Adaptive Behavior

A review of long-term memory in natural and synthetic systems

A review of long-term memory in natural and synthetic systems

Rachel Wood, Paul Baxter and Tony Belpaeme

Abstract
Memory may be broadly regarded as information gained from past experience that is available in the service of ongoing and future adaptive behavior. The biological implementation of memory shares little with memory in synthetic cognitive systems where it is typically regarded as a passive storage structure. Neurophysiological evidence indicates that memory is neither passive nor centralized. A review of the relevant literature in the biological and computer sciences is conducted and a novel methodology is applied that incorporates neuroethological approaches with general biological inspiration in the design of synthetic cognitive systems: a case study regarding episodic memory provides an illustration of the utility of this methodology. As a consequence of applying this approach to the reinterpretation of the implementation of memory in synthetic systems, four fundamental functional principles are derived that are in accordance with neuroscientific theory, and which may be applied to the design of more adaptive and robust synthetic cognitive systems: priming, cross-modal associations, cross-modal coordination without semantic information transfer, and global system behavior resulting from activation dynamics within the memory system.

1 Introduction

There is nothing new in the idea that long-term memory is an important aspect of cognition. However, the question of how memory should be implemented in synthetic cognitive systems remains relatively under-explored. This article reviews the current state-of-the-art in artificial implementations of memory and the extent of their grounding in biology, highlighting a number of shortfalls in the capabilities of of synthetic memory systems. A novel methodology for the application of grounded and contextualized biological inspiration to the construction of artificial memory systems is proposed. While the application of this methodology ultimately seeks to produce more general, adaptive and robust cognitive architectures, in the first instance (and in the context of this article) it enables a reevaluation of existing implementations to inform further development. The principles of the proposed approach are applicable to artificial cognitive systems in general.

Memory is generally discussed in terms of what it does, that is, as a system enabling the storage of information acquired from agent–environment interactions and subsequent recall in the service of behavior. While being very general, this description highlights a number of assumptions that are reflected both in approaches to biological memory and in implementations of memory

for artificial intelligence (AI). The first and perhaps most basic of these assumptions being the division of memory from cognition based on a notion of "storage" as the fundamental purpose of memory. Functionally, memory can be divided into three distinct processes: encoding, storage, and retrieval of information. This paradigm of memory function is ubiquitous in cognitive science and more specifically, in AI, and can be seen as an extension of the implicit but pervasive "mind-as-computer" metaphor (Bickhard, 2009). However, there is a growing body of empirical evidence that is inconsistent with this account and that challenges the assumptions of the "computational mind," rendering it increasingly unsuitable as an approach to understanding what memory is and how it works (e.g., Bar, 2007; Baxter, 2010; Glenberg, 1997). The perspective described in this article builds upon this trend, and proposes an approach to building memory systems that conforms to ethological principles: the explicit

University of Plymouth, Plymouth, UK

Corresponding author:
Rachel Wood, School of Computing and Mathematics, University of Plymouth, Portland Square A216, Plymouth PL4 8AA, UK
Email: rachel.wood@plymouth.ac.uk.

embedding of a model in the context of a complete cognitive system.

In artificial systems where a memory system is explicitly invoked, a *passive storage* structure is typically used. However, memory function may also be *implicit* such as in a learning mechanism (for example using some form of subsymbolic network structure) wherein prior experience shapes future behavior. Interestingly, this type of memory is not often acknowledged as such. However, in both cases memory can be described as present, at least in some sense.

In the domain of human–robot interaction, the requirements for a memory system are more specific. In addition to the capacity to adapt to the characteristics of a user, a nontrivial level of *social* interaction requires that the agent should be able to use both semantic and episodic information (such as remembering a name, and the circumstances of the last interaction). To date, this type of information has typically been encoded in symbolic form, and held in a storage module invoking the functionality of an *episodic memory* system. However, this approach to implementation is arguably inadequate in terms of flexibility and robustness, because the system must generally use a task-specific encoding scheme.

Given that the ultimate aim of this work is the development of synthetic cognitive systems, a methodology is proposed based on the view that simply taking inspiration from the taxonomy of biological memory is insufficient. While modeling plays a central role in understanding the data on biological memory, the

Figure 1. Outline of a methodology: it is proposed that the manner in which biological inspiration is taken has fundamental consequences on how this inspiration is interpreted regarding implementation. Whilst modeling and other neuroethological techniques form an important intermediary stage, it is also necessary to question the implicit assumptions that these methods use. It is in this manner that memory is interpreted in this article

context in which models are constructed is itself an aspect worthy of consideration (Figure 1). Thus, the "traditional approach" may be described as making use of observed biological function at a behavioral level to inform the development of algorithms that replicate this functionality, though with no necessity for mechanistic similarity. The "updated traditional approach" goes a step further by explicitly incorporating aspects of both function and mechanism, with the latter being derived from modeling. Here we propose that the implicit assumptions upon which these models are based also need to be taken into consideration, as these have fundamental consequences for the actual implementation of a synthetic system.

When applied to the problem of building synthetic memory systems this perspective enables a new theoretical framework to be developed and to inform construction of cognitive systems. Thus, properly contextualized modeling of a biological function (such as memory) provides design principles for implementation in a manner consistent with its original biological context, rather than as a monolithic functional module. The benefit of this perspective is that the flexible and robust nature of these functions in biological agents can be exploited in synthetic systems. If memory is considered from the outset as fundamentally embedded in, and integrated with, a wider cognitive architecture, adaptivity is thus a property of the system as a whole, with memory here proposed to be the central component.

Table 1 provides outline definitions of terms that will reoccur throughout this review. These definitions are elaborated further in the relevant sections but are given here in the interests of clarity.

The structure of this article is directly informed by the proposed methodology. Initially, a review of memory in biological systems shows that design principles can be derived from various levels of abstraction and investigation in psychology and neuroscience, and that a growing amount of empirical evidence indicates that memory should be considered as a distributed and active component of cognition (Sections 2 and 3). The methodology described above is applied to the problem of synthetic episodic memory, and a set of operational principles are derived based on the notion of a memory-centered cognitive system, and a more naturalistic model of memory function that is proposed to enable a closer approximation of those desirable properties of biological agents, that is, adaptive, flexible, and robust (Section 5).

2 Memory in biological systems

The neat box-diagram presentation of memory commonly found in psychology textbooks belies the interactive complexity required to implement its functions. The structure and functions of memory are underpinned by interactions between mechanisms distributed

Table 1. Definitions of key terms

Term	Definition
Memory	The capacity to use previous experience to inform subsequent behavior
Long-term memory (LTM)	Temporally indeterminate, independent of specific task demands
Short-term memory (STM)	A functionally distinct memory structure, finite in capacity and retention period, bounded by context and task demands
Working memory (WM)	A variant of the STM concept providing temporary retention of task relevant information
Procedural	Non-consciously accessible memory e.g., skilled motor behavior, habits, and stimulus–response conditioning
Declarative	Consciously accessible information
Episodic	Personally experienced event information, spatially and temporally organized
Semantic	Context independent information, facts and concepts

across multiple brain regions, which have yet to be fully understood. Similarly, these functions can be decomposed in a variety ways, the validity of which are still the subject of considerable debate. A growing body of empirical evidence has led away from notions of memory as a unitary facility toward approaches based on the interaction (and differentiation) of multiple memory subsystems. Human memory is generally characterized in terms of discrete encoding, storage, and retrieval functions with storage structured into dissociable subsystems organized by content. Still, the issue of how to identify the subcomponents of memory and their organization remains an area of much active research. This section will introduce and review the functional subsystems of memory briefly reviewing key themes in long-term memory research and looking at the mechanisms of memory in the brain. Thereafter the discussion will focus on episodic memory, that is, consciously accessible recall of personally experienced events. Episodic memory is widely regarded as a uniquely human cognitive competence[1] and is arguably the most complex form of mnemonic function entailing the spatiotemporal structuring of event information. Episodic memory is also an important factor in complex social behavior, which depends on the capacity to recall and utilize structured *who, what, where, when* information about previous interactions.

The taxonomy of memory is both functional and content-based with systems differentiated according to temporal scale, the extent to which content is consciously accessible and phenomenological characteristics. The primary division imposed on human memory separates processes for short and long-term recall. Hebb (1949) is credited with the original division of memory into long- and short-term systems. Hebb characterized short-term memory (STM) as based on transitory electrical activity in the brain with long-term memory (LTM) consisting in more lasting neurochemical adaptations. This temporal division of memory systems has also been reflected in proposed mechanisms for forgetting, that is, via trace-decay in STM and interference between stored memories in LTM (Baddeley,

2001). The division of long- and short-term memory systems has found empirical support in double dissociation procedures.[2]

Another key distinction in memory function centers on the issue of conscious access to content; Ryle's "knowing how" versus "knowing that" (Squire, 2004). In LTM, non-conscious, procedural (or implicit) memory underpinning skilled behavior, habits, and conditioning, is differentiated from consciously accessible, declarative (or explicit) memory for facts and information (Cohen & Squire, 1980; Graf & Schacter, 1984). The differentiation of declarative and procedural memory originally found support in studies showing intact priming and motor skill in amnesic patients with declarative memory deficits (Baddeley, 2001). However, additional evidence indicates that an apparently disparate collection of procedural memory functions may also be preserved in such patients suggesting that a simple dichotomization of long-term memory may be insufficient. Thus, the procedural–declarative division was further elaborated to provide a framework comprising multiple memory subsystems with the term *non-declarative* covering a set of additional, non-conscious memory functions comprising adaptations to previously learned behavior systems. Non-declarative procedural memory is thus subdivided into four types: skills and habits, priming, classical conditioning, and non-associative learning (Baddeley, 2001).

Declarative memory has been characterized as encoding the unique characteristics of a single input set, while procedural memory generalizes across multiple input sets on the basis of their common elements (Squire, 2004). Squire suggests that the evolution of multiple memory systems reflects different, incompatible purposes, the gradual learning of motor skills having quite different requirements from (for example) one-shot learning of a temporal sequence of sensory inputs. Within declarative memory further distinctions are drawn between systems supporting the retention and retrieval of personally experienced episodes and those predicated on knowledge about the world, that is, semantic content. Finally, psychological and

neuropsychological approaches to memory also highlight functional divisions between processes for detecting previously encountered stimuli (familiarity) and processes for recall of the context in which particular stimuli were encountered (recollection).

Atkinson and Shiffrin (1968) originally proposed a three-stage model of human memory. According to this multi-store account, input passes from sensory memory into a short-term store gated by attention. Items in short-term memory are maintained via rehearsal or forgotten and the length of time an item remains in STM determines the probability of transfer to LTM. The model was influential but is flawed: for example, it suggests that patients with impaired STM should have concomitant deficits in LTM and complex information processing, a prediction that is not supported by the evidence. Retention in LTM has also been found to be dependent on depth of processing rather than length of STM rehearsal (Craik & Lockhart, 1972).

Baddeley and Hitch (1974) subsequently proposed a broadly similar model but with the addition of a working memory module replacing the unitary STM structure with functionally specific subsystems. In their formulation, working memory (WM) comprises three components: the phonological loop for auditory information maintained by rehearsal, the visual–spatial sketch pad storing visual and spatial information and a central executive providing coordination and control of attention. This model was among the first to make explicit the division of memory (as storage) and cognition (control). As a means to address the binding problem and account for "chunking" effects, Baddeley (2000) later reformulated the model to include an extra unit termed the episodic buffer. This unit organizes multi-modal input from WM along with activated content from LTM into "episodes" and mediates their subsequent transfer to and from LTM. Cowan (1988) rejected the notion of WM as a discrete system in favor of a model in which WM comprises an activated section of LTM. In this unified, levels-of-activation model, the content of awareness is a subset of WM, mediated by attention. While each of these models has spurred a great deal of research looking to confirm or discount the functional decomposition of subsystems proposed, psychological and neuroscientific evidence to conclusively verify any one model over and above another has yet to be found.

2.1 Major themes in long-term memory research

Research into LTM has centered on key themes such as the extent to which semantic and episodic memory contents are supported by discrete systems and the level of interaction between subsystems, that is, familiarity detection and recollection in LTM function. Another central question concerns the role of conscious awareness of recall in episodic memory and the capacities of some non-human animals to retain temporally structured information about experienced events.

