A Cognitive Science Based Machine Learning Architecture

Sidney D'Mello[1], Stan Franklin[1], Uma Ramamurthy[1], & Bernard Baars[2]

[1] University of Memphis, Memphis, Tennessee 38152, USA
[2]The Neurosciences Institute, San Diego, California 92121, USA
[1]{sdmello|franklin|urmmrthy}@memphis.edu; [2]baars@nsi.edu

Abstract

In an attempt to illustrate the application of cognitive science principles to hard AI problems in machine learning we propose the LIDA technology, a cognitive science based architecture capable of more human-like learning. A LIDA based software agent or cognitive robot will be capable of three fundamental, continuously active, human-like learning mechanisms: 1) *perceptual learning*, the learning of new *objects*, *categories*, *relations*, etc., 2) *episodic learning* of events, the *what*, *where*, and *when*, 3) *procedural learning*, the learning of new *actions* and *action sequences* with which to accomplish new tasks. The paper argues for the use of modular components, each specializing in implementing individual facets of human and animal cognition, as a viable approach towards achieving general intelligence.

Relevance of Cognitive Science to AI

Dating back to Samuel's checker player (1959), machine learning is among the oldest of the sub-branches of AI with many practitioners and many successes to its credit. Still, after fifty years of effort there are remaining difficulties. Machine learning often requires large, accurate training sets, shows little awareness of what's known or not known, integrates new knowledge poorly into old, learns only one task at a time, allows little transfer of learned knowledge to new tasks, and is poor at learning from human teachers. Clearly, machine learning presents a number of hard AI problems. Can cognitive science help?

In contrast, human learning has solved many of these problems, and is typically continual, quick, efficient, accurate, robust, flexible, and effortless. As an example consider perceptual learning, the learning of new objects, categories, relations, etc. Traditional machine learning approaches such as object detection, classification, clustering, etc, are highly susceptible to the problems raised above. However, perceptual learning in humans and animals seem to have no such restrictions. Perceptual learning in humans occurs incrementally so there is no need for a large training set. Learning and knowledge extraction are achieved simultaneously through a dynamical system that can adapt to changes in the nature of the stimuli perceived in the environment. Additionally, human like learning is based on reinforcement rather than fitting to a dataset or model. Therefore, in addition to learning, humans can also forget. Initially, many associations are made between entities, the ones that are sufficiently reinforced persist, while the ones that aren't decay.

All this suggests a possible heuristic: If you want smart software, copy it after humans. We've done just that. The Learning Intelligent Distribution Agent (LIDA) architecture that we propose here was designed to be consistent with what is known from cognitive science and neuroscience. In addition to being a computational architecture, it's intended to model human cognition. We'll go on to describe the LIDA architecture and its human-like learning capabilities.

Modularity as an Approach for Intelligence

LIDA provides a conceptual and computational model of cognition. It is the learning extension of the original IDA system implemented as a software agent. IDA 'lives' on a computer system with connections to the Internet and various databases, and does personnel work for the US Navy, performing all the specific personnel tasks of a human (Franklin 2001).

The LIDA architecture is partly symbolic and partly connectionist with all symbols being grounded in the physical world in the sense of Brooks (1986; 1990). We argue for unique, specialized mechanisms to computationally implement the various facets of human cognition such as perception, episodic memories, functional consciousness, and action selection. We offer an evolutionary argument and a functional argument to justify the specialized, modular component approach to the design of an intelligent system.

