoding mode. Any retrieved pattern is interpreted as a Place ID. Following memory index theory, the Place ID is output to components in the HippocampalFormation and other cortical areas, where it is stored in association to current activation patterns. Parameters include: size of

the network, minimum number of bits that must be active to constitute successful pattern retrieval.

Petri nets: Petri nets are used to provide a simple model of the Animal's behavior, and to model key events and processes of the overall system. The implementation of Petri nets makes use of prior work, which enabled Petri nets to be composed of other Petri nets in a hierarchy, and to interface to external systems, as documented here: https://github.com/sjdayday/PIPECore/wiki. The source code for that work is here: https://github.com/sarahtattersall/PIPECore/tree/hierarchical-nets. The Petri nets include:

– *placed*: models the orienting behavior of the animal when initially placed into an environment, resulting in a determination that the animal is either in a known place or an unknown place. Invokes *orient*, and sets modulation to recall mode (Figure 6).

– *orient*: models turning through a complete rotation in the current place. Invokes *move*, with parameters Turn, Distance, Speed and Clockwise (Figure 7).

– *move*: models motion, either turning or running. Accepts parameters Distance, Speed, Clockwise or CounterClockwise, Run or Turn, and Simulated, and invokes *run* or *move* (Figure 8).

– *run*: models running straight ahead until Distance is reached, or until an object is sensed. Accepts parameters Distance and Speed (Figure 9).

– *turn*: models turning either Clockwise or CounterClockwise until Distance is reached, or until an object is sensed. Accepts parameters Distance, Speed, Clockwise or CounterClockwise (Figure 10).

– *quiet*: models navigation based on a plan provided by Neocortex. The plan consists of a set of Moves from a known or unknown place to a different place, which will be recognized as either known or unknown, at which point navigation is repeated, and modulation is set to recall mode ("PlaceReadMode") or encoding mode "PlaceUpdateMode"). When energy is depleted or a reward is found, enters a quiet awake state where navigation is simulated, but no physical motion takes place. Invokes *move*, with parameters Simulated, Turn or Run, Distance, Speed, and Clockwise or Clockwise (Figure 11).
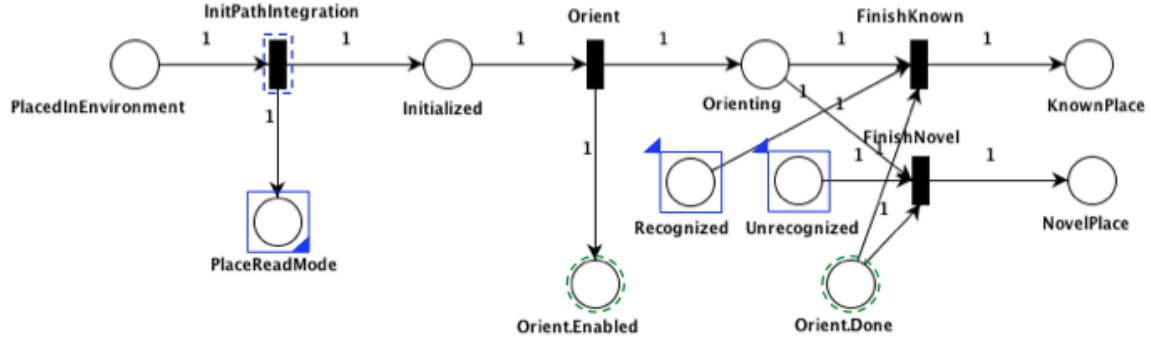
Figure 6. Placed in environment Petri net

The Petri nets interface to the neural components described previously through a combination of ExternalTransitions, and input or output external places. An ExternalTransition is a type of Petri net transition that causes invocation of a Java class, which in turn communicates with an arbitrary external system. An input external place can be marked by an external system; an output external place can be read by an external system. The external system in these cases will be the Matlab classes that implement the neural components.

Model Assumptions

There are number of important simplifying assumptions made in SMS whose implications should be made explicit.

– States in CA3 are modeled in the AutoassociativeNetwork component as binary; an individual node is activated or not activated. This means that concepts that depend on sub-threshold activation cannot be modeled.

– The only output from the PlaceSystem is a place ID, which is stored by other cortical areas in association with the current local cortical state, so that local state may be retrieved later. This is an implementation of memory index theory. One immediate function this enables is continuous attractor dynamics in the HeadDirectionSystem and the GridChartNetworks. Although hippocampal outputs might drive other processing, that is not provided for here.
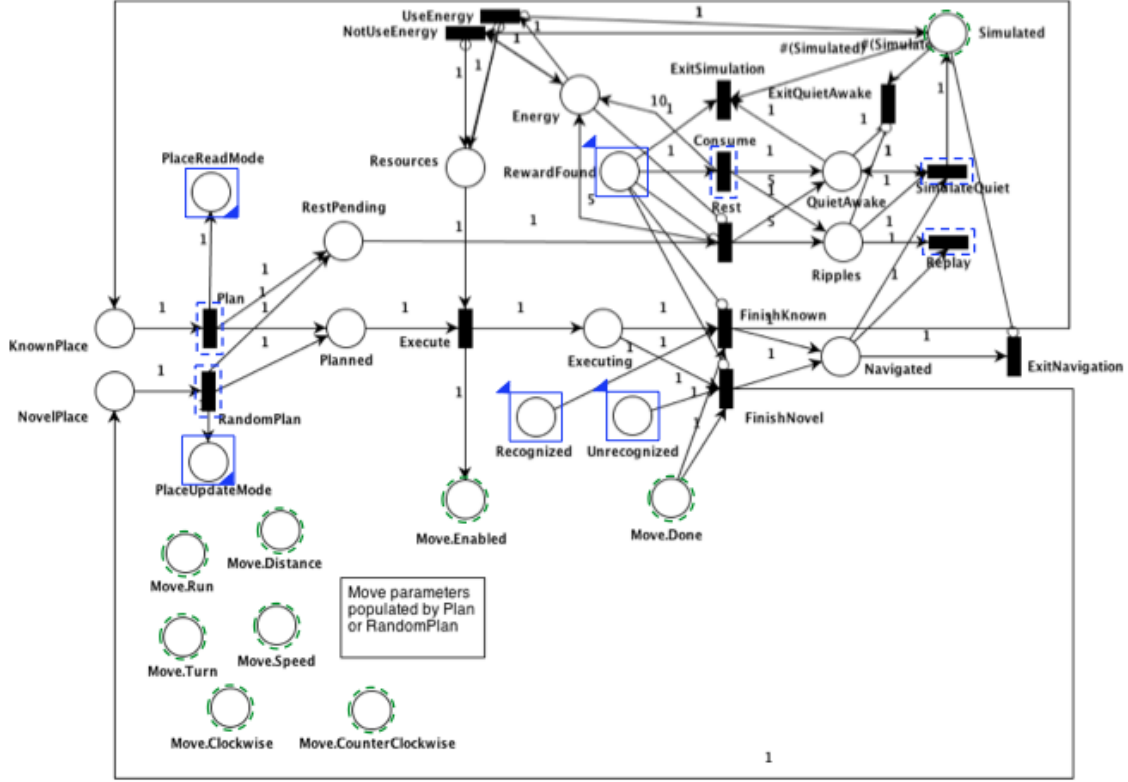
Figure 7. Orient Petri net

Figure 8. Move Petri net

Figure 9. Run Petri net

Figure 10. Turn Petri net

Figure 11. Navigation Petri net, both physical and simulated during quiet awake periods.

– As currently implemented, motion is not continuous. Rather, I assume that
the animal moves for multiple time periods, and then pauses for multiple time
periods. This has the advantage of enabling the inputs and outputs across time
periods to be associated with low error rates. For example, the inputs to the
PlaceSystem from GridChartNetworks and the HeadDirectionSystem at time
$t+1$ will be associated with the place ID output at time t. If the animal is in
the same physical location and orientation, this minimizes the amount of error
in the association. This assumption might be supported by observing that rat
physical motion is frequently discontinuous, with small starts and stops. The
evidence of trajectories during quiet awake periods also does not show contin-
uous motion. Pfeiffer and Foster (2015) suggest the observed discontinuities in
SWR trajectories might have the effect of lowering the error in autoassociative
dynamics, which suggests a relation my assumption. Finally, this assumption
plays an analogous role in SMS to that hypothesized to be played by a visual

saccade (although not in rats) in resetting the visual system such that there is a clear delineation between two successive images.

– As discussed in the description of the LecSystem, I assume that the PlaceSystem stores a canonical local view, consisting of orientation and distance to a cue, associated with a given place field. The canonical local view assumes that the animal's head direction is oriented directly towards the most salient cue, and all other cues are rotated accordingly. This has the advantage of eliminating duplicative information in the place system, representing all possible orientations toward cues. It also supports a notion of object constancy, such that the animal recognizes that there is a single object, not a different object for each possible orientation and distance. This assumption is supported by findings from Navratilova et al. (2012) that in an arena, place fields appear not to have an orientation, i.e., firing rates are similar in all directions. In a linear track environment, differential firing rates develop over time from the initial uniform pattern, suggesting that the two directions develop an independent salience.

