An Allocentric Spatial Model for the Hippocampal Cognitive Map

John O'Keefe

Anatomy and Developmental Biology, University College London, Gower Street, London WC1E 6BT, U.K.

SpatiaL Behaviors vs. SpatiaL Neural Systems

The term "spatial behavior" is ambiguous and needs to be clarified. It might mean behaviors that can best be described with reference to a spatial reference framework, or it might mean behaviors that are dependent on spatial neural systems. An example of the first is the demonstration by Golani (1976) that the seemingly complex movements which form the social interaction between two jackals can be described in several different frameworks, including the use of a framework centered on one animal to describe the movements of the other. Other aspects of behavior will best be described with reference to egocentric frameworks centered on the eye, head or body axis, or axes centered on individual objects or environmental frameworks. It might transpire, however, that the neural systems that generate these behaviors are not in themselves spatial systems, but are acting according to nonspatial principles that conspire with the environment to produce "spatial" behavior. For example, the behavior of an animal approaching a light might best be described as the linear distance between the animal and the object, but might in fact be subserved by systems operating on the principle "increase the intensity of the light," resulting in the animal's moving up an intensity gradient (Barto and Sutton, 1981).

Similarly, as the behaviorists realized, even complex spatial behaviors such as maze learning could, in principle, be dependent on chains of simple associations between stimuli and responses (see stimulus A, activate muscle sequence I; see stimulus B, activate muscle sequence II, etc.). An example of the way in which a "nonspatial" hippocampus might support spatial behavior is the stimulus-response associationist model proposed by McNaughton (1988) and McNaughton and Nadel (1989). In this model the activity of the hippocampal complex-spine cells represents the sensory view from a location; the role of the hippocampus is to associate this view with a particular movement (e.g., turn left) in order to predict the subsequent local view. The power to generate novel behaviors on the basis of such a system depends on the ability to string together stimulus-response-stimulus components and to generate motor equivalences, such as two left turns are equivalent to an about-face in extrahippocampal areas like the parietal cortex. No spatial system is needed.

The second meaning of spatial behavior is behavior generated by information stored in an explicitly spatial system. The strong claims of the cognitive map theory are that one such spatial system represents environments within an allocentric spatial framework, that the ordering principles of the system are the Euclidean distance and direction between places within that framework, and that the system is localized within the nervous system to the hippocampal region.

The role of this hippocampal place system in "spatial" behavior and "spatial" tasks is not straightforward and requires a detailed analysis of each task or behavior in terms of the possible strategies or hypotheses that might underpin it and the environmental manipulations that would either force the animal to select one hypothesis or, alternatively, enable the experimenter to identify which of several possible hypotheses an animal was actually using. Lynn Nadel and I devoted considerable space in our review of the lesion literature to this task, but misunderstanding still persists. For example, the Olton radial arm maze (Olton and Samuelson, 1976) and the Morris water maze task (Morris, 1981) are considered to be "spatial" tasks, and even partial success on either by hippocampal animals is counted against the cognitive map theory. Yet they both can be solved by non-place strategies depending on the particular configuration and procedure used.

The usual version of the Morris task has distinct advantages over the Olton maze in that direction hypotheses can be used to solve the latter but are not very efficient in the water task since the animal is started from several locations and must approach the goal from a different direction on different trials. The water maze can, however, be solved by a nonhippocampal guidance strategy if there is a prominent cue outside the pool behind the goal and the animal is started from the same location on every trial. In this case, use of a guidance strategy would be demonstrated by relocation or removal of the cue once the animals had learned. Eichenbaum et al. (1990) have recently demonstrated guidance learning in the water maze by placing black curtains and a relay rack behind the goal and starting the animals from the same location on each trial. Probe trials confirmed the vital role of the cues in guiding behavior. Damage to the hippocampal system had minimal effect as predicted by the cognitive map theory. Surprisingly, the authors misinterpreted these results as showing that the lesioned animals had demonstrated place learning and concluded that they counted against the cognitive map theory!