2.1.1 Differentiating semantic and episodic memory. The division of declarative memory into episodic and semantic content raises questions of the extent to which these forms of memory are supported by separable mechanisms in the brain. There is behavioral, anatomical, and developmental evidence to support a mechanistic differentiation of episodic and semantic memory. At the behavioral level, anterograde amnesia[3] disrupts memory for new autobiographical information and thus is associated with selective impairment of episodic memory. Bilateral hippocampal damage can result in anterograde amnesia, while damage to other temporal lobe structures has been shown to increase the degree of impairment (Aggleton & Brown, 2006). Anatomical evidence coupled with selective patterns of memory impairment associated with damage to different regions of the medial temporal lobe and diencephalon indicate the specific role of an extended hippocampal pathway in episodic memory. In development, perinatal and childhood hippocampal damage following oxygen deprivation[4] demonstrates a strong dissociation between episodic memory, which is often severely impaired, and semantic memory, which is generally preserved. This pattern is found regardless of the age at which damage was sustained and appears to show episodic memory is functionally dependent on the hippocampus (Vargha-Khadem, Gadian, & Mishkin, 2001; Vargha-Khadem et al., 1997). A double dissociation between declarative memory systems has also been demonstrated with evidence for the opposite pattern of impairment, that is, preserved episodic function with impaired semantic memory (Temple & Richardson, 2004).

Semantic and episodic memory systems can also be differentiated phenomenologically. Tulving (2002) places the recollective experience at the center of the definition of episodic memory by stressing its fundamentally autonoetic quality. Episodic memory concerns personally experienced events occurring in a specific place at a specific time. Minimally, this entails the rememberer being aware that recalled events happened in the past, though more properly, episodic memories can be characterized as events recalled in a temporal framework that preserves their sequence. The autonoetic quality of episodic remembering is inherently perspectival, invoking notions of the self as a locus of experience and episodic memory as self-recall. The content of episodic memory can therefore be differentiated from semantic content by reference to its embedding in time. Episodic content is structured in time both by being explicitly marked as a past state of affairs and by (either explicitly or implicitly) encoding the temporal sequence of events. By this characterization, episodic

memory, with its requirement for introspection and conscious awareness of remembering the past, cannot be proved in non-human species (Aggleton & Brown, 2006). However, episodic-like memory in animals has been investigated by means of studies focusing on simultaneous what, where, when learning (e.g., Clayton, Griffiths, Emery, & Dickinson, 2001; Clayton & Russell, 2009; Griffiths, Dickinson, & Clayton, 1999), thereby providing a purely behavioral perspective on episodic memory function (see Section 3.1).

2.1.2 Recognition memory and the roles of familiarity and recollection in episodic recall.

Much recent research on episodic memory has centered on dissociation of recognition processes supporting familiarity detection (knowing) and recollection (remembering). As a functional subset of declarative memory, recognition refers to the capacity to identify some aspect of the environment as having been encountered before. This type of memory function is composed of two dissociable elements: familiarity detection and recollection. The two components are differentiated experimentally via process-dissociation methods in which subjects are asked to make introspective judgments on the extent to which they "remember" encountering the test stimulus as opposed to "knowing" that the stimulus is familiar. The data indicates that while episodic memory entails recollection, familiarity detection also contributes. Dual process models of recognition memory propose familiarity (knowing) and recollection (remembering) as independent processes served by different brain areas. However, two-process accounts are not universally accepted with single-process models proposing familiarity, that is, the feeling of knowing as simply reflecting a lower degree of confidence than recollection, that is, remembering (Aggleton & Brown, 2006).

Event-related brain potentials (ERPs) have been widely used to investigate the roles played by familiarity and recollection processes in episodic memory. Much data has been obtained through manipulation of "old/ new" effects, contrasting brain activity for stimuli recognized as being old versus correct rejections of new stimuli. The ERP data on recognition memory demonstrates the interaction of physiologically dissociable processes supporting the familiarity detection and recollection (Curran, 2000; Curran, DeBuse, Woroch, & Hirchman, 2006); for a review see Allan, Wilding, and Rugg (1998). Specifically it illustrates the distribution of memory processing in the brain and that the function of recognition is provided dynamically through interaction between differential patterns of frontal and parietal activation. These patterns are not observed in discrete regions reserved exclusively for "memory store" but rather can be observed as activations of regions implicated in numerous aspects of cognitive function.

2.1.3 Episodic-like memory.

If episodic memory is defined by conscious experience of remembering and self-recollection then, because there are no accepted, nonverbal, behavioral markers for conscious experience in non-human species, their capacity for episodic recall cannot be tested. However, Tulving's original (1972) definition characterizes episodic memory as storing temporally organized events in such a way as to also preserve their temporal–spatial relations. By this definition, episodic memory consists of the recollection of linked *what, where* and *when* information, the unique characteristic of this type of memory being the explicit temporal and spatial linking of events, rather than a conscious awareness of their having been personally experienced at some point in the past (Clayton et al., 2001).

Recent research on what is termed *episodic-like* memory in animals, has used paradigms that explicitly test recall of the spatiotemporal components of individual learning episodes. This approach has shown that a number of species do encode, store, and recall episodic event information. For example, brood-parasites (birds which lay their eggs in the nests of others) identify a suitable nest and time the laying of eggs so that they are accepted by the host. This entails recall of information about the state of a nest (e.g., number of eggs already present), its location, and the time visited (so that the parasitizing bird can return to add its own egg to the host clutch at the right time).

Episodic-like memory has also been extensively studied in the context of food-caching behavior in animals such as scrub-jays (Clayton, Dally, & Emery, 2007; Clayton & Dickinson, 1998). The scrub-jay is a scatter-hoarding bird that hides hundreds of individual food items such as seeds over a wide area during times of abundance for retrieval when food is in short supply, weeks or even months later. Food caching behavior is amenable for laboratory study and experiments have demonstrated that scrub jays (among other species) fulfill the behavioral criteria for episodic memory: encoding, storing, and retrieving information about the content, location, and date of caching activities. Laboratory experiments manipulating the decay rates of perishable food items show that scrub-jays learn the decay rates of various items and that this, coupled with information stored during caching episodes, determines the pattern of their behavior during subsequent retrieval opportunities. Thus scrub-jays preferentially retrieve less preferred food items (e.g., seeds over grubs) once the decay period for preferred items has elapsed.

Food caching requires the coupling of existing semantic information (knowledge of decay rates) with episodic content (what, where, when information) from unique caching events. Stronger support for the claim that event memories encode *what* information comes from studies showing that food-caching birds are able to avoid returning to the location of previously
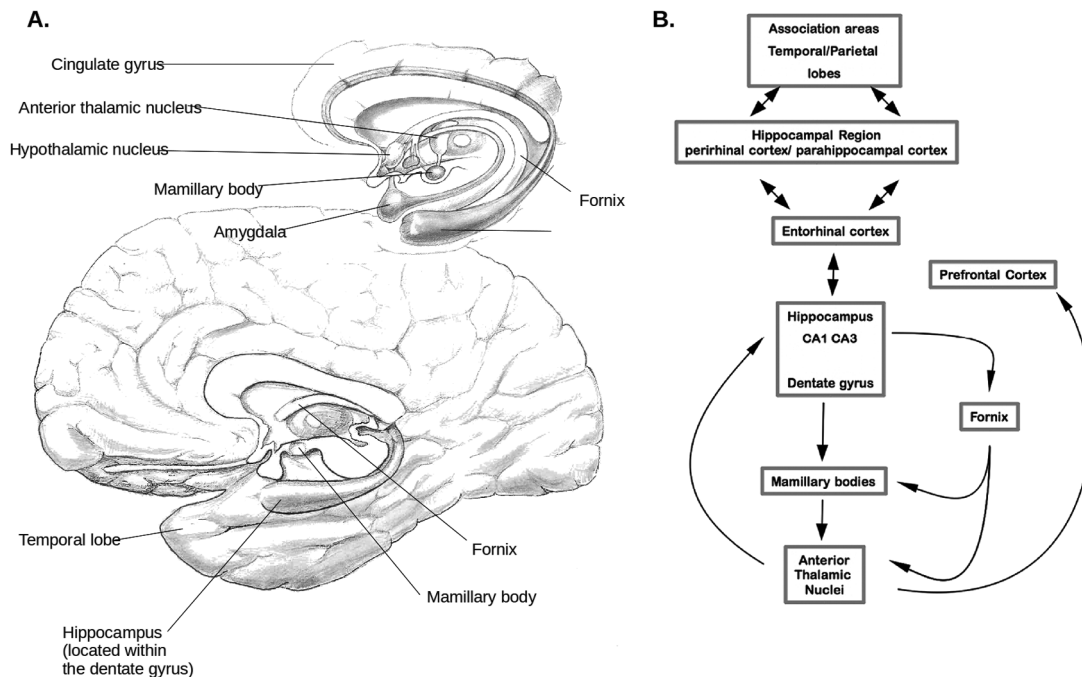
Figure 2. (a) Components of the extended hippocampal–diencephalic system implementing declarative memory. (b) Schematic of the flow of information in recollection memory

retrieved items, demonstrating that they can distinguish memory for the action of placing an item at a particular site from that of retrieving it. These data indicate that food-caching birds can remember storing individual food items and thus fulfill the behavioral criteria for episodic memory, that is, recall of the spatiotemporal organization of events and the temporal order of experiences.

An additional strand of evidence for episodic memory in non-human animals comes from comparative anatomy. The cortical structures implicated in episodic memory are present in all vertebrate brains. In particular, the hippocampus is similarly formed in terms of both structure and connectivity in humans, primates, and (though very much smaller) in rodents (Morris, 2001). In addition the characteristics of activity-dependent synaptic plasticity (a candidate mechanism for memory trace encoding) suggest that the structures in which this mechanism is observed have physiological attributes that are well suited to the formation of an episodic system (Morris, 2001). Animal research has shed light on the neural structures supporting encoding of the what, when, and where components of an event. Recent advances in understanding the functioning of these structures and their interactions have led to models of spatial memory and the proposal that one component process, structural learning, might provide a precursor to episodic memory, through formation of "snapshot" memories (Aggleton & Pearce, 2001).

2.2 Neuroanatomy of memory systems

Procedural memory, that is, skilled behaviors and acquired habits, depend on interactions between cortical areas and the striatum (basal ganglia) with the cerebellum supporting memory for motor learning. During the process of learning new motor acts the motor and somatosensory cortices are active, becoming less so once the new skill is acquired. Connections between the basal ganglia and the cerebellum provide the basic substrate of memory for motor skills with the cerebellum as the long-term storage site for models of sensorimotor mappings (Wolpert, Ghahramani, & Flanagan, 2001). The basal ganglia have primarily been thought to be implicated in motor memory and stimulus response conditioning. However, there is also evidence that this region plays a broader, more cognitive role via interactions with medial temporal lobe regions implicated in, for example, declarative memory function (Packard & Knowlton, 2002).

The extended hippocampal–diencephalic system[5] (see Figure 2.) provides the central substrate for declarative memory (Aggleton & Brown, 1999, 2006; Wolff, Loukavenko, Will, & Dalrymple-Alford, 2008). The hippocampus links to mammillary bodies and thence to the anterior thalamic nuclei via the fornix. Damage to this pathway underlies the core deficits in anterograde amnesia, which lends support for the proposal that medial–temporal and medial–diencephalic interactions are required for efficient encoding of new

event information and thus for normal episodic memory function (Wolff et al., 2008). Aggleton and Brown (1999) propose that the hippocampal–diencephalic system is responsible for putting event information into its temporal and spatial context (the episode) thus aiding retrieval by boosting the discriminability of individual events. Damage to any of the structures in this system produces similar recall deficits.

The structure of the hippocampus[6] is highly specialized with neurons in CA3 each receiving around 10,000 feedback connections from other CA3 neurons (the longitudinal-association pathway) but quite limited numbers of mossy fiber inputs from the dentate gyrus. This very particular architecture raises questions about the processes it implements and how these map to memory function. Rolls and Treves (1998) propose CA3 in the hippocampus as the center of an autoassociative system in which mossy fiber projections may provide a means for gating memory encoding. Activity-dependent synaptic plasticity is a well-studied property of the hippocampus and plasticity mechanisms such as long-term potentiation (LTP) and depression (LTD) are highly plausible candidates for the low-level mechanisms constituting memory. LTP has the characteristics of specificity, associativity and persistence and is thus a good model for the synaptic changes which may underlie the formation of a memory trace (Morris, 2001). In addition, diffuse connectivity between the hippocampus and cortex allows the association of inputs from different cortical regions. Rolls and Treves (1998) describe the hippocampus and neocortex act as "complementary memory systems," with the hippocampus acting as a buffer for temporary storage of unstructured information arising from many different areas of cortex, while the neocortex adapts more slowly on the basis of accumulated information. This view of fast encoding and temporary storage by the hippocampus complementing long-term memory based in posterior neocortex (Mayes & Roberts, 2001) is consistent with evidence suggesting that hippocampal structures are crucial for new memory formation, but less so for long-term storage. However, the hippocampus has also been shown to play a critical role in both encoding and retrieval of spatial and episodic memories and so it remains unclear if it is encoding, consolidation, or storage that is disrupted by hippocampal lesions (Mayes & Roberts, 2001).