– I assume that some neuromodulatory process drives two different modes of processing in the PlaceSystem:

   – *encoding*: novel places are stored, driven by detonator synapses (DG granule cells), and associated to EC inputs

   – *recall*: known places are retrieved from EC inputs; firing of detonator synapses is suppressed.

Model Validation

The behavior of the overall system will be validated by its ability to simulate several typical behaviors and their neural correlates. The principal interactions between components that produce each behavior will be described. This will provide the foundation for explaining how the specific aims will be accomplished.

Repeatable Place Fields. Upon placement in a novel environment the simulated animal will explore randomly. As each place is encountered a unique place ID will be generated. The correspondence between this place ID and its place field (the physical location in the environment), will be noted for later testing. If the animal is removed from the environment and then placed again in the environment, it will orient itself

and then recognize that the environment is known. This recognition can be verified by comparing the place ID that is activated at each place in the environment, and comparing it to the animal's physical location; the discrepancies observed should be small.

Component interaction sequence for creation of novel place fields:

1. Novel locations cause the SubCortex to modulate the HippocampalFormation to encoding mode.

2. As each novel physical location is entered, visual cue information flows from the SensoryCortex to the LecSystem. Also, self-motion cues (linear and angular velocity) flow from the MotorCortex to the HeadDirectionSystem and the GridChartNetworks.

3. The LecSystem retrieves the current head direction from the HeadDirectionSystem, and calculates the canonical view of the visual cues. Each GridChartNetwork calculates the animal's new relative position on its chart.

4. The vector of visual cue information and the vector of position information are input to the PlaceSystem. The inputs are orthogonalized by the OrthogonalizingNetwork and output to the AutoassociativeNetwork. The orthogonalized inputs are then associated with the original input vectors on the direct path to the AutoassociativeNetwork, causing the activation of the output vector of the AutoassociativeNetwork, which is the Place ID.

5. The Place ID is sent back to the GridChartNetworks and the HeadDirectionSystem, where it is associated with the original inputs.

Component interaction sequence for placement in a known environment:

1. Upon placement in an environment, the HeadDirectionSystem and GridChartNetworks are randomly initialized, and the SubCortex modulates the HippocampalFormation to recall mode.

2. The animal orients itself by turning in a circle. Visual cue information enters the SensoryCortex.

3. Visual cue information flows from the SensoryCortex to the LecSystem. Also, self-motion cues (angular velocity) flow from the MotorCortex to the HeadDirectionSystem and the GridChartNetworks.

4. The LecSystem retrieves the current head direction from the HeadDirectionSystem, and calculates the canonical view of the visual cues. Each GridChartNetwork calculates the animal's new relative position on its chart.

5. The vector of visual cue information and the vector of position information are input to the PlaceSystem. Processing of inputs to the OrthogonalizingNetwork is suppressed and generates no output to the AutoassociativeNetwork. The input vectors on the direct path to the AutoassociativeNetwork are used to attempt to recall a Place ID using the point attractor dynamics. The visual cue inputs should match previous visual cue inputs, but the position inputs will vary randomly from what was previously stored for this physical location, causing the network to go through multiple internal iterations of successively lowered inhibition levels to retrieve the Place ID. The retrieval process demonstrates the pattern completion capability of an autoassociative network.

6. The retrieved Place ID is sent back to the GridChartNetworks and the HeadDirectionSystem. Its prior association with a head direction in the HeadDirectionSystem and a position in each GridChartNetwork will drive the continuous attractor dynamics of those networks, causing their activation "bumps" to move to the previously stored locations in the networks. This process may take several cycles. At its completion, the HippocampalFormation state is consistent with its state when the place field was originally established; the animal has "remembered" where it is.

7. As the animal moves to other known locations in the environment, the discrepancy between the place fields of retrieved Place IDs and the animal's physical location should be low.

Global Remapping and Rate Remapping. If the animal is moved from a known environment to a novel environment, there should be no relationship between the place IDs of the two environments. This is equivalent to global remapping. If cues are changed significantly in a known environment, the existing place fields should be

maintained, but the non-positional elements of the place IDs should change. This is equivalent to rate remapping.

The AutoAssociative network uses a binary activation model; a node is active or it is not. In a population coding model, the concept of rate remapping could be approximated by differing levels of activation over some continuum, but the binary model does not support such a continuum. However, it is neurologically plausible that much the same information is available through the binary model, interpreted as follows. Projections from both MEC and LEC synapse across the dendrites of many hippocampal pyramidal cells. When the detonator synapses from a population of dentate gyrus granule cells fire, the synapses from the MEC and LEC projections will be strengthened. This is captured by the AutoassociativeNetwork model, and as outlined above in the sequence for placement in a known environment, presentation of a vector representing a subset of these strengthened synapses will result in retrieval of the entire place ID. As long as the subset is sufficiently large and the network is not too saturated, these point attractor dynamics work regardless of whether the source of the projection is MEC or LEC.

Consider what is observed experimentally. The pattern of hippocampal firing was puzzling for many years, and the subject of much theorizing, until the suggestion by O'Keefe and Dostrovsky (1971) that much of it could be explained by mapping the firing of some units to locations in the environment. Moving to a different environment caused some units to go silent, and others to become active; this came to be termed global remapping. Changing cues within an existing environment caused the firing rates of previously active units to rise or fall; this came to be termed rate remapping. In light of the theory that physical position in the environment is represented by the activation of nodes in MEC, we can interpret these findings as being focused on units whose activation is driven primarily by projections from MEC, and only secondarily, by projections from LEC. As long as the animal stays in a single environment, the MEC signal will be constant for a given place. If the animal changes environments, the MEC signal will change for a given place; some place cells will receive no MEC input and stop firing while others begin firing. This manifests itself as global remapping. When active nodes in MEC and LEC both synapse on the same cells in CA3, a change in cues may modulate the input from LEC, increasing or decreasing the firing rate of the place cells, but firing is unlikely to drop to zero, as the MEC projections are more

proximal to the cell body and therefore contribute more strongly to a base firing rate. This manifests itself as rate remapping.

Note that the account of both global and rate remapping focus on place cells where active MEC inputs are observed. But we can expect that there are other place cells where the only active inputs are from LEC. In CA1, projections from MEC and LEC go to physically separate sub-areas, but even in CA3, we can anticipate that some place cells might be activated solely by virtue of their inputs from LEC. It has proved more difficult to attribute firing patterns in hippocampus to LEC inputs, however, because those inputs are thought to represent multiple dimensions, whose firing patterns do not align so neatly to attributes of the situation as the two-dimensional ordering of place fields. (One exception is the representation of an orientation and distance to a cue (Deshmukh, 2014), which appears to be driven by LEC input, and which is incorporated into the software system design.) But we can imagine a projection from LEC which synapses both on cells with strong MEC inputs, and on cells where only LEC provides strong input. The projection provides the same information in both cases. The first case is more easily observed and interpreted during experiments, as rate remapping. The second case might be interpreted as noise, because its firing is driven only by LEC. For purposes of this model, however, the second case contributes information to the output vector, the place ID, and that information is equivalent to the rate remapping of the first case. Of course, the actual neural system is more complex, and it is quite plausible that different firing rates in the first case have a unique function. For purposes of this model, however, the assumption that the LEC inputs cause separate hippocampal nodes to be activated is an adequate alternate representation of the information conveyed by rate remapping.

Component interaction sequence for global remapping:

1. Upon movement from one environment to a second environment, the inputs from both the LecSystem and the GridChartNetworks to the PlaceSystem differ randomly from their values in the first environment.

2. The PlaceSystem generates or recalls Place IDs in the second environment which differ randomly from those of the first environment. In Figure 5, this would be represented by a different pattern of green dots (MEC-driven nodes) and blue dots (LEC-driven nodes).

Component interaction sequence for rate remapping:

36

1. Upon change to the cue structure within a single known environment at a single place, the inputs from the LecSystem to the PlaceSystem would differ randomly from their previous values, while the inputs from the GridChartNetworks to the PlaceSystem would stay constant.

2. The PlaceSystem generates or recalls Place IDs after the cue structure change where the MEC-driven activation stays constant, but the LEC-driven activation changes. In Figure 5, this would be represented by the same pattern of green dots (MEC-driven nodes) and a different pattern of blue dots (LEC-driven nodes).

Generate Replay Trajectories During Quiet Awake Periods. During periods when the simulated animal is resting or consuming a reward, the animal will generate one or more sequences of place IDs that map to a physically connected set of place fields, a trajectory, that the animal has previously traversed. This replay will be in the order in which the place fields were traversed. Once a set of place fields have been traversed in multiple directions, the replay trajectory may appear to be in either forward or reverse order, or may switch orders in mid-sequence. *This capability and its validation are not absolutely necessary to a minimal implementation of Aim #2, but would support a more full account. It may be dropped if time presses.*

Note that there is no temporal differentiation in the simulated Hippocampal-Formation between physical trajectories and trajectories during quiet awake periods (either replay or simulated motion trajectories). Both consist of a sequence of activations in the AutoassociativeNetwork, and the system does not provide for a way to distinguish them temporally. The only way to distinguish them is to note that physical trajectories are accompanied by movement, and the quiet awake trajectories are not. The simulated system will depict the movements at different rates, but that will be a function of the behavior of MotorCortex, not of the behavior of the HippocampalFormation.

