# The Present and the Future of Hybrid Neural Symbolic Systems
## Some Reflections from the NIPS Workshop

*Stefan Wermter and Ron Sun*

In this article, we describe some results concerning hybrid neural symbolic systems based on a workshop on hybrid neural symbolic integration. The Neural Information Processing Systems (NIPS) workshop on hybrid neural symbolic integration, organized by Stefan Wermter and Ron Sun, was held on 4 to 5 December 1998 (right after the NIPS main conference). In this well-attended workshop, 27 papers were presented, among them were 8 were invited talks in this research area. Authors of invited papers included Gary Cottrell, Joachim Diederich, Jerry Feldman, Lee Giles, Risto Miikkulainen, Noel Sharkey, Lokendra Shastri, Hava Siegelman, David Waltz. Overall, the workshop was wide ranging in scope, covering the essential aspects and strands of hybrid systems research, and successfully addressed many important issues of hybrid system research.

Two panels were also presented. The panel entitled "Issues of Representation in Hybrid Models" was chaired by Sun. It focused on the following issues related to the two different types of representation: How does neural representation contribute to the success of hybrid systems? How does symbolic representation supplement neural representation? How can the two types of representation be combined? How we can use their interaction and synergy? The panel entitled "Hybrid and Neural Systems for the Future" was chaired by

Wermter. The following issues were covered: How can we develop neural and hybrid systems for new domains? What are the strengths and weaknesses of hybrid neural techniques? Are current principles and methodologies in neural and hybrid systems useful? How can they be extended? What will be the impact of hybrid and neural techniques in the future (internet communication; web searching; data mining; integrating image, speech, and language; cognitive neuroscience)?

---

In this article, we describe some recent results and trends concerning hybrid neural symbolic systems based on a recent workshop on hybrid neural symbolic integration. The Neural Information Processing Systems (NIPS) workshop on hybrid neural symbolic integration, organized by Stefan Wermter and Ron Sun, was held on 4 to 5 December 1998 in Breckenridge, Colorado.

---

## A Taxonomy of Architectures

Various classification schemes of hybrid systems were proposed (for example, Sun and Bookman [1994], Wermter [1995]). We present a simplified taxonomy, which reflects the work presented at the workshop and the area of hybrid systems as a whole. Let us divide hybrid systems into two

broad categories: (1) single-module architectures and (2) multimodule architectures (table 1).

For single-module systems, along the representation dimension, there can be the following types of representation (Sun and Bookman 1994): symbolic (as in conventional symbolic models), localist (with one distinct node representing each concept; for example, as presented by Shastri), and distributed (with a set of nonexclusive, overlapping nodes for representing each concept). Usually, it is easier to incorporate prior knowledge into localist models because the structure of the models can be made to directly correspond with that of symbolic knowledge. However, neural learning usually leads to distributed representation, such as in the case of backpropagation. Distributed representation has some unique and useful properties.

Heterogeneous multimodule systems are more interesting. A variety of distinctions can be made. For example, a distinction can be made in terms of representations of constituent modules. In multimodule systems, there can be different combinations of different types of constituent module; for example, a system can be a combination of localist modules and distributed modules (for example, CONSY-DERR, as described by Sun [1994]). Furthermore, it can be a combination of symbolic modules and neural modules (for example, as described by Wermter or by Kraetzschmar, Sablatnoeg, Enderle, and Palm).

## A Taxonomy of Learning

In terms of what is being learned, we can have learning contents (given a fixed architecture) or learning architectures themselves, or we can learn both at the same time. Although most hybrid learning models fall within the two first categories (as described by Frasconi, Gori, and Sperduti and by Mayberry and Miikkulainen), there are some hybrid models that belong to the third category.

In terms of the relation between symbolic and neural knowledge during learning, we have the following possibilities: (1) purely neural learning of symbolic knowledge (for example,

| 1. Homogeneous Single Module | |
| --- | --- |
| Representation | symbolic, localist, distributed |
| Mapping | direct translational, transformational |
| | |
| 2. Heterogeneous Multimodule | |
| Components | localist + distributed, symbolic + neural |
| Coupling | loosely coupled, tightly coupled |
| Granularity | coarse grained, Fine grained |

*Table 1. Classifications of Hybrid Systems.*

the presentations by Mayberry and Miikkulainen and by Morris, Cottrell, and Elman); (2) extraction of symbolic knowledge from neural networks; (3) insertion of symbolic knowledge; or (4) parallel neural and symbolic learning, which includes the use of two separate neural-symbolic algorithms simultaneously, the use of two separate algorithms in succession, or integrated neural-symbolic learning.

