Network: Computation in Neural Systems
December 2006; 447: 447–465

informa
healthcare

Entorhinal cortex grid cells can map to hippocampal place cells by competitive learning

EDMUND T. ROLLS, SIMON M. STRINGER & THOMAS ELLIOT

Oxford University, Centre for Computational Neuroscience, Department of Experimental Psychology, South Parks Road, Oxford OX1 3UD, England

(Received 28 July 2006; accepted 14 October 2006)

Abstract
'Grid cells' in the dorsocaudal medial entorhinal cortex (dMEC) are activated when a rat is located at any of the vertices of a grid of equilateral triangles covering the environment. dMEC grid cells have different frequencies and phase offsets. However, cells in the dentate gyrus (DG) and hippocampal area CA3 of the rodent typically display place fields, where individual cells are active over only a single portion of the space. In a model of the hippocampus, we have shown that the connectivity from the entorhinal cortex to the dentate granule cells could allow the dentate granule cells to operate as a competitive network to recode their inputs to produce sparse orthogonal representations, and this includes spatial pattern separation. In this paper we show that the same computational hypothesis can account for the mapping of EC grid cells to dentate place cells. We show that the learning in the competitive network is an important part of the way in which the mapping can be achieved. We further show that incorporation of a short term memory trace into the associative learning can help to produce the relatively broad place fields found in the hippocampus.

Keywords: *Entorhinal cortex, grid cell, place cell, competitive network, hippocampus, dentate gyrus*

1. Introduction

Neurons have recently been found in a dorsocaudal part of the medial entorhinal cortex (dMEC) that fire in several locations in an environment, with the locations forming a regular pattern as though they were nodes on a triangular grid (Hafting, Fyhn, Molden, Moser & Moser 2005). Different 'grid' cells recorded at the same locations in the medial entorhinal cortex have the same grid spacing and orientation relative to the environment. They differ however in the location of the nodes such that the firing peaks of one neuron are slightly shifted from those of its neighbor, that is there is a spatial phase shift between neurons. The multiple fields of several such cells together cover the environment. In addition, different grid cells have different grid frequencies, and the grid frequencies decrease from dorsal to ventral dMEC. For each cell, the size of the grid appears to be independent of the size or shape of the environment. The orientation of the grid relative to the environment, however, is dependent on the location of a polarizing visual cue on the wall of the enclosure in much the same way as the post-subicular head direction (Taube, Muller & Ranck 1990) and the hippocampal place cells (Muller & Kubie 1987). However, cells in the rat dentate gyrus (DG)

Correspondence: Edmund T. Rolls, Tel: +44 1865 271348, Fax: +44 1865 310447. E-mail: Edmund.Rolls@psy.ox.ac.uk, web: www.cns.ox.ac.uk

ISSN: 0954-898X print / ISSN 1361-6536 online
DOI: 10.1080/09548980601064846

and hippocampal CA3 region typically have place fields, where individual cells are active over only a single continuous portion of the space (O'Keefe & Dostrovsky 1971; O'Keefe 1984; McNaughton, Barnes & O'Keefe 1983; Muller, Kubie, Bostock, Taube & Quirk 1991; Jung & McNaughton 1993; Jeffery, Anderson, Hayman & Chakraborty 2004; Jeffery & Hayman 2004). An important question, then, is how the spatially repeating responses of dMEC grid cells become mapped to the place fields of cells in DG and CA3.

It has been suggested that some type of Hebbian synaptic modification could allow all the grid cells with peak firing at a given location to be connected to place cells whose place fields are centered at that location (O'Keefe & Burgess 2005). However, we are interested in how the DG place cells could be generated by a self-organizing process from grid cells. It has been proposed that the entorhinal to dentate system operates as a competitive network to produce sparse spatial representations (Rolls 1987, 1989a, 1990, 1995, 1996; Rolls & Treves 1998; De Araujo, Rolls & Stringer 2001), which are known to be present in DG as shown by the finding that DG neurons typically have small place fields (Jung & McNaughton 1993). (Competitive network learning is described by Hertz, Krogh and Palmer (1991), Rolls and Treves (1998), and Rolls and Deco (2002). Competitive learning of entorhinal cortex to hippocampal cells has also been proposed by Sharp (1991). McNaughton & Morris (1987) and McNaughton & Nadel (1990) have discussed the formation of sparse representations in dentate granule cells by a Marr codon-like process.) Competitive networks operate using associatively modifiable forward connections (in this case from the preceding dMEC layer to the DG neurons), and recurrent inhibitory feedback mediated by inhibitory interneurons that implements competition between in this case the DG neurons (Hertz et al. 1991; Rolls & Treves 1998; Rolls & Deco 2002) (see Figure 1, and the description under Methods). Previous papers have demonstrated how competitive learning mechanisms can result in place cell responses within the hippocampus based on Hebbian modification of inputs from the

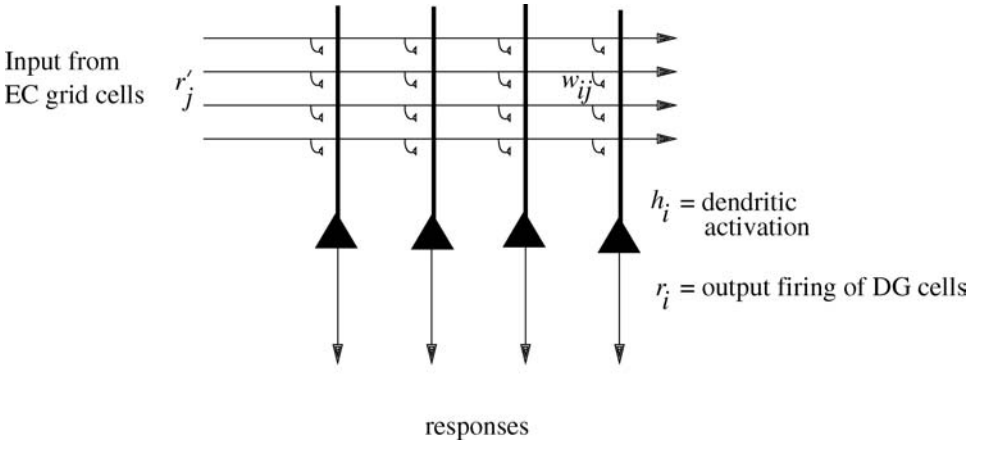

Figure 1. Model neural network architecture. The 2-layer EC/DG model has the architecture of a competitive network. There is an input layer of dMEC grid cells with feedforward associatively modifiable synaptic connections onto an output layer of DG cells. As the agent moves through its environment during training, the dMEC grid cells are set to fire according to their prescribed oscillatory activity profiles. At each location, activity from the EC layer is propagated through the feedforward connections to activate a winning set of cells in the DG layer. Within the layer of DG cells, there is feedback inhibition, which ensures that only a small subset of DG cells remain active. Next, the synaptic weights between the active dMEC grid cells and the active DG cells are strengthened. In this way, DG cells become associated with patterns of activity in the EC layer, which represent particular locations in the environment.

entorhinal cortex (e.g. Sharp (1991)), although the representation in the entorhinal cortex was not of grid cells, which have not been discovered until recently. Shapiro and Hetherington (1993) simulated the formation of place cell responses by error backpropagation using a generalized delta rule in a multilayer network.

In this paper we show how DG place cells could be built from entorhinal cortex grid cells using a competitive learning mechanism to implement self-organization in DG of a place cell representation. We further show that a useful component of the self-organizing process is a short term memory trace that helps the associative learning to build place fields that have the relatively broad place fields of the DG and hippocampus. The short term memory trace could be built into the postsynaptic activation of the DG neurons by processes such as the relatively long time constant of NMDA receptors. Although previous papers have been interested in modelling the entorhinal inputs to the hippocampus (Marr 1971; Sharp 1991; McNaughton & Morris 1987; McNaughton & Nadel 1990; Redish &Touretzky 1998; Redish 1999; Shapiro & Hetherington 1993), this is the first paper we know to demonstrate that the grid cells found in the entorhinal cortex (Hafting et al. 2005) could map to hippocampal place cells by competitive learning.

2. Methods

2.1 The neural network model

The neural network architecture is shown in Figure 1. The fully connected 2-layer EC/DG model has the architecture of a competitive network. There is an input layer of dMEC grid cells with feedforward associatively modifiable synaptic connections onto an output layer of DG cells.