Component interaction sequence for replay trajectories:

1. As a new place ID is created in the AutoassociativeNetwork, it is associated with the place ID from the period before. When the current motor plan is completed or interrupted, e.g., by discovery of a reward, the sequence of associations is terminated.

2. Upon entry into the quiet awake state, SubCortex modulates the HippocampalFormation to recall mode.

3. Upon entry into the quiet awake state, a place ID is retrieved at random from the AutoassociativeNetwork. Any other place ID that has previously been associated with this place ID will then be retrieved, and the process repeats until no further place IDs are retrieved.

4. The retrieved place ID is output to Neocortex to drive update of its representation of the situation. (Note that the output to the HeadDirectionSystem and GridChartNetworks is probably not processed, under the assumption that only physical or simulated motion should change the state of these systems.)

5. Steps 2-4 are repeated some number of times, probably modeled as an integer draw from a Gaussian with a configurable mean in the single digits.

Generate Simulated Motion Trajectories During Quiet Awake Periods. During periods when the simulated animal is resting or consuming a reward, the animal will generate one or more motor plans for simulated motion in the environment, leading to a trajectory in the environment. With probability higher than chance, the subsequent physical trajectory of the animal will be close to the most recent simulated trajectory.

Component interaction sequence for simulated motion trajectories:

1. Upon entry into the quiet awake state, SubCortex modulates the HippocampalFormation to recall mode.

2. Upon entry into the quiet awake state, a motor plan is created at random by Neocortex. Motor cortex receives the motor plan, and a signal to enter simulated mode.

3. Motor cortex sends a copy of the motor plan as a set of angular and linear velocities to the HippocampalFormation, while suppressing the copy that would otherwise be sent to the Animal.

4. The angular and linear velocities are processed by the HeadDirectionSystem and GridChartNetworks, resulting in an update to the current orientations and positions active in those systems. These are output to the PlaceSystem.

5. Visual cues corresponding to the updated positions are not available to the LecSystem, and are therefore not part of the retrieval from the PlaceSystem. This simplification is due to not previously storing the place ID in SensoryCortex in association with the visual cues. In principle, the system could be extended to support such storage and retrieval, but it does not seem critical to the argument, so has not been provided for.

6. The PlaceSystem retrieves a place ID based on the positional inputs, and outputs this to the Neocortex as a hypothesis as to what might be encountered by traveling to the new place.

7. Steps 2-6 are repeated some number of times, probably modeled as an integer draw from a Gaussian with a configurable mean in the single digits.

Research Plan

In this section, I discuss how the Simulated Motion System will be used to accomplish the specific research aims of this work.

Aim #1: Contrast hippocampal and simulated motion hypotheses regarding creation of place fields prior to physical experience

Aim #1 is to assess the support provided by the computational model for for each of two hypotheses regarding the mechanism that generates novel trajectories in the hippocampus during quiet awake periods, prior to physical experience at the locations represented by those trajectories. As noted earlier, there is current debate as to the evidence for representation of locations prior to physical experience, but pending the outcome of that debate, we can use the Simulated Motion System to assess hypotheses regarding the mechanisms that might be at work.

- The **hippocampal** hypothesis, proposed by Dragoi and Tonegawa (2011), suggests that a process within the hippocampus makes populations of cells available for use in representing future locations that have not yet been visited.

- The **simulated motion** hypothesis proposed here is that processes outside the hippocampus drive the generation of trajectories, including novel trajectories. A target location or direction is selected, and simulated motion towards the

target causes updates in the path integration system, thereby activating a new population of place cells in the hippocampus to represent the novel trajectory.

As explained below, by its inherent structure SMS contradicts the hippocampal hypothesis. On the other hand, SMS supports the simulated motion hypothesis with very minor modifications to the capabilities defined earlier. We begin by discussing support for the simulated motion hypothesis.

Support for the Simulated Motion Hypothesis. During the time of initial exploration of a novel environment, the simulated animal may enter the quiet awake state. During such periods, the animal will generate one or more trajectories, some places of which do not yet correspond to any place ID the animal had previously generated. When the animal subsequently moves physically to such a place, the place ID that corresponds to that place should be identical, in its motion-specific components, to the place ID generated during the quiet awake trajectory.

Component interaction sequence for creation of place fields prior to physical experience. Italics represent modifications to processes previously implemented:

1. Upon entry into the quiet awake state, SubCortex modulates the HippocampalFormation to recall mode.

2. Upon entry into the quiet awake state, a motor plan is created at random by Neocortex. Motor cortex receives the motor plan, and a signal to enter simulated mode.

3. Motor cortex sends a copy of the motor plan as a set of angular and linear velocities to the HippocampalFormation, while suppressing the copy that would otherwise be sent to the Animal.

4. The angular and linear velocities are processed by the HeadDirectionSystem and GridChartNetworks, resulting in an update to the current orientations and positions active in those systems. These are output to the PlaceSystem.

5. Visual cues corresponding to the updated positions are not available to the LecSystem, and are therefore not part of the retrieval from the PlaceSystem.

6. *For locations not previously visited, the PlaceSystem fails to retrieve a place ID based on the positional inputs, and the HippocampalFormation enters encoding mode.*

7. *A new place ID is created, driven solely by the inputs from the GridChartNetworks. This would correspond to a place ID in Figure 5 with only green dots (driven solely by MEC).*

8. Steps 2-7 are repeated some number of times, probably modeled as an integer draw from a Gaussian with a configurable mean in the single digits.

Component interaction sequence for retrieval of place fields during subsequent physical experience:

1. SubCortex modulates the HippocampalFormation to encoding mode.

2. Upon exit from the quiet awake state, a motor plan is created at random by Neocortex. Motor cortex receives the motor plan, and a signal to exit simulated mode.

3. Motor cortex sends a copy of the motor plan as a set of angular and linear velocities to the HippocampalFormation and the Animal, resulting in physical motion by the Animal.

4. The angular and linear velocities are processed by the HeadDirectionSystem and GridChartNetworks, resulting in an update to the current orientations and positions active in those systems. These are output to the PlaceSystem.

5. Visual cues corresponding to the updated positions are now available to the LecSystem, and are therefore output to the PlaceSystem.

6. The combined inputs from the LecSystem and GridChartNetworks cause creation of a place ID in the PlaceSystem. As the new place ID is driven by the same inputs from the GridChartNetworks, that portion of the place ID is unchanged. Inputs from LecSystem, however, cause additional nodes to be activated in the AutoassociativeNetwork. This would correspond to a place ID in Figure 5 with the same green dots as created during the simulated motion (driven by MEC), with the addition of a set of blue dots, driven by the new LEC inputs.

41

Support for the Hippocampal Hypothesis. The hippocampal hypothesis suggests that some dynamic local to the hippocampus increases the probability that a particular population of cells will be allocated to represent the next place field to be visited. Dragoi and Tonegawa (2011) suggest that populations whose activation is elevated, but below some threshold, might be the candidates for recruitment for the next place field. From the perspective of SMS, however, the details of the proposed allocation mechanism are irrelevant, because under SMS the place ID to be created for a novel place is driven by the external inputs from the LecSystem and the GridChartNetworks. This is illustrated by contrasting Figures 12 and 13. In Figure 12, an existing active place ID is shown in black. Another population is shown in gray, which is hypothesized to be available to represent the place ID for a novel place field. The details of the internal process that makes this population available are irrelevant however, because under SMS, the allocation of the next place ID is driven entirely by the external inputs from the LecSystem and GridChartNetworks, resulting in the place ID depicted as green and blue dots in Figure 13.

The implementation of the AutoassociativeNetwork uses a binary representation, and therefore does not support modeling sub-threshold activation levels, which makes a direct test of the details of the hippocampal hypothesis infeasible. It should be clear from the above argument, however, that if SMS behaves as expected, *any* hippocampal mechanism will yield the same results – there will be no statistical relationship between the place ID generated by the hippocampal mechanism and the actual place ID driven by the external inputs.

Expected outcomes: Under the simulated motion hypothesis, the place IDs will match with little error between those generated by simulated motion and by subsequent physical exploration. This would provide strong support for the simulated motion hypothesis, and strongly contradict the hippocampal hypothesis. This contradiction covers more than the details of the suggestion by Dragoi and Tonegawa (2011); rather, the external inputs defined by SMS would override any proposed mechanism internal to the hippocampus for the allocation of candidate place cells to as-yet unvisited place fields.

