(A)symmetry × (Non)monotonicity: Towards a Deeper Understanding of Key Cognitive Di/Trichotomies and the Common Model of Cognition

Paul S. Rosenbloom (Rosenbloom@USC.Edu)
Institute for Creative Technologies & Department of Computer Science, University of Southern California
12015 Waterfront Drive, Playa Vista, CA 90094 USA

Abstract

A range of dichotomies from across the cognitive sciences are reduced to either (a)symmetry or (non)monotonicity. Taking the cross-product of these two elemental dichotomies then yields a deeper understanding of both two key trichotomies – based on control and content hierarchies – and the Common Model of Cognition, with results that bear on the structure of integrative cognitive architectures, models and systems, and on their commonalities, differences and gaps.

Keywords: Dichotomies; control; memory; learning; Common Model of Cognition; cognitive architectures.

Introduction

The cognitive sciences embody many dichotomies, with a broad range of work focused on either making a case for one side versus the other of individual dichotomies or on finding a hybrid approach that spans both sides. Here, the focus is on two general clouds of dichotomies – one that is fundamentally reducible to (*a*)*symmetry* and the other to (*non*)*monotonicity* – with the overall aim of understanding them better both individually and jointly. (A)symmetry concerns whether processing – whether conceived of as memory access, derivation, inference or computation – is valid in a single direction versus in arbitrary directions. (Non)monotonicity in its essence concerns whether processing accumulates results versus alters them.

These are not necessarily the most familiar formulations of either dichotomic cloud, but each fundamentally captures the nature of its own cloud in a manner that enables a simple definition and a clear path for mapping the other dichotomies from the same cloud onto it. Although such mappings may at times lose nuances, the main message concerns the commonality at their heart rather than the range of subtleties.

Once the clouds are reduced to the two elemental dichotomies, their cross product yields a 2×2 framework that enables additional analyses. It is first applied to two key cognitive trichotomies that are based, respectively, on control hierarchies – including one implicit in AlphaZero, a system that learns to best humans at challenging board games (Silver et al., 2018) – and content hierarchies. Each trichotomy spans only three of the four cells but together they span all four.

The framework is then applied to the Common Model of Cognition – an attempt to build a community consensus over the structures and processes that define a human-like mind – plus three cognitive architectures that heavily influenced its initial form (Laird, Lebiere & Rosenbloom, 2017): ACT-R (Anderson, 2007), Soar (Laird, 2012) and Sigma (Rosenbloom, Demski & Ustun, 2016). The initial focus here will be on memory and control, with results highlighting one of the major capabilities missing from the Common Model, while clarifying the distinct ways the three architectures span the di/trichotomies. This is followed by an analysis of learning that also includes AlphaZero.

The methodology here is akin in general to the one behind the Common Model – based on abstract analysis and synthesis rather than detailed experiments and models – but the goal is to provide a start at a yet deeper understanding of key parts of cognition at a yet more abstract level of analysis and synthesis. The overall structure of this paper is simple, focused on dichotomies, then trichotomies, and then the Common Model. The results suggest new ways of thinking about existing architectures, models and systems, while also highlighting key commonalities, differences, and gaps.

Dichotomies

(A)symmetry

(A)symmetry fundamentally concerns whether the processing of memory structures is valid in only one direction versus omnidirectionally. For example, consider a rule versus a logical implication. Both can be denoted by arrows, but the former only works moving forward whereas the latter works in both directions, and in fact, can even be replaced by a symmetric connective. Or, consider a feedforward neural network versus a Bayesian network. Here the former also only yields valid results moving forward whereas the latter can be used to infer values in any direction. When reverse processing does happen in asymmetric structures – whether for abduction, planning or learning – it is of a fundamentally different form than the forward processing.

In addition to rules and feedforward (including recurrent) neural networks, additional asymmetric forms include both traditional procedural programs plus more recent AI formulations such as arithmetic circuits (Darwiche, 2009) and sum-product networks (Poon & Domingos, 2011). Beyond logics and graphical models – such as Bayesian or Markov networks and factor graphs – additional symmetric forms also include constraints and Boltzmann machines.

