How Do Symbols and Networks Fit Together

A Report from the Workshop on Integrating Neural and Symbolic Processes

Ron Sun and Lawrence Bookman

■ The Workshop on Integrating Neural and Symbolic Processes (the Cognitive Dimension), sponsored by the American Association for Artificial Intelligence, was held on 16 July 1992 at the San Jose Convention Center, San Jose, California. The workshop addressed the cognitive aspects of integrating neural and symbolic processes through the comparison, the categorization, and the examination of existing and new approaches. The workshop attracted a large audience from both academia and industry. The presentation of 16 papers, 3 invited talks, and a summary panel, as well as open discussions, helped to shed much new light on technical issues and future directions in this area.

There has been a great deal of research in integrating neural and symbolic processes from both cognitive and application viewpoints. The Workshop on Integrating Neural and Symbolic Processes, sponsored by the American Association for Artificial Intelligence (AAAI),[1] was intended to provide a forum for the discussion and the exchange of ideas in this area from a cognitive and connectionistic standpoint. The workshop was a part of the Twelfth National Conference on Artificial Intelligence held at the San Jose Convention Center, San Jose, California, on 16 July 1992. The workshop organizing committee consisted of Ron Sun (chair), Lawrence Bookman (cochair), and Shashi Shekhar.

The organizing committee made a deliberate effort to focus the workshop on the cognitive aspects of integrating neural and symbolic process-es. Integration of connectionist networks and symbolic processing has been attracting interest for a long time, as attested to by the many similar conferences, symposia, and workshops in previous years. However, to date, relatively little effort has been made to compare, categorize, and combine these fairly isolated approaches, especially from a cognitive perspective.

This workshop was the first gathering that specifically addressed the cognitive aspects of this integration: Workshop organizers were concerned more with cognitive architectures, cognitive plausibilities, and new cognitive insights than with either isolated techniques or application-system development. Here, the conception of integration leans more toward integrating symbolic processing capabilities into connectionist network models than toward juxtaposing symbolic codes with neural networks because the former approach is more interesting to cognitive-modeling research.

The workshop set out to address the following questions:

What types of problems are integrated systems suitable for?

What are the problems, difficulties, and outstanding issues in integrating neural and symbolic processes?

What are the relative advantages and disadvantages of each approach to integration?

How cognitively plausible is each approach?

How do we synthesize these existing approaches?

What are the appropriate representational techniques for various tasks?

How do representation and learning interact in integrated systems?

In short, workshop organizers wanted to further the understanding of cognitive and computational architectures for combining symbolic and subsymbolic (neural network–style) processes. In so doing, we needed to look at specific proposed architectures, their cognitive plausibility, and their strengths and weaknesses; such examination can provide the basis for a synthesis of existing divergent approaches as well as insight for further research in this area.

Different Architectural Approaches

The workshop included presentations using different architectural approaches: First, the localist approach is similar to the parallel symbolic system approach, implementing symbolic structures in a (connectionist) network fashion with each node in the network representing a concept. Second, the distributed approach is connectionism in its purest form. Some connectionists believe that simple networks, such as BACKPROP networks, can perform the functional equivalent of symbolic processing (at least to a certain extent), albeit in a holistic way. Third, the combined approach is a juxtaposition of the two types of systems as separate modules. There are a variety of ways to couple these modules: for example, loosely coupled fashion, communicating through an interface; tightly coupled fashion, with a number of different channels for communication; or completely integrated. Fourth, other approaches are also possible, such as incorporating neural networks into symbolic architectures.

Localist Architectures

Chris Lacher presented an approach to reengineering expert systems into

connectionist networks. Such networks (*expert networks*) open the possibility of applying connectionist learning methods to expert systems. Lacher and his colleagues adapted supervised connectionist learning algorithms for use in expert networks. He proposed two possible ways of developing an expert network: top-down engineering from the symbolic to the subsymbolic, such as converting an expert system to a connectionist network, and bottom-up self-organization from the subsymbolic to the symbolic through the use of connectionist learning algorithms. This combined approach is an attempt to deal with the problems associated with building large-scale systems.

Gadi Pinkas briefly described how to implement complete first-order predicate logic in a connectionist network based on the use of energy minimization. Given a logic theory, a symmetric network can be constructed that can search for a proof (of no more than *k* steps) of any query with a normal settling process.

Trent Lange presented a structured localist connectionist network for dynamic inferencing (for natural language understanding). He discussed issues in controlling the activation and inferencing processes in such a network to avoid crosstalk.

Together, these presentations demonstrated the logical or rule-based reasoning capabilities of connectionist models and explored issues in parallel (connectionist) implementations of such capabilities.

