Neurosymbolic Integration: Unified versus Hybrid Approaches

Mélanie Hilario, Yannick Lallement, Frédéric Alexandre

CUI - University of Geneva
24 rue Général-Dufour
CH-1211 Geneva 4
hilario@cui.unige.ch

CRIN-INRIA Lorraine
BP 239 - Campus scientifique
F-54506 Vandœuvre-les-Nancy Cedex
lallemen@loria.fr, falex@loria.fr

Abstract. Since the mid-1980s, researchers have been pursuing the goal of neurosymbolic integration, i.e., the construction of systems capable of both symbolic and neural processing. We distinguish two major avenues toward this goal: the unified and the hybrid approaches. Whereas the unified approach claims that full symbol processing functionalities can be achieved via neural networks alone, the hybrid approach is premised on the necessity and complementarity of symbolic and neural structures and processes. This paper attempts to clarify and compare the assumptions, mechanisms as well as the open problems of both approaches.

1. Definitions

Since the resurgence of connectionist research in the mid-1980s, neurosymbolic integration (NSI)—the incorporation of symbolic and neural processing functionalities in a single system—has been a persistent research goal. Attempts at NSI can be classified into two major approaches according to the particular blend of symbolic and neural structures and processors involved.

In the *unified approach*, more widely known as connectionist symbol processing (CSP) [6], neural networks are used as building blocks to create a cognitive architecture capable of complex symbol processing. The unified approach is premised on the claim that there is no need for symbolic structures and processes as such: full symbol processing functionalities can be achieved using neural networks alone.

The *hybrid approach* integrates *complete* symbolic and connectionist modules: in addition to neural networks, it implements both symbolic structures and processors—e.g., rule interpreters, parsers, case-based reasoners and theorem provers. The hybrid approach rests on the assumption that only the synergistic combination of neural and symbolic structures and processes can attain the full gamut of cognitive and computational powers which is beyond the reach of a single paradigm.

This work was partially supported by the EU under Esprit Basic Research project MIX.

In this paper, we will focus on the basic opposition between the unified and the hybrid approaches. Sections 2. and 3. will elaborate further on each of these approaches, while Section 4. will discuss open problems and research issues.

2. The hybrid approach

Hybrid systems can be classified according to their integration mode, i.e., the way in which the symbolic and neural components are configured in relation to each other and to the overall system. We distinguish four integration modes: chainprocessing, subprocessing, metaprocessing and coprocessing.

2.1. Chainprocessing

In chainprocessing mode, one of the modules—either symbolic or neural—is the main processor while the other takes charge of pre and/or postprocessing tasks. In one example of symbolic preprocessing [9], decision trees are used as feature selectors to limit the number of input nodes in feedforward neural networks and thus accelerate the learning process. The reverse setup is neural preprocessing: in [5], for instance, the main processor is a respiratory monitoring expert system which recommends actions to be taken to avoid breathing complications. Rules are of the form: "If *qualitative-state* then *action*", where *qualitative-state* is a symbolic representation of a change in lung pressure over time (e.g., "pressure is rising rapidly"). Experts found it extremely difficult to formulate rules to determine these qualitative states, so a backpropagation network was used to accomplish this task; its output was then stored as a fact in the expert system's working memory. With neural preprocessing, development time was cut down from 3 to 2 months and accuracy rose from 74.5 % to 97.5 %.

2.2. Subprocessing

In subprocessing mode, one of the two modules is embedded in and subordinated to the other, which acts the main problem solver. In INNATE/QUALMS [2], the symbolic module is the main processor whose task is fault diagnosis in a distillation plant. It is a classical expert system linked to a simulation model of the process to be monitored. To diagnose a process fault, the expert system subcontracts the task of generating a set of candidate faults to the neural component, composed of several multilayer perceptrons cascaded in two layers. Each output neuron is associated with a particular fault and its activation is interpreted as the rating of the corresponding fault. The expert system then either confirms the neural diagnosis or proposes another hypothesis.

2.3. Metaprocessing

In metaprocessing mode, one module is the baselevel problem solver and the other plays a metalevel role (such as monitoring, control, or performance improvement) vis-à-vis the first.

Symbolic metaprocessing is illustrated in the Robotic Skill Acquisition Architecture [4]. Its goal is to develop robots which accomplish complex tasks using designer-supplied instructions and self-induced practice. The architecture's base level is itself hybrid: a rule-base provides a declarative representation of human expert knowledge whereas neural networks embody reflexive procedural knowledge that comes with practice. The metaprocessor is a rule-based execution monitor which supervises the training of the neural network and controls the operation of the system during the learning process.