With respect to actual dichotomies, rules versus logics (with, for example, model-based semantics) is a traditional symbolic AI one that maps directly onto (a)symmetry. In expert systems, a more abstracted variant occurs as rules versus first-principles reasoning (Davis, 1983), with the latter focusing on flexible use of small amounts of general knowledge, whether logical or not, to yield a wide variety of results that might otherwise require many rules. Abstracting this even further, but still within expert systems, yields shallow (or surface) versus deep reasoning (e.g., Hart, 1982).

Function-based versus model-based approaches – where the former may, for example, comprise feedforward neural networks or arithmetic circuits and the latter graphical models such as Bayesian networks – expresses a related dichotomy that arises in probabilistic AI (Darwiche, 2018). Likewise, within neural networks, we get the dichotomy of heteroassociative versus autoassociative networks (e.g., Rizzuto & Kahana, 2001). Feedforward networks are heteroassociative, generating outputs from inputs but not vice versa, whereas Boltzmann machines are autoassociative. It may seem jarring to view this distinction between types of neural networks in a manner akin to that between rules and logics, but that is a clear conclusion from this analysis.

In (machine) learning more broadly, we see classification versus clustering, supervised versus unsupervised learning, and discriminative versus generative learning (e.g., Ng & Jordan, 2001). The first element in each pair acquires a structure that is to be used in only one direction, whereas the second enables processing in arbitrary directions.

A dichotomy familiar in both symbolic AI and cognitive science is procedural versus declarative memory. In a classical cognitive architecture, such as ACT-R or Soar, procedural memory is based on rules and declarative memory on facts. Rules are asymmetric structures. Facts are static structures that don't themselves mandate a direction of processing. However, they do mandate a means for accessing them. Typically, this involves a mechanism for retrieving the best candidate(s) given any set of cues; a form of symmetric processing, whether as partial match, spreading activation, a holographic memory, or an autoassociative network.

It may even be that it is this symmetric processing rather than the nature of the facts themselves that defines declarative memory and distinguishes it from procedural memory; an idea worth capturing as an explicit hypothesis.

> *(A)symmetric Memory Hypothesis*: Procedural and declarative memory are fundamentally distinguished by differences in processing symmetry rather than content.

Particularly attractive about this hypothesis is how simple yet fundamental the underlying distinction is, and how it thus obviates the need for a messier attempt at distinguishing procedural versus declarative content. It also enables directly mapping varieties of neural networks (e.g., heteroassociative versus autoassociative), symbolic structures (e.g., rules versus logics), and probabilistic structures (e.g., arithmetic circuits versus Bayesian networks) onto procedural versus declarative memory, respectively.

Although a difference in (a)symmetry has long been recognized in how knowledge is retrieved from procedural versus declarative memories, the key difference here is that (a)symmetry is proposed as definitional rather than ancillary, yielding a bottom-up mechanistic definition rather than a top-down content-based one. In the process, the hypothesis has direct implications that would be difficult to derive from distinctions concerning memory content.

Given that procedural and declarative memory fully cover the (a)symmetry dichotomy, and that it appears to be a true dichotomy rather than just the endpoints of a more graduated dimension, the possibility is also raised that there is no further conceptual room for other forms of memory along this dimension. There may, however, be variations of these along other dimensions; for example, image memory may simply be a subsymbolic form of symmetric memory, and thus in a deep sense akin to declarative memory. The two may also be combined; for example, both episodic memory and analogy combine symmetric access to memory structures with subsequent asymmetric processing of the structures, via mapping or succession, respectively. One of these memories may even be used to implement or emulate the other, such as when a rule description is stored in declarative memory, retrieved and interpreted to yield procedural behavior; or when an autoencoder is implemented via a pair of feedforward networks. Still, none of this fundamentally changes the essential nature of the dichotomy.

Two additional dichotomies that are sometimes associated with procedural versus declarative memory are procedural versus declarative semantics (in AI) and implicit versus explicit representations (in cognitive science). The former concerns whether or not structures have fixed, a priori semantics, whereas the latter concerns whether or not there is awareness of the structures during processing. Declarative memory does appear to more naturally support both fixed meanings and awareness, but neither is actually inherent to it, nor does either derive directly from symmetry, so an in depth understanding of these dichotomies is left for future work.

