

8-1-2017

# Episodic Memory via Spans and Cospans: A Hierarchy of Spatiotemporal Colimits

Michael John Healy MS  
mjhealy@unm.edu

Thomas Preston Caudell PhD  
tcaudell@unm.edu

Follow this and additional works at: [https://digitalrepository.unm.edu/ece\\_rpts](https://digitalrepository.unm.edu/ece_rpts)



Part of the [Electrical and Computer Engineering Commons](#)

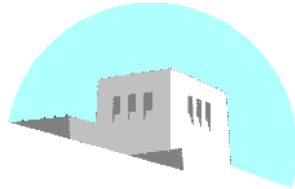
---

## Recommended Citation

Healy, Michael John MS and Thomas Preston Caudell PhD. "Episodic Memory via Spans and Cospans: A Hierarchy of Spatiotemporal Colimits." (2017). [https://digitalrepository.unm.edu/ece\\_rpts/52](https://digitalrepository.unm.edu/ece_rpts/52)

This Technical Report is brought to you for free and open access by the Engineering Publications at UNM Digital Repository. It has been accepted for inclusion in Electrical & Computer Engineering Technical Reports by an authorized administrator of UNM Digital Repository. For more information, please contact [disc@unm.edu](mailto:disc@unm.edu).

DEPARTMENT OF ELECTRICAL AND  
COMPUTER ENGINEERING



SCHOOL OF ENGINEERING  
UNIVERSITY OF NEW MEXICO

**Episodic Memory via Spans and Cospans: A Hierarchy of  
Spatiotemporal Colimits**

Michael J. Healy

e-mail:mjhealy@unm.edu

Thomas P. Caudell

e-mail:tcaudell@unm.edu

Department of Electrical and Computer Engineering  
University of New Mexico  
Albuquerque, New Mexico 87131, USA

UNM Technical Report: ECE-TR-17-0001

Report Date: August 1, 2017

## **Abstract**

We introduce a category-theoretic account of episodic memory as an outgrowth of an existing mathematical theory of the semantics of neural networks. We propose that neural systems which can be said to have episodic memory represent sequences of events and their associated information within a hierarchy of concepts, represented in their neural networks. In the categorical model presented here, the hierarchy is based upon colimits. Colimits “put everything together” mathematically, and appear throughout many categories. The event-sequence colimits can be visualized as assemblies of categorical structures known as spans and cospans. A string of cospans formalizes a hierarchy of overlapping episode segments, with the segments increasing in length by adding a next event as an episode progresses. The concept category can be mapped into a category that expresses the structure and activity of a neural architecture. An episodic sequence is formalized as a string of cospans of its overlapping episodic segments. This kind of neural structure supports the tracing of its event sequence in either the forward or reverse direction during recall, but it also does much more: It allows a holistic access to an episode or entire segments of the episode, it maintains the continuity of that information which is preserved between successive events, and, finally, the cospan cells serve as explicit representatives of the temporal order of events, making a sequence available not only for recall but also for direct access to subsequences of greatest interest. We end with a preliminary sketch of the application of this episodic memory model to understanding the interaction of the hippocampus with other structures of the mammalian medial temporal lobe.

## **Keywords**

category, colimit, concept, cospan, hippocampus, episode, event, neural, span

# 1 Introduction

In the work described here, we develop a mathematical theory of episodic memory which we call the span-cospan model. This theory is part of a larger mathematical theory of concept representation and learning in neural networks which we call the categorical neural semantic theory (CNST). The CNST is based upon a branch of mathematics known as category theory. It is applied here to explain in precise terms the structure and operation of a network which learns and re-enacts temporal sequences of events.

The explanation is based upon four key properties which we propose an episodic memory system must have. For ease of labelling, we shall refer to our proposal that property X is necessary, or key, in an episodic memory system as “Notion X”. The first and most obvious key notion (Notion 1) is that the neural structure must maintain the ability to re-enact the events making up an episode in the order in which they happened when the episode was first stored in memory; in particular, this aids in having a sense of the flow of time. Notion 2 is that the neural structure must maintain access to an episode as a holistic entity. That is, it must enable the episode to be called up as a single unit, not merely accessed a single event at a time. Notion 3 is that the events in an episodic sequence must contain a significant amount of shared information and this sharing must be explicitly represented in the episodic neural structure. And, finally, Notion 4 concerns the flexibility in re-use of an episodic memory. One aspect of flexibility is the ability to re-enact a temporal memory sequence either forward or reversed in time and enter the sequence at more than one point in time. Another aspect is the ability to extract significant information from an episode concerning the objects and situations which appear in it. One aspect of this is that the subsequence of events immediately preceding a particular event such as an encounter with a predator must be immediately accessible, given that the event comprising the encounter itself is active in memory. Another aspect is that information which is “stationary”, that is, common to more than one episode, must be retained. An example is the retention of knowledge about how to ride a bicycle, given the several episodes of past lessons in learning to ride. And finally, there must be neural cells which express in explicit form the event sequence order. Note that this property is not the same as the ability to *re-enact* the sequence in order (Notion 1); instead, it requires declarative representations that *state* the order, say, by expressing which event in a segment of an episodic sequence occurs in the final position. There will be some overlap in the details of these four notions. Repeating for emphasis, we shall refer to these four properties—sequence order, holistic access, the representation of shared information, and flexibility in re-use, as “Notion 1”, “Notion 2”, “Notion 3”, and “Notion 4”, respectively.

The distinction between episodic and other forms of memory was first identified by Endel Tulving [36]. In humans, episodic sequences in memory constitute an autobiographical record of an individual’s life. In nonhumans, they might be useful in capturing important information about the actions of predators, access to food sources, and other information gleaned from experience. Significant information in an episodic memory ap-

parently can be transferred to semantic memory. This is the memory of objects and situations that omits most of the details of the episodic sequences from which the information was initially captured by the memory system. Together, episodic and semantic memory make up declarative memory. Declarative memory is so called because in humans it enables recall, the consciously-accessible description of things experienced in one's past. Since consciousness is at present ill-defined in non-humans, declarative memory is studied in animal models by exploiting its other characteristics. Declarative memories can be acquired based upon a single experience. By contrast, the acquisition of an implicit memory—an acquired habit pattern or a learned skill—has a relatively slow time course, requiring the repetition of similar behaviors over time for its encoding. Another characteristic dissociating declarative and implicit memory is that declarative memories are re-usable and flexible, subject to incorporation in future memories through imagination, the construction of plans, or other ways in which the recall of declarative memories can be used as components in the formation of new ones. Implicit memories are of single use and are, hence, inflexible. Semantic and declarative memory are sometimes regarded as one and the same, but herein we shall regard both the semantic and episodic functions as two divisions of declarative memory.

The distinction between episodic and semantic memory, and their interactions, suggests the existence of subnetworks in the brain effecting two interacting functions within declarative memory. In humans, and also in other primates as well as rats, the hippocampus and the adjoining brain structures in the medial temporal lobe (MTL) seem to be essential to the declarative memory system. Representative work on episodic memory and its neural substrate in the MTL can be found in many sources, some of which are cited here—examples include ([37, 1, 25, 28]). Implicit memory appears to be supported mainly within neocortex, where long-term memory acquisition requires repeated episodes of habituation or practice. Over time, the rapidly-acquired declarative information also appears to become consolidated within neocortex but is thought to continue to require hippocampal involvement for recall and reconsolidation.

The point of this paper is to introduce an episodic memory model based upon the CNST. This model is one of events and their incorporation in episodes. We discuss events, episodes, and the sensor-based and other information that forms the content of the events in terms of concept representation, a way to think about the semantics of neural networks. Certain issues other than concept representation that concern neural networks are discussed often in the neural network and neuroscience literature. We do not address those issues here. Specifically, we do not discuss dynamic systems, statistical learning, probabilistic inference, or other issues which are not essential to a discussion of concept representation. We also note that there is an intimate relationship between semantic and episodic memory, wherein semantic objects are formed from repeated occurrences in episodes and retained for later re-use in forming new episodic memories (note that this is one of the aspects of Notion 4). In order to avoid an overly lengthy exposition in introducing this episodic model, we do not discuss in any detail the relationship between semantic and episodic memory.

In the remainder of this paper, we shall posit an account of episodic memory based upon the aforementioned mathematical theory, the CNST. In previous work ([13, 19]), we presented an initial effort in developing a theoretical model of episodic memory; although partially incorporated in the present model, there are significant new developments. The new model is introduced in Section 6. It provides a framework which we claim more closely satisfies the criteria presented in the first paragraph. A proposed biological implementation of the new model is discussed in Section 7. The categorical background theory and its application in the CNST is discussed in the Appendix and more extensively in the literature, including ([22, 24, 14, 23]). There are other category-theoretic models of neural structure and processing, each taking a unique approach (see for example [9, 12]).

## 2 How Can a Neural Network Represent an Episode?

### 2.1 A putative neural model

Let us begin with a hypothetical neural network architecture for the storage and recall of an episodic memory which maintains the event sequence order. For this example the episode consists of the three events  $e_1$ ,  $e_2$ ,  $e_3$  illustrated in Fig. 1 (a), discrete instances of visual stimuli. The pictures illustrate the events, but mathematically, the events are descriptions in a formal language of the items pictured (think of a very restricted form of language in which the definitions of terms are unambiguous). We refer to these formal descriptions as *concepts*, the same meaning of the word “concepts” used in the Introduction. Notice that as the pictures illustrate, each event is made up of individual objects and their relationships (e.g., a human figure throwing a ball); thus, the concepts have information content, and this content originates either at the sensor level or through the autonomous (creative) construction from the available sensor information. The events occur in a temporal sequence, the steps in a human figure throwing a ball, a bowling ball for example. Questions such as the “smoothness” of motion are not important here; to avoid picturing a “jerky” motion proceeding in abrupt transitions,  $e_1$  to  $e_2$  and then  $e_2$  to  $e_3$ , imagine that the time delay in the perception of closely-spaced events is such that the superposition of  $e_1$  and  $e_2$ , then  $e_2$  and  $e_3$ , allows the brain to provide the illusion of motion flowing along a continuum. This effect is of course the basis for cinema, where a succession of pictures depicting incremental changes in spatial configuration of a set of objects is projected on a screen rapidly enough to allow humans to perceive smooth motion. We shall call this initial episodic neural network model the linear-event-string architecture.

A putative neural network for the storage in memory and recall of the episode is shown in part (b) of Fig. 1. Arguably the simplest idea for an episodic network, it consists of a string of neural cells connected in feedforward fashion through synaptically weighted excitatory connections. Each cell represents the corresponding event in the time sequence of the memory; for example,  $p_1$  represents  $e_1$ . During recall cell  $p_1$ , once active, sends a

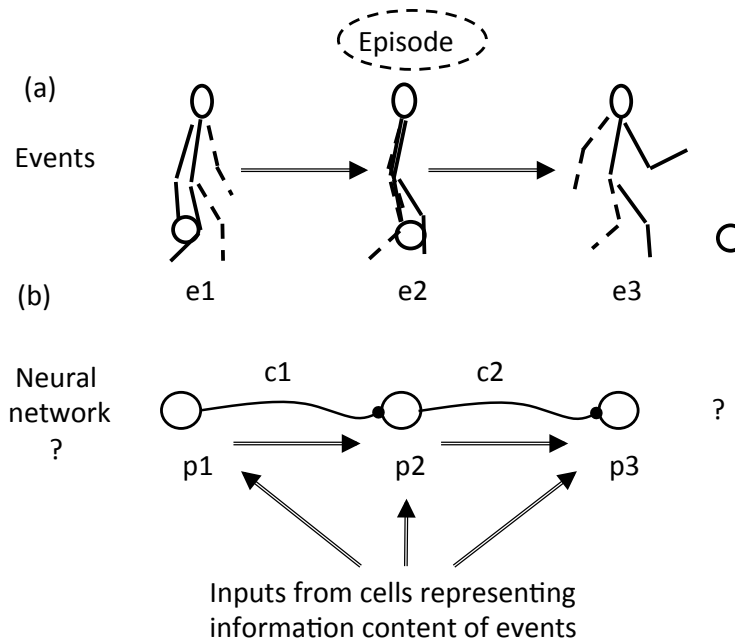


Figure 1: (a) An episodic memory shown as an event stream,  $e_1$ ,  $e_2$ ,  $e_3$ . (b) A hypothetical neural network whose cells represent the events  $e_1$ ,  $e_2$  and  $e_3$ , activated in the order  $p_1$ ,  $p_2$ ,  $p_3$  through the synaptic connections, first  $c_1$ , then  $c_2$ . The lobes at the ends of the connections are excitatory synapses. Notice that the overall network separately represents items making up the content of the events, which are represented through feedforward input.

signal to  $p_2$ , then  $p_2$  sends to  $p_3$ , evoking the events  $e_1$ ,  $e_2$ ,  $e_3$  in that order. Thus, the cell  $p_1$  represents a human figure starting to throw a ball,  $p_2$  represents a human, presumably the same one, at an intermediate stage of a throw, and  $p_3$  represents the human releasing the ball.

We say that the event concepts  $e_1$ ,  $e_2$ , and  $e_3$  illustrated in Fig. 1 (a) describe in unambiguous terms the semantics, or meaning, of the neural cells  $p_1$ ,  $p_2$  and  $p_3$  in Fig. 1 (b). For convenience and to relate to the terminology used by some neuroscientists, we say the the cells  $p_1$ ,  $p_2$  and  $p_3$  *represent* the concepts  $e_1$ ,  $e_2$  and  $e_3$ , respectively. We would like to say that the arrows between the event concepts are also descriptions and are represented by the connections  $c_1$  and  $c_2$  between the corresponding cells. But is any of this correct? For example, how does a cell, which expresses only a level of activation potential, represent a concept in all its complexity? The answer is that none of this is that simple. Neural cells represent concepts only in that they become active when the concepts in question are relevant in the current activity of the neural network. This relevance is brought about by the presence of input connections to the event cells emanating from other neural cells which

represent the information content of the events, that is, their semantics. These connections convey activity from active cells that either sense items in the sensory environment, pre-process the information that is sensed, or supply information that the network concocts autonomously based upon processes operating in other parts of the network. This will be discussed further in the following sections. For the moment, notice the arrows pointing upward from the bottom in Fig. 1 (b), illustrating the presence of these “lower-level” inputs to the event cells.

## 2.2 Deconstructing the neural model

Can this be a viable neural model of memory? Tempting as the parsimony of the model is it is well to ask, for the model has shortcomings. Consider the four key notions we have posited for the structure and operation of an episodic memory neural network.

First of all (Notion1), the neural network shown in Fig. 1 (a) does maintain the order in which the events occurred; the feedforward connections transmit a stimulus from neural cell p1, representing event e1, to p2, representing event e2, and so on along the string. Supposedly, the synapses, or connection weights, in these connections are strong enough to allow each cell to activate its successor, allowing re-enactment of the sequence in its correct order.

However, the second property (Notion 2) is not maintained. The network does not allow access to the episode either in whole or in part as a holistic memory, as would be desired in recalling a familiar route to reach a goal. The network allows access to only one event at a time, consecutively, in temporal order. Grossberg in his theory of human memory [11] gives the example of deciding to take a familiar route from his office to the cafeteria. Suppose cell p3 represents the event of reaching the cafeteria, the “goal” of the episode. If one tries to explain to another person encountered on the way to the cafeteria the reason for following the route, he cannot. This is because the reason, the “goal” of reaching the cafeteria, cannot be retrieved from memory until either he has reached it, or has retrieved and then described each intermediate event in order along the route he plans to follow to reach it.

Third (Notion 3), in order for an event sequence to constitute an episodic memory, the events must be intimately related. This means that they must share a great deal of information concerning one or more objects and their relationships that are present in successive parts of the episode: they cannot be a sequence of situations in which each event comprises a unique collection of objects and relationships and is, hence, unrelated to the other events. In each event of the episode in Fig. 1 (a), the human figure is configured slightly differently for that particular stage of the throw, and the ball is at a different stage of being thrown. However, even though the configurations of the two objects are somewhat changed in each successive event, we understand immediately that the human and the ball are the same two entities throughout the sequence. In this way, the events *cohere* in



memory as a sequence constituting a single episode. But the putative episodic network as shown does not contain an explicit representation of this information. Yet, we propose that the network must have an explicit representation of the shared information in order for an organism—or a neural-network-driven machine—to exhibit behavior consistent with the coherence of the events in an episode.

And, finally (Notion 4), the neural network in Fig. 1 does not provide the flexibility that is supposed to be a property of declarative memory (for example, see [1], esp. sections 2 and 3). The fact that it does not allow access to the entire episode as a single entity is part of the reason for this. In order to re-use the episodic memory in different contexts or to extract information from it such as how one throws a bowling ball, it must be possible to apprehend either the episode as a whole. Another reason for the lack of flexibility is closely related to the foregoing: The linear connectivity of the string directed from  $p_1$  to  $p_3$  dictates that activating any given cell,  $p_1$  for example, results in a spreading activation forward along the string to  $p_2$ , then, without pause,  $p_3$ . There is no representation of an episodic segment, say  $e_1e_2$ , and no way to parse it, to observe that the event we have labelled “ $e_2$ ” is the immediate successor to the one labelled “ $e_1$ ”. Reciprocals to the connections along the string would allow discovering  $e_1$  through retroactivation beginning with  $e_2$  given a stimulus to  $e_2$  through an afferent connection from outside the string. But if instead the task were discovering that  $e_2$  precedes  $e_3$ , a stimulus to  $e_3$  would generate a spreading activation executing the entire sequence in reverse. Recalling  $e_2$  as the predecessor to  $e_3$  could not be done in isolation.

