Annexure I
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 activity 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
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 Dexterous Assembler Robot Working with embodied Intelligence (www.darwin-project.eu).
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 $N \times N$). 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 $V_{\text{new}}$ be a one-dimensional vector representing the activity of $N$ ($N = 1000$) neurons shown as a $20 \times 50$ 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 $1000 \times 1000$, 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. Consider that the episode represented by activity $V_{\text{new}}$ 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:

$$\text{If } V_i = 1 \text{ and } V_j = 1, \text{ then make } T_{ij} = 1 \text{ (regardless of what its value was before). Else, make no change in } T_{ij}. \quad (2.1)$$

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, $V_{\text{new}}$) 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:

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

$$I_{\text{inhib}} = g \left( -\alpha_{\text{in}} V_k + \beta \sum_k V_k \right), \quad (2.2b)$$

$$g(i) = 0, \text{ if } (i < 0), \text{ else, } g(i) = i. \quad (2.2c)$$
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 $K$th 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 $\alpha$ was chosen as 30, $\tau_{rel}$ as 1000, and $\beta$ 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 $20 \times 50$ 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
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 $\circ$ neurons in the object hub (42, here) and $N$ neurons in the episodic memory patch (1000, here), then $W$ is a matrix of $\circ \times N$. $W$ 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

$$W_{hn} = 1 \quad (3.1)$$

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 $20 \times 50$ 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

$$P_c = H \times K,$$

where

$$K_{hn} = W_{hn} \times \Psi_{hn}$$

$$\Psi_{hn} = 1 \text{ if, } M_{hn} = c$$

(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 $\Psi$, 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.

In 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 $\Omega_i$ of the $i$th episodic memory on the hub (without taking into account effect of other competitors) is as follows:

$$\Omega_i = ((W \times EM_i) \times H) \times R_i.$$ (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 $42 \times 1000$ matrix here; $EM_i$ is the reconstructed past experience ($1000 \times 1$ vector representing activations of the $50 \times 20$ episodic memory patch). So the component $W \times E_i$ is a $42 \times 1$ 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 $\Omega_i$ (a $42 \times 1$ 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, $\Omega_i$ 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, $\Omega_i$ 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 $\Omega_i$ 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:

\[ \Omega_i(t) = \Omega_i(t-1) + \frac{1}{\tau} F(\Omega_i(t-1)), \] where \( F(\Omega_i) = \sum_{j=1, j \neq i}^{m} (\Omega_j \ast \Omega_j) \lambda_j, \]

\[ \lambda_j = R_j \sum_{k=1}^{n} (\Omega_j)_k. \] (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 \( \Omega_i \ast \Omega_j \) and (2) how much more a memory \( EM_j \) knows than \( EM_i \) in the present context, which is determined by the scalar \( \lambda_j \) (e.g., if the environment contains spheres, cylinders, and cubes and a remembered past experience, say \( EM_3 \), encodes information related to all of them, then it can inhibit the ones that know only parts like \( EM_1 \) and \( EM_2 \)). In such cases, \( EM_3 \) will inhibit \( EM_1 \) because it knows more in relation to the present context, and both \( EM_3 \) and \( EM_1 \) in turn will cumulatively inhibit \( EM_2 \), which knows the same as \( EM_1 \) but yields even less of a reward. In the present scenario, only case 1 applies since both remembered past experiences \( EM_1 \) and \( EM_2 \) encode knowledge related to spheres and cylinders with \( EM_1 \) (stacking the sphere on the top) anticipating a greater reward (last row). Figure 6B shows five temporal iterations of the top-down influence of \( EM_1 \) and \( EM_2 \) on the object hub. As seen, \( EM_1 \), which fetches a greater reward, quickly manages to inhibit \( EM_2 \) 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 \( EM_1 \). 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
- Recall of past experiences
- Partial cue generation
- Bottom up object hub activity (Cube, Cyl, Large box, Sph)

B. Top down competition between remembered experiences to control the hub (Here, EM3 and EM4 control the hub in a distributed fashion)
- EM1: Cyl on top of Sph
- EM2: Sph on top of Cyl
- EM3: Cube-Cyl-Sph
- EM4: Large Box-Cube

C. Extracting sequence chunks from winning team of remembered memories (EM3 and EM4) and recombining them
- Overlapping object cube

D. Novel action plan synthesized by recombining winning past episodic experiences EM3 and EM4 such that overlap distance is minimum.
- Present behavior: Large box-cube - small cylinder-sphere

E. Reaching cube with the desired map, place
- First move: sequence of pick and place
- Cylinders fitted successfully
- Placing cylinders
- Cylinder placed, mushroom left
- Softly releasing fingers from the mushroom
- With drawing the arm, from the stack
- Initialization, 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.
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 $\tau_{\text{rel}}$, parameters in the inhibitory network $\alpha^m$, and $\beta$. $T$ 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 $\tau_{\text{rel}}$ is varied. Low time constants ($\tau_{\text{rel}} = 40$) 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 $\alpha^m$ is changed from 2 to 50, keeping $\tau_{\text{rel}}$ 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 $\alpha^m$ was chosen for all the experiments reported in this article. On the contrary, the network behavior is more sensitive to $\beta$ 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 $\beta$ is varied.
from 0.5 to 18, keeping $\alpha_{in}$ constant at 5 and $\tau_{rel}$ as 1000. As observed in row C (right panel), very small values of $\beta$ 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 $\beta$ is set to 5, resulting in correct recall of the stored memory from the partial cue. However, further increasing $\beta$ also abnormally affects the retrieval because of the high level of inhibition. Hence, the parameter $\beta$ 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 $\beta$. 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,
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 &
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

The research presented in this article is supported by Istituto Italiano di Tecnologia and the European Union through the FP7 project DARWIN (www.darwin-project.eu, grant FP7-270138). We express our gratitude toward the anonymous reviewers for their enormous patience and constructive feedback to develop the article.

References


Received November 23, 2012; accepted June 15, 2014.
Annexure II
“Connecting Experiences”: Towards a biologically inspired memory for developmental robots

Ajaz Ahmad Bhat, Vishwanathan Mohan, Francesco Rea, Giulio Sandini, Pietro Morasso
Robotics, Brain and Cognitive Sciences department, Istituto Italiano di Tecnologia, Genova, Italy.

Abstract— We all inhabit a continuously changing world where practically neither everything can be known nor can everything be experienced. In order to ultimately survive, diverse ‘context relevant’ chunks of knowledge emerging from one’s multiple past experiences have to be ‘connected’ and exploited flexibly in the context of the present to ensure smooth realization of goals. Understanding how the brain achieves such creative and generative use of experience is a central challenge facing both neuroscience and cognitive robotics today. This article is an exploration in this direction through playful, cumulative experiences of humanoid iCub. A biologically inspired neural framework for organizing, recalling and flexibly combining memories of multiple experiences of the robot under novel situations is proposed. The architecture is validated through both a simple staking task and then a “real world industrial” fuse box assembly task where recall of context specific past experiences trigger simulation of consequent future events in time hence allowing the robot to realize an otherwise unrealizable user goal: Pushing an object that is non-existent in the scene (but could exist if the robot creates it by exploiting its past experiences). Through the resulting behaviors, we emphasize that neural mechanisms for organization of episodic memory is a fundamental design feature necessary for the survival of autonomous robots in a real world ‘where neither everything can be known nor can everything be experienced’.