First we simulate the functionality in a model of 1D space, in order to investigate the principles of how the mapping from a grid-like representation to a place-like representation can be performed in a one-layer competitive network. Then we extend the simulations to a 2D model of space, to demonstrate that grid neurons in the entorhinal cortex which map a 2D space could map by the same principles to the dentate/CA3 representation of 2D place cells.

Our 1D model simulations contain 100 EC grid cells. The grid cells have ten different frequencies, as shown in Figure 2, with 3–12 cycles over the spatial environment. These frequencies reflect the range of grid spacings observed in the neurophysiological studies of

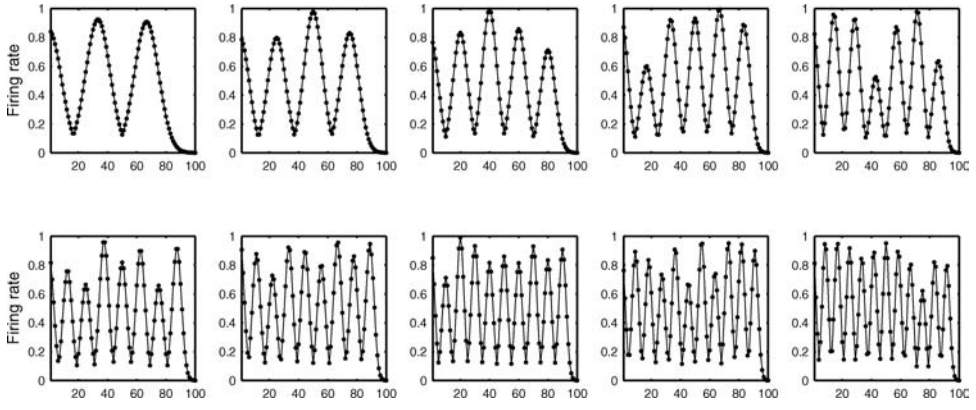

Figure 2. Firing rate profiles of 10 EC grid cells with frequencies 3–12. The standard deviation of the peak heights was set to 0.2

Hafting et al. (2005). For each frequency, there are 10 separate EC grid cells with different phase offsets. There are also 100 DG cells in the output layer. The firing rate of each grid cell is modelled as a series of Gaussians centered with different heights (as specified below to model the properties of the grid cells as recorded) at successive nodal locations. Each Gaussian firing field covered at least several training locations, so that the inherent continuity of the space was adequately represented. Each Gaussian peak of a grid cell is specified as a firing rate r^{EC}

$$r^{\mathrm{EC}} = e^{-s^2/2\sigma^2}, \tag{1}$$

where s is the difference between the actual location x of the animal and the preferred location x for the peak of the grid cell, and σ is the standard deviation specifying the width of the peak.

The space being studied is inherently continuous, because the response fields of nearby neurons overlap in space, and are defined continuously in space. We note that if a continuous space is being modelled on a digital computer, the space must be sampled at a large number of sufficiently close locations in the space, and we ensured that this was the case by using in the 1D simulations 100 locations in which every spatial representation covered a number of the locations.

During training, the agent is moved through a sequence of 100 consecutive locations through the environment. As the agent moves to each location, the dMEC grid cells are set to fire according to their prescribed activity profiles. At each location, activity from the EC layer is propagated through the feedforward synaptic connections to activate a set of cells in the DG layer. (The synaptic weights from the EC to DG cells are initially set to random values. This is standard in competitive networks, and ensures that some firing of output cells will be produced by the inputs, with each output cell likely to fire at a different rate for any one input stimulus.) The activations of the cells in the DG layer are calculated according to

$$h_i^{\mathrm{DG}} = \sum_j w_{ij} r_j^{\mathrm{EC}} \tag{2}$$

where h_i^{DG} is the activation of DG cell i, r_j^{EC} is the firing rate of EC cell j, and w_{ij} is the synaptic weight from EC cell j to DG cell i. Next, mutual inhibition between the DG cells implements competition to ensure that there is only a small winning set of DG cells left active. The mutual inhibition (which would be implemented by inhibitory feedback neurons which implement lateral inhibition in the brain) ensures that the sparseness of the activity in the DG layer is kept to a fixed value, set in the one dimensional simulations to 0.04 and in the two dimensional simulations to 0.02. (In the simulations the competition was achieved by first setting the firing rates to the square of the activations—corresponding to the steeply rising part of a sigmoid activation function—and second by feedback adjustment of the threshold of the firing until the desired sparseness was achieved.[1] The algorithm was thus set to adjust the threshold for the firing of neurons until a sparseness of 0.04 for the one-dimensional and 0.02 for the two-dimensional simulations was reached. With the graded firing rates of the neurons, the result was that typically the proportion of neurons with non-zero firing rates

[1] The sparseness a of the representation can be measured, by extending the binary notion of the proportion of neurons that are firing, as

$$a = \frac{\left(\sum_{i=1}^{N} y_i/N\right)^2}{\sum_{i=1}^{N} y_i^2/N} \tag{3}$$

where y_i is the firing rate of the ith neuron in the set of N neurons.

after the competition was numerically somewhat larger than the sparseness value given as a parameter to the network.

Next, the synaptic weights between the active dMEC grid cells and the active DG cells are strengthened according to the associative Hebb learning rule

$$\delta w_{ij} = k r_i^{\mathrm{DG}} r_j^{\mathrm{EC}} \tag{4}$$

where δw_{ij} is the change of synaptic weight and k is the learning rate constant. k was set to be sufficiently low that weight vectors learned early in training were not overwritten by the end of training. To prevent the same few neurons always winning the competition, the synaptic weight vectors are set to unit length after the learning at each location.[2] Such a renormalization process may be achieved in biological systems through heterosynaptic long-term depression that depends on the existing value of the synaptic weight (Oja 1982; Rolls & Treves 1998; Rolls & Deco 2002). (Heterosynaptic long-term depression was described in the brain by Levy and colleagues (Levy 1985; Levy & Desmond 1985); see Brown et al. (1990).

We also investigated the use of a modified associative learning rule that incorporates a short term memory trace of previous firing in (for example) the postsynaptic term. We hypothesized that use of such a rule might enable close spatial locations (which would be visited at closely related times) to become associated together in the learning process, and thus to potentially help create broader place fields in DG. The trace learning rule is as follows (see further Rolls & Deco (2002) and Stringer, Trappenberg, Rolls & De Arajuo (2002)):

$$\delta w_{ij} = k \overline{r}_i r_j \tag{6}$$

where $\overline{r}$ is a local temporal average or trace value of the firing rate of a cell given by

$$\overline{r}(t + \delta t) = (1 - \eta) r(t + \delta t) + \eta \overline{r}(t) \tag{7}$$

where η is a parameter set in the interval [0,1] which determines the contribution of the current firing and the previous trace. For $\eta = 0$ the trace rule (6) becomes the standard Hebb rule (4), while for $\eta > 0$ learning rule (6) operates to associate together patterns of activity that occur close together in time. (If we think of the simulations as modelling a rat running at a speed of 10 cm/s over a distance of 1 m, which is divided into 100 finely spaced training locations, then with the value of η of 0.8 used in these simulations unless otherwise specified, the time constant of the exponential decay of the trace is 0.5 s.)

The movement of the agent through all 100 locations spanning the environment constitutes one epoch of training. For each simulation, there were 200 training epochs. During training, DG cells learn to respond to patterns of activity in the EC layer, which represent particular locations in the environment.

2.2 Competitive network learning with discrete and continuous representations

Competitive networks are often envisaged as operating to categorise discrete sets of patterns. Each category consists of a discrete set of exemplar patterns that are correlated with each other more than they are correlated with the exemplar patterns in other categories. As the

[2] To implement weight normalization the synaptic weights were rescaled to ensure that for each DG cell i we have

$$\sqrt{\sum_j (w_{ij})^2} = 1, \tag{5}$$

where the sum is over all EC cells j.

competitive net self-organizes, the weight vector of each output neuron moves towards the centre of a set of patterns in the input space (Rumelhart & Zipser 1986; Grossberg 1987; Hertz et al. 1991; Rolls & Treves 1998; Rolls & Deco 2002).

