The involvement of recurrent connections in area CA3 in establishing the properties of place fields: A model.

Szabolcs Káli[1,2] **and Peter Dayan**[1]

[1]Gatsby Computational Neuroscience Unit, University College London, 17 Queen Square, London WC1N 3AR, U.K.

[2]Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, U.S.A.

Abbreviated title: Recurrent model of CA3 place fields

51 text pages, 12 figures, 1 table

206 words in abstract, 474 in introduction, 1454 in discussion

Corresponding author:

Szabolcs Káli
Gatsby Computational Neuroscience Unit
Alexandra House
17 Queen Square
London WC1N 3AR
U.K.

Phone: +44 20 7679 1192

Fax: +44 20 7679 1173

E-mail: szabolcs@gatsby.ucl.ac.uk

Acknowledgements: We are very grateful to Drs. S. Becker, N. Burgess, S. Corkin, K.J. Jeffery, J. O'Keefe, W.E. Skaggs, D.S. Touretzky, and M.A. Wilson for their comments on the manuscript and useful discussions, and our anonymous reviewers for their detailed comments. Supported by the Gatsby Charitable Foundation and NSF grant IBN-9634339.

Abstract

Strong constraints on the neural mechanisms underlying the formation of place fields in the rodent hippocampus come from the systematic changes in spatial activity patterns that are consequent on systematic environmental manipulations. We describe an attractor network model of area CA3 in which local, recurrent, excitatory and inhibitory interactions generate appropriate place cell representations from location- and direction-specific activity in the entorhinal cortex.

In the model, familiarity with the environment, as reflected by activity in neuromodulatory systems, influences the efficacy and plasticity of the recurrent and feedforward inputs to CA3. In unfamiliar, novel, environments, mossy fiber inputs impose activity patterns on CA3, and the recurrent collaterals and the perforant path inputs are subject to graded Hebbian plasticity. This sculpts CA3 attractors and associates them with activity patterns in the entorhinal cortex. In familiar environments, place fields are controlled by the way that perforant path inputs select amongst the attractors.

Depending on the training experience provided, the model generates place fields that are either directional or non-directional, and whose changes when the environment undergoes simple geometric transformations are in accordance with experimental data. Representations of multiple environments can be stored and recalled with little interference, and these have the appropriate degrees of similarity in visually similar environments.

Key words: hippocampus, place cells, CA3, recurrent network, plasticity, familiarity, neuromodulation, directionality, attractor, model

The hippocampus is known to be involved in spatial learning and memory in rodents. Some of the most convincing evidence for this is the presence of place cells in areas CA3 and CA1 of the hippocampus (O'Keefe and Dostrovsky, 1971; O'Keefe, 1976), and of many other types of spatially selective cells in neighboring areas (Jung and McNaughton, 1993; Quirk et al., 1992). Principal neurons in CA3 and CA1 are active only when the animal is located in a well-defined local region of the environment (a place field; Muller et al., 1987), and collectively provide a population code for spatial position (Wilson and McNaughton, 1993). The question we address is how this comes to be in a way that is consistent with the evidence for the involvement of the hippocampus in more general forms of memory.

A key anatomical feature of area CA3 is that its pyramidal cells receive the majority of their inputs from other CA3 pyramidal cells (Amaral and Witter, 1989; Amaral et al., 1990). The resulting recurrent network has been extensively explored as a plastic attractor model of the way that the hippocampus acts as a general memory (Marr, 1971; McNaughton and Morris, 1987; Rolls, 1996; Levy, 1996; Hasselmo et al., 1996), but has been widely ignored by models that are intended to account for various properties of place cells (Zipser, 1985; Sharp, 1991; Touretzky and Redish, 1996; Burgess et al., 1997; but see Battaglia and Treves, 1998).

The model of Samsonovich and McNaughton (1997) was the first to explore the consequences of the CA3 attractor network for the place cell representation. Their model assumes the existence of a collection of independent continuous sets of attractors realized by the CA3 recurrent network, and successfully accounts for some of the basic experimental observations about place cells. However, in a model with fixed, independent sets of attractors, it is hard to explain the recent experimental findings by Skaggs and McNaughton (1998), who found partially overlapping place cell representations in two distinct but similar-looking parts of an apparatus. Such models generally predict either identical or completely

different firing patterns in this situation. In addition, Samsonovich and McNaughton's (1997) model does not address the question as to how the strengths of the CA3 recurrent connections, which are essential for the existence of appropriate attractors, become established. As is critical for models in which the hippocampus acts as a memory, there is substantial evidence for synaptic plasticity in most major hippocampal pathways, including those providing feedforward inputs to area CA3 (Zalutsky and Nicoll, 1990; Breindl et al., 1994) and the CA3 recurrent collateral connections (Zalutsky and Nicoll, 1990; Debanne et al., 1998). These activity-dependent synaptic changes provide the obvious means for setting up the appropriate connection strengths, and, in conjunction with the attractor structure, thereby allow us to relate a major aspect of spatial processing to a major aspect of memory processing.

Brunel and Trullier (1998) and we (Káli and Dayan, 1998) independently implemented models which rely on modifiable recurrent connections in CA3 to explain the differences in the directionality of place cells in different kinds of environment. However, the strongest challenge for models, and particularly models based on attractor networks, comes from data on the behavior of place cells in multiple environments that are similar, or are related by simple geometric manipulations. In this paper, we present an attractor model with appropriate behavior in these cases.

Place field formation in simple environments

Our model is grounded in two assumptions. The first is that observed place cell activity patterns reflect the stable states of the CA3 attractor network, a network whose dynamics is governed by its intrinsic recurrent excitatory connections supplemented by inhibitory feedback (see Figure 1). Inputs to CA3, arriving via learned feedforward connections from entorhinal cortex (EC), are used to select among the stored attractors. We utilize experimental data as well

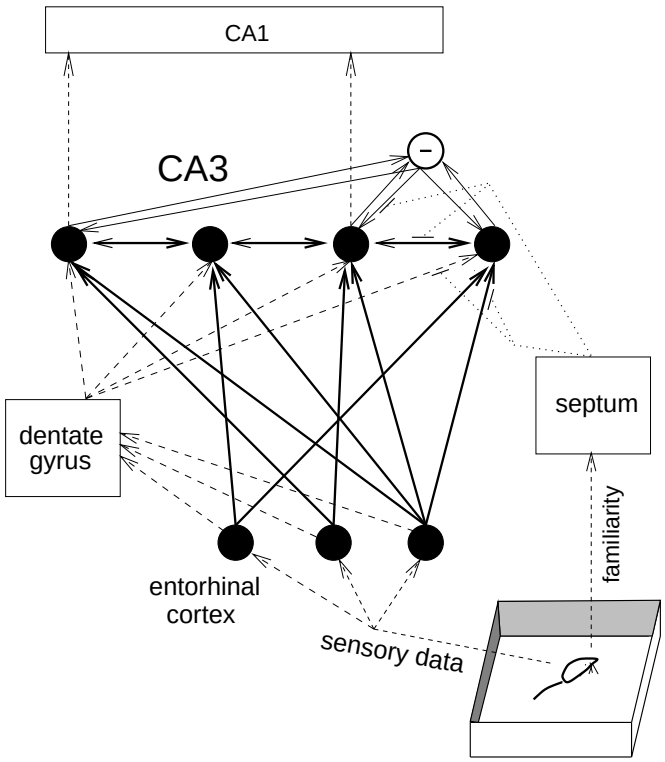

Figure 1: Model architecture. The inputs to the network are the activities of neurons in entorhinal cortex, which are determined by sensory features in the environment. This representation is then transformed by feedforward pathways (the direct perforant path connections to CA3 and the pathway through the dentate gyrus) and recurrent processing in area CA3, which involves lateral connections between CA3 pyramidal cells (filled circles) as well as their connections with an inhibitory neuron (open circle). The solid lines indicate neuronal connections that are modelled explicitly, and the thick ones (the CA3 recurrent connections and the perforant path inputs to CA3) are modifiable. Each type of connection is all-to-all in the model. All inputs to CA3 pyramidal cells are gated by neuromodulatory signals (dotted lines) from septal nuclei, whose activity depends on familiarity with the current environment.

as computational considerations to propose some general constraints on how the EC spatial representation may depend on sensory features of the environment, and also suggest a plausible functional form for this dependence in the simple case that all the information about location that is directly available comes from the walls of the experimental apparatus.

The second basic assumption is that the network establishes new attractors to represent novel situations. This involves an orthogonalization process which

is assumed to take place in the dentate gyrus, as well as on-line modulation of synaptic plasticity and the relative efficacies of the different types of connections, controlled by familiarity with the environment, possibly via neuromodulatory signals from septal nuclei.

In this section, we provide a detailed description of the main components of our model, including the neural architecture (as shown in Figure 1) and dynamics, as well as the input representation. We then demonstrate the basic properties of the model by showing how place fields are generated in the simple case of a single environment surrounded by walls, using an idealized set of weights. In the next section, we tackle the issues related to learning, and introduce a familiarity-based on-line learning process for establishing an appropriate weight structure. The rest of the paper is devoted to modelling a set of more complex experimental paradigms. The values of the parameters used in the simulations are summarized in Table 1 at the end of the description of the model (on page 26).

CA3 neural architecture and dynamics

The main aspect of hippocampal circuitry we actually implement is the CA3 recurrent network (Figure 1). The model CA3 contains a collection of 1200 pyramidal cells, each connected to all the others through modifiable weights. This high degree of connectivity mimics the extensive recurrent collateral connections of CA3 pyramidal neurons (Ishizuka et al., 1990; Li et al., 1994). Due to the relatively small number of neurons in the model, the number of connections per cell is still much lower than in reality, even though the degree of connectivity is higher. This does not pose a problem, though, as long as the cells a particular neuron connects to can be considered from a functional point of view as a random sample, the number of connections per neuron is high enough, and any one connection is weak enough. In this case, neural responses are determined by averaged population effects, and the actual number of connections only enters the calculations as a constant scaling factor for the individual weights.

Local feedback and feedforward inhibition are thought to play an important and complex role in neural dynamics in CA3. Inhibitory interneurons are spatially much less selective than pyramidal neurons, but their activity during locomotion changes periodically at the theta frequency. We ignore this temporal variation as well as the diversity of interneurons and patterns of connectivity, and include in the model a single global inhibitory neuron, which fosters competition between stored patterns and keeps global activity levels approximately constant. This cell receives input from all the excitatory neurons, and provides inhibitory feedback to each which is proportional to the product of the firing rate of the inhibitory neuron and the depolarization of its postsynaptic target. This nonlinear form of inhibition was chosen because our simulations indicated that, compared to more conventional subtractive inhibition, it leads to improved robustness in the network with respect to variations in weight magnitude (for details on networks with shunting inhibition, see Grossberg, 1988). It is also consistent with the observed effect of $GABA_A$ receptor activation. We adapt the equations introduced by Wilson and Cowan (1972) to model the dynamics of the CA3 neural population. The following set of equations describes how the membrane potential of CA3 cells in our model changes over time:

$$\begin{aligned} \tau \dot{u_i} &= -u_i + \sum_j J_{ij} g_u(u_j) - h g_v(v) u_i + I_i^{PP} + I_i^{MF} \\ \tau' \dot{v} &= -v + w \sum_j g_u(u_j), \end{aligned} \tag{1}$$

where u_i is the membrane potential of the i'th pyramidal cell, v is the membrane potential of the global inhibitory cell (all relative to their resting potentials), τ and τ' are the membrane time constants for pyramidal neurons and the inhibitory cell, respectively, J_{ij} is the strength of the connection from neuron j to neuron i, h is the efficacy of inhibition, w represents the strength of the excitatory connection from any one pyramidal cell onto the inhibitory cell, and I_i^{PP} and I_i^{MF} are the inputs to cell i through the perforant path and the mossy fibers, respectively. $g_u(u) = \beta[u - \mu]_+$ is the threshold linear activation function for

the pyramidal cells, where $[...]_+$ makes all negative arguments zero while leaving positive numbers unaffected, μ stands for the threshold and β is the slope of the activation function above the threshold. Similarly, $g_v(v) = \gamma[v - \nu]_+$ for the inhibitory neuron. As will be described in detail later, some of the terms in these equations are assumed to be influenced by neuromodulatory control, and therefore may be absent in certain phases of processing.