It would be desirable, for example, to be able to manipulate the memory so that when the goal,  $e_3$  say, comes to mind, a familiar route to it expressed by the preceding partial string  $e_1e_2$  could immediately be retrieved as a unit without requiring that  $p_1$ , the cell representing its first constituent event  $e_1$ , be active beforehand. Imagine an animal whose memory has stored some representation of an episode resulting in its finding food, or else encountering a predator but surviving the encounter. In either case, it would be advantageous to be able to associate  $e_3$ , the goal (finding food, or encountering a predator), with the sequence of events, such as a route taken through the environment, that led to  $e_3$ . It would also be advantageous to activate any of the cells  $p_1$ ,  $p_2$  or  $p_3$  as a potential goal without activating others, to contemplate one event at a time while maintaining the option to transition to its immediate predecessor or successor. An explicit representation of the predecessor-successor relationship would be advantageous here. Of course, feedforward inputs from cells representing all of the objects making up the event could provide a stimulus sufficient to activate its cell. However, as pointed out in the preceding paragraph with the present network this activation would propagate through the remainder of the episode due to the linear network structure. The presence of reciprocal connections would cause a confusing parallel activation in the reverse direction as well. Of course, adding to this network strong inhibitory connections from other cells could keep all but the desired event cell from being active. In any case, more cells and connections would be required to achieve the desired functionality.

There are additional aspects of flexibility. Inferential memory judgements and reasoning with transitivity, such as the fact that with numbers  $A > B$  and  $B > C$  implies  $A > C$ , require flexibility in the use of memory items [10]. Some authors have other perspectives on this subject and have presented illustrative results [1]. To conclude, the point made in this section is that simply connecting event cells in a feedforward string will not serve as an episodic memory structure.

### 2.3 Adding detail to the events

Although the putative neural episodic model does not allow access to it as a holistic entity (Notion 2), each of its individual events as shown in Fig. 1 *would be* a holistic entity were its information content not supplied through the bottom-up feedforward input at the bottom of the Figure. The absence of this information would be a serious drawback. As Barsalou points out in [3], a representation that is entirely of a holistic nature does not allow the represented entity to be accessible in any detail. That is, it does not allow the expression of its semantics. An analogous situation would occur if one could not discern any of the details in a snapshot of a person throwing a stick for a dog. In such a case, the snapshot would not make sense. This is because neither the person, the stick, the dog, the posture of the person in the act of throwing the stick, the dog jumping into the air after the stick, the grass making up the lawn, a house in the background—none of these things would be perceived even though they appeared in the picture, since they would have no separate representation in the neural system. It must be possible to understand each item separately and in relation to the others to understand what the snapshot is about, and that requires neural cells to represent these items separately as well as together. Hence forth, let us refer to the objects and relations in an event collectively as “objects”.

Ensuring that the events are more than merely holistically represented is easily remedied by adding to the network the cells representing the objects of which they are made. As part of this, each event cell is the target of feedforward connections from those cells. Let us add feedback connections from the event cells to those cells as well. In our previous papers, we have discussed the importance of this feedback. It guarantees that all of the cells whose representations make up the event are active, ensuring that they are accessible along with the event cell itself. Similarly, the object cells are joined to their component stimulus representations by feedforward/feedback representations. Given the activity in all these cells, each event, and hence the episode, can be understood as a combination of its constituents<sup>1</sup>. Referring back to the simple example of Fig. 1, objects labelled “Human figure” and “Ball” are shown with arrows pointing to event  $e_3$  in part (a) of Fig. 2. We call these objects *grounding objects* because they “ground” the event by allowing access to the

---

<sup>1</sup>The constituents of an event must be understood to include relationships between the objects as well as the objects themselves. For simplicity, specific representations of object relationships are omitted from this discussion, with the exception of the “blending” relationships expressed in categorical colimits, discussed in the Appendix.

sensor-derived information that is its semantic content. In this fashion, the separate items and relationships, the grounding objects, comprise the information content of an event. Correspondingly, feedforward/feedback connections from the cells representing these two objects are shown as reciprocal pairs directed to/from event cell p3 in part (b) of Fig. 2. These connections are strengthened when the event is first constructed in memory and aid in the reactivation of the event cell when sufficient activity occurs in its constituent representations. Strong weights in the feedback connections ensure that its constituent object cells are active as well when p3 is active, representing the details of the event during recall. Similar remarks apply to connections joining the object cells to the cells closer to the sensor level representing their component stimuli. Fig. 3 illustrates the grounding objects and events and the corresponding cells and connections from Fig. 2 with arrows indicating that additional inputs are present. Notice that some of the connections emanating from the bottom of Fig. 3 bypass the human-figure- and ball-object cells and synapse directly upon the event cell p3. These connections aid in the activity of commutative diagrams in a neural category, yet to be discussed (see the Appendix, and also [23]).

Let us refer to the neural network regions that represent the grounding objects in support of event representations as *grounding regions*. The processing in these regions constructs and maintains the representations of the objects, behaviors and spatiotemporal relationships that “ground” the neural network by expressing its semantics in full detail. These representations are then used by the various mechanisms for memory such as the construction of the events.

## 2.4 The representation of grounding objects versus Notion 3: shared information

Given this, the episodic network does have the needed representations of the objects in its events. This allows the information content of the events to be made explicit, a requirement for declarative memory. As discussed in Section 5, the representation of the objects includes information that is shared between events. Because of this, the putative episodic network might seem to have all of the necessary architecture for representing the sharing of the information. After all, the grounding object cells represent shared information as well as the information unique to each event. Does this not answer the third requirement of an episodic neural structure, that the continuity of information over successive events be represented? As will be discussed in sections 4 and 5 and is explained in more detail in the Appendix, the inclusion in the network of the shared representations among the grounding cells does not in and of itself constitute a representation of the *sharing* of the information. However, a remedy for this exists in the mathematics of category theory: the presence of *commutative diagrams*, discussed in Section 5.

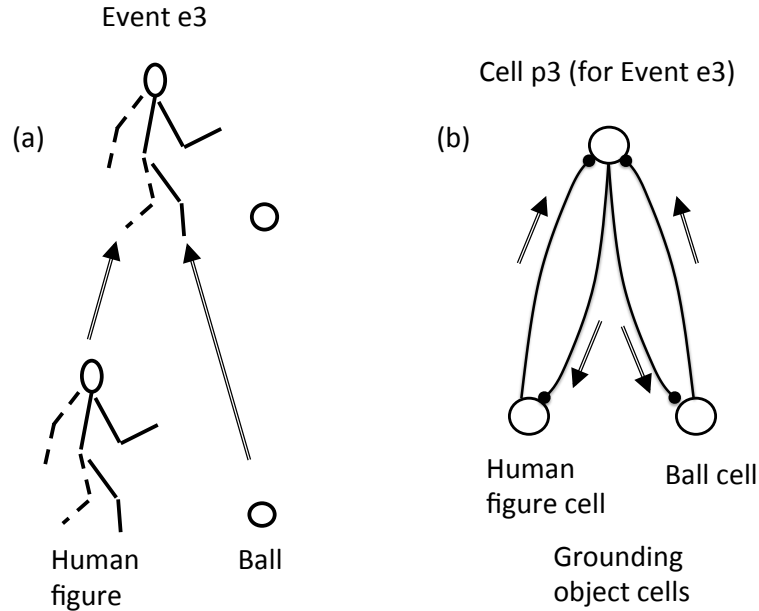


Figure 2: (a) An event and its grounding objects, or stimuli. (b) Feed forward/feedback connections between an event-representing cell and cells representing its individual objects.

## 2.5 Attempting to address Notion 4: flexibility

We can also add to the putative episodic network with extrinsic connections to allow some degree of flexibility. A modified episodic network is shown in Fig. 4 in which all feedforward connections have reciprocals. Also, the event string has been generalized to include an arbitrary number of events, with a focus upon event  $e(n)$  and its immediately adjacent events. As we proceed with this discussion, we shall switch to mathematical symbols with subscripts where needed, such as  $e_n$  in place of  $e(n)$ . Here the full string, which informally we would label as  $e_1e_2e_3 \dots e_{n-2}e_{n-1}e(n)e(n+1) \dots$ , we shall now denote as  $e_1e_2e_3 \dots e_{n-2}e_{n-1}e_n e_{n+1} \dots$ .

To address flexibility, extrinsic connections from other neural network regions facilitate recall in both the forward and backward directions beginning with an arbitrary event in the episodic sequence. To illustrate forward recall (see the top half of Fig.4), a cell which we shall call a “selector cell” provides connections efferent to the cells for events  $e_n$  and  $e_{n+1}$ . The connection to the  $e_{n+1}$  cell is excitatory, initiating the recall of that event. Were that the only connection from the selector cell, a confusing mix of spreading activation would occur, forward through the cells for  $e_{n+1}, e_{n+2}, \dots$  but also backward through the cells for  $e_{n+1}, e_n, e_{n-1}, e_{n-2}, \dots$ . However, the connection to the  $e_n$  cell is

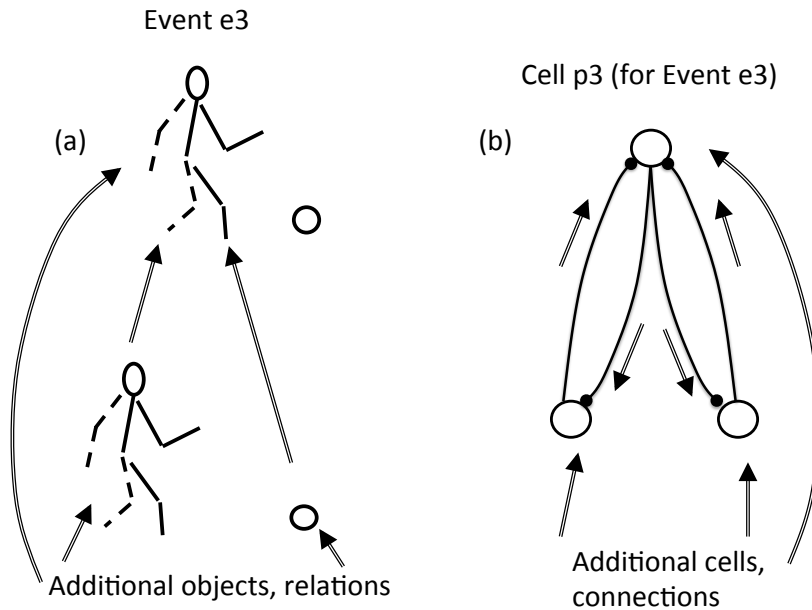


Figure 3: (a) There are additional grounding objects from which events and their grounding objects are formed. Normally, only those necessary for illustration will be shown. (b) Arrows indicate the neural network input connections corresponding to these additional grounding objects.

inhibitory, with synaptic strength sufficient to block the activity in the  $e_n$  cell and, hence, preventing the propagation of activity in the reverse direction. Thus, given the appropriate selector cells and connections, forward recall can be enacted without the confusion caused by a simultaneous backward flow of recall. Similar remarks apply to backward recall using the extrinsic additions to the architecture (see the bottom half of Fig.4). The excitatory afferent to the  $e_{n-1}$  event cell stimulates recall activity beginning with its activation and proceeding backward through the reciprocal connections between the cells for  $e_{n-1}, e_{n-2}, \dots$ , while the inhibitory connection to the  $e_n$  event cell blocks its activation, thereby preventing a confusion of recall that would otherwise occur simultaneously in both directions.

The two forms of recall enabled by selector cells with the appropriate connectivity can support both recall through the re-enactment of an event sequence beginning at an arbitrary event and a sort of retrospective recall as well. But two questions arise. First, where do these selector cells reside and how are *they* selected? And how many of these selector cells are required: Is it necessary to have specialized selector cells of the two different functionalities, forward and backward recall, accompanying each event cell in an episodic string? It would seem likely that more than one selector cell would be involved

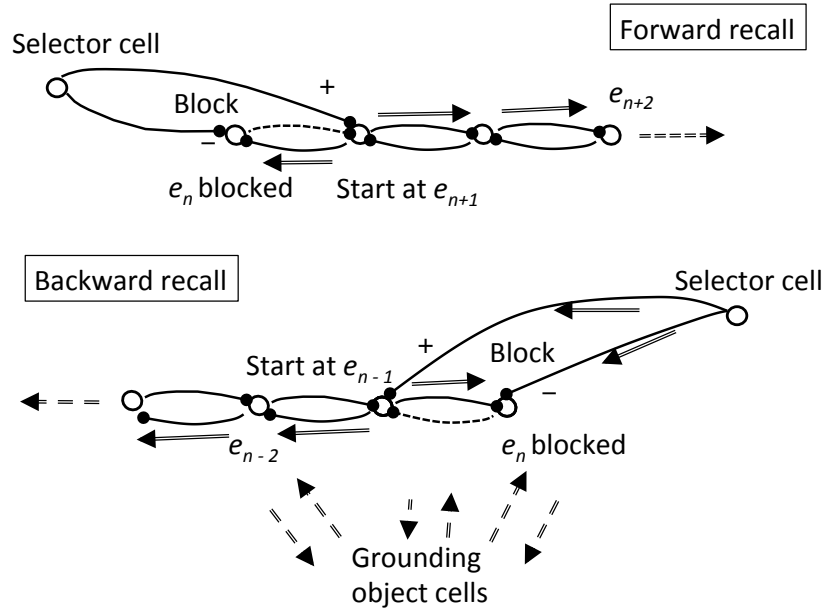


Figure 4: Extrinsic connections enable both forward and backward recall. The excitatory afferent connection to the  $e_{n+1}$  cell for forward recall, and to the  $e_{n-1}$  cell for backward recall, is labelled “+” and the inhibitory afferent to the  $e_n$  cell in either case is labelled “-”. Note: To economize on space, the grounding objects and the bottom-up arrows are shown only once for both the forward and backward recall illustrations.

for a given selection, one providing the excitatory stimulus and the inhibitory blocking as illustrated in Fig.4 and the other providing an inhibitory blockage at a different location to limit the length of the spreading activation along the selected episodic segment in the chosen direction. Also, are the selector cells based at a “higher” level of processing in the overall neural network, or do they reside at a “lower” level, or both, the level being determined in a situation-dependent manner? For example, the ability to recall solely a short subsequence of events immediately preceding an event selected as a goal, such as an awareness of a food source, would be advantageous. Other desirable properties that have been mentioned are the ability to make inferential memory judgements and to reason with numbers by applying properties such as transitivity. In any such cases, event sequences must be enacted within limits upon the flow of memory. Evidently, flexibility in making use of a stored episodic memory sequence requires a rather more complex architecture than the simple linear event sequence shown in Fig. 1.

Second, how is the directionality of recall itself made explicit in the semantics of the neural network? Suppose that recall in the reverse direction of a memory sequence is to

be initiated beginning with some event selected as the current goal. Of course, recall can occur in that direction with our current architecture, event by event through the feedback connections in the string. But what feature of the neural network distinguishes the recall operation going back in time from the recall operation going forward in time? Somewhere in the neural network there must exist the machinery to initiate a “go backward” command through the connections to the episodic memory event sequence discussed in the preceding paragraph. But this requires a representation that *explicitly* expresses an event as a successor to the preceding event in the sequence as opposed to the *implicit* information about the predecessor/successor relation embodied in the reciprocal pathways in the event cell string. An explicit representation of the predecessor/successor relation requires additional neural cells with the appropriate network connectivity. In fact, for full flexibility in re-enacting the sequence at any point either forward or backward in time, a representation of the predecessor/successor relationship must exist for every event and its predecessor. That is, every event cell must have connection pathways shared with additional cells that represent both that event and its immediate predecessor (and successor) cells. Otherwise, how would the neural network express the semantics of “going backward (or forward) in time” so that it could actually command the episodic architecture to initiate the operation? And in humans, how would it be possible for this information to enter one’s awareness?

Some would argue that “causality” explains the representation of recall direction. That is, the occurrence of event  $e_n$  depends upon the immediate past occurrence of event  $e_{n-1}$ . But saying that the semantics of causality are automatically explicit through the dependence of one event upon another is like saying, “The neural network knows full well that event  $e_n$  cannot happen without the immediate past occurrence of event  $e_{n-1}$ , for the latter depends upon the former—for example, one cannot have a ball rolling unless something has initiated the roll”. Obviously, this is common-sense knowledge. But it does not address the question of how this knowledge is represented in the neural network. Furthermore, does “something caused the ball to roll” fully represent the directionality of recall, making the direction of activating events in memory unambiguous? Would an organism with this sequence-string architecture in its brain be aware that it was recalling events backward? Or forward? We shall revisit this question along with all of the others in the final episodic memory architecture to be presented.