Index Terms— Episodic memory, Cumulative learning, iCub humanoid, Embodied simulation, Creativity

I. INTRODUCTION

Indeed, the most striking feature of any cognitive system (natural/artificial) is its ability to keep learning cumulatively (different things at different times, as necessary) and use such experiences creatively when faced with novel situations. As a simple example, consider the case of Betty the new Caledonian crow under the supervision of animal psychologists at Oxford [1]: she had playful experiences of bending flexible pipe cleaners; a year later she faced a nontrivial problem of fetching her dinner basket trapped in a transparent vertical tube; a long wire was available nearby. After a short hesitation, she picked up the wire, fashioned a hook out of it using her beak and used the newly created tool to pull out her dinner basket. The fascinating tinge of creativity demonstrated by Betty involved nontrivial perception-action”, recall of “a specific” learnt past experience in the context of the “present” (making a hook), anticipation of the consequence of pulling the basket up using the newly created “hook tool” connected to its beak and “flexible” integration of all this in the context of the goal. Such generative use of “experience” is important simply because we all ‘inhabit’ a continuously changing world where neither everything can be known nor can everything be experienced. To ultimately survive, diverse ‘chunks of knowledge’ emerging from one’s past experiences have to be “connected” and flexibly exploited in the context of the ‘present’ to ensure smooth realization of goals. Neural mechanisms associated with organization and use of ‘memory’ play a fundamental role here: in connecting ones ‘past’ with the ‘available present’ and ‘possible future’. Indeed, such mechanisms are fundamental for any cumulatively developing robot. Implementing such mechanisms in cognitive robots offers a unique opportunity to reenact the incessant interplay between neural mechanisms related to learning, memory and prospecton from the perspective of an integrated system that “acts, learns, remembers, reasons, makes mistakes”. Situated within such interplay underlies 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, keep learning further. Emerging trends from the fields of connectomics [14,5], functional imaging studies in relation to organization of semantic and episodic memory in the brain and the discovery of the Default Mode network [2-4,13,15] now provide numerous insights to guide development of brain guided computational framework for organization and use of episodic memory in cumulatively learning embodied robots. In particular, there is converging evidence suggesting that there is indeed extensive overlap in the brain networks activated while recalling the past and those engaged during other activities as diverse as simulating the future, goal directed planning, perspective taking and some forms of spatial navigation. Such a perspective urges to view 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. In this article, we attempt to integrate emerging ideas from neurosciences [2-5,13-15], to create a brain guided neural framework for organization and creative use of episodic memory for iCub humanoid. In line with several emerging results from neurosciences, we emphasize that
recall of context relevant past experiences and anticipation of future share computational basis. The architecture is validated through both a simple staking task and then a "real world industrial “fuse box” assembly task where recall of context specific past experiences trigger simulation of consequent future events in time hence allowing the robot to realize an otherwise unrealizable user goal: Pushing an object that is non-existent in the scene (but could exist if the robot creates it by exploiting its past experiences). In this context, the goal of Pushing an object “nonexistent” in the scene faced by iCub approximately mimics the situation encountered by the Caledonian crow Betty: when it has to recall a context relevant past experience, anticipate the consequence of it and initiate real actions to transform the present situation in a way that becomes more conducive towards realization of its goals.

The rest of the paper is organized as follows. Section II describes the computational building blocks, from basic perception-action, to how sensorimotor experiences acquired by the robot form content of its episodic memory, how experiences are encoded and recalled based on context, the underlying biological inspiration can link to emerging trends in neurosciences. Section III validates the framework in two different scenarios, the first related to a simple goal of building the tallest stack and the second Pushing “a nonexistent object” that approximately mimics the Betty scenario. The discussion briefly summarizes the novel features, the biological basis, and ongoing work to extend the framework.

II. THE COMPUTATIONAL BUILDING BLOCKS

The section describes the central building blocks in the proposed computational framework. While subsystems related to basic perception and action are described only briefly to establish context, mechanisms related to learning and organization of episodic memory that form the central core of the framework are described in detail.

Basic Perception-Action: In the context of this article, the starting point in terms of basic perception-action is a functional “Identify-Localize-Reach-Grasp” loop that allows perceiving and manipulating a range of objects necessary for conducting advanced experiments related to the ‘industrial fuse assembly scenario’. Color and Shape based visual analysis is used to recognize a range of objects like cuboids, cylinders, MECCANO blocks, industrial fuse, fuse box etc. The deployed shape analysis system [6] based on chamfer matching combined with the sliding window technique allows real-time recognition of objects in the scene. The output of the shape module is a set of object descriptors ID’s associated with each perceived object in the scene. Primitive motor actions like reaching and grasping are achieved through the Passive motion Paradigm framework (PMP) that coordinates the iCub upper body (left arm-torso-right arm chain). Further details on the basic perception-action system can be found in [11]. Figure 2-4, shows the robot perceiving and manipulating objects used in the experiments described in this paper.

Organization of sensorimotor information through Hubs:

The higher level organization of sensory, motor and cognitive information in the computational architecture is guided by three recent trends in neurosciences: 1) Functional imaging studies providing evidence 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 [10,13]; 2) Studies from the field of connectomics suggesting the presence of a small set of Hubs that enable swift global integration and cross-modal communication between multiple subsystems that process various sensory, motor and cognitive information [5,14]; 3) Converging evidence that a core network of “highly connected” brain areas (Hubs) called as the Default Mode network [3-5] is consistently activated when subjects perform diverse cognitive functions like recalling past experiences, simulating possible future events (or projection), goal directed planning, imagining fictitious scenarios, interpreting perspectives of others. In other words, disparate cognitive functions often treated as distinct, might share common neural processes that crucially deal with the “memories” of one own individual experiences and how it is exploited in various “contexts”. Such a perspective urges to view 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.

Figure 1: A shows a block diagram of how basic information related to perception and action are organized in the proposed computational framework. 1B shows the link between central hubs related to Object, Action and rewards and the neural episodic memory network. Importantly, the content of the episodic memory is the temporal order of activations in the hubs related to perception, action or rewards received when experience is originally gained by the robot.

Figure 1 shows a block diagram of how basic information related to perception and action are organized (1A) and their link to the neural episodic memory of the robot (1B). Bottom up perceptual information coming from the sensory
layer performing visual analysis of the scene activate level 1 SOM’s organized in a property specific fashion (mainly color and shape). The activations in level 1 maps in turn trigger the neurons in the object hub giving rise to a distributed representation of the objects in the scene (for example, red cylinder, yellow mushroom). Note that the connectivity between the maps is bidirectional hence allowing a word inputted by the user like “red container” to generate activations in the Object hub, that in turn can activate the color and shape SOMS, similar to anticipating top down what a red container might be [12]. Similarly, such a two layer, property specific organization also allows to learn which properties are causally dominant in a particular task/goal, for example color of objects do not matter when the goal is to assemble the tallest stack, but shape and size of objects do matter [12]. To clarify, color and shape SOMs are trained from vectors of object descriptors coming from visual analysis and the Object hub trained through the activations in the level 1 maps. The method of training SOM’s is standard [8] and out of scope in the context of this paper. We refer the interested reader to [12] for a detailed description of training different maps in the proposed framework and their links to neurosciences. What is relevant in the context of this paper is that bottom up perceptual analysis trigger different property specific SOM’s ultimately activating the object hub (basically representing: what is it). In relation to organization of Action too, that there is a subtle separation between representation of actions at an abstract level (“what all can be done with and object/tool”) and the details related to action planning, and execution (“how to”) that is further mediated by the body schema (PMP Framework [11]). The abstract layer forms the “action hub” and consists of single neurons coding for different action goals like reach, grasp, stack etc. and grows with time as new primitives are learnt. Single neurons in the connector hub in turn trigger the sub systems that hold detailed information related to execution the action, generation of motor commands for the iCub upper body through PMP. In this sense neurons in the top level “action connector hub” are similar to “canonical neurons” found in the pre-motor [16] that are activated at the sight of objects to which specific actions are applicable. In the present architecture we deal with 10 primitive actions that form the motor vocabulary of the action hub: Reach, Grasp, Release, Push, Find, Pick and Place, Assemble, Observe (the perceptual consequence), ask user help and Quit. Similar to the Object hub, of relevance in the context of this paper is the activations in the Action hub (since details of perceptual analysis and action execution are abstracted and functionally available based on previous works [12]).

**Hubs and their link to iCub episodic memory:** Experiences are either gained through directly interactions between the body and the world or socially through experiences of others, and this must reflect in the episodic memories of a developmental robot too. Here comes the link between Hubs and the neural episodic memory of the robot in our computational framework (figure 1B). We clarify this here and point out the rationality behind our proposition. To decide what information to encode, we looked at emerging results related to the recently discovered default mode network (DMN:[3]). It’s been established through numerous studies that central function of this network is to generate “self-referential episodic simulations”, that includes reconstruction of the past experience in relation to a present context (partial cues), simulation of possible future scenarios and planning [4,5,15], consolidation of new experiences with the existing knowledge, some forms of spatial navigation, perspective taking. Such episodic simulation activates the highest level brain hubs involved in object, action and value processing. This makes sense because while operating in a real world, partial cues (related to perception, action, goals and rewards) emanate from the environment activating bottom up the highest level neural maps (like, object connector hub, action hub in 1B). Thus, there is both biological grounding and computational simplicity in considering that activity in the highest level object-action hubs fill in partial “contextual” information, triggering a reconstruction process of remembering the complete past experiences. Thus, the content of the robot’s episodic memory is the temporal sequence of activity either in the object, action hubs or reward received from user, when real experience was originally gained by the robot and encoded in the episodic memory.