The value of the inhibitory time constant τ' has no effect on the location of the fixed points of the network, although it can change their stability. In the simulations that are described later, we set $\tau' = 0$, so that v is always equal to $w \sum_j g_u(u_j)$. This simplifies the theoretical treatment of the model, and makes the simulations numerically more stable. We conducted simulations to verify that, within a wide range of the parameters, this manipulation does not affect the qualitative dynamical behavior of the model and indeed leads to the same stable patterns of activity. It is worth noting that, in this general class of models (although in a different parameter regime), setting $\tau' > 0$ can give rise to oscillations (which, of course, are consistently observed in the hippocampus during active behavior). Even in an oscillatory regime, however, the mean activities of the units can closely resemble the activities of the units at the fixed points found when $\tau' = 0$ (Li and Dayan, 1999).

Input representation

Instead of building a detailed model of rodent sensory processing, we consider as inputs to our model the firing rates of pyramidal neurons in superficial layers of entorhinal cortex, which provide most cortical input to the hippocampal formation. Unfortunately, there is relatively little direct experimental evidence about the nature of spatial representations in EC, and especially about how these depend on details of the environment. However, there is something of a consensus amongst modelers (e.g., Burgess et al, 1997), which we generally follow. Although entorhinal neurons are found to be spatially selective (Barnes

et al., 1990; Quirk et al., 1992), they appear to be much noisier and more broadly tuned than place cells in the hippocampus. Quirk et al. (1992) also found them to be more "sensory bound" than hippocampal cells in that their firing fields transform in a smooth manner following substantial changes in the shape of the environment. This is very unlike the complete remapping seen in place cells under similar circumstances (Muller and Kubie, 1987). The anatomy of the inputs to EC is rather better understood (Burwell and Amaral, 1998). Many of the inputs to EC come from higher-order association areas, which contain complex representations of the sensory information available to the animal. In particular, cells may convey information about both the identity of a perceived object and its location with respect to the animal, or, to put it differently, about the location of the rat with respect to particular objects in the world. Such information about multiple objects may be combined in EC in order to form a more reliable view-based representation of the animal's location in space. Spatial information derived from path integration may also be available, and may be combined with visual information to determine EC activities.

In the model, each EC cell is assumed to respond to a subset of the available cues. Based on the suggestion that EC is involved in conjunctive coding (Myers et al., 1995), each EC cell in our model combines in a conjunctive manner the sources of spatial information to which it is sensitive. Since the animal's sensory experience depends on both its position and the direction it faces, we assume (in the absence of data either way) that the activity of entorhinal neurons is head direction as well as location dependent. A model EC cell fires maximally when all the cues it is sensitive to are in the position corresponding to the cell's preferred location and orientation, and activity diminishes as some or all of the sources of information signal a different location or orientation. We achieve this by multiplying together Gaussian tuning curves, each of which is tied to the location of a different cue and peaks at the preferred location of the cell. We assume that these individual tuning curves can have different variances.

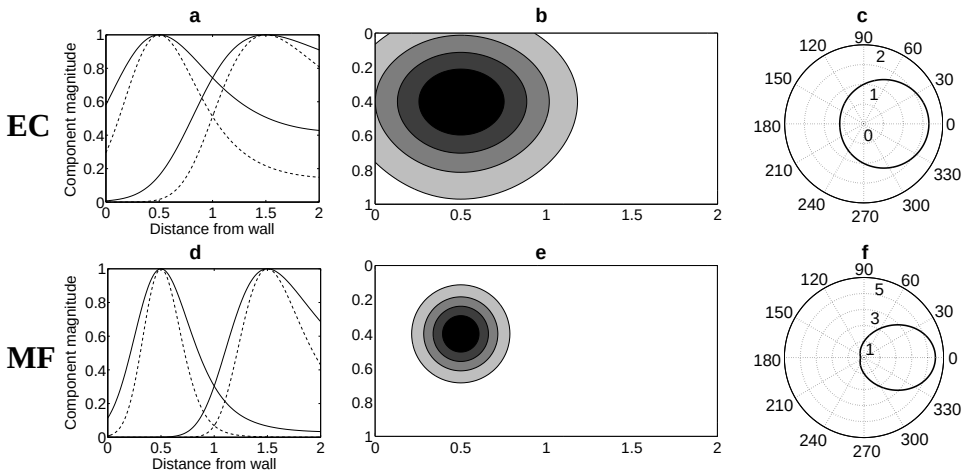

Figure 2: Input components and net spatial and directional tuning. **a**: The dependence of a single (essentially one-dimensional) spatial component of the tuning function of cells in entorhinal cortex (EC) on the distance of the rat from the wall to which that component is tied. Two examples are shown, with preferred distances of 0.5 and 1.5, respectively; for each preferred distance, the solid curve is for the case when the rat is facing the wall, and the dashed curve is for the opposite head direction. Note how the width of the curve changes with preferred distance and actual head direction. **b**: The net two-dimensional tuning of a sample EC neuron in a rectangular box of dimensions 2×1; the preferred location of the cell is (0.5, 0.4); the current heading of the rat at each location is the same as the preferred head direction of the cell. **c**: This polar plot shows the activity of an EC neuron as a function of the difference between its preferred direction and the actual heading of the rat. **d** to **f**: Plots similar to **a-c**, for the mossy fiber (MF) inputs to CA3; note that both the spatial and the directional tuning is much sharper here, due to the orthogonalization property of the dentate gyrus. For all contour plots in this article, darker shading indicates higher activity, and the contour lines are at 20, 40, 60, and 80% of the maximum activity of the given cell or set of cells. Activities are normalized and the absolute values are omitted in most figures since these could be set arbitrarily in the model by changing parameters essentially unconstrained by experimental data.

In cases where the environment has walls, these were found to be important sources of spatial information (O'Keefe and Burgess, 1996). For simplicity, we assume that the activities of EC neurons are completely determined by the rat's position and heading relative to the walls. We also restrict ourselves to rectangular environments, and assume that all cells are sensitive to the position of all four walls (whose allocentric bearings will be referred to as 'north', 'west', 'south', and 'east'). The only difference in the cue selectivity of EC cells in our

model is that they are assumed to be sensitive to spatial information derived from path integration to different degrees. However, since this last property is only expected to be manifested under special circumstances, we actually ignore this variation in most of what follows, and only consider it when we describe the results of our modelling of the experiment of Skaggs and McNaughton (1998). We assume that the tuning curve components tied to the walls of a rectangular apparatus are ridge-like functions with Gaussian dependence on the distance from the wall. The variances of these tuning functions may also depend on the location and heading of the animal – in particular, we assume that the variance is lower if the animal is closer to, or facing *away* from the wall. The latter dependence is based on the influence of a path integration input whose precision is greater when the animal is coming from somewhere nearer the wall and should have been able to maintain its location accurately using path integration.

The total activation of a model EC neuron as a function of the rat's location and heading is described by the following expression:

$$z_k = b\, z_k^N z_k^W z_k^S z_k^E e^{\rho_{EC} \cos(\phi - \phi_k^{EC})}, \tag{2}$$

where k indexes the neuron, b is a constant to set the scale, and z_k^a is the component of the neuron's tuning function tied to wall a. The last term describes the dependence on head direction (which is assumed to be independent from the spatial components) as a circular Gaussian function (with sharpness parameter ρ_{EC}) of the difference between the current head direction ϕ and the cell's preferred heading ϕ_k^{EC}. Equation 2 bears some resemblance to the spatial tuning function used by Touretzky and Redish (1996), in that it also takes the form of a product of terms corresponding to different sources of information. However, they use this tuning function to directly model the spatial response properties of hippocampal place cells, and the parameters change with experience, while our EC representation is always the same for a given location and head direction in any particular environment.

The components of the tuning function tied to particular walls have the following functional form:

$$z_k^a = e^{-\frac{(d_a - d_k^{EC,a})^2}{2\sigma_{EC,a}^2}}, \tag{3}$$

where d_a is the actual distance from wall a (a can be either N, W, S or E), $d_k^{EC,a}$ is the distance from wall a of the neuron's preferred location, and $\sigma_{EC,a}$ is the width of this component, which depends on the current position and heading of the animal according to

$$\sigma_{EC,a} = \sigma_{EC}(1 + 0.35d_a^2)(1 + 0.2\cos(\phi - \phi_a)), \tag{4}$$

where ϕ_a is the direction of wall a (0, $\pi/2$, π, and $-\pi/2$ for N, E, S, and W, respectively), and σ_{EC} is a constant. Equation 3 and the positional dependence in Equation 4 are similar to the expressions describing the spatial tuning of "sensory" cells in the model of Burgess et al. (1997), and "boundary vector cells" in Hartley et al. (2000). The numerical values of the parameters in the above equations have been chosen suitably for environments of around the size employed in most relevant experiments.

Figure 2a shows two examples of the spatial and directional dependence of input components in EC, while Figure 2b and c display the resulting net spatial and directional tuning for a sample EC neuron.

Feedforward connections

There are two separate neural pathways from EC to area CA3 (see Fig. 1), which have quite different characteristics and likely serve different computational purposes (McNaughton and Morris, 1987; Treves and Rolls, 1992). One of these pathways is via the perforant path projection to the dentate gyrus (DG), which in turn provides a set of feedforward inputs to CA3 through the mossy fibers.

Dentate granule cells are spatially selective, and, at least in linearly restricted environments, they have also been found to be sensitive to direction (Jung and McNaughton, 1993). Unlike EC neurons, dentate granule cells have sharper spatial tuning than CA3 place cells, and we assume that they are also sharply tuned for head direction. Episodic memory theories of hippocampal function suggest that an important function of the DG is that of orthogonalization, i.e. reducing the similarity between input patterns in order to facilitate their discrimination (Treves and Rolls, 1994; O'Reilly and McClelland, 1994), and, in keeping with the theme of linking memorial and spatial processing, we assume it plays a similar role for spatially based inputs. One way the DG is thought to decrease pattern overlap is to implement a sparser representation (perhaps through direct competitive interactions), and indeed, the proportion of active cells in the DG at any given time is reported to be only about 0.5% (B.L. McNaughton, cited by O'Reilly and McClelland, 1994; see also Jung and McNaughton, 1993).

A typical CA3 pyramidal cell receives on the order of 50 mossy fiber (MF) inputs, which are thought to be relatively powerful (Yamamoto, 1982; McNaughton and Morris, 1987). Combined with the sparseness of the DG representation, this means that a CA3 neuron is very unlikely to have more than one active mossy fiber input at any given time. In circumstances under which CA3 cells are driven primarily by these inputs, place cells essentially inherit the tuning characteristics of their afferent granule cells. We assume, for simplicity, that each CA3 cell has at most one active MF input in any given environment. This defines the base preferred location and direction for that neuron, which, of course, may then be altered by the recurrent connections in CA3. Multiple active MF inputs may explain why some place cells have multiple place fields even in simple environments (Muller et al., 1987); however, we ignore this complexity for the purpose of this paper. In addition, in order to make better use of the limited number of cells we can implement in our numerical simulations, all our model CA3 pyramidal cells are activated by MF inputs somewhere in any given environment, rather than the 30% or so found in practice (Wilson and McNaughton,

1993).

In its current form, the model considers both the mossy fiber connections and the perforant path connections from EC to DG as being fixed. Since our goal is to model activity in CA3, and that is completely determined by its inputs and internal dynamics, we can therefore skip modeling the dentate gyrus explicitly, and proceed by characterizing how the MF input to CA3 (which results from processing in DG) depends on the characteristics of the environment. We assume that, for any single environment, the MF input to CA3 place cells has a similar functional form to the tuning function of EC cells described in the previous section, but both the spatial and the directional tuning is assumed to be sharper as a result of sparsification and orthogonalization in DG (see Fig. 2d-f). This can be achieved by replacing the spatial spread parameter σ_{EC} with a smaller value, σ_{MF}, and by replacing ρ_{EC}, characterizing the sharpness of directional tuning, with a larger ρ_{MF} in Equations 2 and 4. The proposed orthogonalization property of the dentate gyrus becomes more pronounced when we look at multiple environments. We assume that, except when two environments are quite similar, the MF inputs to CA3 in two different environments are completely unrelated. We will return to the case of exceptionally similar environments in a later section.