2.3 Dynamic networks in memory

Phase synchronization of neural assemblies in different brain regions enables transient interactions between regions, combining multiple processes in large-scale dynamic brain networks; such networks play an important role in cognitive processes including memory. Transient linking through phase synchronized neural oscillations has an amplifying effect on relevant processing and acts to filter irrelevant processing (i.e.,

asynchronous firing). Neural oscillations have been linked to various cognitive functions though there is no direct mapping between oscillatory rhythm and cognitive function. Episodic memory is a complex cognitive function requiring coordinated interaction between brain regions and there is evidence to suggest that gamma (25–100 Hz) and theta (4–8 Hz) oscillations, in particular, play functional roles in episodic memory.

Nyhus and Curran (2010) proposed a model in which gamma and theta oscillations underlie the functioning of a cortical–hippocampal network for encoding and retrieval of episodic memories. Gamma phase synchronization supports binding and encoding of perceptual information from diverse brain regions while theta oscillations provide temporal order to individual episodes and top-down control from frontal cortex to the hippocampus for selective encoding and retrieval of episodic content.

Auto-associative network models of episodic memory are based on the principle that perceptual input is integrated by the rhinal cortices and then encoded by the hippocampus (specifically in CA3). Encoded information is stored as discrete attractor states of the highly recurrent CA3 network. The network can be driven back to a given attractor state by input of any portion of the original input pattern, thus allowing retrieval of the memory and transfer to long-term store in the cortices (Rolls & Treves, 1998).

Fuster (1997) proposed that memory function be understood as a distributed property of the same cortical systems that underpin perception and action. This view is founded on the principle that memory consists of the modulation of synaptic contacts within distributed networks of interconnected cortical cells. Memories are formed by selective facilitation and elimination of synaptic links between neuronal aggregates that activate in response to discrete features of their environment (both within and without the body of the agent) and thus are inherently associative: the information they contain consisting of neuronal relationships. Network memories are constructed via Hebbian synchronous convergence wherein co-activation leads to association via summation of temporally coincident inputs and LTP for reshaping existing networks. These associative processes produce interconnected, functional units of memory (i.e., Hebbian cell assemblies). Thus personal memories can be understood to be implemented across wide areas of the cortex of association. Central to the network memory approach (Fuster, 1997, 2000) is the notion that memory and perception share a common substrate; the same neurons, connections, and networks that provide the functions of perception are also the substrate of memory.

Perception thus consists of object classification achieved through the activation of associative networks that represent those objects in memory. New perceptions are thus expansions of old ones achieved by

adding new associations to preexisting networks. Any cell or assembly can be part of many networks and memory is thus distributed, with a high degree of redundancy and grounded in perceptual and motor experience. Memory networks develop bottom-up via processes of self-organization from sensorimotor areas to the cortex of association. Memory thus has a loosely hierarchical structure with simple sensory and motor memory networks at the bottom, increasing in complexity to encompass multi-sensory and complex motor memory networks, and thence to the cortex of association and networks of declarative and procedural memory.

Fuster extended this principle to propose that the strict dichotomization of long- and short-term memory is unnecessary, with working memory realized through temporary, ad hoc activation of perceptual and motor memory networks. The formation and reactivation of memory depends on association and activation of distributed networks located in the posterior cortex; if the content of memory is associated with action then activation propagates, spreading forward to prefrontal cortex. The central principle of the network memory approach is that memory is a property of the systems it serves and is inseparable from them. Memory thus consists in networks of units, the first level of which is the cortical cell group corresponding to sensorimotor features of the environment. Moving up the scale from the level of single networks there is significant overlap and networks take on a nested, hierarchical structure. Retrieval, in this model, equates with activation while both long- and short-term memory functions activate extensive areas of cortex and share the same substrate.

2.4 Summary

This section has introduced the main epistemological divisions in human memory research and given a brief account of some central themes in the study of long-term declarative memory. The empirical evidence reviewed in this section will allow us to identify a set of attributes that characterize biological memory and some open research questions related to those attributes that can inform models for the construction of synthetic memory systems. The localization of memory processing and storage remains a research challenge that serves to demonstrate the distributed nature of memory systems and of their contents. By extension, we can understand the functions of memory as being integrated with cognitive processing, rather than forming a discrete storage system implemented by the interaction of multiple subsystems. We can also understand memory to be constructed via associative processes operating on multimodal inputs with no inherent difference in the neural systems implementing different types of semantic content.

3 Memory in synthetic systems

Memory clearly forms an important part of synthetic cognitive system implementation, as a means of linking prior experience with ongoing and future behavior. It is therefore clear that many synthetic adaptive systems display the functions of memory, even without explicit implementation of a memory system. This section will review and reevaluate of the use of memory in synthetic computational structures including those where memory is not identified as an explicit component. The aim of this review is not to provide an exhaustive account of all implementations in synthetic systems, but rather to identify the main categories thereof, and to provide typical examples of each. The emphasis is thus on principles of memory function rather than those of architecture (for examples of architecture design see Goertzel, Lian, Arel, Garis, & Chen, 2010; Langley, Laird, & Rogers, 2009; Sun, 2004).

The drawing of a functional and structural distinction between memory and cognition has been primarily derived from psychology, and parallels the more general trend for modular approaches to cognitive function. This perspective has been particularly influential in the area of cognitive modeling and the implementation of synthetic cognitive systems, because it was from this domain that the inspiration for dissociable cognitive processing and memory originated; an example is the cognition as computation metaphor (Bickhard, 2009; Miller, 2003). This division has subsequently remained in computational implementations, and has been argued to be necessary (Sun, 2004).

Furthermore, behavioral evidence has indicated a division of biological memory into a number of

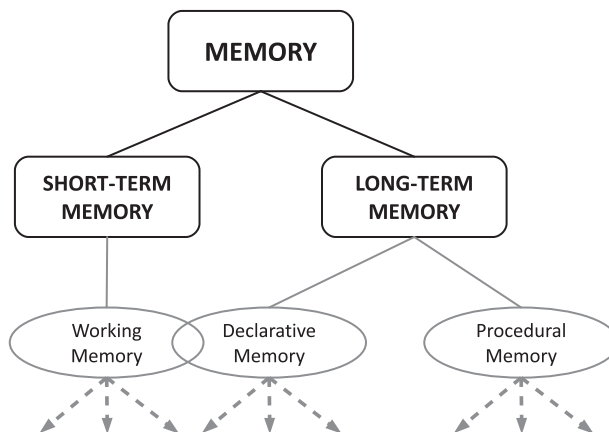**Figure 3.** A coarse taxonomy of functional memory types in synthetic systems, where each may be divided into further memory subsystems. The overlap between working memory and declarative memory indicates that while both systems are part of qualitatively different types of memory (short- and long-term memory respectively), they nominally operate on similar types of information (see Sections 3.1 and 3.2)

functionally distinct components, in particular a separation between short- and long-term memory, and then between different types of long-term memory (Figure 3). This perspective has similarly proven particularly influential in synthetic cognitive systems research. Such systems may thus incorporate a number of discrete, encapsulated memory functions. The previous section has shown that a number of such functionally distinct memory systems may be identified. While there are overlaps between these systems (and further subdivisions may be made within them), this coarse categorization of memory systems forms a suitable basis for examining the (explicit or implicit) use of memory in computational implementations, given the varying degrees and types of biological inspiration upon which they are based. In conducting this examination, a number of implementation characteristics are exposed that reflect implicit assumptions regarding the function of memory systems in biological agents, and which may thus provide the basis of a necessary reevaluation.

3.1 Short-term memory

STM holds information for a finite length of time, commensurate with that of the context to which this information is related. Functionally, the concept of temporary membership of the STM structure is thus apparent, marking the central distinction from LTM. Forming a further functional specification of STM is WM, which is held to be a functionally, and also often structurally, distinct memory system that temporarily stores information in the service of an ongoing task (Cowan, 1999). This perspective on STM store has been influential in the empirical study of human cognition (Section 2), and has subsequently provided fundamental principles of operation, or retrospective justification, for the design of synthetic cognitive architectures.

As such, implementations of STM may be divided into two categories. In the first, computational constructs are implemented with explicit functional characteristics of psychological working memory, typically as part of wider cognitive architectures (for an overview of such architectures see Langley et al., 2009). Here, temporary storage for task-relevant information becomes a workspace-like structure, reminiscent of machine learning blackboard architectures (e.g., Hayes-Roth, 1985). In the second category, it is the property of sustained transient activity that is prevalent, where short-term memory may be described as activated long-term memory, (e.g., Oberauer, 2002). While these implementations have the functional characteristics of STM, they are not generally labeled as such, because this emergent characteristic is not structurally distinct from the wider system. The following paragraphs describe examples

for each of these categories, and provide an outline of further distinctions that may be made.

Unitary and centralized workspaces have been used to implement WM in general cognitive architectures. For example, in the Soar architecture, a central workspace, termed *dynamic memory*, is used as a means of gathering information relevant to a current goal (Newell, 1990; Young & Lewis, 1999). This workspace brings together items drawn from memory (in a procedural-symbolic form, akin to declarative memory, see below) with computation, thereby fulfilling the requirements of a working memory system as described above.

In another formulation of the WM concept, a central workspace is used to provide globally accessible information, rather than as a central computation space. Reminiscent of the global workspace theory of consciousness (Baars, 1988), these architectures allow competition between independent low-level processes to determine the information that should be broadcast across the system; thence, in an adaptive process, the workspace allows this broadcast to occur (Baars & Franklin, 2007, 2009; Shanahan, 2005, 2008), thus acting as a working memory. A similar unitary workspace-based approach is taken in other cognitive architectures, such as the use of blackboard implementations (e.g., Hayes-Roth, 1985; Van der Velde & de Kamps, 2006), to coordinate information across functional components.

Bearing functional but not structural similarity to the the central workspace formulation, the ACT-R framework (Anderson, 1993) explicitly represents the goal of the system, which is used as a filter for the selection of symbolic information and the biasing of sub-symbolic information activation levels (Lovett, Reder, & Lebiere, 1999). The goal representation provides the context for this activation and so, in combination, they encompass the functionality of working memory.

Alternatively, short-term and working memory are not integrated into a single system but, rather, functionality is divided among different modules. For example, in the control architecture for the ISAC (Intelligent Soft Arm Control) humanoid robot, short-term memory allows the passive storage of sensory information with an additional spatial component, whereas working memory contains task-relevant information with no explicit limit on the duration of membership (Kawamura et al., 2004; Peters, Hambuchen, Kawamura, & Wilkes, 2001). Furthermore, a number of architectures implement a multiple-space version of working memory, where each space handles information from a specific functional aspect of the architecture. Working memory is thus in a more distributed state, but still functionally and structurally distinct from the wider cognitive architecture. For example, in the EPIC (Executive Process/Interactive Control) system, various separable working memory components

handle information from different modalities and for different purposes (Kieras & Meyer, 1997; Kieras, Meyer, Mueller, & Seymour, 1999). Similarly, the CAS (CoSy Architecture Schema) groups related processing components. Each group of this kind (a subarchitecture) has a working memory space (Hawes et al., 2007; Wyatt & Hawes, 2008). These working memory spaces allow the subarchitectures to interact and share information (Hawes & Wyatt, 2010). The Polyscheme architecture uses a variation on this approach in that individual functionally specific modules interact using a common ontology (Cassimatis, Trafton, Bugajska, & Schultz, 2004); WM here would consist in the information shared across modules at any given time.

The second category of STM implementation (as transient activation levels in network structures) does not explicitly comply with the notion of binary membership. This is partially evident in the ACT-R framework, where activation levels determine the influence of information on the production of behavior (Lovett et al., 1999). A wide variety of computational systems fulfill this functionality, where the information encoded in the network is in both symbolic and subsymbolic form. For example, in semantic networks (Sowa, 1991, 1992), activation can be set to spread through an existing associative network structure for the retrieval of relevant information, (e.g., Anderson, 1983; Crestani, 1997). In nonsymbolic networks with similar characteristics, such activation flow has been used for a wide variety of purposes, such as the modeling of language learning (Boucher & Dienes, 2003) and developmental systems (Munakata & McClelland, 2003), or the determination of actions for a mobile robot (e.g., Baxter & Browne, 2010; Maes, 1989). Furthermore, the transient activation dynamics seen in recurrent artificial neural networks may also be described as a short-term/working memory system, specifically in terms of the reactivation of those parts of the network structure that are relevant to its input, and the manner in which this influences ongoing behavior (e.g., Wood & Di Paolo, 2007; for a comprehensive introduction to recurrent networks see Haykin, 1999). This reinterpretation of transient activation dynamics as a short-term/working memory is further explored in the following sections in the context of declarative and procedural memory.