However, space is continuous, and neurons representing different locations in the space change their activity continuously as different parts of the space are visited. The spatial fields of nearby neurons overlap in the space. (This is exemplified in the firing of both place cells and grid cells.) In this situation competitive learning distributes the output neurons evenly throughout the space (Hertz et al. 1991), and this effect was present in the model Sharp (1991) described of the formation of hippocampal place cells from entorhinal cortex inputs. However, that model did not have a grid cell representation in the entorhinal cortex, as this was yet to be discovered. We consider grid cells as the input representation in this paper.

3. Results

The simulations of the 1D model contain 100 EC grid cells, with 10 different frequencies corresponding to 3–12 cycles over the spatial environment, and with 10 different phases (or spatial offsets) for each frequency. Figure 2 shows the firing rate profiles of an example 10 of these EC grid cells with different frequencies. The standard deviation of the peak heights was set to 0.2.

We first show, in Figure 3, firing rate profiles of 25 DG cells with no learning, that is, with only the initial random connection weights from EC to DG. The DG cells are generally spatially repeating. There are few DG cells with single-peaked place fields. The 25 DG cells were chosen at random to illustrate the spatial profiles of the 100 DG cells in the simulation. Thus without training, the network does not produce neurons with single place fields in DG.

We next show the effect of training the network with an associative (Hebb) learning rule. The training consisted of the agent moving continuously through the 100 equispaced training locations, and repeating this for 200 such training epochs. (In the simulations described in this paper, we checked that the results were stable for larger numbers of training epochs, as this is an indication of convergence. The training was normally from left to right in each epoch, and we checked that consistent results were obtained when training was from left to right and then right to left in each epoch.) Figure 4 shows firing rate profiles of 25 DG cells after learning with the associative rule (see Eqn. 4), with a learning rate of 0.01. The effect of the associative learning in what is a conventional competitive network but trained with the grid-like patterns of EC cells is to produce some DG cells that have typically a single place field, rather than a spatially repeating field which occurred without training. Thus simple competitive network learning can produce place fields in DG from EC grid cells. Further, this can occur even if there is no variation in the peaks of the firing rates of the EC cells at different grid locations, that is, if the standard deviation of the peaks heights is zero. (The proportion of DG neurons that develop place fields depends on the sparseness of the DG firing, which was set to $a = 0.04$.)

To quantify statistically the effect of the training, we compared the properties of the DG cells after training with the properties without training, by performing eight separate simulations with different random seeds for both the untrained and for the trained conditions. The mean number of DG place cells (with 100 DG cells) with training was 10.25 ± 0.84 (sem), and without training was 0.75 ± 0.25 ($t = 10.84$, $df = 14$, $p < 0.0001$). (A place cell was defined for this statistical analysis as having a single firing rate peak >0.8 and no

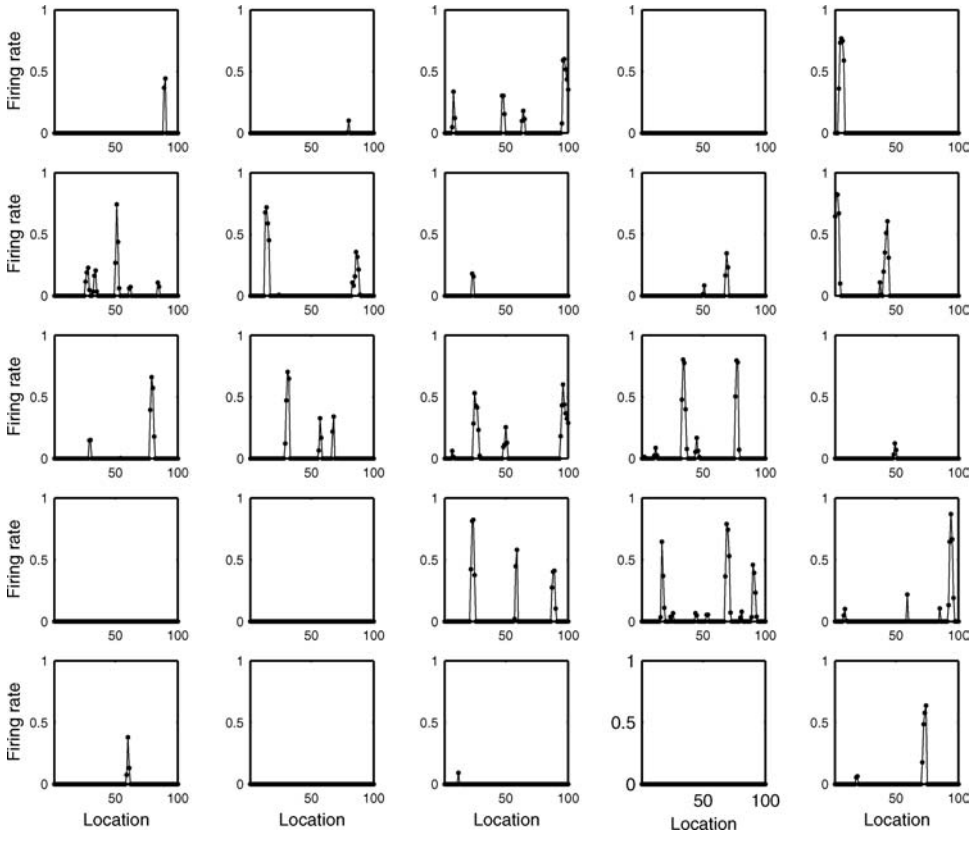

Figure 3. Firing rate profiles of 25 DG cells with no learning. The DG cells are generally spatially repeating, There are few DG cells with single-peaked place fields.

other peak >0.1. In the trained condition, almost all the remaining cells were unresponsive. In the untrained condition, many of the cells had multiple peaks, as illustrated in Figure 3). This analysis shows statistically that training does produce cells with place-like properties (i.e. with one major peak centred on a particular location). (It should be noted that we would not expect all the DG cells to be affected by the learning process, because of the sparseness of the representation used in DG, and because of the finite extent and amount of variation in the input space.)

The place fields that do develop as described above are relatively small (i.e. narrow), covering only a small fraction of the environment. In the following simulations, we explore training regimes that can produce the wider fields (relative to the size of the testing environment) that are more typical of place cells (in DG and CA3) (Jung & McNaughton 1993; O'Keefe & Dostrovsky 1971; O'Keefe 1984; McNaughton et al. 1983; Muller et al. 1991; Jeffery et al. 2004; Jeffery & Hayman 2004). Figure 5 shows firing rate profiles of 25 DG cells after learning with the trace rule, and indicates that use of the trace rule can enable DG cells to have broader place fields. It was found that to facilitate the development of non-repeating DG place fields when the trace was being used, simulations were needed with more variation in the firing rates of the grid cells for the different positions at which they responded. The aim was to enable the DG cells to categorise the space more discretely by providing more variation in the input space, but at the same time, to use the trace rule to help the DG