The neural episodic memory to encode multiple sensorimotor experiences of iCub (or the resulting bottom up activations in Hubs and the rewards acquired through time) is realized using a recently proposed neural framework for organization of memory in iCub [7]. For modeling purposes, we deal with a small patch in the sheet like neocortex, consisting of 1000 pyramidal cells (N=1000). The 1000 neurons are organized in a sheet like structure with 20 rows each containing 50 neurons (see figure 2). Every row may be thought as an event in time (related to activation in object hub, action hub or received end reward) and the complete memory as an episode of experience (for example, picking a cylinder and placing it on a mushroom and getting a reward of 0 from the user: figure 2). Importantly, in the memory network of 1000 neurons, multiple episodic memories can be represented and retrieved (approximately 230 episodes as reported in [7]). So activity in the 1000 neurons are not just used to exclusively represent one episodic memory, but rather many different episodic memories in a cumulative learning set up (for example, building the tallest stack on *day 1* with mushroom and cylinder, playing on *day 2* with cubes and containers, learning to assemble a Fuse box on day 3, so on).

**Encoding experiences in the neural episodic memory:** The memory network is characterized by “all-to-all” connections between the N (1000) excitatory neurons (thus the connectivity matrix is of the order NxN). In addition, there is an inhibitory network that is equally driven by all N excitatory neurons and in turn 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. New experiences are learnt by the robot are encoded by updating connections between the neurons in the following way. Let $V_{\text{new}}$ be a one dimensional vector representing the activity of N (N=1000, here) neurons shown
as a 20x50 matrix (figure 2 top panel). Let $T$ denote the connectivity matrix between the $N$ neurons. Since there are 1000 neurons, the dimensionality of $T$ is 1000x1000, that 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. Consider that a new episode represented by activity $V_{\text{New}}$ has to be stored in the memory network.

\[ V_{ij} = \begin{cases} 1 & \text{if } V_i = 1 \text{ and } V_j = 1 \\ 0 & \text{else} \end{cases} \]

Figure 2: Shows how multiple experiences acquired by the robot while learning to construct the tallest stack (top) and assembly of a fuse box (bottom) are encoded in the neural episodic memory. As seen, the episodic memory network consists of 1000 neurons organized in a sheet like structure with 20 rows each containing 50 neurons. Every row may be thought as an event in time (related to activation in object hub, action hub or received end reward) and the complete memory as an episode of experience (for example, picking a cylinder and placing it on a mushroom and getting a reward of 0 from the user: EM1). Multiple episodes of experiences can be stored in the same memory network using equation 1.

This is done by updating all the connections $T_{ij}$ between the $N$ neurons, using a simple rule summarized below:

- If $V_i = 1$ and $V_j = 1$, then make $T_{ij} = 1$ (regardless of what its value was before);
- Else, make no change to $T_{ij}$.

Equation (1) represent the rule of updating the connectivity matrix $T$. The robot has not yet exhausted all its options, so in episode 2 it attempts to stack the mushroom on top of the cylinder (5B), receives a reward of 2 from the user (for successfully building the tallest stack). This experience is also encoded in the episodic memory (see [7] for further details). This experience is also stored as a new episodic memory. If we compare EM1 and EM2 the difference is that the objects representations swap roles in the temporal order (mushroom moving to row 1 and cylinders to row 3). Similarly, a third experience of assembling a fuse box (picking up a fuse and placing it in the fuse stand) is also encoded in the same network by updating the $T$ matrix using equation 1.

**Remembering past experiences from partial cues:** The interesting feature of the proposed computational framework is that the content of the episodic memory correspond to activity in the object or action hubs (when experience was gained). This has additional advantages in terms of generation of partial cues to trigger recall of context relevant past experiences, because objects in the real world or user goals are often the sources of partial cues. For example, on day X if the green mushroom is placed in front of the robot, bottom up perception will lead activations in the object hub (corresponding to the mushroom) that fill in partial information: or activate a sub set of neurons in the 50x20 neural episodic memory layer. This serves as initial condition to trigger recall of past experiences (in this example, in relation the object mushroom). The network dynamics to recall past experiences from partial cues is given by equation 2: $V_k$ is the activity in the $K$th neuron (a subset of neurons in the 50x20 neural episodic memory layer will be active based on the partial cue or initial condition). $T$ is the connectivity matrix between the neurons learnt using equation 1 when the episodes of experiences are encoded in

\[ \tau_{rel} \dot{V}_k = -V_k + \sum_{j=1}^{N} T_{k,j} V_j + I_{inhib} \]

\[ I_{inhib} = g(-\alpha^{in} + \beta \sum_{k} V_k) \]

\[ g(i) = 0, \text{if } (i < 0), \text{else, } g(i) = i. \]
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 $\alpha^m$ was chosen as 30, $\tau_{rel}$ as 1000 and $\beta$ as 3.5.

In sum, in this section we proposed a brain guided framework for organization of episodic memory in a humanoid robot in a cumulatively learning set up. Multiple experiences of acting and learning from different scenarios, objects over time in a cumulative fashion can be encoded in the same network of 1000 neurons, organized in a 50x20 sheet (every row an event in time). The critical feature of the framework is that the content of the episodic memory corresponds to the temporal order of activations in the object and action hubs or the received reward, when experience is originally gained. The straight forward advantage is that any time in the future, if the hubs are activated bottom up (example, the robot perceives a fuse), the activations in the hubs automatically act as partial cues, providing context and triggering the recall of context relevant past experiences. The next section presents further results exploiting the computational architecture.

III. FROM RECALLING THE PAST TO SIMULATING THE FUTURE IN THE PRESENT CONTEXT

This section explores the interplay between learning, memory and prospection in the proposed episodic memory framework particularly highlighting a shared computational/neural basis for recalling past experiences and simulating the future events hence leading to creative goal directed behaviors in the robot.

III A. Interplay between learning, remembering and planning

Figure 3-E concisely shows the evolution of neural activity in the hubs and the episodic memory network in a simple scenario where the user places the green mushroom in the scene. Bottom up visual analysis activates the object hub, recognizing the object in the scene. Object hub activity directly fills in partial information in the 50x20 episodic memory network triggering the retrieval of related past experiences (through the dynamics of eqn.2). Panels D and E show the two episodes of experiences recalled from the partial cues. As seen, what is recalled is not just information about mushroom, but also the temporal order in which they were stacked along with cylinders in the past, what was the consequence and the reward received (information that was not available in the partial cue itself). Importantly, from the retrieval of the past experiences it is possible to infer which behavior is more rewarding. One may consider that the two recalled “past experiences” are competing with the “anticipated reward” from past experience as the criteria based on which a remembered memory of past experience wins the competition. In other words, given the goal to assemble the tallest stack, the episodic recall of past experience directly leads to the inference that “it is more rewarding to place Mushroom like objects on the top”. This sequence of “stacking the mushroom on top of the cylinder” and anticipating a reward of 2 from the user. Importantly, note that there is no need for an explicit planner: the valuable action sequence is evident in the reconstructed episodic memories.

Figure 3. Panels A-E concisely shows the evolution of neural activity in the hubs and the episodic memory network in a simple scenario where the user places the green mushroom in the scene. Partial cues emerging from bottom up perception activate a sub set of neurons in the episodic memory patch; trigger a retrieval of context relevant episodic experiences (through the dynamics of eqn. 2). Importantly, episodic recall of past experience also leads to inference in terms of which behavior is more rewarding (“it is more rewarding to place Mushroom like objects on the top”). Panel F shows the resulting behavior of the iCub, recalling its past experiences and executing the more rewarding action sequence of “stacking the mushroom on top of the cylinder” and anticipating a reward of 2 from the user. Note that there is no need for an explicit planner: the valuable action sequence is evident in the reconstructed episodic memories.
In other words, memories that win the competition manage to reenact their sensorimotor content through the body (in a way reasserting their value to the organism). Inversely, consistent losers like memory 1 here (the experience of placing the cylinder on the mushroom and failing in constructing the tallest stack) may be forgotten as learning progresses incrementally.