As seen in the paragraphs above, the function of a short-term/working memory system can be found in a wide range of implementations. A distinction between two types of implementation may be made: (1) those directly inspired by psychological concepts of working memory, which generally make use of a structurally discrete component; and (2) those where the functional properties of short-term/working memory are present, but where no explicit provision for such a system is made. While these two types are not inherently mutually exclusive in terms of function, the distinction underlines a difference in design inspiration (biological

or engineering), with clear consequences for the manner of implementation.

3.2 Declarative memory

Declarative memory, as a form of TLM, can be defined in terms of the type of information it encodes. In synthetic systems declarative content is most often symbolic (or based upon symbolic) information that is directly accessible by the system. The distinction between declarative memory and STM, as defined in this article, hinges on the functional role of information in the two systems and their temporal characteristics, rather than the type of information involved. Where memory systems are explicitly included as part of wider computational cognitive architectures, these are typically of the declarative memory type. The structure of the declarative memory system is thus of paramount importance for the efficiency of storage and retrieval processes.

Declarative memory may be further subdivided into episodic and semantic components, a functional distinction supported by considerable empirical evidence (see Section 2). This division is consequently apparent in synthetic implementations: semantic memory generally refers to symbolic information regarding event context-free objects and concepts, and episodic memory explicitly encodes contextual, spatial, and temporal information (i.e., what, where, and when).

The influential physical symbol system hypothesis (Newell & Simon, 1976) states that a physical symbol system has the necessary and sufficient means for general intelligent action. As such, the use of symbolic representation for semantic information is prevalent. The differences in such systems occur in the manner of organization and use of symbols, particularly in cognitive models and architectures. Given that symbols are typically considered to be discrete computational objects (Dorffner, Prem, & Trost, 1993) they can be manipulated independently from one another. Therefore, database-like structures are frequently used because they provide a straightforward way of iterating through the available symbols in the search for matches for retrieval and/or other manipulation (e.g., Kawamura & Gordon, 2006). Alternatively, taxonomic structures are used in attempting to increase the efficiency of storage and retrieval, with variations such as the use of on-line chunking in Soar (Young & Lewis, 1999).

A related set of functions concerns the learning of spatial place maps by mobile robots. In these approaches, a node in a network represents a place, and connecting vertices represent potential routes between nodes (e.g., Kuipers, 2000; Kuipers & Byun, 1991; Mallot & Basten, 2008). Since places in these networks are context independent (i.e., they are applicable regardless of the context in which the node is visited),

they may be regarded as a type of semantic information.

A related approach is that of semantic networks, where symbols (as nodes) are connected in graph form to related symbols (Sowa, 1991, 1992). In this type of structure, the concept of spreading activation may be used to combine symbols in a task-relevant manner, (e.g., Collins & Loftus, 1975; Rumelhart & Norman, 1973). A number of examples of this functionality exist in contemporary computational architectures, including the declarative memory systems of ACT-R (Anderson et al., 2004), and CLARION (Sun, Merrill, & Peterson, 2001).

As these types of semantic memory systems deal primarily with symbolic information, a brief mention should be made of how these symbols relate to the real world (Harnad, 1990). Symbol grounding and tethering are central issues (Dorffner & Prem, 1993; Sloman & Chappell, 2005), which have been, at least in part, addressed with the use of a subsymbolic structure to interfacing between a symbol system and the world (Gardenfors, 2004; Kelley, 2003).

Whilst the implementations discussed above have used semantic information in symbolic form, there have been attempts to use subsymbolic representations which retain some of the properties of symbolic systems For example, Siegelmann (2008) used a reconsolidation attractor network in which the attractor states can be defined and imposed on the network. In this way it can be used as a dynamic symbol-matching system, where a symbol corresponds to a defined attractor. Similarly, in neuroscience-inspired cognitive architectures, semantic information has been included in the form of concepts in a subsymbolic semantic network (Taylor, Hartley, Taylor, Panchev, & Kasderidis, 2009), or implicitly as a consequence of perceptual categorization (McKinstry, Seth, Edelman, & Krichmar, 2008).

In contrast to semantic memory, episodic memory may be characterized as explicitly containing spatial and temporal object information in a single, experience-centered context. Whilst the role of episodic memory has been established in both human (Tulving, 2001) and animal cognition (Clayton & Dickinson, 1998; Clayton & Salwiczek, 2007), this concept has only relatively recently been considered in synthetic architectures, with the acknowledgment that in long-term interaction, the use of previously experienced episodes by an agent is of clear benefit (e.g., Brom & Lukavsky, 2010; Deutsch, Gruber, & Lang, 2008).

Two general approaches may be distinguished in the implementation of episodic memory, although the use of symbolic representation is common to both. In the first approach, episodic memory is implemented in a discrete computational structure, in combination with other functionally distinct memory systems. For example, in the Soar architecture, the purpose of the episodic memory system is to periodically store a snapshot of the contents of the central workspace module; these episodes may be subsequently recalled on a cue-matching basis to influence processing towards the current goal (Nuxoll, 2007; Nuxoll & Laird, 2007). Similarly, in the control architecture for the ISAC humanoid robot (Kawamura et al., 2004), the episodic memory system also contains snapshots of the working memory system, but supplements this with information from an "emotion module" (Dodd & Gutierrez, 2005). Multiple episodes relevant for a current task may be recalled to working memory, but episodes are subject to a decay function, which is biased by the "emotional valence" of the episode. In these examples, and others, methods from the field of machine learning, specifically reinforcement learning (Sutton & Barto, 1998), are used to determine which episodes should be stored, and when they should be retrieved (e.g., Gorski & Laird, 2010).

In the second approach, episodic memory is the primary aspect of the cognitive architecture. The focus of such implementations is typically the domain of human–robot (or other artificial system) interaction due to its potential use in the generation of more naturalistic interaction (e.g., Brom & Lukavsky, 2009; Ho & Watson, 2006). In these approaches, the use of intelligent virtual agents as a means of implementing and testing architectures for episodic memory is common (e.g., Brom & Lukavsky, 2008, 2010; Ho et al., 2009). In terms of the structure of episodic memory, a symbolic representation of information—where hierarchical structures form the majority of implementations (e.g., Kadlec & Brom, 2010; Vargas et al., 2010)—is typically used, with list-like structures also in use, for example, the use of a list of event records as the basis of an autobiographical memory system (Ho, Dautenhahn, & Nehaniv, 2008).

Finally, the role of forgetting[7] and its various implementations in synthetic systems should be discussed. Typically considered in the natural sciences to be a combination of decay and, more importantly, (proactive and retroactive) interference (Wixted, 2004), forgetting is used in synthetic systems as a means to bound the amount of information held in memory for performance reasons, space constraints (e.g., Nuxoll, Tecuci, Ho, & Wang, 2010), or for ethical reasons (Baxter, Wood, Belpaeme, & Nalin, 2011; Vargas, Ho, Lim, Enz, & Aylett, 2009). Due to the symbolic nature of information in these systems, forgetting algorithms are typically based on the principle of decay (temporal) and relevance (some metric of utility; e.g., Nuxoll et al., 2010), rather than interference, because the encapsulation property of symbols makes a scalar comparison (i.e., determination of the degree of overlap) difficult. Even in subsymbolic networks implementations where the concept of forgetting is raised, the mechanism by which this is achieved is typically using a decay-related measure (e.g., Ishikawa, 1996; Sikstrom, 2002).

From the review above, it is apparent that synthetic declarative memory systems are generally explicitly defined structures in their respective computational implementations. Furthermore, the use of symbolic representation of information entails certain types of storage structure and methods for processing information. Where sub- or non-symbolic declarative memory is implemented, it is typically implicit in a wider cognitive architecture rather than as a discrete component. A functional feature common to most implementations of declarative memory is that it has a close relationship with the short-term/working memory system: information is stored in the former for recall to, and processing in, the latter.

3.3 Procedural memory

Procedural memory concerns the storage of non-consciously accessible information. In terms of synthetic implementations, this definition thus encompasses a wide variety of work, including general artificial neural network learning approaches. As with the interpretation of transient activity across distributed networks as STM (see above), the functionality described by procedural memory may be seen in a wide range of computational implementations. A number of types of procedural memory may thus be distinguished. Firstly, rule-based systems, where explicit computational constructs encode a process, operation, or temporal experience. Secondly, the online learning of artificial neural networks based on experience in neuroscience-inspired architectures. Thirdly, developmental robotics models, particularly systems for the acquisition of sensorimotor coordination, are relevant given the explicit necessity of experience.

In certain rule-based systems, explicit computational constructs (rules) are used to encode condition–action pairs, as for example in learning classifier systems (e.g., Bull, 2004; Bull & Kovacs, 2005), or classical expert systems, (e.g., Duda & Shortliffe, 1983). These rules may be generated at system run-time, and also encode statistics related to performance, and thus may be described as being a type of procedural memory. Similar structures exist in general cognitive architectures, such as the Procedural Memory modules of Soar (Derbinsky & Laird, 2010), ATC-R (Anderson et al., 2004), and CLARION (Sun, 2007). In these systems condition–action rules typically involve symbolic manipulations. By contrast, low-level rules may be used to explicitly encode associations as computational units that are subsequently used to define ongoing behavior (Baxter, 2010; Lee, Meng, & Chao, 2007).

Artificial neural networks are typically trained to learn some input–output mapping. Taking, for example, the multi-layered perceptron, the aim of the training process would be to adjust the weights of links in the network such that the desired output is produced for every given input. The multi-layer perceptron thus implements what may be described as implicit memory: the cumulative effect of the training iterations leads to a system capable of applying learned information to future behavior. Similarly, in standard self-organizing maps (Kohonen, 2001), training is achieved through self-organization based on structure in the training data, hence enabling classification based on learned behavior. Whilst these examples are somewhat simplistic, the analogy extends both to other types of artificial neural network and to systems in which the concept of learning is more readily applied than memory.

In accord with the concept of memory being acquired through experience, some artificial neural network implementations enable adaptation without an explicit (and/or separate) training phase. In these implementations, the role of implicit procedural memory is more clearly visible. Wyss, Konig, and Verschure (2006) used a hierarchical neural network with five layers of identical units; input was provided from a camera mounted on a mobile robot engaged in random movements in an enclosed environment. Employing just two computational principles (temporal stability and local memory), the layers of the network self-organized into distinct functional roles, with the top layer (i.e., that furthest removed from incident sensory data) exhibiting behavior akin to place cells, where it is experience that drives this emergence. Because place cells have been centrally implicated in spatial navigation tasks (Burgess, Donnett, Jeffery, & O'Keefe, 1997; Poucet, Lenck-Santini, Paz-Villagran, & Save, 2003), this places procedural memory in an important role. In a similar manner, various implementations of self-organizing maps and related systems have been created that adapt on-line to input information, either by structural reconfiguration (Fritzke, 1994, 1995; Lang, 2007) or by the creation and use of associations across multiple maps (Johnsson & Balkenius, 2009; Luciw & Weng, 2010; Morse, de Greeff, Belpaeme, & Cangelosi, 2010).

Furthermore, mechanisms implicated in procedural memory such as priming and conditioning are also seen in synthetic systems. Priming in this context refers to activation triggered as a result of associations to currently experienced inputs, based on the previous experience of a relationship between the two, and conditioning is the process whereby a repeatedly experienced condition leads to a stereotyped behavior. In Morse and Aktius (2009), both priming and conditioning are demonstrated in a robotic system controlled using a psychologically plausible associative system, using an echo state network (Jaeger, 2002) as a non-linear feature detector. Similarly, the complementary learning systems framework has been used to model memory-based recognition and fear conditioning in an architecture based on processes of slow integration of representations (as a model of neocortical regions) and rapid encoding of arbitrary relationships between

representations (as a model of the hippocampus; O'Reilly, Braver, & Cohen, 1999; O'Reilly & Norman, 2002).

Models of biological nervous systems provide examples of integrated systems that incorporate a wide range of neurophysiological and neuropsychological evidence. For example, the brain-based device series of mobile robots (e.g., Krichmar & Edelman, 2002; Krichmar, Nitz, Gally, & Edelman, 2005; McKinstry et al., 2008) implements large-scale simulations of multiple regions of the mammalian nervous system to demonstrate not only how the functionality implicit to procedural memory described above is instantiated, but also how this can form the basis of more complex behaviors. Other comparable implementations (integrating multiple strands of empirical evidence) similarly aim not only to model biological mechanisms, but also to test means of the generating adaptive behavior in synthetic agents; examples are the GNOSYS architecture (Taylor et al., 2009), a model of the basal ganglia applied to a mobile robot (Prescott, Montes Gonzalez, Gurney, Humphries, & Redgrave, 2006), and the implementation of memory-enhanced obstacle avoidance based on the dynamic field theory approach (Schöner, Dose, & Engels, 1995). In each of these examples, learning is the central feature, but a memory system is implicit in the fact that prior learned associations are the basis for current and future adaptive behavior.

