A mechanism for mental time travel? A critical review of Hasselmo's How we remember: Brain mechanisms of episodic memory

Sarah Robins
Published online: 30 Jan 2014.

PLEASE SCROLL DOWN FOR ARTICLE

REVIEW ESSAY

A mechanism for mental time travel? A critical review of Hasselmo's *How we remember: Brain mechanisms of episodic memory*

Sarah Robins

How We Remember: Brain Mechanisms of Episodic Memory
Michael E. Hasselmo
Cambridge, MA: MIT Press, 2011
384 pages, ISBN: 0262016354 (hbk); $40.00

1. Introduction

Episodic memories are memories for particular past experiences, from walking down the aisle to parking the car outside a store. Recalling these episodes can feel like reliving them: they are accompanied by a level of phenomenological detail that memories for facts, phone numbers, and faces are not. Episodic memory is thus often considered the most important, and yet most difficult, form of memory to explain. Tulving (2002) has described it as "mental time travel," a characterization as apt as it is unilluminating. An *explanation* of episodic remembering requires us to say more.

Neuroscientist Michael Hasselmo's recent book *How We Remember: Brain Mechanisms of Episodic Memory* is an attempt to provide the requisite explanation. The book is rich and ambitious. As this is a philosopher's review of a scientific account of memory, my aim is not to provide an exhaustive critique of Hasselmo's experimental framework. Instead, the review that follows is centered on the question of whether this book lives up to its title. That is, does Hasselmo explain *How We Remember*? Knowing whether he succeeds requires, first, establishing what an explanation of remembering requires.

Sarah Robins is Assistant Professor of Philosophy at the University of Kansas.
Correspondence to: Sarah Robins, Dept. of Philosophy, University of Kansas, Wescoe Hall 3090, 1445 Jayhawk Blvd., Lawrence, KS 66045, USA. Email: skrobins@ku.edu

Mechanistic explanations of memory are familiar territory for philosophers of science (e.g., Bechtel, 2008; Craver, 2007). But while memory serves as a common case study in accounts of mechanistic explanation, little attention has been paid to the unique demands of this explanandum. Hasselmo's book-length treatment of memory mechanisms provides an opportunity to correct for this oversight. Memory involves three processes: encoding, storage, and retrieval. The last of these is often thought to be the most important, and yet the most neglected (e.g., Roediger, 2000). Most mechanistic accounts of memory focus on encoding and storage—explaining how information is acquired and accommodated by particular brain structures, paying little to no attention to how this information is retrieved for subsequent use. They are thus poorly equipped to solve the *retrieval problem*: how does the brain manage to retrieve a particular episode rather than some other?

I begin, in the next section, with a sketch of Hasselmo's account of episodic remembering. In section 3, I highlight three features of Hasselmo's mechanism that speak to current issues and debates in the philosophical literature on mechanistic explanation. In the final section, I introduce the retrieval problem and use it to frame a critique of Hasselmo's account of retrieval. Ultimately, I conclude that while Hasselmo offers a compelling account of how the brain encodes and integrates the spatial and temporal features of an experience, the mechanism he provides cannot solve the retrieval problem. Hasselmo's account thus falls short of explaining how we remember.

2. Hasselmo: Episodic Memory as Spatiotemporal Trajectories

Hasselmo adopts a simple, straightforward definition of episodic memory as "memory of specific events occurring at a specific place and time" (p. 1). His resulting aim is an account of how the brain encodes the spatial and temporal features of an experience, such that they can be connected to meaningful objects and events and then later retrieved. He proposes a phase coding mechanism, whereby oscillations in the rhythmic activity of the hippocampal formation—a part of the brain long known to be involved in memory—are used to record and integrate various dimensions of an experience at multiple spatial and temporal scales.[1]

Hasselmo derives support for this mechanism from neurophysiological research, conducted primarily with rats, and from mathematical models. He builds the account slowly, beginning with a survey of behavioral (chapter 1) and neuroanatomical (chapter 2) research into episodic memory. The review is expansive and impressive, but proceeds at a pace that should be comfortable for any philosopher who has a passing familiarity with neuroscience. He then devotes the central chapters to a discussion of the entities (chapter 3) and activities (chapters 4 and 5) that make up his model of the phase coding mechanism. The final two chapters address pharmacological influences on the mechanism (chapter 6) and the mechanism's role in guiding future behavior (chapter 7). The details of his account require considerably more effort to understand, but should nonetheless be of interest to philosophers of science, psychology, and mind who are interested in mechanistic

explanation and the nature of memory and mental representation. For those interested in learning more about the mathematical models used to simulate episodic remembering, there is a technical—but still eminently readable—appendix.