Potential problems: Should the binary representation be found to be too simplistic or otherwise inadequate to support all the research aims, it should be possible to develop a version of the AutoassociativeNetwork that uses a floating point representation of a continuum of activation levels. Although this is likely to be considerably
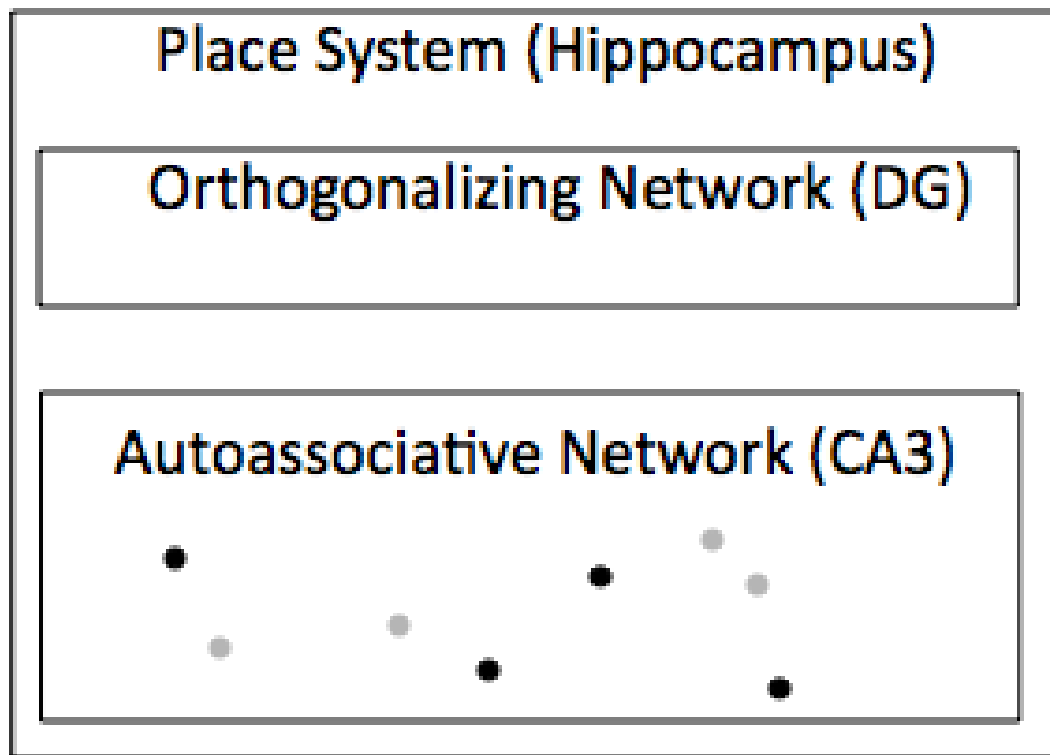
Figure 12. Hypothesized internal dynamics drives allocation of place cells to future place fields
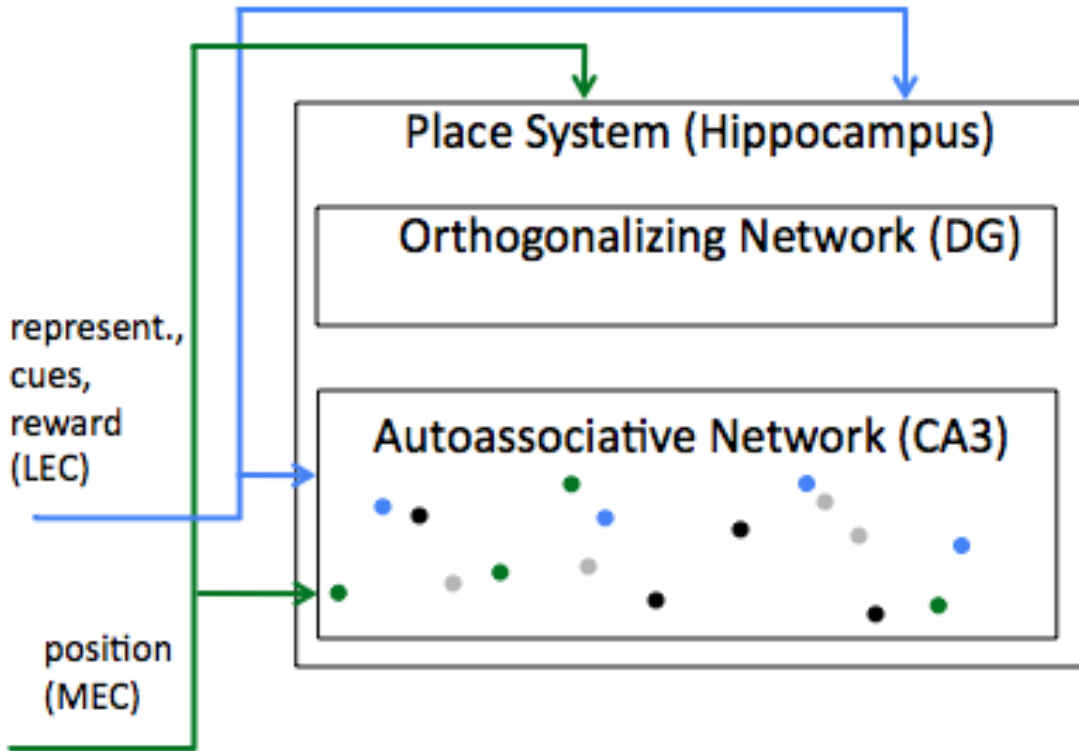
Figure 13. External inputs override hypothesized hippocampal internal dynamics.

more difficult to develop and test, it would have the advantage of making it possible to directly model the details of the hippocampal hypothesis. The expected outcomes would be the same, however.

Aim #2: Contrast target location and target trajectory hypotheses regarding information provided by simulated motion towards a goal location

Aim #2 is to assess the support provided by the computational model for each of two hypotheses regarding the information provided by trajectories in the hippocampus during quiet awake periods. As noted previously, the greater number of home-directed trajectories in the Pfeiffer and Foster (2013) experiment suggests that the animal is aware that the home location may be rewarding. We consider two hypotheses:

– The **target location** hypothesis is that the population of place cells that represent the home location are associated with a higher probability of reward, and there is no additional value in simulating motion towards the goal.

– The **target trajectory** hypothesis, on the other hand, suggests that the sequence of place fields that constitute a trajectory might encode additional information about the probability of reward once the target is reached.

The approach taken toward this aim is to suggest that perhaps the target location and target trajectory hypotheses each describe modes of functioning at different phases during the animal's learning. To motivate this suggestion, we begin by re-visiting the McClelland et al. (1995) theory of complementary learning systems in the hippocampus and neocortex, and then considering the challenges the animal must meet to succeed in the Pfeiffer and Foster experiment.

Complementary learning systems theory suggests that experiences are learned rapidly in the hippocampus, and then replayed multiple times back to the neocortex. This gives the neocortex repeated opportunities to determine the structure of the situation, by slowly altering the weights in a neural network representation of the experience. This theory ascribes to the neocortex the role of learning the structure of a situation, and to the hippocampus the role of remembering what happens at a particular place in a particular context.

Consider the challenges the animal faces in achieving a reasonable mastery of the Pfeiffer and Foster experimental design. The experiment is a type of spatial alternation task, meaning that no particular location is uniformly rewarding. The task is more abstract than spatial alternation in a T-maze, however, where typically there are two specific rewarding locations, and two associated motor plans, turning left or right, to choose between. In this case, the choice is between a home location and an away location. But the away location is abstract; it is some arbitrary location, which is not home, where there is a reward. An optimal representation of the current situation in the experiment would be equivalent to tracking which of two states applies:

– Found Reward Away (FRA): the most recent reward was at an away location

– Found Reward Home (FRH): the most recent reward was at the home location

Corresponding to these representations are two optimal motor plans:

– Move Away (MA): move randomly to an away location

– Move Home (MH): move to the home location

The optimal strategy is to alternate between moving away (MA) when the most recent reward was at the home location (FRH), and moving home (MH) when the most recent reward was at an away location (FRA).

There is one additional level of abstraction – the home location is different on each day of the experiment. Home is defined as the first rewarding location found on the first trial of the day. The first rewarding location is then treated as "home" for all subsequent trials for the day.

Especially given the abstraction implied by the experimental design, it is plausible to suppose that the representations and motor plans necessary to perform optimally in this experiment must compete with various alternatives. An animal might initially associate success in finding a reward with any number of aspects of the situation, e.g., proximity (or distance) from a cue, or a wall, or a corner. Also, the animal might associate success in finding a reward with the motor plan that was executing when the reward was encountered, e.g., approach (or avoid) a cue, or a wall, or a corner, or an egocentrically-defined sequence of motions such as left-right-left, etc.

The training protocol for the experimental animals suggests that it takes some time and different circumstances to distinguish between the various possible representations and motor plans:

1. Traverse 1.8m linear track to receive reward at each end (lesser of 20 minutes or 20 laps, once/day for at least 10 consecutive days).

