A Procedural Learning Mechanism for Novel Skill Acquisition

Sidney D'Mello, Uma Ramamurthy, Aregahegn Negatu, & Stan Franklin
Department of Computer Science & Institute for Intelligent Systems
University of Memphis, Memphis, Tennessee 38152, USA
{sdmello|franklin|urmmrthy|asnegatu}@memphis.edu

Abstract

In this paper we attempt to develop mechanisms for procedural memory and procedural learning for cognitive robots on the basis of what is known about the same facilities in humans and animals. The learning mechanism will provide agents with the ability to learn new actions and action sequences with which to accomplish novel tasks.

1 The LIDA Model

The Learning Intelligent Distribution Agent (LIDA) architecture provides a conceptual and computational model of cognition. She is the partially conceptual, learning extension, of the original IDA system implemented computationally 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 major components of the LIDA architecture are perceptual associative memory, working memory, episodic memory, functional consciousness, procedural memory, action selection, and sensory-motor memory, with the last three being of interest to this paper. LIDA's mechanisms for procedural memory, action selection, and action realization (execution) are inspired by variants of models originally conceived by Drescher's schema mechanism (1991), Maes' behavior network (1989), and Brooks' subsumption architecture (1986) respectively.

Procedural memory in LIDA is a modified and simplified form of Drescher's schema mechanism (1991), the scheme net. 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. The context and results of the schemes are represented by perceptual symbols (Barsalou, 1999) for objects, categories, and relations in perceptual associative memory (not described here). The perceptual symbols are grounded in the real world by their ultimate connections to various primitive feature detectors having their receptive fields among the sensory receptors. The action of a scheme is connected to an appropriate network in sensory-motor memory (described later) that directly controls actuators.

Each scheme also maintains two statistics, a *base-level activation* and a *current activation*. The base-level activation (used for learning) is a measure of the scheme's overall reliability in the past. It estimates the likelihood of the result of the scheme occurring by taking the action given its context. The current activation is a measure of the relevance of the scheme to the current situation (environmental conditions, goals, etc,). 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.

The LIDA architecture employs an enhancement of Maes' behavior net (1989) for high-level action selection in the service of feelings and emotions. The behavior net is a digraph (directed graph) composed of behaviors codelets (a single action), behaviors (multiple behavior codelets operating in parallel), and behavior streams (multiple behaviors operating in an ordered sequence) and their various links. These three entities all share the same representation in procedural memory (i.e., a scheme).

Once an action has been selected, it triggers a suitable sub-network of the sensory-motor memory, modeled after Brook's subsumption architecture (Brooks, 1986). With sensors directly driving effectors, this sub-network effects the selected action.

2 Procedural Learning

Our model of procedural learning is based on functional consciousness, implemented in adherence to Global Workspace Theory (Baars, 1988), and reinforcement learning. Reinforcement is provided via a sigmoid function such that initial reinforcement becomes very rapid but tends to saturate. The inverse of this same sigmoid function 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.

For learning to proceed initially, the behavior network must first select the instantiation of an empty scheme for execution. Before executing its action, the instantiated scheme (activated behavior codelet) spawns a new expectation codelet (a codelet that tries to bring the results of an action to consciousness). 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 that operate in parallel form behaviors. The behavior codelets making up a behavior share preconditions and post conditions. Certain attention codelets (codelets that form coalitions with other codelets to compete for consciousness) notice behavior codelets that take actions at approximately the same time, though in different cognitive cycles (a cyclical process beginning with perception and ending in an action). 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. If a new scheme has to be created, its context is taken to be the union of the contexts of the schemes firing together. The result of the new scheme is the union of the results of the individual schemes. Additionally, this new scheme is provided with some base-level activation and is connected by links to the original schemes it includes. If this composite scheme executes in the future it will pass activation along these links.

Collections of behaviors, called behavior streams 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 in order within some small time span, it attempts to bring this information to consciousness. If successful, it builds a new scheme with links from the first scheme to the second, if such a scheme does not already exist, in which case the existing scheme is simply reinforced. If a new scheme has to be created, its context is the union of the contexts of the first scheme and the second, excluding the items that get negated by the result of the first. Similarly the result of the new scheme formed will be the union of both results, excluding the results of the first that are negated by the results of the second. Using such a learning mechanism iteratively, more complex streams can be built.

3 Discussion

With the continually active, incremental, procedural learning mechanism an autonomous agent will be capable of learning new ways to accomplish new tasks by creating new actions and action sequences. Although our model of procedural learning is motivated by Drescher's schema model (1991), the learning mechanism is different in two significant aspects. First, our approach maintains that functional conscious involvement is a necessary condition for supraliminal learning. The second distinction arises from the fact that while learning in Drescher's system relies on each schema maintaining several reliability statistics, we only use a single, computationally more tractable statistic, the base-level activation modeled by a saturating sigmoid function.

References

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

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

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

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

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

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