A Neural Framework for Organization and Flexible Utilization of Episodic Memory in Cumulatively Learning Baby Humanoids

Vishwanathan Mohan
vishwanathan.mohan@iit.it
Giulio Sandini
giulio.sandini@iit.it
Pietro Morasso
pietro.morasso@iit.it
Robotics, Brain and Cognitive Science Department, Istituto Italiano di Tecnologia, Genova, Italy

Cumulatively developing robots offer a unique opportunity to reenact the constant interplay between neural mechanisms related to learning, memory, prospection, and abstraction from the perspective of an integrated system that acts, learns, remembers, reasons, and makes mistakes. Situated within such interplay lie some of the computationally elusive and fundamental aspects of cognitive behavior: the ability to recall and flexibly exploit diverse experiences of one's past in the context of the present to realize goals, simulate the future, and keep learning further. This article is an adventurous exploration in this direction using a simple engaging scenario of how the humanoid iCub learns to construct the tallest possible stack given an arbitrary set of objects to play with. The learning takes place cumulatively, with the robot interacting with different objects (some previously experienced, some novel) in an open-ended fashion. Since the solution itself depends on what objects are available in the "now," multiple episodes of past experiences have to be remembered and creatively integrated in the context of the present to be successful. Starting from zero, where the robot knows nothing, we explore the computational basis of organization episodic memory in a cumulatively learning humanoid and address (1) how relevant past experiences can be reconstructed based on the present context, (2) how multiple stored episodic memories compete to survive in the neural space and not be forgotten, (3) how remembered past experiences can be combined with explorative actions to learn something new, and (4) how multiple remembered experiences can be recombined to generate novel behaviors (without exploration). Through the resulting behaviors of the robot as it builds, breaks, learns, and remembers, we emphasize that mechanisms of episodic memory are fundamental design features necessary to enable the survival of autonomous robots in a real

world where neither everything can be known nor can everything be experienced.

1 Introduction

Our individual experiences play a fundamental role in leading us to exhibit numerous instances of creativity, rationality, and irrationality in our behaviors. Use of experience to go beyond experience is important simply because we all inhabit a continuously changing world where neither everything can be known nor everything can be experienced. To survive, we must integrate diverse chunks of knowledge emerging from our past experiences and exploit them flexibly in the context of the present to ensure smooth realization of our goals. Neural mechanisms associated with the organization and use of memory play a fundamental role in connecting our past with the available present and possible future. Indeed, such processes are of crucial importance for autonomous robots situated in unstructured environments. Simply put, beyond a point, a software programmer cannot travel the journey of an autonomous robot. Instead, like natural cognitive agents, robots must be endowed with mechanisms that enable them to efficiently organize their sensorimotor experiences into their memories, remember and exploit them effectively when needed, and keep learning cumulatively.

This article is an adventurous exploration in this direction using a playful scenario of the humanoid iCub learning to assemble the tallest possible stack using an arbitrary set of objects available to it: learning progressing cumulatively in an open-ended fashion. There are several causal relations that the robot has to learn, remember, and exploit. For example, nothing can be stacked on top of objects like spheres, mushrooms, or pyramids; it is better to stack large objects at the bottom; the color of objects is not a causally dominant parameter while building stacks (but shape and size do matter); and so on. Importantly, there are no unique solutions to be optimized because the solution itself depends on what objects are available to the robot in the present. Sometimes past experiences may have to be combined with explorative actions on a novel object, and sometimes multiple past experiences could be creatively recombined to generate novel behaviors. In general, this playful scenario allows the investigation of the constant interplay between neural mechanisms related to learning, memory, prospection, and abstraction from the perspective of an integrated system that acts, learns, remembers, reasons, and makes mistakes.

1.1 The Context: Connecting Emerging Trends in Neurosciences to Developmental Robotics. A central challenge for brain science today is to causally and computationally correlate the complex behaviors of animals to the complex activitiy in their brains. Here, emerging empirical studies from the neurosciences connect to developmental robotics that attempts

to understand cognition through a model-building approach that reenacts the gradual process of infant developmental learning through robots. The underlying value is both intrinsic (understanding ourselves) and extrinsic (creating a new generation of autonomous systems that can cognitively assist us in the environments we inhabit and create). Mechanisms related to the organization of memory in the brain have been actively investigated over several decades at multiple levels (Squire & Wixted, 2011) and accompanied by propositions of various computational models (Sederberg & Norman, 2010; Chong, Tan, & Ng, 2007). More recent excitement in this topic is attributable to studies that provide converging evidence for shared neural processes underlying remembering past events and simulating future events. Specifically, converging evidence suggests an extensive overlap in the brain networks activated while recalling the past and those engaged during other activities as diverse as thinking about the future (Addis, Wang, & Schacter, 2007; Szpunar, Watson, & McDermott, 2007; Hassabis, & Maguire, 2011; Schacter et al., 2012; Addis & Schacter, 2012), spatial navigation (Burgess, Maguire, & O'Keefe 2002, Suddendorf, 2013), social cognition (Raichle et al., 2001; Frith & Frith, 2012), and perspective taking (Mason et al., 2007). This network of interacting cortical areas has been termed the default mode network (DMN) (Raichle et al., 2001; Buckner & Carroll, 2007; Buckner, Andrews-Hanna, & Schacter, 2008; Suddendorf, Addis, & Corballis 2009; Bressler & Menon, 2010; Welberg, 2012). While the reviews cited go into precise details, functionally there is consensus that the central function of DMN is to construct self-referential episodic simulations, which include reconstruction of past experiences based on contextual cues, simulation of possible future alternatives, evaluating their desirability, and generating goal-directed plans. What is the underlying computational and neural basis of such processes? Can we emulate such mechanisms in a cumulatively developing robot (here, the humanoid iCub)?

Practically, when a humanoid robot like iCub interacts with various objects in its playground, it is the ongoing sequences of actions on various perceived objects, the ensuing consequences, internal body state, and rewards received that mainly form the content of its experiences. While multimodal elements of sensorimotor experience and their temporal order (i.e., microtime: Eichenbaum, 2004) need to be bound together to create an episodic trace, inversely, partial cues arising from multiple sensorimotor modalities must be able to trigger the recollection of relevant past episodic experiences, filling in the remaining missing information—for example, perceiving a pyramid and recalling that it is more rewarding to place it on the top if the goal is to assemble the tallest stack. Since the real world is the main source of partial cues processed bottom up through the sensory and motor streams, clearly there must be a link between subsystems involved in perception and action, how such information is bound together to form the episodic trace, and mechanisms related to recall, prospection, and goal-directed planning.

To functionally implement such a link in a cognitive robot, we took guidance from multiple emerging results. Recent functional imaging studies have shed light on how conceptual knowledge is organized in the brain (Patterson, Nestor, & Rogers, 2007; Martin, 2007, 2009; Meyer & Damasio, 2009). The main finding is that conceptual information is organized in a distributed fashion in property-specific cortical networks that directly support perception and action (and that were active during learning). The same set of networks is known to be active during real perception and action, imagination, and lexical processing. From a computational perspective, we believe that such organization enables information coming from lower processing areas in the cortical hierarchy (involved in, e.g., color, shape, size, action, sound) to generate partial cues to trigger recall of context-relevant past experiences and facilitates learning which properties are causally dominant for a specific task (e.g., the color of objects is not a causally dominant property while constructing the tallest stack). At the same time, information processed by subsystems organized in a distributed property-specific fashion must be coherently integrated both to form the episodic trace and facilitate critical top-down, bottom-up interactions during learning, recall, prospection, and forgetting. Findings from the field of connectomics, specifically in relation to small-world properties, provide valuable clues in this direction. Small worlds are complex systems involving a large number of individual members (e.g., people, neurons, computers) that form tightly knit local communities (high clustering) and are characterized by very short path lengths (globally accessible in a very few hops). Since the seminal work of Watts, Strogatz, and Barabási (Watts & Strogatz, 1998; Barabási & Albert, 1999; Barabási, 2003); it is now established that several complex systems (e.g., the Internet, power grids) exhibit the small world property (Barabási, 2012). More recent attempts to map the large-scale structural architecture of the cerebral cortex (Haggman et al., 2008; Sporns, 2010) have revealed that cortical networks in the brain also exhibit small-world property, specifically pointing to the existence of a small set of hubs (highly connected clusters) that mediate global traffic, facilitating swift integration and in turn forming a core network of interacting cortical areas (Van den Heuvel & Sporns, 2013; Bressler & Menon, 2010).

Guided by these studies, our working hypothesis was that while the distributed property-specific organization brings in a level of functional segregation enabling efficient organization of sensorimotor information, the small-world property enables global integration between them and facilitates the emergence of a *small set of hubs* that together form a higher-level cognitive network (like DMN). In this sense, the proposed neural framework both connects and embodies these emerging trends in neurosciences. As seen in Figure 1, there is a distributed property-specific organization of sensorimotor information, integrated through a small set of hubs. The temporal order of activations in hubs while experience is being gained forms the core content of the robot's episodic memory, duly supplemented by
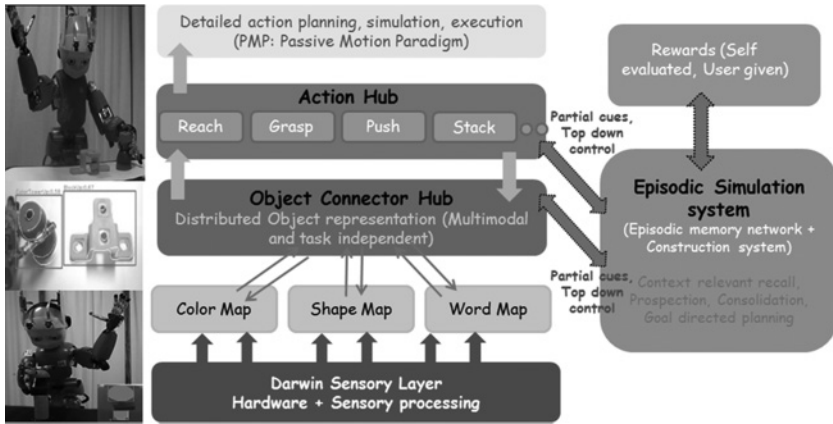
Figure 1: (Right) Block diagram of how information related to perception and action is organized and the link to the episodic simulation system. There is a distributed property-specific organization of sensorimotor information, integrated through a small set of hubs. The temporal sequence of activations in the hubs when experience is originally gained is used to form the episodic memory. At the same time, bottom-up activations in the hub provide partial cues to trigger context-related recall. Activations in the episodic memory network in turn modulate top down the hubs to mediate fundamental processes like combining past experiences with exploration, flexibly connecting multiple experiences in a novel situation, consolidation, and forgetting. (Left) Snapshots of the basic perception-action loop at work: the robot perceiving objects through color and shape and performing basic motor actions necessary to interact with objects to kick-start the learning-memory-prospection-consolidation loop.

mechanisms that enable context-specific recall, combining past experiences with explorative actions, creative plan formation, and forgetting.

1.2 Aims and Scopes. The emerging trends in neuroscience coupled with inherent difficulties faced while enabling robotic systems to exhibit brainlike resourcefulness, purposefulness, and adaptivity in their behaviors call for novel frameworks for cumulative development going beyond conventional engineering and machine learning techniques. In this article, we integrate emerging ideas from neuroscience, to create a brain-guided framework for the organization and creative use of episodic memory in a cumulatively developing humanoid. Both the proposed computational framework and the results are described in a cumulative fashion as learning progresses gradually. The goal for the robot is to learn to build the tallest possible stack given an arbitrary set of objects. Each episode of play may involve objects that have been experienced previously along with novel ones. Furthermore, there is no unique solution, as the solution itself depends on

the objects available in the now and what the robot knows about them. Hence, both learning and reasoning take place in an open-ended setup where the robot is continuously pushed to both exploit what it "knows" from its past experiences in the context of new situations and at the same time learn by exploring novel objects, remember its own mistakes, and perform better in the future. The simple, playful scenario is both novel and fitting to explore complex open issues that lie at the intersection of learning, memory, and prospection planning when any autonomous robot learns incrementally in an unstructured setup. Using this scenario, we explore the computational mechanisms related to organization and utilization of episodic memories in a cumulatively learning humanoid and specifically try to address the following open questions:

- What are the basic neural mechanisms underlying storage and recall of past experiences based on the present context in an open-ended cumulatively learning setup?
- How can remembered past experiences be combined with explorative actions to learn and memorize something new?
- How can multiple remembered experiences be recombined to generate novel behaviors in a new situation (without the need for explorative actions)?
- What is the relationship between the robot's episodic memories and the core subsystems directly involved in perception and action when experience is gained originally?
- The neural basis for forgetting: How do multiple episodic memories compete to survive in the neural space and thus not be forgotten?
- Putting it all together: What are the basic computational processes governing the incessant interplay between learning, memory, prospection, and abstraction in a cumulatively developing system?

We next present a brief overview of the robot and existing sensorimotor infrastructure.