I would like to suggest that the use of the ambiguous terms "spatial behavior," "spatial learning," and "spatial task" be discontinued and that, in future, predictions about the behavioral effects of hippocampal lesions be derived in the manner that Nadel and I advocated and that I have outlined above. In particular, they should be derived from theories that have as their core a model that can be related to the anatomy and physiology of the hippocampal system. I am convinced that, without such a constraint, the cognitive map, as with so many ideas in this field, reduces to just one more vague notion that each player is free to interpret in his or her own way, and over which endless semantic battles can be fought. In the remainder of this article I will sketch out one version of such a computational model.

Hippocampal Units and the Map Model

As Lynn Nadel mentioned in the target article, the genesis of the cognitive map theory was the discovery of the spatial cells of the hippocampus in freely moving rats (O'Keefe and Dostrovsky, 1971). In addition, however, many of the prop-
erties attributed to the cognitive mapping system were derived from the properties of these and other cell types in the hippocampus. Three types of cells were described: place cells, displase cells, and misplace cells. These three classes of neurons had comparable counterparts in cell types described by Ranck (1973) in his coeval study, although he did not emphasize their spatial characteristics. The early non-computational models (Nadel and O'Keefe, 1974; O'Keefe and Nadel, 1974) were based on the assumption that each place cell represented a patch of an environment and that these patches were connected together by the displace cells, which were suggested to code for the distance between places on the basis of the animal's movements. The notion of places in an allocentric space derived from the observation that although some of the place cells had directional firing fields, others were omnidirectional, firing equally well regardless of the direction in which the animal faced within the field (O'Keefe, 1976). Further, non-directional local cues (olfactory and tactile) could be ruled out as sole determinants of place field firing by substitution of a different arm for the usual arm. This dominance of extramaze over intramaze cues in the control of place firing has been shown repeatedly for elevated mazes (Kubie and Ranck, 1983; Olton et al., 1978), in contrast to enclosed boxes in which intramaze cues predominate (Kubie and Ranck, 1983). The importance of the directionality correlate for the development of models continues to the present. McNaughton et al. (1983) found that most of the place cells were directional when animals were tested on an eight-arm radial maze with narrow arms. Bostock et al. (1988) have confirmed this but also found that none of the 13 place cells tested with the rats in a cylinder had directional fields. McNaughton (1988) has suggested that the strength of the directionality factor in the radial maze might mean that the cells are in fact responding to the local view from the field location; he has proposed a theory of hippocampal function based on this finding (see above). In contrast, I prefer to believe that the directionality of the place field firing depends on the shape of the environment and/or is imposed by the direction cells in the presubiculum.

The notion that there existed a subsystem for changing the representation within the map from the current location to the next location following a movement derived from the finding that the displace cells, like the EEG theta waves (Van derwolf, 1969), were clearly related to movements but not to the direction of the movement or the specific muscle groups involved. It was noted that in order to predict the next location, directional as well as distance information would be necessary, and the absence of any pure directional information in the hippocampal cellular response was considered problematic for the mapping theory; the existence of such information was predicted (O'Keefe, 1976; O'Keefe and Nadel, 1978). The major discovery of the head direction cells by Ranck and his colleagues (Ranck, 1984; Taube et al., 1990) in the dorsal presubiculum has removed the last major obstacle in the construction of a computational theory for the cognitive map.

The idea that learning changes within the mapping system were not driven by biological needs or reinforcers but by mismatches between the current sensory inputs and those expected on the basis of the latest map of that environment derived from several unit observations. Primary amongst these was the existence of misplace cells, which signaled the absence of expected items or the presence of unexpected items in a particular place (O'Keefe, 1976). Furthermore, place cells could be found in nongoal areas of an environment or, indeed, in environments entirely bereft of rewards and punishments. Recently, we (Speakman and O'Keefe, 1990) have strengthened this postulate of a cognitive motivation by showing that relocating the reward in a + maze task had no effect on the fields of almost all of the hippocampal place cells recorded during the task.