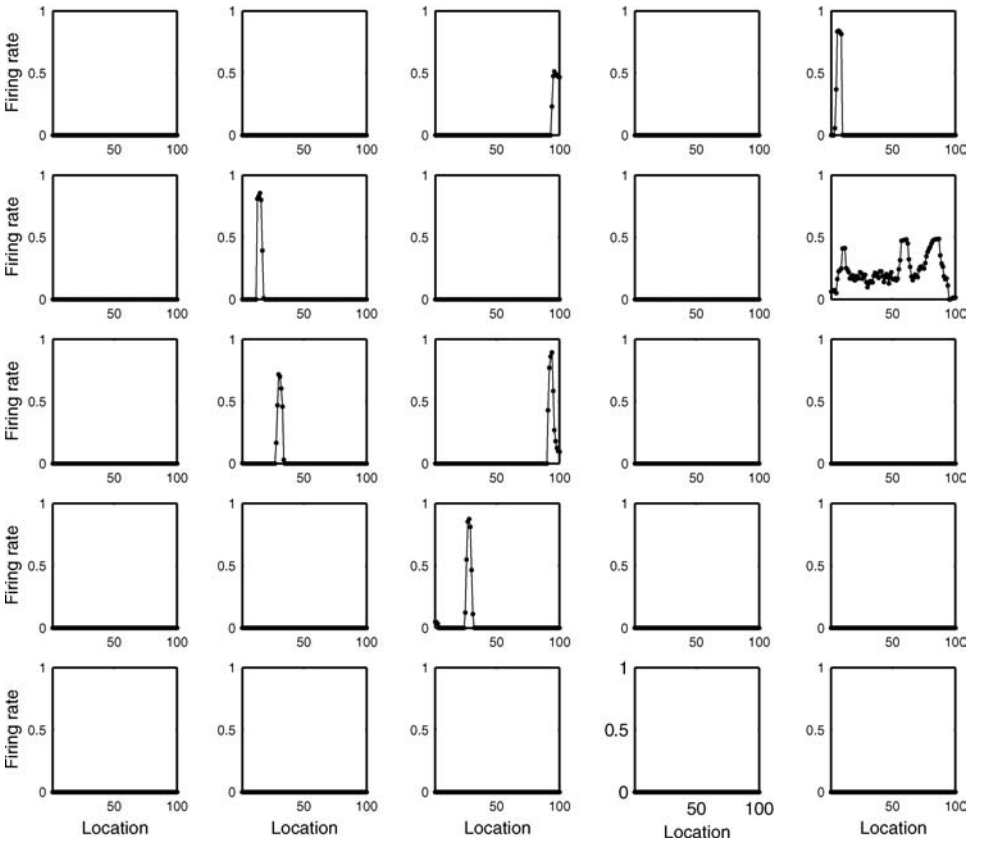

Figure 4. Firing rate profiles of 25 DG cells after learning with the Hebb rule, with learning rate 0.01. There is less spatial repetition among the DG cells, and some DG cells have developed place fields.

place fields to become wide. Figure 6 shows the firing rate profiles of 10 EC grid cells with frequencies 3–12, and with the standard deviation of the peak heights set to the larger value of 0.6 used for the results shown in Figure 5.

Other simulations (not illustrated for space reasons) indicated that the trace rule did not work as well if the variation in the EC grid cell firing was small. This was shown in simulations in which the trace rule was used but the standard deviation of the peak heights was left at 0.2. For example a sd of 0.2 tended to produce repeating fields with the trace rule. This was because there was insufficient variation in the inputs to produce peaks sufficient to fractionate the space. Under these conditions, the low frequency EC cells dominated the repeating spatial fields in DG, with the trace effectively smoothing out the effects of the higher spatial frequencies in the EC representation, thus leaving the low spatial frequency EC cells dominating the firing of DG cells, and thus leading to repeating place fields in DG if there was little variation in the peaks of EC cells.

Training with the temporal trace learning rule can, as shown in Figure 5, help broader place cells to form. The trace rule effectively enables positions that are visited close together in time, and are thus spatially close, to be linked by temporal associative learning. (Of course, the trace would have the effect of associating more distant spatial locations if the running speed was high.) The temporal trace can become large for a location where the inputs produce

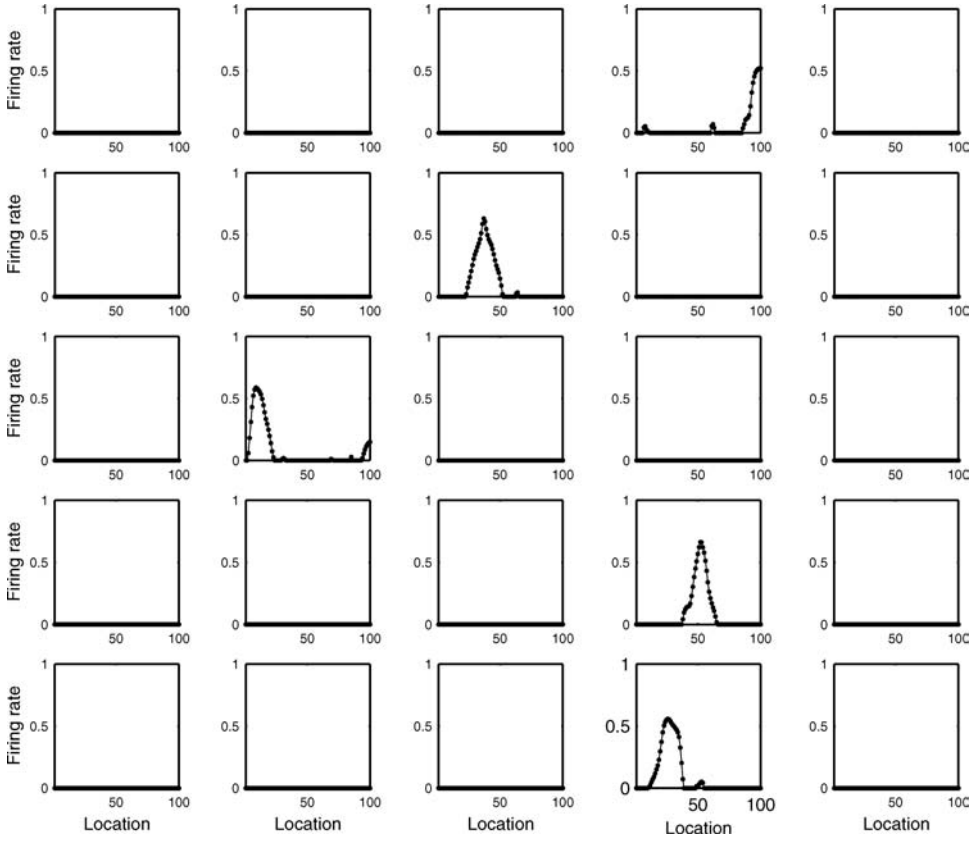

Figure 5. Firing rate profiles of 25 DG cells after learning with the trace rule, with $\eta = 0.8$ and learning rate 0.01. In this simulation, the standard deviation of the EC peak heights was set to 0.6 to provide greater variation. In this case, it is evident that there is less spatial repetition in the DG cells, and some DG cells have developed broad place fields.

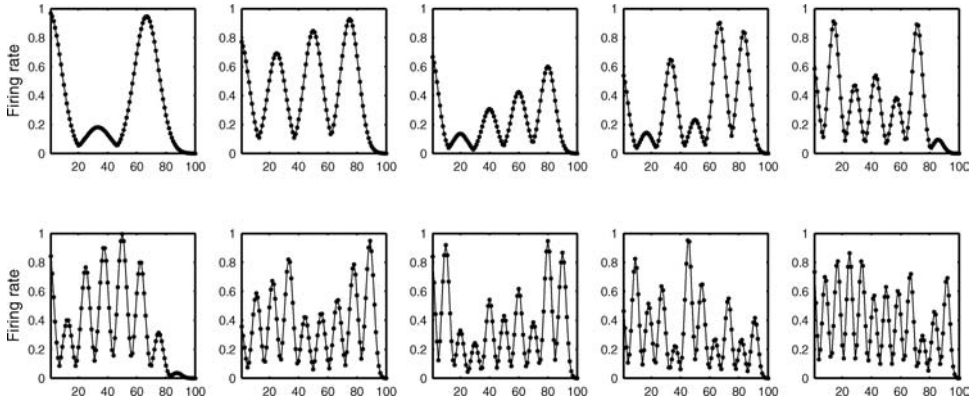

Figure 6. Firing rate profiles of 10 EC grid cells with frequencies 3–12. The standard deviation of the peak heights was set to 0.6.

high firing, and the persistence of the trace as the agent moves in the next few timesteps helps nearby cells to also learn, because the postsynaptic term in the learning rule continues for a period determined by η (see Eqn. 6). Figure 5 shows firing rate profiles of 25 DG cells after learning with the trace rule, with $\eta = 0.8$ and learning rate 0.01. In this simulation, the standard deviation of the EC peak heights was set to 0.6 to provide greater variation. Some DG cells have developed broad place fields. Further simulations showed that the value of η is a major factor that controls the size of the DG place field, with larger values of η that produce longer temporal traces producing broader place fields. The longer temporal trace means that as the agent traverses the environment, the firing at more distant locations can become associated by the temporal trace effect, and it is this that results in the useful contribution of trace learning to the formation of larger place fields than would otherwise occur.

The field of the neuron in the second row fifth column of Figure 4 shows a high firing rate throughout much of the space. Fields of this type occur as a result of continuous spatial transformation learning utilizing a purely associative synaptic modification rule, in which overlapping input representations in a space can all come to activate the same postsynaptic neuron, as described in detail elsewhere (Stringer, Perry, Rolls & Proske 2006). Factors that can minimize the occurrence of continuous spatial transformation learning include low learning rates, and a low probability of connections in the competitive network. In the simulations described here, it was rarely found with the full range of spatial frequencies used for Figure 4 (3–12 peaks in the space), and only became somewhat more frequent if the high spatial frequency EC cells were removed leaving only those with frequencies of 3–7 peaks in the space. With the normal full range of spatial frequencies present, there was sufficient irregularity in the input representations to the DG cells that the space was fractured sufficiently to minimize continuous spatial transformation learning.