1.3 The iCub Humanoid and the Underlying the Perception-Action Loop. The iCub is a small humanoid robot of the dimensions of a three-and-half-year-old child and designed by the Robot Cub consortium (www.icub.org). The 105 cm tall robot is characterized by 53 degrees of freedom: 7 DoFs for each arm, 9 for each hand, 6 for the head, 3 for the trunk and spine, and 6 for each leg. The iCub body is also endowed with a range of sensors for measuring forces, torques, joint angles, inertial sensors, and tactile sensors in the hand and arms and three axis gyroscopes, cameras, and microphones for visual and auditory information acquisition. With a special focus on manipulation and interaction of the robot with the real world, iCub is characterized by highly sophisticated hands, a flexible oculomotor system, and sizable bimanual work space. Figure 1 shows a block diagram of how the perception-action related information is

organized. At the bottom is the Darwin sensory layer that includes the sensors, associated communication protocols, and algorithms to analyze properties of the objects—mainly color, shape, and size.[1] Results of perceptual analysis activate various neural maps (property-specific SOMs in layer 1), ultimately leading to a distributed representation of the perceived object in the connector hub. (Interested readers may refer to Mohan, Morasso, Sandini, & Kasderidis, 2013, for a detailed description of the sensorimotor organization and learning). The kind of distributed property-specific organization and global integration through hubs is in line with emerging results from neuroscience discussed in section 1.1. What is relevant as far as this article is concerned are mainly that (1) bottom-up processing leads to a distributed representation of the perceived objects in the object connector hub (i.e., "what is it?"), and due to reciprocal connectivity between the hubs and layer 1 SOMS, it becomes possible to learn which properties are causally dominant in a particular task (we explore this issue in subsequent sections).

In relation to the organization of action, there is a subtle separation between the representation of actions at an abstract level ("what can be done with an object") and the action planning details ("how to do"). While the former relates to the affordance of an object, the latter relates to procedural memories of motor skills. The abstract layer forms the action hub and consists of single neurons coding for different action goals like reach, grasp, push, and stack and grows with time as new skills are learned. In this sense, neurons in the top-level action connector hub are similar to canonical neurons found in the premotor cortex (Murata, Fadiga, Fogassi, Gallese, Raos, & Rizzolatti, 1997) that are activated at the sight of objects to which specific actions are applicable. The action hub in turn provides motor goals to the action generation layer that is responsible for the details of motion planning, and synthesis of motor commands to perform the requisite action. The passive motion paradigm framework (Mohan & Morasso, 2011; Mohan et al., 2011), coordinating the iCub upper body is used to generate all motor actions relevant to this article. To summarize, we begin the tallest stack task with a functional identify-localize-reach-grasp loop. Figure 1 also illustrates the link between the core hubs and the episodic simulation system that forms the locus of investigation in this article. The temporal sequence of activations in the hubs when experience is originally gained is used to form episodic memory. At the same time, bottom-up activations in the hub provide partial cues to trigger context-related recall. Activations in the episodic memory network in turn modulate top down the hubs to mediate fundamental processes like combining past experiences with exploration, flexibly connecting multiple experiences in novel situations, and the role

---

[1]The acronym Darwin stands for the ongoing EU-funded project **D**exterous **A**ssembler **R**obot **W**orking with embodied **In**telligence (www.darwin-project.eu).
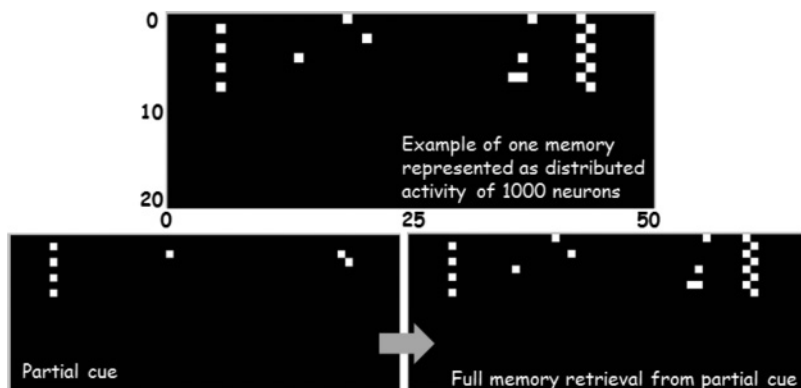
Figure 2: (Top) An example of one memory organized as a distributed activity in 1000 neurons arranged in a sheetlike structure with 20 rows, each containing 50 neurons. Activity in every row may be thought an event in time. The complete sequence is considered as an episode of experience. (Bottom) Reconstruction of the complete episodic memory triggered by a partial cue coming from the environment (e.g., perceiving a cylinder and remembering the past experience of stacking the cylinder on top of a pyramid and failing to receive any reward).

of consolidation and forgetting as learning progresses cumulatively. These topics form the central core of the rest of this article.

2 A Basic Implementation of Episodic Memory

In this section, we briefly summarize a recently proposed excitatory-inhibitory neural network of autoassociative memory (Hopfield, 2008). This network that deals with basic storage and retrieval mechanisms will be taken as a starting point and further enriched in the context of a cumulative developmental learning and reasoning framework where experiences are cumulatively acquired by the robot by interacting with the world; the number of memories grows with time, some eventually forgotten, some consolidated; and multiple memories of past experiences retrieved based on context and goals may have to be causally combined to generate novel creative behaviors. For modeling purposes in the context of this article, we deal with a small patch in the sheetlike neocortex, consisting of 1000 pyramidal cells ($N = 1000$). For simplicity in visualization, the 1000 neurons are organized in a sheetlike structure with 20 rows, each containing 50 neurons. An example is shown in the top panel of Figure 2; activity in every row may be thought of as an event in time and the complete memory as an episode of experience. We are mainly dealing with objects, actions and rewards as these are different aspects relevant to the tallest stack assembly scenario. But in general, anything worth remembering can be represented in

such neural activity. Importantly, in the memory network of 1000 neurons, multiple episodic memories can be encoded and retrieved—for example, playing on day 1 with cubes and pyramids; playing on day 2 with spheres, cubes, and containers; and so on. At the same time, given a partial cue (the robot perceives a red pyramid on day 3), the complete past experience that it had on day 1 (or other days) can be recalled from this partial cue. The memory circuit is characterized by all-to-all connections between the N excitatory neurons (thus, the connectivity matrix is of the order N × N). Memories are stored in the network by updating the connections between different neurons using Hebbian learning. In addition, there is an inhibitory network equally driven by all N excitatory neurons that inhibits equally all excitatory units. A rate-based model is used in which the instantaneous firing rate of each neuron is a function of its instantaneous input current. The procedures for storage and recall are as follows:

● *Memorizing an episode:* Let V_{new} be a one-dimensional vector repre-senting the activity of N ($N =1000$) neurons shown as a 20 × 50 matrix (see Figure 2, top panel). Let T denote the connectivity matrix between the N neurons. Since there are 1000 neurons, the dimensionality of T is 1000 × 1000, which represents the strength of the connection between any neuron i to any neuron j. T is a null matrix to start with, as nothing is known. Con-sider that the episode represented by activity V_{New} has to be stored in the memory network. This is done by updating all the connections T_{ij} between the N neurons in a Hebbian fashion, using a very simple rule:

If $V_i = 1$ and $V_j = 1$, then make $T_{ij} = 1$ (regardless of what its

value was before). Else, make no change in T_{ij}. (2.1)

Starting with $T = 0$, as newer and newer experiences are gained, forgotten, or consolidated, the connectivity matrix gradually is updated dynamically.

● *Network dynamics to remember an episode from partial cue:* To reconstruct the complete memory (say, V_{new}) from a partial cue (e.g., the next day the robot perceives one of the objects it has played with in the past), such partial cues (see Figure 3 bottom left) or initial conditions V are impressed on the memory network and the network is allowed to evolve according to the equation of motion:

$$\tau_{rel}\dot{V}_k = -V_k + \sum_{j=1}^{N} T_{k,j}V_j + I_{inhib},$$ (2.2a)

$$I_{inhib} = g\left(-\alpha^{in} + \beta \sum_{k} V_k\right),$$ (2.2b)

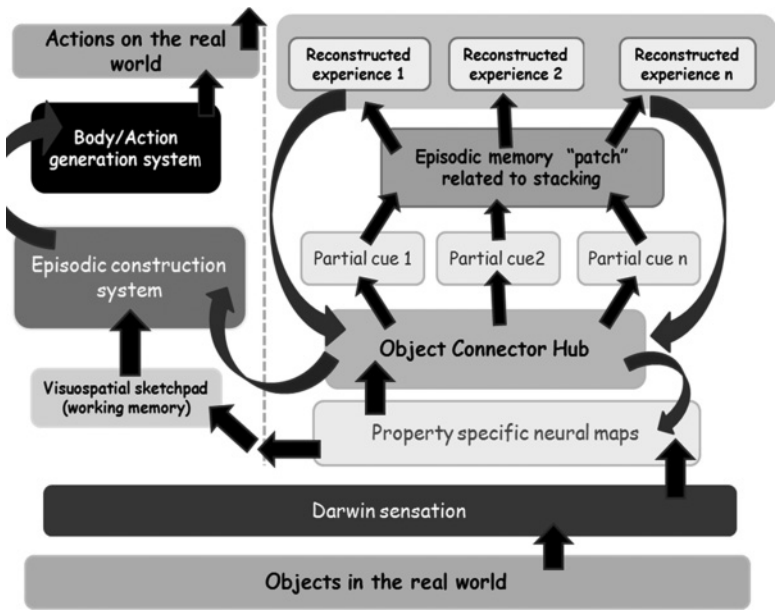$$g(i) = 0, if (i < 0), else, g(i) = i.$$ (2.2c)

Figure 3: A global picture of bottom-up and top-down information flow between the subsystems involved in perception-action, the hubs, and the episodic simulation system. Note the direct link between the hubs and the episodic memory network. Bottom-up information flow (black arrows) gives rise to partial cues triggering the recall of context-relevant past experiences. However, as the robot learns cumulatively, there will be several remembered experiences. Thus arises the need to both filter out the most valuable "team" of past experiences relevant to the present context or goal and at the same time gradually consolidate or forget some of these stored memories. This is functionally implemented by the top-down information flow (the arrows) through a survival of the fittest–like competition mechanism.

V_K is the activity in the Kth neuron. T is the connectivity matrix between the neurons learned using equation 2.1 when the memory is stored in the network. I is the current coming from the inhibition network that is modeled as a single neuron. The function of the inhibitory network is to keep the excitatory system from running away, to limit the firing rate of the excitatory neurons. At low levels of excitation the inhibitory term generally vanishes. For all experiments α^{in} was chosen as 30, τ_{rel} as 1000, and β as 3.5. As seen in Figure 3 (bottom right panel) triggering the memory network with a partial cue and allowing it to evolve in the dynamics described in equation 1.2, it is possible to reconstruct the complete episode. Multiple episodes around 200 to 250 (Hopfield, 2008) can be simultaneously stored and correctly retrieved

in a network of 1000 neurons. In the sections that follow, we start from zero and gradually present results related to:

- How the robot learns cumulatively about different objects and their affordances in the context of enabling it to assemble the tallest stack
- How the robot combines recalled past experiences with explorative actions to learn further or causally connects multiple remembered experiences to generate novel behaviors
- Survival of the fittest like competition between multiple stored experiences and the ensuing process of growth, forgetting and assimilation of episodic memories as learning progresses cumulatively

3 Storing and Remembering Experiences from Partial Cues

In the beginning, the robot has no memory of any past experience ($T =$ Null). Only the robot is able to execute primitive sensory and motor actions like identifying objects in the scene and generating reach and grasp actions. With this starting point, iCub is presented with 2 objects to explore: a green mushroom and a yellow cylinder.

3.1 Content of iCub's Episodic Memories: Top Down–Bottom Up Interactions Between Hubs and the Episodic Simulation System. We noted that episodic memories of iCub are organized as activations in a 1000-neuron patch arranged in the form of a 20×50 sheet. However, we did not clarify what those activations meant. We clarify this here before proceeding with the first episode of learning. In the proposed framework, the content of the robot's episodic memory is the temporal sequence of activity in the object, action hubs, or reward received when experience was originally gained by the robot and encoded in the neural connectivity (using equation 2.1). Every row (in the 20×50 sheet of neurons) is a discrete event in time and the complete sequence an episode of experience (like stacking a cylinder on top of the mushroom and receiving a reward of 0). Hence, there is a direct relation between activity in the hubs and the activity in the episodic memory network. There is both biological grounding (see section 1.1) and computational simplicity behind this proposition. The crucial advantage is that such a scheme allows both bottom-up activation of the hub to generate partial cues, thus triggering a recall of past experiences, and inversely, the possibility of such remembered episodic experiences to modulate the hub's top-down facilitating core processes related to combining past experiences with explorative actions, creative plan formation, and forgetting. Both of these issues will be addressed in detail gradually with numerous examples in this letter. Figure 3 gives a global picture of bottom-up and top-down interactions between the subsystems involved in perception-action, the hubs, and the episodic simulation system. Objects present in the world activate

the object hub, bottom up (black arrows) through the perceptual streams processing color, shape, and size-related information. The distributed activity in the object hub is the source of partial cues. From partial cues, context-relevant past experiences are recalled (using equation 2.2). However, as the robot learns cumulatively, there will be several remembered experiences. Thus arises the need to both filter out the most valuable "team" of past experiences relevant to the present context or goal and at the same time gradually consolidate or forget some of these stored memories. This is functionally implemented by the top-down information flow through a survival of the fittest–like competition mechanism. Only memories that gain top-down control over the hubs enter the construction system and get their content reenacted again through the body thus reasserting the value of their content to the organism. This ensures their longevity. Memories that never win the top-down competition are either consolidated or eventually forgotten. In sum, bottom-up activation of the hub is equivalent to what is there in the world (this is also the input to the visuospatial sketch pad, a component of the working memory that keeps track of things in the present). Top-down activation of the hub is equivalent to what is known from experience and plays a crucial role in facilitating how past experiences are combined with explorative actions on novel objects (see section 4) or recombining multiple past experiences to generate novel goal-oriented behaviors (see section 5) or consolidation and forgetting (see section 6).