The perforant pathway (PP) also provides a direct connection between EC and CA3, and has a large degree of divergence and convergence. Thus, CA3 cells can sample the EC representation very effectively. In the model, we implement this property using all-to-all connections between EC and CA3 neurons, although this is obviously a simplification. This pathway is also known to be capable of long-term synaptic plasticity (Breindl et al., 1994). In the model, the strengths of these connections (denoted by W_{ik} for the connection from entorhinal cell k to CA3 cell i) are initially set to zero, and they are assumed to be modifiable by associative Hebbian learning.

Network dynamics

Although we will shortly be interested in the spatial representation that results from on-line learning during exploration, we first test our model using an idealized set of connection strengths in order to gain some insight into its dynamical behavior. For this, we just assume that the weights result from an idealized form of Hebbian associative learning, and thus reflect the correlations between connected neurons. It has been noted (Muller et al., 1991; Shen and McNaughton, 1996) that such an associative learning process for spatially selective neurons can lead to connections whose strength is a function of the distance between the preferred locations of the pre- and post-synaptic neurons, exactly the sort of connections that can support a place-field-like attractor structure in CA3 (Samsonovich and McNaughton, 1997). Here we assume that the CA3 recurrent weights are determined by the correlations between the mossy fiber inputs to the cells, and the perforant path weights between EC and CA3 are given by the correlations between EC activities and MF inputs to CA3. These correlations are calculated as spatial averages (which assumes spatially homogeneous exploration) over all locations and head directions in the part of the environment where the postsynaptic cell is active, resulting in the following expressions for the recurrent weights J_{ij} and perforant path weights W_{ik}:

$$J_{ij} = \kappa \frac{\iiint\limits_{I_i^{MF}>0} I_i^{MF} I_j^{MF}\, dx\, dy\, d\phi}{\iiint\limits_{I_i^{MF}>0} dx\, dy\, d\phi} \quad \text{and} \quad W_{ik} = \kappa \frac{\iiint\limits_{I_i^{MF}>0} I_i^{MF} z_k\, dx\, dy\, d\phi}{\iiint\limits_{I_i^{MF}>0} dx\, dy\, d\phi} \tag{5}$$

where I_i^{MF} is the mossy fiber input to neuron i in CA3, z_k is the activity of neuron k in entorhinal cortex, and κ sets the learning rate.

Using these expressions, we can calculate the weights resulting from even exposure to a rectangular box (with one side twice as long as the other). Then, letting $I_i^{PP} = \sum_k W_{ik} z_k$ and $I_i^{MF} = 0$, where z_k is the EC activity pattern corresponding to a particular location and heading in the environment, we sim-

ulate the neural dynamics described by the full Equations 1 for a fixed number of iterations (using Euler's method). We find that, within a broad range of model parameters, the network always settles into a stable state by the end of the iterations. Furthermore, for most initial CA3 activity patterns, the same final state is reached for given feedforward inputs. This shows that these states are actually attractors of the neural dynamics, and that they have suitably large basins of attraction. The final state of the network was determined for different input patterns in EC, representing different positions and head directions of the animal over a grid that covered the whole environment. The firing rate map for a given cell is defined as the final activity of that cell as a function of the actual location and head direction of the animal.

Throughout the paper, two different kinds of plots are used to display the activities of neurons (and their inputs). Quantities characterizing single cells as a function of actual position and heading (such as firing rate maps) are shown in "single cell plots", which is the kind of plot traditionally used to describe the spatial activity patterns of place cells. A single cell plot may contain multiple subplots to represent different headings at any given location. The second kind of plot we use is the "population plot", which describes the behavior of *all the cells* with the actual position and heading of the animal kept fixed. In the population plot, we arrange cells with a given preferred direction on an imaginary plane according to their preferred locations (for CA3 place cells, this is defined as the preferred location of their active mossy fiber input). A complete population plot would include 8 subplots, one for each population of cells with a different preferred direction, but we typically show only 1, 2, or 4 of these, depending on the degree of variation with preferred direction in that particular case. Population plots in this paper are marked with 'P' in the lower right corner for easy identification.

The results of the simulations with the "ideal" weights are summarized in the population plots of Figure 3, which display activities in EC, net perforant path

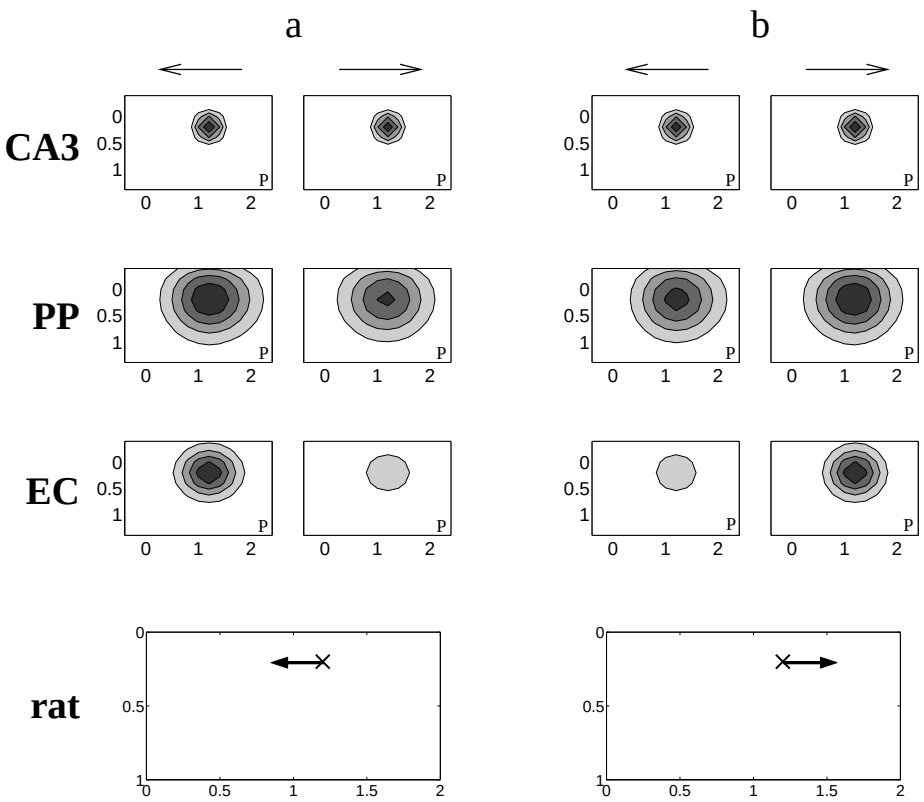

Figure 3: The formation of non-directional place fields. **a,b**: The bottom plot in each case shows the actual position (indicated by the cross) and head direction (indicated by the arrow) of the rat in the environment. The other plots are population plots (as defined in the text, and marked with 'P'), and they show, at the location and direction in the bottom plot, the activities of cells in entorhinal cortex (EC), the net perforant path (PP) inputs to CA3 neurons (I_i^{PP}), and the final activities of the same place cells (marked CA3), as a function of the preferred location of the neuron; the two columns in both **a** and **b** are for cells with preferred head direction indicated by the arrow above each column.

inputs and final activities in CA3 for all cells with two particular (opposite) preferred head directions, when the model rat is at a given location, facing in a particular direction. Figure 3 shows that the final states of the model CA3 network resemble thresholded two dimensional Gaussian bumps of activity in the population plot. This type of solution can emerge spontaneously from the network dynamics even in the absence of external inputs, in which case the location of the bump is random – i.e., determined by the initial neural activities, as well as various other factors including the distribution of preferred locations

and directions of the neurons. Even though the network only has a finite number of point attractors (possible stable activity patterns) in the absence of input, when there is even a small perforant path input to CA3, the location of the bump is determined by this input so that the activity profile provides the best possible fit to the input. The position of the peak varies continuously, and the shape of the activity profile is essentially constant. This holds in our model if the net feedforward input to the most active CA3 neurons is between roughly 1% and 30% of the summed input they receive from other CA3 cells; in most simulations we set the relative efficacies of perforant path and recurrent synapses so that this ratio is about 5%.

Figure 3a illustrates how inputs are used by the network to effectively select one of the possible final states. First of all, the EC activity pattern (which is determined by sensory features in the environment as already described) gives rise to a pattern of perforant path inputs to CA3 which is centered on neurons with preferred locations close to the actual position of the rat, although the profile is even broader than the activity profile in EC. This is the consequence of plasticity of the perforant path in the learning phase, which establishes an association between EC cells and CA3 neurons with similar preferred locations and head directions. Based on the learned weights, the PP projection also reduces directionality substantially, so that inputs to CA3 already depend less on the preferred head direction of the cell than neuronal activities in EC. The shape of the final activity profile across place cells is, however, essentially determined by the CA3 internal dynamics, resulting in a spatial activity profile which is much more sharply peaked than the feedforward inputs. Further, the final activities of the cells are essentially independent of their preferred head direction. The resulting model place fields (i.e., single cell activity maps) possess many of the characteristics of real place cell firing patterns recorded in open environments. As we will see later (e.g., in Figure 6a), they are unimodal, approximately Gaussian with circular symmetry, and essentially non-directional.

Figure 3 reveals how non-directional place fields result despite the directional input representation in EC. The two parts of the figure compare the activities of EC neurons, the PP inputs to CA3 place cells, and the final activities of place cells as the model rat faces in two opposite directions at the same location. Due to the properties of the PP projection discussed above, place cells receive relatively similar inputs in the two cases. More importantly, however, this leads to the emergence of the same, non-directional, attractor in CA3, making the place fields independent of head direction. It should be emphasized that, given the dominance of internal connections in determining the final state of the system, even PP inputs as similar as those in Figure 3 a and b could easily lead to fundamentally different patterns of final activity if the two input patterns biased the system towards different attractors. Indeed, these same two EC input patterns do actually give rise to two very dissimilar final patterns if the weights are set up during a directed search task like the one described later (instead of the omnidirectional random exploration assumed here).

On-line learning of attractors

So far we have assumed that weights proportional to the spatially averaged correlations between cells had been established by an appropriate learning procedure before spatial activity patterns are measured. We have not yet shown that a neurobiologically standard Hebbian learning rule, applied to the activity patterns occurring in the network during random exploration of an environment, is capable of establishing this kind of weight structure, within the time window during which place fields are seen to develop in experiments (on the order of 5 minutes; Wilson and McNaughton, 1993).

A general property of attractor networks is that, in order to store more than a single pattern, the recurrent connections need to be suppressed while new patterns are learned. Experimental data and theoretical considerations have been

adduced to justify models of CA3 in which the relative strengths and adaptability of mossy fiber input and perforant path and recurrent collateral input is different between initial learning about an environment and recall of information within a familiar environment. We adopt the suggestion of Hasselmo et al. (1996) which is based on experimental data on the effects of septal (cholinergic and $GABA_B$-receptor-mediated) modulation in the hippocampus.

In particular, substances that activate muscarinic cholinergic receptors or $GABA_B$ receptors in the hippocampus were found to selectively suppress excitatory recurrent synapses in area CA3 compared to feedforward excitatory connections (Ault and Nadler, 1982; Hasselmo et al., 1995). In addition, cholinergic input to the hippocampus has been shown to enhance long-term synaptic plasticity (Burgard and Sarvey, 1990; Huerta and Lisman, 1993), and leads to the suppression of inhibition (Pitler and Alger, 1992) and the direct depolarization of hippocampal pyramidal neurons (Benardo and Prince, 1982). These effects of cholinergic modulation create exactly the right circumstances for the learning of new information in the hippocampus while minimizing interference from previously stored information. This is convincingly illustrated by the associative memory model of Hasselmo et al. (1995), where several moderately overlapping input patterns can be stored and recalled successfully using feedback cholinergic modulation of network parameters.

It turns out that attractor networks with continuous attractors, such as ours, face a more stringent requirement for learning because of potential bias in the sampling of a continuous set of input patterns, and we therefore consider a slightly different model of neuromodulatory control. In the resulting on-line learning procedure, plasticity is gated by familiarity, and we show that it leads to weights similar to those in the "ideal" model described above, and thus a place cell representation similar to ones observed experimentally.