Hasselmo's account of episodic memory builds on the well-known *cognitive map theory* of the hippocampus, according to which spatial memory is explained by maps of the organism's environment that are stored in the hippocampus. This theory has been used to capture a range of activities, from maze-running in rats to the navigational expertise of London taxi drivers (Maguire, Frackowiak, & Frith, 1997; O'Keefe & Dostrovsky, 1971). Central to this account is the discovery of place cells: neurons in the hippocampus that fire preferentially—exhibiting a burst of action potentials—in response to a specific location in the organism's environment. When a rat navigates a maze, for example, some place cells fire at the beginning of the maze, others at the first fork, still others at the second fork, and so on. These place cells are organized topographically, so that their collective firing pattern reflects the rat's route. After the maze has been run, the pattern is rehearsed, establishing a "map" that allows the rat to navigate this environment more easily the next time it is encountered.

Hasselmo expands the range of entities and activities included in this mechanism, thereby refining the standard view of what these maps represent. He does so by reasoning backward—exploring the functional properties of cells in other parts of the hippocampal formation that drive the activity of place cells in the hippocampus. Specifically, he appeals to two new entities: head-direction cells in the postsubiculum and grid cells in the entorhinal cortex. Large-amplitude oscillations, known as theta rhythms, are the activity at the heart of this mechanism. These rhythms allow cells throughout the hippocampal formation to maintain stable rates of firing for extended periods of time. As activation spreads throughout the mechanism—from head-direction cells to grid cells and then to place cells, as described below—the rhythm of the activated cells is altered. Hasselmo proposes that these relative changes in firing pattern serve a phase code via which the spatial and temporal features of the experience are encoded.

The encoding of an episode begins when the organism begins moving (e.g., when the rat begins exploring the environment). Head-direction cells are sensitive to this movement; each head-direction cell has a particular direction of movement toward which it is most responsive. This activation then spreads to the entorhinal cortex, disturbing the rhythm of the activated cells, causing them to go out of phase with other cells nearby. When these cells resynchronize, their joint firing produces a signal strong enough to cause a third cell to fire, marking a particular coordinate in the environment. This third cell is a grid cell. Theta rhythms govern these interactions, ensuring that these cells go in and out of phase at regular, evenly spaced intervals. The resulting pattern is (as their name suggests) a grid, with clusters of activity occurring in a periodic array across the space explored.

This process of grid generation is repeated several times throughout the entorhinal cortex. The resultant grids are not fully redundant, however. Each grid represents the environment at a scale larger than the last, made possible by the gradual change in theta rhythms across the entorhinal cortex. That is, as one moves from the dorsal to

ventral end of the entorhinal cortex, theta oscillations slow. The slower the oscillations, the longer it takes the cells to resynchronize. Although the grids generated by each rhythm are distinct, there will be places in the environment where their receptive fields overlap. Hasselmo proposes that whenever three or more grids intersect, the signal produced by their synchronized firing will be strong enough to activate a place cell in the hippocampus (p. 117). As the organism continues moving, these coordinated firings will recur, activating a series of place cells along its path. In this way, grid cells provide the coordinate system upon which place cell trajectories are built. According to Hasselmo, the hippocampus produces a series of cognitive maps, representing the environment in more or less detail, as a function of the scale imposed by the grid cells from which they are activated.

Most versions of the cognitive map theory have focused on cells in the dorsal hippocampus, where the receptive fields of place cells are relatively small. Correspondingly, cells in the ventral hippocampus have been thought to serve a different function. In opposition to this view, Hasselmo suggests that cells in the ventral hippocampus are also place cells, only with much larger receptive fields. Maps featuring place cells with the smallest receptive fields represent the organism's immediate surroundings in detail, whereas larger maps are produced by place cells with larger receptive fields, situating the experience within its (increasingly broad) spatial and temporal context.