2. Arena, all wells filled and re-filled immediately, to encourage exploration (30 minutes, once/day for 2 days)

3. Arena, one well filled per quadrant; another random well in same quadrant re-filled once animal had found reward in another quadrant (30 minutes, once/day for 3 days)

4. Arena, same as experimental procedure, except two random away wells filled to make finding reward easier (30 minutes, once/day for 3 days)

5. Arena, same as experimental procedure (lesser of 30 minutes or 30 trials, once/day until criterion reached of 30 trials in less than 30 minutes for 3 consecutive days)

With this context, I suggest there may be three phases to the animal's learning and behavior. Although these modes are treated as discrete, it is more likely that all exist at any given time, with different probabilities.

1. The first phase begins with the animal's initial exposure to the experimental situation. During this phase, the animal exhibits certain typical behaviors that are repeated in any novel environment (Eilam & Golani, 1989), including random exploration and returning to the initial point of entry into the environment ("home base"). During quiet awake periods in the arena, the complementary learning systems theory suggests the hippocampus replays recent experience to help the neocortex infer structure. As discussed earlier, there is ample evidence of replay activity in a number of experimental situations, but there is no recording of hippocampal activity during the animal's training in the Pfeiffer and Foster experiment, so we may only infer the presence of replay activity.

2. In the second phase, the animal's performance has met the criteria for commencement of the experimental trials. By this point, under complementary learning systems theory, we suppose that the animal has developed representations of the experimental design in the neocortex, and has stored hippocampal memories of each location in the arena, including the representation that was active at the time of the experience at a given place. Under the target trajectory hypothesis, during quiet awake periods, the animal uses its cortical understanding of the experimental design to explore scenarios, in effect, asking what would happen if it moved in the direction of a cue, or a corner, or towards home. By simulating motion towards the target, a hippocampal memory of each location along the trajectory is retrieved, along with a memory of its rewarding characteristics, as well as the currently active representation. If the result of the simulated motion meets some threshold for likely reward, the motor plan that generated the simulated trajectory is selected as the next motor plan for physical execution. Executing several simulated motions enables the animal to test the likely results from different motor plans, at lower cost than physical motion.

The suggestion is that the results from Pfeiffer and Foster were generated by animals in this second phase of learning, consistent with the target trajectory hypothesis.

3. In the third phase, the animal's representation of the experiment has further improved, to the point where the animal's responses approach the level of habit. We might conjecture that the animal develops stimulus-response associations between occurrence of the Found Reward Away representation and the Move Home motor plan, and between occurrence of the Found Reward Home representation and the Move Away motor plan. Having established these associations, the animal no longer obtains value from simulating its next motion, and therefore executes the associated motor plan immediately. A similar pattern has been observed in experiments that studied virtual trial and error (VTE), which I have hypothesized may use the same simulated motion mechanism. Johnson and Redish (2007) found that the incidence of VTE drops over the course of a test session as the animal learns the task better.

The suggestion is that if the Pfeiffer and Foster experiment were repeated, but the animals were trained longer, perhaps their responses would have achieved habitual status, and the amount of simulated motion would drop. This is consistent with the prediction of the target location hypothesis.

Cortical Representation of the Experiment. To begin with, we might suppose that a cortical representation might be a neural network comprised of places, representations, motor plans, and outcomes as in Figure 14. As suggested by memory index theory, these types of nodes are stored in neocortical autoassociative memory associated with hippocampal place IDs, such that retrieval of a place ID causes retrieval of the associated cortical representation (see Figure 15). Then, as trials of the simulated experiment are run, information from the hippocampus and from experience would be used to update the weights in this network. Initial random weights are gradually adjusted during the first phase of random exploration, based on actual experience and augmented by replay from the hippocampus. Under the target trajectory hypothesis, in phase 2 the animal begins to make predictions about the outcomes of simulated motions, and use those to choose its next physical trajectory. This corresponds to activation of a place and a representation, in conjunction with choice of a motor plan, and then simulating the motion. If the simulated trajectory
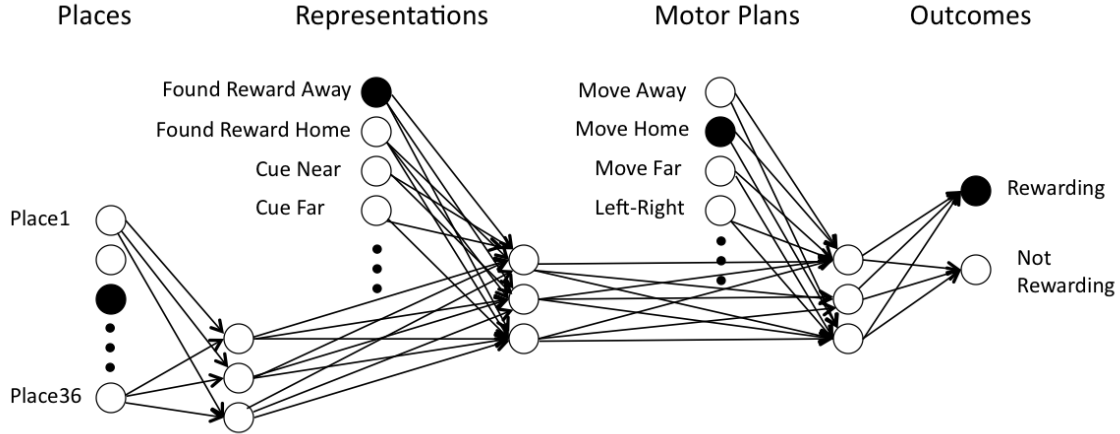
48

Figure 14. Cortical neural network linking places, representations, motor plans, and outcomes

results in retrieving a rewarding place memory, that is treated as the predicted outcome, and the motor plan is executed physically. The weights in the network continue to be adjusted during phase 2. Under the target location hypothesis, in phase 3, the animal begins to choose its next motor plan directly, without simulated motion. This corresponds to activation of a place, a representation, and the rewarding outcome, and using those to predict the best motor plan. The motor plan is then selected for physical execution.

There are at least two simplifications in this approach that may not be justifiable. Representing a motor plan as a single node is problematic; more realistic would be depicting a motor plan as a finite state machine or Petri net. This will be explored further, but for now, the argument can be explained by treating a motor plan as an atomic node. Also problematic is the implicit suggestion that the structure of the network is fixed *a priori*, that the candidate nodes are known prior to the experiment. Perhaps it is sufficient to state that SMS does not attempt to suggest a theory of how the neocortex develops representations or motor plans. Rather, I will assume as simple a model as feasible of cortical representation, and examine how information from the hippocampus supports changes in the cortical representation over time, consistent with the three phases proposed.
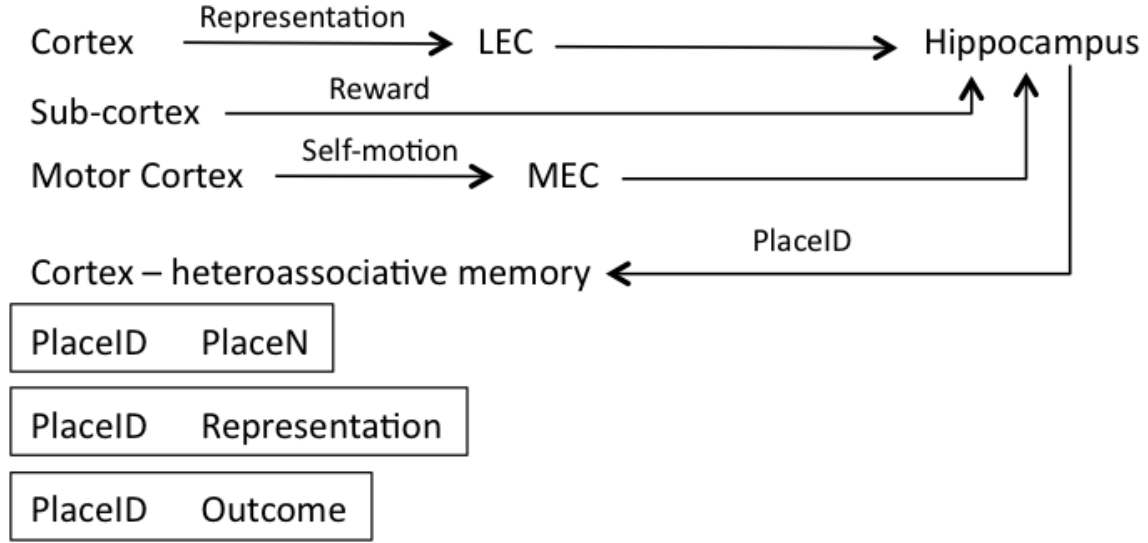
Figure 15. Hypothesized associations between hippocampal places and cortical network nodes

Support for Phase One: Replay. The component interaction sequence for replay of trajectories has been outlined previously. In step 4 the retrieved place ID is output to Neocortex to drive update of its representation of the situation. Specifically, the weights connecting the place, representation and outcome are adjusted. As currently modeled, there is no motor plan associated with replay, so those weights are left unadjusted.