In our model, the hippocampal network has two modes of operation. When the rat first encounters a new environment, learning in both the PP inputs to CA3

and the CA3 recurrent synapses is enabled, synaptic transmission through the recurrent connections is suppressed, inhibition in CA3 is reduced, and inputs through the mossy fiber connections dominate. This state of the network is called "learning mode". On the other hand, when the rat is in a highly familiar environment, no learning takes place in any of the connections, the MF inputs are relatively less effective than the PP connections and CA3 recurrent synapses, and the intrinsic dynamics of the recurrent network dominates activity in CA3, leading to previously established attractors. This is called "recall mode".

Initial learning in a novel environment is essentially input-driven due to the suppression of recurrent activity, but this phase is responsible for setting up the attractors and feedforward associative projections which determine the patterns of place cell activity seen subsequently. Note that synapses are modified even when their efficacy is reduced to zero by neuromodulation, i.e., when the post-synaptic effect of perforant path and recurrent connections is negligible in the learning phase.

Perforant path and recurrent weights are acquired during the learning phase. The neural dynamics described by Equations 1 is simplified substantially in this phase by making the recurrent connections ineffective and neglecting inhibition, leaving $\tau \dot{u_i} = -u_i + I_i^{MF}$. Assuming that the MF inputs change more slowly than the membrane time constant, the membrane potential of CA3 place cells during the learning phase is given by $u_i = I_i^{MF}$. Application of a Hebbian learning rule (with the addition of weight decay to prevent weights from growing indefinitely) to these activities leads to weights which are proportional to the temporally averaged correlations between pre- and postsynaptic cells. The only difference between these weights and the "ideal" ones used in earlier simulations is that while the "ideal" weights were obtained by averaging the product of pre- and post-synaptic activities across spatial locations and headings, this on-line method calculates averages across time. The two processes become exactly equivalent if we assume that, during initial exploration in the environment, the

rat receives even exposure to all combinations of location and head direction allowed by the apparatus (and the movement pattern followed).

However, there are two potential differences between uniform spatial averaging and temporal averaging over random exploration, the first coming from any systematic spatial bias (which depends on the exploration strategy), and the second coming from random deviations from this ideal, if biased, exploration. Figure 4a shows a sample path from a simulation of a common exploration paradigm (which is essentially equivalent to experiment 1 in Markus et al., 1995). This shows the first five minutes of exploration in a new environment whilst a rat chases food pellets thrown into random locations in a rectangular apparatus. Once it has retrieved one pellet, the next one is thrown in at random. We assume that the rat runs at a constant speed V, and it always heads essentially in the direction of the next food pellet, with random fluctuations in direction. The exact movement laws and parameters were taken from the model of Brunel and Trullier (1998).

Even at a first glance, this exploration strategy clearly results in an inhomogeneous sampling of the environment. We quantify variations in exposure to different locations and directions in the apparatus by convolving the sample path with a Gaussian, yielding the function

$$g(\mathbf{x}^*, \phi^*, t) = \int_0^t e^{-\frac{(\mathbf{x}(t')-\mathbf{x}^*)^2}{2\sigma^2}+\rho\cos(\phi(t')-\phi^*)}\,dt' \tag{6}$$

where t is time since the beginning of exploration, $\mathbf{x}(t)$ and $\phi(t)$ are the rat's position and heading at time t, σ is the width of the spatial Gaussian, and ρ is the sharpness of the circular Gaussian applied to differences in direction. This measures sampling density as a function of position and direction, and an example (after 5 minutes of exploration, averaged over all directions) is shown in Figure 4b. There is clear deviation, both random and systematic, from a uniform sampling density. The random aspect of the deviation turns out to be benign, since it does not destroy the overall structure of the attractors.

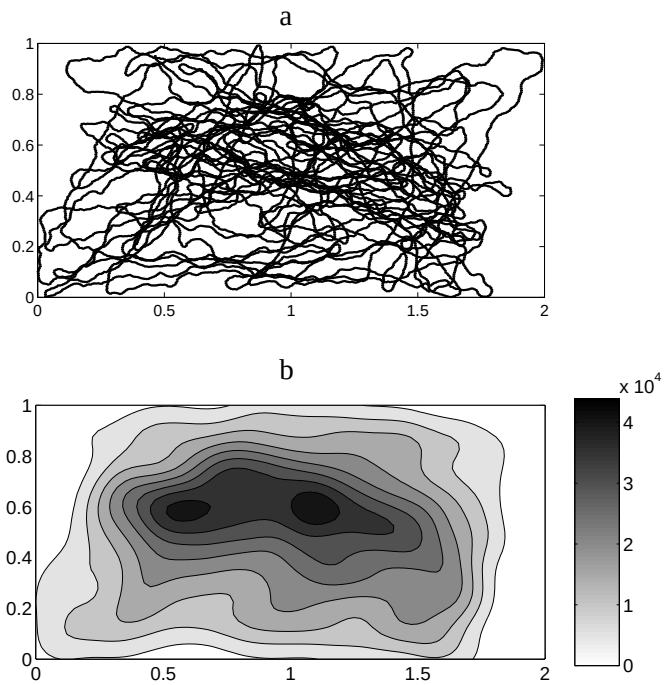

Figure 4: Non-uniform sampling of the environment during random exploration. **a**: An example trajectory, showing the first 5 minutes of exploration in our simulation of a common paradigm in which the rat chases food pellets thrown into random locations in the environment. Note that some parts of the environment are visited much more often than others. **b**: Convolution of the path in **a** with a 2-dimensional Gaussian ($\sigma = 0.075$), which measures exposure to locations in the apparatus ($g(\mathbf{x}^*, \phi^*, t)$), summed over all directions, after 5 minutes of exploration.

However, the fact that, on average, the animal spends several times as much time at a location near the center of the apparatus than at a location near the edges, causes the naive on-line Hebbian learning procedure to produce a non-uniform weight structure, resulting in a very poor place cell representation. An example of this is given in Figure 5a – the network possesses just two or three distinct attractors, and only neurons which are active in one of these attractors ever become active in this environment. This effect cannot be mitigated by increasing exploration time, and is also persistent with respect to the specifics of the movement laws. In particular, even though rats have a tendency to stay close to the walls of the apparatus (e.g., Muller et al., 1987), this is unlikely to precisely counterbalance the effect described above and result in spatially and

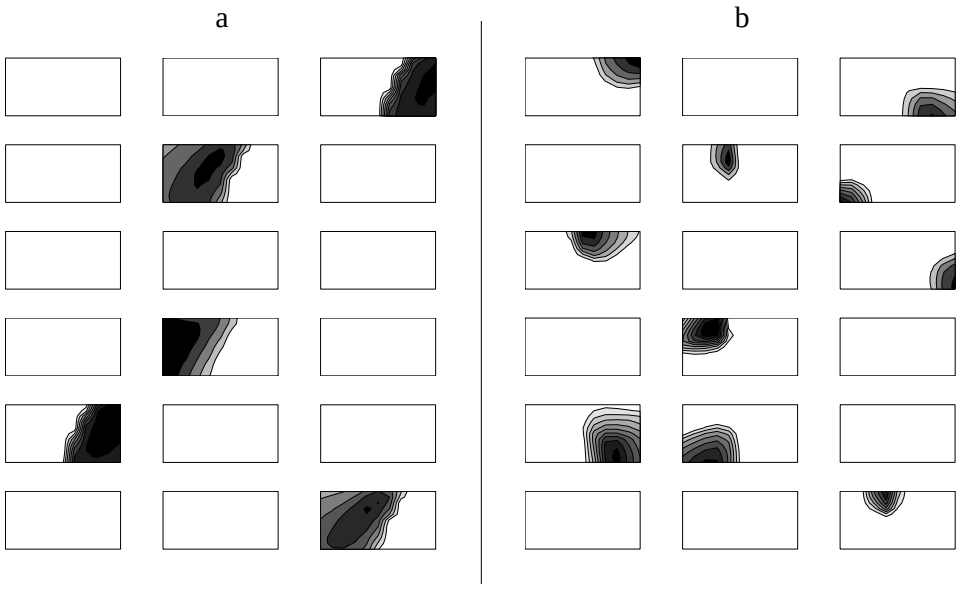

Figure 5: Place cell firing patterns during recall, after using different learning procedures. The figure shows the firing rate maps of 12 randomly selected CA3 place cells after the exploration shown in Figure 4a, **a**: using simple Hebbian learning, and **b**: using the familiarity-based learning procedure to establish the weights. The place fields in **(b)** closely resemble experimental place fields, and provide good coverage of the whole environment. Conversely, the spatial firing patterns in **(a)** reflect essentially two different attractor states containing only a small proportion of the neurons, perturbed to some extent by the feedforward input.

directionally unbiased exploration.

Systematic differences in sampling density have a profound effect on the resulting attractor structure because of the continuous nature of the set of patterns that need to be represented by the network. This requires the set of recurrent weights to be such that the activity patterns corresponding to all different positions in the environment are equally stable. Continuous attractor networks like ours are generally known to be very sensitive to the regularity of the recurrent weight structure (Zhang, 1996; Pouget et al., 1998), and most such previous models were forced to set these weights by hand.

Using all patterns indiscriminately during on-line learning is also questionable from a computational point of view, especially in the presence of substantial

sampling bias. Learning should be gated by familiarity – the more familiar a part of the environment, the less about it that should be learned. Figure 4a shows that familiarity is actually a graded quantity – since the animal has more exposure to the center of the environment than the perimeter. Therefore, we use a graded familiarity signal, like the one proposed by Hasselmo et al. (1995). Note that it is not clear how familiarity is measured; for instance, Hasselmo et al. (1996) even suggest that a feedback loop involving the septal nuclei and the hippocampus itself might be responsible. We adopt the simple procedure of using the exposure measure of Equation 6, gating learning according to

$$\dot{J}_{ij} = e^{-\alpha g(\mathbf{x}(t),\phi(t),t)}\ u_i(t)\ g_u(u_j(t)) \quad \text{and} \quad \dot{W}_{ik} = e^{-\alpha g(\mathbf{x}(t),\phi(t),t)}\ u_i(t)\ z_k(t) \tag{7}$$

where σ, ρ and α were determined so that the amount of learning that occurs in different parts of the apparatus is as uniform as possible after 5, or more, minutes of exploration. The application of this learning procedure results in an attractor structure not very much different from the one defined by the "ideal" weights described earlier, and, as shown in Figure 5b, leads to a good place cell representation after just 5 minutes of exploration, in agreement with experimental data. The weight structure becomes increasingly uniform, given more exploration, and the place fields duly become increasingly regular.

Although the efficacies of the different types on inputs to CA3 cells may also be modulated in a graded fashion (this may even involve the same signals which modulate plasticity), we currently use a simple heuristic based on the notion of two distinct processing modes as described above. Recall mode is entered after a fixed amount of exploration per unit area of the environment (by which time learning has saturated essentially everywhere in the environment), or immediately upon entry into the environment if it is similar enough to an environment already explored; otherwise, learning is initiated. More precisely, we skip learning in a new environment only if it shares most sensory features with an

Number of CA3 place cells	N	1200
EC tuning amplitude	b	100
EC spatial tuning width	σ_{EC}	0.4
EC directional sharpness	ρ_{EC}	0.5
MF spatial tuning width	σ_{MF}	0.2
MF directional sharpness	ρ_{MF}	1.5
Excitatory time constant	τ	100
Inhibitory time constant	τ'	0
Inhibitory feedback weight	h	3
Pyramidal-to-inhibitory weight	w	0.005
Excitatory gain	β	1
Excitatory threshold	μ	80
Inhibitory gain	γ	1
Inhibitory threshold	ν	12
Running speed	V	0.3
Spatial spread of familiarity	σ	0.075
Directional spread of familiarity	ρ	3
Novelty decay rate	α	250

Table 1: Model parameters. The table displays the values of model parameters used in the simulations. σ_{MF}, σ_{EC}, and σ are in units such that the shorter side of the rectangular environment used in most simulations is of unit length. τ and τ' are given in time steps used during simulations of recall, and all other quantities are in their natural units. Note that since the parameters only appear in certain combinations in the equations, some groups of parameters can be changed together appropriately without affecting the behavior of the model.

environment which is completely familiar to the animal, i.e., one that has been thoroughly explored.