The result is a mechanism that produces representations rich enough to support episodic remembering. The existence of multiple maps allows for a single episode to be recorded at several "scales of experience" (Hasselmo, 2008), capturing the episode as occurring not only at a particular place and time but as associated with various objects and events. The mechanism even captures the phenomenological features characteristic of episodic memory, offering insight into what it might mean to engage in mental time travel. Head-direction cells are sensitive to direction of movement, which Hasselmo suggests as a way to encode information about the organism's point of view and thus explain the perspectival feature of these memories (p. 67). And the use of phase coding in the entorhinal cortex provides a representation of space and time as continuous, preserving the flow of the experience and thereby accounting for the sensation that remembering an experience is somewhat akin to its reliving (p. 83).

3. Memory Mechanisms

Mechanistic explanations of memory are familiar territory for philosophers of science (e.g., Bechtel, 2008; Craver, 2007). The strength of Hasselmo's account lies in the degree to which it complements and extends this tradition. In this section, I pause to highlight some of the ways in which Hasselmo's proposal intersects with current debates over the nature and significance of mechanist explanation.

First, the mechanism of episodic memory that Hasselmo proposes is an information-processing mechanism. Memory for particular episodes is explained as

the encoding and retrieval of a mental representation of a spatiotemporal trajectory of events, built up out of more basic representations of location, context, and perspective (as provided by grid, place, and head-direction cells, respectively). Hasselmo's account thus echoes Bechtel's (2008, 2009) recent refrain that descent into mechanisms should not be seen as a replacement for—or reduction of—explanations in terms of mental representations and information-processing.

Second, by characterizing this mechanism as representational, Hasselmo is able to sidestep an entrenched debate over whether non-human animals possess episodic memory. Episodic remembering is often thought to require a particular type of awareness—what Tulving has called "autonoetic consciousness" (2002). Tulving believes that this form of awareness underwrites a particular type of self-knowledge that is distinctively human. Thus, he claims, "they [animals] too are capable of learning about and from experiences of the past, but without autonoetic awareness that they are doing so" (2002, p. 7). While some memory researchers agree with Tulving (e.g., Suddendorf & Corballis, 2007), others point to episodic behavior in animals as a suggestion that this form of remembering is more pervasive. There are, for example, food-caching scrub jays, who can remember what, where, and when they have hidden food (Aggleton & Pearce, 2002; Clayton, Griffiths, Emery, & Dickinson, 2002; Morris, 2002). Those on both sides of this debate have struggled to explain why certain animal behaviors are, or are not, sophisticated enough to warrant attribution of this form of conscious awareness. Hasselmo provides an intriguing work-around: start from the assumption that episodic memory requires a continuous representation of space, time, and context, then ask how the brain could build such a representation, and finally, attribute episodic memory to any organism that possesses the neural mechanisms that produce such representations.

Finally, Hasselmo demonstrates the complementary roles of molecular neuroscience and mathematical models in mechanism discovery. Mechanists (e.g., Kaplan & Craver, 2011; Piccinini & Craver, 2011) and Dynamicists (e.g., Chemero & Silberstein, 2008) disagree about the explanatory status of dynamical equations like those that Hasselmo uses to model episodic encoding and retrieval. Dynamicists maintain that these equations are (or at least can be) fundamental, whereas Mechanists claim that they are best understood as mechanism sketches, replete with black boxes that must someday be replaced with a detailed account of entities and activities.