3.2 Day 1: Playing with a Green Mushroom and a Yellow Cylinder. In episode 1, the robot is presented with a green sphere (with a flat base like a mushroom; see Figure 4) and a yellow cylinder. Since there is no past experience, the connectivity matrix T is null. Considering that nothing is known, the only option is to explore. Randomly the robot chooses to stack the mushroom on top of the cylinder. The sequence of activation in various neural maps (color, shape, word, and hub) as a function of time when the sphere is stacked on top of the cylinder is shown in the top panel of Figure 4A. The yellow cylinder is identified and localized (sensory streams trigger different property-specific maps processing color and shape information leading to activation in the object hub in relation to the yellow cylinder). Since the goal is to stack and this comes directly from the user, the single neuron coding for stacking in the action hub is activated. Next, attention is focused on the mushroom, activating the hub in relation to the sphere that is stacked on top of the yellow cylinder. Finally, the user/teacher gives a reward (a number entered by keyboard) to the robot. In this case, the reward received is 2 because two objects were stacked successfully. This temporal sequence forms the basis of our first episodic memory, say, EM1, shown in Figure 4B. Every row in the 20×50 memory represents activity in the object hub, action hub, or reward received (that terminates the sequence). In the case of episode 1, the first row corresponds to activity in the object hub
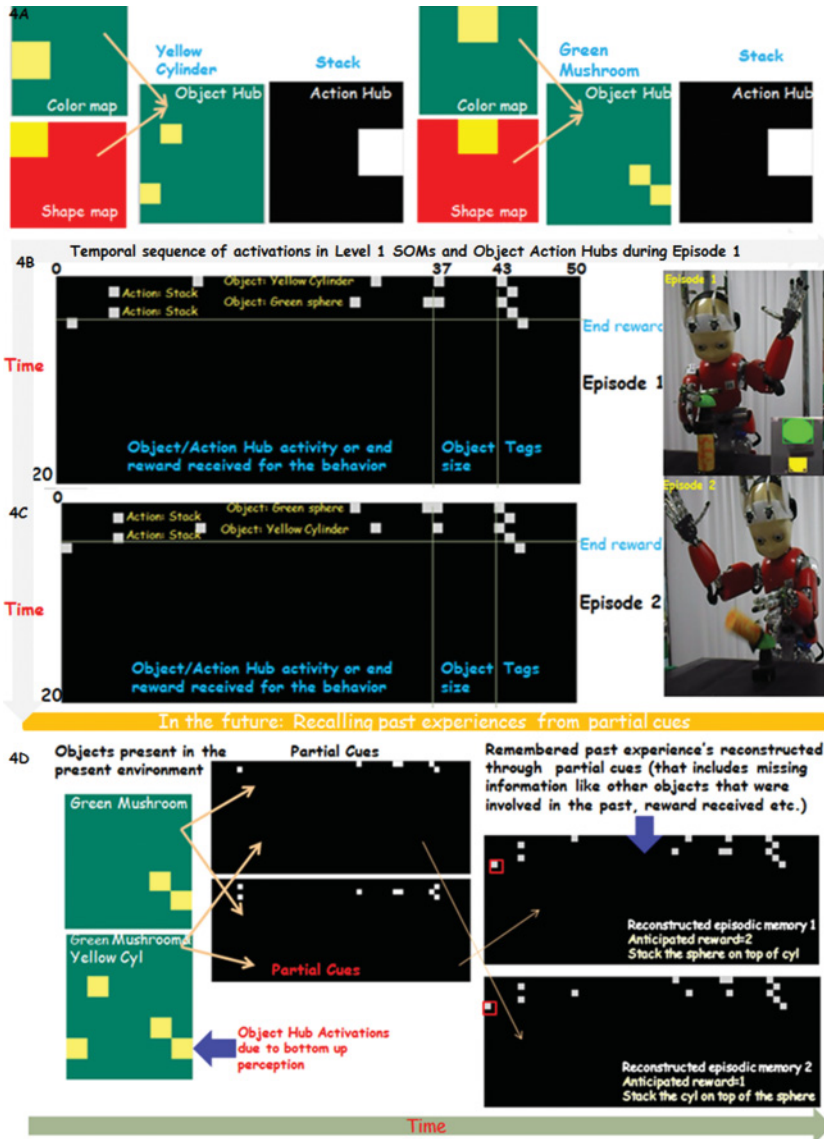
in relation to the yellow cylinder, the second row corresponds to activity in the action hub related to action taken (stack), the third row the activity corresponding to the green sphere, the fourth row corresponding to action hub activity, and the fifth indicating the reward received. Columns 43 to 45 in each row code the identity of the hub to which the information in the row is related (object, action, or value). EM1 is stored in the memory based

on the learning rule of equation 2.1 to update the connectivity matrix T. The robot has not yet exhausted all its explorative options. In episode 2, it attempts to stack the cylinder on top of the sphere (4C). If we compare episodic memory 1 and 2, the difference is that the object representations swap roles (spheres moving to row 1 and cylinders to row 3). This turns out to be a disaster, and the user rewards the robot with just 1 (row 5). Episode 2 is also impressed in the neural network and stored as a new memory. So now the robot has two episodic memories of its explorative experiences: sequences of actions on different objects with reward received at the end.

3.3 Generation of Partial Cues. What happens even after these two initial episodes of explorative sensorimotor experience is interesting. Two cases are shown in Figure 4D. In the first case (scenario 1), a green mushroom is presented to the robot. Perception of the green mushroom generates two partial cues from which the past experiences related to it (episodes 1 and 2) can be recalled from memory (using the dynamics of equation 2.2). In short, what is remembered is that "in the past, I have seen this object coming along with yellow cylinders and stacking the spherical object on the top was more rewarding." While equations 2.1 and 2.2 describe storage and retrieval mechanisms of the episodic memory, we now describe the computational basis of how partial cues are generated. This is a nontrivial problem in a cumulative learning setup where the robot gradually gains

---

Figure 4: (A) Temporal sequence of activations in various neural maps (color, shape, object, and action hubs) when the robot by random exploration stacks the mushroom on top of the cylinder. The content encoded in the episodic memory network is the temporal sequence of activity in the object, action, and value hubs, when the robot gains experiences. (B) The complete temporal sequence of bottom-up activity in the object-action hubs and rewards received when experience is acquired (Panel A) as represented in the 50×20 episodic memory network. (C) The similar encoding of episode 2 where the robot stacks the cylinder on top of the mushroom, receiving the lesser reward (as the tallest stack was not built). Note that the activations in rows 1 and 3 are swapped in Panels B and C, reflecting the temporal sequence of activations when experience is gained. Rewards received are based on the robot's success in building the tallest stack and changes dynamically with the situation. (D) The behavior immediately after two episodes of experience are encoded. Bottom-up perception of the mushroom activates the object hub and fills in partial information in the episodic memory network, leading to recall of past experiences associated with it. What is recalled filling in all missing information is a valuable inference that it is more rewarding to stack mushroom-like objects on the top. The emphasis in the preliminary example is that valuable action sequences are implicitly evident in the episodic recall of past experiences.

new experiences, new memories are formed, and some forgotten, and the same objects may be a part of multiple episodic memories (in combinations with other objects and rewards received). Partial cues basically come from objects perceived in one's immediate environment and action-related goals (to build a stack) that activate the top-level object action hubs (bottom-up information Flow of Figure 3). To generate partial cues in the episodic memory network based on bottom-up activations in the hubs, we introduce three new variables.

The first variable, C, is a scalar counter that keeps track of the number of episodic experiences stored in the memory. C starts from zero and is incremented when a new memory is stored and decremented when memories are forgotten ($C = 2$ at present because two episodes, EM1 and EM2, are stored in memory).

W, the second variable, encodes connections between neurons in the object hub and episodic memory network. If there are $\odot$ neurons in the object hub (42, here) and N neurons in the episodic memory patch (1000, here), then W is a matrix of $\odot \times N$. W is also null to start with and learned in a Hebbian fashion at the same time when a particular episode is stored in the memory. The learning rule is that if a neuron i in the object hub activates a neuron j in the episodic memory patch, then reinforce the connection between them. For example, if neuron h in the object hub and neuron n in the episodic memory patch (in rows that relate to activity in the object hub) are concurrently active and W_{hn} denotes the link between them, then set

$$W_{hn} = 1 \tag{3.1}$$

The net effect is that any time in the future when the same neurons in the object hub are activated due to bottom-up perception (like seeing a sphere on day 2), specific neurons in the 20×50 episodic memory network are activated as modulated by W, giving rise to partial cues to trigger retrieval of the complete past experience (through the dynamics of equation 2.2).

However, one further issue must be dealt with: the connectivity matrix W encodes all possible partial cues that could be triggered by a perceived object. Hence, there is a need to bring in additional context that must have the effect of switching on only a subset of W that relates to the generation of partial cues for retrieving one episodic memory and not all of them at the same time.

This is done by introducing a local parameter, M_{hn}, associated to every W_{hn}, that encodes the identity of the episodic memory during which W_{hn} was adapted (using equation 3.1). For example, if a connection between a neuron h in the hub and a neuron n in the episodic memory patch, W_{hn}, was learned while memorizing episodic memory c, then M_{hn} is set to c. In this way, the connectivity matrix W_{hn} can be further modulated to enable generation of partial cues related to retrieval of specific episodic memories.

If there are C episodic memories stored in T (connectivity matrix of the 1000-neuron-patch encoding past experiences), then the partial cue P_c related to an episodic memory c, based on present context (i.e. object hub activity H), can be generated using

$$P_c = H \times K, \text{ where } K_{hn} = W_{hn} \times \Psi_{hn}$$
$$\Psi_{hn} = 1 \text{ if, } M_{hn} = c \tag{3.2}$$

As depicted in equation 3.2, partial cues generated are a function of three components: the activity in the hub H, which represents objects perceived in the present; W, which encodes all possible partial cues that can be generated by all the objects active in the hub; and Ψ, which has the net effect of switching on a partial cue related to a specific episodic memory c, and not all of them at the same time. When the robot is presented with a green sphere (see Figure 5D), two partial cues are generated—one in relation to EM1 and other in relation to EM2. The dynamics of equation 2.2 triggers a pattern completion process by which the full experience in relation to the partial cue is reconstructed. What is then known at the end is not just information about spheres, but also how they were stacked along with cylinders in the past and what the consequence was (information that was not available in the partial cue itself). Retrieved past experiences are transferred to temporary buffers (working memory) and begin their life in the system. This summarizes the bottom-up information flow of Figure 3.

3.4 Valuable Action Sequences Are Evident in the Episodic Reconstruction. As seen in the right panel of Figure 4D, from the retrieval of the past experiences, it is possible to infer which behavior is more rewarding. This is the simplest example to illustrate the use of episodic reconstruction of the past toward planning actions in the present. One may also envision that the two remembered "past experiences" are competing to survive (as depicted in Figure 3), with the "losers" gradually forgotten. In this simplest case, anticipated reward is the criterion based on which a reconstructed memory of past experience wins the competition. Note that there is no need for an explicit planner; the valuable action sequence is evident in the reconstructed episodic memories that win the top-down competition (in this case, EM1, which anticipates greater rewards). Memories that win the top-down competition manage to reenact their sensorimotor content through the body (in a way, reasserting their value to the organism). Inversely, consistent losers like EM2 may be forgotten as learning progresses incrementally. We elaborate these topics in detail with examples in the sections that follow.

3.5 Causally Irrelevant "Properties" Can Be Eliminated During the Assimilation of Episodic Memories. Before introducing new objects in
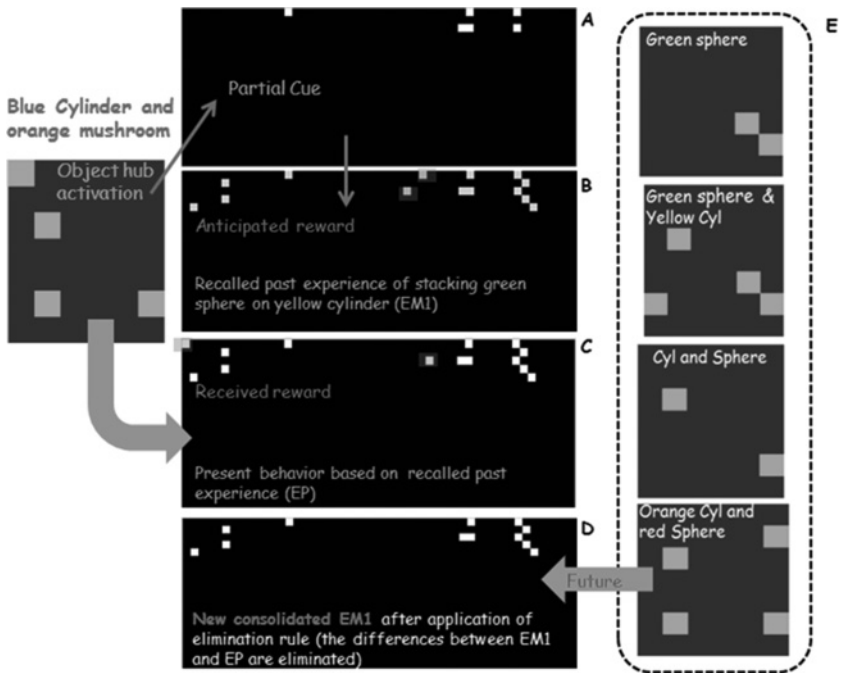
Figure 5: When iCub is presented with a blue cylinder and orange sphere, partial cues are generated (A), leading iCub to recall its past experiences of exploring with the green sphere and yellow cylinders (B). Only the more rewarding experience EM1 is shown (B). (C) The present behavior of the robot of stacking the orange sphere on top of the blue cylinder. Note that this episode is different from the recalled past experience (panel B). The difference between panels B and C is highlighted and mainly corresponds to change in activations related to color in the object hub. Nevertheless, the consequence in terms of reward received is as anticipated. Thus, instead of storing the present behavior as a new memory, EM1 is consolidated by eliminating the difference between recalled past experience and present behavior. In other words, the robot encodes that color of objects do not matter when building the tallest stack. Thus, not every episode of experience is stored. Only those that contain information that is not available in the retrieved past experiences are stored. In the future, when any combination of spheres and cylinders (E) is encountered, the new consolidated episodic memory is remembered and used to guide the present action plan.

the environment, we describe an interesting consequence of distributed property-specific organization of objects in our computational framework. It becomes possible to go beyond object-action and learn which properties are causally dominant in a particular task. How can we abstract which property is causally dominant for a specific task by playing and learning

incrementally with objects in the real world? We briefly address this topic here in the context of stacking. Considering that the robot has past experiences with the green sphere and yellow cylinder, the teacher now presents the robot with a blue cylinder and orange sphere. Bottom-up visual analysis of the scene activates the object hub and leads to the generation of partial cues (see Figure 5A). Note that the generated partial cue is different and contains lesser information as compared to the partial cues in Figure 4D. This is because the objects in the scene that caused the generation of partial cues are also different: they share similarity in shape but not in color. From the partial cue, the past experiences of playing with the green sphere and yellow cylinder is recalled successfully. Only the more rewarding memory (i.e., placing the green sphere on top of the yellow cylinder) EM1 is shown (see Figure 5B).