## What Does Symbolic and Neural Representation Contribute?

The panel discussions on representation in hybrid models highlighted the views on multiple levels of cognitive modeling and the reductions among them. In Feldman and Bailey's talk, it was proposed that there are the following distinct levels: cognitive linguistic, computational, structured connectionist, computational biological, and biological. A condition for this vertical hybridization is that it should be possible to bridge the different levels, and the higher levels should be reduced to, or grounded in, lower levels. A top-down research methodology is advocated and examined for concepts toward a neural theory of language. Although the particulars of this approach are not universally agreed on, the workshop participants generally accepted the overall idea of multiple levels of cognitive modeling.

This same panel also discussed the issue of mixing levels. One view among connectionists argues that in modeling, there should not be a mixing of levels. That is, models should be constructed entirely of neural components; both symbolic and subsymbolic processes should be implemented in neural networks. This approach is referred to as vertical hybridization. Another view, horizontal hybridization, argues that it might be beneficial, and sometimes crucial, to mix levels, so that we can make better progress toward understanding cognition. This latter view is based on a realistic assessment of the state of the art of neural model development and the need to focus on the essential issues (such as the synergy between symbolic and subsymbolic processes), rather than nonessential details of implementation. Especially for real-world hybrid systems, for example, in speech-language analysis, horizontal approaches have been used successfully.

Representation, learning, and their interaction represent some of the major issues for developing symbol-processing neural networks. Connectionist networks designed for symbolic processing often involve complex internal structures consisting of multiple components and several different representations. Thus, learning is made more difficult. We need to address the problems of what types of representation to adopt in such systems, how the representational structure is built up, how the learning

processes involved affect the representation acquired, and how the representational constraints might facilitate or hamper learning.

## Directions for Hybrid Neural Systems

The issues described earlier are important for making progress in the theories and applications of hybrid systems. Currently, there is not yet a coherent theory of hybrid systems. There has been some preliminary early work toward a theoretical framework for neural-symbolic representations, but to date, there is still a lack of an overall theoretical framework that abstracts away from the details of particular applications, tasks, and domains. One step toward such a direction might be the research into the relationship between automata theory and neural representations.

Natural language processing has been and will continue to be an important test area for exploring hybrid neural architectures. At the workshop, David Waltz argued that language is the quintessential feature of human intelligence. Although certain learning and architectures in humans might be innate, most researchers of neural networks argue for the importance of development and environment during language learning. For example, at the workshop, it was argued that syntax is not innate; it is a process rather than a representation, and abstract categories, such as subject, can be learned bottom up. The dynamics of learning natural language is also important for designing parsers using techniques such as SRN and RAAM. SARDSRN and SARDRAAM were presented in the context of shift-reduce parsing (by Mayberry and Miikkulainen) to avoid the problem of losing constituent information associated with SRN and RAAM. Furthermore, it has been argued that compositionality and systematicity in neural networks arise from an associationistic substrate acquired from evolution.

Several other tasks were addressed. Research into improving web search engines by using learning and possibly neural networks was presented by Lee Giles. Although currently most search

engines only use fairly traditional search strategies, machine learning and neural networks could improve processing of heterogeneous unstructured multimedia data. Another important promising research area is knowledge extraction from neural networks to support text mining and information retrieval. A lot of the data are noisy and enormous in size. Therefore, inductive learning techniques from neural networks and symbolic machine-learning algorithms could be combined to generate underlying rules for such data.

A crucial task for applying neural systems, especially for applying learning distributed systems, is the design of appropriate vector representations for scaling up to real-world tasks. A simple but general model was proposed to represent words. Based on a large usenet corpora, global co-occurrence statistics are computed based on 300 million words to generate context vectors of 150 elements, illustrating how to compute continuous representations based on symbolic input. Computing with large patterns is generally seen as an important precondition to scaling up neural processing as well as binding and merging operations on such patterns. Large context vectors are also essential for learning document retrieval (as in, for example, the work of Gallant). Because of the size of the data, only linear computations are useful for full-scale information retrieval. Vector representations are often restricted to co-ocurrences rather than syntax, discourse, logic, and so on. Complex representations might be analyzed using fractal approaches.