The modifiability of the place cells and their suitability as memory elements was first suggested by the observation that they changed their response to environmental alterations such as turning the lights off or rotating the maze. Usually on the very first exposure to the change, the place field firing disappeared, but the cell quickly "learned" to maintain its field after one or two experiences of the altered environment. The responses of these cells to repeated exposure to the dark has been examined more extensively by Jones-Leonard et al. (1985) and by Quirk et al. (1990). Both found that the cells maintained their fields if the rat was in the environment when the lights were turned off. This ability of these cells to remember the spatial properties of the environment after a brief exposure has been explored in experiments recently published by Andrew Speakman and myself (O'Keefe and Speakman, 1987). We studied the spatial representation in the hippocampus of an environment in which we controlled the extramaze spatial cues by rotating them as an ensemble from trial to trial. The rat had to solve a simple place learning task on an elevated + maze where the goal was identified by its constant relationship to these cues. Hippocampal complex spike cells had place fields that maintained their location relative to the controlled spatial cues, replicating the previous findings of O'Keefe and Conway (1978). By exposing the rat to a particular configuration of cues at the beginning of a trial and then removing them before it was allowed to run to the goal arm, we established both that the place cells maintained the appropriate firing fields during this memory period and that the rat remembered where the goal was at the end of each trial. This knowledge was usually not revealed in any aspect of the animal's behavior in the start arm. The rat appeared to have a map of the cue-controlled environment, which it could orient in four positions relative to the laboratory frame, and which it could maintain following the removal of the spatial cues. Detour trials in which the rat was forced to go to a nongoal arm during the memory period showed that once the appropriate hippocampal representation had been set up, it could be continuously updated in the absence of the controlled cues. Control trials in which the spatial cues were absent throughout the trial showed that the rat still entertained an internally consistent map in a particular orientation and chose the goal arm that was consistent with this map, but that the orientation could change from one control trial to the next. The rat went where it thought the goal was. That the goal was subordinate to and located within the map, and not vice versa, was demonstrated by a subsequent experiment (Speakman and O'Keefe, 1989). Here, it was shown that changing the food location relative to the controlled cues half-way through the recording period and re-
training the animal to the new goal location did not alter the spatial firing fields of most cells (see Breese et al., 1989 for an alternative view). Experiments in a cylindrical environment in which the fields could be rotated by rotation of a polarizing white card demonstrated identical effects, although it was not possible to test the animal's knowledge of its environment since there was no explicit goal (Muller et al., 1987; Muller and Kubie, 1987). If anything, it could be presumed that in this latter experiment, as in an earlier experiment by McNaughton et al. (1983) on a forced choice radial arm maze, the rats were not using a place strategy to solve the task. The hippocampus appears to track the animal's location even when not required to do so.

My recent efforts have been directed toward exploring how the hippocampus might provide an egocentric spatial framework for locating the array of objects within a head-centered framework (O'Keefe, 1988) or an allocentric framework centered on the environment (O'Keefe, 1990; 1991). I will briefly describe the major components of this latter model and discuss how it addresses some of the claims that have acted as a barrier to acceptance of the cognitive map theory, in particular, the existence of apparently nonspatial correlates of the single unit responses in the hippocampus of awake animals and the claim that hippocampal damage affects nonspatial memory.

THE CENTROID MODEL FOR THE HIPPOCAMPAL COGNITIVE MAP THEORY

The primary requirements of a spatial system are a coordinate system that serves as a reference framework within which the animal, the objects contained in the environment, and all other places occupied or unoccupied can be located. The model should, in addition, describe the machinery for changing the location of the representation of the animal during movement, for storing the locations of incentives or punishments, and for calculating the translation necessary to get from the current location (or any location) to a desired location. In addition, the model should explain how environmental representations are constructed and changed, and how the correct representation for the current environment is retrieved from all of the previously stored areas.