(Non)monotonicity

(Non)monotonicity fundamentally concerns whether processing is additive, cumulative or increasing versus modifiable, retractable or reducible. For example, one of the core pieces of the Common Model is a cognitive cycle that runs at ~50 msec in humans. In Soar and Sigma this cycle is structured as a (mostly) monotonic elaboration phase during which new information is added about the current situation, followed by a nonmonotonic decision (or adaptation) phase during which the situation is actually changed.

This dichotomy also maps to a distinction in cognitive science between automatized versus controlled behavior (Schneider & Shiffrin, 1977), with monotonic processing safely allowed to proceed automatically, while controlled decisions are needed to determine which nonmonotonic change to make. Taking this a step further, it maps onto the dichotomy of parallel versus serial processing, where the absence of interactions or conflicts in monotonic processing authorizes parallelism whereas the need for control and the possibility of interactions among nonmonotonic options implies a need for seriality The mapping for both of these dichotomies is not perfect, as control may be needed to limit parallelism and parallelism may be possible for noninteracting nonmonotonic components; however, the essential commonalities are again what matter here.

Aligning these last two dichotomies yields one form of processing that is autotomized and parallel, plus a second that is controlled and serial. This aggregate dichotomy clearly maps onto both the dichotomies of reactive versus deliberative behavior in cognitive control and fast (System 1) versus slow (System 2) behavior in Kahneman (2011). It has also been characterized in terms of knowledge versus search, or a bit more precisely, as knowledge (K) search versus problem space (PS) search, with the former being monotonic search over what is already known and the latter nonmonotonic problem-space search over the space of combinatoric possibilities (Newell, 1990).

A key takeaway for cognitive science from this is again worth capturing as an explicit hypothesis.

(Non)monotonic Control Hypothesis: Reactive (System 1) and deliberative (System 2) are fundamentally distinguished by differences in processing monotonicity.

Shifting from cognitive science to the cognitive sciences more broadly, and in particular to various subfields of AI, a number of additional variations on this same dichotomy can be found. In constraint solving, there is monotonic propagation (where existing constraints on some variables induce additional constraints on others) versus nonmonotonic conditioning (where hypothetical commitments are made to particular variable values) (Dechter, 2003)). In causal reasoning, the first two steps on the Ladder of Causation (Pearl & Mackenzie, 2018) are association (monotonic probabilistic reasoning) and intervention (nonmonotonic action changes). In logic, the distinction between monotonic and nonmonotonic logics depends on whether inferences made remain valid forevermore versus being retractable. Finally, in search over multimodal spaces, making monotonic moves that never decrease the current value only guarantees a local optimum whereas reaching a global optimum may require interim nonmonotonic moves to lower-valued states.

(A)symmetry × (Non)monotonicity

The cross product of these two elemental dichotomies yields the 2×2 framework outlined in Table 1. Other such cross products have previously been explored in cognitive science, such as one in ACT-R and CLARION (Sun, 2016) that spans (a)symmetry – under two different names – and (sub)symbolic. However, replacing (sub)symbolic with (non)monotonicity in the analysis yields new opportunities for a deeper understanding.

Table 1: 2×2 Framework.

	Asymmetry	*Symmetry*
Monotonicity		
Nonmonotonicity		

Of particular interest here is how this 2×2 framework structures cognitive architectures, models and systems, and how it reveals commonalities and differences among them. With one last explicit hypothesis, it also helps reveal gaps in them.

(A)symmetric×(Non)monotonic Necessity Hypothesis: General intelligence necessitates appropriate processing and learning in all four cells of the (a)symmetry × (non)monotonicity framework.

Initial evidence for this hypothesis will, in what is to come, take the form of how all four cells are required to handle both trichotomies, plus the three architectures that most influenced the initial form of the Common Model of Cognition.