The evolutionary argument draws support from the sustained efforts by neuroscientists and brain physiologists in mapping distinct functions to different areas in the brain. In many ways the brain can be viewed as a kluge of different mechanisms. For example, parts of perceptual associative memory are believed to be in the perirhinal cortex (Davachi, Mitchell & Wagner 2003), while some of the neural correlates of autobiographical memory have been identified as the medial frontal cortex and left hippocampus (Conway & Fthenaki 2000; Maguire 2001). In addition to the neuroscience evidence, there are developmental arguments for a distinct mechanism for perceptual memory. Infants who have not yet developed object permanence (any episodic memory) are quite able to recognize and categorize (Mandler 2000). Other arguments come from studies of human amnesiacs with significant loss of declarative memory, but mostly intact perceptual associative memory and learning (Gabrieli et al. 1990, Fahle & Daum 2002). Perhaps the most convincing argument comes from experiments with rats in a radial arm maze. With four arms baited and four not (with none restocked), normal rats learn to recognize which arms to search (perceptual associative memory) and remember which arms they have already fed in (episodic memory) so as not to search there a second time. Rats with their hippocampal systems excised lose their episodic memory but retain perceptual associative memory, again arguing for distinct mechanisms (Olton, Becker, & Handelman 1979). Similarly, arguments for finer distinctions between the various episodic memory systems have been made. Episodic memories are memories for events of the what, where, and when. Conway (2001) argues for a unique memory system for recent, highly specific, sensory-perceptual information, than autobiographical memory, on the basis of different functions, knowledge stored, access techniques, phenomenology, and neurology. Additionally, Moscovitch et al. (2005) offer neuroimaging evidence for a finer grained component analysis for semantic and spatial memories.

The functional arguments in support of specialized modular components for general intelligence are derived from the need for primitives for any agent capable of robust autonomy. An initial set of primitive feature detectors for perception and a set of primitive effectors for action execution are a computational necessity for any autonomous agent, natural or artificial, software or robotic (Franklin 1997). Additionally, as widely recognized in humans, the need for primitive motivators implemented as feelings and emotions may also be required for any cognitive system that attempts to stimulate general intelligence. In humans, primitive feature detectors for vision include neurons in the primary visual cortex (V1) detecting line segments at various orientations, while primitive effectors include neuronal groups controlling individual muscles. Higher level visual perception such as categorization, object detection, etc, is realized from associations between the various primitive feature detectors. Similarly, more complex actions and action sequences are learnt by associating the primitive effectors. These functional requirements that differ between perception and action strengthen the argument for specialized modular components.

The LIDA Architecture

On the basis of the arguments for specialized, modular components as an approach for intelligence, the LIDA architecture operates by the interplay of unique mechanisms implementing the major facets of human cognition. The mechanisms used in implementing the several modules have been inspired by a number of different 'new AI' techniques (Drescher 1991; Hofstadter & Mitchell 1994; Jackson 1987; Kanerva 1988; Maes 1989; Brooks 1986). We now describe LIDA's primary mechanisms.

Perceptual Associative Memory. LIDA perceives both exogenously and endogenously with Barsalou's perceptual symbol systems serving as a guide (1999). The perceptual knowledge-base of this agent, called perceptual associative memory, takes the form of a semantic net with activation called the slipnet, a la Hofstadter and Mitchell's Copycat architecture (1994). Nodes of the slipnet constitute the agent's perceptual symbols, representing individuals, categories, and perhaps higher-level ideas and concepts. Pieces of the slipnet containing nodes and links, together with perceptual codelets (a codelet is a small piece of code running independently; perceptual codelets are a special type of codelet designed for perceptual tasks such as recognition) with the task of copying the piece to working memory, constitute Barsalou's perceptual symbol simulators (1999). Together they constitute an integrated perceptual system for LIDA, allowing the system to recognize, categorize and understand.

Workspace. LIDA's workspace is analogous to the preconscious buffers of human working memory. Perceptual codelets write to the workspace as do other, more internal codelets. Attention codelets (codelets that form coalitions with other codelets to compete for functional consciousness) watch what is written in the workspace in order to react to it. Items in the workspace decay over time, and may be overwritten.

Another pivotal role of the workspace is the building of temporary structures over multiple cognitive cycles (see below). Perceptual symbols from the slipnet are assimilated into existing relational and situational templates while preserving spatial and temporal relations between the symbols. The structures in the workspace also decay rapidly.