### 3 An Architecture for Notion 2: Holistic Access to an Episode

Let us next examine a reformulation of the episodic neural network architecture to address Notion 2, the importance of access to an episode as a holistic item. Here, we ignore all of the issues discussed previously and focus solely upon the holistic access notion. Fig. 5 (a) illustrates a scheme for accessing an episodic memory as a holistic entity,  $E$  representing the entire sequence  $e_1e_2e_3$ . Fig. 5 (b) shows a neural network layout for this scheme. The

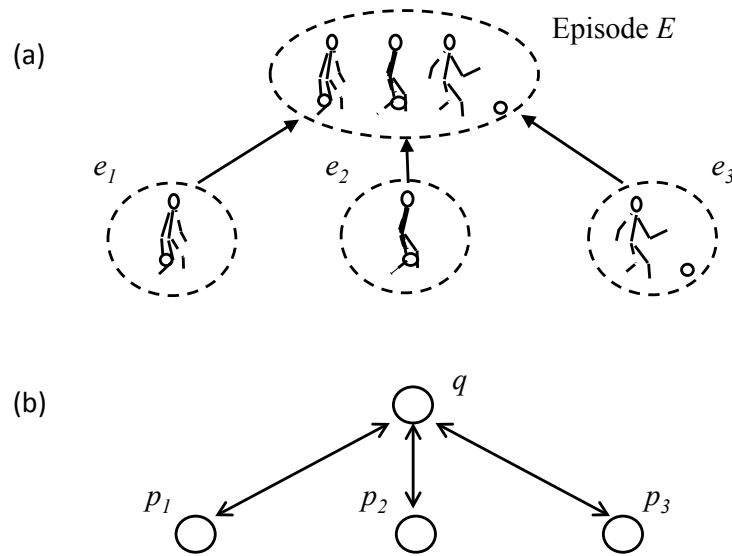


Figure 5: (a) An episode organized as a hierarchy in which the separate events are connected to a central entity, the episode. (b) A hierarchical neural network in which a cell  $q$  representing the episode  $E$  connects separately with the cells  $p_1, p_2, p_3$  representing its constituent events.

idea is simple: Instead of connecting the event cells in a string, connect the event cells  $e_1, e_2$  and  $e_3$  to the episode cell  $E$ . The cell  $E$  remains active during the recall of the whole episode or any part of its time span, while the event cells are active when their time segment relative to the whole is undergoing recall. The idea is not new (see for example [11]) and was a part of the episodic network architecture presented in ([13, 19]).

Notice that the arrows connecting the  $E$  cell to its individual event cells  $e_1, e_2, e_3$  in Fig. 5 (b) are double-ended, as are those between each event cell and its grounding cells in Fig. 2 (b). This serves to illustrate that both the event cell/episode cell connections and the grounding cell/event cell connections exist as reciprocating pairs. Having such connections in both the feedforward and feedback directions enables activity to propagate in both directions: for example, sufficient activity in the array of event cells stimulates an episode cell, bringing it to an active state. Reciprocally, the feedback connections are intended to have sufficient strength to enable an episode cell to activate its individual event cells. On the other hand, the arrows connecting  $E$  with its events  $e_1, e_2, e_3$  in the conceptual semantic illustration of Fig. 5 (a) point in only one direction, from each event to the episode. Similarly, the arrows connecting each event with its grounding objects point in only one direction, from each grounding object to the event in the conceptual semantic



illustration of Fig. 2 (a). The idea here is that as we transition into the mathematics, episodes such as  $E$ , events such as  $e_1, e_2, e_3$ , and the events' grounding objects and, in fact, all conceptual objects that are represented by neural cells, will become mathematical objects which we shall refer to as *concepts*. As such, the human figure grounding object is a subconcept of each event  $e_1, e_2, e_3$ , hence the upward direction of the arrows in Fig. 2 (a). Similarly, each event  $e_1, e_2$  and  $e_3$  is a subconcept of the episode  $E$ . The next section contains a more detailed discussion of these ideas.

Part (a) of the two Figs. 2 and 5 illustrates a revised model of episodic memory, one which replaces the event-to-event string with a hierarchy in which the events relate to the episode as subconcepts. Correspondingly, the neural architectures in part (b) of each figure present a hierarchical picture of grounding object and event cells connected in both feedforward and feedback fashion to their target cells, the event cells and the episode cell, respectively. The feedback connections reciprocate the feedforward connections all through the hierarchy and serve concept representation and other functions not addressed here. The feedback connections enable the propagation of activity in either direction in an episodic sequence; in many cases, they also enable cells to become fully active through the recurrent trading of stimuli between source and target cells. As Barsalou points out, reciprocal connections ensure *concept stability* [3]. This means that a cell's fully-active state occurs in parallel with the activity of its full complement of afferents which represent the subconcepts of its concept. This ensures that its concept representation persists and does not undergo either a catastrophic or fluctuating degradation because doing so would entail a concomitant change in its input sources.

Although it addresses the issue of holistic access to an episode, this architecture does so at the cost of sacrificing the explicit enforcement of the order of events which was inherent in the linear-event-string architecture. To accomplish the recall of events in order, either forward or backward in time, a control mechanism for sequencing the events must be added. As was the case with achieving flexible re-use of a stored memory in the first episodic memory architecture, this architecture requires extrinsic connections to enact the flow of events. A potential advantage of this second model, however, is that flexible re-use is more readily achieved because activating an event cell will not automatically cause activity to spread to other events. Activating the other events instead requires explicit action, either by bringing the episode cell to full activity or by providing stimuli to the other event cells by a different route.

The neural cells and connections required for such a control mechanism have the added task of capturing the correct sequence order adaptively, for the information about the order is not known to the neural network until the sequence first occurs. As a consequence, a method such as a specialized sort of weight adaptation is required to capture the sequencing information; previous papers have addressed this to some extent using this same holistic modeling approach (see the last paragraph of the Introduction). However, the question of how sequence order is captured and made use of in this holistic architecture remains an open question. As with the first neural episodic model, a great deal of infrastructure must

be added, such as something functionally equivalent to the selector cells referred to earlier. The loss of an explicit neural mechanism for sequence order complicates the need for added structure. An approach to the problem of holistic access that respects the episodic sequence order will be described in the presentation of the span-cospan model in a later section.

## 4 Concepts as Objects, Concept Relationships as Morphisms

Before continuing, a brief digression on the mathematical model is appropriate. The CNST was developed as a means of expressing the semantics of neural networks with mathematical rigor. The Appendix contains a brief introduction to category theory and its use in this semantic model. Here, we provide an overview.

Expressing the semantics of a neural network requires a form of mathematics allowing human-understandable, precise descriptions of the meaning implicit in its structure and activity. In formal logic, a precise mathematical description of a domain, or entities and situations in which they play a role, is called a *theory*. We also refer to theories as *concepts*, which is more suggestive of a description of the kind of stimuli that a particular cell in a neural network responds to. A cell responds based upon its connectivity with other cells, which may include sensor elements. This suggests that the concepts associated with interconnected cells are related by virtue of their connectivity and the relative strengths of the synaptic connection weights, whose magnitudes affect the transfer of stimulus activity.

Now, the mathematical expressions of the CNST are diagrammatic, diagrams being the expressions of category theory. The diagrams are somewhat in the nature of figures such as Fig. 5 (a), as they involve the *objects* and the *arrows* or *morphisms* in a *category*. Mathematically, arrows such as those in Fig. 5 (a) express concept inheritance relationships in a concept category **Concept** which has concepts as objects. For ease of visualization, we illustrate the concepts with drawings of the things they describe. A morphism  $A \rightarrow B$  in the concept category expresses the truth-preserving transformation of a concept  $A$  to a subconcept of a concept  $B$ . This formalizes the inheritance by  $B$  of the properties expressed in  $A$ , adapted to the form of expression in  $B$ . This formalism expresses the descriptive statement of each domain as a theory in a formal logic, allowing a precise symbolic description of the items and their properties. Similarly, a morphism formalizes the inheritance of concepts from their subconcepts. By this means, the category formalizes the notion of a concept hierarchy. For example, in the events  $e_1, e_2, e_3$  in the episode  $E$  (see Fig. 6) of our running example, the concept of a ball becomes, say, the concept of a bowling ball at the various stages of being thrown.

A particular neural cell, when active, is signifying that a particular stimulus to which the cell responds is present, signifying the presence of a particular item in the neural net-

work's input environment. Alternatively, the stimulus could be autonomously-generated. For example, a neural network with the capability to generate plans of action would synthesize concept representations of objects and relationships at the various stages of a plan. In either case, a concept associated with a cell describes the item at some level of detail; this description is an explicit account of the cell's semantics. To make the semantics of neural processing fully explicit, categories  $\mathbf{N}_w$  which we call neural categories formalize a neural network's structure and possible states of activity at a given stage of connection weight adaptation, based upon the current array  $w$  of connection weights (or in a biological neural network, synaptic strengths). A mapping between categories called a *functor* can be used to fully formalize the semantics of the neural network. The functor does this by mapping concepts and their relationships to states of neural activity in cells and the instances of activity flow between cells. Formal transitions between neural categories and functors express the learning that occurs through the synaptic adaptation. In neuroscience terms, the concept category is a formalized version of the hierarchy of representation, from sensor level to the "higher" levels at which recognition, recall and other cognitive operations take place. More mathematical detail is contained in the Appendix and a complete exposition resides in some of the references, principally ([22, 14, 23]).

The hippocampus and its closely-related structures such as the entorhinal, perirhinal and parahippocampal cortices making up the medial temporal lobe (MTL) evidently supports such a hierarchy for declarative memory ([26, 31]). Thus, the concepts and concept relationships at this level of processing express the facts, events and episodes of declarative memory. Other concepts and relationships express the items in nondeclarative or implicit memory.

At this point, we issue a word of caution. The discussion from this point increases in its intensity. Occasionally, to maintain a readable text, this intensity forces us to overload certain symbols and phrases with multiple meanings. We shall try to avoid this, but sometimes it will happen. Thus, for accuracy we would like to distinguish between informal discussions of the events and episodic segments on the one hand, and the concepts that express them mathematically on the other. This seems unnecessary in this paper because we are not providing a detailed expression of concepts and concept relationships as mathematical objects and morphisms in a category of theories; we simply mention the category and define colimits, commutative diagrams, and spans and cospans in categories, largely in the Appendix. More serious is the fact that the concepts, or events and episodic segments and grounding objects, are encoded or represented by neural cells and their connectivity, and sometimes the intensity of discussion makes it difficult to distinguish concept and relationship from neural cell and connection. Worse yet, the neural cells and connections and their activity, taken together, are formalized mathematically as a category of neural objects and morphisms, and we sometimes discuss the neural network in categorical terms, as in "cospan cells" or " $E_n$  cell". The cospans are in categories, not systems of episodic segments and events or neural networks and connectionist activity, but we find it necessary to simplify by merging the informal with the formal. Hopefully, this merging will be noticeable only to those concerned with the subtleties in having a mathematical model.

## 5 An Architecture for Notion 3: an Explicit Representation of Shared Information

Having proposed episodic models that address notions 1 and 2, we proceed to Notion 3. It is based upon the observation that the events in an episodic sequence are mutually consistent. That is, they somehow *cohere* in memory: Successive events share information by maintaining some grounding objects in passing from one event to the next. Imagine what a memory of a sequence of events might be like if the grounding objects or their relationships were to change abruptly in the transition from one event to another. Consider an example:

Suppose that in event  $e_1$  a human started throwing a ball, but in event  $e_2$  the human was replaced by a wagon and the ball became a goat towing the wagon with a rope. In this example, the stringing-together of such images as  $e_1$  and  $e_2$  would of course result in utter confusion, or else register as the kind of abrupt transition that occurs only in dreams or surrealistic films. In the latter case, the human brain apparently maintains an expectation that such transitions make sense in some way, perhaps in a symbolic interpretation. The point here is that to make everyday sense as an episodic memory, an event sequence must consist of events having grounding objects and their relationships in common, or at least in a close similarity relationship. A transition between objects such as a human figure becoming a wagon does not meet this criterion, nor does a transition between relationships in which object  $A$  throwing object  $B$  becomes object  $A$  being pulled by object  $B$ .

A question then arises: How does a neural network dissociate related and unrelated event pairs, and, hence, sequences that constitute episodes from those that do not? For example, how can we be aware that a memory makes sense because it has continuity but a dream sequence with abrupt transitions does not? Or distinguish memories that have been formed before a sleep period from those that are formed after? Another consideration is the ability of the brain to reconstruct images and other complex information as a contiguous whole, attending to the overlap of shared information between components so that there is no confusing duplication or uncertainty in the arrangement of the components.

In answer to these considerations, we hypothesize that successive events which are closely related by content, such as the grounding objects shared by the events in Fig. 5 (a), must be explicitly represented as such in a corresponding neural architecture—but this is not apparent in the architecture of Fig. 5 (b). This suggests a principle which we might call *the shared subconcept principle*, requiring that the common content of two concepts be made explicit in their neural representation. This principle, a part of the CNST, is based upon the category-theoretic structure called a *colimit*.

The architectures illustrated in Figs. 2 (b) and 5 (b) can be incorporated in an architecture that addresses the shared subconcept principle. This architecture is suggested conceptually in Fig. 6, where the grounding objects are shown as shared subconcepts of all three events  $e_1$ ,  $e_2$  and  $e_3$  and these in turn are subconcepts of the episode  $E$ . Cor-

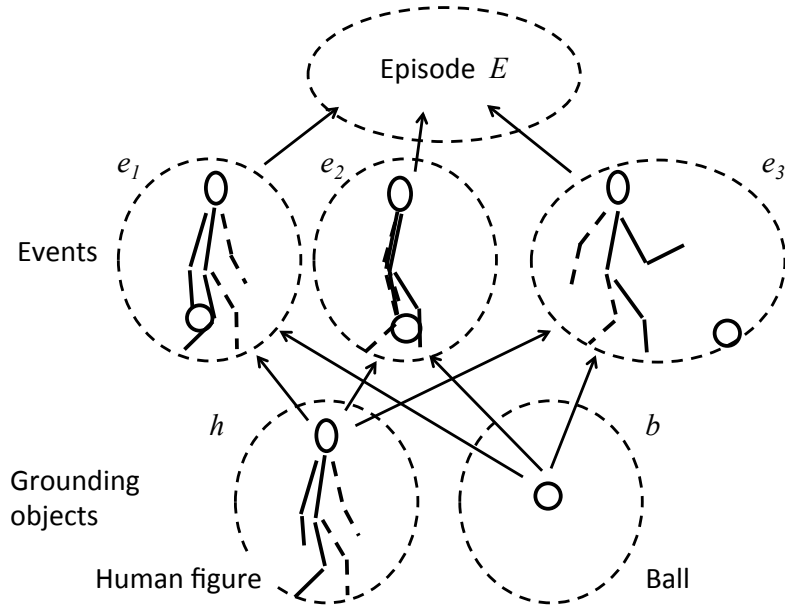


Figure 6: **An episode organized as a hierarchy in which the grounding objects are shared subconcepts of the separate events, which are in turn subconcepts of the episode.**

respondingly, one might expect there to be separate neural network connections for each grounding object cell—the human figure, and the ball—to/from each event cell and for each event cell to/from the episode cell. Now, this connectivity was mentioned previously; we stated then that the fact of such connectivity of ground object cells to event cells was not sufficient for an explicit representation of the sharing of grounding objects.

To address this insufficiency, notice first that (see Fig. 7) each grounding object cell is connected through three different feedforward connection pathways to the episode cell. For example, for the human figure concept cell, one pathway consists of its connection with the  $e_1$  cell,  $p_1$ , and that cell with the episode  $E$  cell. The two connections actually consist of reciprocating pairs, annotated with arrows to show the direction of signal propagation in each connection (also, the lobes at the terminal ends of the connections show the synapses, again indicating the direction of signal propagation). Together the reciprocating pairs comprise a feedforward path from the human figure cell to the  $E$  cell and also the accompanying feedback path. Similarly, the human figure cell also connects with the  $E$  cell through both  $p_2$  and  $p_3$  via the respective feedforward/feedback pathways. In line with the discussion accompanying Fig. 3, additional objects and relationships and their corresponding cells and connections which are not shared by the three events are not shown.