Trichotomies

Tri-level Control Hierarchy

The (non)monotonic dichotomy by itself provides a classic two-level control hierarchy, whether one thinks of it as reactive versus deliberative or System 1 versus System 2. However, a number of approaches go beyond this to three levels. One canonical form spans reactive (immediate response), deliberative (action sequences), and reflective (metacognition), which when mapped to the 2×2 framework bends the normal linear trichotomy into an L shape (Table 2).

Table 2: 2×2 Mapping of Tri-Level Control Hierarchy.

	Asymmetry	*Symmetry*
Monotonicity	Reactive	
Nonmonotonicity	Deliberative	Reflective

The vertical leg retains the general mapping from earlier of reactive onto monotonic and deliberative onto nonmonotonic but restricts them to the corresponding asymmetric cells. The reactive level in control hierarchies unsurprisingly focuses on procedural rather than declarative memory, due to the former's focus on control, and thus maps to the top-left cell. Declarative memory can clearly play a role in control, but this is typically ignored in control trichotomies.

At the elbow of the L is deliberative processing, consisting of a controlled action sequence that yields a single asymmetric path through situations in the world. Following the horizontal leg to the right yields reflective use of action models to explore simulated paths between arbitrary states – that is, models of situations – thus yielding the ability to search omnidirectionally in a metacognitive problem space.

Tables 3-4 show how this all works for two tri-level control hierarchies from very different contexts: a classical robot control approach (Bonasso et al., 1997); and the AlphaZero approach to board games. Although these examples are, respectively, from robotics and (neural) ML/AI, and each implements the cells in the hierarchy differently, they both fit this same trichotomic framework, as do also the three cognitive architectures that are analyzed later.

Table 3: 2×2 Mapping of the 3T Architecture.

3T Architecture	*Asymmetry*	*Symmetry*
Monotonicity	Skill Manager	
Nonmonotonicity	Sequencer	Planner

Table 4: 2×2 Mapping of AlphaZero.

AlphaZero	*Asymmetry*	*Symmetry*
Monotonicity	Neural Networks	
Nonmonotonicity	Game Moves	Monte Carlo Tree Search

Tri-level Content Hierarchy

Tri-level content hierarchies are less common than tri-level control hierarchies, but they do exist, and bear an interesting relationship to the other. One version of this can be seen in Table 5, for affective content (Ortony, Norman & Revelle, 2005). The development of this hierarchy began with a tri-level control hierarchy, but then the distinct nature of the emotional content at each level was identified. As in control, both nonmonotonic cells are filled, but with emotional content. The larger difference, however, is that the monotonic level is now symmetric rather than asymmetric, corresponding to declarative rather than procedural memory.

Table 5: 2×2 Mapping of the Affect Hierarchy.

Affect	*Asymmetry*	*Symmetry*
Monotonicity		Proto-Affect
Nonmonotonicity	Primitive Emotions	Cognitively Elaborated Emotions

Another tri-level content hierarchy, but from AI, is the Ladder of Causation mentioned earlier (Table 6). The tri-level content hierarchy here includes Bayesian reasoning (association level), reasoning about actions (intervention level), and hypothetical, or metacognitive, reasoning (counterfactual level). One major point of Pearl's work is that causal reasoning isn't all just (monotonic) Bayesian.

Table 6: 2×2 Mapping of the Ladder of Causation.

Causality	*Asymmetry*	*Symmetry*
Monotonicity		Association
Nonmonotonicity	Intervention	Counterfactuals

The asymmetric monotonic cell, where procedural memory resides, is unsurprisingly blank in both of these content hierarchies. As with the corresponding gap in control hierarchies, the missing memory could be used, but it is at best of secondary importance, and thus not typically a focus.

Common Model of Cognition

The Common Model of Cognition is being developed as an evolving community consensus concerning the structures and processes that yield human-like minds, in service of creating a cumulative reference point for the field while guiding efforts to both extend and break it. The question of interest here is to what extent the 2×2 framework can help to better understand the Common Model. The first step involves a mapping of its memory and control aspects (Table 7), followed by corresponding mappings of ACT-R, Soar and Sigma (Tables 8-10). Learning is then mapped, with AlphaZero added to the mix for this analysis.