III B Episodic simulation through time: Emulating the ‘Betty’ scenario in a cognitive humanoid

This subsection presents an interesting scenario where episodic recall directly facilitates simulation into the future and enables the robot to realize and otherwise unachievable user goal. Panels in figure 4 show the temporal evolution of the behavior of the robot (indicated by the red arrow). The user issues a goal to Push the “Fuse Box.” As seen, a fuse and fuse stand is present in the scene and perceived by the robot. However, the requested object on which the primitive behavior “Push” needs to be executed is not present in the scene thus rendering the user goal not directly realizable. In this sense, the goal of Pushing an object “nonexistent” in the scene faced by iCub approximately mimics the situation encountered by the Caledonian crow Betty: when it has to recall a context relevant past experience, anticipate the consequence of it and initiate real actions to transform the present situation in a way that becomes more conducive towards realization of its goals. As seen in figure 4, though the user goal of pushing the fuse box is not directly realizable, the objects present in the scene are perceived through vision and activate “bottom up” the object hub (two objects fuse and the fuse stand are present). Object hub activations directly generate partial cues for the episodic memory network and trigger recall of past experiences in relation to the context. As seen, from the partial cue, the robot recalls its past experience of picking up the fuse and placing it into the stand (EM3: Figure 2), the consequence being the creation of the fuse box. Note that this is directly inferred through the reconstruction of past experience from the partial cue. In other words, recall of the past experience also directly allows simulation into the future allowing the robot to infer that the actions of picking and placing the fuse into the fuse stand will result in creation of the ‘fuse box’. In such a modified world, the goal requested by the user is indeed realizable. Thus even though the user goal was to Push the fuse box, the robot picks up the fuse and places it in the stand, thus creating the Fuse box and then executes the Pushing. Thus, exploiting its past experience, the available objects in the present and the anticipation of the resulting consequence of its actions, the robot thought its actions modifies the world appropriately to create the fuse box and realize an otherwise unachievable user goal.

IV. DISCUSSION

“It’s a wrong sort of memory that only works backwards”, remarked the White Queen in Lewis Carroll’s 1865 novel Alice in Wonderland. Emerging studies from the fields of connectomics [14,5], organization of semantic and episodic memory in the brain and the discovery of the Default Mode network [2-4,13,15] interestingly provide converging evidence suggesting that there is indeed extensive overlap in the brain networks activated while recalling the past and those engaged during other activities as diverse as simulating the future, goal directed planning, perspective taking and some forms of spatial navigation. Such a perspective urges to view 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. It is infact fundamental that cumulatively developing robots learning and experiencing different things (building the tallest stack, fuse box assembly, others) at different time instances (days) must also be equipped with a powerful “biologically inspired” memory architecture: that allows them to remember their past experiences based on context and further exploit them flexibly in novel circumstances. This article was both an adventurous and preliminary exploration in this direction. The proposed

Figure 4. Panels show the temporal evolution of the behavior of the robot (indicated by the red arrow) when the user issues a goal to Push the “Fuse Box”: a nonexistent object in the present scene. The panels depict how episodic recall of context relevant past experience also indirectly enables simulation into the future, enabling the robot to initiate actions that “modify the existing world” to make more conducive towards realization of its internal goals. As seen, even though the user goal was to Push the fuse box, the robot picks up the fuse and places it in the stand, thus creating the Fuse box and then executing the Pushing.
framework for organization and flexible exploitation of episodic memories in iCub is (a) Embodied: in the sense that the content of the robots episodic memory directly corresponds to the temporal order of activations in the object, action hubs and the rewards received when experience is originally gained. The crucial advantage of this feature is also that, in future any bottom up activation in the perception, action systems (robot perceives a fuse on Day X), automatically generated partial cues that trigger recall of context related past experiences (Example: picking up the fuse and placing it on the stand results in a fuse box, Placing a cylinder on top of the mushroom does not result in tallest stack); (b) Takes into account the need to “learn, remember, realize goals” in a cumulative setting: in the sense that multiple experiences acquired during different time instances (stacking on day 1, assembly of a fuse box on day 2) can be encoded in the same network and recalled in the future based on partial cues emerging from the present environment; (c) In line with emerging results from neurosciences, emphasizes that recall of context relevant past experiences and anticipation of future share computational basis. Two different scenarios of how recall of past episodic experiences leads to inferences and goal directed behaviors in the context of the “present” were illustrated. Ongoing works to build up the proposed architecture attacks further issues mainly 1) Combining past experiences with explorative actions to learn further: for example, cuboids are introduced as novel object along with mushroom and cylinder to build the tallest stack (this creates situation where the robot has partial knowledge of mushrooms and cylinders from past experience); 2) Combining multiple recalled experiences to generated novel behaviors (this requires taking into account overlaps between remembered experiences and is work underway); 3) Incorporating mechanisms related to forgetting: there is already a built in mechanism for this in our architecture (for example, the memory of stacking the cylinder on the mushroom will never win the completion as it is less rewarding and over time can be eliminated). In other words, past experiences that manage to reenact their content through the actor in the future (like placing the mushroom on top of the cylinder, creating the fuse box) re-emphasize their relevance hence retained, while consistent losers (like EM1 of figure 2) are gradually eliminated. While is support for this opinion from functional imaging studies that the act of recalling refreshes an episodic trace a new [16], there is no clear consensus on what are the underlying computational mechanisms and what are their advantages. We believe the proposed computational framework provides a foundational basis to further investigate and implement such mechanisms in cumulatively developing robots.

In Sum, Betty the New Caledonian crow, managed to recall her past experience of creating a hook out of plastic wires experienced an year back and exploit this experiences to pull out her dinner basket trapped in a vertical tube [1]. iCub recalls its past experience of assembling a “fuse box”, and exploits this experience to initiate actions that modify the present world (creating a fuse box from the available components in the present situation) in order to realize an otherwise unrealizable user goal (Pushing the Fuse box, that is not available in the scene). While it might be very tough to understand what is going on in Betty’s brain, cognitive robots like iCub coupled with emerging trends in neurosciences provides a platform to computationally “reenact” and causally understand the mechanisms that underlie such behaviors: from the perspective of an integrated system that “perceives, acts, learns, remembers, reasons, forgets and makes mistakes”. This article is a preliminary exploration in this direction, towards creating a brain guided organization of episodic memory and its flexible utilization in developmental robots, with an emphasis that it’s a wrong sort of memory that only works backwards.

ACKNOWLEDGMENT

The research presented in this article is supported by Robotics, Brain and Cognitive sciences dept. IIT, the EU FP7 project DARWIN (www.darwin-project.eu, Grant No: FP7-270138) and US Dept. of defence Grant (W911QY-12--C-0078).

REFERENCES


Copyright ©2014 IEEE
Annexure III
‘Muscleless’ motor synergies and Actions without Movements

Vishwanathan Mohan, Giulio Sandini, Pietro Morasso

Robotics, Brain and Cognitive Science department,
Istituto Italiano di Tecnologia,
Via Morego 30, Genova, ITALY
(vishwanathan.mohan, pietro.morasso, giulio.sandini)@iit.it

Corresponding Author: Vishwanathan Mohan, PhD
Robotics, Brain and Cognitive science department,
Istituto Italiano di Tecnologia,
Via Morego 30, 16163 Genova, Italia.
Phone: +39 01071781473

Article submitted to Trends in Cognitive sciences: Version 2
‘Muscleless’ motor synergies and Actions without Movements
Vishwanathan Mohan, Pietro Morasso & Giulio Sandini

The functional role of a “plastic, expandable, configurable” internal representation of the body as a connecting link facilitating the seamless continuum between real and imagined action while we “act, interact, anticipate and understand” is highlighted, revisiting old ideas from neural control of movement like Equilibrium point hypothesis and synergy formation in the renewed context of emerging trends in motor neurosciences.

Real and Imagined actions: Siamese twins?