Modeling more complex paradigms

So far we have shown that an attractor-based model, using weights defined by correlations between the feedforward activations of cells, can account for many of the experimentally observed basic properties of the CA3 spatial representation. We have also described a two-mode on-line learning process which computes an approximation to these "ideal" weights, and results in a very similar, although slightly less regular, place cell representation. In this section, we show how our model can also account for experimental results in a number of more complex paradigms, including the task-dependence of place field directionality, the co-

existence of several orthogonal representations for very different environments as well as overlapping representations for very similar environments, and the transformations of place fields following manipulations of the environment. We ran all simulations using both idealized, correlation-based weights and those resulting from on-line learning, and got qualitatively similar results in all cases. Most figures display results obtained using the "ideal" weights, because these tend to illustrate our points more clearly due to the lack of randomness.

Task-dependence of directionality

We have already described how random exploration in an open environment can lead to non-directional place fields (an example of which is shown in Figure 6a), through the establishment of appropriate attractors in CA3. In agreement with the recent modelling study by Brunel and Trullier (1998), we found that the ability of the recurrent network to suppress the directionality of the inputs depends critically on the set of locations and head directions experienced by the rat during learning. Place cells become direction independent only in situations in which the animal is exposed to a wide range of directions at a particular location. On the other hand, when the behavioral task or the environment itself constrains the set of directions experienced at a given location, as in a radial maze or when the rat is required to follow a specific route in an open field, place cells retain their intrinsic directionality. Even in these cases, the width of directional tuning can, however, be modified by the recurrent network. These results are in good agreement with experimental findings (Muller et al., 1994; Markus et al., 1995). The dependence of directionality on movement patterns is illustrated in Figure 6b, which shows the place field of the same model CA3 cell that appears in Figure 6a, for a rat which has performed a different behavioral task in the same environment. In this task, which can be thought of as a simplified version of the directed search task described by Markus et al. (1995), the rat is required to run back and forth between the two shorter walls of the

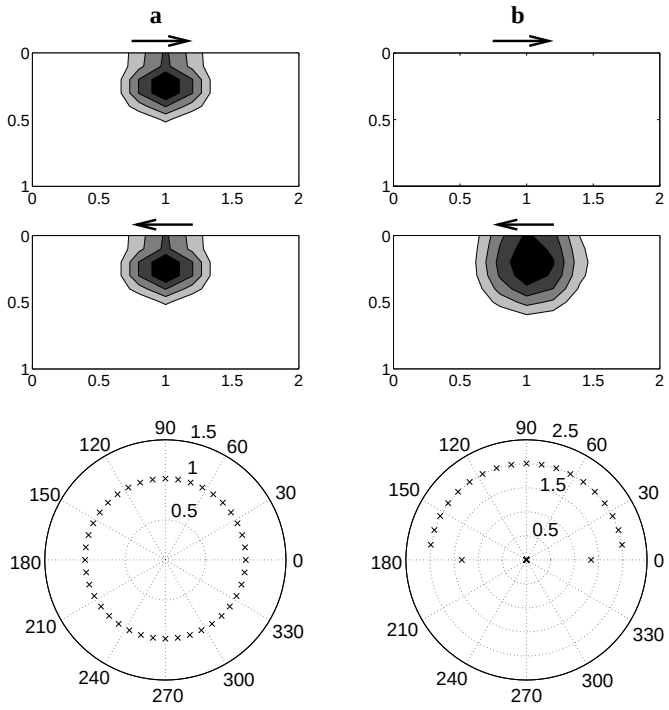

Figure 6: The task dependence of directionality. The contour plots show the place field of a CA3 cell which prefers the left direction, when the rat faces in the direction indicated by the arrows, and the polar plots show the maximum firing rate (indicated by the crosses, and relative to the maximum rate when averaged across directions) of the same neuron as a function of head direction; **a**: in a model rat which explored the environment randomly during the learning phase; **b**: in a model animal which always ran in one of the directions parallel to the long walls of the box during learning. The top plot is empty in **b** because the cell does not fire at all in that direction in this case. The effect of the attractor dynamics is very prominent in the all-or-none nature of activity in the directional plot in **b** (all the points in the bottom half of the plot collapsed to the origin). The maxima of the top and bottom contour plots correspond to the crosses at 270 and 90 degrees, respectively, in the polar plots.

environment to obtain reward. For the idealized case, we model this by assuming during exploration that the rat is now exposed only to the two directions parallel to the long walls instead of all directions at each location. Everything else in the simulations is left the same. This change affects the correlations between place cells in the learning phase, resulting in altered weight structure, which, in turn, changes the attractors. In agreement with experimental data, the new attractors do not eliminate the directionality of the inputs to the place cells. In fact, two

very distinct sets of attractors are established, one corresponding to each of the two directions sampled during learning. To illustrate this point, in the bottom half of Figure 6, we plotted the maximum activity of the place cell shown at the top of the figure, as a function of the animal's heading. In the rat trained using random exploration, the activity of the cell is essentially direction independent; however, if we train the rat in the shuttling task instead, the cell fires at a high rate for all directions with a westward component, and is completely silent for all directions with an eastward component. Cells which prefer direction east behave in exactly the opposite way. Once more, the results of simulations with the on-line learning procedure are similar to those obtained using the "ideal" weights, although the differences in directionality between the two training paradigms are generally somewhat reduced, and activity changes with head direction tend to be more graded. (For further discussion on how different behavioral paradigms might lead to spatial representations with different degrees of directionality, see Brunel and Trulier (1998)).

Very different environments

Experiments in which the firing rate maps of place cells are recorded in multiple environments which are similar to a controlled degree can provide valuable information about how input representations depend on details of the environment, how they are transformed into the place cell representation, and also about possible interference between representations of different environments realized by the same network of place cells. The general pattern of results is that radically different environments give rise to very different, and apparently unrelated place cell representations (O'Keefe and Conway, 1978; Muller and Kubie, 1987; Bostock et al., 1991). On the other hand, when a previously familiar environment is subjected to subtle alterations, the place cell representation often stays basically the same (O'Keefe and Conway, 1978; Bostock et al., 1991), or changes according to the transformation of the environment (Muller and Kubie, 1987;

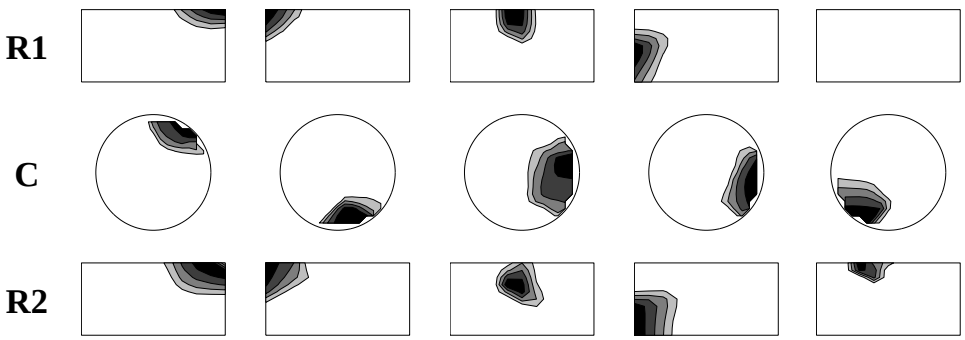

Figure 7: Very different environments. This figure shows the place fields of 5 selected place cells in a rectangular and a circular apparatus which have very different sensory features. The top row (marked R1) shows the place fields after learning in the rectangular apparatus but before any experience in the circular one, and the other two rows show the place fields in the circular and rectangular environments (marked C and R2, respectively) after the rat has become familiar with both. There is no obvious relation between place fields of the same cell in the two environments. The effect of encoding a second environment on the place cell representation in the first environment can be assessed by comparing the top and bottom rows of this figure. Although there are some visible changes, these tend to be small, and do not affect the general structure of the spatial representation. One of the few exceptions is shown on the far right, where a place cell which had been silent in the rectangular environment becomes active there after experience in the circular environment.

O'Keefe and Burgess, 1996).

In order to test our model in the first type of situation, we added another model environment to the one described in the previous section, and tested whether these two environments can be learned and recalled simultaneously without interference. The two environments are very different in terms of visual appearance; the new environment has a circular shape, and is assumed to carry visual features that are dissimilar to the ones in the rectangular box. Therefore, we assume that the spatial characteristics of both EC neuronal activities and mossy fiber inputs to CA3 as well as their relations are completely independent in the two environments; i.e., for instance, knowing the relative locations of maximum activity for two EC neurons in one environment carries no information about the relation of their preferred locations in the other environment. However, as a worst case scenario, we use exactly the same neuronal populations to represent

the two environments; if these populations are distinct to any extent, this can only improve the separability of the two environments. Since we are interested in interactions between different environments, and not in extending our input model to curved walls and other cues, we derive the inputs in the circular environment assuming that there is a very salient square box (which looks very different from the rectangular box) surrounding the circular arena so that the inputs are determined by distances from the walls of the square box in the same way as before.

Initial learning in the rectangular environment is performed using the on-line procedure described in the previous section, and the resulting place cell firing patterns are determined as before. Then the weights are modified by running a learning phase in the circular environment, and spatial firing distributions during recall are determined in both environments to assess interference caused by exposure to the other environment.

Figure 7 shows the firing rate maps of 5 model CA3 cells in the rectangular apparatus before any exposure to the circular environment (top row of figure), and in the rectangular and the circular apparatus after learning in both environments (middle and bottom rows). In general, there is no systematic relation between the location of place fields in the two different environments, which indicates that several different sets of attractors can be stored and recalled independently in the model.

Comparing the top and bottom rows of Figure 7 reveals that most place cells have very similar firing rate maps in the rectangular box before (R1) and after (R2) training in the circular environment. In particular, for the majority of CA3 cells, the location of maximal firing, the size, shape, and directionality (not shown) of the place field are all virtually unchanged. Consequently, the overall structure of the spatial representation is essentially unaffected by exposure to a different environment. However, for a minority of place cells, experience in the circular environment resulted in a more radical change in the firing rate map in

the rectangular box (as in the last example in Figure 7). The most commonly observed types of change were the appearance of a new place field and the disappearance of one previously present. These probably occurred when the changes in the net input received by the cell – resulting from the weight changes that took place in the other environment – caused the neuron to cross the dynamic threshold for activation. Learning to represent a new, "orthogonal" environment can be thought of as introducing noise into both the feedforward and the recurrent weights as far as the representation of the original environment is concerned.

In order to quantify the change caused by exposure to a different environment, we computed the overlap between the overall CA3 spatial representations in the rectangular box before and after learning in the circular environment. To obtain a scale against which we can measure differences in overlap, and also to facilitate direct comparison with experimental data, we generated from our firing rate maps a large number of spike count samples, assuming independent Poisson noise for all cells and bins. Maximum firing rates, bin sizes, and session time were similar to those in experiments (e.g., Muller et al., 1987). The correlation coefficient between samples from the R1 and R2 spatial representations was found to be 0.754 ± 0.001 (mean and standard deviation), which is significantly lower (t-test, $P < 0.0001$) than the correlation between different samples from R1 (0.911 ± 0.001), but significantly higher ($P < 0.0001$) than the correlation between samples from R1 and a version of R1 where place cells have been randomly reshuffled (-0.005 ± 0.0005). These figures confirm our observation that, although there is a certain degree of degradation, the spatial representation after learning in an "orthogonal" environment remains quite similar to the original one. Furthermore, since the number of neurons and connections is much larger in the real hippocampus than in the model, and not all neurons are active in any particular environment, interference between representations of different environments is likely to be less severe, and the number (and perhaps the spatial extent) of environments that can be stored is probably larger.

Finally, our model would also produce orthogonal place cell representations for environments that differ only in shape (Muller and Kubie, 1987), even from non-orthogonal input representations (Quirk et al., 1992), provided that the DG can separate the input patterns effectively, and the two environments are perceived as different so that learning is initiated in both environments.