Episodic Memory. Episodic memory in the LIDA architecture is composed of a declarative memory (DM) for the long term storage of autobiographical and semantic information as well as a short term transient episodic memory (TEM) similar to Conway's (2001) sensory-perceptual episodic memory with a retention rate measured in hours. LIDA employs variants of sparse distributed memory (SDM) to computationally model DM and TEM (Kanerva 1988; D'Mello, Ramamurthy, & Franklin 2005). SDM is a content addressable memory that, in many ways, is an ideal computational mechanism for use as an associative memory system.

Functional Consciousness. LIDA's 'consciousness' module implements Global Workspace (GW) theory's (Baars 1988) processes by codelets, small pieces of code each running independently. These are specialized for some simple task and often play the role of a daemon watching for an appropriate condition under which to act. The apparatus for functional 'consciousness' consists of a coalition manager, a spotlight controller, a broadcast manager, and of attention codelets that recognize novel or problematic situations.

Procedural Memory. Procedural memory in LIDA is a modified and simplified form of Drescher's schema mechanism (1991), the scheme net. Like the slipnet of perceptual associative memory, the scheme net is a directed graph whose nodes are (action) schemes and whose links represent the 'derived from' relation. Built-in primitive (empty) schemes directly controlling effectors are analogous to motor cell assemblies controlling muscle groups in humans. A scheme consists of an action, together with its context and its result. At the periphery of the scheme net lie empty schemes (schemes with a simple action, but no context or results), while more complex schemes consisting of actions and action sequences are discovered as one moves inwards. In order for a scheme to act, it first needs to be instantiated and then selected for execution in accordance with the action selection mechanism described next.

2.2.6. Action Selection. The LIDA architecture employs an enhancement of Maes' behavior net (1989) for high-level action selection in the service of feelings and emotions. Several distinct feelings and emotions operate in parallel, perhaps varying in urgency as time passes and the environment changes. The behavior net is a digraph (directed graph) composed of behaviors (instantiated action schemes) and their various links. As in connectionist models, this digraph spreads activation. The activation comes from four sources: from pre-existing activation stored in the behaviors, from the environment, from feelings and emotions, and from internal states. To be acted upon, a behavior must be executable, must have activation over threshold, and must have the highest such activation.

The LIDA Cognitive Cycle

Since the LIDA architecture is composed of several specialized mechanisms, each implementing various facets of human cognition, the need for a continual process that causes the functional interaction among the various components becomes paramount. We offer the cognitive cycle as such an iterative, cyclical, continually active process that brings about the interplay among the various components of the architecture. Cognitive cycles are flexible, serial but overlapping cycles of activity usually beginning in perception and ending in an action. We suspect that cognitive cycles occur five to ten times a second in humans, cascading so that some of the steps in adjacent cycles occur in parallel (Baars & Franklin 2003). Seriality is preserved in the conscious broadcasts. We now describe the cognitive cycle dividing it into nine steps.

1) Perception. Sensory stimuli, external or internal, are received and interpreted by perception producing the beginnings of meaning.

2) Percept to preconscious buffer. The percept, including some of the data plus the meaning, as well as possible relational structures, is stored in the preconscious buffers of LIDA's working memory.

3) Local associations. Using the incoming percept and the residual contents of the preconscious buffers of working memory, including emotional content, as cues, local associations are automatically retrieved from transient episodic memory (TEM) and from declarative memory (DM) and stored in long-term working memory.

4) Competition for consciousness. Attention codelets view long-term working memory, and bring novel, relevant, urgent, or insistent events to consciousness.

5) Conscious broadcast. A coalition of codelets, typically an attention codelet and its covey of related information codelets carrying content, gains access to the global workspace and has its contents broadcast. In humans, this broadcast is hypothesized to correspond to phenomenal consciousness.

6) Recruitment of resources. Relevant schemes respond to the conscious broadcast. These are typically schemes whose context is relevant to information in the conscious broadcast. Thus consciousness solves the relevancy problem in recruiting resources.

7) Setting goal context hierarchy. The recruited schemes use the contents of consciousness, including feelings/emotions, to instantiate new goal context hierarchies (copies of themselves) into the behavior net, bind their variables, and increase their activation. Other, environmental, conditions determine which of the earlier goal contexts receive additional activation.