The current model suggests that the allocentric framework is a polar coordinate system whose center is the geometrical center of moments of the cues in the environment (centroid) and whose principal axis is the direction given by a measure of the departure of these cues from a symmetrical distribution (slope). The egocentric representation of cues on which the computations are based is shown in Figure 1A. The centroid is the average of the egocentric vectors from the animal's current location to each of the cues in the environment (Fig. 1C). The resultant vector identifies a location in the environment that is independent of the animal's location. The major axis for the allocentric framework is provided by the average slopes between all of the cue pairs in the environment (Fig. 1C). The computation of the slope between each cue pair is performed by subtraction (rather than addition) of their cue vectors. The average of these slopes gives an overall direction that can serve as the second major component of the mapping framework.

The calculation of the centroid is attributed to the CA3 pyramidal cells, each of which makes an estimate of the centroid on the basis of the limited subset of cues available to it. These mini-centroids form a matrix that uniquely identifies each place in every environment. There are several ways in which the place cells might encode this information. One possibility currently under investigation is that each vector is represented in phasor notation where the length of the vector or distance to the cue is represented by the amplitude of a sine wave, and the angle in egocentric space is represented by the phase of the sine wave relative to a reference (Fig. 1B). The theta system is viewed as a compound waveform (Vanderwolf and Leung, 1983), the function of one component being to provide the standard sinusoidal reference wave against which phase angles are measured. Vector addition in such a system is performed by linear summation of the sinusoids. Movements that translate and/or rotate the animal in the environment are also represented as vectors; this information is carried by a second component of the theta system.

The direction that serves as the allocentric axis is coded as a matrix of minislopes in the dorsal presubiculum where Rancck and his colleagues (Rancck, 1984; Taube et al., 1990) have reported finding the head-direction cells. In a fashion analogous to the use of the displacement component of the theta vector to update the location matrix during movement, the current heading direction is updated by rotational movement cues, such as those provided by head angle and joint receptors and the vestibular system. It is assumed that these are converted to the phasor notation and added to the current heading using linear addition. Thus, the system has two independent measures derived from the environment, the direction, which is independent of the animal's location in the environment, and the location, which is independent of the direction. Together they provide an allocentric reference framework, which fulfills all of the properties of a cognitive map.

Some of the criticisms of the theory can be addressed by the centroid model. For example, critics who seek to find a broader function for the hippocampus cite evidence of simple cue responses in hippocampal cells in support. O'Keefe (1976) reported a small number of hippocampal cells with non-spatial sensory responses, including several sniffing units, but chose not to emphasize them. For example, one cell fired when the animal sniffed at its water bowl, while another fired when the rat sniffed at the experimenter's hand. Eichenbaum et al. (1987) have reported that hippocampal cells fired during sniffing in an olfactory discrimination task and had spatial fields. These sniffing cells may be spatially dependent and, therefore, identical to the misplace cells described above, or they may in fact be pure olfactory cells, as Eichenbaum et al. suggest. The latter need not mean that the hippocampus has nonspatial as well as spatial functions. One version of the spatial model proposed above suggests that the firing rate of a place cell is an inverse function of the distance to the centroid. In this model, one would expect to find some cells that acted as though they were responding to individual sensory cues. Recall that each hippocampal complex-spike cell calculates an estimate of the centroid of the environment on the basis of the subset of cues available to it. In general, cells will have inputs from more than one cue in an environment
Fig. 1. Calculation of the centroid and slope of an environment on the basis of egocentric cue information. (A) The representation of two cues in an egocentric polar framework centered on the head. Each cue is represented by a vector (A and B), whose lengths RA and RB are related to the distance of the cues and whose angles (α and β) correspond to the angles in the head framework. Averaging of two or more cue vectors yields a vector (C) that points to the geometrical center (centroid) of the cue array. (B) One way in which the vector calculations might be performed in the hippocampal system. In phasor notation, each cue vector can be represented by a sinusoid in which the amplitude of the wave is related to the length of the vector RA, and the phase shift of the wave relative to a clock wave is related to the vector angle (α). The linear average of the sinusoids gives the vector average of the vectors. (C) The centroid and the slope of an environment with 5 cues, A–E. The length of the centroid vector is invariant with rotations by the animal in one location but varies with translations. The angle between the slope and the rat’s egocentric axis is invariant with translations but changes with rotations. Together, the centroid and the slope provide the basis for an allocentric coordinate framework.
and will compute a centroid vector pointing to a location devoid of cues. However, for those neurons that receive inputs from only one cue, the centroid estimate will be identical to the cue location, and the neuron will fire maximally to the cue itself. As the number of cues in an environment is reduced, these sensory responses should become more predominant.