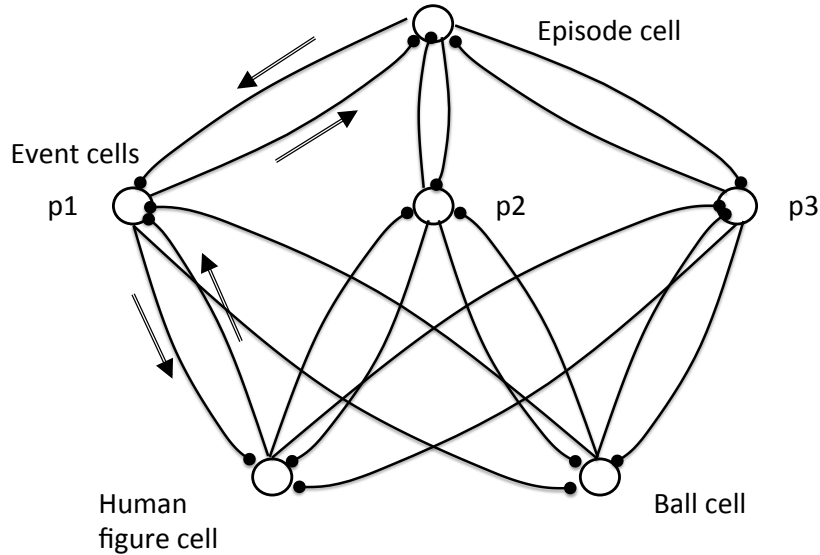


Figure 7: **An architectural layout for the episodic network, organized as a hierarchy of cells through reciprocating pairs of connections.**

How is the neural network in Fig. 7 to respect the human figure as a common presence in the events if the event cells and the human figure cell are not all active simultaneously? After all, the simultaneity of activity among cells is basic to a neural network's representation of objects that are part of a more complex whole. The required simultaneity of activity is illustrated in Fig. 8 where the heavy arrows along the two paths (human figure cell) $e_1$ (episode cell) and (human figure) $e_2$ (episode cell) is meant to indicate that the two paths always act together. Given our use of notation, with the human figure denoted by  $h$ , we can write the two paths from  $h$  to  $E$  as  $he_1E$  and  $he_2E$ . The implication of this is that if  $e_2$  is the current event during either an original encoding of the sequence or during recall, we require that there remain some persistent activity in the  $e_1$  cell while the cells for  $h$ ,  $e_2$  and  $E$  are active. This simultaneity of activity in two pathways with a common source and target corresponds to the mathematical notion of a *commutative diagram* in a *neural category*. The notion of a commutative diagram is a general one in category theory and is one of two requirements defining a *colimit*, of which our object sharing by events is an example. The activity in the neural architecture corresponding to the heavy arrows in Fig. 8 is illustrated by the arrows accompanying the appropriate connections in Fig. 9. Of course, the ball cell will also be active, and so although this is not illustrated, the simultaneous activity in the connections with the ball cell indicate that the two shared objects, human figure and ball, are being highlighted simultaneously, as would happen in

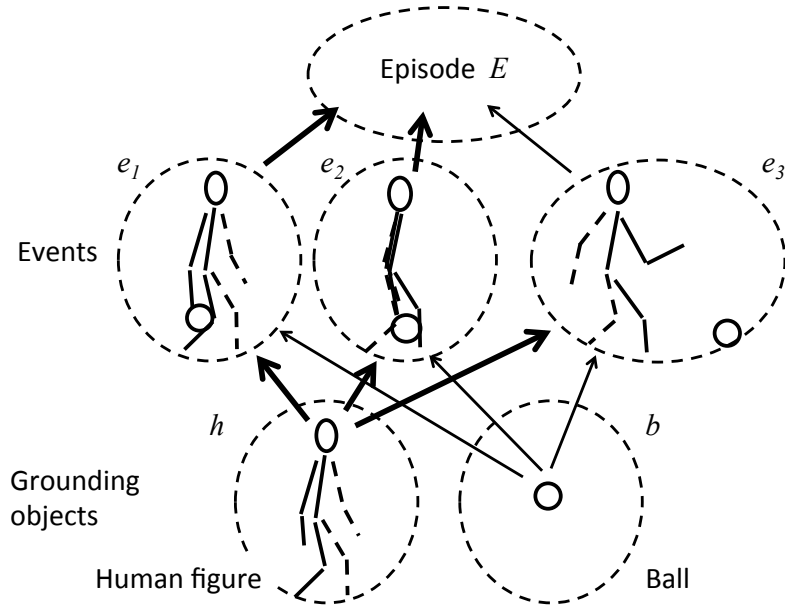
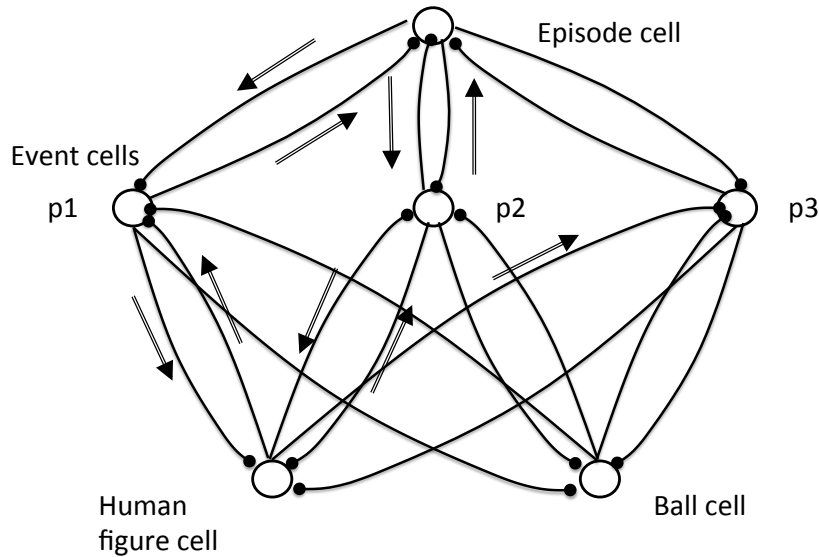


Figure 8: **Simultaneous activity in the two pathways from the human figure cell to the episode cell enforces the sharing of the human figure object between the events  $e_1$  and  $e_2$ . This property enables the architecture to address Notion 3: shared-object representations.**

the events in question. In fact, the commutative diagram for this episode extends to not only the ball and its connection pathways with the episode cell, but also the connections involving event  $e_3$ .

The important point is that in an episode in which a grounding object is present in more than one event, the connections joining the grounding object cell to each such event cell and the event cells to the episode cell are active simultaneously. Yet, the event cells are to be active successively, not all at once. How is it possible, then, to claim simultaneity of activation? As explained in [13], we have proposed an extended definition of simultaneity for what we call *temporal morphisms*, as follows. The simultaneity of activity in a set of pathways with a common source and target cell means that neural activity and signaling through the connections occurs in each pathway of the set *at some time* during a time interval that defines the set, while *the common source and target cells of the pathways are persistently active* during the entire interval. The complete definition of the set of pathways for the commutative diagram encompasses all the pathways for the episode in our example, since they all have the same sources, the  $h$  and  $b$  cells. In summary, the cells for  $h$ ,  $b$  and  $E$  persist in activity during the entire time interval of the episode while the pathway involving each individual event is active during the appropriate subinterval of



**Figure 9: A neural architecture for Notion 3. Simultaneous activity in the two neural network pathways from the human figure cell to the episode cell enforces the sharing of the human figure object between the events e1 and e2. Notice that the pathways referred to in the discussion have both feedforward and feedback connections active.**

the time interval defining the episode. This example illustrates the third episodic memory architecture, enacting both the holistic episode and the sharing requirement. This type of architecture was described in [13], where it was pointed out that it utilizes the categorical notion of colimits. A simulation model for it which incorporates a capable classification network from the literature is described in [19], which also reports simulation experiments carried out with the model. Commutative diagrams and colimits are explained in more detail in the Appendix.

There is a final step in architectural design. This final architecture addresses all four of the notions about episodic memory architectures. To review, these are the Notion 1, the maintenance of the event representations in sequence; Notion 2, holistic access to an episode; Notion 3, shared-object representations; and the notion not yet fully addressed in concert with the others, Notion 4, flexibility. Notion 4 proposes the necessity of separate access to partial event strings, or episodic segments, and requires that there be cells harboring an explicit representation of the predecessor/successor relationships between the events in an episodic sequence. Our fourth design addresses all of these issues in a single episodic memory architecture called the *span-cospan model*.



## 6 A New Episodic Memory Architecture for Notion4: flexibility

This section introduces the subject of this paper, the span-cospan episodic memory model. Some reference will be made to mathematical structures including the objects and morphisms that form a category, commutative diagrams, colimits, spans and cospans, and functors that map one category into another. These will be explained in summary form as they appear; some have already been introduced. For a more detailed exposition, see the Appendix.

### 6.1 Spans and cospans

To address all of the shortcomings of the previous episodic memory models, we propose a fourth. This model fully addresses Notion 1, 2, and 3, and also fully addresses Notion 4, flexibility in re-use. A key part of this is that it provides an explicit representation of sequence order, the precedence-successor relationships between events. It does so by taking advantage of two categorical structures which have not yet been introduced: spans and cospans. For those not familiar with category theory, a reading of the sections of the

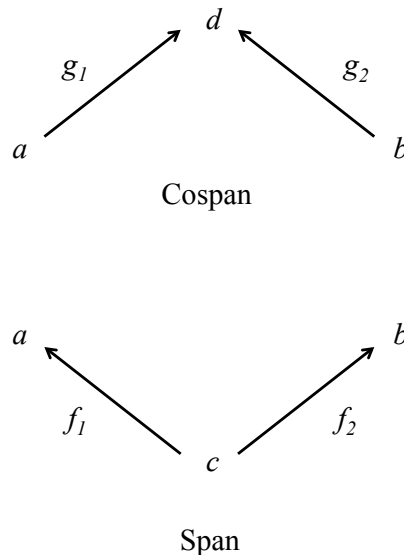


Figure 10: A span and a cospan.

Appendix that refer to morphisms, commutative diagrams, and colimits will greatly aid in following this discussion.

Examples of the two structures, span and co-span, are shown in Fig. 10. The structure containing the three objects  $a, b, c$  and the two morphisms  $f_1: c \rightarrow a$  and  $f_2: c \rightarrow b$  is called a *span*; it has two morphisms with the same domain  $c$  but different codomains  $a$  and  $b$ . An example of a span in the category **Set**, where sets are the objects and functions are the morphisms, is a relation  $R: a \rightarrow b$  between sets  $a$  and  $b$  represented by functions  $f_1$  and  $f_2$ , with  $c \subseteq a \times b$ . The structure containing  $a, b, d$  and the two morphisms  $g_1: a \rightarrow d$  and  $g_2: b \rightarrow d$  is called a *cospan*; it has two morphisms with different domains  $a$  and  $b$  but the same codomain  $d$ . Interestingly, fusing the two structures in Fig. 10 along the objects  $a, b$  produces a diagram with a diamond shape. This same shape characterizes a simple type of colimit called a pushout, in which the diamond is a commutative diagram (see Appendix). In fact, a pushout is shown in the Appendix as Fig. 19; here, the colimit is shown with all of its cocone morphisms including the one directed up the diagonal of the diamond shape. Since the diagram commutes and, hence, the diagonal morphism equals the composition along either side of the diamond (i.e., the resultant morphisms of the two compositions are equal) it is often not shown. It is worth showing it here because our neural models include such diagonal morphisms explicitly as connections.

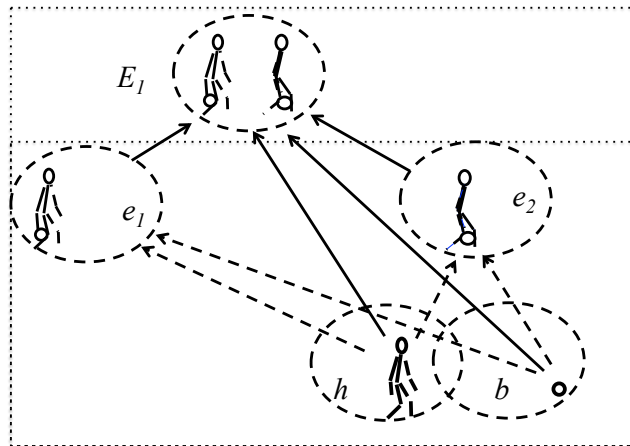
## 6.2 Episodic segments and overlapping colimits

The fourth episodic model consists of a neural structure that represents an incremental buildup of overlapping *episodic segments*. As shown in Fig. 11, the episode of our ball-throwing example is formed incrementally with the successive, overlapping colimit defining diagrams labelled Step 1 and Step 2. Another way of looking at this is illustrated in Fig. 12. In similarity with the discussion in the preceding section, analyzing a pushout as an overlapping span and cospan, the upper part of the diamond-shaped diagram of a pushout (omitting the middle morphism from the picture) is a cospan. In this alternative view, then, omitting for simplicity of illustration the lower half of each colimit diagram, which are the spans involving the grounding objects with the events, yields a picture of two overlapping cospans, labelled Cospan 1 and Cospan 2 in Fig. 12. Cospan 1 has at its apex the first episodic segment, the first two steps in throwing the ball. Cospan 2 joins this first segment with the final step in throwing the ball.

The span-cospan architecture has the following operational characteristics. First, at each step in the formation of a new memory a cospan cell is recruited to be the apex of a colimit defining diagram that blends an episodic segment (the previous episodic segment cospan object) and the next event in order using the grounding objects as blending objects. In the corresponding neural architecture, this yields a new episodic segment cospan cell further along in the memory sequence. The new episodic segment serves as a holistic representation of the memory sequence through that step. The corresponding neural cell has access to the preceding episodic segment cospan cell as well as its terminal event cell

via feedback through the reciprocals to its feedforward cospan connections from those cells. This allows a memory to be parsed in either forward or reverse order, proceeding iteratively through feedback and feedforward reciprocals. This allows an episodic segment to be separated into its preceding segment and its final event for applications such as the determination of what led to some important event.

Step 1



Step 2

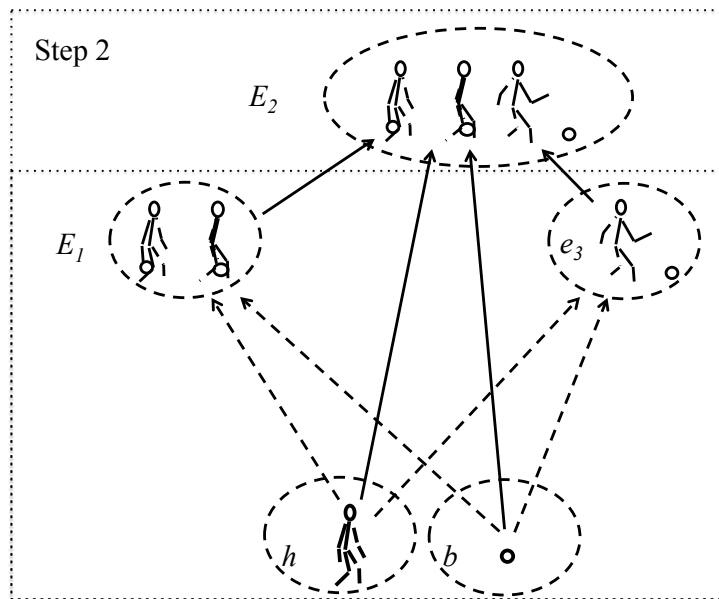


Figure 11: **The incremental formation of an episode in segments. The second colimit builds upon the first.**

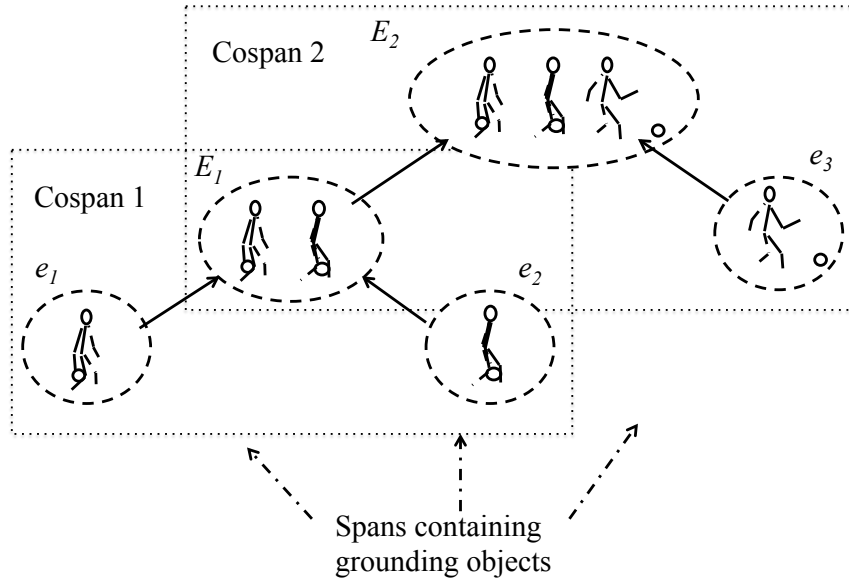


Figure 12: **The result is a string of linked cospans. Each episodic segment incorporates the preceding one and adds in a new event.**

This is illustrated in Fig. 13, where the episodic segment  $e_1e_2e_3$ , or  $E_2$ , of Cospan 2 is parsed into its two components, the Cospan 1 episodic segment  $e_1e_2$ , or  $E_1$ , and the event  $e_3$ . The same process can parse episodic segment  $E_1$  to recall  $e_1$  and  $e_2$  simultaneously as separate representations. In this way, any of the holistic memory segments can be separated into its end state and the episodic segment immediately preceding the end state. The neural architecture corresponding to figs. 12 and 13 is illustrated in Fig. 14. Recall in the forward direction involves the reverse process, with each event in turn stimulating through its feedforward connection its cospan cell, followed by a cospan-to-cospan cell morphism and the activation of the next event cell via the reciprocal to the cospan connection from that cell. Both the forward and backward recall operations are illustrated in Fig. 15.