Presently, there is growing consensus that cortical networks in the predominantly motor areas are activated in several other contexts related to ‘action’ that do not cause any overt movement. Distributed, multi-centered neural activity is consistently detected during different conditions like imagination of movement, observation/imitation of other’s actions, and comprehension of language [1-4]. The general insight emerging is that the fundamental problems of shaping motor output during action execution and providing the self with critical information related to feasibility, consequence and understanding of potential actions (of oneself or others) are closely intertwined than previously conceived. From an evolutionary perspective too, in organisms with complex bodies and inhabiting unstructured environments, actions are “goal oriented” and not just stimulus oriented, fundamentally requiring ‘covert simulation and overt execution of action’ to seamlessly alternate during the maturation of purposive behavior and social interaction: to maximize success, ensure survival. In this sense, real and imagined actions are like Siamese twins, inseparable but to some extent independent. How is this delicate balance realized in the brain? How are cortical substrates basically involved in organization of overt action functionally ‘recycled’ in other contexts (i.e. actions without movements)? Revisiting older ideas from motor control like the Equilibrium Point Hypothesis (EPH: 5-6) and synergy formation in the context of emerging trends in neuroscience, we explore the functional role of a “plastic, expandable” internal representation of the body, namely the ‘body schema’, as a fundamental connecting link facilitating the seamless continuum between actions “with and without movements” while we effortlessly ‘act, interact, anticipate and understand’.

Muscle Synergies and Equilibrium point hypothesis (EPH): Why Now?

Given the complexity of our bodies, even during the generation of simple movements, the brain has to deal with two typically contrasting requirements: the need to choose ‘one’ from infinite possible solutions (Bernstein’s Degrees of Freedom problem) and the need to produce ‘one solution’ in an infinite number of ways (Lashley’s Principle of Motor equivalence). As a simple example, consider that even the task of aimlessly moving your finger from one point to another can be achieved in an infinite number of ways (motion trajectories, joint rotations). An implicit solution of how the brain might deal with this problem of ‘plenty’ was posited several decades ago by the EPH. The basic tenet was that complex actions can indeed be achieved without a complex, high dimensional optimization process (i.e. the need to choose one from many), by simply allowing the intrinsic dynamics of the neuromuscular system to seek its equilibrium state when trigged by “intended” goals. The EPH idea exploited two beneficial properties of the neuromuscular apparatus of the body: 1) to induce, locally (in a muscle-wise manner), an instantaneous disturbance compensation action, and 2) to induce, globally (in a total body-wise manner), a multi-dimensional force field with attractor dynamics. Numerous studies conducted with intact and spinalized animals [7] show that complex motor behaviours can indeed be constructed by muscle synergies. Innovative aspects of the decades old EPH idea were its strong grounding on the biomechanics of the body and the apparent computational simplicity in solving the degrees of freedom problem. At the same time, only recently

Article submitted to Trends in Cognitive sciences: Version 2
advanced brain imaging techniques have allowed to gain direct access to cognitive/mental states in absence of overt behavior, thus making clear that actions involve a covert stage. In this renewed context, it is worth pondering how the computational principles captured by the EPH idea proposed for coordination of overt movement could be recycled to explain actions without movements and without muscle contractions. Undoubtedly, there must be some continuity with observable neural correlates. This line of thought interestingly leads towards two consequences: 1) the format of spatio-temporal patterns of purposive actions, namely the organization of the synergy formation process, must be shared by covert and overt actions; 2) this format cannot be strictly dependent upon the dynamics of the body (in particular, the neuromuscular system, as attributed by EPH), because in covert actions there is no motion of body masses or contraction of the muscles. Otherwise, as pointed out by Martin [8], reading words like “lick, pick, and kick” would result in licking picking and kicking. While ‘motor synergies’ have traditionally been associated with “muscles”, a way to resolve this conundrum is to take the EPH idea beyond its manifestation ‘in flesh’ and look instead at “muscleless” motor synergies. This is indeed the motivation of our proposal, connecting results from functional imaging, neuro-rehabilitation, tool-use in animals and cognitive robotics.

**EPH beyond Flesh: Muscleless motor synergies and the Body Schema**

A plausible way to unify covert simulation and overt generation of action (critical during the evolution of purposive behavior and social interactions) is to posit that even real movements are consequences of an “internal simulation”, an idea formulated in its basic essence by Marc Jeannerod [9]. Such a provocative proposal might imply reassessing our understanding of neural coordination of overt action itself: an issue that has been in recent debate [10-11,1]. Note that, even during the generation of overt actions as posited by EPH, the “compositionality” of the muscle synergies is ultimately made possible by the “compositionality” of the underlying force fields and attractor dynamics. A plausible extension is to consider that what occurs in the brain during both mental simulation and overt execution of action reflects an endogenous cortical dynamics very similar to the physical dynamics implicit in EPH, but realized through “animation” of a ‘flexible, plastic and configurable’ internal representation of the body, with the attractor dynamics of force fields induced by the intended goal. This line of thought is not new but emerged during the 80’s taking shape as the the Passive Motion Paradigm (PMP) [11-12] formally extending the EPH idea from muscles to the body schema (Box 1). *Importantly, PMP emphasizes that covert and overt stages must represent a continuum, such that every overtly executed action implies the existence of a covert stage, whereas a covert action needs not necessarily turn out into an overt action.* The synergistic interaction between the body schema and the attractor dynamics induced by the intended goal could be a means to functionally realize such continuum. That humans have an integrated, internal representation of their body is strongly suggested by the variety of pathological conditions which can only be explained by a deficient internal representation [13] or by sensory illusions [14]. Modern neuroscience has greatly enriched the concept, with numerous studies [15] identifying cortical areas in fronto-parietal cortex integrating proprioceptive and exteroceptive sensory information to maintain a coherent /updated internal representation of the spatio-temporal organization of the body. However, the role of the body schema in synergy formation has not been investigated in depth. Intriguing insights are now emerging in this direction, particularly in support of body schema being the connecting link between overt and covert action. As [16] demonstrate, stimulating right inferior parietal regions (in patients undergoing awake brain surgery) triggers a strong intention to move the contralateral hand, arm, or foot with participants believing that they performed the movements, although *no movement was performed and even no electromyography activity was detected.* Conversely, stimulating the premotor region triggered overt limb movements, though the patients firmly denied that they moved. While such emerging results are beginning to highlight, on one hand, the coupling between motor intentions and the internal representation of the body, and, on the other hand, the link

Article submitted to Trends in Cognitive sciences: Version 2
between actions with and without movements, PMP as an extension to the EPH idea provides a computational basis with the “body schema” as a central core: that separates motor cognition from motor control but allows a mutual, bidirectional interaction in an action-oriented manner during goal directed behavior and social interactions.

Acting ‘Green’: Quest for Computational Simplicity, Configurability and Recycling

Interestingly, dexterity in overt movement, purposive behavior with anticipation of the consequences of one’s actions, and social intelligence are critical desirable features if robots are to become commonplace assistants in numerous application domains: domestic, industrial, elderly care to mention a few. Prevalent computational modelling approaches generally converge on the role of forward models, but diverge on the perspective of how they might be realized in the brain or modelled computationally [1]. Our proposal brings back the critical role of body and its internal representation in the brain as a means to both simplify the computational process of coordination of action and at the same time recycle the same mechanism to engage in diverse forward simulations. Consider that the explosive growth in the complexity of the body in a limited number of species is typically concentrated in two body parts (the hand and the vocal tract) that do not have a specific function but are general-purpose tools (for manipulation and communication, respectively) to be employed in infinite numbers of possible manners and purposes. In this sense, given the brain’s constraints, critical desirable features for any biologically plausible realization of forward models is computational simplicity (to tackle the curse of dimensionality) and goal-specific configurability (to ensure flexibility, at the same time reusing/recycling the same hardware as far as possible). In relation to the former aspect, PMP networks operate with “well posed” computations, hence avoiding the critical need to choose one from many through a computationally expensive optimization process (box 1). Our proposal for muscleless motor synergies is closely connected to forward models emerging from predictive coding and active inference (AI) [1]. As Friston points out [10], the idea of modeling movements ‘caused by’ changes in the intrinsic space (joints, muscles,) as movements that ‘cause changes’ in the intrinsic space is the common feature between active inference based forward models and the body schema based PMP. The direction of causality has significance in terms of ‘computational cost’, because it circumvents the need for kinematic inversions and cost function optimizations, that can be prohibitive in practical situations like coordinating a 53 DoF humanoid [11]. At the same time, the body schema and force field metaphor is closer to the biomechanics and cybernetics of action as compared to the AI metaphor. Based on the nature of the goal, associated configuration of the body (connected tools), diverse forward/inverse models can be composed on the fly recycling the same neural/computational substrate (Box 1).