Like the earlier trichotomies, the Common Model is incomplete, spanning only three of the framework's cells. However, in contrast to the two trichotomies, the Common Model spans both monotonic cells while omitting a metacognitive, or reflective, capability in the symmetric nonmonotonic cell. This lack. However, reflects that a consensus is needed rather than that there is a consensus against such a capability (Kralik, et al., 2018).

Table 7: 2×2 Mapping of the Common Model.

Common Model	*Asymmetry*	*Symmetry*
Monotonicity	Procedural	Declarative
Nonmonotonicity	Action Selection & Execution	

Three Cognitive Architectures

The Common Model, as a partial consensus over cognitive architectures, lacks aspects such as metacognition that may exist in the architectures from which it is derived. So, as a follow up step, it is useful to extend this analysis to the three architectures that heavily influenced its initial development – ACT-R, Soar and Sigma (Tables 8-10) – each of which includes some form of metacognition.

Table 8: 2×2 Mapping of ACT-R.

ACT-R	*Asymmetry*	*Symmetry*
Monotonicity	Rule Match	Facts
Nonmonotonicity	Selection & Execution	Imaginal Buffer

Table 9: 2×2 Mapping of Soar.

Soar	*Asymmetry*	*Symmetry*
Monotonicity	Parallel Rule System	Facts & Episodes
Nonmonotonicity	Selection & Execution	Reflection

Table 10: 2×2 Mapping of Sigma.

Sigma	*Asymmetry*	*Symmetry*
Monotonicity	Asymmetric Graphs	Graphical Models
Nonmonotonicity	Selection & Execution	Reflection

In conjunction with Table 7, these mappings show how the three architectures fulfill the Common Model's requirements for its three cells, and fill in its blank cell, while highlighting the diverse ways they implement such capabilities.

All three architectures support rules in procedural memory, but in ACT-R it is only their match process that is monotonic, with a rule then being selected to yield a nonmonotonic action execution. Soar matches and fires its rules in parallel, making the whole rule system – but not final action (or operator) selection – part of procedural memory. Sigma uses a unidirectional extension of its graphical models that subsumes not only parallel rules but also feedforward neural networks (Rosenbloom, Demski & Ustun, 2017) and sum-product networks (Joshi, Rosenbloom & Ustun, 2018); with action/operator selection also separated out.

One implication of this analysis of procedural memory is that the 2×2 framework may draw boundaries that are somewhat askew from those found in standard analyses. The approach here splits off nonmonotonic aspects that would traditionally be considered part of procedural memory and includes them instead as part of action selection and execution. Whether this is ultimately the correct view remains to be seen, but either way, such boundary shifts are an important part of what falls out of these analyses.

In declarative memory, all three architectures can represent facts, although Soar decomposes this general memory capability into distinct semantic and episodic memories, and Sigma's factor graphs provide a broader range of possibilities that includes not only these two but also other forms of hard and soft constraints. All three architectures also support selecting the best partial match from declarative memory, but this does involve asymmetric and nonmonotonic processing.

Soar and Sigma are similar in the nonmonotonic layer, due to Sigma's approach being based on Soar's, with the asymmetric cell being action related and impasse-driven reflection providing the ability to leverage models for search within the symmetric cell. However, Sigma's selection process for declarative memory shares much with its procedural selection, whereas in Soar they are distinct, including an asymmetrical form of spreading activation. In ACT-R, rule selection and action execution provide its asymmetric component, while its symmetric component is based on an imaginal buffer that can represent hypotheticals.

Learning

Table 11 shows an abstract mapping of forms of learning that blends terms from the tri-level control hierarchy and the Common Model. Combining this with an extension, to all four cells, of the Common Model's notion that structure and parameter learning are needed in both procedural and declarative memory, we can jointly analyze learning in the Common Model, ACT-R, Soar, Sigma and AlphaZero to better understand its overall structure, how the approaches compare and contrast, and what gaps may show up in them.

Table 11: 2×2 Mapping of Learning.