### 6.3 Holistic access, blending along shared concepts, and intrinsic flexibility

The architecture in Fig. 14 maintains the episodic sequence order. It does this by allowing recall through the feedforward connections between successive cospan cells and the ability to parse the represented segments by successively activating the cell representing the terminal event of each segment via the feedback connections to the event cells. This feedforward/feedback interplay occurs in the direction of the episodic sequence order, thereby

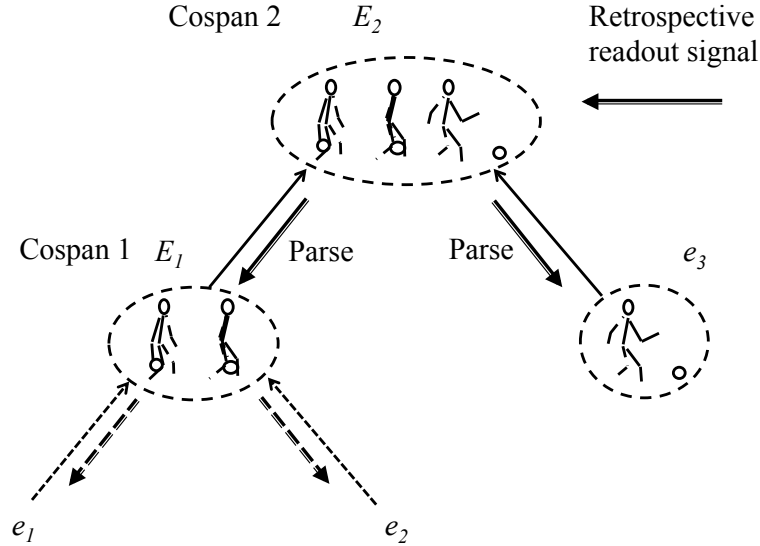


Figure 13: **Parsing an episodic segment by separating the cospan representation  $E_2$ , then  $E_1$ , into their components. The episodic segment  $E_2$  separates into its predecessor segment  $E_1$  and terminal event  $e_3$ . Next,  $E_1$  separates into its predecessor segment, here consisting of the single event  $e_1$ , and its terminal event  $e_2$ .**

addressing Notion 1. Based upon this parsing capability, each cospan cell provides holistic access to each episodic segment, including the entire episode. This more than addresses Notion 2. The presence of the spans, whose cells represent the grounding objects which are the blending objects in each colimit diagram for an episodic segment, partially addresses Notion 3, the representation of the shared information between events that enforces the continuity of an episode. The representation of shared information is completed by the fact that a cospan and the spans connected to it together form the defining diagram of a colimit, illustrated in Fig. 21. This means that the diagram commutes, and therefore the represented episodic segment is a properly-blended composite of the events within it along their shared grounding objects.

Finally, the span-cospan episodic memory model transcends the first, second and third models because of the intrinsic flexibility it has in the storage and recall of episodic memories, addressing Notion 4 as well as notions 1, 2 and 3. Adaptation to form the episodic segment cospans in feedforward fashion allows a memory to be learned and stored in the neural network. The memory can be recalled by activating the cell for a chosen event through its afferent connections from grounding object cells, through a feedback connection from the episodic segment cospan cell with which it forms the cospan, or through a

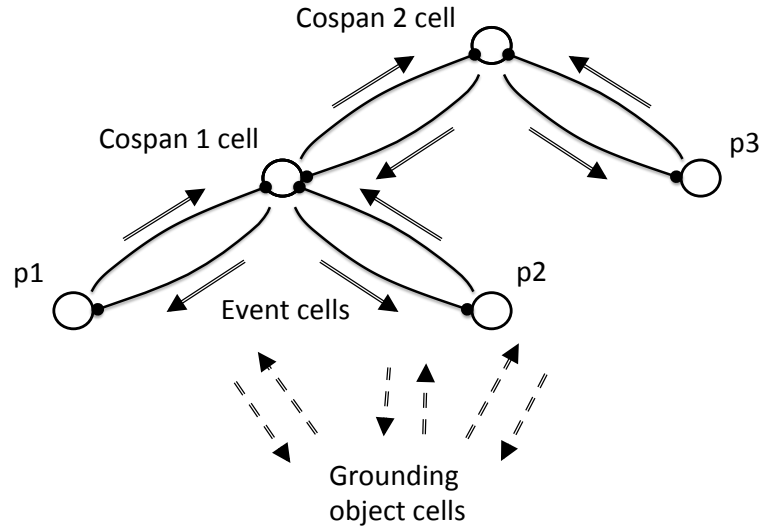


Figure 14: A neural architecture for the span-cospan episodic memory model. The Cospan 2 cell represents Episodic segment  $E_2$ ; parsing  $E_2$  consists of tracing through the feedback connections to activate the Cospan 1 cell (representing Episodic segment  $E_1$ ) and also the cell  $p_3$  (representing the terminal event  $e_3$ ). Parsing  $E_1$  is done in a similar fashion, activating the cells  $p_1$  and  $p_2$  (representing the events  $e_1$  and  $e_2$ , respectively). The cells and connections representing the grounding objects provide input to the event cells; the arrows represent the reciprocal pairs of connections between the grounding objects and the event cells  $p_1$ ,  $p_2$  and  $p_3$ .)

connection from some other part of the network. Feedforward activity from cospan cell to cospan cell activates the event cells via the reciprocals to their connections to the cospan cells. Backward tracing of a memory through this reciprocal parsing strategy has already been described. Finally, a sequence of cospans cells can be entered at any place in the neural cospan array. One pathway for this is through feedforward connections from the grounding objects cells to an event cell at that place, and then forward to the cospan cell it joins. For example, seeing a familiar face might trigger an episodic recall at one or more places in an episode in which the person was last seen. Another potential pathway is through connections formed between an episodic segment cospan cell and cells in some other part of the neural network.

Presently, we shall explain the major contribution of the span-cospan model, the explicit representation of concepts which express the episodic sequence order. This requires a discussion of the episodic segments in the concept hierarchy of an episodic sequence and

their representation by the cospan cells.

Before discussing the latter property of the span-cospan model, let us examine the neural network and its operation in more detail. Notice that in the neural model illustrated in Fig.14 all feedforward connections have reciprocals which enable the parsing of the episodic segments. In the example, the episodic segment  $E_1$  of Cospan 1, which immediately precedes the goal state event  $e_3$ , can be recalled separately from  $e_3$  once the  $e_3$  cell has been stimulated into full activity by an extrinsic stimulus. Suppose that the  $e_3$  cell receives an extrinsic stimulus, a command sent through a neural connection from some other part of the neural network. This cell first stimulates the  $E_2$  cell for Cospan2. The Cospan 1 cell is then stimulated via signaling activity through the reciprocal connection from the Cospan 2 cell to the Cospan 1 cell representing  $E_1$ . This enables the  $E_2$  cell, whose episodic segment representation includes the goal state  $e_3$ , to stimulate the  $E_1$  cell, whose episodic segment precedes the goal state. By virtue of the other cospan reciprocal connection, the  $E_2$  cell also reinforces the goal state  $e_3$  cell's activity, or activates it in case it was the  $E_2$  cell instead which received an extrinsic stimulus. Thus, as mentioned before both the preceding episodic segment and the goal state can be recalled into working memory simultaneously as separate representations in an enactment of the parsing strategy. The "Retrospective readout signal" in Fig. 13 illustrates an example in which it is the  $E_2$  cospan cell that receives the extrinsic stimulus, which then parses the episodic segment  $E_2$  into its two components  $E_1$  and  $e_3$ .

The forward and backward recall operations are illustrated for the general case beginning at an arbitrary event  $e_n$  in Fig. 15. An excitatory connection is shown emanating from a "Selector cell" in "Step 1" and terminating upon the  $e_n$  cell. In the "Forward recall" case a separate, inhibitory connection terminates upon the cospan cell labelled " $E_{n-2}$ ", blocking the activation of that cell and, hence, any progression in the backward direction along the episodic sequence. In "Step 2", the excitatory stimulus activates the  $e_n$  cell, allowing it to activate its episodic segment  $E_{n-1}$  cell through a feedforward connection. Further in the feedforward direction in "Step 3" the  $E_{n-1}$  cell stimulates the  $E_n$  cell, which then in "Step 4" activates the cell representing its terminal event  $e_{n+1}$  and separately through a feedforward connection the  $E_{n+1}$  cell, and so forth until an inhibitory input blocks some further-along cell. Notice that in a separate branch of "Step 3" the  $E_{n-1}$  cell attempts to stimulate the  $E_{n-2}$  cell through a feedback reciprocal to the feedforward  $E_{n-2} \rightarrow E_{n-1}$  connection, but this is blocked by the inhibitory connection from the selector cell. The operation of the network in the "Backward recall" case proceeds in an analogous fashion. Retrospective memory can be seen as an elaboration of backward recall in which the activation of the selected event  $e_n$  cell activates the  $E_{n-1}$  cospan cell, calling up its episodic segment as before, but in this case recall is forced to proceed in the backward direction. The Figure provides some indication of the full flexibility available in this fourth episodic memory model.

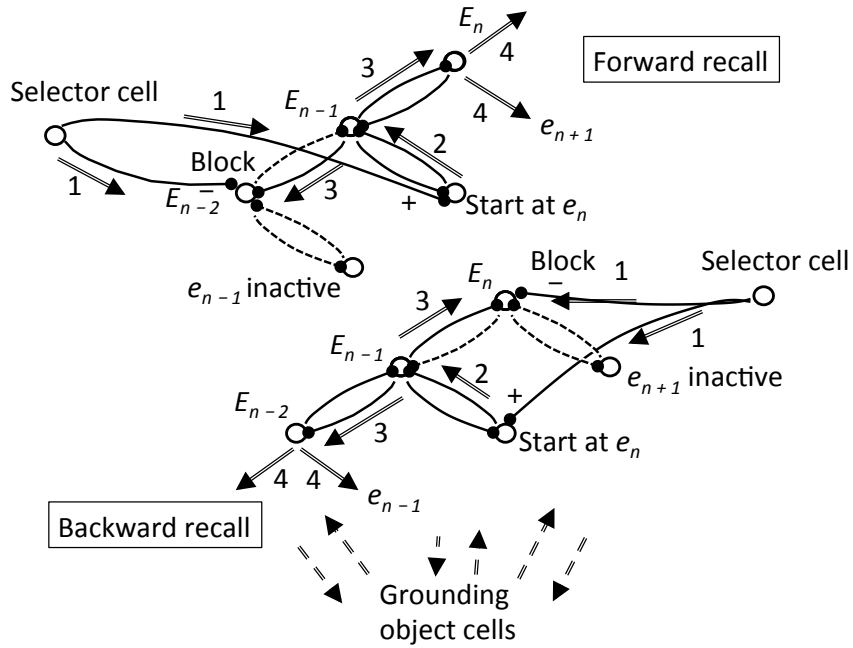


Figure 15: **Two basic recall operations of the span-cospan episodic memory model. For ease of reading the Figure, the word “cell” is omitted from the caption for each cell, as for example “ $E_n$ ” in place of “ $E_n$  cell”.**

## 6.4 The explicit representation of sequence order

The unique contribution of the span-cospan episodic memory model in transcending the previously-described models is the explicit representation of sequence order. The cospan cells represent episodic segments, which build up incrementally by adding the next event at each step to form a new segment. This is the basis for an unambiguous, explicit representation of sequence order. First, the succession of cospans cells through their feedforward connections to their successor cospan cells maintains the sequence order. The growing accumulation of events during the formation of an episodic memory is represented by the cospan cells, recruited successively as events occur. During this phase, some event cells are themselves newly recruited, some event cells are re-used but have their representations modified by new information, and some are simply re-used. During the learning of an episode, each cospan cell represents not only the current event, but the episodic segment recording the history of the episode up to that point in time. The addition of a new event at each successive cospan corresponds to the successor operation. The feedback reciprocals between cospan cells facilitate the parsing of each episodic segment into its predecessor episodic segment and final event, where the feedback link to its predecessor is a manifestation of the predecessor operation.



However, the fact of an operable architecture does not in and of itself express the sequence order explicitly, for this requires that the order relation be represented explicitly in one or more concepts. Although an episodic segment concept is a colimit, the order relation is not an intrinsic property of colimits. The event concepts are subsumed in the episodic segment concepts but this is a superposition of concepts, not an ordered list. How, then, can we claim that the span-cospan architecture provides an explicit representation of sequence order as a concept?

Previously, we have stated that the cells which are afferent to a neural cell  $p$  must remain active at some level in order for its concept  $T$  to be fully represented. Any inputs that are not active will cause their corresponding representations to be absent, resulting in the representation of a part of  $T$  only in the most abstract fashion, that is, without any defining detail. It is then correct to say that when some of its original inputs—for example, cells representing some of the grounding objects for an event, or previous events and episodic segments for a cospan cell—are inactive, the cell is not at that time representing  $T$ , but an abstraction  $T'$ , a subconcept that has some of  $T$ 's component objects represented with less detail or not at all. This applies in particular to a colimit cell representation: its defining diagram cells, and in fact *their* inputs, must have persistent activity during the time interval in which the colimit concept representation is active in the neural network in order for all of the details of the colimit concept to be fully expressed. Each cospan cell  $E_n$  is one such colimit cell. Hence, for its representation to be specific to  $E_n$ , the cells in its defining diagram must persist in activity sometime during the interval of activity of the cospan cell. The diagram cells include the cospan cell representing the previous episodic segment  $E_{n-1}$ , the cell representing the current event  $e_{n+1}$ , which is the final event of  $E_n$ , and the grounding cells representing the concepts that  $E_{n-1}$  and  $e_{n+1}$  have in common. That  $e_{n+1}$  is represented separately from the previous episodic segment  $E_{n-1}$  in the diagram implies that the active cells represent not only the current episodic segment  $E_n$ , but also, *separately*, its predecessor episodic segment  $E_{n-1}$  and the current event  $e_{n+1}$ . This hierarchy is illustrated by the layout in Fig. 12. This is one way in which the sequence order is made explicit: An episodic segment is displayed simultaneously with its terminal event and the episodic segment that is its predecessor.

Further, previously we have defined temporal neural morphisms, which allow for the separate connection pathways to be active one at a time during separate subintervals of an interval in which the common source and target of the pathways is persistently active (see for example [13]). This property therefore allows the sequential activation of the pathways in the diagram for the cospan colimit cell representing  $E_n$ . This allows first the  $E_{n-1}$  cell and then the  $e_{n+1}$  cell to become active during the activation of the  $E_n$  cell, yet this sequential activity is regarded as simultaneous because of the persistent activity of the  $E_n$  cell and the blending cells representing the grounding objects in its colimit defining diagram. Nevertheless, the sequential activation of the  $E_{n-1}$  and  $e_{n+1}$  cells provides a second expression of sequence order, occurring within the duration of their temporal-morphism simultaneity.

It is worth mentioning here that the activation of the  $E_{n-1}$  and  $e_{n+1}$  cells can co-occur; an example of this is an overlap during the sequential activation of these cells. Or, alternatively, the order of activation of the two cells can be reversed. Which of the alternatives takes place depends upon the activity of other cells which have connections terminating on the two in question. In fact, fluctuating activity in the grounding cells, including those not in the colimit diagram for  $E_n$ , can cause the activities of the  $E_{n-1}$  and  $e_{n+1}$  cells to oscillate. This does not imply that these cells alternately shut down and then spring into full activity; although we have not yet discussed the graded nature of activity possible in the cells, the oscillatory activity mentioned can occur in an increase/decrease rather than on/off fashion.

In connection with the graded activity of neural cells, we mention a third mechanism for the explicit representation of sequence order in the concepts forming the episodic segments. This mechanism is a progressive decrease in the specificity, or the amount of detail, in the information that is the content of a neural cell's concept representation when the activity of a cell undergoes decay. It can occur as an episodic sequence is either learned through synaptic connection-weight adaptation in an episodic network, or is explored through recall. A decrease in the specificity of a concept corresponds to an increase in its abstract nature. Neural mechanisms for an increasing abstraction in the concepts represented in an episodic sequence are discussed in the next two sections. Because the events undergo increasing abstraction during later events as an episodic sequence is encoded in memory, the earlier an event is in an episodic segment, the more abstract its representation. To greater or lesser extent depending upon how rapidly abstraction occurs, this allows for the representation of order in the events encoded in an episodic segment as a colimit.