Given the prevalence of modeling in Hasselmo's proposal, one might assume that it would serve as a case in point for the Dynamicist, but Hasselmo's loyalties are clearly mechanistic. He does not address the Mechanist-Dynamicist debate directly, but he is explicit about his "reductionist approach" from the outset: behavioral patterns must be mapped onto anatomical structures and cellular physiology (p. 9). This allegiance is displayed throughout the book. In building the model, he offers an explicit and detailed account of how each anatomical and functional feature of the mechanism is characterized mathematically. I will not detail these intricacies here. Put simply, the model works by representing neural population activity as a vector, and the connections between grid cell vectors and place cell vectors are updated via appeal to equations expressing Hebbian learning rules. The models are trained to reflect

behavioral data, derived mostly from rats. Figures throughout the book thus often display results from the model alongside experimental findings. Further, Hasselmo and colleagues pursue a series of additional experiments, meant to operate in parallel with the model, which provide support for the model's implementational details. For example, as discussed above, Hasselmo supposes that the function of grid cells is dependent upon theta rhythm oscillations in the entorhinal cortex. Hasselmo and colleagues have tested this assumption experimentally, showing how pharmacological disruption of theta rhythms interferes with grid cell creation in rats (Brandon et al., 2011). And finally, even though Hasselmo advances a particular view of how the phase coding of episodic memories is implemented—the persistent spiking model—he takes time to explore other ways that this phase code could be implemented, surveying the advantages and limitations of each (e.g., via oscillatory interference or attractor dynamics). This is easily read as a concession that the phenomena captured by the equations of his model are, at present, mechanism sketches.

4. Explaining Retrieval

Much of Hasselmo's effort in crafting this model of episodic memory has been focused on encoding—explaining how the brain creates and stores spatiotemporal trajectories of past experiences. Accounting for how we remember requires an explanation of retrieval as well. Explaining retrieval requires thinking of the episodic memory in question as one stored episode among many. It requires an account of how the brain manages to retrieve this particular episode rather than some other. Episodic memories may be encoded as spatiotemporal trajectories of events, but how does the brain know which such trajectory to select? Call this the *retrieval problem*.

Hasselmo does not address the retrieval problem directly. This concern about remembering has been (mostly) a preoccupation for philosophers, as I will explain below. Hasselmo does, however, provide an account of retrieval. In what follows, I introduce philosophical approaches to the retrieval problem and use this discussion to frame Hasselmo's proposal. Doing so makes it possible to see that his account of retrieval is unworkable, and also *why* it does not work. Attention to the retrieval problem reveals the limitations of not only Hasselmo's particular mechanism, but of the mechanistic approach to retrieval more generally.

4.1. The Retrieval Problem

Traditionally, the retrieval problem was posed as a conceptual challenge to theories of memory that appealed to memory traces to explain remembering (Heil, 1978; see also Bennett & Hacker, 2003). The problem is that appeal to stored representations of past experiences (i.e., memory traces) appears to presuppose remembering rather than explain it. If remembering always requires a trace, then the memory of how to locate the right trace must itself have a trace, one that explains when to search for the trace and how to find it. Explaining the retrieval of this second trace requires the

introduction of yet another memory, and in turn, another trace. And so on. By appealing to traces to explain retrieval, the trace theory of memory enters a regress. If, at any point, the trace theorist attempts to halt the regress by stipulating that some forms of remembering do not involve traces, then the initial motivation for inviting traces into an account of remembering is lost. Memory traces cannot explain how retrieval is possible because their role in explaining remembering is unsustainable.

Proponents of mechanistic accounts of memory have assumed that this challenge can be avoided if retrieval is characterized as an automatic process rather than an explicit search. On mechanistic accounts, remembering begins with a match between an environmental cue and a feature of the episode stored in memory. Characterizing retrieval in this way obviates the search through a set of stored trajectories, circumventing rather than solving the retrieval problem. Mechanistic solutions to the retrieval problem can be divided into two types, distinguished by the way that the match between the memory system and the cue occurs: those that construe memory as *index-addressable* and those that construe it as *content-addressable* (Bechtel & Abrahamsen, 2002). I explain each briefly below.

Index-addressable memory systems characterize memory as a filing system, one from which individual memories are accessed by a guided scan of the memory store. Much as one might fetch a book from the library by jotting down its call number and walking through the stacks to reach the desired location, so too index-addressable accounts characterize remembering as the act of scanning one's mental repository to arrive at a specified address. Index-addressable views work by supposing that each episode stored in memory has a unique call number, or address, to which it is indexed. The cue that initiates retrieval is either a restatement of this address or a cue tied to the address directly, ensuring that this "search" finds its target. Homage to this claim can be seen in historical accounts of memory, as well as in contemporary computational models of memory from cognitive science. Anderson's ACT-R model of cognitive processing, for example, operationalizes remembering as a scan through a mental database, following a set of pre-established if-then search rules (Anderson, 1976; Anderson, Bothell, Lebiere, & Matessa, 1998).