Returning to the characterization of procedural memory as inherently requiring an account of the history of interaction of an agent with its environment, ontogenetic development becomes relevant. Where the general architectures in the previous paragraph assumed some determined structure of cognitive processing (and hence no specific ontogenetic development), developmental robotics attempts to address this assumption by exploring how complex cognitive structures may emerge through an ontogenetic process (Lungarella, Metta, Pfeifer, & Sandini, 2003; Weng et al., 2001). Developmental robotics employs a wide range of methodologies where the role of memory is not made explicit (Weng, 2007), with some exceptions, such as the storage of interaction histories in Mirza, Nehaniv, Dautenhahn, and te Boekhorst (2007), and the description of memory as distributed and subsymbolic in the dynamic field theory (Schöner et al., 1995; Schutte, Spencer, & Schöner, 2003). In the specific context of the development of sensorimotor coordination, a number of implementations have departed from the neural-system inspired approach, and made use of explicit representations of associations (Baxter & Browne, 2011; Lee et al., 2007; Quinton & Buisson, 2008). This explicit form of association representation allows the memory-based interpretation to become clear: interaction (i.e., experience on the part of the agent) enables associations to be formed, and thence to

influence subsequent behavior, that is, memory is playing an active role in ongoing action selection.

Finally, a brief mention should be made of some of the approaches used in the application of previous learning to ongoing adaptive behavior, that is, how to use memory for a current task. Biologically derived theory has increasingly emphasized that emotion (and more generally valence) plays a central role in the production of behavior (Damasio, 1994; Pessoa, 2008). This functionality has thus also become important in synthetic implementations (Arbib & Fellous, 2004; Canamero, 2005). A variety of methodologies exist for incorporating these principles into the types of neural-inspired systems described above, including the use of artificial endocrine system-like functionality (Neal & Timmis, 2003), analogues of neuromodulation (Baxter & Browne, 2009; Krichmar, 2008; Ziemke & Thieme, 2002), and homeostatic mechanisms (Di Paolo & Iizuka, 2008; Wood & Di Paolo, 2007; Ziemke, 2008). These approaches provide candidate mechanisms for the application of previously acquired information (i.e., memory) to a current behavioral context.

The reinterpretation of existing subsymbolic implementations of neuroscientifically inspired principles in the context of procedural memory reveals that this form of memory is implicit and pervasive in the such systems—to a much greater extent than declarative memory. Furthermore, whilst this review has considered (albeit briefly) a very diverse set of research approaches, procedural memory is typically regarded as being fundamentally a subsymbolic system subject to various biasing and modulatory influences in the service of adaptive behavior.

3.4 Summary

Whilst a wide range of synthetic memory systems are available, a number of broad characteristics are apparent, which to a large degree reflect the underlying assumptions of such implementations. These assumptions have in part been drawn from the psychological literature, and in part from the computational sciences (computer theory, and the consequences of this for the cognitive sciences). Most fundamental of these assumptions is that cognition (as computation) and memory (as storage) are both functionally and structurally dissociable. An extension of this perspective to the functional fractionation of memory itself is also in evidence, most notably in the use of a working memory subsystem which forms the bridge between the computation and storage aspects: this typically takes the form of a discrete workspace structure. Another of these issues regards the nature of information representation/encoding and use, with symbolic structures used for declarative memory systems, and sub- or non-symbolic structures generally used for what may be described as procedural memory systems. In general, whilst

psychological evidence has been used in support of many of the discussed implementations of memory, there is only a limited amount of such evidence that conforms to the more recent neuroscientific perspective that memory systems are inherently distributed and integrated with cognition, as reviewed above (Section 2), with implementations of implicit procedural memory providing the examples closest to this perspective.

4 Rethinking memory: a new approach for synthetic cognitive systems

The first part of this article presented a review of psychological and neuroscientific perspectives, largely based on research into human memory function, intended to highlight some broad principles characterizing biological memory. A significant research effort has gone into the localization of memory and finding the brain areas responsible for particular facets of memory function. This undertaking has been successful in that the neural connectivity underpinning some aspects of mnemonic function have been mapped (for example the hippocampal–diencephalic pathways involved in recognition memory) and there is agreement on the importance of particular brain regions for broad classes of memory function. Yet there are still many open questions about the interactions between the areas, because the evidence (e.g., amnesias following focal lesions) can be contradictory. The specific patterns of impairment associated with particular lesion sites have been shown to be less clearly defined than previously assumed and new experimental paradigms have shown areas of preserved function with lesion sites which have been previously associated with impairment.

Furthermore, the patterns of neural activity associated with memory do not, in themselves, differentiate types of content: the neuronal response associated with memory of an acoustic signal is not inherently different from that for a tactile stimulus or for a lexical item. The "neural code" does not make semantic differentiations between items or modalities. Physiologically, all memories are equal and consist of activation dynamics distributed over networks of neurons. These overlapping networks are configured and reconfigured on an ad hoc basis driven by the flow of inputs to the system and so there is no real sense in which memory can accurately

be considered as a static storage system or that particular patterns of activation can be said to represent, or code, specific items of content.

The structure of memory is fundamentally associative so that rather than coding information explicitly in some symbolic fashion, memory is inherently *relational* with semantic content existing in the form of patterns of activation within associative networks. Biological memory has the additional property of being fundamentally integrated within cognition rather than being an adjunct storage system. Memory primes cognitive function and constrains learning providing a structure for the acquisition of new information. Being inherently amodal, the operations of memory must necessarily be closely integrated with other cognitive modalities.

4.1 Memory in synthetic systems: a reinterpretation

The characteristics of synthetic memory systems described in the second part of this article highlight the differences in inspiration from biology, particularly in relation to the function that the memory system is proposed to serve. This is important because function to a large extent defines mechanism. Given the general acceptance of the encode–store–recall paradigm, the contents of the memory system are defined by the task that it is to be engaged in. In the majority of explicit memory system implementations, this results in the use of a symbolic representation scheme, with discrete symbols as "chunks" of information being manipulated in accordance with the mind-as-computer paradigm. While there are some exceptions to this characterization which do begin to approach certain aspects of low-level biological mechanisms, these systems are typically not embedded into general cognitive architectures, and thus cannot inform these questions of theory directly; rather, they provide examples of how such mechanisms may be implemented.

Four characteristics may thus be distinguished from the review of synthetic systems (see Table 2). Firstly, implementations of memory are either *modular* themselves, or part of an inherently modular architecture. The second and third characteristics are related to this point. Memory is treated as a *passive storage structure*, in which information may be placed, remaining static

Table 2. A summary of the main distinctions between the prevalent view of memory in synthetic systems and the approach proposed here

The prevalent view: memory…	The novel proposal: memory is…
…has a modular structure	…integrated and distributed
…is a passive storage device	…an active cognitive component
…makes use of a global ontology for information transfer	…associative and inherently amodal
…entails externally controlled information flow	…a coordinator of information across modalities (priming)

during the storage period, for recall by a specific function at some point in the future. The third characteristic is that the format of information stored in memory is typically based on a *global ontology*, which ensures that information is in a common format that can be handled by all of the computational (or cognitive) functions. Memory thus contains information that has been abstracted away from low-level sensorimotor data. The fourth characteristic is that with memory as a purely passive storage module, specific computational mechanisms are required to handle the *information flow* to and from memory, with centralized workspaces being the typically used structure. While the workspace paradigm as used (for example) in Soar does provide a level of integration between elements of memory function and more general cognitive processing this still falls short of the notion of memory as the active substrate of cognition that is proposed here.

These four characteristics, whilst arising from attempts to replicate biological memory function, demonstrate fundamental theoretical and mechanistic differences. The biological data used in support of synthetic system design is often decontextualized and relatively selective (typically being drawn from the psychological literature, with less impact of mechanistic neuroscience). As a result, considerable divergence has developed between biological mnemonic function and memory in synthetic systems. As a result of this dichotomy, synthetic systems are "missing out" on essential features of the biological systems that enable adaptive and robust behavioral competencies.

The first challenge that becomes apparent in seeking to bridge the gap between what is known of how biological memory works and attempts to replicate its functions artificially is that of operationalizing psychological concepts. This process involves taking what are often relatively high-level characterizations of memory function and interpreting them in terms of structures that can be constructed. There are obviously many potential pitfalls in such a process. One primary issue is the separation of what memory does (the empirical data on memory function) and what memory is (structures and mechanisms). One approach is to implement the functional characteristics of memory piecemeal, to replicate one or more empirical findings about biological memory. The problems with this method are two-fold. (1) It is not straightforward to hand design a mechanism for a particular function; the history of cybernetics and AI has shown that this approach can often lead to brittle suboptimal results. (2) The empirical data does not necessarily provide an accurate picture of what a cognitive function does; that is, experiments give a snapshot of function under specific, very constrained circumstances.

These problems become more difficult when the function in question is most clearly characterized in psychological terms and the mapping between neural activity (i.e., mechanism) and cognitive function is ill-defined. For some aspects of perception and action the neural dynamics underpinning function are relatively well understood, for example, the neural pathways enabling particular motor acts are relatively well-defined. Similar mappings cannot be made for cognitive functions such as memory. It is known that localized lesions induce specific mnemonic deficits and that particular neural dynamics may be associated with a given component of memory function. However, the data do not yet permit the derivation of prescriptive models for the construction of artificial systems with either the same functional or structural characteristics that are found in biological memory systems.

In the last two decades AI research has moved away from top-down approaches to the implementation of intelligent behavior toward more biologically informed, bottom-up methods such as "computational (or synthetic) neuroethology" (Beer, 1990; Cliff, 1991). This approach has generated insights into behavior generation at many levels, from the abstract principles of dynamic interaction between context, morphology, and nervous system to specific, empirically testable predictions concerning the mechanisms underpinning behaviors such as navigation, locomotion, and learning. Such models may be more or less specific to the characteristics of a particular target species (for a discussion of the value of abstract models in the study of adaptive behavior see Webb, 2009; also Husbands, 2009; Seth, 2009), but there is a common focus on the construction of model behavior generating systems that emphasize embodiment and agent–environment interaction.

Such models provide a method for exploring candidate mechanisms for the production of a particular behavior and thence understanding the nature of the function itself (e.g., Slocum, Downey, & Beer, 2000). However, these methods are less often applied to cognitive function and particularly not to putative higher level functions. It is not necessarily obvious how an embodied, contextually embedded approach can contribute to the understanding of a function such as memory that might be characterized as encapsulated information processing. To date the most common approach to implementing memory systems in silico has been to isolate memory from general cognition and to treat memory store as a passive component in a cognitive architecture. A neuroethological approach, by contrast, situates memory within cognition and thus within the perception–action cycle of adaptive behavior in an embodied agent.

This perspective also broadens the scope of inquiry from attempts to replicate the characteristics of human memory to modeling the neural mechanisms and activation-based dynamics required to achieve the functions of memory as they can be understood across species. This expansion of focus arguably supports an operationalization of memory function that is amenable to modeling and thus affords an incremental

process of model building and empirical hypothesis testing by which new insights can be gained.

Synthetic neuroethological modeling forces a commitment to system design on a *process* rather than a *functional* basis, in the context of the wider system. A modular design focused on individual functions (such as memory) lacks the system-level interactive context required for fully integrated functioning. Furthermore, by incorporating this neuroethological perspective with the biologically inspired principles derived from traditional empirical approaches (reviewed in the first part of this article), a set of properties can be identified that capture the functional significance of memory, but do so within the context of a wider cognitive architecture. Memory should therefore not be defined on the basis of putative cognitive processes (or vice versa); rather, the properties of each should be considered as interdependent. For example, in synthetic cognitive architectures such as Soar (see Sections 3.1 and 3.2 for details), the computational requirements (i.e., cognition as symbol manipulation) determines the structure and use of memory. In this way, memory and cognition certainly have a close relationship, but they are not integrated (i.e., interdependent) because the former is specified in terms of the operational constraints of the latter. These four properties of memory, being fully integrated with cognition, may be used as the basis of a synthetic architecture that conforms more closely to the biological theory, and in so doing, enables more of the desirable properties of biological function to be captured, namely, adaptivity, generality, and robustness.

4.2 Proposed properties and mechanisms of memory

Four properties of memory function may be identified, which are in opposition to the characteristics of current memory system implementations (Table 2).

Firstly, memory is *fundamentally associative*. This implies that there is no requirement for a global ontology for the integration of domain-specific information, and thus that there is no requirement for semantic information (i.e., modality-specific "meaningful" information) to be explicitly transferred via the memory system—the process of association being itself considered inherently amodal. All semantic information remains in the domain-specific processing modules, thus negating the necessity for an explicit mechanism for the creation of amodal symbolic representations.

Secondly, the memory system *does not require the explicit storage of semantic information*. This storage aspect of the memory system is thus not a functional module, but a property of the memory system as embedded in the wider cognitive architecture. Associations in this context may be regarded as multimodal constructs that are not meaningful in themselves, but only by virtue of the correlations in activation that they encode.