Geometric Manipulations

We also investigated what happens to place fields in our model if the environment undergoes some simple geometric transformation. We chose to model the experiment of O'Keefe and Burgess (1996) because of its relatively complex pattern of results. In this experiment a rat, which has been thoroughly familiarized with a rectangular box, is transferred into a new box that differs from the original one only in the length of one or both sides. We will concentrate on the case when the second environment is a larger square box which can be obtained by stretching the original box by a factor of two. In this case, stretching the environment had one of the following general effects (O'Keefe and Burgess, 1996): some fields remained fixed with respect to one of the walls of the apparatus; some changed their location and/or shape in correspondence with the transformation of the box; others developed a second peak in the direction of stretching. Many of the cells with two-peaked or stretched fields also developed directional dependence; i.e., the location of maximum activity depended on the heading of the rat, usually in the way that the subfield closer to a wall was more active when the rat was facing away from that wall.

We assume that learning is triggered by exposure to the novel situation of the initial, rectangular box, and that the transformed environment in this case is similar enough to the original one so that no significant learning occurs subsequently. Therefore, the attractors established in the first environment are the final states of the network dynamics in the new environment as well, and place fields are determined by the way that the inputs (as a function of location

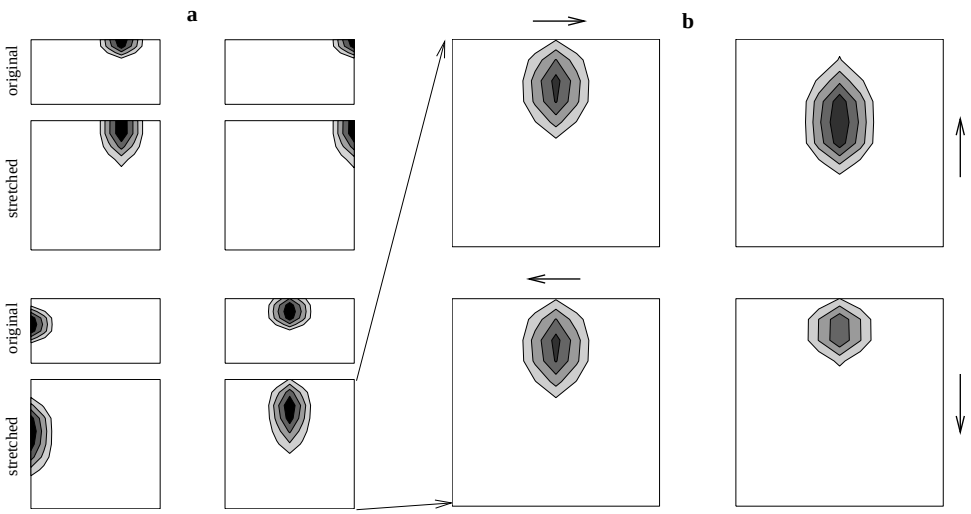

Figure 8: Place fields in transformed environments. **a**: The place fields of four selected cells in the original and the stretched environment in our simulation of the experiment by O'Keefe and Burgess (1996); the firing rates shown are averages over all head directions. **b**: Directionality of the place field shown in the bottom right corner of part a; the place field depends on the heading of the rat (indicated by the arrows). This dependence on head direction is induced by the transformation of the environment; place fields in the original environment are essentially non-directional (like the one shown in Figure 6a).

and direction) in the new environment *select* attractors established in the old environment.

Figure 8a shows the place fields of four model CA3 neurons in the rectangular box which was used during initial learning, and in the larger square box. The place fields follow the transformation of the box; that is, their centers remain at the same relative distance from opposite walls, and their shapes become elongated along the direction of stretching. As revealed by Figure 8b, the fields consist of directional subcomponents with the observed relation between subfield position and preferred direction.

We can understand some of the characteristics of transformed place fields by looking at attractor selection in our model. Attractors have a regular, compact shape if place cells are characterized by their preferred locations in the original environment; on the other hand, we have no a priori knowledge about what they

look like as a function of preferred locations in the new environment. Thus, it is much easier to understand the transformations occurring in the system if we look at activities in the new environment (the square box) as a function of the neurons' preferred coordinates in the old environment (the rectangle). This is illustrated in Figure 9, which shows the activities of EC neurons, PP inputs to CA3, and final activities (after recurrent processing) in CA3 at three different locations in the square box, all as functions of preferred locations in the rectangular box. The activities of EC cells are determined by multiplying together (Gaussian-tuned) components whose activities depend on the animal's heading, and its position with respect to the walls. Since the walls have moved relative to each other, the different components lead to different estimates of position in the old coordinate system. Combining such inputs conjunctively leads to an EC activity profile which peaks somewhere between the positions indicated by individual walls. For instance, when the rat is halfway between the two walls that have been moved apart, listening to one of these walls would indicate that the animal is located at the opposite wall, and the resulting EC activity profile is centered on neurons which like the middle of the rectangular box (see the bottom left contour plot in Figure 9). Since the PP connections were established in the rectangular box, the PP input pattern to CA3 cells is centered around the same location as the EC activity pattern if both are viewed as a function of preferred coordinates in the rectangle (compare the first and the second columns of Figure 9). The recurrent connections then sharpen the activity profile considerably, but leave the location of the bump (in the old coordinate system) essentially unchanged. The final activities of CA3 cells as a function of location in the square box define the place fields in the new environment. We can see that as the rat moves around in the new environment, the activity packet also moves smoothly on the plane defined by the preferred locations of place cells in the rectangular box. This results in a smooth transformation of place fields between the two environments. In addition, the activity packet moves more slowly in the stretched direction in the

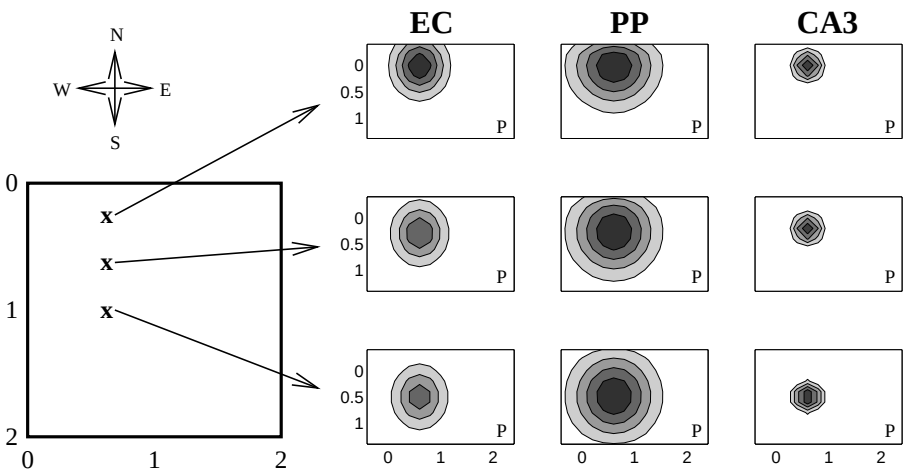

Figure 9: Place field stretching. The population plots of this figure show the neuronal activities in entorhinal cortex (EC), the perforant path (PP) inputs to place cells, and the CA3 final activities as a function of the preferred location of the neuron in the original, rectangular box, for three different positions of the rat in the square box, indicated by the crosses in the plot on the left. The plots only show cells with preferred direction 'north', and the model rat faces 'west' in all cases.

old coordinate system than the actual speed of the rat in the new environment, or, in other words, the rat needs to travel about twice as much in the square box than in the rectangular box for the activity profile to shift by the same amount; consequently, place fields become elongated in the direction of stretching.

The emergence of directional subcomponents can be understood by looking at how the activities of EC cells and the resulting activities of CA3 neurons depend on the head direction of the rat. This is depicted in Figure 10, which shows that due to the dependence on head direction of the rat's confidence in the inputs from different walls (as described earlier), conflicting sources of information are weighed differently depending on which way the rat faces. The EC activity profile and, consequently, the CA3 activity profile, shift as the rat turns around in the square box, and the result is that a given place cell fires maximally at different locations depending on head direction. In the example shown in Figure 10, the activity profile shifts 'north' when the rat faces 'north', and shifts 'south' when the rat faces 'south'. In an apparent paradox, from the

perspective of a single place field, this actually has the opposite effect (shown in Figure 8b), that the center of the place field is further 'south' when the animal faces north, and vice versa. The easiest way to see this is to ask where in the environment the rat has to be when it is facing in a particular direction, to arrange for exactly the population activity across CA3 shown in the middle picture of Figure 10. The answer to this will tell us how the favored location of the most active cell in this population depends on direction. When the rat faces 'north', the activity profile shifts 'north', so the rat must be displaced relatively 'south' in order to compensate for this. Thus, the location for the cell's peak response is shifted 'south'. The converse is also true – when the rat actually faces 'south', then the place field moves 'north'.

Very similar environments

Skaggs and McNaughton (1998) conducted an experiment designed specifically to probe the relation between spatial representations in environments with a high degree of similarity. In this experiment, animals explored an apparatus which consisted of two visually identical boxes connected by a corridor. Many place cells were found to have similar place fields in the two regions, whereas others had uncorrelated place fields. This finding challenges the idea that there is a predefined set of uncorrelated attractors wired into the recurrent connections in CA3 (Samsonovich and McNaughton, 1997), because such a model would predict either identical or orthogonal firing patterns in different environments or different parts of the same environment. This particular problem may be solved by postulating a hierarchy of fixed attractors with various degrees of overlap (Samsonovich et al., 1998); however, it still remains to be explained why similar representations are selected in very similar environments. On the other hand, the attractors established in our model are input dependent, which in principle allows attractors with an arbitrary degree of similarity, and directly defines the association between attractors and environments. Therefore, we

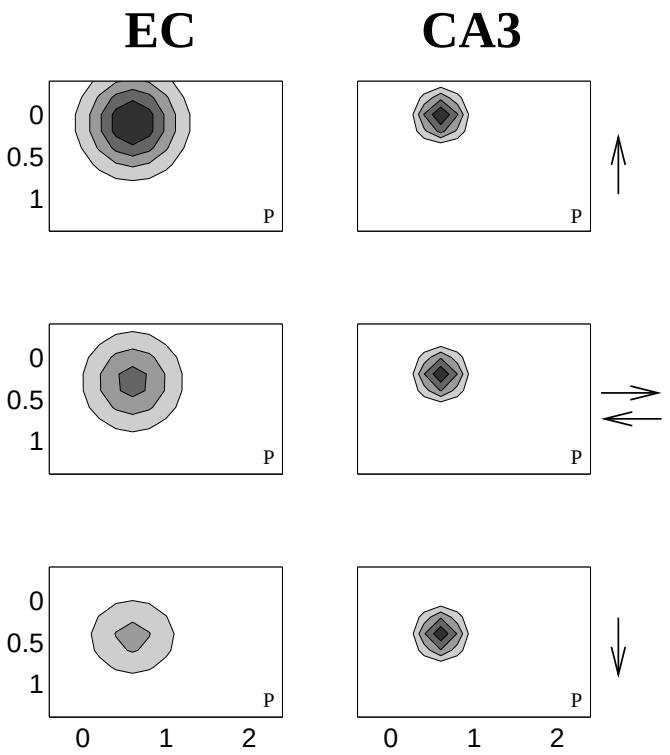

Figure 10: Directionality of stretched place fields. Population plots of EC neuronal activities and CA3 final activities (of the same sets of cells as in Figure 9), as a function of the neurons' preferred locations in the rectangular box, for different headings of the rat (indicated by the arrows) at a single location in the square box (marked by the middle cross in Figure 9). The middle row of plots is for both directions 'east' and 'west' as these lead to the same activities for the neurons displayed here. The position of the input peak changes as the rat faces in different directions (due to the dependence on head direction of the breadths of the input components tied to different walls), and the position of the final activity profile in CA3 changes accordingly. This shift can be compensated for by changes in location as seen in Figure 9, resulting in the directional subfields shown in Figure 8b.

simulated the experiment by Skaggs and McNaughton (1998) in our model to study the spatial representations in very similar environments.