Models such as the above should also help to clarify the circumstances under which the hippocampus might become involved in the solution of nonspatial tasks, such as those designed by Olton and his colleagues (Olton and Fisteus, 1981; Raphaele and Olton, 1988), in which several enclosed boxes, each containing a different set of cues, are interchanged within a trial so that the location of the boxes relative to the room or to one another is not a possible means to solution. While it is correct to conclude that no overall integrated map of the situation is possible under these circumstances, one cannot rule out the use of the hippocampal system in a reduced capacity if no easier or more efficient strategy is available. Perhaps the animal builds several independent maps, each based on the intrabox cues restricted to one of the boxes, and simply codes for the presence or absence of food in that environment or the recency of a visit to that environment. A similar comment applies to the configurational association ideas of Sutherland and Rudy (1989). They suggest that the hippocampus acts to recognize configurations of cues. Again, a hobbled hippocampus might be able to act in such a capacity in the absence of more natural solutions. I would predict the existence of place fields related to the intrabox cues in these nonspatial environments.

My final remarks concern the adequacy of the cognitive map model in explaining the role of the hippocampus in human amnesia. Other contributors to this exercise will undoubtedly marshal arguments in favor of the idea that amnesia is due to damage to a combination of structures in the temporal lobe, as originally suggested by Mishkin (1978), and that the contribution of the right hippocampus to this overall deficit is primarily or entirely spatial (Smith and Milner, 1981; 1989). My greater concern is with the apparently nonspatial role of the human left hippocampus. Our original idea of a semantic map (O'Keefe and Nadel, 1978), which uses an essentially spatial framework to organize linguistic entities into narratives, has not generated much attention or comment in the psycholinguistic literature. There are, however, continuing attempts to generate space grammars (Langacker, 1982) based to some extent on the notion that all sentences describe relations in symbolic spaces. For example, Langacker notes that prepositions have a primary meaning that is spatial and derive their secondary symbol meanings from this primary spatial root. While this seems obvious for prepositions such as to, at, or with, Langacker has even suggested that the use of the preposition by in passive sentences functions in a quasi-spatial way in the deep structural representation of those sentences.

I believe that the way forward here will also depend on the development of an adequate computational model for the human. Here, one can only speculate on some of the changes that would have to be incorporated into the rat model before it could begin to be applied to the human hippocampus. The idea that the left human hippocampus stores narratives commits one to the notion of a temporal dimension to the maps. In addition, the model will need to incorporate the notion of imagined movement within the mapping space in the absence of physical movement, the concept of an agent as a map entity that can act in the absence of external forces and that contains its own embedded maps as the wellsprings of action, and the notion of self-reference in which there is a representation of the organism in its own map. The language function of the human left hippocampus needs to be explored within this elaborated version of the cognitive mapping system. On this model, language is the means for transferring information from the speaker’s to the listener’s map. 

ACKNOWLEDGMENTS

The research reported in this article was supported by the Medical Research Council of Britain.

References


Muller, R. U., and J. L. Kubie (1987) The effects of changes in the