8) Action chosen. The behavior net chooses a single behavior (scheme, goal context), from a just instantiated behavior stream or possibly from a previously active stream. Each selection of a behavior includes the generation of an expectation codelet (see the next step).

9) Action taken. The execution of a behavior (goal context) results in the behavior codelets performing their specialized tasks, having external or internal consequences, or both. LIDA is taking an action. The acting codelets also include at least one expectation codelet whose task it is to monitor the action bringing to consciousness any failure in the expected results.

Human like Learning in Machines

For agents immersed in simple, mostly static, domains it is quite plausible for the system designer to architect the required knowledge so that the agent can effectively pursue its agenda. This approach, though tedious in its undertaking, has worked quite successfully for knowledge based systems such as expert systems, etc. However, as the complexity of the agents' world increases, the onus of the requisite knowledge engineering proves to be an extremely daunting task. In an attempt to circumvent these knowledge engineering problems we argue for mechanisms that support a developmental period, one of rapid learning, in the "lives" of both software agents and robots. Such a developmental period would circumvent the necessity of designing and implementing a complex ontology, a clear pragmatic advantage. In complex, dynamic environments, the learned ontology can be expected to out perform one designed and built in, and to do so with much less human effort.

The LIDA model realizes three fundamental learning mechanisms that underlie much of human learning: 1) perceptual learning, the learning of new objects, categories, relations, etc., 2) episodic learning of events, the what, where, and when, 3) procedural learning, the learning of new actions and action sequences with which to accomplish new tasks. Although, the type of knowledge retained due to these three learning mechanisms differ, the mechanisms are founded on two basic premises. The first premise states that conscious awareness is sufficient for learning. Although subliminal acquisition of information appears to occur, the effect sizes are quite small compared to conscious learning. In a classic study, Standing (1973) showed that 10,000 distinct pictures could be learned with 96% recognition accuracy, after only 5 seconds of conscious exposure to each picture. No intention to learn was needed. Consciously learned educational material has been recalled after 50 years (Bahrick, 1984). No effect sizes nearly as long-lasting as these have been reported in the subliminal learning literature (Elliott & Dolan 1998). Conscious access greatly facilitates most types of learning. The second premise that is shared among the various learning mechanisms is that the learning is modulated by feelings and emotions, i.e. the learning rate varies with arousal (Yerkes & Dodson 1908).

Perceptual Learning

Perceptual associative memory (PAM) is implemented in the LIDA architecture as a slipnet, a semantic net with passing activation (Hofstadter and Mitchell 1994). Perceptual learning in the LIDA model occurs with consciousness. This learning is of two forms, the strengthening or weakening of the base-level activation of existing nodes, as well as the creation of new nodes and links. Any existing concept or relation that appears in the conscious broadcast (Step 5 of the cognitive cycle) has the base-level activation of its corresponding node strengthened as a function of the arousal of the agent at the time of the broadcast. The base-level activation curve of a slipnet node is modeled along a saturating, sigmoid curve,

A new individual item that comes to consciousness results in a new node being created, together with links into it from the feature detectors of its features. Such a new item gets to consciousness by means of some *new-item attention codelet* that notices a collection of active features in the percept without a common object of which they are features. Such a new item-attention codelet might be looking for such features as spatial contiguity, common motion, and persistence over time. Here's how a new category may be formed. If a *similarity-attention codelet* notices in long-term working memory (see Step 4 of the cognitive cycle) two items with several common features, and succeeds in bringing this similarity to consciousness, a new category is created by the perceptual learning mechanism with *is-a* links into the new category from each of the items. New relation nodes occur similarly in a manner suggested by the work of Doumas and Hummel (2005). New relations are learned into nodes in PAM from structures built in the preconscious working memory buffers by perceptual codelets, that instantiate existing relation nodes from PAM and bind objects to their arguments. These new relation nodes are learned when *relation-noting-attention codelets* succeed in bringing the new relations to consciousness, that is when the relation 'pops into mind' or 'occurs to me.' New links are learned along with the nodes. The initial base-level activation of new nodes, be they object, category or relation nodes, are assigned as a function of arousal at the time of the conscious broadcast (Step 5 of the cognitive cycle above).