The blurred distinction between tool, body and others

The argument of configurability and recycling is interesting because it can be extended to coordinated tools and other bodies. As demonstrated by [17], with practice, a rake becomes a part of the acting monkey body schema. Recording from monkeys trained to use pliers to grasp otherwise unreachable food rewards, Umiltà et al [18] demonstrated that the end effect of learning skilled tool use was the transfer of the temporal discharge patterns that control ‘hand grasping’ (area F5) to the tool, as if the tool was the hand of the monkey and its tips were the monkey’s fingers. The body schema based PMP framework closely resonates with these results [11]. The tool space is represented exactly in the same manner as any other intrinsic motor space and during coordination the body and the tool act as one cohesive unit, to realize a goal at hand or afford mental simulations related to ensuing consequences of possible use of the tool itself (box 1). An intriguing idea proposed by Iriki [17] is that the ability to literally incorporate external objects into one’s own body schema and the ability to ‘objectify’ other bodies, thus recycling the same neural machinery, are two sides of the same coin. Several studies now provide converging evidence supporting
this perspective: that when observing actions, people recruit motor representations as if they were themselves acting [1-4]. Simply put, understanding may be conceived as an internal simulation that entails the reuse of our own ability to act with our bodily resources in order to functionally attribute meaning to ‘others’ actions, recycling some of the same cortical substrates the enable us to act ourselves.

In sum, overt actions are just the tip of the iceberg, under the surface is hidden a vast territory of actions ‘without movement’, understanding the neural/computational basis, which is the essence of motor cognition.

References

Acknowledgement
VM and PM are supported by EU FP7 grant DARWIN (FP7-270138) and US Dept. of Defense grant (W911QY-12--C-0078).

Article submitted to Trends in Cognitive sciences: Version 2
**Panel A**: Pictorially depicts the basic idea of synergistic interaction between the internal representation of the body and the attractor dynamics of force fields induced by the intended goal, as a means to describe both covert and overt action. Based on the intended task, diverse forward/inverse models can be composed dynamically based on the associated configuration of the body, coupled tools: two examples shown in panels B and C. Panel B: it shows a simple forward inverse model for one arm. In this simple case there are basically two motor spaces hand (or end-effector) space and joint space, both represented by two displacement nodes ($dX$: hand displacement, $dq$: joint rotation) and corresponding force nodes ($dF$: hand force, $dT$: joint torque). Vertical links characterize the elastic element (i.e. stiffness $K$ and admittance $A$) while horizontal link characterize the geometric element i.e. the Jacobian encoding the relationship between hand displacement and joint rotation (that can be learnt by babbling). The goal ($X_d$) induces a disturbance compensation action, eliciting a reconfiguration of the body to a new equilibrium so as to minimize the prediction error between the forward model output ($x$) and the goal ($X_d$). In this sense the computational basis closely resonates with predictive coding and free energy principle. The network can be used to derive the necessary motor commands ($dq$) for over action or anticipate the consequence. This provides vital information to either trigger overt action or reason about exploiting affordances in the environment like a rake tool (panel E). Panel D: it shows a composite forward/inverse model, for bimanual coordination with a connected tool (like a steering wheel or a toy crane in panel F). Here the network of panel B for one arm acts as a sub component, with the new addition of tool space represented in the same fashion with a force and displacement node. The tool Jacobian that encodes relation between the end effector and tool effector has to be learnt. Note that all computations are well posed: from the joint angles to the simulated position of the hand to from the force exerted by the two hands to the position of the tool effector (while the inverse will lead to infinite solutions, necessitating complex optimizations).
Annexure IV
From object-action to property-action: Learning causally dominant properties through cumulative explorative interactions

Vishwanathan Mohan *, Ajaz Bhat, Giulio Sandini, Pietro Morasso

Robotics, Brain and Cognitive Science Dept, Istituto Italiano di Tecnologia, Genova, Italy

Received 7 November 2014; accepted 7 November 2014

Abstract
Emerging studies from neuroscience in relation to organization of semantic memory in the brain provide converging evidence suggesting that conceptual knowledge is organized in a distributed fashion in ‘property specific’ cortical networks that directly support perception and action (and were active during learning). Though learning ‘object-action’ affordances and using such knowledge for prediction and planning is an active topic in cognitive robotics, this article urges to go beyond and look at ‘property-action’ networks instead. To this effect, a brain guided framework for distributed property specific organization of sensorimotor knowledge for humanoid iCub is presented. Two simple learning rules namely ‘elimination’ and ‘growth’ are proposed to compare top down anticipation and bottom up real experience to abstract underlying causal relations. An engaging scenario how the robot cumulatively learns and abstracts causally dominant properties that influence motion of various objects when forces are exerted on them is used to validate the neural architecture. The implicit advantage is that such learnt ‘property-action’ relations can be effortlessly generalized to a domain of objects for which the robot need not have any past experience/learning but nevertheless share the ‘property’. Further, the study has relevance in both better understanding how common causal relations can be cumulatively learnt, represented and exploited, to providing novel embodied frameworks for analogical reasoning.

© 2014 Published by Elsevier B.V.

Introduction

Affordances are the seeds of action. Being able to identify and exploit them opportunistically in the 'context' of other-
wise unrealizable goals is a sign of cognition. Despite incessantly encountering novel exemplars of objects we have never interacted with before we can often anticipate about what can be done with them and even exploit them to realize our goals. Based on her past experience of bending flexible pipe cleaners about a year back, Betty the new Caledonian crow quickly fashioned a hook out of a piece of wire to pull her dinner basket trapped in a transparent vertical tube (Weir, Chappell, & Kacelnik, 2002). Studies from animal cognition (Visalberghi & Tomasello, 1997; Whiten, Mcguigan, Marshall-Pescini, & Hopper, 2009) indicate that several other primates are able to flexibly reason about physical causality, exploit inherent properties of objects available in the environment: for example choosing a tool of a right length to push out a trapped reward (among others). How by cumulative and explorative interactions with different objects in the world, task relevant physical and causal relations are both abstracted and exploited flexibly in novel contexts is a problem that presents challenges both in terms of representation and learning. In this context, cognitive robots offer a unique opportunity to reenact the gradual process of cumulative learning and investigate the underlying computational basis. The value is both intrinsic i.e. better understanding our own selves and extrinsic i.e. creating a range of artifacts that can flexibly assist us in the environments we inhabit and create.

Learning ‘object-action’ relations and using such knowledge for prediction and planning is becoming an active topic of study in embodied robotics with approaches ranging from probabilistic Bayesian models, to neural associative networks and symbolic formalisms (Krüger et al., 2011; Montesano, Lopes, Bernardino, & Santos-Victor, 2008; Montesano, Lopes, Bernardino, & Santos-Victorm, 2007). Despite intriguing attempts, both applicability and generalization of the methods to novel contexts and the necessity to facilitate cumulative learning (like natural cognitive agents) have been known bottlenecks. In parallel, emerging results from functional imaging are beginning to provide useful information as to how conceptual knowledge about object concepts and causal relations is organized in the brain (Bressler & Menon, 2010; Buckner, Andrews-Hanna, & Schacter, 2008; Martin, 2007, 2009; Meyer & Damasio, 2009; Patterson, Nestor, & Rogers, 2007). The main findings emerging from these results is that conceptual information is grounded in a “distributed fashion” in “property specific” cortical networks (Martin, 2009; Patterson et al., 2007) that directly support perception and action (and that were active during learning). Same set of networks are known to be active both during real perception/action, imagination or lexical processing (Martin, 2007; Meyer & Damasio, 2009). Further, there is a fine specialization of areas representing conceptual information related to animate vs. inanimate objects as evident from functional imaging and TMS studies on both normal subjects and semantic dementia patients (Buckner et al., 2008; Patterson et al., 2007). It is also now known that “retrieval” or reactivation of the conceptual representation can be triggered based on partial cues coming from “multiple modalities”: for example sound of a hammer retro activates its shape representation (Meyer & Damasio, 2009), presence of a real object (banana) or a 2D picture of it can still activate the complete network associated with the object (and that was active during learning of it in the first place). These results provide valuable insights to guide development of computational frameworks for organizing information related to perception—action and foster learning of causal relations, importantly in an embodied and cumulative learning setup. The present article is an ambitious adventure in this direction.