Support for Phase Two: Target Trajectory Hypothesis. The component interaction sequence for generation of simulated motion trajectories has been outlined previously, and is repeated here, with modifications in italics.

1. Upon entry into the quiet awake state, SubCortex modulates the HippocampalFormation to recall mode.

2. Upon entry into the quiet awake state, *Neocortex draws a motor plan from a distribution representing the relative weights of all the motor plans in the cortical network.* Motor cortex receives the motor plan, and a signal to enter simulated mode.

3. *Neocortex sends the currently active representation of the situation to the HippocampalFormation.*

4. Motor cortex sends a copy of the motor plan as a set of angular and linear velocities to the HippocampalFormation, while suppressing the copy that would otherwise be sent to the Animal.

5. The angular and linear velocities are processed by the HeadDirectionSystem and GridChartNetworks, resulting in an update to the current orientations and positions active in those systems. These are output to the PlaceSystem.

6. Visual cues corresponding to the updated positions are not available to the LecSystem, and are therefore not part of the retrieval from the PlaceSystem. This simplification is due to not previously storing the place ID in SensoryCortex in association with the visual cues. In principle, the system could be extended to support such storage and retrieval, but it does not seem critical to the argument, so has not been provided for.

7. *LecSystem sends the representation of the situation to the PlaceSystem.*

8. The PlaceSystem retrieves a place ID based on the positional inputs *and situation representation*, and outputs this to the Neocortex as a hypothesis as to what might be encountered by traveling to the new place.

9. *Using the returned place ID, the Neocortex retrieves the associated outcome from its autoassociative network. If the retrieved outcome is rewarding, the current motor plan is chosen as the next motor plan to be executed. Otherwise, steps 2-8 are repeated some number of times, probably modeled as an integer draw from a Gaussian with a configurable mean in the single digits.*

During this processing, the cortical network is being used for inference, so its weights are not updated during the simulated motion. Once a motor plan is selected and physically executed, the experienced results are used to update the weights.

Support for Phase Three: Target Location Hypothesis. The component interaction sequence for navigation has been outlined previously as part of placement in a known environment. What is added is the selection of motor plan for physical execution, which is done by retrieving a motor plan from the neural network based on

the activations of the current place, the current representation, and the rewarding outcome.

Expected outcomes: The evaluation of the target trajectory and target location hypotheses is expected to show a trade-off between the extra cost of simulation under the target trajectory hypothesis against the cost of unproductive physical motion under the target location hypothesis. At a high level, this trade-off is depicted in Figure 16. Over time, the difference should grow between the mean weights of the network and the weights of the most effective representations of the situation, FRA and FRH. For small differences, there should be an advantage to simulating motion, as a way to avoid the frequent penalty of unproductive physical motion. For large differences, corresponding to a strengthening stimulus-response association, there should be an advantage to selecting the motor plan for immediate execution, and avoiding the cost of simulated motion.

Potential problems: There are a number of potential problems in demonstrating this aim. First, the expected trade-off between the target trajectory and target location hypotheses may not emerge. One hypothesis may always have an advantage over the other. Verifying such an asymmetry would require some searching of the parameter space of the relative costs of simulated motion against physical motion, as well as the number of simulated trajectories per quiet awake period, and the number of representation nodes and motor plan nodes in the neural network.

Failure to find an area in the parameter space where the trade-off exists might then suggest that the neural network model of the cortical representation is not appropriate. In particular, the simplifying assumption of pre-defining the possible representations and motor plans might skew the outcome toward one hypothesis. Perhaps the more realistic case, where representations and motor plans emerge through some dynamics would lead to the hypothesized trade-off. However, there are practical limits on how realistic a model can be introduced here.

Finally, it may be the case that the target trajectory hypothesis cannot be supported by the approaches suggested here. This would mean that the simulated motion system, although capable of generating behaviors compatible with the observed behaviors in the Pfeiffer and Foster experiment, does not yet explain why simulated motion might be advantageous. The central observation remains, however, that animals appear to be simulating their future activity. The failure to demonstrate Aim #2 would leave additional room for other computational theories of simulation, e.g.,
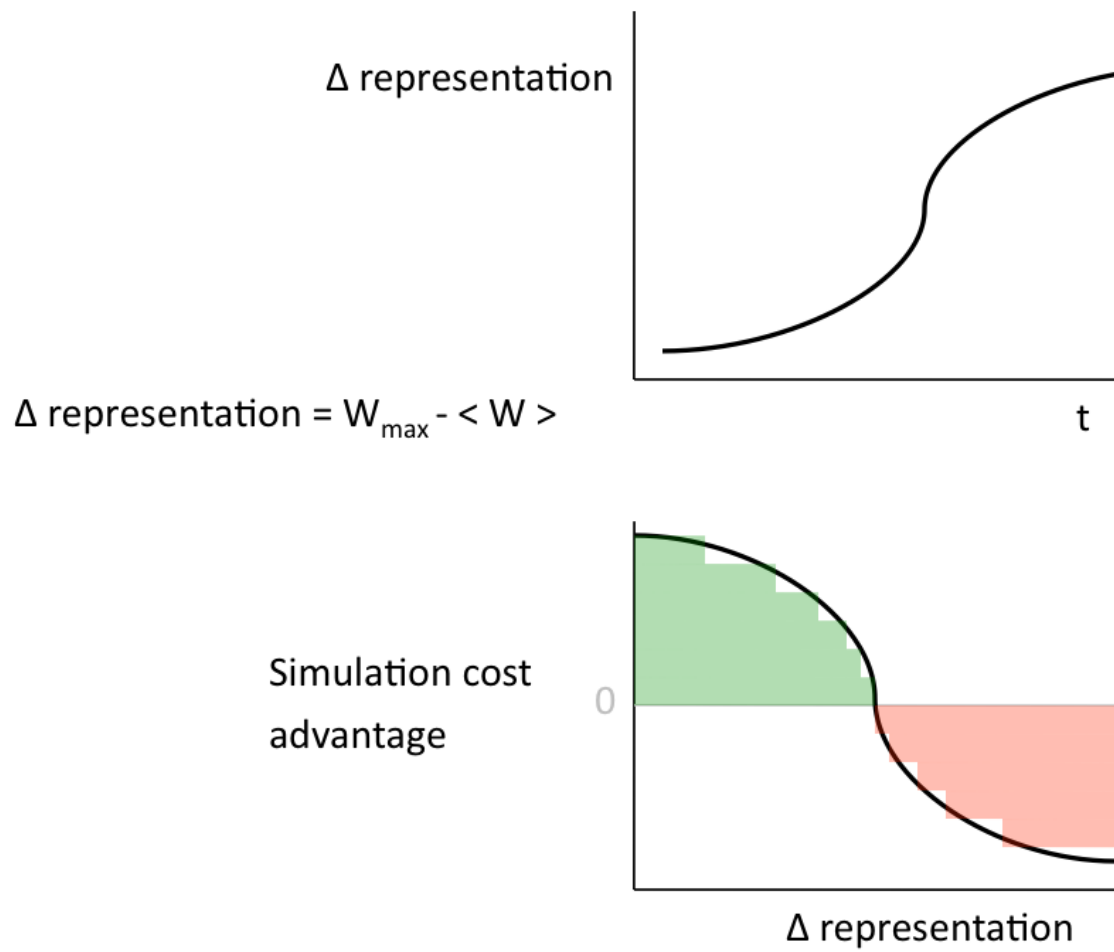
Δ representation

Δ representation = W_{max} - $<W>$

t

Simulation cost advantage

0

Δ representation

Figure 16. Evaluating the tradeoff between the target trajectory and target location hypotheses

that it is driven out of internal processes within the hippocampus. Regardless, a definitive account of simulation will be driven by additional experimental evidence, allowing the evaluation of the testable predictions made earlier.

Aim #3: Assess the usefulness of Petri nets as an approach for modeling complex neural processes

Aim #3 is to assess the usefulness of Petri nets in support of the simulated motion system, as an example of modeling neural processes. SMS works at the level of population models and connections among components. It is not primarily a mathematical model, nor is it as detailed as a spiking neuron model. At this middle level of abstraction, Petri nets are a way to accomplish several goals:

– model relatively fixed routines, such as motor programs

– model the overall behavior of the subject process, i.e., behavior of the simulated animal in the experimental setting

– explicitly model key states and events

As discussed earlier, Petri nets are for the most part neurologically plausible. Their pictorial representation allows straightforward evaluation of their functioning. This should lower the reader's concerns about the amount of unrealistic model complexity that might otherwise be buried in source code.

With the use of the extensions to enable Petri nets to drive and be driven by external systems, discussed earlier under prior work, Petri nets serve as a mechanism to tie together the interaction of multiple components. It is a general software principle, as well as an apparent attribute of neural systems, that complex behavior is a result of the interaction of multiple relatively simple components. Petri nets provide good support for such models.