Learning	*Asymmetry*	*Symmetry*
Monotonicity	Procedural	Declarative
Nonmonotonicity	Deliberative	Reflective

Procedural – i.e., asymmetric monotonic – learning includes rule creation via composition/chunking (Common Model, ACT-R and Soar) and parameter learning via backpropagation (Sigma and AlphaZero). None of these models/systems are thus complete with respect to procedural learning. The Common Model is described in a way that appears to be complete, but that is due to considering reinforcement learning (RL) – which learns to select actions from experience with action sequences – as procedural. But, by the analysis here, RL is an asymmetric nonmonotonic form of learning, and thus belongs instead in that cell.

On the positive side, by including RL, all five models/systems do thus span asymmetric nonmonotonic parameter learning. None of them, however, learns new primitive actions, although Soar at least learns new high-level actions by combining primitive ones (Mohan & Laird, 2014).

For declarative – i.e., symmetric monotonic – learning, the Common Model acquires facts and the quantitative metadata that facilitates their use. Both ACT-R and Soar directly implement such a combination. In Sigma, facts are instances of predicates with typed arguments. The only actual structure learning at present is type extension, whereas quantitative metadata is learned via Hebbian-style symmetric learning. Adding facts to declarative memory occurs not by structure learning but by raising probabilities above 0. AlphaZero has no declarative memory, and thus no role for its learning.

Symmetric nonmonotonic, or reflective, learning can be thought of as the acquisition of models and their parameters. The Common Model does not include these forms of learning due to its general lack of metacognition, even though all three of the architectures mentioned do embody some form of it. AlphaZero uses action models in model-based RL, but it does not appear to learn these models.

Conclusion

The first step in this paper was to reduce two distinct clouds of dichotomies down to simple computational forms. In the process it was hypothesized that the distinction between procedural and declarative memory – along with many other dichotomies (Table 12, left column) – can be grounded in the more elemental terms of (a)symmetry. The possibility was even raised that although there may be other variants or combinations of these two basic types of memory, there may be no further basic types along this dimension. It was then also hypothesized that the distinction between reactive and deliberative behavior can be grounded in the more elemental terms of (non)monotonicity, also along with many other dichotomies (Table 12, right column).

The cross product of these dichotomies yields a 2×2 framework that enables analyzing two key trichotomies and the Common Model of Cognition, providing a common means for understanding and comparing across divergent integrations of cognitive capabilities. It also identifies gaps, when accompanied by a hypothesis relating to the processing and learning that is necessary in all four cells. It further helps understand how apparently ad hoc but highly successful systems such as AlphaZero can fit within the same coherent framework for memory, control and learning as more traditional cognitive architectures, models and systems.

Table 12: Summary of Dichotomy Mappings.

(A)symmetry	*(Non)monotonicity*
Rules vs. Logic	Elaboration vs. Decision
Rules vs. First Principles	Automatized vs. Controlled
Shallow vs. Deep	Parallel vs. Serial
Function vs. Model	Fast vs. Slow
Hetero. vs. Autoassociative	System 1 vs. System 2
Classification vs. Clustering	K vs. PS Search
Supervised vs. Unsupervised	Propagation vs. Conditioning
Discriminative vs. Generative	Association vs. Intervention
Proc. vs. Decl. Memory	Mon. vs. Nonmon. Logic
	Local vs. global

In future work, this analysis needs to be extended to more systems and architectures, to more precise mappings onto the framework, and to a deeper level of understanding of the full dichotomic clouds. A complete analysis of cognition should also ultimately provide a coherent story over all relevant dichotomies and their combinations. Additional dichotomies of relevance may include discrete versus continuous, central versus peripheral, explicit versus implicit, symbolic versus subsymbolic, conscious versus subconscious, and short-term versus long-term. Additional combinations of dichotomies will also be of central importance; possibly even eventually up to a full combination of all relevant, distinct dichotomies.

Acknowledgments

The work described in this article was sponsored by the U.S. Army. Statements and opinions expressed may not reflect the position or policy of the United States Government, and no official endorsement should be inferred.

References

Anderson, J. R. 2007. *How Can the Human Mind Exist in the Physical Universe?* New York: Oxford Univ. Press.

Bonasso, R. P., Firby, R. J., Gat, E., Kortenkamp, D., Miller, D. & Slack, M. (1997). Experiences with an architecture for intelligent reactive agents. *Journal of Experimental and Theoretical Artificial Intelligence*, *9*, 237-256.