Content-addressable approaches, in contrast, characterize memory as a distributed network of features and retrieval as the activation of a particular pattern among a subset of the nodes in this network.[2] These networks do not contain nodes corresponding to particular episodes encoded. Instead, these networks contain a node corresponding to each basic idea that the subject has encountered and these nodes are connected to one another in ways that reflects various associations between these ideas. When a particular node is active, this activation spreads through the network via its connections to other nodes. The memory for a particular episode is thus distributed throughout the network, represented as a particular pattern of associations between various representations. Organizing memory in this way shapes the process of retrieval. There is no longer a unique address for each episode stored in memory; activation of an episode can begin from the activation of any of the nodes in its pattern. The memory system is thus *content-* (rather than *index-*) addressable. From this initial activation, retrieval of the full episode is a matter of pattern

completion—activation spreads from this first node to the other nodes via the pre-established associative connections.

While there are mechanistic accounts of memory as both index- and content-addressable, there is an emerging consensus that content-addressable systems are preferable (Bechtel & Abrahamsen, 2002; Michaelian, 2011; Sutton, 1998). This preference is based on considerations of descriptive adequacy: content-addressable systems are fault-tolerant and they offer multiple retrieval cues per episode. The latter feature is often thought to be particularly important. Index-addressable systems provide only a single, pre-established route to the retrieval of an item from memory. But, at least for many of the things we can remember, there are multiple ways to initiate retrieval. Recall of an episode can begin from consideration of its location, time, events, or significance. One could, for example, think about her eighth birthday party and recall that it was the year she received her first bicycle, or think about her first bicycle and recall that she received it on this birthday.

4.2. Hasselmo's Account of Retrieval

Hasselmo's approach to retrieval is driven by the same concerns that motivate the content-addressable approach. In Hasselmo's mechanism, retrieval involves a match between a feature of the environment and its representation in the hippocampal map; it begins with the automatic activation of place cells when the organism reenters a particular place or reencounters a particular object. Further, he stipulates that there are as many routes to the retrieval of an episode as there are features of the episode stored in the map. Any of these spatial, temporal, or contextual features can serve to initiate recall, allowing the rememberer to travel (mentally) along this trajectory in either direction (p. 121).

In short, Hasselmo's mechanism presents retrieval as encoding in reverse. The encoding of an episode begins with the activation of head-direction cells. This activity spreads to cells in the entorhinal cortex, creating grids and maps in turn. The co-activation of these cells strengthens the connections between them, establishing an activation pattern unique to this episode.[3] When the organism re-enters the environment, or encounters a cue (i.e., a feature of the remembered event) stored in one of the contextual maps, this activates the corresponding place cell. Retrieval is a matter of pattern completion from this initial activation, spreading to the other associated place cells, and then to grid and head-direction cells, respectively, via the connections established during encoding (p. 123).

In building this account of retrieval, Hasselmo runs into a problem. Place cells connect a particular location in the environment with a particular cognitive map. So long as retrieval is characterized as beginning with the activation of place cells, then it follows that the organism will only be able to remember one spatiotemporal trajectory associated with a given environmental feature—the trajectory encoded most recently (p. 130). Even worse, if the trajectory is self-crossing (i.e., if the organism's path across the environment involves revisiting particular places), then only the more recent activity at points of overlap will be retained. Such a model not only fails to capture the

behavior exhibited by maze-running rats, who are capable of remembering multiple overlapping routes, but appears unpromising for scaling up to capture human forms of episodic remembering as Hasselmo intends. When remembering, the same, or similar, retrieval cues can be used to retrieve distinct episodes. After all, we can remember multiple events that took place in the same location (e.g., various meetings held in the same office) or that involve the same contextual cues (e.g., making dinner).