Episodic Learning

In the LIDA model memory is hypothesized to interact with conscious events for its normal functioning. Within the context of episodic memory we are concerned with *interpreting the contents of 'consciousness' so as to be able to encode the what, where and when* of each cognitive cycle into episodic memory. LIDA is endowed with two types of episodic memories, one with a small capacity for short term retention of detailed sensory-perceptual information (transient episodic memory, TEM) and the other with a large capacity for possibly long term storage of lifelong events and facts (declarative memory, DM).

Episodic learning in the LIDA architecture results from events taken from the contents of 'consciousness' being encoded in our modified sparse distributed memory (SDM) representing TEM (D'Mello, Ramamurthy, & Franklin 2005). Perceptual symbols (slipnet nodes) making up an event in 'conscious' contents are traced back along slipnet links to primitive feature detectors. These activation patterns of these primitive feature detectors are encoded into TEM thus effecting episodic learning within each cognitive cycle. The recall of events (local associations) from TEM and from DM will require routing the retrieved activation trace through perceptual associative memory so as to recover the corresponding perceptual symbols.

In addition to the encoding of sensory perceptual details of each episode manifested through the contents of consciousness, this episodic learning includes encoding of feelings and emotions, and of actions taken by the agent.

Periodically, and offline, the not yet decayed contents of TEM are consolidated into declarative memory (DM) (autobiographical memory plus semantic memory) which is also implemented as a modified SDM system.

Procedural Learning

We propose a combination of instructionalist as well as selectionist motivated agendas (not discussed here) for procedural learning in the scheme net (see procedural memory above), with functional consciousness providing reinforcement to actions. Reinforcement is provided via a sigmoid function such that initial reinforcement becomes very rapid but tends to saturate. The inverse of the sigmoid function that produces the reinforcement curve, serves as the decay curve. Therefore, schemes with low base-level activation decay rapidly, while schemes with high (saturated) base-level activation values tend to decay at a much lower rate.

Empty schemes, each containing only a primitive action with no context and results, lie at the periphery of the scheme net. Each scheme has two types of activation, a base-level activation and a current activation. The base-level activation measures the reliability of the scheme, that is, the likelihood of its result occurring when the action is taken within its context. The current activation measures the applicability of the scheme to the current situation. For learning to proceed initially, the behavior network must first select an instantiation of an empty scheme for execution. Since an empty scheme has no context, it is assumed to be always applicable to the situation at hand. Before executing its action, the instantiated scheme (behavior codelet) spawns a new expectation codelet. After the action is executed, this newly created expectation codelet focuses on changes in the environment as a result of the action being executed, and attempts to bring this information to consciousness. If successful, a new scheme is created, if needed. If one already exists, it is appropriately reinforced. Conscious information just before the action was executed becomes the context of this new scheme. Information brought to consciousness right after the action is used as the result of the scheme. The scheme is provided with some base-level activation, and it is connected to its parent empty scheme with a link.

Collections of behavior codelets form behaviors. The behavior codelets making up a behavior share preconditions and post conditions. Certain attention codelets notice behavior codelets that take actions at approximately the same time, though in different cognitive cycles. These attention codelets attempt to bring this information to consciousness. If successful, a new scheme is created, if it does not already exist. If it does exist, the existing scheme is simply reinforced, that is, its base-level activation is modified.

Collections of behaviors, called behavior streams, with activation passing links between them can be thought of as partial plans of actions. The execution of a behavior in a stream is conditional on the execution of its predecessor, and it directly influences the execution of its successor. When an attention codelet notices two behavior codelets executing within some small time span, it attempts to bring this information to consciousness. If successful, it builds a new scheme if such a scheme does not already exist.