Hard, real-world applications are important. A system was built for foreign exchange-rate prediction that uses a self-organizing map for reduction and generates a symbolic representation as input for a recurrent network that can produce rules. Other related applications include Bayesian ensemble classifier models applied to credit scoring. Another self-organizing approach for symbol processing was described for classifying usenet texts and presenting the classification as a hierarchical two-dimensional map. Neural network representations also have been used for vision.

*It was argued by David Waltz that in 20 years, computer power will be sufficient to match human capabilities, at least in principle, but meaning and deep understanding would still be lacking.*

Finally, there is promising progress in neuroscience. Computational neuroscience is still in its infancy, but it might be relevant to the long-term progress of hybrid symbolic neural systems. Related to that, more complex high-order neurons might be one possibility for building more powerful functions (for example, the presentation of Lipson and Siegelmann). Another way would be to focus more on global brain architectures, for example, for building more biologically inspired robots with rooted cognition (for example, the work of Sharkey and Ziemke). It was argued by David Waltz that in 20 years, computer power will be sufficient to match human capabilities, at least in principle, but meaning and deep understanding would still be lacking. Other important questions are perception, situation assessment, and action, and ultimately, language is the quintessential feature of human intelligence.

## Concluding Remarks

Work toward a theory of hybrid neural symbolic integration is needed. There is promising work toward relating automata theory with neural networks, or logics with such networks. Furthermore, the issue of representation, in particular, vector representations, has been emphasized. To tackle larger real-world tasks, as in information retrieval or large-scale classification using neural networks, an underlying vector representation is important. Vertical forms of neural-symbolic hybridization are widely used in cognitive processing, logic representation, and language. Horizontal forms of neural-symbolic hybridization exist for larger tasks, such as speech-language processing, knowledge engineering, and intelligent agents. Furthermore, computational neuroscience might offer further ideas and constraints for building more realistic forms of neural symbolic models.

### References

Sun, R. 1994. *Integrating Rules and Connectionism for Robust Commonsense Reasoning.* New York: Wiley.

Sun, R., and Bookman, L., eds. 1994. *Architectures Incorporating Neural and Symbolic Processes.* Boston: Kluwer.

Wermter, S. 1995. *Hybrid Connectionist Natural Language Processing.* London: Chapman and Hall.

**Ron Sun** is an associate professor of computer science at the University of Missouri and a consulting scientist at NEC Research Institute. He received his Ph.D. in computer science in 1991 from Brandeis University. Sun's research interest centers on the studies of intelligence and cognition, especially in the areas of commonsense reasoning, human and machine learning, and hybrid connectionist models. His web page is www.cecs.missouri.edu/~rsun. His e-mail address is rsun@cecs.missouri.edu.

**Stefan Wermter** holds a research chair in intelligent systems at the University of Sunderland, United Kingdom. He holds a Ph.D. from the University of Hamburg, Germany, in computer science. His research interests are in AI, neural networks, hybrid systems, language processing, and cognitive neuroscience, and fuzzy systems as well as the integration of speech, language, and image processing.

# $C$omputation $\&$ $I$ntelligence

### *Collected Readings*

EDITED BY

GEORGE F. LUGER

This comprehensive collection of 29 readings covers artificial intelligence from its historical roots to current research directions and practice. With its helpful critique of the selections, extensive bibliography, and clear presentation of the material, *Computation and Intelligence* will be a useful adjunct to any course in AI as well as a handy reference for professionals in the field.

# Case-Based Reasoning

## Experiences, Lessons, & Future Directions

### *Edited by David B. Leake*



Case-based reasoning is a flourishing paradigm for reasoning and learning in artificial intelligence, with major research efforts and burgeoning applications extending the frontiers of the field. This book provides an introduction for students as well as an up-to-date overview for experiences researchers and practitioners. It examines the field in a "case-based" way, through concrete examples of how key issues — including indexing and retrieval, case adaptation, evaluation, and application of CBR methods — are being addressed in the context of a range of tasks and domains.