We still do not model the different sources of spatial information explicitly. We assume that there are some inputs (e.g., signals derived from path integration) which allow the two boxes to be distinguished, while other inputs to the system (e.g., local visual cues) are identical at corresponding locations in the two boxes. Since cells in EC are assumed to respond to different inputs to a randomly

varying extent and to encode these inputs conjunctively, we applied the following scheme to determine activities in EC at locations inside the two boxes. EC cells are now characterized by a preferred location (and also a preferred head direction) based on visual inputs (this is now actually a set of two locations, one in each box), as well as a polarization index (P), which is defined as the maximum firing rate for the cell in the north box minus the maximum firing rate in the south box, divided by the maximum rate in any of the boxes. P is always between -1 and 1, its magnitude indicates how much that particular cell is influenced by cues that distinguish the two boxes, and its sign shows which box the neuron prefers. We assign P values to EC cells randomly from a uniform distribution. The firing rate of an EC neuron is then given by $z_k = (1 + P_k)z_k{}'$ in the north box and $z_k = (1 - P_k)z_k{}'$ in the south box, where $z_k{}'$ is a function of coordinates within the current box, and it depends on spatial position and head direction the same way as z_k in Equation 2. We assume that the MF inputs to CA3 can be characterized similarly; however, due to the orthogonalizing properties of the dentate gyrus, P values do not vary continuously, but only take the values -1, 0, and 1, each with probability 1/3. This means that there is a population of cells in CA3 which receives the same input at corresponding locations in the two boxes during learning, while another population receives different inputs. Since the first time the rat is introduced into the apparatus it is allowed to explore it entirely, we do not treat the two halves of the environment differently during the learning phase.

Some examples of the place fields that develop in this model are shown in Figure 11. There are cells which have similar firing rate patterns in the two boxes, while others are active in only one of the boxes, in accordance with experimental observations. In other words, our model has no difficulty storing and recalling partially overlapping spatial representations. In the model, the degree of overlap is determined by the extent of orthogonalization occurring in DG, i.e., what proportion of granule cells distinguishes between the two boxes – CA3 cells simply inherit the selectivity of their MF inputs as attractors are

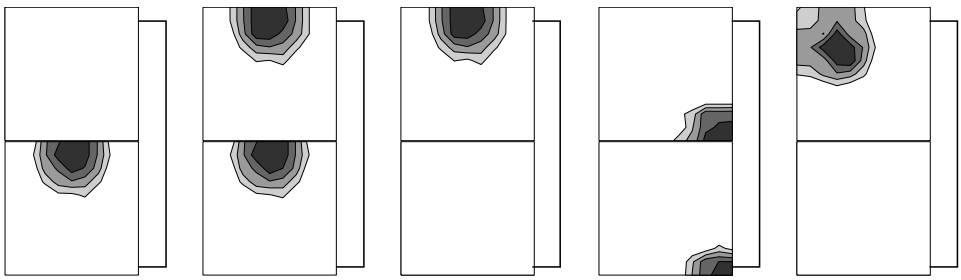

Figure 11: Place fields in our simulation of Skaggs and McNaughton (1998). The figure shows the place fields of five CA3 place cells in the two identical boxes; activity in the corridor connecting the boxes was not simulated.

established during the learning phase. Most EC neurons are active in both boxes, although to a different extent (see Figure 12a). Consequently, all CA3 cells that are active in this environment get a substantial PP input in both boxes (Figure 12b); however, the activity patterns encoded during learning are restored by the recurrent connections and feedback inhibition, and the PP input only determines which of these patterns emerges. The figures also show that although EC neurons have relatively broad tuning curves, and this results in CA3 cells receiving feedforward input that is even more broadly tuned, the final tuning of CA3 neurons is considerably sharper due to recurrent activity. The attractor network also renders place cells directionally nonselective, just as before.

Discussion

Principal findings

We have presented a plastic attractor network model of CA3 place cells which describes how a conjunctive representation of location- and direction-specific sensory information in entorhinal cortex can be transformed by feedforward pathways and recurrent processing in the hippocampus, into a place cell representation whose properties match a wide range of experimental observations. In

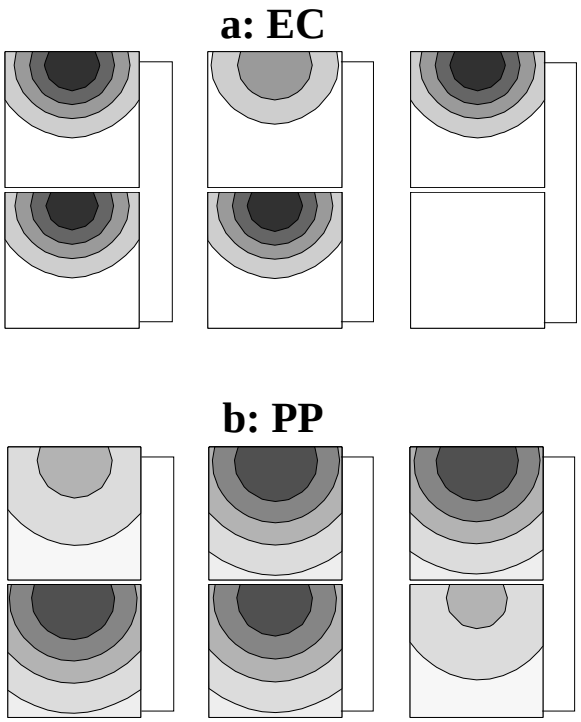

Figure 12: Input representation and inputs to CA3 in our simulation of Skaggs and McNaughton (1998). **a**: The activities, as a function of location in the apparatus, of 3 entorhinal neurons which have the same preferred (visual) location within the boxes, but different degrees of polarization (as defined in the text; the polarization indices are -0.01, 0.25, and -0.70, respectively). **b**: This part of the figure, which displays the perforant path inputs to the first three CA3 place cells of Figure 11, shows that, as a result of learning in the perforant pathway, some place cells receive similar inputs at corresponding locations in the two boxes, while others receive inputs of different magnitudes, setting the stage for the CA3 recurrent network which makes these differences much more pronounced (as seen in Figure 11).

particular, our model (1) accounts for the head direction independence of place cells in open environments as well as their directionality in linearly restricted environments, (2) demonstrates how several different environments can be stored and recalled independently by the CA3 recurrent network, (3) produces place cell activity patterns with an appropriate degree of overlap in visually similar environments, and (4) correctly captures the transformations of place fields after simple geometric manipulations of the environment.

Further, we have shown that the neural connections required for this spatial

representation can be computed from the correlations between the input-driven feedforward activations of neurons during initial exploration of the environment, using a familiarity-based on-line learning procedure.

Although the representations formed may be useful for spatial tasks such as navigation (Burgess et al., 1997; Foster et al., 1998, 2000), a major goal for our model was to show how ideas about how non-spatial information is processed by the hippocampus are in accordance with data on place fields.

Components of the model

The idea of using attractor networks for computations has been applied in various settings (Somers et al., 1995; Zhang, 1996; Pouget et al., 1998); such networks have been shown to be capable of amplifying certain facets of their inputs (Ben-Yishai et al, 1995) as well as creating invariance (Chance et al., 1999). Our model (and Brunel and Trullier's, 1998) displays both behaviors simultaneously; the recurrent network enhances the spatial tuning of place cells while suppressing their directional tuning in open field environments. Under attractor dynamics (which we assume characterizes well the average behavior of CA3 across theta oscillations), it is unwise to invent rules describing how individual place cells respond in various situations; rather, the system is better described collectively, by identifying the attractors and specifying which attractor gets selected for any particular input. The attractor concept also helps explain the persistence of spatial firing patterns in the face of environmental manipulations such as cue removal or cue rotation (O'Keefe and Conway, 1978; Muller and Kubie, 1987) as well as the abrupt changes that ensue for changes of other kinds (e.g., changing the shape of the environment from circular to square; Muller and Kubie, 1987) or of a larger magnitude. Feedforward models (*eg* Sharp, 1991; Burgess et al., 1997; Hartley et al., 2000), albeit ignoring the recurrent connections, can also be made to exhibit many of the properties we have demonstrated. We have not yet modeled the pathway from CA3 to CA1, assuming that the spatial properties of

the latter faithfully reflect those of the former, assuming normal plasticity. CA1 is, of course, the source of the bulk of the experimental data on place fields.

The learning rule was chosen as a crude model of experimental long-term synaptic plasticity (LTP), and we have ignored most empirical complexities. We have not taken into account the fact that the sign and magnitude of long-term synaptic modification depends on the relative timing of pre- and postsynaptic activity (Levy and Steward, 1983; Markram et al., 1997), which has been suggested as a mechanism underlying a navigational role of place cells (Blum and Abbott, 1996). Indeed, the recurrent weights in our model ultimately learn a weight structure similar to the "cognitive graph" described by Muller et al. (1991, 1996).

Similar proposals to ours have been put forward in associative memory models of the hippocampus (Treves and Rolls, 1992) as to the separate roles for the indirect pathway to CA3 via the dentate gyrus (which defines attractors during the learning mode) and the direct perforant path (which selects attractors during recall mode). However, the activity patterns representing location and direction are intrinsically continuous, and thus strongly overlapping, so the patterns that are retrieved can differ in systematic ways from all the patterns encountered during learning (*eg* being insensitive to head direction in open field environments). The relationship between attractor networks storing discrete vs. continuous sets of patterns, particularly regarding their storage capacity, has been studied by Samsonovich (1997) and Battaglia and Treves (1998).

Maintaining such overlapping attractors requires a learning rule which compensates for systematic spatial biases during exploration by gating learning through familiarity in a *graded* rather than a *binary* fashion, a subtlety not necessary for the very distinct attractors assumed by memory models. Gating of synaptic effectiveness (among the MF, PP, and recurrent collateral connections) and plasticity may either be mediated by closely coupled or by more distinct mechanisms (e.g., acetylcholine vs. GABA; Hasselmo, 1999; Sohal and Hasselmo,

1998), but there is very little evidence to distinguish between these possibilities at this point.

Exactly how entorhinal and dentate neurons encode features in the environment, and how they respond to manipulations of the environment, is not experimentally clear. Our choice was necessarily somewhat arbitrary – the aim has been to show that there exists at least one reasonable choice that results in place fields consistent with experimental data in a wide range of experimental situations. Our entorhinal representation is similar to that of Burgess et al. (1997), except that their units are directionally non-selective, and each is tied to exactly two orthogonal walls of the environment. In their model, place cell firing patterns are then determined through the feedforward weights connecting EC to CA3; these weights are set up using a competitive learning scheme similar to the one used by Sharp (1991) to model the formation of place fields. Competitive learning supports the separation of different input patterns in these models; in our model, the same task is thought to be accomplished through processing by the dentate gyrus. It is possible, of course, that both of these processes contribute to pattern separation in the hippocampus.

Comparison with other models

Apart from Brunel and Trullier (1998), the models of Samsonovich and McNaughton (1997) and Burgess et al. (1997) are closest to ours. The most important distinction from Samsonovich and McNaughton (1997) is that it relates the position of the activity profile (or 'packet') in CA3 (as an attractor network) to external coordinates in a different way, assuming a hard-wired system which is capable of updating the position of the CA3 activity packet based on self-motion information, and a learned association with sensory representations which can be used to correct for accumulated errors in path integration. Learning works differently in our model, and the metric of the place cell representation reflects the way in which the EC representation depends on external

coordinates, including sensory features of the environment and, to account for the formation and maintenance of place fields in darkness, self-motion information. The direct involvement of the hippocampus itself in path integration is controversial (Alyan and McNaughton, 1999; Maaswinkel et al., 1999).

Burgess et al.'s (1997) model also accounts for some of O'Keefe and Burgess's (1996) data. Their results are complementary to the ones we presented here, in that their model captures the behavior of those place cells which remain fixed with respect to one wall or develop a second place field after stretching the environment, while our model correctly describes those place fields that follow the transformation of the environment, and also explains the acquired directionality of stretched place fields in the transformed environment. A modified version of our model, which incorporates random variations in the extent to which input cells respond to different spatial cues, reproduces all the observed classes of place field transformation. Due to its randomness, it offers less insight into the underlying mechanisms than the model described here. Our model also accounts for other properties of place cells, such as directionality and non-directionality.

