Proceedings of the AAAI-10 Workshop on Statistical Relational AI, 2010

An Architectural Approach to Statistical Relational AI

Paul S. Rosenbloom

Department of Computer Science and Institute for Creative Technologies
University of Southern California
13274 Fiji Way, Marina del Rey, CA 90292 USA
rosenbloom@usc.edu

Abstract

The architectural approach to AI focuses on the fixed structure underlying intelligence. Applying it to statistical relational AI should further stimulate the application of statistical relational techniques across AI, while focusing research on their commonalities, (in)compatibilities and integration. It could also yield new architectures that are simpler yet more comprehensive than today's best.

One fundamental approach to fostering sharing, cross-pollination and integration across the diverse subfields of AI is the pursuit of *architectures* that embody hypotheses about the fixed structure underlying intelligent behavior (Langley, Laird and Rogers 2009). At a minimum, architectures typically comprise mechanisms for memory, learning, interaction, and decisions. Each of these mechanisms may be simple and uniform or complex and varied. Decisions, for example, could simply be based on a preference-based or decision-theoretic choice algorithm, or may involve more complex forms of planning and reasoning. Depending on the intended architectural scope, other mechanisms may also be included, for example, in support of perception, motor control, reflection, motivations and emotions. Such architectures may be known as cognitive architectures, architectures for human-level AI or intelligent agent architectures.

When knowledge and goals are added, architectures yield behavior that is intended to model human behavior and/or yield artificially intelligent behavior. Architectures define languages and provide tools for developing intelligent systems, but they aren't simply languages or toolkits. Because architectures embody hypotheses about intelligent behavior, issues of necessity (minimality) and sufficiency (completeness) become relevant. Because these hypotheses include assumptions about how intelligent behavior is and/or should be generated – constraining both what kinds of behaviors are possible and how these behaviors can be produced – they embody claims about the scope of intelligent behavior. Because they combine a range of capabilities in support of a diversity of intelligent behaviors, they focus research on what is common across capabilities and behaviors rather than on what distinguishes them, and on how mechanisms integrate together to yield a system that is more than the sum of its parts rather than on how mechanisms may be optimized in isolation. In general, architectures must integrate together a sufficient set of high-enough level mechanisms to automatically yield appropriate behavior given goals and knowledge. They must also sufficiently constrain these mechanisms and their interactions so as to eschew inappropriately dysfunctional behavior.

Although statistical relational AI – the combination of logic (or, at least, symbolic relations) and probabilities in artificial intelligence – has been explored in the context of individual mechanisms within architectures, such as in (Iklé and Goertzel 2008), the larger potential of architectures in stimulating the exploration of statistical relational AI remains largely untapped, as does the potential impact of statistical relational AI on the evolution of architectures. Consider what might be expected from an effort to build architectures based entirely on statistical relational techniques. It would force examination of the techniques' applicability across all areas of AI and across a wide range of task domains. It would encourage understanding the commonalities and (in)compatibilities among the resulting mechanisms and applications in aid of combining them within individual architectures. It would also raise the possibility of radically new architectures that are significantly more functional than today's best.

One important step in this direction has been the development of languages and toolkits for statistical relational AI, such as Alchemy (Domingos and Dowd 2009) and BLOG (Milch et al. 2007). Such systems provide a broadly applicable, although generally low-level, functionality that can encourage exploration of statistical relational mechanisms across a wide range of intelligent capabilities and tasks. They may also facilitate developing and integrating higher-level architectural mechanisms (as in Domingos and Dowd's discussion of an *interface layer* for AI). However, they still fall short of what architectures provide in forcing explorations across the full scope of AI, in providing and constraining the requisite higher-level mechanisms and their integration, and in encouraging a focus on commonality and compatibility.

Over the past couple of years I have been rethinking architectures from the ground up based on *graphical models* (Koller and Friedman 2009). Graphical models are particularly intriguing from an architectural perspective because they can produce state-of-the-art algorithms across symbol, probability and signal processing from a single representation and inference algorithm (*summary product*). While other significant approaches to graphical inference do exist – such as sampling algorithms – even without them we get state-of-the-art algorithms such as the Rete production match algorithm (symbol processing), loopy belief propagation in Bayesian networks (probability processing), and Kalman filters and the forward-backward algorithm in hidden Markov models (signal processing). More broadly, graphical models have become the standard paradigm in both probability and signal processing, and may potentially become so in symbol processing as well.