Darwiche, A. (2009). *Modeling and reasoning with Bayesian networks.* New York: Cambridge University Press.

Darwiche, A. (2018). Human-level intelligence or animal-like abilities? *Communications of the ACM*, *61*, 56-67.

Davis, R. (1983). Reasoning from first principles in electronic troubleshooting. *International Journal of Man-Machine Studies*, *19*, 403-423.

Dechter, R. (2003). *Constraint Processing*. San Francisco, CA: Morgan Kaufmann.

Hart, P. (1982). Directions for AI in the eighties. *ACM SIGART Bulletin*, *79*, 11-16.

Joshi, H., Rosenbloom, P. S. & Ustun, V. (2018). Exact, tractable inference in the Sigma cognitive architecture via sum-product networks. *Advances in Cognitive Systems*, *7*, 31-47.

Kahneman, D. (2011). *Thinking Fast and Slow*. New York: Farrar, Straus and Giroux.

Kralik, J. D., Lee, J. H., Rosenbloom, P. S., Jackson, P. C., Epstein, S. L., Romero, O. J., Sanz, R., Larue, O., Schmidtke, H. R., Lee, S. W. & McGreggor, K. (2018). Metacognition for a Common Model of Cognition. *Procedia Computer Science*, *145*, 740-746.

Laird, J. E. 2012. *The Soar Cognitive Architecture*. Cambridge, MA: MIT Press.

Laird, J. E., Lebiere, C. & Rosenbloom, P. S. (2017). A Standard Model of the Mind: Toward a Common Computational Framework across Artificial Intelligence, Cognitive Science, Neuroscience, and Robotics. *AI Magazine*, *38*, 13-26.

Mohan, S. & Laird, J. E. (2014). Learning goal-oriented hierarchical tasks from situated interactive instruction. *Proc. 28th AAAI Conf.* (pp. 113–130).

Newell, A. (1990). *Unified Theories of Cognition*. Cambridge, MA: Harvard University Press.

Ng, A. Y. & Jordan, M. I. (2001). On discriminative vs. generative classifiers: A comparison of logistic regression and naive Bayes. *Proc. 14th International Conf. on Neural Information Processing Systems* (pp. 841-848).

Ortony, A., Norman, D.A. & Revelle, W. (2005). Affect and proto-affect in effective functioning. In J.M. Fellous & M.A. Arbib (Eds.), *Who Needs Emotions? The Brain Meets the Machine*. New York: Oxford Univ. Press.

Pearl, J. & Mackenzie, D. (2018). *The Book of Why: The New Science of Cause and Effect*. New York: Basic Books.

Poon, H., & Domingos, P. (2011). Sum-product networks: A new deep architecture. *Proc. IEEE International Conf. on Computer Vision Workshops* (pp. 689-690).

Rizzuto, D. S. & Kahana, M. J. (2001). An Autoassociative Neural Network Model of Paired-Associate Learning. *Neural Computation*, *13*, 2075–2092.

Rosenbloom, P. S., Demski, A. & Ustun, V. (2016). The Sigma cognitive architecture and system: Towards functionally elegant grand unification. *Journal of Artificial General Intelligence*, *7*, 1-103.

Rosenbloom, P. S., Demski, A. & Ustun, V. (2017). Toward a neural-symbolic Sigma: Introducing neural network learning. *Proc. 15th Annual Meeting of the International Conference on Cognitive Modeling*.

Schneider, W. & R. M. Shiffrin. (1977). Controlled and automatic human information processing: 1. Detection, search, and attention. *Psychological Review*, *84*, 1-66.

Silver, D., Hubert, J., Anonoglou, I., Lai, M., Guez, A., Lanctot, M., Sifre, L., Kumaran, D., Graepel, T., Lillicrap, T., Simonyan, K. & Hassabis, D. (2018). A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. *Science*, *363*, 1140-1144.

Sun, R. (2016). *Anatomy of the Mind: Exploring Psychological Mechanisms and Processes with the Clarion Cognitive Architecture*. New York: Oxford Univ. Press.