Critical experiments

Various experiments could, in principle, test the key assumptions and predictions of our model. First, pharmacological or molecular biological blockade of plasticity in the CA3 recurrent connections should prevent the formation of a new representation in a novel situation. According to our model, the system would either remain trapped in learning mode, which would be indicated by, among other things, retained directionality of place fields in an open field, or recall attractors from one or more environments explored before the blockade, resulting in irregular or fragmented place fields. Direct manipulations of the neuromodulatory control mechanisms governing the choice of learning versus recall mode should have a similar effect. Unfortunately, there exist many different forms of experimental plasticity, and it is not clear which in particular are

most relevant for learning in vivo.

Our model predicts that the CA3 place cell representation should be different during the first few minutes of exploration in a new environment from the time after the animal has become familiar with its surroundings. In particular, place cells are expected to be directional in any novel environment immediately after entry, and become non-directional later in open environments.

Analysis of our model also indicates that the amount of training in a given environment might have a significant effect on the place cell representation in a similar environment encountered subsequently, since only well-established attractors are assumed to be capable of being recalled. For instance, we would expect to see a less obvious relation between place fields in different environments in the experiment of O'Keefe and Burgess (1996) if, instead of training the rat in one size of box before allowing it to explore the others, they had made it explore all four environments in quick succession, especially if the rat is prevented from using extramaze cues.

References

[1] Alyan S, McNaughton BL (1999) Hippocampectomized rats are capable of homing by path integration. Behav Neurosci 113:19-31.

[2] Amaral DG, Witter MP (1989) The three-dimensional organization of the hippocampal formation: a review of anatomical data. Neuroscience 31:571-591.

[3] Amaral DG, Ishizuka N, Claiborne B (1990) Neurons, numbers and the hippocampal network. Prog Brain Res 83:1-11.

[4] Ault B, Nadler JV (1982) Baclofen selectively inhibits transmission at synapses made by axons of CA3 pyramidal cells in the hippocampal slice. J Pharmacol Exp Ther 223:291-297.

[5] Barnes CA, McNaughton BL, Mizumori SJY, Leonard BW, Lin L-H (1990) Comparison of spatial and temporal characteristics of neuronal activity in sequential stages of hippocampal processing. Prog Brain Res 83:287-300.

[6] Battaglia FP, Treves A (1998) Attractor neural networks storing multiple space representations: A model for hippocampal place fields. Phys Rev E 58:7738-7753.

[7] Ben-Yishai R, Bar-Or RL, Sompolinsky H (1995) Theory of orientation tuning in visual cortex. Proc Natl Acad Sci USA 92:3844-3848.

[8] Benardo LS, Prince DA (1982) Ionic mechanisms of cholinergic excitation in mammalian hippocampal pyramidal cells. Brain Res 249:333-344.

[9] Blum KI, Abbott LF (1996) A model of spatial map formation in the hippocampus of the rat. Neural Comput 8:85-93.

[10] Bostock E, Muller RU, Kubie JL (1991) Experience-dependent modifications of hippocampal place cell firing. Hippocampus 1:193-205.

[11] Breindl A, Derrick BE, Rodriguez SB, Martinez JL Jr (1994) Opioid receptor-dependent long-term potentiation at the lateral perforant path–CA3 synapse in rat hippocampus. Brain Res Bull 33:17-24.

[12] Brunel N, Trullier O (1998) Plasticity of directional place fields in a model of rodent CA3. Hippocampus 8:651-665.

[13] Burgard EC, Sarvey JM (1990) Muscarinic receptor activation facilitates the induction of long-term potentiation (LTP) in the rat dentate gyrus. Neurosci Lett 116:34-39.

[14] Burgess N, Donnett JG, O'Keefe J (1997) Robotic and neuronal simulation of hippocampal navigation. Phil Trans Roy Soc B 352:1535-1543.

[15] Burwell RD, Amaral DG (1998) Cortical afferents of the perirhinal, postrhinal, and entorhinal cortices of the rat. J Comp Neurol 398:179-205.

[16] Chance FS, Nelson SB, Abbott LF (1999) Complex cells as cortically amplified simple cells. Nat Neurosci 2:277-282.

[17] Debanne D, Gähwiler BH, Thompson SM (1998) Long-term synaptic plasticity between pairs of individual CA3 pyramidal cells in rat hippocampal slice cultures. J Physiol (Lond) 507:237-247.

[18] Foster DJ, Morris RGM, Dayan P (1998) Hippocampal model of rat spatial abilities using temporal difference learning. In Jordan, M.I., Kearns, M.J., Solla, S.A., eds, Advances in Neural Information Processing Systems 10:145-151.

[19] Foster DJ, Morris RGM, Dayan P (2000) A model of hippocampally dependent navigation, using the temporal difference learning rule. Hippocampus 10:1-16.

[20] Grossberg S (1988) Nonlinear neural networks: principles, mechanisms, and architectures. Neural Networks 1:17-61.

[21] Hartley T, Burgess N, Lever C, Cacucci F, O'Keefe J (2000) Modeling place fields in terms of the cortical inputs to the hippocampus. Hippocampus, in press.

[22] Hasselmo ME (1999) Neuromodulation: acetylcholine and memory consolidation. Trends Cognit Sci 3:351-359.

[23] Hasselmo ME, Schnell E, Barkai E (1995) Learning and recall at excitatory recurrent synapses and cholinergic modulation in hippocampal region CA3. J Neurosci 15:5249-5262.

[24] Hasselmo ME, Wyble BP, Wallenstein GV (1996) Encoding and retrieval of episodic memories: Role of cholinergic and GABAergic modulation in the hippocampus. Hippocampus 6:693-708.

[25] Huerta PT, Lisman JE (1993) Heightened synaptic plasticity of hippocampal CA1 neurons during a cholinergically induced rhythmic state. Nature 364:723-725.

[26] Ishizuka N, Weber J, Amaral DG (1990) Organization of intrahippocampal projections originating from CA3 pyramidal cells in the rat. J Comp Neurol 295:580-623.

[27] Jung MW, McNaughton BL (1993) Spatial selectivity of unit activity in the hippocampal granular layer. Hippocampus 3:165-182.

[28] Káli S, Dayan P (1998) The formation of direction independent place fields in area CA3 of the rodent hippocampus using Hebbian plasticity in a recurrent network. Soc Neurosci Abstr 24:931.

[29] Levy WB (1996) A sequence predicting CA3 is a flexible associator that learns and uses context to solve hippocampal-like tasks. Hippocampus 6:579-590.

[30] Levy WB, Steward O (1983) Temporal contiguity requirements for long-term associative potentiation/depression in the hippocampus. Neuroscience 8:791-797.

[31] Li Z, Dayan P (1999) Computational differences between asymmetrical and symmetrical networks. Network 10:59-77.

[32] Li X-G, Somogyi P, Ylinen A, Buzsáki G (1994) The hippocampal CA3 network: An in vivo intracellular labeling study. J Comp Neurol 339:181-208.

[33] Maaswinkel H, Jarrard LE, Whishaw IQ (1999) Hippocampectomized rats are impaired in homing by path integration. Hippocampus 9:553-561.

[34] Markram H, Lubke J, Frotscher M, Sakmann B (1997) Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. Science 275:213-215.

[35] Markus EJ, Qin Y-L, Leonard B, Skaggs WE, McNaughton BL, Barnes CA (1995) Interactions between location and task affect the spatial and directional firing of hippocampal neurons. J Neurosci 15:7079-7094.

[36] Marr D (1971) Simple memory: a theory for archicortex. Philos Trans R Soc Lond B 262:24-81.

[37] McNaughton BL, Morris RGM (1987) Hippocampal synaptic enhancement and information storage within a distributed memory system. Trends Neurosci 10:408-415.

[38] Muller RU, Bostock E, Taube JS, Kubie JL (1994) On the directional firing properties of hippocampal place cells. J Neurosci 14:7235-7251.

[39] Muller RU, Kubie JL (1987) The effects of changes in the environment on the spatial firing of hippocampal complex-spike cells. J Neurosci 7:1951-1968.

[40] Muller RU, Kubie JL, Ranck JB Jr (1987) Spatial firing patterns of hippocampal complex-spike cells in fixed environment. J Neurosci 7:1935-1950.

[41] Muller RU, Kubie JL, Saypoff R (1991) The hippocampus as a cognitive graph. Hippocampus 1:243-246.

[42] Muller RU, Stead M, Pach J (1996) The hippocampus as a cognitive graph. J Gen Physiol 107:663-694.

[43] Myers CE, Gluck MA, Granger R (1995) Dissociation of hippocampal and entorhinal function in associative learning: A computational approach. Psychobiology 23:116-138.

[44] O'Keefe J (1976) Place units in the hippocampus of the freely moving rat. Exp Neurol 51:78-109.

[45] O'Keefe J, Burgess N (1996) Geometric determinants of the place fields of hippocampal neurons. Nature 381:425-428.

[46] O'Keefe J, Conway DH (1978) Hippocampal place units in the freely moving rat: why they fire where they fire. Exp Brain Res 31:573-590.

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

[48] O'Reilly RC, McClelland JL (1994) Hippocampal conjunctive encoding, storage, and recall: avoiding a tradeoff. Hippocampus 4:661-682.

[49] Pitler TA, Alger BE (1992) Cholinergic excitation of GABAergic interneurons in the rat hippocampal slice. J Physiol (Lond) 450:127-142.

[50] Pouget A, Zhang K, Deneve S, Latham PE (1998) Statistically efficient estimation using population coding. Neural Comput 10:373-401.

[51] Quirk GJ, Muller RU, Kubie JL, Ranck JB Jr (1992) The positional firing properties of medial entorhinal neurons: description and comparison with hippocampal place cells. J Neurosci 12:1945-1963.

[52] Rolls ET (1996) A Theory of Hippocampal Function in Memory. Hippocampus 6:601-620.

[53] Samsonovich A (1997) Attractor map theory of the hippocampal representation of space. PhD thesis, University of Arizona.

[54] Samsonovich A, McNaughton BL (1997) Path integration and cognitive mapping in a continuous attractor neural network model. J Neurosci 17:5900-5920.

[55] Samsonovich A, McNaughton BL, Nadel L (1998) Hierarchical multichart model of the hippocampal cognitive map. Soc Neurosci Abstr 24:931.

[56] Sharp EP (1991) Computer simulation of hippocampal place cells. Psychobiology 19:103-115.

[57] Shen B, McNaughton BL (1996) Modeling the spontaneous reactivation of experience-specific hippocampal cell assembles during sleep. Hippocampus 6:685-692.

[58] Skaggs WE, McNaughton BL (1998) Spatial firing properties of hippocampal CA1 populations in an environment containing two visually identical regions. J Neurosci 18:8455-8466.

[59] Sohal VS, Hasselmo ME (1998) GABA(B) modulation improves sequence disambiguation in computational models of hippocampal region CA3. Hippocampus 8:171-193.

[60] Somers DC, Nelson SB, Sur M (1995) An emergent model of orientation selectivity in cat visual cortical simple cells. J Neurosci 15:5448-5465.

[61] Touretzky DS, Redish AD (1996) A theory of rodent navigation based on interacting representations of space. Hippocampus 6:247-270.

[62] Treves A, Rolls ET (1992) Computational constraints suggest the need for two distinct input systems to the hippocampal CA3 network. Hippocampus 2:189-199.

[63] Treves A, Rolls ET (1994) A computational analysis of the role of the hippocampus in memory. Hippocampus 4:374-391.

[64] Wilson HR, Cowan JD (1972) Excitatory and inhibitory interactions in localized populations of model neurons. Biophys J 12:1-24.

[65] Wilson MA, McNaughton BL (1993) Dynamics of the hippocampal ensemble code for space. Science 261:1055-1058.

[66] Yamamoto C (1982) Quantal analysis of excitatory postsynaptic potentials induced in hippocampal neurons by activation of granule cells. Exp Brain Res 76:170-176.

[67] Zalutsky RA, Nicoll RA (1990) Comparison of two forms of long-term potentiation in single hippocampal neurons. Science 248:1619-1624.

[68] Zhang K (1996) Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: a theory. J Neurosci 16:2112-2126.

[69] Zipser D (1985) A computational model of hippocampal place fields. Behav Neurosci 99:1006-1018.