Comparison with Current Technologies

Computational models have long been a major, and perhaps indispensable, tool in cognitive science. Many of these model some psychological theory of a particular aspect of cognition, attempting to account for experimental data. Others, such as the construction-integration model (Kintsch 1998), SOAR (Laird, Newell, & Rosenbloom 1987), and ACT-R (Anderson 1990), aspire to be general computational models of cognition. The LIDA architecture is a general computational model of cognition.

The specialized, modular approach to LIDA's architecture constitutes the major difference between LIDA and SOAR or ACT-R. Additional differences with SOAR arise from the fact that it (SOAR) does not provide a consciousness mechanism and that it is a purely symbolic, rules-based system unlike LIDA's hybrid, non-rules-based approach. Furthermore, the percept-motor interface in SOAR for interacting with the environment is entirely domain-specific, whereas the LIDA model allows for domain-independent perceptual and procedural learning mechanisms. ACT-R shares several similarities to the LIDA model such as the incorporation of both symbolic and sub-symbolic levels of abstraction. However, there is a very distinct separation between the two in ACT-R. The symbolic level is more concerned with production rules and chunks, while the sub-symbolic level consists of several parallel processes modeled by a set of equations.

Discussion

With the design of three continually active incremental learning mechanisms we have laid the foundation for a cognitive architecture capable of human like learning. We feel that integrating these three types of learning modulated by feelings and emotions in machines is an example of how cognitive science principles can be applied towards the hard problems of AI. No large training sets would be required. New knowledge would be easily integrated into old. Several tasks could be learned concurrently with transfer of knowledge to new tasks. A number of the old, hard AI problems would potentially be solved.

We conclude by speculating on the usefulness of AI towards obtaining a better understanding of human cognition. We believe that large scale working models of cognition such as the LIDA model can be useful as a tool to guide some of the research of cognitive scientists and neuroscientists by generating testable hypotheses about human (and animal) cognition. The LIDA model generates hypotheses about human cognition by way of its design, the mechanisms of its modules, their interaction, and its performance (Baars & Franklin 2003: Franklin et al 2005). While, all of these hypotheses are, in principle, testable, due to the relatively fine-grained level of analyses required to either confirm or deny them, more sophisticated brain and behavioral assessment technologies may be in order.

Using the LIDA architecture to implement cognitive robots and software agent performing real-world tasks will produce working models of cognition. Though experiments provide the gold standard for scientific evidence, it is not possible to test all parameters of actual working models of cognition. Experiment-based models typically have too few variables to accomplish real-world perception or control of action. Simulations based only on experimental evidence would simply fail in the real world. Hence, workability should be combined with experimental evidence as desirable features of cognitive models.

References

Anderson, J. R. (1990). *The Adaptive Character of Thought*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Baars, B. J. 1988. *A Cognitive Theory of Consciousness*. Cambridge.: Cambridge University Press.

Baars, B., and Franklin, S. 2003. How conscious experience and working memory interact. *Trends in Cognitive Science* 7:166-172.

Barsalou, L. W. 1999. Perceptual symbol systems. *Behavioral and Brain Sciences* 22:577–609.

Bahrick, H. P. 1984. Semantic memory content in permastore: fifty years of memory for Spanish learned in school. J. Exp. Psychol. Gen. 113:1-29.

Brooks, R. A. 1986. A robust layered control system for a mobile robot. *IEEE Journal of Robotics and Automation*. RA-2:14-23.

Brooks, R. A. 1990. Elephants Don't Play Chess. *Robotics and Autonomous Systems* 6: 3–15.

Conway, M. A. & Fthenaki, A. 2000. Disruption and loss of autobiographical memory. In *Handbook of neuropsychology: memory and its disorders, 2nd edition* (ed. L. Cermak), Ansterdam: Elsevier.

Conway, M. A. 2001. Sensory-perceptual episodic memory and its context: autobiographical memory. In *Episodic Memory*, ed. A. Baddeley, M. Conway, and J. Aggleton. Oxford: Oxford University Press.

Davachi, L., J. P. Mitchell, and A. D. Wagner. 2003. Multiple routes to memory: Distinct medial temporal lobe processes build item and source memories. Proc. *Natl. Acad. Sci*. USA, 100:2157-2162.