Although the robot knows nothing about stacking blue cylinders and orange spheres, it knows something about yellow cylinders and green spheres and the fact that it was more rewarding in the past to place the sphere on top of the cylinder. EM1, the more rewarding action sequence, is once again executed, and it turns out that the consequence (in terms of reward received) is the same as anticipated. This new episode generated by the robot is shown in Figure 5C. Note that this is different from the recalled past experience but results in same consequence (the difference, which is highlighted, mainly deals with different activity in terms of the color in the object hubs). Does this new episode also have to be stored in the memory by updating the T matrix? Not really, because we can come up with an elimination rule that compares a reconstructed past experience with the present experience: if a change in property results in no change in anticipated consequence, then the property that has changed is not causally dominant for the task being learned. Hence, the nondominant property can be eliminated.

Thus, instead of storing episode 3, the knowledge that the color of objects does not matter while building stacks can be assimilated into the previously stored episodic memory by inhibiting the ability of the color map to activate the object hubs in the context of stacking (this will ensure that color-related activations do not trigger the partial cues related to stacking). The consolidated memory is shown in Figure 5D. Thus, instead of memorizing the new episode, the robot has implicitly learned that the color of objects does not affect the way they should be stacked. Hence, not every episode is encoded in the memory. Only those that contain information that is not available in the retrieved past experiences are stored (we see this in the next section when cubes are introduced).

4 When Memories of Past Experiences Compete to Become Alive Again: Introducing Cuboids as Novel Objects

Cubes are introduced as novel objects along with spheres and cylinders. Now there is an interesting combination, because the robot has incomplete knowledge: it knows something about cylinders and spheres but has never

experienced the effect of cubes in the context of assembling the tallest stack. This is the simplest case where exploration and experience have to be combined. In the sections that follow, we incrementally propose a number of ideas related to this topic, implementation of the necessary subsystems, and experimental results.

4.1 Top-Down Information Flow: What Does It Take for Past Experiences to Become Alive Again. "Becoming alive again" refers to the ability of a remembered memory trace to get its content reenacted by the actor (body), hence reasserting its value to the organism. To functionally implement this, we introduce a survival of the fittest–like top-down competition between remembered episodic memories to gradually retain the valuable ones and forget consistent losers. The schematic representation of this process is also shown in Figure 3. In our framework, of all the remembered experiences in relation to the present context, only a small subset that manage to gain control over the object hub top down get access to the construction system (and the body). Gaining access to the construction system basically means that either the complete remembered experience or a part of it will be used or reenacted in the "now," hence ensuring the longevity of that memory trace. This in fact is the beauty of top-down and bottom-up driving each other. The only way for a memory to stay alive is to go through the same process that gave birth to it in the first place: control the object hub top down. Whosoever manages to do so enters the construction system, has an opportunity to reenact their content through the body, reasserts their value, and ultimately survives longer. We believe mechanisms related to interleaving of top-down and bottom-up control of hubs may be crucial in the efficient exploitation, growth, and assimilation of memory importantly when acquired by a process of cumulative learning through playful sensorimotor interactions.

A subtle point to note here is that episodic memories of past experiences that manage to enter the construction system may involve actions on several objects that may not actually be present in the now and hence cannot be acted on (e.g., when the robot is presented with a green sphere, the past experience that was remembered involved both green sphere and yellow cylinder: see Figure 4D). To eliminate such elements of the past that are not relevant in the now and extract only the doable actions, we need another subsystem that represents just objects in the now and is not corrupted by top-down activity. To this subsystem we add the visuospatial sketchpad (VSSP), an element of the working memory. Though it has several cognitive functions, we consider for simplicity that VSSP represents perceived objects that are available in the now. VSSP itself is refreshed through bottom-up perception as the robot perceives objects present in front of it and has similar representations as the bottom-up activity of the object hub. The only difference between VSSP and object hub is that VSSP holds only context-dependent information, while object hub may be activated even top down

(by reconstructed memories of past experiences). So an object that is not present in the environment but is internally simulated manages to activate the object hub top down but not VSSP (VSSP in this sense represents objects on which real actions can take place).

4.2 Combining Exploration and Past Experience to Create Plans. Figure 6 shows the temporal evolution of the system when cubes are introduced as novel objects along with spheres and cylinders. For clarity, we break this scenario into three phases.

4.2.1 Bottom-Up Information Flow: From Objects (and Goals) in the World to Remembering Past Experiences Encountered with Them. Figure 6A shows the bottom-up information flow. Objects present in the world are analyzed by perceptual modules, ultimately activating the object hub. Bottom-up activation of the object hub (which indicates the recognized objects in the world) is also transferred to the VSSP. As a result of bottom-up information flow, both VSSP and object hub activity show the presence of three objects. Object hub activations generate partial cues (as in Figure 4D, hence not shown here) leading to retrieval of two past experiences the robot has had in the past: EM1, stacking the sphere on top of the cylinder and receiving a reward of 2, and EM2, stacking the cylinder on the sphere and receiving a reward of 1. To summarize, bottom-up information processing first refreshes the VSSP (what objects are there) and then gives rise to partial cues that lead to the retrieval of relevant past experiences ("what I have done in the past in relation to the present situation").

4.2.2 Top-down Inhibitory Competition Between Multiple Remembered Episodic Memories to Assert Their Significance with Respect to Others in the Present Context. Remembered episodes of the past now compete and inhibit each other in an attempt to control the hub in a top-down way. Which episodic memories (among all those remembered) win the competition are based on two factors:

1. The anticipated reward that could be obtained by the robot if the content encoded by the remembered episodic memory (or a part of it) is reenacted to realize the goal at hand.
2. The exclusivity of the knowledge they encode in the context of the goal. This implies that there need not be one winning past experience; multiple experiences may reach the construction system by controlling parts of the object hub. Hence, the hub can be controlled in a distributed fashion by multiple reconstructed episodic memories. This is because different past experiences may encode different kinds of knowledge that could contribute to realizing the present goal. In such cases, it is like a team of past experiences connected together in the context of the present situation to realize
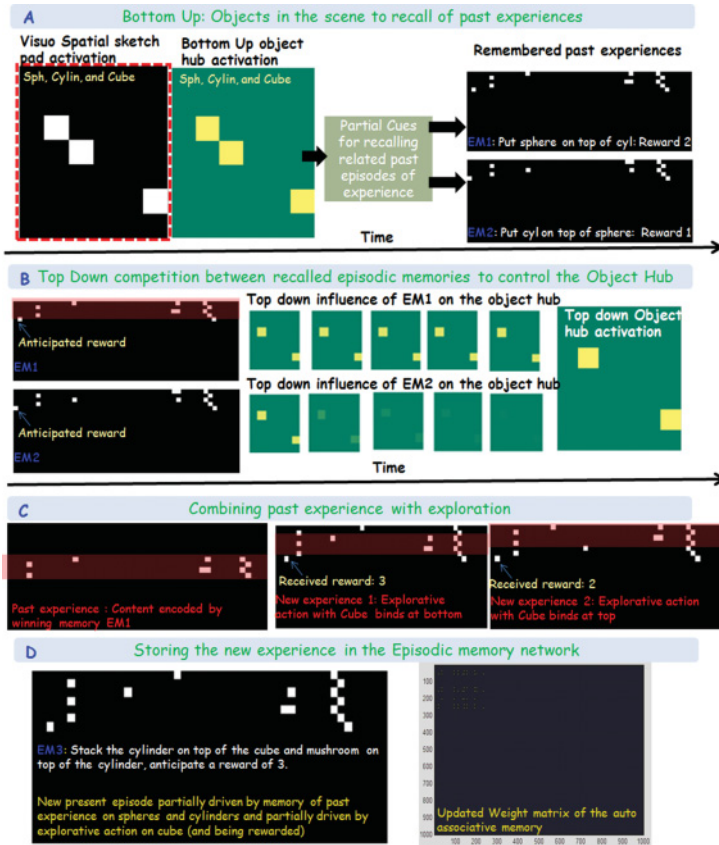
Figure 6: The temporal evolution of the system when cubes are introduced as novel objects along with spheres and cylinders. (A) Bottom-up hub activation through perceptual streams shows the presence of three objects. This information is also copied into the VSSP. Hub activity generates partial cues leading to recall of past experiences (EM1 and EM2 that encode knowledge related to spheres and cylinders). (B) EM1 and EM2 compete to gain top-down control of the hub. Temporal evolution of the top-down influence of these competing memories on the object hub is shown. EM1 emerges the winner in this simple case because of the greater reward fetched by it and gains full control of the hub. Note that bottom-up and top-down object hub activity are different, indicating that not everything about the present world is known through past experience. (C) Action sequence encoded by winning memory EM1 enters the construction system and is combined with two explorative actions: either binding at the end or at the beginning of the chunk that comes from past experience. The episode that fetches greater reward is stored as a new memory (placing the cube at the bottom, stacking the cylinder on top of it and sphere on top of the cylinder: EM3). The updated connectivity matrix T is shown in the right corner.

the goal. This interesting issue is elaborated in the next section when the robot is presented with a large box, cube, cylinder, and sphere.

Top-down inhibitory competition between episodic memories to gain access to the hub is implemented as follows. Let H be the bottom-up activation of the hub (present situation), and let there be m episodic memories reconstructed through partial cues competing to gain top-down access to the hub. Then the individual value or top-down influence Ω_i of the ith episodic memory on the hub (without taking into account effect of other competitors) is as follows:

$$\Omega_i = ((W \times EM_i) \times H) \times R_i. \tag{4.1}$$

The component $((W \times EM_i) \times H)$ determines how much information is known in memory EM_i in the context of the present situation (i.e., bottom-up hub activation H). W is the hub-to-episodic-memory connection learned whenever any memory is stored using equation 3.1. W is the 42×1000 matrix here; EM_i is the reconstructed past experience (1000×1 vector representing activations of the 50×20 episodic memory patch). So the component $W \times E_i$ is a 42×1 vector and determines all possible top-down influences caused by EM_i. To bring in the present context, we multiply every element in $W \times E_i$ with the corresponding element in bottom-up hub activity H (which encodes the present environmental situation). The result is weighted by a scalar R_i that denotes normalized reward fetched by this past experience (e.g., R_i for EM1 is 1 and for EM2 is 0.5) giving rise to Ω_i (a 42×1 matrix that captures the initial top-down influence of memory E_i on every neuron in the hub).

Equation 4.1 accounts only for the influence of normalized reward fetched by E_i in its ability to inhibit other competitors E_j. This suffices for simple cases where all competing memories encode the same knowledge but yield different rewards, as in the present case (both EM1 and EM2 encode knowledge related to cylinders and spheres but yield different rewards). In addition, we also need to take into account case 2: exclusivity of knowledge encoded in the competitors. Hence in our scheme, Ω_i is basically the initial condition with respect to the net top-down influence of the memory EM_i to begin with. Starting from the initial condition, Ω_i related to every episodic memory EM_i evolves in time based on its own value in the present context and inhibitory effects of other competitors EM_j (where $j \rightarrow 1:m$ episodic memories retrieved through partial cues).