Expected outcomes: The accomplishment of this aim will be assessed in qualitative terms. If the development and validation of the SMS, and accomplishment of the first two aims proceeds as outlined above, this aim will be taken as successfully achieved. If any of these conditions is not met, the cause of the failure will be analyzed. If it is due in part to the deficiencies of the Petri net models, then this aim will not have been achieved.

Potential problems: There might be two problems in achieving this aim. First, there could be unanticipated problems in integrating Petri nets with Matlab. Although unlikely, as prior work with Python models showed the basic feasibility of the approach, the fallback position would be to code the entire system in Matlab. This would be at a cost of visibility to the component integration, but should at least enable the other aims to be accomplished. Secondly, there could be an unacceptable performance penalty of the Petri net integration. Although use of distributed or large-scale computing environments is certainly possible, as a practical matter, it is desirable for the model to run comfortably on a modern laptop or desktop. This should increase the potential usefulness of the system for others. Again, the fallback position might be to code the entire system in Matlab, which should lower the performance cost. Another alternative, if performance is marginal, is to accept that obtaining simulation results might require long periods of processing, and to adjust the experimental schedule accordingly.

References

Bergen, B., Chang, N., & Narayan, S. (2004). Simulated action in an embodied construction grammar. In *Proceedings of the twenty-sixth annual conference of the cognitive science society* (pp. 108–113).

Bergen, B. K. (2012). *Louder than words: The new science of how the mind makes meaning*. Basic Books.

Brandon, M. P., Koenig, J., Leutgeb, J. K., & Leutgeb, S. (2014). New and distinct hippocampal place codes are generated in a new environment during septal inactivation. *Neuron*, *82*(4), 789–796.

Buzsáki, G. (1986). Hippocampal sharp waves: their origin and significance. *Brain research*, *398*(2), 242–252.

Buzsáki, G., Horvath, Z., Urioste, R., Hetke, J., & Wise, K. (1992). High-frequency network oscillation in the hippocampus. *Science*, *256*(5059), 1025–1027.

Buzsáki, G., S., L. L.-W., & Vanderwolf, C. H. (1983). Cellular bases of hippocampal {EEG} in the behaving rat. *Brain Research Reviews*, *6*(2), 139–171. Retrieved from `http://www.sciencedirect.com/science/article/pii/0165017383900371` doi: http://dx.doi.org/10.1016/0165-0173(83)90037-1

Chrobak, J. J., & Buzsáki, G. (1994). Selective activation of deep layer (V-VI) retrohippocampal cortical neurons during hippocampal sharp waves in the behaving rat. *The Journal of neuroscience*, *14*(10), 6160–6170.

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

Colgin, L. L., Denninger, T., Fyhn, M., Hafting, T., Bonnevie, T., Jensen, O., ... Moser, E. I. (2009, November). Frequency of gamma oscillations routes flow of information in the hippocampus. *Nature*, *462*(7271), 353–357. Retrieved from `http://dx.doi.org/10.1038/nature08573`

Colgin, L. L., Leutgeb, S., Jezek, K., Leutgeb, J. K., Moser, E. I., McNaughton, B. L., & Moser, M.-B. (2010). Attractor-map versus autoassociation based attractor dynamics in the hippocampal network. *Journal of neurophysiology*, *104*(1), 35–50.

Couey, J. J., Witoelar, A., Zhang, S.-J., Zheng, K., Ye, J., Dunn, B., ... Others (2013). Recurrent inhibitory circuitry as a mechanism for grid formation. *Nature neuroscience*, *16*(3), 318–324.

Davidson, T. J., Kloosterman, F., & Wilson, M. A. (2009). Hippocampal replay of extended experience. *Neuron*, *63*(4), 497–507.

Deshmukh, S. S. (2014). Spatial and nonspatial representations in the lateral entorhinal cortex. In *Space, time and memory in the hippocampal formation* (pp. 127–152). Springer.

Diba, K., & Buzsáki, G. (2007). Forward and reverse hippocampal place-cell sequences during ripples. *Nature neuroscience*, *10*(10), 1241–1242.

Dragoi, G., & Tonegawa, S. (2011, January). Preplay of future place cell sequences by hippocampal cellular assemblies. *Nature*, *469*(7330), 397–401. Retrieved from `http://dx.doi.org/10.1038/nature09633http://www.nature.com/nature/journal/v469/n7330/abs/10.1038-nature09633-unlocked.html\#supplementary-information`

Dragoi, G., & Tonegawa, S. (2013). Distinct preplay of multiple novel spatial experiences in the rat. *Proceedings of the National Academy of Sciences*, *110*(22), 9100–9105.

Eilam, D., & Golani, I. (1989). Home base behavior of rats (Rattus norvegicus) exploring a novel environment. *Behavioural brain research*, *34*(3), 199–211.

Ferbinteanu, J., Shirvalkar, P., & Shapiro, M. L. (2011). Memory modulates journey-dependent coding in the rat hippocampus. *The Journal of Neuroscience*, *31*(25), 9135–9146.

Foster, D. J., & Wilson, M. A. (2006). Reverse replay of behavioural sequences in hippocampal place cells during the awake state. *Nature*, *440*(7084), 680–683. doi: 10.1038/nature04587

Frank, L. M., Brown, E. N., & Wilson, M. (2000). Trajectory encoding in the hippocampus and entorhinal cortex. *Neuron*, *27*(1), 169–178.

Fries, P. (2005). A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends in cognitive sciences*, *9*(10), 474–480.

Gothard, K. M., Skaggs, W. E., & McNaughton, B. L. (1996). Dynamics of mismatch correction in the hippocampal ensemble code for space: interaction between path integration and environmental cues. *The Journal of neuroscience*, *16*(24), 8027–8040.

Guanella, A., Kiper, D., & Verschure, P. (2007). A model of grid cells based on a twisted torus topology. *International journal of neural systems*, *17*(04), 231–240.

Hafting, T., Fyhn, M., Molden, S., Moser, M.-B., & Moser, E. I. (2005). Microstructure of a spatial map in the entorhinal cortex. *Nature*, *436*(7052), 801–806.

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the national academy of sciences*, *79*(8), 2554–2558.

Johnson, A., & Redish, A. D. (2007). Neural ensembles in CA3 transiently encode paths

forward of the animal at a decision point. *The Journal of neuroscience*, *27*(45), 12176–12189.

Jung, M. W., Wiener, S. I., & McNaughton, B. L. (1994). Comparison of spatial firing characteristics of units in dorsal and ventral hippocampus of the rat. *The Journal of neuroscience*, *14*(12), 7347–7356.

Karlsson, M. P., & Frank, L. M. (2009). Awake replay of remote experiences in the hippocampus. *Nature neuroscience*, *12*(7), 913–918.

Knierim, J. J., Kudrimoti, H. S., & McNaughton, B. L. (1995). Place cells, head direction cells, and the learning of landmark stability. *The Journal of Neuroscience*, *15*(3), 1648–1659.

Lee, A. K., & Wilson, M. A. (2002). Memory of sequential experience in the hippocampus during slow wave sleep. *Neuron*, *36*(6), 1183–1194.

Leutgeb, J. K., Leutgeb, S., Moser, M.-B., & Moser, E. I. (2007). Pattern separation in the dentate gyrus and CA3 of the hippocampus. *science*, *315*(5814), 961–966.

Leutgeb, S., Leutgeb, J. K., Barnes, C. A., Moser, E. I., McNaughton, B. L., & Moser, M.-B. (2005). Independent codes for spatial and episodic memory in hippocampal neuronal ensembles. *Science*, *309*(5734), 619–623.

Louie, K., & Wilson, M. A. (2001). Temporally structured replay of awake hippocampal ensemble activity during rapid eye movement sleep. *Neuron*, *29*(1), 145–156.

Maurer, A. P., VanRhoads, S. R., Sutherland, G. R., Lipa, P., & McNaughton, B. L. (2005). Self-motion and the origin of differential spatial scaling along the septo-temporal axis of the hippocampus. *Hippocampus*, *15*(7), 841–852.

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychological review*, *102*(3), 419.

McNaughton, B. L., Barnes, C. A., Gerrard, J. L., Gothard, K., Jung, M. W., Knierim, J. J., . . . Others (1996). Deciphering the hippocampal polyglot: the hippocampus as a path integration system. *The Journal of Experimental Biology*, *199*(1), 173–185.

McNaughton, B. L., Battaglia, F. P., Jensen, O., Moser, E. I., & Moser, M.-B. (2006). Path integration and the neural basis of the'cognitive map'. *Nature Reviews Neuroscience*, *7*(8), 663–678.

McNaughton, B. L., Chen, L., & Markus, E. J. (1991). Dead reckoning, landmark learning, and the sense of direction: a neurophysiological and computational hypothesis. *Cognitive Neuroscience, Journal of*, *3*(2), 190–202.

McNaughton, B. L., Leonard, B., & Chen, L. (1989). Cortical-hippocampal interactions

and cognitive mapping: A hypothesis based on reintegration of the parietal and inferotemporal pathways for visual processing. *Psychobiology*, *17*(3), 230–235.