The analysis of how place cells in the dentate gyrus/CA regions could be produced from EC grid cells presented above has been with one-dimensional simulations, as these are tractable and allow the factors that influence the mapping to be explored efficiently. Having understood some of these properties, we next tested that the same mapping principle can indeed provide an account for how the grid cells recorded in the entorhinal cortex with a representation in 2D space map to the dentate/CA representations also in 2D space. The 2D space was 100 $\times$ 100 training and testing locations. There were 125 EC cells with frequencies of 3, 4, 5, 6 and 7 cycles (to limit the computational time, and to capture the frequencies described neurophysiologically (Hafting et al. 2005)) and 25 phases for each frequency. (A phase is defined as an offset in the X and Y directions, and five offset values were used in each direction.) There were 100 DG cells, with the sparseness of the representation set to $a = 0.02$, and the learning rate set to 0.0001. Figure 7a and b show examples of two 2D EC grid cell (with frequencies of 4 and 7 cycles along the X axis). As in the neurophysiological recordings (Hafting et al. 2005), the firing peaks occur at the vertices of a grid of equilateral triangles. The standard deviation of the peak heights was set to 0.6 for the 2D simulations illustrated. Figure 7c and d show the firing rate profiles of two DG cells without training. Multiple peaks in the response profiles are evident, and this shows that the contrast enhancement and the competitive interactions in the DG layer were not sufficient without training to produce place-like fields in DG. The training consisted of the agent moving continuously through the 100 $\times$ 100 equispaced training locations (one column at a time from bottom to top in a column, followed by the same movement along the next column to the right), and repeating this for 100 such training epochs. We checked that the exact path followed during training is not critical to the mappings obtained. The mappings were influenced primarily by the response profiles of the grid cells.

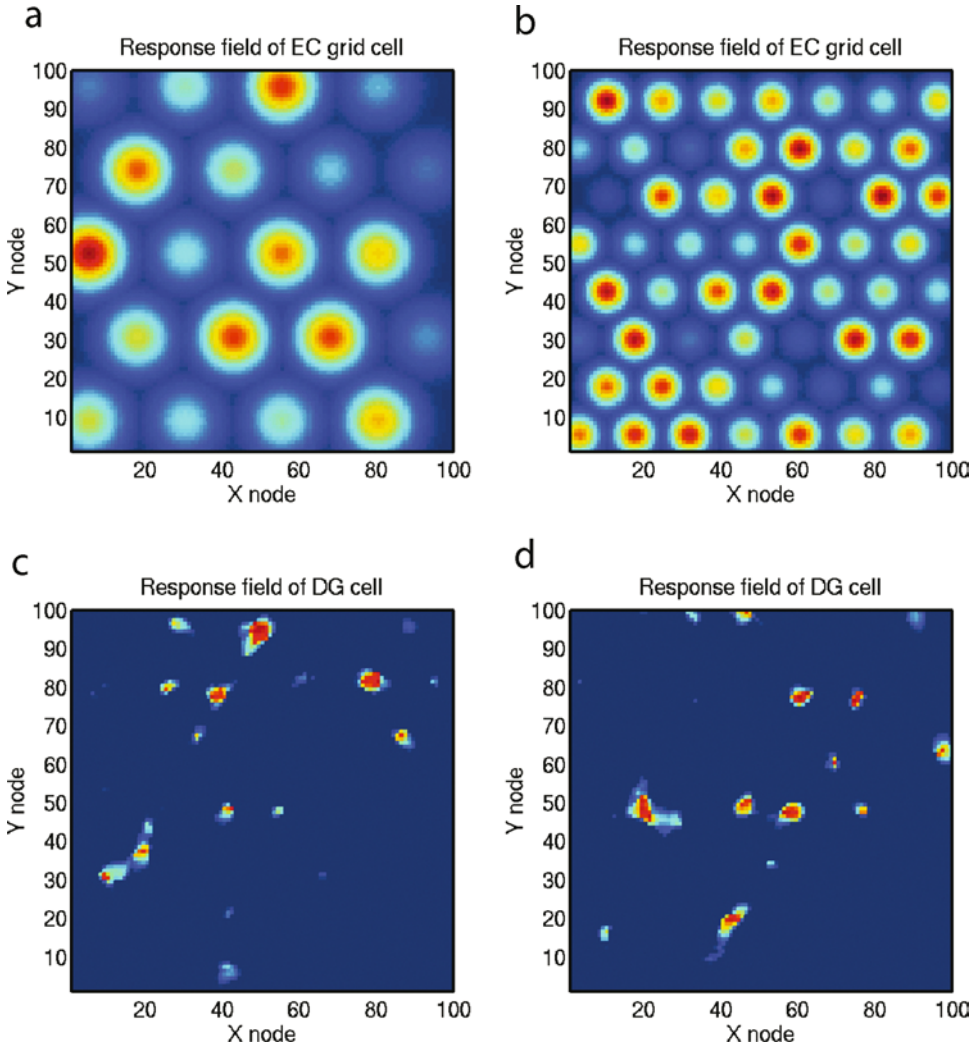

Figure 7. 2D simulation of place cells. a and b. Firing rate profiles of 2 EC grid cells with frequencies of 4 and 7 cycles. The standard deviation of the peak heights was set to 0.6. c and d. Firing rate profiles of two DG cells with no training. Linear mapping from the firing rates to the representation in the Figure was used.

Figure 8a and b show the firing rate profiles of two DG cells after training with the Hebb rule. This type of relatively small 2D place field was frequently found. Although the DG cells illustrated had one main place field, some evidence of some remaining grid-like firing was sometimes evident in the place fields with both Hebb and trace rule training, and indeed is evident in Figure 8c and d. These are also characteristics of dentate granule cells recorded under the same conditions as the EC grid cell experiments (E. Moser, personal communication) (see also (Jung & McNaughton 1993)). These results were found only with training, with untrained networks showing repeating peaks of firing in large parts of the space as illustrated in Figure 7c and d.

In extensive simulations with 6 separate runs for each condition (no training, training with the Hebb rule, and training with the trace rule), the following results were found. The mean number of peaks of each DG cell with training with the Hebb rule was 2.63 ± 0.10 (sd), and