Hasselmo attempts to solve this problem via the introduction of *arc length cells*: cells that encode the organism's overall trajectory across the episode. He uses the example of parking one's car in a parking lot on two consecutive days to explain how these cells work. The spatiotemporal trajectories for these episodes will have significant overlap: they involve the same location (the parking lot) and are of comparable duration. Each map will have a place cell whose receptive field corresponds to the lot's entrance. When this cell is activated, what determines which of the two maps will be activated? If the mechanism relies on place cells exclusively, the map for the second route through the lot will be activated. When this pattern was created, it replaced any previous associations with this place. The addition of arc length cells allows the two trajectories to be distinguished. The arc length cell for day 1 tracks the person's trajectory from the entrance to east side of the lot, say, while the arc length cell for day 2 tracks the trajectory from the entrance to a western parking spot. Adding these cells allows the person to retrieve multiple trajectories from the same entrance point, so long as each arc length cell is paired with a distinct cue (p. 135). If, for example, day 1 is paired with a blue-shirted parking lot attendant and day 2 is paired with a green-shirted one (or an altogether different attendant), then the person has an option as to which parking trajectory is retrieved when she returns to the lot entrance. Starting from the blue-shirted attendant cue prompts the trajectory leading to the east side of the lot; starting from the green-shirted one leads to the west.

Before we can evaluate this patch to Hasselmo's account of retrieval, we must ask a more basic question: do arc length cells exist? This is a serious question. Hasselmo's decision to include arc length cells in his mechanism is driven by the needs of his model rather than the neuroanatomical features of the system he is modeling. That is to say, arc length cells are "theoretical" (p. 130). They earn a spot in the mechanism because they provide a way to explain retrieval; there is no accompanying evidence that such cells exist.

This revelation may not strike the reader as problematic. "Arc length cells" may serve as a stand-in for a set of to-be-discovered entities and activities. Mechanistic explanations often involve such black boxes; they are what distinguish mechanism sketches from completed mechanisms.[4] So long as they are identified as such, perhaps there is no need to be concerned about them. The worry is not that Hasselmo has yet to identify the neural entity (or entities) that plays this role, but rather that the role is such that no entity could.

In previous work, Hasselmo has suggested that the discovery of "splitter cells" in the rat hippocampus can be used to support the stipulation of arc length cells (Hasselmo, 2007). Splitter cells are characterized by preferential firing for movement in one of two directions when rats are presented with a two-direction T-maze. These cells allow the rat to represent two routes through the same maze over a short period of time. At first

pass, appeal to splitter cells appears to bolster confidence in the existence of arc length cells. The task facing the rat who utilizes splitter cells parallels the task facing the driver who has parked twice in the same lot. But the appeal to splitter cells mischaracterizes, in an overly simplistic way, the role that arc length cells are asked to play.

Consider again what arc length cells are asked to do. These cells provide a blueprint for the episode in question, a recipe for reconstructing the right pattern among a given set of features. To succeed in disambiguating between overlapping episodes, they must be able to differentiate several episodes, each on multiple dimensions. The cells must do much more than split. A person can park in the same lot more than twice and park in other lots as well. What's more, the relevant feature for distinguishing among episodes will sometimes be direction, but in other cases it might be time of day, duration of the episode, or any of a number of contextual factors (e.g., whether he carpooled, which vehicle he drove, what he was wearing, etc.). Hasselmo allows that the cells used to guide retrieval could be time cells rather than arc length cells (i.e., cells that code for episode duration rather than trajectory; p. 130). The issue is not simply that there could be cells that could code for these features, but that the decision about which feature will be the relevant one for retrieval must be determined in advance. This code must be established during the episode and be connected to the episode's spatiotemporal trajectory via a predetermined cue (p. 135).