The temporal evolution of the top-down control Ω_i exerted by every episodic memory EM_i on the hub (i.e., $\tau\dot{\Omega}_i = F(\Omega_i)$) is implemented by

means of a Euler integration step:

$$\Omega_i(t) = \Omega_i(t-1) + \frac{1}{\tau}F(\Omega_i(t-1)), \text{ where } F(\Omega_i) = \sum_{j=1, J \neq i}^{m} (\Omega_i. * \Omega_j).\lambda_j,$$

$$\lambda_j = R_j \sum_{k=1}^{\eta} (\Omega_j)_k. \tag{4.2}$$

The function F basically generates a net inhibitory effect on memory EM_i coming from all other m competing episodic memories and takes into account (1) how much knowledge is common between any EM_i and EM_j determined by the term $\Omega_i. \times \Omega_j$ and (2) how much more a memory EM_j knows than EM_i in the present context, which is determined by the scalar λ_j (e.g., if the environment contains spheres, cylinders, and cubes and a remembered past experience, say EM3, encodes information related to all of them, then it can inhibit the ones that know only parts like EM1 and EM2). In such cases, EM3 will inhibit EM1 because it knows more in relation to the present context, and both EM3 and EM1 in turn will cumulatively inhibit EM2, which knows the same as EM1 but yields even less of a reward. In the present scenario, only case 1 applies since both remembered past experiences EM1 and EM2 encode knowledge related to spheres and cylinders with EM1 (stacking the sphere on the top) anticipating a greater reward (last row). Figure 6B shows five temporal iterations of the top-down influence of EM1 and EM2 on the object hub. As seen, EM1, which fetches a greater reward, quickly manages to inhibit EM2 and gains complete control of the hub.

4.2.3 Combining Task-Relevant Action Sequence Known from the Past and Combining It with Explorative Actions (to Come Up with New Plans and Learn Further). Note that the top-down activity in the object hub (see Figure 6B, right corner) is different from activity in the VSSP (which holds bottom-up object hub activation). This is because there is no experience related to cubes encoded in the winning episodic memory EM1. In other words, directly comparing the VSSP and top-down hub activity, it is possible to infer that past experience is not sufficient to realize the goal in the present context, thus requiring explorative actions to be combined with what is known from past experience. The inverse of this argument is even more intriguing and will be addressed in section 5.

Now we are left with the problem of connecting the explorative stacking action on the cube with the partial actions sequence that comes from memory of past experience (EM1). This is straightforward: the explorative stacking action binds at either the end or the beginning of the chunk that comes from past experience, so the robot tries out two different action sequences. In the first episode (new experience 1), the robot places the cube
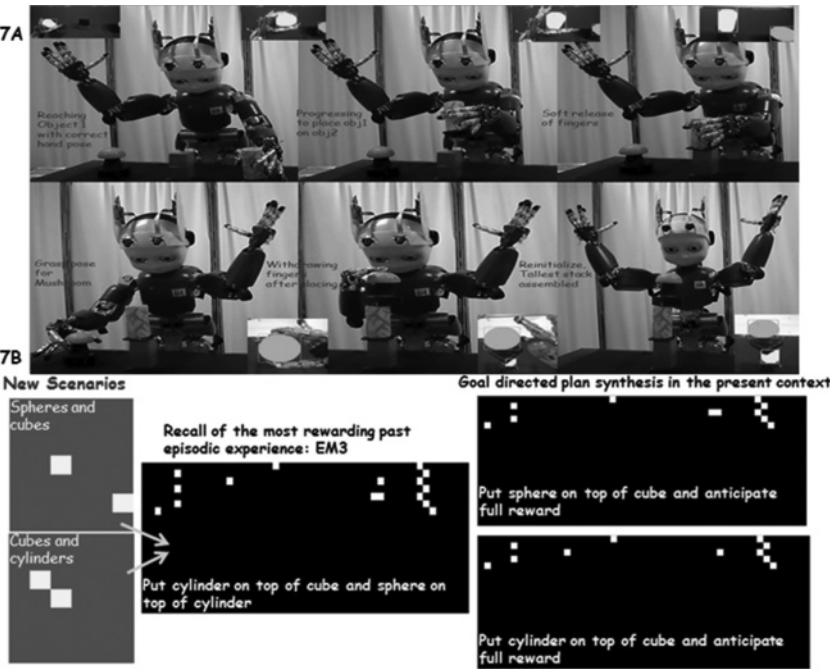
Figure 7: (A) Snapshots of the robot building the tallest stack as a result of combining past experience (with cylinders and mushrooms) and exploration (on the novel object: cubes). (B) Novel scenarios where this new experience is exploited to give rise to novel action sequences. Both of these action sequences generated by the robot are new and related to achieving the goal in the new situation (a novel combination of objects). Neither learning nor planning is needed. The correct action sequence is implicitly embedded in the remembered past experience.

at the bottom (explorative action) and then places the cylinder on top of the cube and sphere on top of the cylinder. It is rewarded fully "i.e., 3" (all objects are stacked correctly). Exploring further (experience 2), the robot tries to put the cube on top of the sphere but does not succeed in getting the full reward (for obvious reasons). The more rewarding experience is now encoded as a new episodic memory (EM 3) as shown in Figure 6C. The connectivity matrix T is shown in the right corner. As seen, beginning from a null matrix, it has slowly started to grow. Figure 7A shows snapshots of the robot combining exploration with past experience to build the tallest stack using cubes, spheres, and cylinders. Figure 7B shows novel scenarios where no further learning is needed to come up with the correct stacking plan. In the first novel scenario, cubes are presented with spheres. The same

bottom up–top down activity flow (see Figures 6A and 6B) ensues, and the new memory EM3 (which encodes the knowledge related to spheres, cubes, and cylinders) controls the complete hub and enters the construction system. Since the cylinder is not represented in the VSSP, the action chunk related to the cylinder is not possible and is deleted from the action sequence encoded by EM3. The robot stacks the sphere on top of the cube and anticipates the full reward. The same applies to the second scenario. Note that both of these action sequences generated by the robot are new, and related to achieving the goal in the new situation (one not encountered previously). Neither learning nor planning is needed. The correct action sequence is implicitly embedded in the remembered past experience that is, the "winning" episodic memory.

4.3 Introducing Large Objects. Before moving to the next level of complexity, we introduce one more object category: a large box. This section may also serve as a case that summarizes all that has been said so far. In the next episode of experience, the robot is given a large box and a small cube. The temporal evolution of the behavior is shown in Figure 8A. The bottom-up information flow leads to neural activations in the object hub and VSSP. From object hub activity, partial cues are generated, reconstructing the most relevant past experience in the context of the present situation. EM3 (the previous episode related to stacking cubes, cylinders, and spheres) emerges as the winner and controls the hub top down. Note again that top-down hub activation differs from bottom-up hub activations because the past experience itself is not sufficient (there is a new object of which nothing is known). The winning past episodic experience enters the construction system where the task-specific chunk is extracted (cylinders and spheres are not present in the world or VSSP, hence vanish); only the cube remains. The robot explores by placing the large box at the bottom and placing the cube on top of it and is rewarded fully (as seen in the last row of explorative binding 1); explorative binding 2, putting the novel object on top of the cube, fails (and hence yields a lesser reward). The more rewarding action sequence is now stored as EM4 by updating the T matrix.

5 How Novel Action Sequences Emerge out of Multiple Past Experiences (Without Exploration)

The user puts all the objects (cube, small cylinder, large box, and sphere) in front of the robot to assemble the tallest stack. Note that iCub has isolated past experiences with all of them. However, it has never encountered all of them together. This is an interesting scenario because none of the past experiences of the robot has the full information to deal with all these objects (all of them have partial chunks of sequences), but if the robot is able to
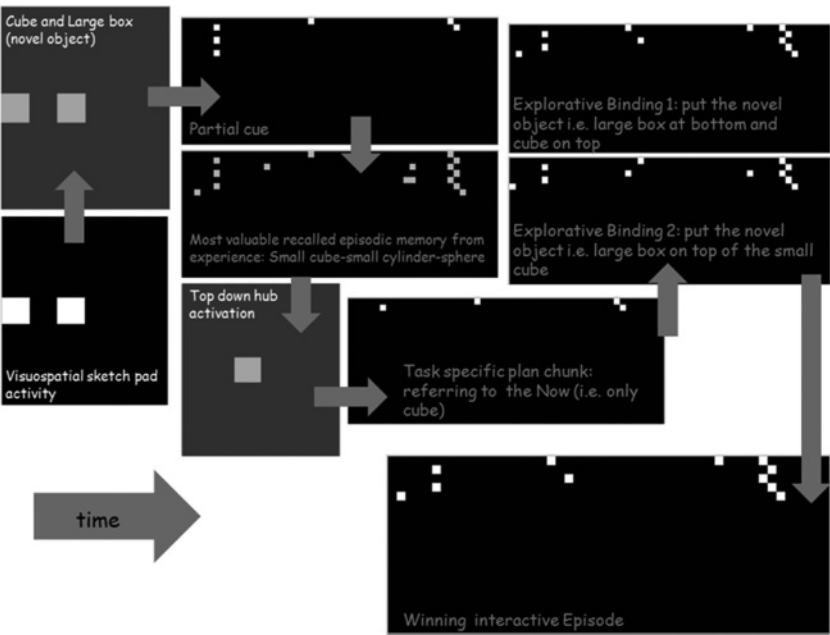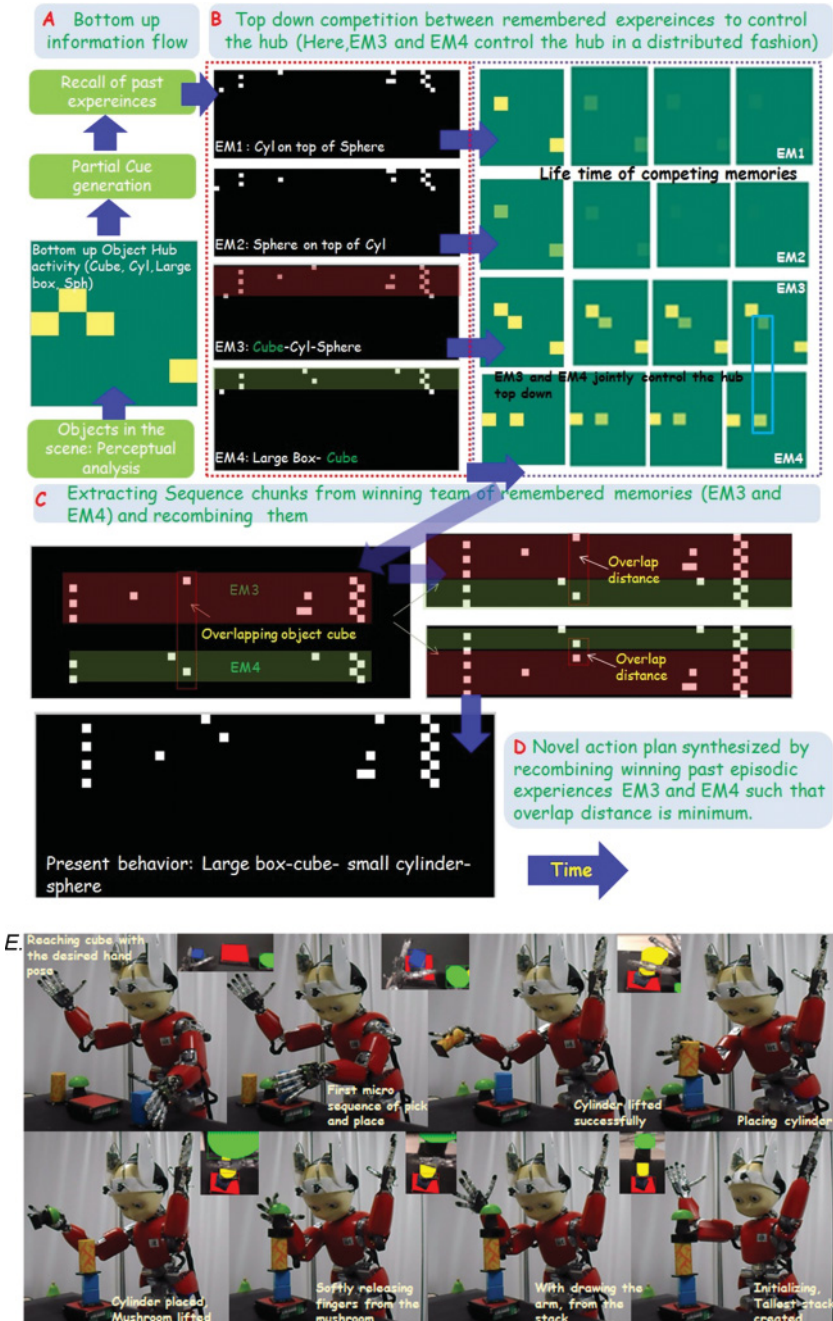
Figure 8: This episode is similar to the one presented in Figure 6, but with one more new object: a large box. The temporal evolution of the system dynamics is shown. Bottom-up information flow leads to the generation of a partial cue and recall of past experiences. EM3 (the previous episode related to stacking cubes, cylinders, and spheres) emerges as the winner and controls the hub top down. Note that top-down hub activation differs from the bottom-up hub activation because the past experience itself is not sufficient. The winning episodic memory enters the construction system where the task-specific chunk is extracted (cylinders and spheres are not present in the world or VSSP, hence vanish); only the cube remains. The robot explores by placing the large cylinder at the bottom and placing the cube on top of it and is rewarded fully (i.e., 2, as seen in the last row of explorative binding 1); explorative binding 2 of putting the novel object on top of the cube fails. The more rewarding action sequence is now stored as EM4 by updating the T matrix.

combine knowledge from multiple experiences to come up with a novel action sequence without any further learning, it is indeed interesting. With the help of Figure 9, we discuss how multiple past experiences remembered in the context of the present can be recombined to generate novel behavior (without any exploration).