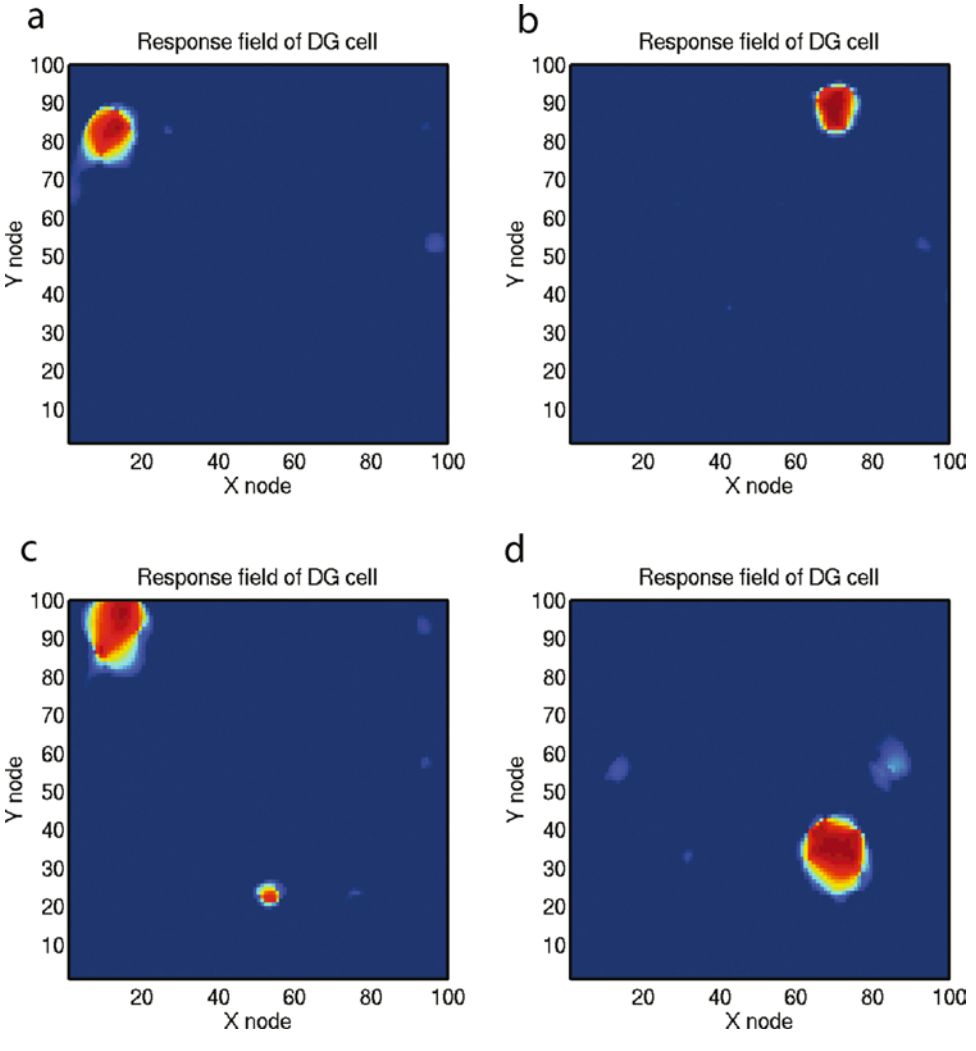

Figure 8. 2D simulation of place cells. a and b. Firing rate profiles of two DG cells after training with the Hebb rule. c and d. Firing rate profiles of two DG cells after training with the trace rule with $\eta = 0.8$. Linear mapping from the firing rates to the representation in the Figure was used.

without training was 6.81 ± 0.17 ($t = 17$, $df = 874$, $P < 10^{-57}$). The distributions were that with Hebb training, 36.5% of the DG cells had one peak, 23.1% had 2 peaks, and 13.5% had three peaks, whereas without training the distributions were 9.6%, 9.0% and 6.4%. (These distributions were significantly different, chi-square = 209, df = 3, $P \ll 0.001$.) In addition, the mean size of the DG fields after Hebb training, 112 ± 126 (sd) training locations, was larger than with no training, 22 ± 24 ($t = 40.7$, $df = 4625$, $P < 10^{-64}$). This analysis shows statistically that training using competitive learning does produce cells with place-like properties (i.e. with one or a few major peaks, as is typical of dentate granule cells, as noted below; and with much larger fields that were found in DG without training).

Figure 8c and d show the firing rate profiles of two DG cells after training with the trace rule with $\eta = 0.8$. Training with the trace rule produced significantly larger DG fields (166 ± 243 training locations) than did training with the Hebb rule ($t = 5.3$, $df = 1298$, $P = 1.4 \times 10^{-7}$).

It is known that the responses of rat place cells settle down in a new environment in a few minutes (Wilson & McNaughton 1993). It was therefore of interest to assess how rapidly the self-organizing learning of place fields described in this paper might take. The number of training epochs used for the results described in Figure 8 was 100. In further simulations, it was shown that when training was with 5 epochs, the results were very similar to those obtained with 100 epochs, and very different to those obtained without training and illustrated in Fig. 7c and d. Even with one training epoch (and with the learning rate increased to 0.01), considerable self-organization was found, with the resulting spatial fields showing place-like properties, though less perfectly than with training with 5 or 100 epochs, in that some neurons showed evidence of multiple small peaks, in addition to the main place field. The time it might take for a rat to be in one place in a real environment and for synaptic modification to occur might be in the order of 1 s. For the rat to move through a reasonable part of the environment might take 10–20 s, and this would correspond to one training epoch in the type of simulation described here. The implication is that if a rat were in each place in a new environment for several seconds, this would be sufficient for the formation of place fields using the type of self-organizing competitive learning described here to occur. This is consistent with the time course of what is found experimentally (Wilson & McNaughton 1993).

3. Discussion

The results described in this paper show that self-organizing competitive learning can account for how place fields in the dentate gyrus (DG) are generated from grid cells in the dorsocaudal medial part of the entorhinal cortex (dMEC). (A similar process could occur at the entorhinal to CA3 (perforant path) synapses, as this part of the connectivity has the architecture of a competitive network.) Further, we showed that to generate relatively broad place fields typical of those found in the hippocampus and DG, utilization of a modified associative rule that incorporates a temporal trace of preceding neural activity can be helpful. This effect emerges because as the animal traverses space, nearby places will tend to become associated. Distant places will not be temporally associated, and this help to prevent cells responding to multiple places in an environment.

The competitive network learns this mapping correctly by taking advantage of several properties of entorhinal cortex grid cells. One is that different grid cells (at different dorso-ventral positions in the entorhinal cortex) have different grid frequencies. We suggest that the reason that there are different frequencies is that a combination of cells with different frequencies and phases provides a unique representation of position. (If only one frequency was present, the firing would increase then decrease repeatedly as the animal locomoted, providing no indication of the absolute position. By including cells with different, and especially low, grid frequencies, a particular combination of grid cells firing each with different frequencies and phases effectively specifies a unique position.) The competitive network takes advantage of this, and also of the different peak heights within each cell, to map this highly distributed encoding of position into a much more sparse representation, in which a dentate cell fires primarily to one position in space. The competitive network solves this not just by the competition it implements between its neurons, but also by synaptic modification, which as shown in Figs. 3 and 4 and by the statistical analyses in the Results, is important for the network to achieve primarily unimodal place cells. Although as shown in Figures 3 and 7 and as also shown by Fuhs & Touretzky (2006) and McNaughton, Battaglia, Jensen, Moser & Moser (2006) simply setting up a sparse representation without training in DG can transform the

representation from grid cells, the synaptic modification inherent in a competitive network allows much better place-like fields to develop, as shown in Figs. 4 and Figure 8, and by the statistical analyses we present. (Fuhs & Touretzky (2006) used no learning, synapses of the same unitary strength but with low connectivity in their analysis; McNaughton et al. (2006) also used no learning, and suggested that place cells in the hippocampus might arise due to beat effects between the different spatial resolution representations in the entorhinal cortex.)

The investigations described here thus first emphasize the importance of learning, but second also emphasize the importance of the variation in the peak heights of a grid cell, which over a large population of grid cells helps to break the symmetry in the mapping from the entorhinal cortex to a place cell with a single field. Indeed, McNaughton et al. (2006) assumed that the peaks heights of an individual grid cell were similar, and hence in the cell they show with place-like properties in their Figure 6, the symmetry is not broken, and the place-like field retains evidence for symmetry. The third component that we suggest helps a unique mapping to occur to place cells is the low connection probability from the entorhinal cortex to the dentate granule cells, which results in different dentate / hippocampal cells responding to unique combinations of entorhinal cell firing, and therefore tending to have more unimodal and asymmetric fields. In addition to these three important contributions, a fourth is setting the representation to be sparse in the dentate / hippocampal cells. A fifth factor is that trace learning may help to increase the size of dentate / hippocampal place fields. Sixth, we note that because the grid cells with the largest spacing between peaks can have the peaks separated by quite large distances (E.Moser, personal communication), these low frequency grid cells may be particularly important in contributing to relatively unimodal place fields even in large environments. An additional factor that might help would be low frequency beats arising from the different fundamental frequencies of different grid cells, as suggested by McNaughton et al. (2006), though whether this would occur in practice in the context of the variability in the peaks heights of individual grid cells remains to be determined.

The present investigation is the first to show that grid cells can be mapped to place-like cells using a competitive network, and shows that the learning in the competitive network makes an important contribution to the mapping. The value for hippocampal computation of having a sparse representation of place (with place cells with unimodal fields) is it is suggested because this sparse representation can then be associated with other inputs in the hippocampus to memories, with a high capacity for the number of memories that can be stored and retrieved (Rolls 1989b; Treves & Rolls 1994; Rolls & Kesner 2006). Moreover, the pattern separation achieved by the dentate competitive network may be useful not only for forming place cells, but also for forming relatively orthogonal representations of other inputs about for example objects in the hippocampus received from the entorhinal cortex, so that episodic memories, with a high capacity, and representing for example object-place (Rolls, Xiang & Franco 2005; Rolls & Xiang 2006) or reward-place associations (Rolls & Xiang 2005), can be formed in the hippocampus.