A biologically inspired framework for distributed property specific organization of sensorimotor knowledge for humanoid iCub is presented with an emphasis on learning “property-action” relations. The implicit advantage is that such learnt “property-action” relations can be effortlessly generalized to a do-main of objects for which the robot need not have any past experience/learning but nevertheless share the “property”. An engaging scenario how the robot cumulatively learns and abstracts causally dominant properties that influence motion of various objects when forces are exerted on them is used to validate the neural architecture. It is known from studies on animal behavior that different species have different levels of understanding of the causality related to this task (Visalberghi & Tomasello, 1997; Whiten et al., 2009). In addition to the multiple utilities of the “push/pull” action itself in the context of day to day interactions with objects, what makes it interesting is the sheer range of physical concepts that have to be “learnt” and “abstracted”. For example, it has to be learnt that contact is necessary to push, object properties influence pushability (balls roll faster than cubes and it does not matter what is the color of the ball or the cube), pushing objects gives rise to path of motion in specific directions (the inverse applies for goal directed pushing), pushing can be used to support grasping, bring objects to proximity. The requirement to capture/learn such a wide range of physical concepts through cumulative “playful interactions” of the robot with different objects makes this task both interesting and challenging.

The rest of the article is organized as follows: Section ‘Distributed “property specific” organization of sensorimotor information and the basic pushing forward/inverse model’ describes the organization of sensorimotor information in iCub taking guidance from emerging results from neurosciences. How the basic Pushing forward/Inverse model i.e. anticipating how objects move and inversely generating goal directed pushing actions is learnt is illustrated. Section ‘Cumulatively abstracting “causal dominant properties” related to pushing’ introduces two learning rules namely elimination and growth that augment the distributed property specific organization of sensorimotor information to facilitate the robot to abstract properties that are causally dominant through cumulative explorative interactions. A discussion concludes.

Distributed “property specific” organization of sensorimotor information and the basic pushing forward/inverse model

Emerging trends from the fields of connectomics, functional imaging studies in relation to organization of semantic and episodic memory in the brain (Bressler & Menon, 2010; Buckner et al., 2008; Martin, 2009; Patterson et al., 2007) now provide numerous insights to guide development of brain guided computational framework for organization
and use of memory in cumulatively learning embodied robots. Fig. 1 shows a block diagram of how sensorimotor information is organized in the proposed computational framework with the focus of the “Pushing” sub-network. 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 (Cai, Werner, & Matasm, 2013). Word information is an additional input coming from the teacher either to issue user goals or interact with the robot. Results of perceptual analysis activate various neural maps organized in a property specific fashion ultimately leading to a distributed representation of the perceived object in the connector hub (layer 2). In this sense, a red cylinder and red cube will have identical activity in the color map, but different activity in the shape map, ultimately leading to a distributed representation in the connector hub. How these self-organizing maps are learnt is out of scope for this article, but uses standard techniques (Fritzke, 1995; Kohonen, 1995). Interested reader is referred to a recent article (Mohan, Morasso, Sandini, & Kasderidis, 2013) that goes into formal details of the algorithms with experimental results.

An interesting aspect of such kind of organization is that as we move upwards in the hierarchy information becomes more and more integrated and multimodal and as we move downwards information is more and more differentiated to the level of perceived properties. The connectivity between hubs and property specific maps is essentially bidirectional hence allowing information to move “top down, bottom up or in cross modal” fashion. The pushing sub-network, connected to the object hub consists of two networks: one a growing SOM encoding average displacement of an object per unit force exerted (averaged over several trials of interactions) and another representing distributed coding of direction. These networks together enable learning and representing the forward/inverse model for pushing while the property specific organization facilitates learning which properties are causally dominant. Right panel shows the robot interacting with different objects during the episodes of cumulative learning.

44 V. Mohan et al.

Fig. 1 Left panel shows a block diagram of how sensorimotor information is organized. Results of perceptual analysis activate the layer 1 neural maps organized in a property specific fashion ultimately leading to a distributed representation of the perceived object in the connector hub (layer 2). An interesting aspect of such kind of organization is that as we move upwards in the hierarchy information becomes more and more integrated and multimodal and as we move downwards information is more and more differentiated to the level of perceived properties. The connectivity between hubs and property specific maps is essentially bidirectional hence allowing information to move “top down, bottom up or in cross modal” fashion. The pushing sub-network, connected to the object hub consists of two networks: one a growing SOM encoding average displacement of an object per unit force exerted (averaged over several trials of interactions) and another representing distributed coding of direction. These networks together enable learning and representing the forward/inverse model for pushing while the property specific organization facilitates learning which properties are causally dominant. Right panel shows the robot interacting with different objects during the episodes of cumulative learning.

The acronym Darwin stands for the ongoing EU funded project Dexterous Assembler Robot Working with embodied Intelligence (www.darwin-project.eu).
i.e. the red container to be). If top down activation of the property specific maps resonates with their bottom up activation (through perceptual stream) this is sufficient to lead to the inference that the novel object being perceived is most probably the one the user requested to grab (Mohan et al., 2013). What is relevant as far as the present article is concerned are mainly two things: (1) The bottom up processing leads to a distributed representation of the perceived objects (in relation to its perceptual properties color, shape, size) in the object connector hub that identifies the object (i.e. in other words coding for 'what is it'). (2) Due to reciprocal connectivity between the hubs and property specific maps, it becomes possible to go beyond ''object-action'' and learn things at the level of ''property-action'' instead.

Different objects move in different ways when force is exerted on them, some do not move too. By interacting with various objects, the goal of the robot is to learn a general forward/inverse model for 'pushing action': i.e. being able to predict how an object will move when pushed (forward model) and being able to generate goal directed pushing actions in order to displace an object to a desired location. When presented with any 'object', different property specific neural maps are activated bottom up leading to a distributed representation of the concerned object in the object connector hub. Since object properties influence pushing, activity in the object connector hub influences the pushing forward/inverse model and hence is bidirectionally connected to it (connectivity learnt by experience). The push sub network is represented using two neural maps:

1. A growing SOM learning ''average displacement of an object per unit force'' and
2. That represents a distributed coding of direction in which the object is moving.

The former neural map is empty to start with and is gradually grown as the robot interacts with different objects. Note that ''average displacement per unit force'' basically measures in abstract terms the ''mobility'' of the object when a certain amount of force is exerted on it. Inversely, this information allows the robot to predict how an object will move when force is exerted on it (useful while generating goal directed pushing). For every object presented, the robot is allowed to explore displacing it to a distance of 15 cm (in eight different directions. Averaging the result of this experience, the parameter $P_t$ (average displacement/unit force) for the neuron ''i'' coding for a particular object is estimated. Growth in this neural SOM only takes place if there is a contradiction between ''the robots anticipation of how a novel object might move'' and ''how it actually moves in reality''. Cubes, cylinders and balls of different colors and sizes (some heavy ones), few MECCANO blocks (from the MECCANO 2+ kit for 2 year olds) were presented gradually.

The inverse operation i.e. pushing an object to a desired location can be achieved through the following 4 steps (that depict how the learnt forward model (how something moves) can be used to perform inverse operation (how the object can be moved as desired):

1. Detect and localize the current position $X(x, y, z)$ of the object and the target $X_t(x_t, y_t, z_t)$ (where the object has to be displaced).
2. Compute the desired direction ''$\theta$'' to push using information on $X$ and $X_t$. This activates the neurons in the motor map responsible for directional coding. Based on the instantaneously computed direction, we also see distributed activity the 8 neurons coding for different directions.
3. Anticipate the average displacement of the object in the desired direction for an incremental iteration where unit force is applied on it ($P_t$).
4. Iterate steps 1–3, to get a virtual trajectory in space. The synthesized virtual trajectory can then be fed as a moving point attractor to the iCub action generation system (Mohan & Morasso, 2011) to compute the motor commands for the body chain executing the goal directed pushing. The process (1–4) is like smooth sliding of an object along a predicted trajectory created due to past experience. Fig. 2 shows the virtual trajectories (moving attractors) and real trajectories during goal directed pushing of a red cuboid and a ball. Activity in the neurons responsible for distributed coding of direction during the synthesis of the motor actions is shown at the top. Note that while pushing a ball the end effector needs to be displaced just by a small amount along an estimated virtual trajectory (green trajectory) like kicking a football to the goal. Cubes move more uniformly with the displacement of the end effector and have to be pushed gradually to the destination. Note that, time is also implicitly represented. This is because every ''dot'' in the virtual trajectory represents iteration in time: the virtual trajectory is very short while pushing the ball, almost uniform while pushing the small cube and longer when pushing the heavy/large objects.