## 6.5 Simultaneity and the sequential activation of cospan cells

Just as the activity of the  $E_{n-1}$  cell of Fig. 15 requires the activity of the cells  $E_{n-1}$  and  $e_{n+1}$ , sequential or otherwise, and persistent activity in the blending cells in its colimit diagram, the activity of the  $E_{n-1}$  cell requires the activity of the cells  $E_{n-2}$  and  $e_n$  and persistent activity in the blending cells of *its* diagram in turn. Utilizing the notion of temporal morphisms, however, only the activity of the  $E_{n-1}$  cell and its blending cells is required during the entire interval of its activation. The foregoing argument applies also to  $E_{n-2}$ , which is a component in the colimit defining diagram of  $E_{n-1}$ . The requirement for simultaneity of activation would seem to lead to a chain of activity all along the string of cospan cells as well as in the events contained in their diagrams and, hence, all grounding objects, whether they are contained in all events or not. But again, by virtue of the temporal morphisms, the entire hierarchy of cells does not have to be active at once; our extended notion of simultaneity in a temporal morphism implies that each cospan cell, hence, the blending cells in its diagram, are the only ones required to be active at once, and that only for the duration of that cell's colimit representation and its required presence during part of the next cospan cell's representation. Hence, the episodic sequence of cells need not all

fire as one.

The question arises, however: Is it necessary for all of the cospan cells preceding the  $E_n$  cell and their diagrams to become sequentially active if  $E_n$  is the current focus of attention, say, in determining what immediately preceded the event  $e_{n+1}$ ? The same question arises in parsing the episodic sequence representation in the backward direction; would the requirement for traversing the full sequence in the forward direction cause confusion? The abstraction of concept representations through decreasing activity in cells provides an answer. For, the full activation of a cospan cell, representing  $E_n$ , say, is not required if we are willing to allow abstraction to occur in the events within the episodic segment representation. One way this occurs is in the limitation of the length of a parsing in the backward direction, where the operation starts at the  $E_n$  cell but ends at a cell representing an earlier episodic segment  $E_{n-k}$ . This can be brought about through the inhibition of the  $E_{n-k}$  cell by a selector cell; selector cells are extrinsic to the episodic sequence in the neural network as shown in Fig. 15. Another way that abstraction can occur is through the decay of an active neural cell's activity. The farther back an event is in the episodic sequence, the more abstract its representation by the  $E_n$  cell because that event cell is registering a low level of activity. Hence, in actuality the  $E_n$  cell is representing an episodic segment  $E'_n$  whose events register this backward sense of increasing abstraction. If the decay in activity of the cells is rapid, it will reach a sub-threshold level for earlier cells and, hence, the actual length of an episodic segment can be reduced. Mechanisms for this backward-decrease in activity of the cells representing an episodic sequence are discussed in the next section.

## 6.6 Time and the abstraction-specialization hierarchy

The explicit representation of sequence order and the requirement for simultaneity in the activation of preceding cospan cells when a cospan cell representing an episodic segment is the focus of activity are two considerations we addressed in sections 6.4 and 6.5. In addition to providing more detail, this section addresses yet another consideration: How does a sense of the flow of time become associated with an episode? The latter issue can be addressed with a proposal which if implemented also leads to an additional means of addressing the simultaneity of activity in the components of an episodic segment.

One mechanism for the "time-stamping" of the events within a single episodic segment is the decrease in the strength of each event's representation within that segment. The activity decay of active neural cells was used in an earlier episodic model we published, in which the cells represented the events ([13, 19]). This earlier episodic memory model is a combination of those which address Notion 2: holistic access and Notion 3: the representation of shared information, presented here (see Fig. 21). In the earlier published model, the activity of a cell representing an early event in an episode has lower activity at a later time than the activity of a later event (see Fig. 16). This results in a gradient of activation over the events in which the cell activities increase with the events in order of

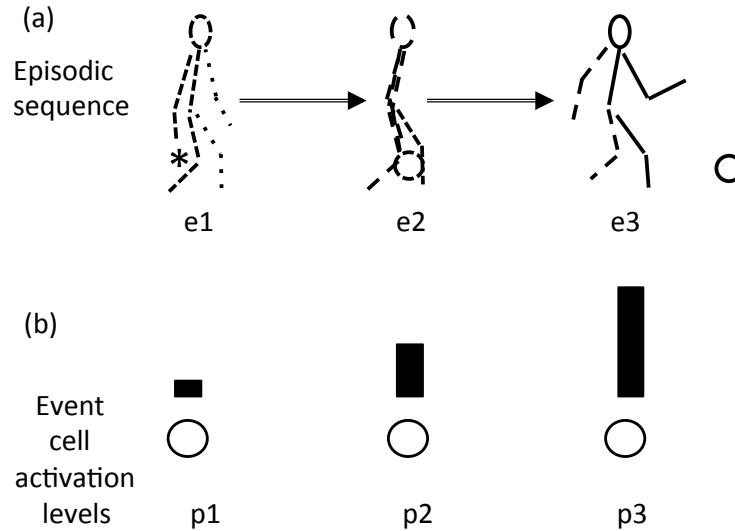


Figure 16: **An activation gradient over the event cells has been proposed as a way to “time-stamp” the events in an episodic sequence.**

occurrence, a *recency gradient*. This would affect the representation of a cospan cell in that the event concepts would have different types of representation depending upon their recency, the ones with less-active cells being more abstract at a given step in the event sequence.

Other alternatives are available for the “time-stamping” of events in a cospan representation via a gradient of abstraction backward in time. To illustrate such an alternative, Fig. 17 shows the cospan cells in the span-cospan model interconnected by feedforward connections with decreasing strength. The connection weight value  $1/2$  is used as an example. Let the feedback connections have connection weight unity. Also, let the reciprocating connections between each event cell and the cospan cell it connects to have unit weight as shown in Fig. 17. Then, a cospan cell receives only half the activation value of its predecessor cospan cell as input (to simplify, we shall not be concerned with the connection weight of its input from the relevant event cell). However, through the reciprocal feedback connection, it supplies the full strength of its activation to its predecessor cospan cell. The effect these weight values have is as follows. First, the cumulative weakening of input from episodic segments further back in time through successive low-weight connections establishes an effective recency gradient within each episodic segment. Second, the full-strength feedback connections maintain the ability of each cospan cell to parse its episodic segment into the previous segment and its final event by activating them with

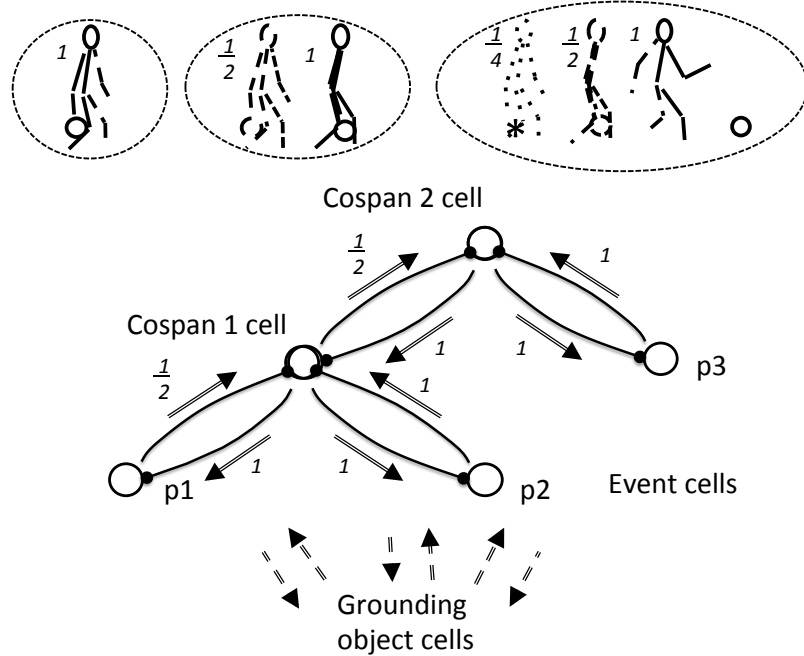


Figure 17: A hierarchical gradient of representation over the events within each episodic segment results from the cumulative decrease effect due to weights of less than unity strength between cospan cells. Notice the gradient of event representations from abstract to specific in each segment.

their full strength. The weakening effect of event representations further back in time occurs because of the successive weak cospan-to-cospan connections; note that cospan cells occurring in a string before a given one provide a lower level of activation input than does its final event cell, which provides its input at full strength through a feedforward connection with weight unity. This effects a separation into a fully specific representation of a currently-active event cell and a more abstract representation of the preceding episodic segment. Finally, we can claim that this *hierarchical gradient* in the representation of events within an episodic segment corresponds to a notion of time, since it represents the order of the events.

Taken together, there are three components of the episodic memory model for the explicit representation of the event sequence order in an episodic sequence. One lies in the structure of the neural network, based upon the separate representations of  $E_n$ ,  $E_{n-1}$  and  $e_{n+1}$  at each step of the sequence. The second is the sequential activation of events along with the episodic segments that contain them during succeeding (even perhaps overlapping) subintervals of the interval of activation of the  $E_n$  cell. And in this section we discussed a third component, which can be implemented through two mechanisms. One mechanism is the decay in activity of neural cells representing earlier events in an episodic

sequence. The second is a decrease in the transmission of activity from the cospan cell representing the episodic segment  $E_{n-1}$  to the cospan cell representing the episodic segment  $E_n$  due to a connection weight strength less than unity, of magnitude  $1/2$ , for example. This decrease becomes compounded at the next stage; for example, in a linear model of activity transmission one could multiply weights and find the activity transmitted is at  $1/4$  its original strength, and so on. The two mechanisms operating together would cause a rapid progression of abstraction going backward along the episodic sequence. Based upon these considerations, we claim that the information concerning the sequence order is explicitly represented in accessible form at each step in an episode.

With holistic access, shared representations, an intrinsic flexibility in recall and, finally, the explicit representation of sequence order, the span-cospan episodic memory model has, as far as we are aware, capabilities that no other episodic model in the literature does. The CNST expresses the correspondence between a span-cospan episodic memory structure in the concept category and the neural architecture for it mathematically as a functor. The functor maps the entire concept structure—sensor stimulus theories, grounding object theories, event theories, and episodic segment theories and the morphisms connecting this hierarchy of concepts—into a category representing a neural network which has all the required neural structure at a given stage of synaptic learning. As noted earlier, functors are discussed in previous papers.

To end this paper, we ask whether any of the neural network structures for episodic memory are consistent with known neural science. For example, what about the human brain and its declarative memory system?

## 7 Proposed Biological Correlates of the Span-Cospan Episodic Memory Model

Based partly upon the span-cospan model of episodic memory architecture as proposed here, and partly upon the available data from years of neuroscience research, we propose the following, speculative biological model of episodic memory. We call it a speculation because it is based upon evidence from neuroscience research on the hippocampus and related structures which is at present, as far as we know, too sparse to confirm or deny the span-cospan model. First, we present a "broad-brush" overview of what is known about the biology, and following that we "take the plunge", describing how we think our categorical span-cospan model of episodic memory explains what the biology is doing.

### 7.1 The biology: a brief overview

For this section, we rely heavily on certain references which have provided us with a relatively clear overview of contemporary structural and functional models of the medial

temporal lobe (MTL) declarative memory system. It is fairly well established that the hippocampus and its interactions with other brain structures are essential for the learning, consolidation, and recall of declarative memory, that is, semantic and episodic memory ([34, 1, 25, 28, 6, 30, 35]). An overall description of the interconnection pathways within the system comprising the entorhinal cortex (EC) and the hippocampus, and other information about the the MTL and declarative memory, is in [35] and the other references cited here (for example, see [26, 32, 31]). Our discussion of neuron-and-synapse function and the dynamics of interaction of the hippocampus with its principle input/output region in MTL relies heavily on the simulation model of [6].

In particular, the entorhinal cortex (EC), lying adjacent to the hippocampus in the medial temporal lobe (MTL), provides the main cortical region for input and output connections to the dentate gyrus (DG) and CA1 and CA3 fields of the hippocampus. As such, EC is the final convergence site (in the sense of Damasio [7]) for multimodal representations of sensory-derived representations ([26, 32]) (visual, auditory, proprioceptive, somatosensory, nociceptory (heat/cold, pain), gustatory and olfactory) and internal information from the amygdala and certain other neocortical regions such as the prefrontal lobe. In fact, the organization of information represented in the MTL appears to be hierarchical in nature [31], with the stimuli to which the cells respond increasing in complexity going from the convergence zones in neocortex to parahippocampal cortex (PHC), through perirhinal cortex (PRC), the EC, then from there to the hippocampus. The hippocampus exclusively contains the full richness of representation for the MTL declarative memory system. Practically all the connections from neocortex to PHC and PRC, from PHC to PRC, and from these structures (mainly PRC) and the others mentioned to EC, have reciprocals, or feedback connections. In the CNST, these reciprocal connections are essential in representing a concept hierarchy together with its dual, a hierarchy of the possible things that the neural network is sensing (e.g., objects in the environment of an organism) or creating (e.g., through forming a plan). There are also reciprocals for the direct connections from EC to the hippocampus. We regard their role as the incremental learning of sequences of related events in the form of the span-cospan system.

Fig. 18 accompanies the following discussion; the references, principally [35], contain much more comprehensive illustrations that more directly show the anatomical layout of MTL. The hippocampus comprises the granule cell layer of the dentate gyrus (DG) and the CA3, CA2 and CA1 pyramidal cell fields (some authors may disagree with our assignment of the DG to the hippocampus proper). The granule and pyramidal cells are called *principle cells* because they are the focus of the neural computations that enact the functions performed by the anatomy; for example, their excitatory projections to other principle cells across the various cell layers and regions bring about the spreading activation of the computations. That said, the many inhibitory interneurons that accompany the pyramidal cells exert control over them and, hence, are essential to the functional roles played by their activity.

The CA1 pyramidal cells and their connectivity with the EC (and probably the subicu-

lum) are the nexus of the MTL declarative memory system and are the main output neurons of the hippocampus. Projections from the EC to the hippocampus are organized in two pathways (the *perforant pathways*). The monosynaptic input pathway (the *direct pathway*) consists of connections emanating from cells within layer III of the EC and synapsing directly upon the distal portion of the CA1 pyramidal cell apical dendrites, farthest from the soma of each cell. The trisynaptic pathway (the *indirect pathway*) emanates from cells within layer II of the EC and reaches the proximal portion of the CA1 dendrites, close to the soma. The trisynaptic pathway is so called because it synapses first on the DG granule cells in the dentate gyrus, then through connections called *mossy fibers* efferent from the granule cells synapses upon the CA3 cell dendrites (there are also direct EC-to-CA3 connections). Finally, the CA3 cells project to the CA1 cells through connections called *Schaffer collaterals*, synapsing on the proximal portion of the CA1 dendrites. This convergence of two input pathways upon CA1 neurons is suggestive of a convergence of two kinds of information, and indeed it is variously described as an input/output or stimulus/context association [26]. Because the layer III cells seem to be more directly related to the current stimulus, the monosynaptic pathway is regarded as having the role of the stimulus, whereas the substantial processing associated with the EC layer II, DG and CA3 cells and connectivity suggests that the trisynaptic pathway takes on the role of a specified context for the stimulus. On the other hand, the timing of the arrival of the respective EC-DG-CA3-CA1 and EC-CA1 inputs to the CA1 proximal and distal locations on the apical dendrite of a CA1 cell, together with the dynamics of interaction between the pyramidal cells and interneurons, seem to suggest that the direct EC inputs trigger a subset of CA1 cells which then determine which of the subsequently-arriving CA3 inputs are relevant (e.g., see [8]). The EC layer III inputs to the selected CA1 cells thus are regarded as outputs of the hippocampus during future recalls, where the CA3 inputs become the inputs. In the latter sense, the two pathways impinging upon CA1 could be associating a spatial and/or temporal context with a stimulus. This suggests a representation of a new event in consistency with the situation in which an episode is occurring, or an association of facts with relevant background knowledge, which can be seen as the modification of a concept representation or the formation of a new concept in declarative memory. Finally, there are connections from CA1 to the deep layers V and VI of EC. Through internal EC interconnectivity, these close the EC-CA1 processing loop.

Two cycles are essential in the timing of the two pathways of the EC-hippocampal system: The theta rhythm (4–7 Hz), which is most strongly implicated in learning, and the gamma cycle (30–100 Hz), using values from Cutsuridis *et al.* [6]. Each theta cycle has two half-cycles within which memory storage and recall alternate, and the gamma rhythm serves as a sort of clock cycle for the hippocampus. The cells responsible for storage and recall of items within a theta cycle are active within a particular gamma cycle. Inhibitory interneurons mediate this process. The details of the dynamics of a theta cycle are important but not essential to this presentation, more a consequence of the present state of the span-cospan model than of current knowledge of the biology. Similarly, we shall omit other details which can be found in the references given.