It is an advantage of the proposed mechanism that the same competitive network operation implemented from EC to DG is potentially useful also for non-spatial representations, which become categorised, more orthogonal, and sparse in DG in preparation as input to the CA3 network, for which such an input representation is efficient in terms of storage capacity of the CA3 system (Rolls 1987; Rolls 1989a; Treves & Rolls 1992; Rolls 1996; Rolls & Treves 1998). It is thus a unifying principle, and consistent with earlier work, to propose that a competitive network is used to perform the mapping from entorhinal grid cells to dentate/ hippocampal place cells. The advantage of generating sparse representations in preparation

for a later autoassociation process has also been described by (Marr 1971), and (McNaughton & Nadel 1990).

As noted in the Methods, competitive networks are often envisaged as operating to categorise a discrete set of patterns. As the competitive net self-organizes, the weight vector of each output neuron moves towards the centre of a set of patterns in the input space (Rumelhart & Zipser 1986; Grossberg 1987; Hertz, Krogh & Palmer 1991; Rolls & Treves 1998; Rolls & Deco 2002). The competition between neurons ensures that each output neuron responds to a different set of patterns. The sparseness of the output representation could be as sparse as a winner-take-all representation with just one neuron representing the category, or the sparseness could be less, with distributed representations in the output (Rolls & Treves 1998; Rolls & Deco 2002). The number of categories formed depends on the number of output neurons, and on the sparseness of the output representation (Rolls & Treves 1998; Rolls & Deco 2002). The important issue in the current context is that the inputs are normally conceived of as being separable into discrete categories, in that the vector correlations between some of the exemplars or training patterns are higher than the correlations between other training patterns.

What happens if there are no discontinuities in the input space, so that all the patterns overlap continuously with each other? This is what might be considered prototypical of spatial representations, where as one moves through the space, there is continuous overlap between the spatial fields of neurons. When continuous patterns of this type are used to train a competitive network, the neurons tend to distribute themselves evenly throughout the space (Hertz et al. 1991). Sharp (1991) took advantage of this property in her model of the formation of hippocampal place cells from entorhinal cortex inputs, but this was before grid cells had been discovered in the entorhinal cortex. In this paper, we have shown that a competitive network can account for the formation of dentate and hippocampal place cells from entorhinal cortex grid cells, and have also investigated how a temporal trace can influence the learning of the mapping.

The 2D simulations illustrated in Figure 8 show that the competitive learning principles we describe in this paper can account for the mapping of 2D grid cells in the entorhinal cortex to 2D place cells in the dentate gyrus / CA regions. The competitive learning principle that can account for this mapping could be implemented at the entorhinal cortex to DG cell synapses, in that small place fields (Jung & McNaughton 1993) with some evidence for weak grid-like properties (E. Moser, personal communication) are found in DG. However, some place fields are found in CA1 even when this is disconnected from DG and CA3 (Brun, Otnass, Molden, Steffenach, Witter, Moser, & Moser 2002). We believe that the same competitive network functionality operating in effectively the same type of architecture, with EC cells making direct synaptic connections to CA1 cells, could account for the CA1 place fields recorded in these conditions. Having the dentate gyrus as a first stage may help to produce particularly small place fields in DG (Jung & McNaughton 1993), and this may be part of its contribution to spatial pattern separation, allowing distinct memories to be formed of nearby spatial positions (Rolls & Kesner 2006). We also note that the entorhinal cortex has direct connections to the CA3 cells, and these could operate to help in the formation of CA3 place cells by competitive learning. However, in this paper we have focussed more on the formation of dentate gyrus place cells, because this is the first stage following the entorhinal cortex at which neurons with place fields are found. The dentate cells then directly project to the CA3 cells.

Part of the interest of the theory described in this paper is that a process as simple as competitive learning could account for the mapping from entorhinal cortex grid cells to hippocampal place cells. This is an elegant solution, which is consistent with prior concepts that

the entorhinal cortex to dentate projection is a competitive network that prepares the representation for associative learning in CA3 by making the representations more orthogonal and sparse (Rolls 1989a, 1989d, 1989c, 1990; McNaughton & Nadel 1990; Roll & Kesner 1998) thus producing spatial pattern separation (Gilbert, Kesner & Lee 2001; Rolls & Kesner 2006). We know of no other demonstration of how entorhinal cortex grid cells could map to hippocampal place cells using competitive learning. In addition, the theory does make predictions. One is that learning of the type implemented in competitive networks is required for the correct mapping. We demonstrated that without learning, the mapping achieved still has multiple peaks. The prediction thus is that, if the associative synaptic modification that is needed for competitive learning (Hertz et al. 1991, Rolls & Deco 2002) was impaired by interference (e.g. a knockout) with the NMDA receptors on the receiving neurons (e.g. dentate granule cells), then the dentate spatial fields would not have a single place represented. A further prediction is that because the competitive network learns the mapping in an unsupervised way, dentate / CA3 place fields would still develop normally even in the absence of the visual room-based inputs (which can influence hippocampal neurons (Shapiro & Eichenbaum 1999)). A third prediction is that, as some dMEC cells have conjunctive encoding of grid and head direction properties (Sargolini, Fyhn, Hafting, McNaughton, Witter, Moser & Moser 2006), then we predict with the type of model we describe that some head direction tuning of dentate and thus hippocampal CA3 and CA1 place cells might occur (even when this is not encouraged by training in directional environments such as a linear runway).

We finish this paper by commenting on the importance of transforming from a grid cell representation in the entorhinal cortex to a place cell representation (or in primates a spatial view cell representation) in the hippocampus. At least one important function this allows is the formation of memories, formed for example between an object and a place, which is proposed to be a prototypical function of the hippocampus that is fundamental to episodic memory (Rolls 1996; Rolls & Treves 1998; Rolls & Kesner 2006; Rolls & Xiang 2006; Rolls 2007). For the formation of such object-place memories in an associative network (in particular, an autoassociative network in the CA3 region of the hippocampus), the place must be made explicit in the representation, and moreover, for high capacity, that is the ability to store and retrieve many memories, must be sparse. By made explicit, we mean that information can be read off easily from the firing rates of the neurons, with different places producing very different sets of neurons firing, so that different representations are relatively orthogonal. The entorhinal cortex representations are not only not sparse (in our simulations, the sparseness was 0.54), but in addition the typical overlap between the sets of firing of the neuronal populations representing two different places was high (with a mean cosine of the angle or normalized dot product 0.540). In contrast, the sparseness of the representations formed in the DG cells with Hebb rule training was 0.024, and a typical cosine of the angle was 0.000). (Low values of this measure of sparseness, a, defined in Eqn (3), indicate sparse representations. The cosine of the angle between two vectors takes the value 1 if they point in the same direction, and the value 0 if they are orthogonal.) These sparse and orthogonal representations are what is required for high capacity storage of object and place, object and reward, and of in general episodic memories, and this is the function, we believe, of the mapping from entorhinal cortex to hippocampal cells for which we have produced a computational model here. This concept maps well onto utility in an environment, for it is the place where we are located or at which we are looking in the world with which we wish to associate objects or rewards, and this is made explicit in the DG / hippocampal representation (Rolls 1999; Rolls, Xiang & Franco 2005; Rolls & Xiang 2005; Rolls & Kesner 2006; Rolls & Xiang 2006). In contrast, the entorhinal cortex may

represent self-motion space in a way that is suitable for idiothetic path integration in *any* environment.

Acknowledgements

This research was supported by the Medical Research Council.

References

Brown TH, Kairiss EW, Keenan CL. 1990. Hebbian synapses: Biophysical mechanisms and algorithms. Annual Review of Neuroscience 13:475–511.

Brun VH, Otnass MK, Molden S, Steffenach HA, Witter, MP, Moser, MB, Moser, EI. 2002. Place cells and place recognition maintained by direct entorhinal–hippocampal circuitry. Science 296:2243–2246.

De Araujo IET, Rolls ET, Stringer SM. 2001. A view model which accounts for the response properties of hippocampal primate spatial view cells and rat place cells. Hippocampus 11:699–706.

Fuhs MC, Touretzky DS. 2006. A spin glass model of path integration in rat medial entorhinal cortex. Journal of Neuroscience 26:4266–4276.

