

# Integrated Connectionist Models: Building AI Systems On Subsymbolic Foundations

Risto Miikkulainen  
Department of Computer Sciences  
The University of Texas at Austin  
risto@cs.utexas.edu

Symbolic artificial intelligence is motivated by the hypothesis that symbol manipulation is both necessary and sufficient for intelligence [7]. In symbolic systems, knowledge is encoded in terms of explicit symbolic structures, and inferences are based on handcrafted rules that sequentially manipulate these structures. Such systems have been quite successful, for example, in modeling in-depth natural language processing, episodic memory, and symbolic problem solving. However, much of the inferencing for everyday natural language understanding appears to take place immediately, without conscious control, apparently based on associations with past experience. This type of reasoning is difficult to model in the symbolic framework.

In contrast, subsymbolic (distributed connectionist) networks represent knowledge in terms of correlations, coded in the weights of the network. For a given input, the network computes the most likely answer given its past experience. A number of human-like information processing properties such as learning from examples, context sensitivity, generalization, robustness of behavior, and intuitive reasoning emerge automatically in subsymbolic systems. The major motivation for subsymbolic AI, therefore, is to give a better account for cognitive phenomena that are statistical, or intuitive, in nature.

It is not immediately obvious that large-scale AI systems can be built from distributed connectionist networks. Most of the research in subsymbolic AI so far has concentrated on isolated, small, low-level tasks, and relied heavily on pre- and postprocessed data (see e.g. [8] for an overview). In many cases, the problem is reduced to learning a simple mapping. However, even if subsymbolic AI is based on a new foundation, as can be argued, the symbolic and subsymbolic paradigms eventually have to address the same basic high-level issues. Many of the results obtained through symbolic research are still valid, and could be used as a guideline for developing subsymbolic models of cognitive processes.

One example of this approach is the DISCERN system [5], a distributed connectionist model of script-based story understanding. DISCERN is purely a subsymbolic model, but at the high level it consists of modules and information structures similar to those of symbolic systems, such as scripts, lexicon, and episodic memory. Moreover, DISCERN is an integrated connectionist architecture. Independent subsymbolic models of the various subtasks are brought together into a single system capable of performing the high-level natural language processing task. The component models are designed as independent cognitive components that by themselves account for interesting perception, language-processing, reasoning, and memory phenomena. Combining these models into a single, working system is one way of validating them. In an integrated system, the components are not just models of isolated language processing phenomena; they are shown to be sufficient constituents for generating complex high-level behavior.

The integrated connectionist approach primarily aims at building complete systems that perform well in high-level tasks, and in this sense, the approach is very similar to traditional artificial intelligence. In addition, integrated models aim at showing how certain parts of human cognition could actually be put together. As a tool for building practical natural language processing systems, however, there is a long way to go before integrated connectionist modeling will rival the traditional symbolic techniques. Subsymbolic systems are very good at dealing with regularities and combining large amounts of simple pieces of evidence, but they do not easily lend themselves to processing complex knowledge structures and unusual and novel situations that often arise in the real world.

In designing subsymbolic models that would scale up, we are faced with two main research problems: (1) how to represent and learn abstractions, and (2) how to implement control of complex processing strategies. Distributed neural networks are currently used as

statistical pattern transformers. They are trained to compute smooth functions between patterns, and can only interpolate between the patterns on which they were trained. However, processing novel input would require that the network represents meta-level information about the structure and relations of the data. Without such information, the networks are limited to processing different versions of familiar I/O situations. Also, subsymbolic systems currently do not have representations concerning the nature of their own representations and processes. As a result, they cannot employ high-level strategies to control their processing, and their behavior is limited to a series of reflex responses.

While solutions to these issues are being worked out, it is possible to build hybrid symbolic/connectionist systems that attempt to combine the advantages of both approaches. Several such systems already exist, and have been quite successful, for example, in robust speech understanding [3], symbol grounding [2], noun-phrase analysis [10], and story understanding [1,4,9]. Through the hybrid approach it is possible to gain understanding in the interactions between symbolic and subsymbolic processes, and gradually work towards more robust, human-like integrated connectionist models.

## References

- [1] Dolan, C. P. (1989). *Tensor Manipulation Networks: Connectionist and Symbolic Approaches to Comprehension, Learning and Planning*. Doctoral thesis, Department of Computer Science, University of California, Los Angeles.
- [2] Dyer, M. G and Nenov, V. I. (1993). Learning Language via Perceptual/Motor Experiences. In *Proceedings of the 15th Annual Conference of the Cognitive Science Society*, 400-405. Hillsdale, NJ: Erlbaum.
- [3] Jain, J. N (1991). *PARSEC: A Connectionist Learning Architecture for Parsing Spoken Language*. Doctoral thesis, Department of Computer Science, Carnegie Mellon University.
- [4] Lee, G. (1991). *Distributed Semantic Representations for Goal/Plan Analysis of Narratives in a Connectionist Architecture*. Doctoral thesis, Department of Computer Science, University of California, Los Angeles.
- [5] Miiikkulainen, R. (1993). *Subsymbolic Natural Language Processing: An Integrated Model of Scripts, Lexicon, and Memory*. Cambridge, MA: MIT Press.
- [6] Nenov, V. I. (1991). *Perceptually Grounded Language Acquisition: A Neural/Procedural Hybrid Model*. Doctoral thesis, Department of Computer Science, University of California, Los Angeles.
- [7] Newell, A. (1980). Physical symbol systems. *Cognitive Science*, 4:135-183.
- [8] Rumelhart, D. E., McClelland, J. L., and the PDP Research Group (1986). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Cambridge, MA: MIT Press.
- [9] Sharkey, N. E. (1989). A PDP Learning Approach to Natural language Understanding. In Alexander, I. (ed.) *Neural Computing Architectures: The Design of Brainlike Machines*, 92-116. Cambridge, MA: MIT Press.
- [10] Wermter, S. and Lehnert, W. G. (1989). A Hybrid Symbolic/Connectionist Model for Noun Phrase Understanding. *Connection Science* 1, 255-272.