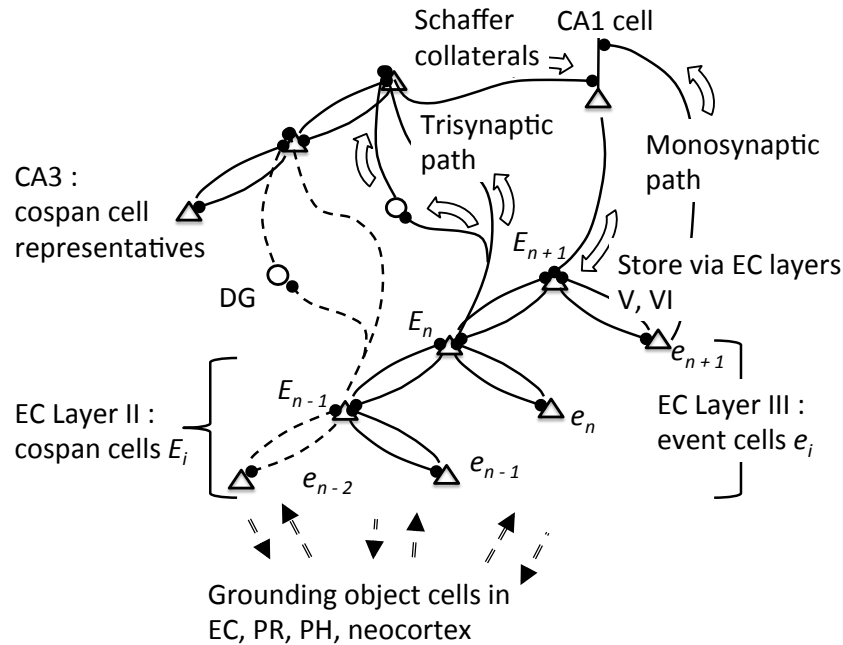


Figure 18: A sketch of a biological model correlate for the span-cospan episodic memory model involving regions such as CA1 and CA3 of the hippocampus and the entorhinal cortex (EC) and other regions of the medial temporal lobe and neocortex. The dentate gyrus is not shown because of the incomplete state of associating its function with the span-cospan model, although it is thought to organize previous episodic segment representations for CA3 which are then blended with the current event at CA1.

The most important fact for relating the span-cospan model to the biology at present is the apparent existence of closed loops in the connectivity between EC and CA1. Connections from CA1 to EC have the characteristic of feedback projections, terminating in the deep layers V and VI [26]. These together with connections from the deep to the superficial layers II and III apparently serve to close parallel EC  $\rightarrow$  CA1  $\rightarrow$  EC loops ([26, 32, 34]). Further, reciprocal connections between neocortex, PRC, PHC, and the EC appear to extend these loops to involve neocortical object representations (The subiculum is also involved and may well be important, but its functionality does not seem to be well established at present.)

## 7.2 A span-cospan biological model

We think this has the following consequences (again, refer to Fig. 18). First, we note that the layer II EC pyramidal cells are fairly large [4] and are the ones projecting to DG in the

hippocampus. These include the EC cospan cells, with the event-encoding and possibly also the blending cells (maintaining continuity between successive events) located in layer III. The function of the DG cells in the trisynaptic pathway is obviously essential but is not yet well enough understood (by us, at least) to include in the span-cospan model. The connectivity of DG with EC layer II evidently provides a sort of working memory for the string of cospan cells in CA3 encoding the currently-active episode. Thus, we posit that the large pyramidal cells in EC layer II for the same recurrent loops that have been observed in CA3, so that the string of segments in an episodic sequence eventually become resident in EC, more or less permanently, yet mutable through recall and recoding. The CA3 cells encoding successive segments of the episode in concert with the cells in EC layer II are connected by recurrent loops as mentioned. It is possible that during future recall operations, Tulving's "encoding specificity" [36] results in the modifications to episodic memories that have been observed and that this recoding originates in the interaction of EC layer III and DG-CA3 inputs to CA1. A selection of one or more CA1 pyramidal cells are active in the current monosynaptic-trisynaptic interaction through EC layer III input at the distal synapses on CA1 apical dendrites and layer II inputs from CA3 through the Schaffer collaterals. Apparently, the selection of CA1 cells is driven by the fact that only those CA1 cells have strong synapses at the connection terminals from a given set of EC layer III cells. This would allow the active sensory representations in neocortex to give rise to the current event representation contributing, via EC layer III, to the encoding of the current episodic cospan representation. The encoding, which combines the current layer III event representation with the preceding episodic segment cospan representation in layer II via CA3, occurs through the monosynaptic-trisynaptic interaction on each selected CA1 cell's dendrite. This encoding at the currently active CA1 cells results in the encoding of a new (the "next") episodic segment cospan representation by a newly-selected cell in layer II. The encoding occurs through synaptic potentiation in the CA1-to-layer V, VI feedback and layer V, VI-to-layer II connectivity. Thus, the completion of the EC-CA1-EC loops enables the cells selected by inputs to layer III and persistent activity in layer II to have their representations combined into new episodic segment cospans in newly-selected layer II cells.

As we mentioned, there is much more biological detail that could be explicated here. However, the foregoing is sufficient to the purpose of relating to the span-cospan model as it exists at present. To summarize the CNST span-cospan episodic memory model, then, we propose that the brains of at least some mammals—rodents and primates in particular—have the following concept structure represented in their central nervous systems. First, concept colimits and possibly limits (see the prior references for the CNST) are represented in neocortex. This and the MTL representations are modeled in the CNST as a system of functors from the concept category to a neural category. Natural transformations between the functors model the interconnectivity of representations in different brain regions via white-matter connections (we have referred to this aspect of representation as *knowledge coherence*; see for example [22]). Different neural categories correspond to the same neural network at different stages of synaptic learning, each neural category be-

ing the target of a different system of functors and natural transformations. Recall that the structure of a neural category depends upon both the neural cells and their connectivity and the current values for the connection weights, or synaptic strengths, of the network. The connectivity structure and connection weights of a neural network determine its possible activity states in response to input stimuli, and this determines the corresponding neural category. The span-cospan model for episodic memory is an outgrowth of these ideas.

The span-cospan episodic memory model is a continuation of the foregoing mathematical model, but especially structured to capture the special properties of declarative memory, such as instantaneous learning as opposed to gradual memory acquisition over time and flexibility of subsequent use. Characteristics of declarative memory storage and recall such as these require a rather special architecture. The hippocampal DG, CA1 and CA3 cell fields and their interaction with the EC and via the EC with other structures fill that requirement. We have proposed a somewhat sketchy, and definitely preliminary, formalization of this via the span-cospan model. We offer the model for consideration by cognitive neuroscientists for their further investigations in episodic memory.

## APPENDIX: Category Theory and the CNST

A category  $C$  contains two kinds of entities: *objects*, which can be anything representable mathematically, and *arrows*, also called *morphisms*, which are relationships between pairs of objects. A morphism  $f: a \rightarrow b$  has a *domain* object  $a$  and a *codomain* object  $b$ , and serves as a directed relationship between  $a$  and  $b$ . There can be many or no arrows in either direction between a given pair of objects, depending upon the category, so it has an underlying directed multi-graph. For example, sets are objects in a category called **Set** in which the morphisms are functions, or many-to-one mappings, between sets. The significance of this notion of a morphism, and what distinguishes a category from a directed multigraph, is that a category has the property of *compositionality*. That is, in a category  $C$ , each pair of arrows  $f: a \rightarrow b$  and  $g: b \rightarrow c$  (where the codomain  $b$  of  $f$  is also the domain of  $g$  as indicated) has a *composition* arrow  $g \circ f: a \rightarrow c$  whose domain  $a$  is the domain of  $f$  and whose codomain  $c$  is the codomain of  $g$  (note that the order of appearance of the labels “ $f$ ” and “ $g$ ” is reversed in the “ $\circ$ ” expression). The composition of functions in **Set** is an example. If  $x$  is an element of the domain of  $f$ , then  $(g \circ f)(x) = g(f(x)) = y$ , yielding an element  $y$  of the codomain  $c$  of  $g$ . Another example, if one were to regard people as mathematical objects, would be the descendant-of relation: if John is a son of Maria and Domingo is a son of John, then John is a descendant of Maria and Domingo is a descendant of John and, therefore, Domingo is a descendant of Maria: the composition of two descendant morphisms yields a descendant morphism. Some references for category theory are ([5, 27, 29, 33, 2]); the books by Lawvere and Schanuel, Pierce, and Barr and Wells are the most accessible to non-mathematicians.

One of the two significant properties of composition is that it is *associative*, meaning that for three arrows  $f: a \rightarrow b$ ,  $g: b \rightarrow c$  and  $h: c \rightarrow d$ , the result of composing them is order-independent, with  $h \circ (g \circ f) = (h \circ g) \circ f$ . The second significant property is that for each object  $a$  there is an *identity morphism*  $\text{id}_a: a \rightarrow a$  such that for any arrows  $f: a \rightarrow b$  and  $g: c \rightarrow a$  associated with an object  $a$ ,  $\text{id}_a \circ g = g$  and  $f \circ \text{id}_a = f$ . Thus, composition distinguishes a category from a directed multi-graph. This fact is significant in making category theory an appropriate mathematical modeling vehicle for knowledge representation and semantics.

Any discussion of category theory involves *commutative diagrams*, graph-like structures containing objects and morphisms within which morphisms along different pathways from one object to another are the same morphism under composition, and *functors*, mappings of one category to another which preserve commutative diagrams. For example, the morphisms  $f: a \rightarrow b$ ,  $g: a \rightarrow c$  and  $h: c \rightarrow b$  form a diagram. The diagram commutes if it happens that  $f = h \circ g: a \rightarrow b$ , that is, if the two pathways from  $a$  to  $b$  through the diagram arrows yield one and the same morphism arrow.

In several papers including [22] and [14], we have shown how colimits model the learning of complex concepts through the re-use of simpler concepts already represented in the connection-weight memory of a neural network. Colimits and functors are both

essential in the mathematical modeling of this idea. We shall omit categorical constructs that have been described in detail in previous publications. This includes the detailed description of limits, neural categories, and functors. However, it is important to include a brief description of concept colimits. Finally, the roles of two categorical constructs, *spans* and *cospans*, in concept representation in neural networks are introduced for the first time here. They will be introduced in the section following as part of the episodic memory model.

## Colimits in a category of concepts

The categorical formalization of concepts and their neural representation begins with a concept category. This is actually a category of *theories* and *theory morphisms*. The theories, expressed in a formal logic, are a mathematical formalization of the concepts which we have been illustrating with figures. The concept morphisms are theory morphisms. These express the subconcept-concept (subtheory-theory) relationships mathematically as symbol mappings which preserve the logical properties of a subtheory as part of a theory. Commutative diagrams and other categorical constructs involving theories are defined accordingly. For example, the concept  $T_{\text{colim}}$  in Fig. 19 is more complex than the others, which is not surprising in view of the fact that the latter are the domains of morphisms for which  $T_{\text{colim}}$  is the codomain. This expresses the inheritance by  $T_{\text{colim}}$  of the information in the theories  $T$  and  $T'$ . Likewise, both  $T$  and  $T'$  inherit the information in  $T_\ell$ . But of course there is more to be said about this hierarchical relationship among the theories, for it is an example of “theory blending” and this involves a very important categorical construct known as a *colimit*.

Briefly, a colimit is a special diagram of concepts and morphisms, part of which is called its *base diagram* and the other part its *cocone*. In Fig. 19, the base diagram lies inside the dashed box. The conelike structure is the cocone, an array of arrows having the base diagram objects as their domains but all having a common codomain lying outside the base diagram. In this case, the common codomain is the colimit concept  $T_{\text{colim}}$ . The colimit concept is often called, simply, the “colimit”. The composite diagram (all of Fig. 19) is called the *defining diagram* of the colimit. What makes this so special?

Mathematically, the defining diagram of a colimit is special because it has two key properties. The two properties are that, first of all, the defining diagram commutes. In the concept category, commutativity means that some base diagram concepts are merged properly. Second, the cocone is *initial*. Initiality means that the colimit object represents only that which is expressed in the base diagram objects and morphisms: It does not add any extraneous information.

Together, the two properties of a colimit ensure that it expresses combinations of items along with the information about how they are “pasted” or “blended” to create a single item. In this example, a line segment  $L$  is an item in a concept which is mathematically

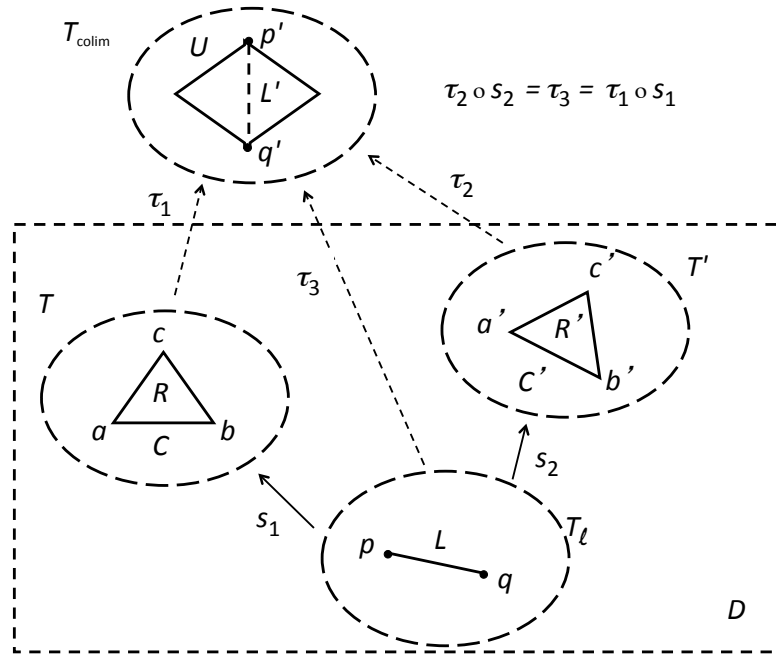


Figure 19: A correct scheme for combining concepts. The base diagram arrows  $s_1$  and  $s_2$  specify that the two concepts  $T$  and  $T'$  are to be merged along a common subconcept  $T_\ell$  in the colimit concept  $T_{\text{colim}}$ .

expressed as a theory  $T_\ell$  in the formal language. The line segment is mapped mathematically via the theory morphism  $s_1: T_\ell \rightarrow T$  to the side  $C$  of the triangle  $R$  expressed in the theory  $T$ . It is also mapped via the morphism  $s_2: T_\ell \rightarrow T'$  to the side  $C'$  of the triangle  $R'$  expressed in the theory  $T'$ . Together, the three theory objects and the two morphisms form the base diagram  $D$  in Fig. 19. The three morphisms  $\tau_1: T \rightarrow T_{\text{colim}}$ ,  $\tau_2: T' \rightarrow T_{\text{colim}}$  and  $\tau_3: T_\ell \rightarrow T_{\text{colim}}$  and the theory object  $T_{\text{colim}}$  at the apex form a cocone. A cocone for the diagram  $D$  has the property that the two “triangle-shaped” diagrams involving the cocone “leg” morphism  $\tau_3$  and either  $\tau_1$  and  $s_1$  or  $\tau_2$  and  $s_2$  commute. The triangle with sides  $\tau_1, s_1, \tau_3$  is highlighted with boldface objects and arrows in Fig. 20.

Since each of the “triangle-shaped” diagrams must commute and share a common edge (the morphism  $\tau_3$ ), we have  $\tau_1 \circ s_1 = \tau_3$  and  $\tau_2 \circ s_2 = \tau_3$ . This implies that  $\tau_1 \circ s_1 = \tau_2 \circ s_2$ ; that is, all three pathways having domain  $T_\ell$  and codomain  $T_{\text{colim}}$  are one and the same morphism when we form the available compositions. This is important: It means, mathematically, that the line segments  $C$  and  $C'$  must be blended as one and the same line segment  $L'$  in the combination of the two triangles  $R$  and  $R'$  because there is only one way that the theory  $T_\ell$  is mapped into the theory  $T_{\text{colim}}$ . Therefore, the theory  $T_{\text{colim}}$  expresses the quadrilateral  $U$  with the two triangles blended together along the line segment  $L'$ . Finally, the initiality property mathematically implies that the apical object  $T_{\text{colim}}$  of the

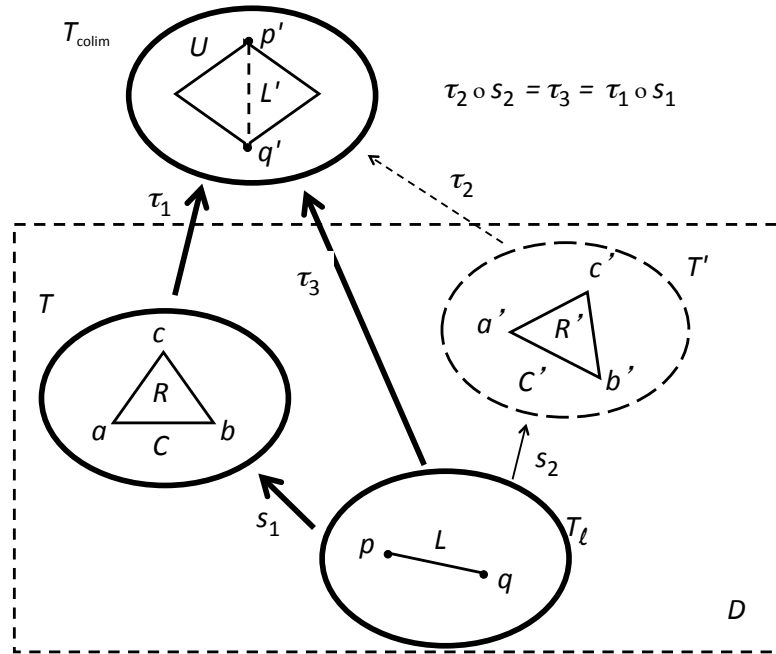


Figure 20: **One of the commutative triangle diagrams for the blending of the two concepts  $T$  and  $T'$  is highlighted in boldface.**

cocone expresses  $U$  and only  $U$ . The colimit concept  $T_{\text{colim}}$  consists of the theories  $T$  and  $T'$  blended around their shared subtheory  $T_\ell$ . Because of the role it plays in the colimit, we refer to  $T_\ell$  as a *blending object*

To put this another way, the colimit object  $T_{\text{colim}}$  is a sort of least upper bound in the concept hierarchy for all concepts that include the concepts in  $D$  blended as described. If a neural network can adapt its synaptic connection weights to form colimits of concepts it already represents, it has a means of incrementally forming new, more complex concept representations from its existing representations through adaptation based upon its current stimuli.