Thirdly, the memory system is a *functionally distributed* system (in that associations link information held in multiple modalities); this again negates the necessity for a global ontology into which all modality-specific information needs to be encoded, and also the necessity for a separable storage device. While each association can be treated in the same manner, the modal-specific information being associated differs, such that these associations take on differing functional roles.

On the basis of these implementational properties, the fourth property may be defined, describing a fundamental function of the cognitive system in which the memory system operates: *priming* (be it intra- or cross-modal). Given the proposed lack of a global ontology by which modal-specific information is integrated, the associative structure of memory serves as a coordinator of information across modalities rather than a simple storage device. Further, given that (symbolic) semantic information does not pass through the memory system, cross-modal coordination is proposed to occur through the propagation of activation through the substrate of associative memory. Thus, the use of information in one modality (be it recognition in a sensory modality, or the execution of a motor act) gives rise to activation that may propagate from one modality to another via the associative memory network. The significance of memory as fundamentally associative in structure and mechanism thus becomes clear: associations can only form in terms of a history of activation profiles, which then become the substrate for future cross-modal coordination. Thus previously acquired associations (based on a history of perception–action activity) modulate the control of ongoing and future behavior, that is, memory as an active system fully integrated with behavior generation.

Thus the role played by priming is proposed to be fundamental in the contribution of memory to the cognitive system as an integrated whole. Cross-modal priming involves the spread of activation in one modality to another and has a material effect on the processing in this second modality (with this effect being bidirectional). Note that this coordination of activity across modalities does not require the transfer of modality-specific information, but does assume the presence of the associative network-like structure that forms the substrate for the priming functionality. This proposal is in accordance with a wide range of empirical evidence on the role played by priming in cognitive function, but importantly, the postulated associative substrate provides a mechanistic account of how this function may be fulfilled in synthetic systems.

4.3 Episodic memory: a case study

Given that episodic memory is arguably the most complex (though certainly the least understood) aspect of

Table 3. A comparison of the central characteristics of current synthetic implementations of episodic memory (Section 3.2) with the mechanisms proposed to fulfill this functionality from the novel memory interpretation perspective (Section 4.2)

Episodic memory in current synthetic implementations	The novel proposal: the mechanisms that provide the functionality of episodic memory
Modular implementation of episodic memory with respect to other memory systems, and cognitive processing	What/when/where aspects encoded and served by recurrent neural network activation dynamics (fundamentally associative)
Episodic memory comprised of discrete, passive and symbolic structures	Embedding in the system's associative network structures means that this functionality is inherently integrated in ongoing behavior (memory is functionally and structurally distributed)
Episode structure and the contents thereof are determined by the application context	The functionality of episodic memory is an inherent part of non-task-specific mechanisms (no explicit storage of semantic information)
Specific mechanisms required for encoding, recall, and forgetting	Context-dependent partial reactivation of recurrent network structures enables encoded information to play an active role in ongoing processing (priming); and the subsymbolic nature of this information enables interference-based dynamics to apply (as proposed by biological theory)

mnemonic function (Section 2), and has generally been considered as a uniquely human competence, it forms an ideal test case for the proposed reinterpretation of general long-term memory. This case study attempts to provide the foundation for such an account, that emphasizes integrated, rather than discrete, structures and functions.

It has been shown that the basic components of episodic memory (i.e., the spatiotemporal structuring of event information) are present in the behavior of several non-human species (Section 2.1.3). These behaviors provide a model for the functional, non-phenomenological components of episodic memory and enables the derivation of an operational concept of episodic memory with which to test putative mechanisms. For example, by modeling episodic-like memory in food-caching birds we can generate minimal candidate mechanisms for the functions required of a simple what–when–where system. Such a system should allow the generation of new hypotheses for testing in empirical populations and also an incremental approach to increasing the complexity of the model and the task domain.

This approach to modeling complex adaptive function can contribute directly to the design and construction of artificial cognitive systems by generating and testing ecologically grounded mechanisms. Where a given function is to be integrated into a cognitive architecture and is required to support the production of adaptive behavior, it is a necessity for that function to be modeled in a way that takes account of its embedding in a larger system of interacting parts. It is therefore necessary to account for the functionality of episodic memory using the proposed properties of memory (Section 4.2).

Current implementations of episodic memory in synthetic systems display similar characteristics to those that describe general memory system implementations (Section 3.2), as summarized in Table 3. By starting

with the properties of general long-term memory (described in Section 4.2) rather than the specific functions of episodic memory, it is possible to provide an account of episodic memory that is inherently embedded within a wider system (Table 3). This account is task independent, and may be at least partially achieved through the application of existing implementation methodologies, such as those used to implement procedural-like synthetic memory systems (Section 3.3). While these approaches typically emphasize the learning aspect, a reinterpretation of their function in terms of the properties of memory enable a naturalistic integration of the range human mnemonic competences.

4.4 Summary

The examination of various implementations of memory in synthetic systems has shown that they do not embed memory function into the wider context of cognitive processing. By applying the neuroethological approach of contextualized modeling, a set of principles of operation for memory have been obtained that emphasize memory as a fundamental aspect of cognitive processing, rather than as an incidental process.

5 Conclusion

Memory can be understood in terms of dynamics grounded in the ongoing activities and previous history of an agent and thus has an inherent developmental trajectory (Baxter & Browne, 2010; Wood, Baxter, & Belpaeme, 2010). This essentially constructivist and enactive view of memory is a rich area for future research and, by extension, requires a significant reinterpretation of the use of memory in synthetic cognitive systems, if the wide range of adaptive and flexible features of biological cognition are to be captured.

Memory systems as implemented in current synthetic cognitive systems only reflect the biological properties of memory in a very superficial manner: the conceptual gap between the two is actually based on fundamental inconsistencies. The application of the neuroethological methodology to this problem provides new tools to enable a set of principles and mechanisms to be formed that begin to reconcile this dichotomy. These principles are consistent with the biological theory, but also propose a set of mechanisms that may be directly applied to computational cognitive architectures.

A definition of memory in the context of adaptive cognitive systems can now be ventured. Functionally, memory is an active process that serves current and future adaptive behavior, based on previously acquired information. At an implementation level, the processes of memory are amodal and operate on multi-modal information, inherently associative (rather than semantic), fundamentally distributed (rather than centralized) and dynamic (rather than static). This new characterization of memory for synthetic cognitive systems has fundamental implications for the design and implementation of these systems that have only just begun to be explored.

Notes

1 Though not universally so, see Section 3.1.
2 Experimental manipulations which selectively operate on one or other of two mechanisms and thus demonstrate their functional independence.
3 Anterograde amnesia is characterized by a deficit in mnemonic function for information acquired post injury.
4 Hypoxic-ischemic encephalopathy (HIE).
5 Including the fornix, mamillary bodies, and the anterior thalamic nuclei.
6 Fields 1–4 of the Cornu Ammonis, the dentate gyrus and the subiculum.
7 Forgetting is mentioned here rather than in the section on procedural memory, because it is to declarative memory that theories of forgetting are usually applied, particularly because of recourse to the concept of explicit recall, which does not apply to procedural memory. This is not to say that forgetting in the context of procedural memory does not occur, merely that the relevant theories are more developed for declarative memory.

Funding

References

Aggleton, J., & Brown, M. (1999). Episodic memory, amnesia, and the hippocampal–anterior thalamic axis. *Behavioural and Brain Sciences*, *22*, 425–489.

Aggleton, J., & Brown, M. (2006). Interleaving brain systems for episodic and recognition memory. *Trends in Cognitive Sciences*, *10*, 454–463.

Aggleton, J., & Pearce, J. (2001). Neural systems underlying episodic memory: insights from animal research. *Behavioural and Brain Sciences*, *356*, 1467–1482.

Allan, K., Wilding, E., & Rugg, M. (1998). Electrophysiological evidence for dissociable processes contributing to recollection. *Acta Psychol.*, *98*, 231–252.

Anderson, J. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior*, *22*.

Anderson, J. (1993). *Rules of the mind.* Lawrence Erlbaum Associates, Inc.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, *111*, 1036–1060.

Arbib, M. A., & Fellous, J.-M. (2004). Emotions: from brain to robot. *Trends in Cognitive Sciences*, *8*, 554–561.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: a proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation* (pp. 89–195). New York: Academic Press.

Baars, B. (1988). *A cognitive theory of consciousness.* Cambridge University Press.

Baars, B., & Franklin, S. (2007). An architectural model of conscious and unconscious brain functions: Global workspace theory and IDA. *Neural Networks*, *20*, 955–961.

Baars, B., & Franklin, S. (2009). Consciousness is computational: the LIDA model of global workspace theory. *International Journal of Machine Consciousness*, *1*, 23–32.

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, *4*, 417–423.

Baddeley, A., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 8, pp. 47–89). New York: Academic Press.

Baddeley, A. D. (2001). The concept of episodic memory. In A. Baddeley, J. Aggleton & M. Conway (Eds.), *Episodic memory: new directions in research* (pp. 1–10). Oxford: Oxford University Press.

Bar, M. (2007). The proactive brain: using analogies and associations to generate predictions. *Trends in Cognitive Sciences*, *11*, 280–289.

Baxter, P. (2010). *Foundations of a constructivist memory-based approach to cognitive robotics.* Unpublished doctoral dissertation, University of Reading.

Baxter, P., & Browne, W. (2009). The role of internal value systems for a memory-based robotic architecture. In L. Cañamero, P.-Y. Oudeyer & C. Balkenius (Eds.), *Proceedings of the Ninth International Conference on Epigenetic Robotics* (Vol. 146, pp. 195–196). Lund: LUCS.

Baxter, P., & Browne, W. (2010). Memory as the substrate of cognition: a developmental cognitive robotics perspective. In B. Johansson, E. Sahin & C. Balkenius (Eds.), *Proceedings of the Tenth International Conference on Epigenetic Robotics* (Vol. 147, pp. 19–26). Lund: LUCS.

Baxter, P., & Browne, W. (2011). Memory-based cognitive framework: a low-level association approach to cognitive architectures. In G. Kampis, I. Karsai & E. Szathmáry (Eds.), *European Conference on Artificial Life (ECAL 09)*, LNCS (Vol. 5777, pp. 402–409). Berlin: Springer-Verlag.

Baxter, P., Wood, R., Belpaeme, T., & Nalin, M. (2011). On memory systems for companion robots: implementation methodologies and legal implications. In W. Ho, M. Lim

& C. Brom (Eds.), *Proceedings of the 2nd AISB Symposium on Human Memory for Artificial Agents* (pp. 30–34).

Beer, R. (1990). *Intelligence as adaptive behaviour*. New York: Academic Press.

Bickhard, M. H. (2009). The biological foundations of cognitive science. *New Ideas in Psychology*, *27*, 75–84.

Boucher, L., & Dienes, Z. (2003). Two ways of learning associations. *Cognitive Science*, *27*, 807–842.

Brom, C., & Lukavsky, J. (2008). Episodic memory for human-like agents and human-like agents for episodic memory. In *AAAI Fall Symposium 2008: Biologically Inspired Cognitive Architectures* (pp. 42–47) Menlo Park, CA: AAAI Press.

Brom, C., & Lukavsky, J. (2009). Towards virtual characters with a full episodic memory II: The episodic memory strikes back. In K. Decker, J. S. Sichman, C. Sierra & C., Castelfranchi, (Eds.), *Proceedings of the 8th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2009)* (pp. 1–9), Richland, SC: AAMAS.

Brom, C., & Lukavsky, J. (2010). Episodic memory for human-like agents and human-like agents for episodic memory. *International Journal of Machine Consciousness*, *2*, 227–244.

Bull, L. (2004). Applications of learning classifier systems. In L. Bull (Ed), *Applications of Learning Classifier Systems* (pp. 260–275), Berlin: Springer-Verlag.

Bull, L., & Kovacs, T. (2005). Foundations of learning classifier systems: An introduction. In L. Bull & T. Kovacs (Eds.), *Foundations of learning classifier systems* (pp. 1–20). Berlin: Springer-Verlag.

Burgess, N., Donnett, J. G., Jeffery, K. J., & O'Keefe, J. (1997). Robotic and neuronal simulation of the hippocampus and rat navigation. *Philosophical Transactions of the Royal Society of London B*, *352*, 1535–1543.

Canamero, L. (2005). Emotion understanding from the perspective of autonomous robots research. *Neural Networks*, *18*, 445–455.

Cassimatis, N. L., Trafton, J. G., Bugajska, M. D., & Schultz, A. C. (2004). Integrating cognition, perception and action through mental simulation in robots. *Robotics and Autonomous Systems*, *49*, 13–23.