The process initiates with bottom-up information coming from the world activating the object hub—generation of partial cues that enable recall of all four past experiences (EM1–EM4) stored so far in the episodic memory. This

A Bottom up information flow

B Top down competition between remembered expereinces to control the hub (Here, EM3 and EM4 control the hub in a distributed fashion)

Recall of past expereinces

Partial Cue generation

Bottom up Object Hub activity (Cube, Cyl, Large box, Sph)

Objects in the scene: Perceptual analysis

EM1: Cyl on top of Sphere

EM2: Sphere on top of Cyl

EM3: Cube-Cyl-Sphere

EM4: Large Box- Cube

Life time of competing memories

EM1

EM2

EM3

EM3 and EM4 jointly control the hub top down

EM4

C Extracting Sequence chunks from winning team of remembered memories (EM3 and EM4) and recombining them

EM3

Overlapping object cube

EM4

Overlap distance

Overlap distance

Present behavior: Large box-cube- small cylinder- sphere

D Novel action plan synthesized by recombining winning past episodic experiences EM3 and EM4 such that overlap distance is minimum.

Time

E. Reaching cube with the desired hand pose

First micro sequence of pick and place

Cylinder lifted successfully

Placing cylinder

Cylinder placed, Mushroom lifted

Softly releasing fingers from the mushroom

With drawing the arm, from the stack

Initializing, Tallest stack created

is because all these memories have some information related to a subset of objects present in the world. This summarizes the bottom-up process, from objects in the world to remembering past experiences encountered with them (see Figure 9A). At the same time, not all of these episodic memories enter the construction system; for this, they have to assert their significance by controlling the hub either fully or partially. The temporal evolution of the top-down influence of these competing memories on the hub is shown in Figure 9B. Note that EM1 and EM2 are completely wiped out in the competition because there are other competitors that know more (in the context of the present situation). EM3 encodes information related not just to cylinders and spheres (encoded by EM1 and EM2) but also about cubes, and hence is a stronger competitor. But in addition to EM3, EM4 also manages to stay alive (it knows something about large objects that none of the others know anything about). Furthermore, we see also that EM3 and EM4 know something in common (i.e., cubes) to control, which are basically inhibiting each other (the overlapping neuron is shown in the box. Note that it is approximately, 50% controlled by EM3 and 50% controlled by EM4). Note that in this interesting case, the sum of the activity imposed top down on the hub by EM3 and EM4 is equal to the activity in bottom-up object hub activation (unlike the cases of Figures 7 and 8, where there was a difference because there was a novel object for which there was no experience).

Figure 9: The cube, cylinder, large box, and sphere are presented. The robot has had isolated experiences with all of these objects, but none of its past experiences encode the complete solution to build the tallest stack now. The figure shows how the robot achieves this by recombining its multiple past experiences; the arrow shows the temporal evolution of this process. (A) Bottom-up perception leads to generation of partial cues and the retrieval of relevant past experiences (note that all past episodes, EM1 to EM4, are remembered as they all contain some information relevant to the present context). Remembered episodic memories compete to control the hub top down (temporal evolution of the top-down influence of these memories on the hub shown in panel 9B). EM1 and EM2 are eliminated, EM3 and EM4 both jointly control the hub partly competing for the common element both of them know, the cube (EM4 exclusively encodes experience related to large objects). The net top-down activity of the hub is identical to the bottom-up activity, which indirectly implies that the complete solution is available in the isolated past experiences (without the need for exploration). (C) Action sequence chunks of EM3 and EM4 enter the construction system, with two ways to bind these sequences. The preferred solution is the one in which overlapping elements of knowledge encoded by different experiences are brought as close as possible (overlaps in this sense playing the role of a subgoal). (D) The final solution: a large box-cube-cylinder-sphere with anticipation of a full reward that is given. (E) Novel sequence of actions generated to assemble the tallest stack using four available objects in the scene.

This implies that the complete action sequence to solve the problem is already available in the isolated past experiences that won the competition, and this applies independent of how many past experiences claim their control over the hub. Either the most valuable action sequence is directly available (in a single episodic memory) or multiple past experiences may have to be combined in a novel fashion to generate a new behavior. In any case, if the net top-down hub activity is equivalent to the bottom-up hub activity (or, equivalently, VSSP), then even if the environment is "novel" (as in the present case), the robot can infer that its past experiences contain enough information to realize the goal (by optimally combining these past memories into a novel sequence). So action sequence chunks encoded by EM3 and EM4 enter the construction system (see Figure 9C), the overlapping object cube shown. The overlaps in knowledge between different remembered experiences are advantageous because they help to connect the experiences together. The construction system employs one simple rule to achieve this: if there are overlaps in knowledge encoded by different winning past experiences, bring them as close as possible. In this sense, the overlapping element is similar to an intermediate subgoal (a point of intersection between two different past experiences).

As seen in Figure 9C (right panels), binding the sequence encoded by EM4 before EM3, the overlaps are closest. This is the one enforced by the construction rule. The other alternative is also shown but does not make cognitive sense because we believe that overlaps in knowledge related to past experiences in general play the function of subgoals (or points where one chunk of knowledge of memory connects to another). When isolated memories of past experiences are combined, a novel sequence emerges (see Figure 9D): stack the large box at the bottom, then the cube, the small cylinder on top of the cube, and the sphere on top of the small cylinder and anticipate a full reward for this. Indeed full reward is given!

6 Effects of Key Parameters, Change in Order of Objects During Cumulative Learning, and Mechanisms for Forgetting

We have gradually described the neural episodic memory of the robot in section 2, explorative learning and recall of relevant past experiences based on partial cues to generate goal-oriented behaviors in section 3, combining explorative actions with past experiences in section 4, and combining multiple past experiences to generate novel behaviors in section 5. In this section, we quantitatively analyze the global behavior of the proposed computational framework under dynamic conditions like change in key parameters of the episodic memory network, change in order of the presentation of objects during cumulative learning, consolidation mechanisms to minimize similar memories being encoded, and mechanisms related to forgetting and the ensuing computational advantages.
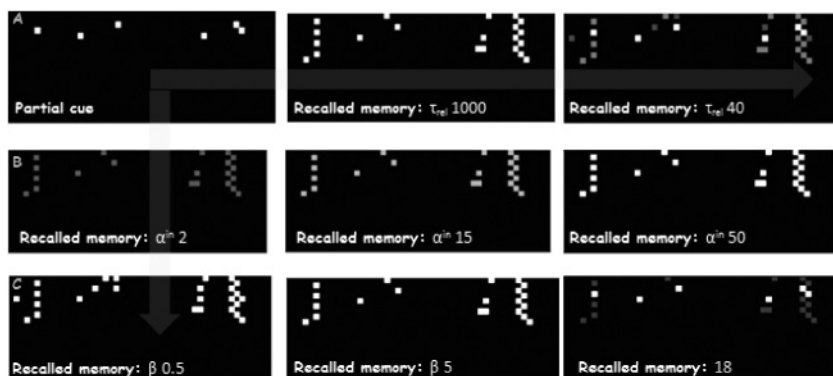
Figure 10: The retrieval performance of the episodic memory network when parameters τ_{rel} (row A), α^{in} (row B), and β (row C) are varied.

6.1 Effects of Change in Parameters in the Episodic Memory Network.

As described in section 2, the dynamics of the episodic memory network is dependent on some key parameters like T (learned network connectivity matrix), the time constant of the relaxation τ_{rel}, parameters in the inhibitory network α^{in}, and β. T is null to begin with and cumulatively learned from scratch. It changes dynamically as new information is encoded or forgotten. The other parameters are constant and are set empirically. Here we briefly investigate the behavior of the episodic memory when these parameters are modified. Figure 10 shows the effects of variations in the parameters and the effect on the retrieval performance of the episodic memory network. Row A (left corner) is the partial cue used to trigger retrieval in all the cases. The middle and right panels of row A show the recalled patterns as a result of the dynamics of equation 2.2, when τ_{rel} is varied. Low time constants ($\tau_{rel} = 40$) adversely affect the convergence, resulting in both reduced activations in the neurons and spurious retrieval (row A, right panel). Nominally, a time constant of 1000 is sufficient to ensure stable and robust recall from partial cues in real time (on a quad core laptop).

The middle row shows the retrieval performance from the same partial cue (row A, left panel), when α^{in} is changed from 2 to 50, keeping τ_{rel} as 1000. As observed, a change in this parameter does not significantly affect the retrieval performance. Instead, increasing it from 2 to 15 has a gradual scaling effect on the activation of the neurons. This behavior was also indicated by Hopfield (2008). Increasing this parameter beyond a certain value (>20) has no significant effect (row B, right column). A nominal value of 5 for α^{in} was chosen for all the experiments reported in this article. On the contrary, the network behavior is more sensitive to β as it has an effect on the inhibitory current to the neurons. Row C shows the retrieval performance for the same partial cue (row A, left panel) when β is varied

from 0.5 to 18, keeping α^{in} constant at 5 and τ_{rel} as 1000. As observed in row C (right panel), very small values of β result in a very low inhibition current. As a result, we can see spurious activations and incorrect retrieval. The middle panel shows the retrieval when β is set to 5, resulting in correct recall of the stored memory from the partial cue. However, further increasing β also abnormally affects the retrieval because of the high level of inhibition. Hence, the parameter β must be neither too small (resulting in very low inhibitory current) nor too large (resulting in very high inhibition). In all our experiments, a nominal value of 3.5 was set for β. The retrieval performance can also be affected as more memories are stored, but as estimated by Hopfield (2008), around 250 episodes can be simultaneously stored and correctly retrieved in a network of 1000 neurons. Further, the proposed framework also includes mechanisms related to both consolidation and forgetting, which have an effect of either merging multiple memories into one or eliminating them altogether.

6.2 Effects of Change in the Order of Presentation of Objects During Cumulative Learning. While section 6.1 dealt with variations to key parameters in the episodic memory effect on recall of encoded experiences, we now go to the next level: change in order of presentation of different objects during cumulative learning and the resulting effect on the behavior of the robot. Through Figure 11, we also revisit sections 3 to 5 (exploration, combining past experiences with exploration to learn further, combining multiple past experiences to generate novel behavior) when the order of presentation of objects to the robot is changed. Rows A to D show four cases of different orders of presentation of objects and the resulting behavior of the robot under situations described in sections 3 to 5 (columns in Figure 11). "EM" stands for episodic memory encoded during the particular stage. For example, in row C, at the beginning the robot is presented with a large box and a cylinder leading to the formation of EM1. In the next episodes, cuboids and mushrooms are introduced as novel objects (with the robot already having past experience with a large box and cylinder) leading to the formation of EM2 and EM3. The next column shows the behavior when all the objects are presented together to construct the tallest stack. In this situation, multiple past experiences have to be combined to generate novel behavior (see section 5). In this case, EM1 and EM3 win the top-down competition and control the hub. Merging EM1 and EM3 through the process described in section 5 leads to generation of the novel behavior. As also seen in the other cases (row A, B, and D), change in the order of presentation of objects mainly affects the content of the episodic memory encoded during the learning process. Despite this, the novel behavior generated by combining multiple past experiences to construct the tallest stack using all the objects is the same (all rows, right column).
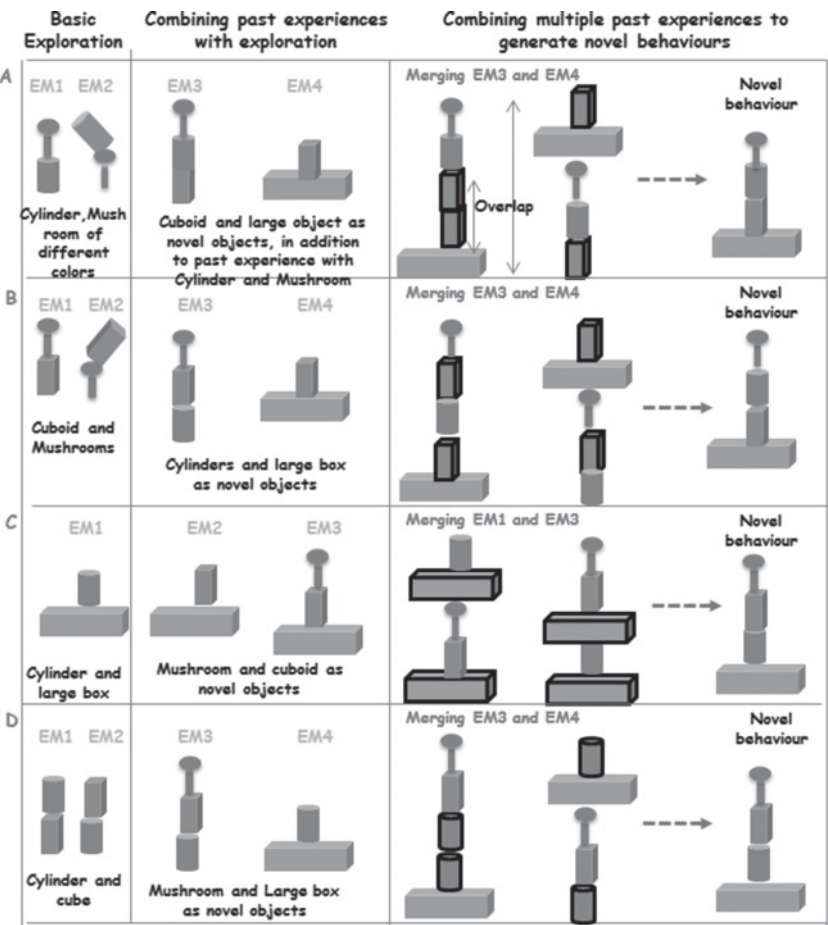
Figure 11: Rows A–D show the gradual progression of learning and formation of episodic memories (EM) with the order of presentation of objects varied. Each row is divided into three phases: basic exploration, combining past experience with exploration, and combining multiple past experiences to generate novel behavior (that captures the essence of sections 3–5). In the right column, despite the changes in order of presentation of the objects, behaviors generated to build the tallest stack in different conditions do not change.