Gilbert PE, Kesner RP, Lee I. 2001. Dissociating hippocampal subregions: Double dissociation between dentate gyrus and CA1. Hippocampus 11:626–636.

Grossberg S. 1987. Competitive learning: From interactive activation to adaptive resonance. Cognitive Science 11:23–63.

Hafting T, Fyhn M, Molden S, Moser MB, Moser EI. 2005. Microstructure of a spatial map in the entorhinal cortex. Nature 436:801–806.

Hertz JA, Krogh A, Palmer RG. 1991. Introduction to the Theory of Neural Computation, Addison-Wesley, Wokingham, UK.

Jeffery KJ, Anderson MI, Hayman R, Chakraborty S. 2004. A proposed architecture for the neural representation of spatial context. Neuroscience and Biobehavioral Reviews 28:201–218.

Jeffery KJ, Hayman R. 2004. Plasticity of the hippocampal place cell representation. Reviews in the Neurosciences 15:309–331.

Jung MW, McNaughton BL. 1993. Spatial selectivity of unit activity in the hippocampal granular layer. Hippocampus 3:165–182.

Levy WB. 1985. Associative changes in the synapse: LTP in the hippocampus, in WB Levy, JA Anderson & S Lehmkuhle (eds.) Synaptic Modification, Neuron Selectivity, and Nervous System Organization, Erlbaum, Hillsdale, NJ, chapter 1, pp. 5–33.

Levy WB, Desmond NL. 1985. The rules of elemental synaptic plasticity, in WB Levy, JA Anderson & S Lehmkuhle (eds), Synaptic Modification, Neuron Selectivity, and Nervous System Organization, Erlbaum, Hillsdale, NJ, chapter 6, pp. 105–121.

Marr D. 1971. Simple memory: A theory for archicortex. Philosophical Transactions of The Royal Society of London, Series B 262:23–81.

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

McNaughton BL, Battaglia FP, Jensen O, Moser EI, Moser M-B. 2006. Path integration and the neural basis of the hippocampal map. Nature Reviews Neuroscience 7:663–676.

McNaughton BL, Morris RGM. 1987. Hippocampal synaptic enhancement and information storage within a distributed memory system. Trends in Neuroscience 10:408–415.

McNaughton BL, Nadel L. 1990. Hebb-Marr networks and the neurobiological representation of action in space, in MA Gluck & DE Rumelhart (eds). Neuroscience and Connectionist Theory, Erlbaum, Hillsdale, NJ, pp. 1–64.

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

Muller RU, Kubie JL, Bostock EM, Taube JS, Quirk GJ. 1991. Spatial firing correlates of neurons in the hippocampal formation of freely moving rats, in J. Paillard (ed.). Brain and Space, Oxford University Press, Oxford, pp. 296–333.

Oja E. 1982. A simplified neuron model as a principal component analyzer. Journal of Mathematical Biology 15:267–273.

O'Keefe J. 1984. Spatial memory within and without the hippocampal system, in W. Seifert (ed.). Neurobiology of the Hippocampus, Academic Press, London, pp. 375–403.
O'Keefe J, Burgess N. 2005. Dual phase and rate coding in hippocampal place cells: theoretical significance and relatioship to entorhinal grid cells. Hippocampus 15:853–866.
O'Keefe J, Dostrovsky J. 1971. The hippocampus as a spatial map: preliminary evidence from unit activity in the freely moving rat. Brain Research 34:171–175.
Redish AD. 1999. Beyond the Cognitive Map: From Place Cells to Episodic Memory, MIT Press, Cambridge, Massachusetts.
Redish AD, Touretzky DS. 1998. The role of the hippocampus in solving the Morris water maze. Neural Computation 10:73–111.
Rolls ET. 1987. Information representation, processing and storage in the brain: analysis at the single neuron level, in J.-P. Changeux & M. Konishi (eds). The Neural and Molecular Bases of Learning, Wiley, Chichester, pp. 503–540.
Rolls ET. 1989a. Functions of neuronal networks in the hippocampus and neocortex in memory, in J. Byrne & W. Berry (eds). Neural Models of Plasticity: Experimental and Theoretical Approaches, Academic Press, San Diego, chapter 13, pp. 240–265.
Rolls ET. 1989b. Information processing and basal ganglia function, in C. Kennard & M. Swash (eds). Hierarchies in Neurology, Springer-Verlag, London, chapter 15, pp. 123–142.
Rolls ET. 1989c. Parallel distributed processing in the brain: implications of the functional architecture of neuronal networks in the hippocampus, in RGM Morris (ed.). Parallel Distributed Processing: Implications for Psychology and Neurobiology, Oxford University Press, Oxford, chapter 12, pp. 286–308.
Rolls ET. 1989d. The representation and storage of information in neuronal networks in the primate cerebral cortex and hippocampus, in R. Durbin, C. Miall & G. Mitchison (eds). The Computing Neuron, Addison-Wesley, Wokingham, England, chapter 8, pp. 125–159.
Rolls ET. 1990. Theoretical and neurophysiological analysis of the functions of the primate hippocampus in memory. Cold Spring Harbor Symposia in Quantitative Biology 55:995–1006.
Rolls ET. 1995. A model of the operation of the hippocampus and entorhinal cortex in memory. International Journal of Neural Systems 6, Supplement:51–70.
Rolls ET. 1996. A theory of hippocampal function in memory. Hippocampus 6:601–620.
Rolls ET. 1999. Spatial view cells and the representation of place in the primate hippocampus. Hippocampus 9:467–480.
Rolls ET. 2007. Memory, Attention and Decision-Making, Oxford University Press, Oxford.
Rolls ET, Deco G. 2002. Computational Neuroscience of Vision, Oxford University Press, Oxford.
Rolls ET, Kesner RP. 2006. A theory of hippocampal function, and tests of the theory. Progress in Neurobiology 79:1–48.
Rolls ET, Treves A. 1998. Neural Networks and Brain Function, Oxford University Press, Oxford.
Rolls ET, Xiang J-Z. 2005. Reward–spatial view representations and learning in the primate hippocampus. Journal of Neuroscience 25:6167–6174.
Rolls ET, Xiang J-Z. 2006. Spatial view cells in the primate hippocampus, and memory recall. Reviews in the Neurosciences 17:175–200.
Rolls ET, Xiang J-Z, Franco, L. 2005. Object, space and object-space representations in the primate hippocampus. Journal of Neurophysiology 94:833–844.
Rumelhart DE, Zipser D. 1986. Feature discovery by competitive learning. in DE Rumelhart & JL McClelland (eds). Parallel Distributed Processing, Vol. 1, Foundations, MIT Press, Cambridge, Mass, pp. 151–193.
Sargolini F, Fyhn M, Hafting T, McNaughton BL, Witter MP, Moser MB, Moser EI. 2006. Conjunctive representation of position, direction, and velocity in entorhinal cortex. Science 312:758–762.
Shapiro ML, Eichenbaum H. 1999. Hippocampus as a memory map: synaptic plasticity and memory encoding by hippocampal neurons. Hippocampus 9:365–384.
Shapiro ML, Hetherington PA. 1993. A simple network model simulates hippocampal place fields: parametric analyses and physiological predictions. Behavioral Neuroscience 107:34–50.
Sharp PE. 1991. Computer simulation of hippocampal place cells. Psychobiology 19:103–115.
Stringer SM, Perry G, Rolls ET, Proske J.H. 2006. Learning invariant object recognition in the visual system with continuous transformations. Biological Cybernetics 94:128–142.
Stringer SM, Trappenberg TP, Rolls ET, De Araujo IET. 2002. Self-organizing continuous attractor networks and path integration: One-dimensional models of head direction cells. Network: Computation in Neural Systems 13:217–242.
Taube JS, Muller RU, Ranck Jr, JB. 1990. Head-direction cells recorded from the postsubiculum in freely moving rats. II. Effects of environmental manipulations. Journal of Neuroscience 10:436–447.

Treves A, Rolls ET. 1992. Computational constraints suggest the need for two distinct input systems to the hippocampal CA3 network. Hippocampus 2:189–199.

Treves A, Rolls ET. 1994. A computational analysis of the role of the hippocampus in memory. Hippocampus 4:374–391.

Wilson MA, McNaughton BL. 1993. Dynamics of the hippocampal ensemble code for space. Science 261:1055–1058.