An example of *neural metaprocessing* is a system in which the baselevel symbolic processor solves high school physics problems, and the connectionist component enforces search control [3]. Object-level rules encode equations such as velocity = initial-velocity + acceleration x time. If several of the needed variables are unknown, the metalevel connectionist module is called on to make a choice. This is a previously trained net whose inputs are the current goal variable as well as the known and unknown variables; it outputs the next (unknown) variable to solve for, in effect guiding the problem-solving process.

2.4. Coprocessing

In coprocessing, the symbolic and the neural modules are equal partners in the problem solving process: each can interact directly with the environment, each can transmit information to and receive information from the other. Neurosymbolic coprocessing has been applied to the improvement of arrhythmia diagnosis via intracardial defibrillators, devices implanted in people with heart disorders [7]. These devices detect abnormal rhythms or arrhythmias, which can be clustered into three groups depending on the type of action they call for: continue monitoring, pace the heart, or apply high-voltage electric shock. Classification accuracy is of course crucial in this application. In one hybrid model used, incoming signals are channelled to a decision tree which acts as a timing classifier and to a multilayer perceptron which performs morphology-based classification. The outputs of both modules are fed into an arbitrator which determines the class of the arrhythmia. On a multipatient database, the decision tree/neural network hybrid attained an accuracy rate of 99%.

3. The unified approach

Unified neurosymbolic models aim at performing all cognitive tasks using neural networks alone. The unified approach comes in two flavors, depending on where the bulk of the research effort is placed. The first is biologically oriented: its main goal is neurophysiological plausibility. It is still in its infancy, and though successes have been scored at the perceptual level, it will take some time before a frontal attack on symbol processing can be envisaged. The second, symbol processing version will be the focus of the following subsections. Here, the main emphasis is on building connectionist architectures for symbol processing. Depending on the underlying knowledge representation scheme, these architectures can be classified as localist (each node in the system corresponds to a single concept, and vice-versa), distributed (each concept is encoded using several nodes, and each node participates in the representation of several concepts), or combined localist/distributed.

3.1. Localist models

A connectionist architecture for rule-based reasoning and variable binding is described in [1]. Knowledge is represented in a localist fashion using several types of nodes: predicate nodes (for predicate names), role nodes (for predicate arguments or variables), and filler nodes (for possible role values). Rules are encoded by interconnection patterns between predicate and role nodes. Variables and constants are bound if their associated role and filler nodes fire in synchrony as activation spreads throughout the network. This idea of using synchronous or in-phase activation to represent the binding of distinct features and concepts is supported by neurophysiological evidence. However, the model has significant difficulties in scaling up to complex, high-level reasoning; more importantly, it fails to give any convincing account of the learning process.

3.2. Distributed models

BOLTZCONS [11] uses distributed representations of linked lists to implement symbolic structures like trees and stacks. One cell of a linked list is encoded as a triple of symbols of form `(tag, car, cdr)`. Since memory is assumed to be sparse, only a fraction of these triples can be present in memory at any one time. A coarse coding technique, the description of which is beyond the scope of this paper, reduces the number of units required to store a triple while adding a measure of representational redundancy. Pointer traversal is implemented via associative retrieval, and a connectionist maintenance system adds, deletes and updates memory triples. Fault tolerance and robustness are side benefits of this distributed representation: like biological networks, BOLTZCONS will not crash abruptly if some cells cease to function.

3.3. Localist/distributed models

CONSYDERR [10] is a unified system aimed at modelling common-sense reasoning (e.g., *Can a duck fly? A duck is like a sparrow, so I guess it can fly.*). It uses a dual but purely connectionist representation scheme. At the conceptual level, domain concepts are encoded as nodes in a localist network and rules are represented by links between these nodes. At the subconceptual level, a distribution network performs finer-grained and less structured processing, i.e., associations and pattern matching. Each concept known to the system is represented in the two networks. The localist network performs classical rule-based inferencing while the distributed network performs similarity-based reasoning. With these two levels, CONSYDERR can handle a number of difficult issues, such as partial or inexact information, in one integrated framework.

4. Open research issues

Unified and hybrid approaches share the same assumption: symbol processing functionalities are needed to perform high-level cognitive tasks. They also share the same goal: to overcome the well-known shortcomings of symbolic and connectionist AI models. In this section, we briefly review the strengths and weaknesses of both approaches, identify the key issues involved, and compare the results obtained so far.