6.3 The Computational Advantage of Forgetting: When, Why, and What. As evident in Figures 6 to 9, memories related to episodes 1 and 2 (i.e., Figure 4) no longer win the top-down competition to control the object hub and get their content reenacted. New episodic memories that in fact originated through their support now exert greater influence on the hub,
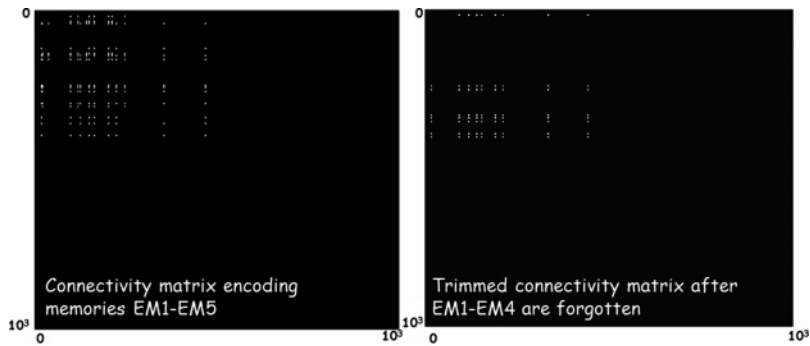
Figure 12: Starting from null, the left panel shows the growth of the connectivity matrix encoding memories EM1–EM5 acquired cumulatively by the robot. The right panel shows the connectivity matrix after EM1–EM4 is gradually forgotten and only the memory trace EM5 encoding the cumulative knowledge of all the experience gained so far is retained.

inhibiting them. In the proposed framework, memories that consistently lose the competition are forgotten because there is a new "competitor" that overshadows them by not only encapsulating the knowledge they encode but going beyond and extending the knowledge (to newer objects experienced cumulatively). Since EM5 encapsulates all the knowledge related to large objects, cubes, cylinders, and spheres, it is retained and all others (EM1–EM4) are forgotten. Figure 12 shows the T matrix before and after the assimilative process (forgetting EM1–EM4 and storing only EM5). As seen, the result is that now we have a trimmed T matrix as compared to the previous case (where all memories were stored in the connectivity matrix). This is because now there is one big memory encapsulating everything instead of five different isolated sequences. So what is the computational advantage of such forgetting and assimilation? When should it take place? The "when" part is when older ones no longer win the access to the hub because a new competitor encapsulates the knowledge they encode and goes even beyond. Now regarding "what is the advantage" part of the question, there are two central advantages:

1. Forgetting decreases patterns that are too close to each other, making retrieval more efficient (when triggered by partial cues) and increasing the storage capacity of the episodic memory network;
2. Reduces computational load: Instead of retrieving several isolated experiences that then compete against each other top down to control the hub, there is a minimal set of a few winners that encode all necessary information to synthesize any goal-directed behavior (the

global system at the same time open to further learning and formation of new memories).

Thus, forgetting is indeed advantageous, and we have further shown "when" memories are forgotten (i.e., when they do not win access to hubs for long time and hence are unable to reenact their plans through the body), how they are forgotten (by duly updating the connectivity matrix of the memory network), and hinted at the computational advantages of such processes (i.e., more efficient retrieval, increase in storage capacity, reduction in computational load).

After the phase of forgetting, the memory network contains only one episodic trace, EM5, that encodes the cumulative knowledge of everything experienced and learned so far. Before we conclude, we add one more object that is commonly found: the container. The robot is presented a combination of cubes and containers. Figure 13A shows the resulting behavior, which is similar to what happened when cubes and large objects were introduced (see Figure 5). Bottom-up activation of the object hub gives rise to partial cues. Importantly, note that now there is only one winner that encodes a large sequence (because of the effect of forgetting). Only one winner means no competition to control the hub top down; EM5 is the winner and gets access to the hub. However, comparison between bottom-up and top-down hub activation indicates that not everything is known from past experience. The task-relevant chunk of the EM5 gains access to the construction system and becomes part of the plan, with the rest to be developed by exploring with the new object. Putting the container on top of the cube leads to greater reward. This episode of playing with containers and cubes becomes our new memory, EM6. So now there are two episodic memories: one that encodes knowledge about large cylinders, cubes, cylinders, and spheres and one that knows something about cubes and containers. Figure 13B presents the response when the robot is presented now with four objects in a novel combination: a large object, a small cylinder, a small cube, and a container. Similar to the situation encountered in section 5, the robot has isolated experiences with all these objects but none of the memories encode the complete solution. The hub is controlled in a distributed fashion. As we can see, the sum of the top-down hub activity imposed by competing episodic memories (EM5 and EM6) is equal to the bottom-up hub activity (resulting through bottom-up sensory stream). This implies that the complete knowledge to solve the problem is embedded in the isolated past experiences if recombined creatively. A novel behavior emerges and brings the full reward!

7 Discussion

"It's a wrong sort of memory that only works backwards," remarked the White Queen in Lewis Carroll's *Alice in Wonderland*. Interestingly, emerging
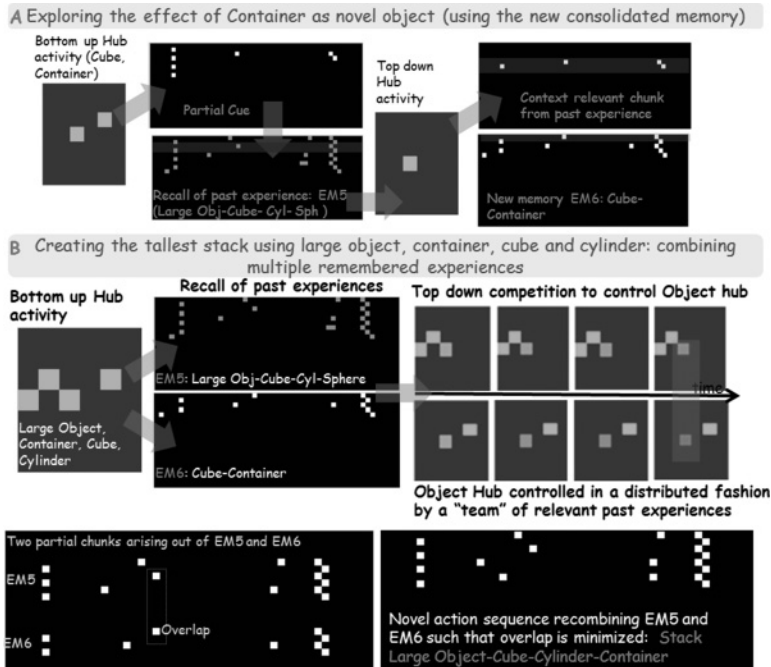
Figure 13: (A) The temporal evolution of the behavior when containers are introduced as novel objects along with cubes. Partial cues lead to recall of EM5, the only existing experience in the new consolidated memory network. Note that top-down hub activity as modulated by EM5 differs from bottom-up hub activity, indicating that not everything is known from past experience (similar to Figure 6). The task-relevant chunk of EM5 needs to be combined with explorative actions on the container. Stacking the container on top of the cube brings greater reward. This experience is encoded as a new memory, EM6. (B) The behavior when a large object, a cube, a cylinder, and a container are presented together. As seen, both context-relevant experiences EM5 and EM6 are recalled and now control the hub in distributed fashion (as described in section 5). Note that in contrast to Figure 13A, the sum of the down hub activity imposed by competing episodic memories (EM5 and EM6) is equal to the bottom-up hub activity. This implies that the complete solution is available in the isolated past experiences if they are recombined flexibly. A novel behavior emerges. The robot places the cube on the large box, the cylinder on the cube, and the container on top of the cylinder and anticipates a reward of 4.

trends in neurosciences, in particular the discovery of the DMN, now provide converging evidence suggesting an extensive overlap in the brain networks activated while recalling the past and those engaged during activities as diverse as simulating the future, goal-directed planning, perspective

taking, and some forms of spatial navigation. Such a perspective urges viewing memory not just as past oriented but also future oriented—in other words, as a key component of the prospective brain that actively facilitates simulation of future events, formation of flexible plans, and predictions— the essence beautifully captured in Carroll's novel.

While the computational bases of such mechanisms are still elusive, it is imperative that cognitive robots envisaged to assist us in the unstructured environments we inhabit must be equipped with a powerful biologically inspired memory architecture that allows them to remember their past experiences based on context and exploit them flexibly in novel situations. This article was an exploration in this direction, capturing in a simple way the constant interplay between neural mechanisms related to learning, memory, prospection, and abstraction in a cumulatively developing humanoid robot. In this section, we briefly summarize the general perspective we have gained by teaching a baby humanoid to build the tallest stack and how we are taking the framework ahead in near future.

7.1 "No Traveler, No Travel." In a seminal article Tulving (1972) suggested that retrieval of one's own past experiences involves a conscious reliving of past events, like a mental journey into the past. In recent years, evidence has accumulated that such time travels are also responsible for simulating the possible future in order to facilitate flexibly goal-directed behaviors in the present. Indeed, if the sole function of episodic memory mechanism was to record the past, it might be expected to function in a reproductive manner, similar to a video recorder (Suddendorf & Corballis, 1997). Instead, it functions in a constructive fashion, where multiple experiences can be retrieved, eliminated by competition and at the same time creatively recombined to facilitate the survival of the "traveler" in the dynamically changing unstructured world. While much of the discussion on mental time travel has been centered around whether nonhuman animals possess this ability (Tulving 2002; Suddendorf & Corballis, 2007), attempts to emulate such mechanisms on cumulatively learning embodied robots have been negligible to our knowledge. Such an exercise may give rise to novel computational insights and at the same time aid the creation of better cognitive artifacts. This article goes in this direction. Both when a stack of objects gets destroyed or built successfully, iCub learns something from them and uses such memories in the future. A time travel to its past explorative interactions with the world and resulting consequences enables it to do so. At the same time, had it not experienced these events gradually in time through direct sensorimotor interactions, it would not have been able to encode such diverse experiences into its episodic memory or use them in the future. Time travel needs an active traveler, and this directly resonates with the concept of embodiment and the emergence of representational content as a consequence of sensory-motor interactions of the agent with its environment (Wiener, 1961; Gibson, 1966, 1979; Maturana &

Varela, 1980; Clark, 1997). We show in this article how such continuous exchange of signals between the brain, the body, and the environment leads to the formation and flexible use of episodic memories in an embodied robot. Given the diversity of the world, the travels of different travelers are indeed unique, and this is reflected in the diversity in individuals behaviors and preferences. Some of the facts may eventually get assimilated into semantic knowledge (iCub learns that color of objects do not affect construction of tallest stacks). But the diversity in our behaviors is in many ways attributable to our own unique episodic experiences. This may reflect also in the behaviors of different iCub robots, each learning cumulatively and guided by its own episodic memories. In general, just as we all have to travel our own journey, software programmers cannot travel the journey for an autonomous robotic assistant expected to inhabit an unstructured world. Instead they must keep learning cumulatively in time and use their experiences effectively in the future (Georg Stork, 2012). In this context, mechanisms related to episodic memory as addressed in this article serve as a central design feature of the prospective brain, and there is a need to push further the state of the art in relation to creation and use of such mechanisms in cognitive robots.

7.2 Why Top Down and Bottom Up Must Share Neural Substrates. A central feature in our computational framework is the innovative use of top-down and bottom-up information flows that share neural substrates. Numerous studies from functional imaging and embodied cognition provide direct evidence for this (Hesslow, 2002; Grafton, 2009; Martin, 2009; Bressler & Menon, 2010; Gallese & Sinigaglia, 2011). However it is not clear what the computational advantages are, how cognitive architectures for embodied robots must exploit this idea, and how much of the computational and neural substrates are eventually shared (this is the more recent debate between hard embodiment vs. soft embodiment; see Martin, 2009). In our framework, the higher-level maps related to perception and action (see Figure 1) can be activated both top down and bottom up (while both early stages of perception like sensory processing and late stages of action at the level of motor commands are not involved). In general, sharing of computational substrates between top down and bottom up gives rise to two main advantages that we have exploited in our framework. First, it simplifies comparison between what has been experienced in the past (i.e., reconstructed through memory) with what an embodied agent is presently experiencing, since both mechanisms are brought down to a common platform (i.e., the shared computational/neural substrate: hub). Such comparisons play a crucial role in both inference and assimilation. The former utility is fairly straightforward: the resonance between top down and bottom up directly indicates that the world is working as anticipated (and the inverse is true if there is dissonance). In other words, sharing of neural substrates between top down and bottom up can be effectively used to close

the loop between learning and reasoning in an open-ended setup: more learning leading to better reasoning, inconsistencies in reasoning leading to greater learning. As seen in section 3.5, direct comparison between remembered past experience (of stacking green sphere on yellow cylinder) with the present behavior (of stacking blue cylinder and orange sphere) and their resulting consequences is sufficient to infer that color is not a causally dominant property as far as the goal of creating the tallest stack is concerned. Hence, instead of storing the new episode in the memory, the ability of color map to activate the object hub in the case of stacking was reduced.

The second application of this idea in our article is subtle and relates to control of the object hub bottom up by real perception and top down by multiple competing episodic memories. In our framework, top-down activation of the hub when compared with the bottom up gives rise to three crucial pieces of information as demonstrated by numerous examples in sections 4 to 6:

1. It clusters what is known from past experience about the present situation and what is unknown because of novelty in the environment. This facilitates combining past experiences with explorative actions to learn further (see Figures 6, 8, and 13).