Distributed Architectures

Ronald Sumida discussed a distributed multiple-module connectionist network that integrates symbolic and PDP features: Individual PDP networks are used to represent concepts and their associated roles, with each instance (filler) of a concept or a role represented by a set of distinctive activation patterns. Because of the distributed nature of each network, a filter is used to direct the flow of information in ways that are suitable for natural language–processing tasks. This form of integration helps to overcome the limited representa-

tion capability of structured knowledge usually found in PDP systems through the encoding of specific knowledge that is necessary for ambiguity resolution and dynamic inferencing.

Andreas Stolcke and Dekai Wu explained how tree matching can be done with a recursive distributed representation. They believe that recursive distributed representations are important for bridging the gap between connectionist models and higher-level cognitive functions, and they hope to extend this framework by investigating how complex tasks (such as tree unification) can be performed.

These presentations demonstrated how distributed connectionist networks can be used to perform certain aspects of symbolic processing, although most tasks involved are relatively simple and limited.

Combined Localist and Distributed Architectures

Lawrence Bookman presented a two-tier framework for representing semantic memory (which was used to deal with text comprehension): (1) a relational tier that provides for the systematic connections between concepts and their case roles and (2) an analog semantic feature (ASF) tier that encodes the background knowledge associated with these concepts. Bookman argued that the background frame details (corresponding to the encoding of a concept through ASFs) are nonsystematic in nature and are best expressed as a set of statistical associations and implemented as a distributed network. This integration of local and distributed knowledge provides a model for both coarse- and fine-grained views of comprehension.

Ron Sun presented a two-level dual-representation framework that contains both localist and distributed representations. The localist network performs rule-based reasoning, which is essential to commonsense reasoning tasks, and the distributed network encodes similarities with the feature-based representation. Based on the combination of similarity-based reasoning and rule-based rea-

soning, many difficult patterns in commonsense reasoning emerge without being explicitly put into it, which demonstrates the utility of this system.

Furthermore, to argue generically for such an architectural approach, four requirements were established in Sun's work to narrow the choice: (1) direct accessibility of concepts, (2) direct inaccessibility of similarity matching, (3) linkages from concepts to features, and (4) linkages from features to concepts. From these four requirements, it is natural to devise this two-level architecture, which also suggests that such a framework can be extended or altered for other cognitive tasks and can serve as a basis for building more flexible and powerful intelligent systems. Work in this area can be viewed as exploring the synergy between different types of components in an integrated system to better deal with cognitive problems. It is clear from these presentations that there is still a long way to go to really understand various alternatives in integrating different representations.

Other Architectures

The connectionist building block approach basically adopts a symbolic architecture, such as a semantic network or a parallel production system, but instead of using symbolic components, neural networks are used in their place to obtain adaptability and partial match capability. Lacher's work and Sumida's work, discussed earlier, also embody this approach. In contrast, another architectural approach is exemplified by Stefan Wermter's architecture, which has symbolic programs and connectionist networks as separate components.

Learning and Representation

Representation, learning, and their interaction represent some of the major issues for developing symbolic processing connectionist networks. Connectionist networks designed for symbolic processing often involve complex internal structures consisting of multiple components and sev-

eral different representations. Thus, learning is more difficult; there is a need to address what type of representation to adopt, how the representational structure in such systems is built, how the learning processes involved affect the representation acquired, and how the representational constraints might facilitate or hamper learning.

Several presentations addressed the question of representation. John Barnden discussed alternatives in belief (propositional attitude) representation, ranging from metalogical approaches to more general cases of metarepresentational approaches. Having examined the advantages and pitfalls of each approach, he favored the metalinguistic approach, in which the representational objects are natural language sentences or utterances.

J. G. Wallace discussed representation in terms of semantic transparency. He proposed the idea of utilizing brain-monitoring data to determine semantic transparency of representations in various cognitive tasks. Such experimental results have a clear relevance to building more powerful and more complete intelligent systems that use hybrid representations in a cognitively realistic way. In terms of theoretical studies of distributed connectionist representation, Noel Sharkey reviewed some experiments to show the functional roles of constituent structures in distributed connectionist representations. He argued that distributed representations can do more than what is typically attributed to them (see the previous section). He then stated that such representations should play a big part in building cognitive systems.

Some of these theoretical threads can clearly be traced to the works of Smolensky (1988), who argues for the dichotomy of conceptual and subconceptual processing, and Dreyfus and Dreyfus (1987), who put forward the distinction between analytic thinking and intuitive thinking. These two pieces of work lay the foundation for many of the presentations and the discussions at the workshop.