4.3. Retrieval Mechanisms

Arc length cells are also problematic as an approach to explaining retrieval more generally. To see this point, consider how these cells guide retrieval. The addition of arc length cells works by transforming Hasselmo's memory mechanism from a content-addressable system to an index-addressable one. Arc length cells succeed by restricting the number of routes into the memory of a particular episode. It is no longer the case that the memory can be accessed by any of the place cells in any of its corresponding maps. Instead, one must activate the arc length cell via a predetermined cue with which it is paired. Place cells, which appeared so prominent in Hasselmo's account earlier, have been demoted. Prior to the invocation of arc length cells, Hasselmo's model had the advantage of allowing an episodic memory to be triggered by reactivation of any of the place cells along its spatiotemporal trajectory. This was important because it reflected our ability to retrieve a memory in any number of ways, from a range of distinct cues. This kind of advantage is characteristic of content-addressable approaches to memory (Bechtel & Abrahamsen, 2002, p. 50). The appeal to arc length cells removes this advantage.

Should we accept that, despite previous appearances to the contrary, this transition reveals that it is preferable to characterize retrieval as index-addressable rather than content-addressable? No. By making the mechanism index-addressable, Hasselmo has swapped one set of difficulties for another. Arc length cells ensure that there is only one route by which each particular episode can be retrieved and that route must be determined in advance. Allowing for multiple routes into an episode looks ad hoc, and adds an additional explanatory burden. Stipulating the existence of cells that code for

episode-level features, whether one or many, does not answer the question of how we remember a particular episode, it merely postpones it. If we explain how a person retrieves parking trajectory 1 rather than parking trajectory 2 by appealing to the cues associated with each arc length cell, a new question arises: how does the person remember which trajectory is associated with which cue? Answering this question requires postulating an index where such cue-episode pairings are recorded. Having multiple, index-addressable cues would only require postulating more indexes (and then prompt further questions of how the person knows which index to consult).

Hasselmo's shift from a content-addressable system to an index-addressable one does not reflect a mere preference or quirk on his part. Rather, it is an adjustment demanded by the phenomenon Hasselmo is trying to explain. What Hasselmo's difficulty with the retrieval problem reveals is that neither of the two mechanistic accounts of retrieval is well equipped for explaining how we remember. Each suffers from serious limitations in descriptive adequacy. Index-addressable accounts guarantee access to a given memory, but do so by limiting the means of access to a particular, pre-determined route. Content-addressable accounts offer many more routes to retrieval, but fail to differentiate between two memories with overlapping features. What is needed is an account of retrieval that can explain our ability to retrieve distinct episodes with overlapping features from a range of cues or prompts. Given Hasselmo's (at least implicit) awareness of the conflict between these constraints on retrieval, there is reason to hope that Hasselmo's research will help lead the way toward a hybrid model. Either way, these difficulties reveal that the retrieval problem requires the attention of both philosophers and scientists who espouse memory mechanisms.

5. Conclusion

In raising these worries, I do not mean to detract from the significant achievements reflected in this book. Hasselmo has done a great deal to advance our understanding of how the brain encodes memories of particular events, and the insights gained from his approach will no doubt continue for years to come. While his account struggles to explain retrieval, the retrieval problem is not new, nor is it Hasselmo's alone. The persistence of this problem should serve to remind us of why we find this capacity so intriguing and why research like Hasselmo's is so important.

Notes

[1]   The hippocampal formation includes the hippocampus itself, as well as the dentate gyrus, the entorhinal cortex, and other parahippocampal regions of the medial temporal lobe.
[2]   Content-addressable models of memory are often described in terms of a semantic network whose nodes are connected by principles of spreading activation (e.g., Collins & Loftus, 1975; Plaut, 1995; Ratcliff & McKoon, 1994).
[3]   Hasselmo does not offer much detail as to how these connections are formed. Instead, he relies on standard models of synaptic modification by long-term potentiation (pp. 50–54).
[4]   I am grateful to Daniel Burnston for raising this issue.

References

Aggleton, J. P., & Pearce, J. M. (2002). Neural systems underlying episodic memory: Insights from animal research. In A. Baddeley, J. P. Aggleton, & M. A. Conway (Eds.), *Episodic memory: New directions in research* (pp. 204–231). New York: Oxford University Press.

Anderson, J. R. (1976). *Language, memory, and thought*. Hillsdale, NJ: Erlbaum.

Anderson, J. R., Bothell, D., LeBiere, C., & Matessa, M. (1998). An integrated theory of list memory. *Journal of Memory and Language, 38*, 341–380.