The hybrid approach is based on a psychologically plausible distinction between two types of cognitive operations: automatic, reflexive or low-level (e.g., perception) *vs* controlled, deliberative or high-level (e.g., reasoning) [8]. Typically, in hybrid systems, reflexive tasks are assigned to the connectionist subsystem and deliberative tasks to the symbolic subsystem. Another advantage of the hybrid approach is that it can directly benefit from previous work in both symbolic and connectionist processing: the construction of a hybrid model will will not always have to start from scratch. However, hybrid systems do not reflect biological reality since, at a sufficiently fine grain level, the architecture of the brain is uniform.

On the contrary, the unified approach—whether in its biologically oriented or in its symbol-processing oriented version—can legitimately claim a certain degree of biological plausibility, and further progress of the approach could be triggered by new biological data. Another strength of the unified approach is its use of massively parallel processing with its dual advantage—biological fidelity as well as computational efficiency. However, unified systems inherit a recurrent problem of connectionism, namely its lack of scalability: as applications grow, neural networks rapidly reach their asymptotic performance level.

Both approaches are faced with an impressive array of research issues, of which we briefly mention three. A fundamental issue is knowledge representation: whereas the aim of the unified approach is to find new ways of representing knowledge, the hybrid approach strives to interface known connectionist and symbolic representational schemes in order to facilitate communication or knowledge-sharing between the two subsystems. A specific subproblem is the transition between opposite representation modes such as qualitative/quantitative, discrete/continuous, deterministic/statistical. Another important issue is that of reasoning power: neural networks have yet to demonstrate reasoning capacities at least equal to that of first-order predicate logic. On this point, the hybrid position is certainly more convenient, as symbolic modules can be used wherever neural nets fall short of the reasoning power required by a given task. Learning is a third key issue: the unified approach bears the burden of showing how the diversity of symbolic machine learning methods can be implemented in neural networks, whereas hybridists are faced with the task of integrating symbolic and neural learning strategies into a viable and hopefully synergistic whole. These three issues remain widely open on either side.

When we compare the results of the two approaches, it seems that the hybrid approach has led to more convincing results so far: some industrial applications are already operational while unified models are still far from achieving the representational power of symbolic systems. However, considering biological evidence, it is certain that the unified approach has not said its last word. Thus, it is not clear at the moment which approach will turn out to be the most fruitful. The hybrid approach may be viewed as a short-term engineering expedient, whereas the unified approach may be viewed as leading to fundamental breakthroughs which are, however, still out of reach.

References

[1] V. Ajjanagadde and L. Shastri. Rules and variables in neural nets. *Neural Computation*, (3):121–134, 1991.

[2] W. R. Becraft, P. L. Lee, and R. B. Newell. Integration of neural networks and expert systems for process fault diagnosis. In *Proc. of the 12 th IJCAI*, pages 832–837, 1991.

[3] Matthias Gutknecht and Rolf Pfeifer. An approach to integrating expert systems with connectionist networks. *AICOM*, 3:116–127, 1990.

[4] D.A. Handelman, S.H. Lane, and J.J. Gelfand. Robotic skill acquisition based on biological principles. In A. Kandel and G. Langholz, editors, *Hybrid Architectures for Intelligent Systems*, pages 301–327. CRC Press, 1992.

[5] S. Hayes, V.B. Ciesielski, and W. Kelly. A comparison of an expert system and a neural network for respiratory system monitoring. Technical Report TR #92/1, Royal Melbourne Institute of Technology, 1992.

[6] G. Hinton, editor. *Connectionist Symbol Processing.* MIT/Elsevier, 1991.

[7] M. Jabri, S. Pickard, P. Leong, Z. Chi, B. Flower, and Y. Xie. ANN-based classification for heart difibrillators. In *Advances in Neural Information Processing 4*, pages 637–644. Morgan-Kaufmann, 1992.

[8] A. Newell. *Unified Theories of Cognition.* Harvard University Press, Cambridge, Mass., 1990.

[9] S. Piramuthu and M.J. Shaw. On using decision tree as feature selector for feed-forward neural networks. In *International Symposium on Integrating Knowledge and Neural Heuristics*, pages 67–74, 1994.

[10] R. Sun. *Integrating Rules and Connectionism for Robust Reasoning : A Connectionist Architecture with Dual Representation.* PhD thesis, Brandeis University, Waltham, MA 02254, 1991.

[11] D. S. Touretzky. Boltzcons : Dynamic symbol structures in a connectionist network. *Artificial Intelligence*, 46(1-2), 1990.