Criticisms, Directions, and Common Themes

The focus of the workshop presentations was various architectural approaches to the integration of symbolic and neural processes, although several of the talks focused more on the engineering aspects of the problem than on the cognitive aspects. Jack Gelfond suggested that any proposed model should be able to explain or account for its behavior along the following cognitive dimensions: (1) frequency—one does better with things that one does often; (2) context priming—the patterns processed earlier have considerable influ-

ence on subsequent patterns; (3) the short-term versus long-term distinction—separate structures are needed for short- and long-term memory; (4) reduced representations—humans can use reduced representations and still manage to understand one another; (5) automaticity—something that is done often is usually done in an abbreviated fashion; (6) attention—as humans learn, their representation shifts from explicit to implicit; and (7) learning—a system must be capable of learning from its experience.

An even more pressing set of concerns was raised by Jim Hendler. It was argued that what is important is not whether we should integrate or use a particular model or particular techniques but what we can learn from the computational model. For example, how does the model generalize? Does the model scale? In addition, it is not sufficient to argue that because the model works, one's approach is justified. Rather, one has to explain why it works. Can it predict some testable behavior that can lead to new insights?

Some critiques were given by Stuart Dreyfus (jointly with Hubert Dreyfus) of what he saw from the current approaches as potentially promising for the future. Dreyfus argued, based on Heideggarian philosophy (the work of German philosopher Heiddeggar; see Dreyfus and Dreyfus [1987]), that "the good old-fashioned AI" lacks the ability to deal with ongoing involvement with the world—what comes under the rubric as commonsense or skillful coping. He then critically analyzed three currently promising approaches as ways to address the deficiencies of "good old-fashioned AI": (1) interactive AI, (2) artificial neural networks, and (3) reinforcement learning. In interactive AI, there is an attempt to deal with purposeful action that is missing from other AI approaches. With artificial neural networks, although there has been some success, there is still the problem of generalization; that is, many instances are required, but in many learning situations, only a few instances are actually needed. Dreyfus suggested

that to learn as humans do, we might need a network as large as the brain itself. In principle, nothing is wrong with this. It is more a practical issue: These neural networks will need to deal with massive complexity and large sizes. Finally, Dreyfus pointed out that although reinforcement learning fits well with certain particular cognitive phenomena, it, too, has some problems. For example, (1) reinforcement learning cannot avoid the generalization issue because it must generalize to other situations as humans do and (2) the issue of what is salient is not dealt with.

In abstracting from all the presentations and papers, we see two differing viewpoints, which can be classified as (1) the integration of symbolic structures into connectionist architectures and (2) "connectionist to the top." In the first view, the representations and techniques from both symbolic processing models and neural network models are used in a hybrid system to tackle problems that neither model can, by itself, handle well. Such problems might include modeling cognition that requires an ability to deal with human inferencing and reasoning capabilities and with the ability to perceive. Several researchers at the workshop argued that both of these capabilities are better implemented with more integrated symbolic and neural components.

In the second view (as explicitly and eloquently argued for by Jerome Feldman and implicitly argued for in many of the workshop papers), there is no need for symbolic structures as such. Instead, one can perform complex symbolic processing using various mappings. (This approach was also exemplified in Amit Almor's presentation). Feldman argued that the L0 project at The International Computer Science Institute is an example of a task that no symbolic system could ever represent because the geometric constraints in the task cannot be represented in symbolic systems. Sharkey also argued against the traditional hybrid view in which one does the fast memory access with neural nets and leaves the other work to a symbolic system. Instead, he argued for the

expanding role of superpositional representations in hybrid systems because such representations are functionally compositional and can enable holistic structure-sensitive operations.

Concluding Remarks

This workshop indeed served its purpose by providing a forum for an exchange of ideas, methodologies, and techniques as well as for the presentation of the research undertaken by the individual authors. New light was shed, some warnings were issued, and some promising approaches were spotlighted. It is our hope that after this workshop new thinking will be produced that can advance the state of the art in AI and cognitive science.

Acknowledgments

Note

1. For more information regarding the workshop and the related issues, contact Ron Sun at the University of Alabama, Department of Computer Science, Tuscaloosa, AL 35487, e-mail: rsun@athos.cs.ua.edu.

References

Dreyfus, H., and Dreyfus, S. 1987. *Mind over Machine*. New York, N.Y.: The Free Press.

Smolensky, P. 1988. On the Proper Treatment of Connectionism. *Behavioral and Brain Sciences* 11(1): 1–74.

Ron Sun is currently an assistant professor of computer science at the University of Alabama. His research interest centers on the studies of intelligence and cognition, especially in the areas of reasoning, learning, and connectionist models.

Lawrence A. Bookman is currently a research scientist at Sun Microsystems Laboratories in Billerica, Massachusetts. His interests include computational models for understanding text and connectionist and neural models of cognition.