**Cumulatively abstracting ''causal dominant properties'' related to pushing**

The previous section described how the robot by interacting with different objects can learn the Pushing forward inverse model i.e. predicting how an object will move when force is exerted on it or inversely generating goal directed pushing actions to displace an object to a goal location. In this section, we will present results on how the proposed bio inspired framework for sensorimotor organization further allows the robot to additionally learn and abstract ''which properties are causally dominant'' for the task. To this effect we introduce two simple learning rules that facilitate abstraction of causally dominant properties by comparing the present experience with a recalled past experience. Let us consider that $\Delta$ property is the difference in activity in a property specific map (Fig. 1) when activated bottom up through sensory layer during present experience and when activated top down from the pushing SOM while recalling a past experience. Let $\Delta$ Contradiction be the difference between the robots anticipation of how an object might behave and the real observed behavior. Then the two learning rules are as follows
Elimination rule: If change in "property" does not cause "contradiction" between anticipation and real experience, then that property is not causally dominant. The result is a drastic reduction in the connection strength between associated maps (hence reducing the capability of activating each other).

Growth rule: If change in "property" causes "contradiction" between anticipation and observed behavior, then that property is causally dominant. Contradiction implies that there is something new that has not been learnt in the past episodes of experiences.

Two different scenarios are presented in the subsections that exploit the elimination and growth rule to enable the robot to abstract causally dominant properties by cumulative interactions with objects.

**Small cubes: learning that "color" of objects does not affect the way they move**

Fig. 3 shows the activations in the various neural maps when the robot explores pushing cubes. The first 3 columns show the activity in the color-word-shape maps. The fourth column shows the distributed activity in the object hub representing in a distributed manner "what the object is". In episode 1, there is no anticipated activity in the pushing SOM (since there are no connections between the Pushing SOM and the connector hub because no experience has been gained before and nothing is known about the behavior of the cube). After experience of pushing red cube in different directions, there is one neuron (winner) in the pushing SOM that is coding the average displacement of the red cube when pushed by the robot. For cubes this ratio is approximately 0.8, averaged over 10 trials of pushing. The connections between the active neurons in the "object hub" (column 4) and the winner in the pushing SOM are also built. This is standard Hebbian learning with the effect that in future if a "small red cube" is presented, the robot can both anticipate how it will move and also generate goal directed actions to push it to the desired location (based on the steps outlined in the previous section). In the next episode the robot is presented with a "small blue cube". Bottom up activity form the sensory layer activates different neural SOM's (episode 2, column 1–3) which activate the object hub (column 4). Because of episode 1, there exists now some connectivity between the object hub and the Pushing SOM, enabling the object hub to activate the "only" existing winner in the pushing SOM (the robot does not know anything about small blue cubes, but it knows something about small red cubes). So it anticipates that blue cubes should also behave the same way and this indeed turns out to be the case from real experience. The
anticipation of the robot correlates with what is experienced by direct interaction with the new object. Due to reciprocal connectivity between the pushing SOM and the property specific maps, it becomes possible to move top down from the pushing SOM and reconstruct from memory the activity in the property specific maps that were active when this experience was originally encoded. As seen in row 3, the past experience of pushing the red cube is reconstructed successfully. Comparing row 2 (bottom up: present experience with the blue cube) and the past experience with the red cube (reconstructed top down), it is easy to observe that elimination rule applies (change in property not causing a contradiction in anticipation).

What the robot abstracts is that color of objects do not affect their mobility when forces are exerted on them (row 4). The net effect of assimilation is that the connectivity of the push SOM with the color SOM is now reduced drastically and they no longer retroactivate each other. Row 4, shows the activity in different maps when we move top down from the pushing SOM after elimination rule is applied. Note that the color map is no longer activated as the robot has now learnt that color is a irrelevant property as far as how objects move when pushed. This is in fact equivalent to querying the robot ‘which object moves like this?’ note that the word map generates linguistic outputs. Before assimilation the answer would have been ‘small red cubes’, after assimilation the answer is just ‘small cubes’!

Small cylinders: learning that ‘Shape’ is a dominant property

In Section ‘Small cubes: learning that ‘color’ of objects does not affect the way they move’ the robot learnt colors of objects do not causally affect their mobility when forces are exerted on them. Now the robot is presented with a small green cylinder. The loop form bottom up and top
down is shown in row 1–2 of Fig. 4. The flow of information is indicated by arrows. Row 1 Bottom UP: To start with we move from property specific maps to the object hub and then to the Pushing SOM. As seen in column 5, the robot anticipates that small cylinders may also move like small cubes that are all it knows so far. But real experience with cylinders reveals a different behavior, they roll much faster, displacement/unit force is much higher as compared to cubes. But we need to know “what caused this contradiction?”. The trick is to go top down to reconstruct from memory the past experience that resulted in the wrong anticipation and compare it with the new experience (row 2). Note that color map is no longer activated top down, due to assimilation in the past episode. Further directly comparing top down with bottom up it is possible to infer which property is causing the contradiction (shape map shows change). Growth rule applies here because a change in property causes contradiction between anticipation and observed behavior. As a result of application of growth rule, there is now a new neuron in the Pushing SOM that is coding for how cylinders move when force is exerted on them. Appropriate connections between the new winner in the push SOM and the object hub are developed.

In sum, cumulatively interacting with different objects and exploiting the dissonance or resonance between real experience and anticipations, it becomes possible to learn which properties are causally dominant for a particular task. Applying the elimination and growth rule, the robot in few episodes of interaction learnt that color of objects do not matter, but shape does and cylinders roll much faster than cubes when forces are exerted on them. Further, the forward model anticipating how objects move naturally solves the inverse problem of generating goal directed pushing of any object to a target location as seen in Section “Distributed “property specific” organization of sensorimotor information and the basic pushing forward/inverse model”. The same mechanism can be used to eliminate any irrelevant property like the end effector used (stick, left hand, right hand etc.), the starting and initial positions from where the object is in fact pushed and so on. All of these are not “casually dominant” as far as the task is concerned, the learner can abstract all of this by the process of assimilating what is experienced in the present with what was experienced in the past.

**Discussion**

To wipe off a spider web on the top most corner of a room, it does not matter if a red colored broom or yellow colored broom or even a long stick is deployed. Any object that has
the specific property "length" in this case will suffice. Similarly, colors of objects do not affect the way they move when pushed but shape and size do. At the same time, color does play a causally dominant role in numerous other tasks (for example, traffic rules). The article explores the issue of how causally dominant properties relevant to different tasks can be learnt through cumulative episodes of interactions. Part of the problem is solved by taking guidance from emerging results from brain science providing converging evidence 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. To supplement such property specific organization, two learning rules namely elimination and growth where proposed, that modulate the connectivity between associated neural maps based on contradictions between the robots top down anticipations and the bottom up experience. A well investigated scenario from animal cognition (Visalberghi & Tomasello, 1997) i.e. pushing was used to illustrate the basic framework. Both how the robot acquires the forward model i.e. anticipating how objects move when forces are exerted and the inverse model i.e. generating goal directed pushing actions with different objects was described. Going beyond the acquisition of the basic forward/inverse model, how causally dominant properties can be abstracted within the proposed framework was illustrated. An interesting aspect of the proposed framework is that it computationally explains why top down and bottom up information flow must share computational/neural substrates. Several studies from functional imaging related to organization of conceptual information in the brain support that the same set of cortical networks are known to be active both during real perception/action, imagination and lexical processing. Our framework points out such a shared basis provides a computationally efficient means to directly compare and put into context the present experience with what has been learnt in the past. At the same time, learning and anticipation goes hand in hand (and continuously in life time of the learner): more experience driving better anticipation and inconsistencies in reasoning driving new learning.

Acknowledgements

The research presented in this article is supported by Robotics, Brain and Cognitive sciences dept. IIT, the EU FP7 project DARWIN (www.darwin-project.eu, Grant No: FP7-270138) and US Dept. of defence Grant (W911QY-12-C0078).

References