Clayton, N., Dally, J., & Emery, N. (2007). Social cognition by food-caching corvids: the western scrub jay as a natural psychologist. *Philosophical Transactions of the Royal Society of London B*, *362*, 507–522.

Clayton, N., & Dickinson, A. (1998). Episodic-like memory during cache recovery by scrub jays. *Nature*, *395*, 272–274.

Clayton, N., Griffiths, D., Emery, N., & Dickinson, A. (2001). Elements of episodic-like memory in animals. In A. Baddeley, M. Conway & J. Aggleton (Eds.), *Episodic memory: new directions in research* (pp. 232–248). Oxford: Oxford University Press.

Clayton, N., & Russell, J. (2009). Looking for episodic memory in animals and young children: Prospects for new a minimalism. *Neuropsychologia*, *47*, 2330–2340.

Clayton, N., & Salwiczek, L. (2007). Episodic memory. *Current Biology*, *17*, R189–R191.

Cliff, D. (1991). Computational neuroethology: A provisional manifesto. In J.-A. Meyer & S. Wilson (Eds.), *From Animals to Animats: Proceedings of the First International Conference on the Simulation of Adaptive Behavior* (pp. 29–39), Cambridge, MA: MIT Press.

Cohen, N., & Squire, L. (1980). Preserved learning and retention of pattern analyzing skill in amnesia: Dissociation of knowing how and knowing that. *Science*, *210*, 207–209.

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

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin*, *104*, 163–191.

Cowan, N. (1999). An embedded-processes model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 62–101). Cambridge University Press.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*, 671–684.

Crestani, F. (1997). Application of spreading activation techniques in information retrieval. *Artificial Intelligence Review*, *11*, 453–482.

Curran, T. (2000). Brain potentials of recollection and familiarity. *Memory and Cognition*, *28*, 923–938.

Curran, T., DeBuse, C., Woroch, B., & Hirchman, E. (2006). Combined pharmacological and electrophysiological dissociation of familiarity and recollection. *Journal of Neuroscience*, *26*, 1979–1985.

Damasio, A. (1994). *Descartes error: emotion, reason and the human brain*. New York: Grosset/Putnam.

Derbinsky, N., & Laird, J. E. (2010). Extending Soar with dissociated symbolic memories. In M. Lim & W. C. Ho (Eds.), *Remembering Who We Are: Human Memory for Artificial Agents Symposium, AISB 2010* (pp. 31–37) SSAISB.

Deutsch, T., Gruber, A., & Lang, R. (2008). Episodic memory for autonomous agents. In *2008 Conference on Human System Interactions* (pp. 621–626). Los Alamos, CA: IEEE.

Di Paolo, E., & Iizuka, H. (2008). How (not) to model autonomous behaviour. *BioSystems*, *91*, 409–423.

Dodd, W., & Gutierrez, R. (2005). The role of episodic memory and emotion in a cognitive robot. In *IEEE International Workshop on Robot and Human Interactive Communication, RoMan 2005* (pp. 692–697). IEEE.

Dorffner, G., & Prem, E. (1993). Connectionism, symbol grounding, and autonomous agents. In *Proceedings of the 15th annual meeting of the Cognitive Science Society* (pp. 144–148), Hillside, NJ: Lawrence Erlbaum Associates.

Dorffner, G., Prem, E., & Trost, H. (1993). Words, symbols, and symbol grounding. In R. Casati, B. Smith & G. White (Eds.), *Philosophy and the cognitive sciences: Proceedings of the 16th international Wittgenstein symposium*. Vienna: Hölder-Pichler-Tempsky.

Duda, R., & Shortliffe, E. (1983). Expert systems research. *Science*, *220*, 261–268.

Fritzke, B. (1994). Growing cell structures–A self-organizing network for unsupervised and supervised learning. *Neural Networks*, *7*, 1441–1460.

Fritzke, B. (1995). A growing neural gas network learns topologies. In G. Tesauro, D. Touretzky & T. Leen (Eds.), *Advances in neural information processing systems 7*. Cambridge, MA: MIT Press.

Fuster, J. (1997). Network memory. *Trends in Neurosciences*, *20*, 451–459.

Fuster, J. (2000). Cortical dynamics of memory. *International Journal of Psychophysiology*, *35*, 155–164.

Gardenfors, P. (2004). *Conceptual Spaces: the geometry of thought*. Cambridge, MA: MIT Press.

Glenberg, A. M. (1997). What memory is for. *The Behavioral and Brain Sciences*, *20*, 1–19 [discussion 19–55].

Goertzel, B., Lian, R., Arel, I., Garis, H. D., & Chen, S. (2010). World survey of artificial brains, Part II: Biologically inspired cognitive architectures. *Neurocomputing*, *74*, 30–49.

Gorski, N. A., & Laird, J. E. (2010). Learning to use episodic memory. *Cognitive Systems Research*, *12*(2), 1–10.

Graf, P., & Schacter, D. (1984). Implicit and explicit memory for new associations in normal and amnesic subjects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*, 501–518.

Griffiths, D., Dickinson, A., & Clayton, N. (1999). Episodic memory: what can animals remember about their past? *Trends in Cognitive Sciences*, *3*, 74–80.

Harnad, S. (1990). The symbol grounding problem. *Physica D*, *42*, 335–346.

Hawes, N., Sloman, A., Wyatt, J., Zillich, M., Jacobsson, H., Kruijff, G.-J. M., et al. (2007). Towards an integrated robot with multiple cognitive functions. In R. C. Holte & A. Howe (Eds.), *Proceedings of the Twenty-Second AAAI Conference on Artificial Intelligence (AAAI 2008)* (pp. 1548–1553). Menlo Park, CA: AAAI Press.

Hawes, N., & Wyatt, J. (2010). Engineering intelligent information-processing systems with cast. *Advanced Engineering Informatics*, *24*, 27–39.

Hayes-Roth, B. (1985). A blackboard architecture for control. *Artificial Intelligence*, *26*, 251–321.

Haykin, S. (1999). *Neural networks* (2nd ed.). New Jersey: Prentice-Hall.

Hebb, D. (1949). *The organization of behavior*. New York: John Wiley and Sons, Inc.

Ho, W. C., Dautenhahn, K., Lim, M. Y., Vargas, P. A., Aylett, R., & Enz, S. (2009). An initial memory model for virtual and robot companions supporting migration and long-term interaction. In *RO-MAN 2009 - The 18th IEEE International Symposium on Robot and Human Interactive Communication* (pp. 277–284). IEEE.

Ho, W. C., Dautenhahn, K., & Nehaniv, C. (2008). Computational memory architectures for autobiographic agents interacting in a complex virtual environment: a working model. *Connection Science*, *20*, 21–65.

Ho, W. C., & Watson, S. (2006). Autobiographic knowledge for believable virtual characters. In *Intelligent Virtual Agents (IVA 2006)*, LNAI (pp. 383–394). Berlin: Springer.

Husbands, P. (2009). Never mind the iguana, what about the tortoise? Models in adaptive behavior. *Adaptive Behavior*, *17*, 320–324.

Ishikawa, M. (1996). Structural learning with forgetting. *Neural Networks*, *9*, 509–521.

Jaeger, H. (2002). *Short term memory in echo state networks* (Tech. Rep. GMD 152). Fraunhofer Institute for Autonomous Intelligent Systems.

Johnsson, M., & Balkenius, C. (2009). Associative self-organizing maps. In A. Dourado, A.C. Rosa & K. Madani (Eds.), *Proceedings of the International Joint Conference on Computational Intelligence* (pp. 363-370), Setubal: INIS-TICC Press.

Kadlec, R., & Brom, C. (2010). I can (almost) remember what you are doing: from actions to tasks. In M. Lim & W. C. Ho (Eds.), *Remembering Who We Are - Human Memory for Artificial Agents Symposium, AISB 2010* (pp. 27–30). AISB.

Kawamura, K., & Gordon, S. (2006). From intelligent control to cognitive control. In *Eleventh International Symposium on Robotics and Applications (ISORA 2006)*.

Kawamura, K., Peters, R. A., Bodenheimer, R. E., Sarkar, N., Park, J., Clifton, C. A., et al. (2004). A parallel distributed cognitive control system for a humanoid robot. *International Journal of Humanoid Robotics*, *1*, 65–93.

Kelley, T. D. (2003). Symbolic and sub-symbolic representations in computational models of human cognition: What can be learned from biology? *Theory and Psychology*, *13*, 847–860.

Kieras, D. E., & Meyer, D. E. (1997). An overview of the EPIC architecture for cognition and performance with application to human-computer interaction. *Human–Computer Interaction*, *12*, 391–438.

Kieras, D. E., Meyer, D. E., Mueller, S., & Seymour, T. (1999). Insights into working memory from the perspective of the EPIC architecture for modelling skilled perceptual-motor and cognitive human performance. In A. Miyake & P. Shah, *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 183–223). Cambridge: Cambridge University Press.

Kohonen, T. (2001). *Self-organizing maps* (3rd ed.). Berlin: Springer-Verlag.

Krichmar, J. L. (2008). The neuromodulatory system: A framework for survival and adaptive behavior in a challenging world. *Adaptive Behavior*, *16*, 385–399.

Krichmar, J. L., & Edelman, G. M. (2002). Machine psychology: Autonomous behavior, perceptual categorisation and conditioning in a brain-based device. *Cerebral Cortex*, *12*, 818–830.

Krichmar, J. L., Nitz, D. A., Gally, J. A., & Edelman, G. M. (2005). Characterising functional hippocampal pathways in a brain-based device as it solves a spatial memory task. *Proceedings of the National Academy of Sciences of the USA*, *102*, 2111–2116.

Kuipers, B. (2000). The spatial semantic hierarchy. *Artificial Intelligence*, *119*, 191–233.

Kuipers, B., & Byun, Y.-T. (1991). A robot exploration and mapping strategy based on a semantic hierarchy of spatial representations. *Robotics and Autonomous Systems*, *8*, 47–63.

Lang, R. (2007). *Dynamic neural networks*. Unpublished doctoral dissertation, University of Reading.

Langley, P., Laird, J. E., & Rogers, S. (2009). Cognitive architectures: Research issues and challenges. *Cognitive Systems Research*, *10*, 141–160.

Lee, M. H., Meng, Q., & Chao, F. (2007). Developmental learning for autonomous robots. *Robotics and Autonomous Systems*, *55*, 750–759.

Lovett, M. C., Reder, L. M., & Lebiere, C. (1999). Modeling working memory in a unified architecture. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 135–182). Cambridge: Cambridge University Press.

Luciw, M., & Weng, J. (2010). Top-down connections in self-organizing Hebbian networks: Topographic class

grouping. *IEEE Transactions on Autonomous Mental Development*, 2, 248–261.

Lungarella, M., Metta, G., Pfeifer, R., & Sandini, G. (2003). Developmental robotics: a survey. *Connection Science*, 15, 151–190.

Maes, P. (1989). How to do the right thing. *Connection Science*, 1, 291–232.

Mallot, H. A., & Basten, K. (2008). Embodied spatial cognition: Biological and artificial systems. *Image and Vision Computing*, 27, 1658–1670.

Mayes, A., & Roberts, N. (2001). Theories of episodic memory. *Philosophical Transactions of the Royal Society of London B*, 356, 1395–1408.

McKinstry, J. L., Seth, A. K., Edelman, G. M., & Krichmar, J. L. (2008). Embodied models of delayed neural responses: Spatiotemporal categorization and predictive motor control in brain based devices. *Neural Networks*, 21, 553–561.

Miller, G. (2003). The cognitive revolution: a historical perspective. *Trends in Cognitive Sciences*, 7, 141–144.

Mirza, N., Nehaniv, C., Dautenhahn, K., & te Boekhorst, R. (2007). Grounded sensorimotor interaction histories in an information theoretic metric space for robot ontogeny. *Adaptive Behavior*, 15, 167–187.

Morris, R. (2001). Episodic-like memory in animals: psychological criteria, neural mechanisms and the value of episodic-like tasks to investigate animal models of neurodegenerative disease. *Philosophical Transactions of the Royal Society of London B*, 356, 1453–1465.

Morse, A., & Aktius, M. (2009). Dynamic liquid association: Complex learning without implausible guidance. *Neural Networks*, 22, 875–889.

Morse, A. F., de Greeff, J., Belpaeme, T., & Cangelosi, A. (2010). Epigenetic robotics architecture. *IEEE Transactions on Autonomous Mental Development*, 2, 325–339.

Munakata, Y., & McClelland, J. L. (2003). Connectionist models of development. *Developmental Science*, 6, 413–429.

Neal, M., & Timmis, J. (2003). Timidity: A useful emotional mechanism for robot control? *Informatica*, 4, 197–204.

Newell, A. (1990). *Unified theories of cognition* (2nd ed.).Cambridge, MA: Harvard University Press.