Fig. 21 illustrates a mathematical formalization of the colimit for the third episodic model previously illustrated in Fig. 6 as a conceptual model. One of the commutative triangles is highlighted with boldface arrows. the events  $e_1, e_2, e_3$  are theories along with the objects  $h$  (for human figure) and  $b$  (for ball) inside the base diagram (dashed box) of the colimit. Appropriately,  $e_1, e_2, e_3$  are called event theories while  $h$  and  $b$  are called blending theories. As indicated by the morphisms (solid arrows) inside the dashed box,  $h, b$  are shared subtheories of  $e_1, e_2, e_3$ ; for example, the three base diagram morphisms that have  $b$  as their shared domain are  $p:b \rightarrow e_1, q:b \rightarrow e_2, r:b \rightarrow e_3$ . The cocone consists of the five morphisms  $i, j, k, l, m$  (dashed arrows, except for  $i$  and  $l$  which are

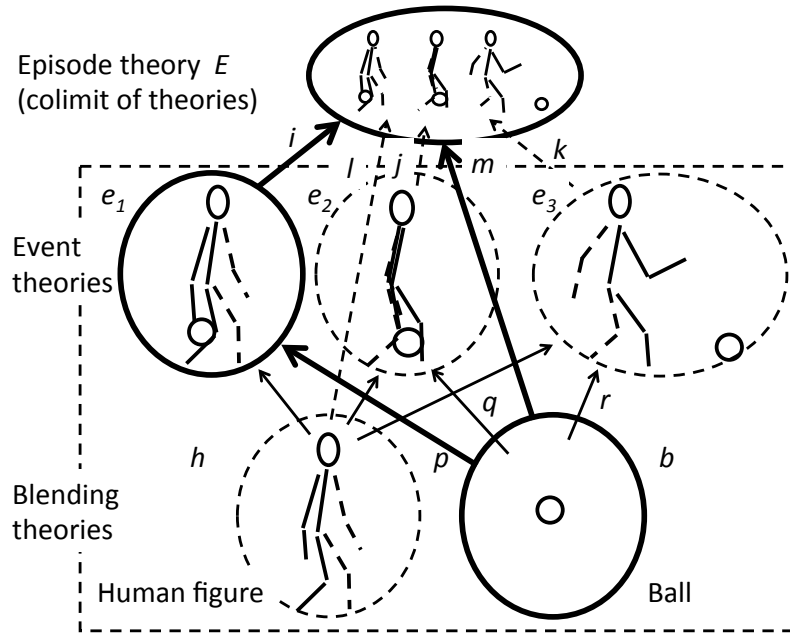


Figure 21: **The third episodic model as a colimit. The episode previously denoted  $e_1e_2e_3$  is now the mathematical theory  $E$ . A typical commutative triangle of theories and theory morphisms is highlighted.**

in the highlighted commutative triangle), where  $i:e_1 \rightarrow E$ ,  $j:e_2 \rightarrow E$ ,  $k:e_3 \rightarrow E$ ,  $l:h \rightarrow E$ ,  $m:b \rightarrow E$ . Here  $E$ , the apical object of the cocone, is the formalization of the episode ( $e_1e_2e_3$  in Fig. 6) as a theory. The requirement for a cocone is the commutativity of the triangular diagrams within the resulting defining diagram; for example, for the highlighted commutative triangle  $i \circ p = m$ , so that the two pathways through the arrows from  $b$  to  $E$  represent the same morphism. The cocone is shown as an initial cocone, meaning as before that the colimit object  $E$  adds no extraneous information.

## Specialization and abstraction

We refer to the concept colimit process, building new concepts from existing ones via blending along shared subconcepts, as *specialization*. In the direction going “up” the hierarchy, the concepts become more specialized because they contain more detailed descriptions and, hence, apply to fewer and fewer instances in the neural network inputs because of their increased complexity.

Another category-theoretic construct, the *limit*, is dual to the colimit. A limit is the common domain of arrows directed *into* its base diagram as opposed to outward and,



hence, these *cone leg morphisms* have the base diagram objects as their codomains. Again the commutativity property characterizes the resulting defining diagram, but the initiality property is replaced by *terminality*. In effect, this means that in the concept category a limit concept is a sort of “greatest lower bound” for the base diagram. Another way to look at this is that it constructs a concept that contains the maximum shared information from an array of concepts. We therefore call the limit process *abstraction*. A limit concept is abstract in that it has fewer details in its description and, hence, applies to a greater number of neural network input instances. In a sense, specialization “pushes out” the concept representation hierarchy of a neural network while abstraction “pulls back” on the hierarchy. In this paper, we focus upon colimits, specialization through concept blending in a hierarchy that begins at the level of simple stimuli and culminates in episodic event sequences.

## Neural categories and functors

Morphisms and commutative diagrams involve categories, and this brings us to the notion of *neural categories* and a notion of *simultaneity* in neural processing. As these notions are addressed in our previous writings, we present here a brief summary.

We posit that the learning of both the specialized and the abstract concepts occurs through the incremental formation of colimits and limits in the brains of animals. This is effected by synaptic potentiation and depotentiation based upon sensory input stimuli and also upon autonomous activity within a brain’s neural network. As described in previous papers, the CNST formalizes all this in expressions of colimits and limits in *neural categories* which consist of *neural objects* and *neural morphisms* (the name CNST appears only in the more recent papers). A neural morphism between two neural objects formalizes the *simultaneous* activity in connection pathways between a pair of neural cells. The connection pathways of a neural morphism  $m$  between two cells  $p_0, p_3$  are shown in Fig. 22. The cells  $p_1, p_2$  are intermediate cells along two of the paths each of which involves two connections in this example. In every instance of persistent activity in any one of the pathways, all three must share in the activity. Notice the feedback connections from  $p_3$  to the other cells defining the morphism; these connections aid in ensuring simultaneity.

Descriptions of neural colimits and limits and functors mapping the concept category into neural categories to formalize the representation of concept structures in neural networks are described in some detail in ([15, 16, 17, 21, 20, 22, 18, 13, 14, 19, 23]). An experiment in applying colimits to model the similarity judgements made by human subjects viewing pairs of visual objects is described in [24]. Applications of concept and neural colimits and limits to improve upon a well-known neural network from the literature are described in ([20, 38, 14]). A type of neural morphism expressing the sequential activation of event representations is described in ([13, 19]). This type of morphism, called a *temporal morphism*, extends the notion of simultaneity to any time interval during which the source and target cells of the multiple pathways of a neural morphism exhibit persistent

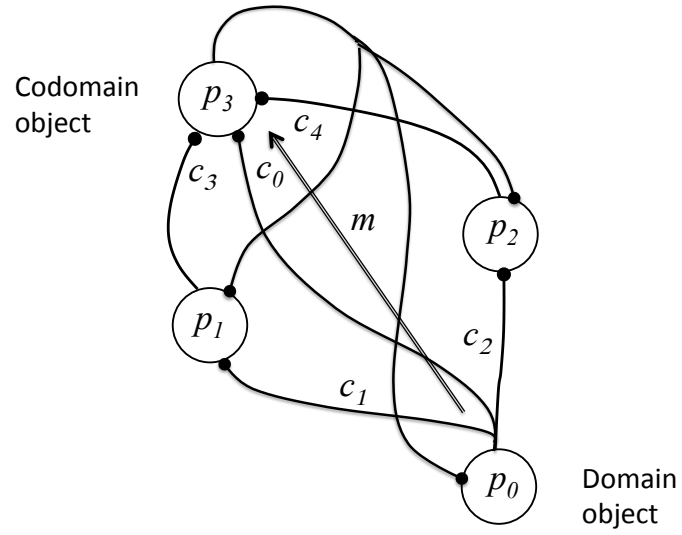


Figure 22: A neural morphism with connection pathways  $c_0, c_1; c_3, c_2; c_4$ .

activity.

## References

- [1] Hisham E. Atallah, Michael J. Frank, and Randall C. O'Reilly. Hippocampus, cortex, and basal ganglia: Insights from computational models of complementary learning systems. *Neurobiology of Learning and Memory*, 82:253–267, 2004.
- [2] Michael Barr and Charles Wells. *Category Theory for Computing Science*. Centre de Recherches Mathématiques, third edition, 1999.
- [3] Lawrence W. Barsalou. Perceptual symbol systems. *Behavioral and Brain Sciences*, 22:577–660, 1999.
- [4] Cathrin B. Canto, Floris G. Wouterlood, and Menno P. Witter. What does the anatomical organization of the entorhinal cortex tell us? *Neural Plasticity*, 2008(Article ID 381243), 2008.
- [5] Roy L. Crole. *Categories for Types*. Cambridge University Press, 1993.
- [6] Vassilis Cutsuridis, Stuart Cobb, and Bruce P. Graham. Encoding and retrieval in a model of the hippocampal CA1 microcircuit. *Hippocampus*, In press, 2009.
- [7] Antonio Damasio. Time-locked multiregional retroactivation: A systems-level proposal for the neural substrates of recall and recognition. *Cognition*, 33:25–62, 1989.
- [8] Joshua T. Dudman, David Tsay, and Steven A. Siegelbaum. A role for synaptic inputs at distal dendrites: Instructive signals for hippocampal long-term plasticity. *Neuron*, 56:866–879, 2007.
- [9] A. C. Ehresmann and J.-P. Vanbremeersch. Information processing and symmetry-breaking in memory evolutive systems. *BioSystems*, 43:25–40, 1997.
- [10] Howard Eichenbaum. Hippocampus: Cognitive processes and neural representations that underlie declarative memory. *Neuron*, 44:109–120, 2004.
- [11] Stephen Grossberg. A theory of human memory. In *Studies of Mind and Brain*, pages 498–639. D. Reidel Publishing Company, Dordrecht, Holland, 1982. 1988 paperback edition.
- [12] H. Gust and K.-U. Kühnberger. Learning symbolic inferences with neural networks. In *Proceedings of the XXVII Annual Conference of the Cognitive Science Society (CogSci2005)*. Cognitive Science Society, 2005.
- [13] M. J. Healy and T. P. Caudell. Temporal semantics: An extended definition for neural morphisms. In *2009 International Joint Conference on Neural Networks*, page paper no. 278. IEEE, INNS, 2009.

- [14] M. J. Healy, R. D. Olinger, R. J. Young, S. E. Taylor, T. P. Caudell, and K. W. Larson. Applying category theory to improve the performance of a neural architecture. *Neurocomputing*, 72:3158–3173, 2009.
- [15] Michael J. Healy and Thomas P. Caudell. A categorical semantic analysis of ART architectures. In *IJCNN '01: International Joint Conference on Neural Networks, Washington, DC*, volume 1, pages 38–43. IEEE, INNS, IEEE Press, 2001.
- [16] Michael J. Healy and Thomas P. Caudell. Aphasic compressed representations: A functorial semantic design principle for coupled ART networks. In *The 2002 International Joint Conference on Neural Networks (IJCNN '02). Honolulu. (CD-ROM proceedings)*, page 2656. IEEE, INNS, IEEE Press, 2002.
- [17] Michael J. Healy and Thomas P. Caudell. From categorical semantics to neural network design. In *The Proceedings of the IJCNN 2003 International Joint Conference on Neural Networks. Portland, OR, USA. (CD-ROM proceedings)*, pages 1981–1986. IEEE, INNS, IEEE Press, 2003.
- [18] Michael J. Healy and Thomas P. Caudell. Generalized lattices express parallel distributed concept learning. In Vassilis G. Kaburlasos and Gerhard X. Ritter, editors, *Computational Intelligence Based on Lattice Theory*, volume 67 of *Studies in Computational Intelligence*, pages 59–77. Springer, 2007.
- [19] Michael J. Healy and Thomas P. Caudell. Temporal Sequencing via Supertemplates. Technical Report EECE-TR-10-0001, Department of Electrical and Computer Engineering, University of New Mexico, July 2010.
- [20] Michael J. Healy, Richard D. Olinger, Robert J. Young, Thomas P. Caudell, and Kurt W. Larson. Modification of the ART-1 architecture based on category theoretic design principles. In *IJCNN 2005 Conference Proceedings: International Joint Conference on Neural Networks. Montreal, Quebec, Canada. (CD-ROM proceedings)*, pages 457–462. IEEE, INNS, IEEE Press, 2005.
- [21] Michael John Healy and Thomas Preston Caudell. Neural networks, knowledge, and cognition: A mathematical semantic model based upon category theory. Technical Report EECE-TR-04-020, Department of Electrical and Computer Engineering, University of New Mexico, June 2004.
- [22] Michael John Healy and Thomas Preston Caudell. Ontologies and worlds in category theory: Implications for neural systems. *Axiomathes*, 16(1-2):165–214, 2006.
- [23] Michael John Healy and Thomas Preston Caudell. The Neural Representation of Concepts at The Sensor Level. Technical Report ECE-TR-16-0001, Department of Electrical and Computer Engineering, University of New Mexico, March 2016.

- [24] Michael John Healy, Thomas Preston Caudell, and Timothy E. Goldsmith. A model of human categorization and similarity based upon category theory. Technical Report EECE-TR-08-0010, Department of Electrical and Computer Engineering, University of New Mexico, June 2008.
- [25] Ole Jensen and John E. Lisman. Hippocampal sequence-encoding driven by a cortical multi-item working memory buffer. *TRENDS in Neurosciences*, 28(2):67–72, 2005.
- [26] Pierre Lavenex and David G. Amaral. Hippocampal-neocortical interaction: A hierarchy of associativity. *Hippocampus*, 10:420–430, 2000.
- [27] F. W. Lawvere and S. H. Schanuel. *Conceptual Mathematics: A First Introduction to Categories*. Cambridge University Press, 1995.
- [28] P. A. Lipton and H. Eichenbaum. Complementary roles of hippocampus and medial entorhinal cortex in episodic memory. *Neural Plasticity*, 2008:1–8, 2008.
- [29] Saunders Mac Lane. *Categories for the Working Mathematician*. Springer-Verlag, 1971.
- [30] A. Rory McQuiston. Cholinergic modulation of excitatory synaptic input integration in hippocampal CA1. *J. Physiol.*, 588(19):3727–3742, 2010.
- [31] Florian Mormann, Simon Kornblith, Rodrigo Quian Quiroga, Alexander Kraskov, Moran Cerf, Itzhak Fried, and Christof Koch. Latency and selectivity of single neurons indicate hierarchical processing in the human medial temporal lobe. *The Journal of Neuroscience*, 28:8865–8872, 2008.
- [32] Pieterke A. Naber, Fernando H. Lopes da Silva, and Menno P. Witter. Reciprocal connections between the entorhinal cortex and hippocampal fields CA1 and the subiculum are in register with the projections from CA1 to the subiculum. *Hippocampus*, 11:99–104, 2001.
- [33] B. C. Pierce. *Basic Category Theory for Computer Scientists*. MIT Press, 1991.
- [34] Lokendra Shastri. Episodic memory and cortico-hippocampal interactions. *TRENDS in Cognitive Sciences*, 6:162–168, 2002.
- [35] Steven A. Siegelbaum and Eric R. Kandel. Prefrontal Cortex, Hippocampus, and the Biology of Explicit Memory Storage. In Eric R. Kandel, James H. Schwartz, Thomas M. Jessell, Steven A. Siegelbaum, and A. J. Hudspeth, editors, *Principles of Neural Science*, chapter 67. McGraw-Hill, New York, fifth edition, 2013.
- [36] Endel Tulving. *Elements of Episodic Memory*. Oxford University Press, New York, 1983.

- [37] Endel Tulving. What is episodic memory? *Current Directions in Psychological Science*, 2:67–70, 1993.
- [38] Robert J. Young, Mike Ritthaler, Peter Zimmer, John McGraw, Michael J. Healy, and Thomas P. Caudell. Comparison of adaptive resonance theory neural networks for astronomical region of interest detection and noise characterization. In *2007 IEEE-INNS International Joint Conference on Neural Networks. Orlando, FL. (CD-ROM proceedings)*, page No. 1541. IEEE,INNS, IEEE Press, 2007.