Bechtel, W. (2008). *Mental mechanisms: Philosophical perspectives on cognitive neuroscience*. New York: Lawrence Erlbaum.

Bechtel, W. (2009). Looking down, around, and up: Mechanistic explanation in psychology. *Philosophical Psychology, 22*, 543–564.

Bechtel, W., & Abrahamsen, A. (2002). *Connectionism and the mind: Parallel processing, dynamics, and evolution in networks* (2nd ed.). Malden, MA: Blackwell.

Bennett, M., & Hacker, P. M. S. (2003). *Philosophical foundations of neuroscience*. Oxford: Basil Blackwell.

Brandon, M. P., Bogaard, A. R., Libby, C. P., Connerney, M. A., Gupta, K., & Hasselmo, M. E. (2011). Reduction of theta rhythm dissociates grid cell spatial periodicity from directional tuning. *Science, 332*, 595–599.

Chemero, A., & Silberstein, M. (2008). After the philosophy of mind: Replacing scholasticism with science. *Philosophy of Science, 75*, 1–27.

Clayton, N. S., Griffiths, D. P., Emery, N. J., & Dickenson, A. (2002). Elements of episodic-like memory in animals. In A. Baddeley, J. P. Aggleton, & M. A. Conway (Eds.), *Episodic memory: New directions in research* (pp. 232–248). New York: Oxford University Press.

Collins, A. M., & Loftus, E. F. (1975). A spreading activation theory of semantic processing. *Psychological Review, 82*, 407–428.

Craver, C. F. (2007). *Explaining the brain: Mechanisms and the mosaic unity of neuroscience*. Oxford: Clarendon.

Hasselmo, M. E. (2007). Arc length coding by interference of theta frequency oscillations may underlie context-dependent hippocampal unit data and episodic memory function. *Learning & Memory, 14*, 782–794.

Hasselmo, M. E. (2008). Neuroscience: The scale of experience. *Science, 321*, 46–47.

Heil, J. (1978). Traces of things past. *Philosophy of Science, 45*, 60–72.

Kaplan, D., & Craver, C. F. (2011). The explanatory force of dynamical and mathematical models in neuroscience: A mechanistic perspective. *Philosophy of Science, 78*, 601–627.

Maguire, E. A., Frackowiak, R. S. J., & Frith, C. D. (1997). Recalling routes around London: Activation of the right hippocampus in taxi drivers. *Journal of Neuroscience, 17*, 7103–7110.

Michaelian, K. (2011). Is memory a natural kind? *Memory Studies, 4*, 170–189.

Morris, R. G. M. (2002). Episodic-Like memory in animals: Psychological criteria, neural mechanisms, and the value of episodic-like tasks to investigate animal models of neurodegenerative disease. In A. Baddeley, J. P. Aggleton, & M. A. Conway (Eds.), *Episodic memory: New directions in research* (pp. 181–203). New York: Oxford University Press.

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

Piccinini, G., & Craver, C. F. (2011). Integrating psychology and neuroscience: Functional analyses as mechanism sketches. *Synthese, 183*, 283–311.

Plaut, D. C. (1995). Semantic and associative priming in a distributed attractor network. *Proceedings of the 17th Annual Conference of the Cognitive Science Society* (pp. 37–42). Mahwah, NJ: Lawrence Erlbaum.

Ratcliff, R., & McKoon, G. (1994). Retrieving information from memory: Spreading activation theories versus compound cue theories. *Psychological Review, 101*, 177–184.

Roediger, H. L. (2000). Why retrieval is the key process to understanding human memory. In E. Tulving (Ed.), *Memory, consciousness and the brain: The Tallinn conference* (pp. 52–75). Philadelphia, PA: Psychology Press.

Suddendorf, T., & Corballis, M. C. (2007). The evolution of foresight: What is mental time travel and is it unique to humans? *Behavioral and Brain Sciences*, 30, 299–351.

Sutton, J. (1998). *Philosophy and memory traces: Descartes to connectionism*. Cambridge: Cambridge University Press.

Tulving, E. (2002). Episodic memory: From mind to brain. *Annual Review of Psychology*, 53, 1–25.