D'Mello S., Ramamurthy U., and Franklin S. 2005. Encoding and Retrieval Efficiency of Episodic Data in a Modified Sparse Distributed Memory System. Proceedings of the *27th Annual Meeting of the Cognitive Science Society*. Stresa, Italy

Doumas, L. A. A., and J. E. Hummel. 2005. A Symbolic-Connectionist Model of Relation Discovery. Proceedings of the *27th Annual Meeting of the Cognitive Science Society*. Stresa, Italy

Drescher, G. 1991. *Made Up Minds: A Constructivist Approach to Artificial Intelligence*, Cambridge, MA: MIT Press.

Elliott, R., and R. J. Dolan. 1998. Neural response during preference and memory judgments for subliminally presented stimuli: a functional neuroimaging study. *J. Neurosci*. 18:4697-4704.

Fahle, M., and Daum, I. 2002. Perceptual learning in amnesia. *Neuropsychologia* 40:1167–1172.

Franklin, S. 2001. Automating Human Information Agents. *Practical Applications of Intelligent Agents*, ed. Z. Chen, and L. C. Jain. Berlin: Springer-Verlag.

Franklin, S., Baars, B.J., Ramamurthy, U., and Ventura, M. 2005. The Role of Consciousness in Memory. Brains, Minds and Media, Vol.1, bmm150 (urn:nbn:de:0009-3-1505)

Franklin, S., and A. C. Graesser. 1997. Is it an Agent, or just a Program?: A Taxonomy for Autonomous Agents. In *Intelligent Agents III*. Berlin: Springer Verlag. 21-35.

Gabrieli, J. D., W. Milberg, M. M. Keane, and S. Corkin. 1990. Intact priming of patterns despite impaired memory. *Neuropsychologia* 28:417–427.

Hofstadter, D. R., and M. Mitchell. 1994. The Copycat Project: A model of mental fluidity and analogy-making. In *Advances in connectionist and neural computation theory*, Vol. 2: logical connections, ed. K. J. Holyoak, and J. A. Barnden. Norwood N.J.: Ablex.

Jackson, J. V. 1987. *Idea for a Mind*. Siggart Newsletter, 181:23–26.

Kanerva, P. 1988. *Sparse Distributed Memory*. Cambridge MA: The MIT Press.

Kintsch, W. 1998. *Comprehension*. New York: Cambridge University Press.

Laird, E., Newell, A. & Rosenbloom P. 1987. SOAR: An Architecture for General Intelligence. *Artificial Intelligence* 33:1-64.

Maes, P. 1989. How to do the right thing. *Connection Science* 1:291-323.

Maguire, E. A. 2001. Neuroimaging studies of autobiograpical event memory. *Phil. Trans. R. Soc. Lond*. B 356:1441-1451

Mandler, J. 2000. Perceptual and Conceptual Processes in Infancy. *Journal of Cognition and Development* 1:3-36.

Moscovitch, M., Rosenbaum, R.S., Gilboa, A., Addis, D.R., Westmacott, R., Grady, C., McAndrews, M.P., Levine, B., Black, S.E., Winocur, G. & Nadel, L. 2005. Functional neuroanatomy of remote episodic, semantic and spatial memory: A unified account based on multiple trace theory. *Journal of Anatomy*, 207, 35-66.

Olton, D. S., J. T. Becker, and G. H. Handelman. 1979. Hippocampus, space and memory. *Behavioral and Brain Sciences* 2:313-365.

Samuel, A. L. 1959. Some Studies in Machine Learning Using the Game of Checkers. *IBM J. Res. Develop.* 3:210-229

Standing, L. 1973. Learning 10,000 pictures. *Quarterly Journal of Experimental Psychology* 25:207-222.

Yerkes, R. M., and J. D. Dodson. 1908. The Relationship of Strength of Stimulus to Rapidity of Habit Formation. *Journal of Comparative Neurology and Psychology* 18:459–482.