McNaughton, B. L., & Morris, R. G. M. (1987). Hippocampal synaptic enhancement and information storage within a distributed memory system. *Trends in neurosciences*, *10*(10), 408–415.

McNaughton, B. L., & Nadel, L. (1990). Hebb-Marr networks and the neurobiological representation of action in space. *Neuroscience and connectionist theory*, 1–63.

Mehta, M. R., Barnes, C. A., & McNaughton, B. L. (1997). Experience-dependent, asymmetric expansion of hippocampal place fields. *Proceedings of the National Academy of Sciences*, *94*(16), 8918–8921.

Mittelstaedt, M.-L., & Mittelstaedt, H. (1980). Homing by path integration in a mammal. *Naturwissenschaften*, *67*(11), 566–567.

Navratilova, Z., Hoang, L. T., Schwindel, C. D., Tatsuno, M., & McNaughton, B. L. (2012). Experience-dependent firing rate remapping generates directional selectivity in hippocampal place cells. *Frontiers in neural circuits*, *6*.

Navratilova, Z., & McNaughton, B. L. (2014). Models of path integration in the hippocampal complex. In *Space, time and memory in the hippocampal formation* (pp. 191–224). Springer Vienna.

Norimoto, H., Matsumoto, N., Miyawaki, T., Matsuki, N., & Ikegaya, Y. (2013). Subicular activation preceding hippocampal ripples in vitro. *Scientific reports*, *3*.

O'Keefe, J. (1976). Place units in the hippocampus of the freely moving rat. *Experimental neurology*, *51*(1), 78–109.

O'Keefe, J., & Burgess, N. (1996). Geometric determinants of the place fields of hippocampal neurons. *Nature*, *381*(6581), 425–428.

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*(1), 171–175.

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

O'Keefe, J., & Recce, M. L. (1993). Phase relationship between hippocampal place units and the EEG theta rhythm. *Hippocampus*, *3*(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), 1–27.

Ólafsdóttir, H. F., Barry, C., Saleem, A. B., Hassabis, D., & Spiers, H. J. (2015). Hippocampal place cells construct reward related sequences through unexplored space. *eLife*, *4*. doi: 10.7554/eLife.06063

O'Reilly, K. C., Alarcon, J. M., & Ferbinteanu, J. (2014). Relative contributions of CA3

and medial entorhinal cortex to memory in rats. *Frontiers in behavioral neuroscience*, *8*. doi: http://dx.doi.org/10.3389/fnbeh.2014.00292

Ormond, J., & McNaughton, B. L. (2015). Place field expansion after focal MEC inactivations is consistent with loss of Fourier components and path integrator gain reduction. *Proceedings of the National Academy of Sciences*, *112*(13), 4116–4121. Retrieved from `http://www.pnas.org/content/112/13/4116.abstract` doi: 10.1073/pnas.1421963112

Paller, K. A. (1997). Consolidating dispersed neocortical memories: the missing link in amnesia. *Memory*, *5*(1-2), 73–88.

Peyrache, A., Lacroix, M. M., Petersen, P. C., & Buzsaki, G. (2015, April). Internally organized mechanisms of the head direction sense. *Nat Neurosci*, *18*(4), 569–575. Retrieved from `http://dx.doi.org/10.1038/nn.396810.1038/nn.3968http://www.nature.com/neuro/journal/v18/n4/abs/nn.3968.html\#supplementary-information`

Pfeiffer, B. E. (2015). Personal correspondence.

Pfeiffer, B. E., & Foster, D. J. (2013). Hippocampal place-cell sequences depict future paths to remembered goals. *Nature*. Retrieved from `http://www.nature.com/nature/journal/v497/n7447/full/nature12112.html`

Pfeiffer, B. E., & Foster, D. J. (2015). Autoassociative dynamics in the generation of sequences of hippocampal place cells. *Science*, *349*(6244), 180–183. Retrieved from `http://www.sciencemag.org/content/349/6244/180.abstract` doi: 10.1126/science.aaa9633

Ranck Jr, J. B. (1985). Head Direction Cells in the Deep Cell Layer of Dorsal Postsubiculum in Freely Moving Rats. *Electrical Activity of the Archicortex*, 217–220.

Redish, A. D., Elga, A. N., & Touretzky, D. S. (1996). A coupled attractor model of the rodent head direction system. *Network: Computation in Neural Systems*, *7*(4), 671–685.

Samsonovich, A., & McNaughton, B. L. (1997). Path integration and cognitive mapping in a continuous attractor neural network model. *The Journal of neuroscience*, *17*(15), 5900–5920.

Sargolini, F., Fyhn, M., Hafting, T., McNaughton, B. L., Witter, M. P., Moser, M.-B., & Moser, E. I. (2006). Conjunctive representation of position, direction, and velocity in entorhinal cortex. *Science*, *312*(5774), 758–762.

Schacter, D. L., Addis, D. R., Hassabis, D., Martin, V. C., Spreng, R. N., & Szpunar, K. K. (2012). The future of memory: remembering, imagining, and the brain. *Neuron*, *76*(4), 677–694.

Sharp, P. E., Kubie, J. L., & Muller, R. U. (1990). Firing properties of hippocampal neurons in a visually symmetrical environment: contributions of multiple sensory cues and mnemonic processes. *The Journal of Neuroscience*, *10*(9), 3093–3105.

Silva, D., Feng, T., & Foster, D. J. (2015). Trajectory events across hippocampal place cells require previous experience. *Nature neuroscience*, *18*(12), 1772–1779.

Singer, A. C., & Frank, L. M. (2009). Rewarded Outcomes Enhance Reactivation of Experience in the Hippocampus. *Neuron*, *64*(6), 910–921. Retrieved from `http://www.sciencedirect.com/science/article/pii/S089662730900899X` doi: http://dx.doi.org/10.1016/j.neuron.2009.11.016

Skaggs, W. E., Knierim, J. J., Kudrimoti, H. S., & McNaughton, B. L. (1995). A model of the neural basis of the rat's sense of direction. *Advances in Neural Information Processing Systems 7*, *7*, 173.

Skaggs, W. E., McNaughton, B. L., Wilson, M. A., & Barnes, C. A. (1996). Theta Phase Precession in Hippocampal Neuronal Populations and the Compression of Temporal Sequences. *Hippocampus*, *6*, 149–172.

Stensola, H., Stensola, T., Solstad, T., Frø land, K., Moser, M.-B., & Moser, E. I. (2012). The entorhinal grid map is discretized. *Nature*, *492*(7427), 72–78.

Taube, J. S., Muller, R. U., & Ranck, J. B. (1990a). Head-direction cells recorded from the postsubiculum in freely moving rats. I. Description and quantitative analysis. *The Journal of Neuroscience*, *10*(2), 420–435.

Taube, J. S., Muller, R. U., & Ranck, J. B. (1990b). Head-direction cells recorded from the postsubiculum in freely moving rats. II. Effects of environmental manipulations. *The Journal of Neuroscience*, *10*(2), 436–447.

Terrazas, A., Krause, M., Lipa, P., Gothard, K. M., Barnes, C. A., & McNaughton, B. L. (2005). Self-motion and the hippocampal spatial metric. *The Journal of neuroscience*, *25*(35), 8085–8096.

Teyler, T. J., & DiScenna, P. (1986). The hippocampal memory indexing theory. *Behavioral neuroscience*, *100*(2), 147.

Trappenberg, T. (2010). *Fundamentals of computational neuroscience* (second ed.). OUP Oxford.

Treves, A., & Rolls, E. T. (1992). Computational constraints suggest the need for two distinct input systems to the hippocampal CA3 network. *Hippocampus*, *2*(2), 189–199.

Van Cauter, T., Poucet, B., & Save, E. (2008). Unstable CA1 place cell representation in rats with entorhinal cortex lesions. *European journal of Neuroscience*, *27*(8), 1933–1946.

Wang, Y., Romani, S., Lustig, B., Leonardo, A., & Pastalkova, E. (2015). Theta sequences are essential for internally generated hippocampal firing fields. *Nature neuroscience*, *18*(2), 282–288. Retrieved from `http://www.nature.com/neuro/journal/v18/n2/full/nn.3904.html\#methods`

Wilson, M. A., & McNaughton, B. L. (1993). Dynamics of the hippocampal ensemble code for space. *Science*, *261*(5124), 1055–1058. Retrieved from `http://www.sciencemag.org/content/261/5124/1055.abstract` doi: 10.1126/science.8351520

Wilson, M. A., & McNaughton, B. L. (1994). Reactivation of hippocampal ensemble memories during sleep. *Science*, *265*(5172), 676–679.

Wood, E. R., Dudchenko, P. A., Robitsek, R. J., & Eichenbaum, H. (2000). Hippocampal neurons encode information about different types of memory episodes occurring in the same location. *Neuron*, *27*(3), 623–633.

Zhang, K. (1996). Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: a theory. *The journal of neuroscience*, *16*(6), 2112–2126.

Zilli, E. A. (2012). Models of grid cell spatial firing published 2005–2011. *Frontiers in neural circuits*, *6*.