The goals behind this rethinking have been to evaluate graphical models as a uniform *implementation level* for exploring the space of architectures, reconstructing and better understanding existing architectures, and developing new architectures that are both simpler and more functional than existing ones (Rosenbloom 2009). Progress to date has focused on the implementation of a hybrid (discrete and continuous) mixed (Boolean and Bayesian) memory architecture that provides the kinds of memories embodied in two leading cognitive architectures – ACT-R (Anderson 2007) and Soar 9 (Laird 2008) – while also going beyond them in significant ways. This work is based on factor graphs in which both factor functions and messages are represented as general N dimensional continuous functions (approximated as piecewise linear functions over rectilinear regions). The domains of these functions can be discretized and the ranges Booleanized to support discrete distributions and symbols. From this uniform base, a working memory is defined along with multiple long-term memories: classical procedural (rule) and declarative (semantic and episodic) memories plus a constraint memory. Interaction among the memories is also grounded in the shared graphical implementation level.

The ultimate goal of this work is to develop complete architectures that are simpler yet more comprehensive than today's state of the art, combining their current strengths with tightly integrated probability and signal processing, and breaking down the traditional wall between (symbolic) central cognition and (continuous) peripheral perception and motor control. This will help force the development of graphical mechanisms across AI, while adding a strong focus on their integration. Initial thought has already, for example, gone into how to extend the implemented memory architecture with capabilities for decisions, reflection, imagery, perception, theory of mind, and learning. Beyond just developing graphical approaches to these capabilities, much of the effort goes into the discovery and leveraging of commonalities across them, and understanding and working through incompatibilities among them. To mention just one example, uniformly implementing and combining together procedural and declarative memories has revealed several subtle inconsistencies requiring resolution, such as the closed world assumption embodied in rules versus the open world assumption at the heart of retrieval from semantic memory.

The focus of this rethinking has not been strictly on the combinations of logic and probability that are the norm in statistical relational AI; e.g., rules underlie symbol processing here rather than logic. Yet there is much in common, such as the centrality of graphical models in implementing mixed (and hybrid) processing. More broadly, the potential utility of an architectural approach is independent of the details of symbol processing. An architectural approach, with its emphasis on breadth and integration, should in general have much to offer the study of statistical relational AI, just as its study also appears to have much to offer architectures.

Acknowledgements. This effort has been sponsored by the USC Institute for Creative Technologies and the U.S. Army Research, Development, and Engineering Command (RDECOM). Statements and opinions expressed do not necessarily reflect the position or the policy of the United States Government, and no official endorsement should be inferred.

References

Anderson, J. R. 2007. *How Can the Human Mind Occur in the Physical Universe?* Oxford University Press.

Domingos, P. and Lowd, D. 2009. *Markov Logic: An Interface Layer for Artificial Intelligence*. Morgan & Claypool.

Iklé, M. and Goertzel, B. 2008. Probabilistic quantifier logic for general intelligence: An indefinite probabilities approach. In *Artificial General Intelligence 2008: Proceedings of the First AGI Conference*. IOS Press.

Koller, D. and Friedman, N. 2009. *Probabilistic Graphical Models: Principles and Techniques*. Cambridge, MA: MIT Press.

Laird, J. E. 2008. Extending the Soar cognitive architecture. In *Artificial General Intelligence 2008: Proceedings of the First AGI Conference*. IOS Press.

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

Milch, B. Marthi, B., Russell, S. Sontag, D., Ong, D. L., Kolobov, A. 2007. BLOG: Probabilistic models with unknown objects, in L. Getoor, B. Taskar (Eds.), *Introduction to Statistical Relational Learning*, Cambridge, MA: MIT Press.

Rosenbloom, P. S. 2009. Towards a new cognitive hourglass: Uniform implementation of cognitive architecture via factor graphs. In *Proceedings of the 9th International Conference on Cognitive Modeling*.