2. It separates out experiences that are valuable in the present context from the set of all remembered episodic memories, which are competing to control the hub top down. The bottom-up activation of the hub represents the present context and helps generate of partial cues that lead to the retrieval of multiple related past experiences. Which past experiences win the competition and control the hub top down are based on the anticipated rewards they fetch (information that is filled in during episodic reconstruction) and the exclusivity of the knowledge they encode (i.e., knowing something about the present situation that no competitor knows). In sum, memories are reconstructed bottom up through partial cues because they are relevant in the present context. But this is not enough; to survive, they have to compete and demonstrate that they are valuable in comparison with others. The hub is the arena where both top-down and bottom-up processes culminate.

3. If the net top-down hub activity is equivalent to the bottom-up hub activity, then even if the environment is novel, the robot can infer that its past experiences contain enough information to realize the goal (sometimes though generation of a novel behavior). Either the most valuable action sequence is directly available (in a single episodic memory: see Figures 5 and 7) or multiple past experiences may have to be combined in a novel fashion to generate a new behavior (see Figures 9 and 13). In the latter case, it is like a team of past experiences reassembled together in the context of the goal. As a final remark,

we chose not to involve initial layers of sensation and final stages of action because otherwise it would be impossible to distinguish imagination from reality (imagine activating the joints every time you read words like *lick, kick*). There is evidence from functional imaging too to justify this assumption (Martin, 2009). We believe millions of years of evolution have managed to strike the right balance on the set of neural substrates shared by top down and bottom up processing.

7.3 Survival of the Fittest May Apply to Memory Too. While there is some support for the opinion that the act of recalling refreshes an episodic trace anew (Dudai, 2006), there is no clear consensus on what the underlying computational mechanisms are and what their advantages are. We were forced to consider this topic when the robot was experiencing different episodes of interactions with objects on different days. Some of them encoded partially the same content but with different consequences; some of them included knowledge related to objects that others did not encode but coincided partially with others. Even assuming that all such memories may be recalled accurately based on partial cues from the present (that is not true because there are indeed errors in retrieving similar patterns), not all of them can be used at the same time. Hence we introduced the idea that only the fittest memories—-those that win the competition and manage to control the hub top down (even partially)—are refreshed. We found that the inverse naturally leads to a mechanism of forgetting: the only way for an episodic trace to survive is by reenacting its content (even partial) through the body. Conversely, consistent losers are eliminated. Just like the old EM1 to EM4 were forgotten and new EM5 and EM6 took over (themselves constructed using the old experiences) one day someone else who knows more or reaps greater rewards may eventually replace them as learning goes on. We believe that this is a consequence of a natural process of cognitive development of a cumulatively learning agent: as it encounters new things, old things have to be put in context and some of them get eliminated, their knowledge encoded in a new competitor who goes beyond. We also showed that such a scheme is healthy in the sense that it decreases patterns that are too close to each other, hence making retrieval more efficient and in turn also increasing storage capacity of the memory network, and it reduces computational load by decreasing drastically the number of isolated experiences that have to be remembered and then compete against each other to control the hub top down. In this sense, we believe survival of the fittest applies even in the mental space with a direct implication for efficient management of computational resources, reduction in computational load, hence fast reaction times, and having implications on the growth of the cognitive agent. It may be interesting to see what happens if this capacity is deactivated in iCub. We look forward to this in our future works.

7.4 Counting Comes Before Calculus: Role of the Teacher. We are social agents, and helping and seeking help is undoubtedly cognitive. The obvious reason is that it has a minimizing effect related to efforts that an agent needs to direct toward exploration (that can be expensive energetically). In moving from basic counting to the complexity of calculus, often the helping hand of the good teacher helps. And it does so for an embodied robot that learns cumulatively. If the "user" can support its development by creating scenarios of gradually increasing complexity for it to act and learn, we believe it may minimize the need for engaging in needless exploration. Readers might have noticed that while the robot itself learns cumulatively, the teacher has also intelligently introduced various objects in the environment cumulatively in time (see sections 3 to 6) in order to intentionally cause contradictions and trigger explorations or generation of novel behaviors'. We believe the introduction of such social context and soft user guidance moves in the direction of a middle path that both minimizes excessive exploration by the robot on one hand and eliminates hard coding by the programmer on the other. Human infants often go through this phase where toys of different levels of complexity are introduced gradually to play with (even categorized approximately in age groups). The same applies to a baby humanoid learning cumulatively, but unlike a human infant, this helps looking deeper into the underlying computational principles as we users are ourselves are at the receiving end and are constantly learning. So closing the loop between robot and an intelligent teacher/user can make learning more productive. To sum up, infants often learn in a social environment where the parent/teacher plays a key role in nurturing the developmental curve. Cognitive robots/assistants are also envisaged to exist in a shared environment with its user, and it is up to the user to train it in the tasks in which he or she needs assistance. We have attempted to incorporate such an aspect into our ongoing efforts to develop iCub cognition. In general, the teacher/user plays three crucial roles:

1. *Motivate:* Set goals and create rich sensorimotor worlds where the robot can get diverse experiences; at the same time ensure that the environment is within the zone of proximal development (Vygotsky, 1978) of the robot.
2. *Demonstrate:* This deals with imitation learning to acquire motor skills, for example learning to use common day-to-day tools found in domestic and industrial setups. This issue is ongoing, and recent results have been addressed elsewhere (Mohan & Morasso, 2011, 2012).
3. *Reinforce:* Rewards and penalties coming from the teacher/user aid the value-dependent learning process of the robot. This feature contributes toward creating contradictions between what the robot anticipates getting and what it gets, hence driving it to learn what was wrong and help it to reason better next time. Perhaps a

"humanlike" touch to machine learning is the need of the times if we are to see the emergence of machines that can assist us flexibly in the environments we inhabit and create.

Acknowledgments

References

Addis, D. R., & Schacter, D. L. (2012). The hippocampus and imagining the future: Where do we stand? *Frontiers in Human Neuroscience*, *5*, 173. doi:10.3389/fnhum.2011.00173

Addis, D. R., Wong, A. T., & Schacter, D. L. (2007). Remembering the past and imagining the future: Common and distinct neural substrates during event construction and elaboration. *Neuropsychologia*, *45*, 1363–1377.

Barabási, A. L. (2003). *Linked: The new science of networks*. Cambridge Perseus Books.

Barabási, A.-L. (2012). The network takeover. *Nature Physics*, *8*, 14–16.

Barabási, A.-L., & Albert, R. (1999). *Emergence of scaling in random networks. Science*, *286*, 509–512.

Bressler, S. L., & Menon, V. (2010). Large-scale brain networks in cognition: Emerging methods and principles. *Trends in Cognitive Sciences*, *14*, 277–290.

Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The brain's default network: Anatomy, function, and relevance to disease. *Ann. NY Acad. Sci.*, *1124*, 1–38.

Buckner, R. L., & Carroll, D. C. (2007). Self-projection and the brain. *Trends in Cognitive Science*, *2*, 49–57.

Burgess, N., Maguire, E. A., & O'Keefe, J. (2002). The human hippocampus and spatial and episodic memory. *Neuron*, *35*, 625–641.

Chong, H. Q., Tan, A.-H., & Ng, G.-H. (2007). Integrated cognitive architecture: A survey. *Art. Intelligence Review*, *28*, 103–130.

Clark, A. (1997). *Being there: Putting brain, body and world together again*. Cambridge, MA: MIT Press.

Dudai, Y. (2006). Reconsolidation: The advantage of being refocused. *Curr. Opin. Neurobiol.*, *16*, 174–178.

Eichenbaum, H. (2004). Hippocampus: Cognitive processes and neural representations that underlie declarative memory. *Neuron*, *44*, 109–120.

Frith, C. D., & Frith, U. (2012). Mechanisms of social cognition. *Annu. Rev. Psychol.*, *63*, 287–313.

Gallese, V., & Sinigaglia, C. (2011). What is so special with embodied simulation. *Trends in Cognitive Sciences*, *15*, 512–519.

Georg Stork, H. (2012). Towards a scientific foundation for engineering. *Biologically Inspired Cognitive Architecture*, *1*, 82–91.

Gibson, J. J. (1966). *The senses considered as perceptual systems*. Boston: Houghton Mifflin.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.

Grafton, S. T. (2009). Embodied cognition and the simulation of action to understand others. *Annals of the New York Academy of Sciences*, *1156*, 97–117.

Hagmann, P., Cammoun, L., Gigandet, X., Meuli, R., Honey, C. J., Wedeen, V. J., & Sporns, O. (2008). Mapping the structural core of human cerebral cortex. *PLoS Biology*, *6*, e59.

Hassabis, D., & Maguire, E. A. (2011). The construction system of the brain. In M. Bar (Ed.), *Predictions in the brain: Using our past to generate a future*. New York: Oxford University Press.

Hesslow, G. (2002). Conscious thought as a simulation of behavior and perception. *Trends in Cognitive Sciences*, *6*, 242–247.

Hopfield, J. J. (2008). Searching for memories: Sudoku, implicit check bits, and the iterative use of not-always-correct rapid neural computation. *Neural Computation*, *20*, 512–519.

Martin, A. (2007). The representation of object concepts in the brain. *Annual Review of Psychology*, *58*, 25–45.

Martin, A. (2009). Circuits in mind: The neural foundations for object concepts. In M. Gazzaniga (Ed.), *The cognitive neurosciences* (4th ed., pp. 1031–1045). Cambridge, MA: MIT Press.

Mason, M. F., Norton, M. I., Van Horn, J. D., Wegner, D. M., Grafton, S. T., & Macrae, C. N. (2007). Wandering minds: The default network and stimulus-independent thought. *Science*, *315*, 393–395.

Maturona, H. R., & Varela, F. J. (1980). *Autopoiesis and cognition*. Dordrecht: D. Reidel.

Meyer, K., Damasio, A. (2009). Convergence and divergence in a neural architecture for recognition and memory. *Trends in Neuroscience*, *32*, 376–382.

Mohan, V., & Morasso, P. (2011). Passive motion paradigm: An alternative to optimal control. *Front. Neurorobot.*, *5*, 4. doi:10.3389/fnbot.2011.00004

Mohan, V., & Morasso, P. (2012). *How past experience, imitation and practice can be combined to swiftly learn to use novel "tools": Insights from skill learning experiments with baby humanoids*. Intl. Conf. on Biomimetic and Biohybrid Systems: Living Machines 2012, Barcelona, Spain.

Mohan, V., Morasso, P., Sandini, G., & Kasderidis, S. (2013). Inference through embodied simulation in cognitive robots. *Cognitive Computation*, *5*, 355–382. doi:10.1007/s12559-013-9205-4

Mohan, V., Morasso, P., Zenzeri, J., Metta, G., Chakravarthy, V. S., & Sandini, G. (2011). Teaching a humanoid robot to draw "Shapes." *Autonomous Robots*, *31*(1), 21–53.

Murata, A., Fadiga, L., Fogassi, L., Gallese, V., Raos, V., & Rizzolatti, G. (1997). Object representation in the ventral premotor cortex (area f5) of the monkey. *Journal of Neurophysiology*, *78*, 2226–2230.

Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews Neuroscience*, *8*(12), 976–987.

Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman, G. L. (2001). A default mode of brain function. *Proc. Natl. Acad. Sci. U.S.A.*, *98*, 676–682.

Schacter, D. L., Addis, D. R., Hassabis, D., Martin, V., Nathan, R. S., & Szpunar, K. K. (2012). The future of memory: Remembering, imagining, and the brain. *Neuron*, *76*, 677–694.

Sederberg, P. B., & Norman, K. A. (2010). Learning and memory: Computational models. In G. F. Koob, M. Le Moal, & R. F. Thompson (Eds.), *Encyclopedia of behavioral neuroscience*. Oxford: Academic Press.

Sporns, O. (2010). *Networks of the brain*. Cambridge, MA: MIT Press.

Squire, L. R., & Wixted, J. (2011). The cognitive neuroscience of human memory since H.M. *Annual Review of Neuroscience*, *34*, 259–288.

Suddendorf, T. (2013). Mental time travel: Continuities and discontinuities. *Trends in Cognitive Sciences*, *17*, 151–152.

Suddendorf, T., & Corballis, M. (1977). Mental time travel and the evolution of the human mind. *Genetic, Social, and General Psychology Monographs*, *123*, 133–164.

Suddendorf, T., & Corballis, M. (2007). The evolution of foresight: What is mental time travel and is it unique to humans? *Behavioral and Brain Sciences*, *30*, 299–313.

Suddendorf, T., Addis, D. R., & Corballis, M. C. (2009). Mental time travel and the shaping of the human mind. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *364*, 1317–1324.

Szpunar, K. K., Watson, J. M., & McDermott, K. B. (2007). Neural substrates of envisioning the future. *Proc. Natl. Acad. Sci. U.S.A. 104*, 642–647.

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organisation of memory* (pp. 381–403). Orlando, FL: Academic Press.

Tulving, E. (2002). Episodic memory: From mind to brain. *Annu. Rev. Psychol.*, *53*, 1–25.

Van den Heuvel, M. P., & Sporns, O. (2013). Network hubs in the human brain. *Trends Cogn. Sci., 17*, 683–696.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Watts, J. D., & Strogatz, S. (1998). Collective dynamics of small world networks. *Nature*, *393*(6684), 440–442.

Welberg, L. (2012). Neuroimaging: Rats join the "default mode" club. *Nature Reviews Neuroscience*, *13*, 223–223, doi:10.1038/nrn3224

Wiener, N. (1961) Cybernetics: Or control and communication in the animal and the machine (2nd rev. ed.). Cambridge, MA: MIT Press.