Newell, A., & Simon, H. A. (1976). Computer science as empirical inquiry: Symbols and search. *Communications of the ACM*, 19, 113–126.

Nuxoll, A. (2007). *Enhancing intelligent agents with episodic memory*. Unpublished doctoral dissertation, University of Michigan.

Nuxoll, A., & Laird, J. (2007). Extending cognitive architecture with episodic memory. In R. C. Holte & A. Howe (Eds.), *AAAI Conference on Artificial Intelligence* (pp. 1560–1565). AAAI Press.

Nuxoll, A., Tecuci, D., Ho, W. C., & Wang, N. (2010). Comparing forgetting algorithms for artificial episodic memory systems. In M. Lim & W. C. Ho (Eds.), *Remembering Who We Are - Human Memory for Artificial Agents Symposium, AISB 2010* (pp. 14–20), SSAISB.

Nyhus, E., & Curran, T. (2010). Functional role of gamma and theta oscillations in episodic memory. *Neuroscience and Biobehavioural Reviews*, 34, 1023–1035.

Oberauer, K. (2002). Access to information in working memory: Exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28, 411–421.

O'Reilly, R. C., Braver, T. S., & Cohen, J. D. (1999). A biologically based computational model of working memory. In A. Miyake & P. Shah (Eds.) *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control* (pp. 375–411), Cambridge: Cambridge University Press.

O'Reilly, R. C., & Norman, K. A. (2002). Hippocampal and neocortical contributions to memory: advances in the complementary learning systems framework. *Trends in Cognitive Sciences*, 6, 505–510.

Packard, M., & Knowlton, B. (2002). Learning and memory functions of the basal ganglia. *Annual Review of Neuroscience*, 25, 563–593.

Pessoa, L. (2008). On the relationship between emotion and cognition. *Nature Reviews: Neuroscience*, 9, 148–158.

Peters, R., Hambuchen, K. E., Kawamura, K., & Wilkes, D. M. (2001). The sensory ego-sphere as a short-term memory for humanoids. In *Second IEEE-RAS International Conference on Humanoid Robots* (pp. 451–459).

Poucet, B., Lenck-Santini, P.-P., Paz-Villagran, V., & Save, E. (2003). Place cells, neocortex and spatial navigation: a short review. *Journal of Physiology*, 97, 537–546.

Prescott, T. J., Montes Gonzalez, F. M., Gurney, K., Humphries, M. D., & Redgrave, P. (2006). A robot model of the basal ganglia: Behavior and intrinsic processing. *Neural Networks*, 19, 31–61.

Quinton, J.-C., & Buisson, J.-C. (2008). Interactivist sensorimotor learning: Computational implementation and parallel optimization. In *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2008)*.

Rolls, E., & Treves, A. (1998). *Neural networks and brain function*. Oxford: Oxford University Press.

Rumelhart, D. E., & Norman, D. A. (1973). Active semantic networks as a model of human memory. In *Third International Joint Conference on Artificial Intelligence* (pp. 450–457) San Francisco: Morgan Kaufmann.

Schöner, G., Dose, M., & Engels, C. (1995). Dynamics of behavior: Theory and applications for autonomous robot architectures. *Robotics and Autonomous Systems*, 16, 213–245.

Schutte, A. R., Spencer, J. P., & Schöner, G. (2003). Testing the dynamic field theory: Working memory for locations becomes more spatially precise over development. *Child Development*, 74, 1393–1417.

Seth, A. K. (2009). Don't throw the baby iguana out with the bathwater. *Adaptive Behavior*, 17, 338–342.

Shanahan, M. (2005). Consciousness, emotion, and imagination: A brain inspired architecture for cognitive robotics. In *AISB '05 Workshop: Next Generation Approaches to Machine Consciousness* (pp. 26–35). SSAISB.

Shanahan, M. (2008). A spiking neuron model of cortical broadcast and competition. *Consciousness and Cognition*, 17, 288–303.

Siegelmann, H. (2008). Analog-symbolic memory that tracks via reconsolidation. *Physica D: Nonlinear Phenomena*, 237, 1207–1214.

Sikstrom, S. (2002). Forgetting curves: implications for connectionist models. *Cognitive Psychology*, 45, 95–152.

Slocum, A., Downey, D., & Beer, R. (2000). Further experiments in the evolution of minimally cognitive behaviour:

From perceiving affordances to selective attention. In J.-A. Meyer, A. Berthoz, D. Floreano, H. Roitblat & S. Wilson (Eds.), *From Animals to Animats 6: Proceedings of the Sixth International Conference on Simulation of Adaptive Behavior* (pp. 430–439). Cambridge, MA: MIT Press.

Sloman, A., & Chappell, J. (2005). The altricial-precocial spectrum for robots. In *19th International Joint Conference on Artificial Intelligence IJCAI'05* (Vol. 1, pp. 1187). San Francisco: Morgan Kaufmann Publishers.

Sowa, J. F. (Ed.). (1991). *Principles of semantic networks: explorations in the representation of knowledge*. San Mateo, CA: Morgan Kaufmann.

Sowa, J. F. (1992). Semantic networks. In *Encyclopedia of artificial intelligence*, Volume 5. John Wiley and Sons, Inc.

Squire, L. (2004). Memory systems of the brain: A brief history and current perspective. *Neurobiology of Learning and Memory*, *82*, 171–177.

Sun, R. (2004). Desiderata for cognitive architectures. *Philosophical Psychology*, *17*, 341–373.

Sun, R. (2007). The importance of cognitive architectures: an analysis based on clarion. *Journal of Experimental and Theoretical Artificial Intelligence*, *19*, 159–193.

Sun, R., Merrill, E., & Peterson, T. (2001). From implicit skills to explicit knowledge: A bottom-up model of skill learning. *Cognitive Science*, *25*, 203–244.

Sutton, R., & Barto, A. (1998). *Reinforcement learning: An introduction*. Cambridge, MA: MIT Press.

Taylor, J., Hartley, M., Taylor, N., Panchev, C., & Kasderidis, S. (2009). A hierarchical attention-based neural network architecture, based on human brain guidance, for perception, conceptualisation, action and reasoning. *Image and Vision Computing*, *27*, 1641–1657.

Temple, C. M., & Richardson, P. (2004). Developmental amnesia: A new pattern of dissociation with intact episodic memory. *Neuropsychologia*, *42*, 764–781.

Tulving, E. (2001). Episodic memory and common sense: how far apart? *Philosophical Transactions of the Royal Society of London B*, *356*, 1505–1515.

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

Van der Velde, F., & de Kamps, M. (2006). Neural blackboard architectures of combinatorial structures in cognition. *The Behavioral and Brain Sciences*, *29*, 37–70.

Vargas, P. A., Freitas, A. A., Lim, M., Enz, S., Ho, W. C., & Aylett, R. (2010). Forgetting and generalisation in memory modelling for robot companions: a data mining approach. In M. Lim & W. C. Ho (Eds.), *Remembering Who We Are - Human Memory for Artificial Agents Symposium, AISB 2010* (pp. 42–45) SSAISB.

Vargas, P. A., Ho, W. C., Lim, M., Enz, S., & Aylett, R. (2009). To forget or not to forget: Towards a roboethical memory control. In G. Michaelson & R. Aylett (Eds.) *Proceedings of the Symposium Killer Robots vs Friendly Fridges: The Social Understanding of Artificial Intelligence* (pp. 18-23), SSAISB.

Vargha-Khadem, F., Gadian, D., & Mishkin, M. (2001). Dissociations in cognitive memory: The syndrome of developmental amnesia. In A. Baddeley, M. Conway & J. Aggleton (Eds.), *Episodic memory: New directions in research*. Oxford: Oxford University Press.

Vargha-Khadem, F., Gadian, D., Watkins, K. E., Connelly, W., Van Paesschen, W., & Mishkin, M. (1997). Differential effects of early hippocampal pathology on episodic and semantic memory. *Science*, *227*, 376–380.

Webb, B. (2009). Animals versus animats: Or why not model the real iguana? *Adaptive Behavior*, *17*, 269–286.

Weng, J. (2007). On developmental mental architectures. *Neurocomputing*, *70*, 2303–2323.

Weng, J., McClelland, J., Pentland, A., Sporns, O., Stockman, I., Sur, M., et al. (2001). Autonomous mental development by robots and animals. *Science*, *291*, 599–600.

Wixted, J. T. (2004). The psychology and neuroscience of forgetting. *Annual Review of Psychology*, *55*, 235–269.

Wolff, M., Loukavenko, E., Will, B., & Dalrymple-Alford, J. (2008). The extended hippocampal-diencephalic memory system: Enriched housing promotes recovery of the flexible use of spatial representations after anterior thalamic lesions. *Hippocampus*, *18*, 996–1007.

Wolpert, D., Ghahramani, Z., & Flanagan, J. (2001). Perspectives and problems in motor learning. *Trends in Cognitive Sciences*, *5*, 487–494.

Wood, R., Baxter, P., & Belpaeme, T. (2010). A developmental perspective on memory centered cognition. In B. Johansson, E. Sahin & C. Balkenius (Eds.), *Proceedings of the Tenth International Conference on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems* (Vol. 149, pp. 183–184). Lund: LUCS.

Wood, R., & Di Paolo, E. (2007). New models for old questions: Evolutionary robotics and the 'A not B' error. In F. Almeida e Costa, L. M. Rocha, E. Costa, I. Harvey & A. Coutinho (Eds.), *Advances in Artificial Life: Proceedings of the 9th European Conference on Artificial life., LNCS* (Vol. 4648, pp. 1141–1150). Berlin: Springer-Verlag.

Wyatt, J. L., & Hawes, N. (2008). Multiple workspaces as an architecture for cognition. In A.V. Samsonovitch (Ed) *2008 Fall Symposium on Biologically Inspired Cognitive Architectures,* Technical Report (AAAI) FS-08-04 (pp. 201–206), Menlo Park, CA: AAAI Press.

Wyss, R., Konig, P., & Verschure, P. (2006). A model of the ventral visual system based on temporal stability and local memory. *PLoS Biology*, *4*(5), e120, 1–8.

Young, R. M., & Lewis, R. L. (1999). *The Soar cognitive architecture and human working memory*, (pp. 224–256). Cambridge: Cambridge University Press.

Ziemke, T. (2008). On the role of emotion in biological and robotic autonomy. *BioSystems*, *91*, 401–408.

Ziemke, T., & Thieme, M. (2002). Neuromodulation of reactive sensorimotor mappings as a short-term memory mechanism in delayed response tasks. *Adaptive Behavior*, *10*, 185–199.

About the Authors

Rachel Wood received a Ph.D. in computer science and artificial intelligence from the University of Sussex in 2008. She is currently a research fellow in the Centre for Robotics and Neural Systems at the University of Plymouth working on the FP7 project "Adaptive Strategies for Sustainable Long-Term Social Interaction" (ALIZ-E). Her research currently focuses on the role of memory in embodied social interaction. She has published articles in *Artificial Life* and *Phenomenology and the Cognitive Sciences* among other journals. *Address*: School of Computing and Mathematics, University of Plymouth, Portland Square A216, Plymouth PL4 8AA, UK. E-mail: rachel.wood@plymouth.ac.uk
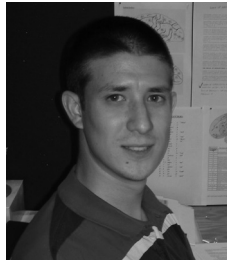
Paul E. Baxter has been a research fellow in the Centre for Robotics and Neural Systems at the University of Plymouth since 2010. He received an M.Eng. (Hons) in electronic engineering from the University of Nottingham, UK (2005), and completed a Ph.D. on the formulation of a memory-based perspective on cognition for robotics within a developmental framework in the School of Systems Engineering at the University of Reading, UK (2010). His current research seeks to build upon this memory-based cognitive robotics perspective, within the context of long-term any-depth human–robot interaction, as part of the EU FP7 ALIZ-E project. *Address*: School of Computing and Mathematics, University of Plymouth, Portland Square A216, Plymouth PL4 8AA, UK. *E-mail*: paul.baxter@plymouth.ac.uk.

Dr Tony Belpaeme is Associate Professor in Intelligent Systems at the University of Plymouth, where he has been since 2005. He received his Ph.D. in 2003 from the Vrije Universiteit Brussel on a study of cultural transmission of perceptual and color categories. Dr Belpaeme's research is broadly concerned with understanding natural cognition and building artificial cognition on robots, the building of social robots, and the understanding of the impact of language on natural and artificial cognition. His work is funded by the EPSRC and the EU FP7 program through several large-scale research projects, such as ITALK and ALIZ-E. *Address*: School of Computing and Mathematics, University of Plymouth, Portland Square A216, Plymouth PL4 8AA, UK. *E-mail*: tony.belpaeme@plymouth.ac.uk.