4

COGNITIVE ASPECTS OF NEUROSYMBOLIC INTEGRATION

Yannick Lallement and Frédéric Alexandre

CRIN-CNRS-INRIA
BP 239 – Campus scientifique
54506 Vandœuvre-ls-Nancy Cedex
France

1 INTRODUCTION

The ultimate goal of AI is to emulate the human cognitive processes. During the brief history of AI, two main schools have competed with each other in this endeavor, namely the symbolic and the connectionist schools. Hubert and Stuart Dreyfus have summarized this division by their famous phrase: *making a mind* vs *modeling the brain* [6]. So far, neither of these schools has achieved the final goal. A new research field is growing more and more important at present, one which aims at combining the two schools into *neurosymbolic* models, with the capacities of both the symbolic and connectionist models.

In this chapter, we present some representations to analyze different trends of research in artificial intelligence today. We propose to define two spectra: the AI spectrum representing a variety of AI models, and the cognitive spectrum representing a variety of human cognitive operations. We show how the relationship between these two spectra can be interpreted as a strong argument for the necessity of neurosymbolic models, integrating connectionist and symbolic capabilities. We identify two major possibilities for the construction of neurosymbolic models: the unified approach and the hybrid approach.

2 COGNITIVE ASPECTS OF AI

In this section, we define two representations which will be useful to draw a link between AI and human cognitive processes: the AI spectrum and the cognitive spectrum.

2.1 The AI spectrum

Two different schools are often distinguished in AI research: the symbolic and the connectionist ones. The two schools were born approximately at the same time, in the early 1950s [4].

The symbolic approach, *making a mind*, starts at the higher cognitive level: symbol manipulation and language. The fundamental hypothesis of this approach is that the brain is not the only physical substrate able to produce thought processes. Its typical realizations are expert systems, and such famous natural language handling programs as Shrdlu or Eliza [26].

Contrasting with this, the connectionist approach, *modeling the brain*, starts at the lowest cognitive level: the neurons and the connections between them. The fundamental hypothesis of this approach is that the architecture of the brain has an influence on the thought processes. Its typical realizations are concerned with pattern matching systems (speech or image processing) [19].

It is widely admitted that these two main schools have different and complementary capabilities. Connectionist models often offer good performance in pattern recognition and generalization, and offer features such as natural learning ability, noise tolerance and graceful degradation. They use a non-propositional, distributed, and, thus not easily understandable knowledge representation. In contrast, symbolic models often present a complementary profile: they offer good performance in reasoning and deduction, and present such features as natural symbolic manipulation and explanation ability. They use a propositional, localized, explicit knowledge representation. The complementarity between the two paradigms is often discussed in great detail, see for example [8, 10, 9]. Table 1 gives a brief list of the capabilities usually attributed to connectionist and symbolic models.

However, this is only a superficial view of the field of AI, and things are in fact more complicated. Actually, when we look more closely at various AI models, it often seems difficult to classify them as either connectionist or symbolic: most models will present some typically symbolic features *and* some typically connectionist features. As AI research progresses and new models appear, the symbolic – connectionist duality seems less valid. Moreover, it is not easy to clearly distinguish even between symbolic and connectionist features themselves: some features traditionally attributed to connectionism, such as learning capabilities, are sometimes present in symbolic models, whereas some features traditionally attributed to symbolic models, such as data structure represen-

	Conn. AI	Symbolic AI
Learning	Easy	Difficult
Explanation	None	Good
Performance degrad.	Slow	Fast
Algorithmics	Parallel	Sequential
Noise resistance	Good	Weak
Generalization	Easy	Difficult
Data str. representation	Difficult	Easy

Table 1 Some features of typical connectionist and symbolic models

Figure 1 The AI spectrum

tation, are sometimes present in connectionist models too [13]. Thus, there is not anymore a simple dichotomy between symbolic AI and connectionist AI, and it may be hard to classify a given model as either symbolic or connectionist. A particular model will be *mostly* connectionist or *mostly* symbolic, rather than purely connectionist or purely symbolic. To represent this "continuous dichotomy", we propose to draw a continuum leading from a purely-connectionist side to a purely-symbolic side. According to its respective shares of symbolic and connectionist parts, each AI model can be positioned somewhere on that continuum. We call this continuum the *AI spectrum* (see figure 1).

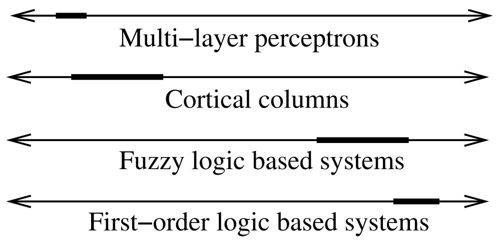

Figure 2 Some AI models

For example (see figure 2), a multi-layer perceptron is clearly on the very connectionist side (a kind of pure connectionist model), and a first-order logic-based expert system on the very symbolic side (a kind of pure symbolic model).

Some models like cortical columns [2] or fuzzy-logic-based expert systems are neither purely connectionist nor purely symbolic, and thus are located somewhere in-between, according to their respective shares of numeric and symbolic parts.

2.2 The cognitive spectrum

From a cognitive point of view, we can notice that several types of mental processes exist and each of them plays a role in human intelligence [3, 24, 23]. In particular, consider the following two contrasting kinds of operations: *synthetic operations* versus *analytical operations* (table 2).

Figure 3 The cognitive spectrum

Synthetic	Analytic
Perception	Reasoning
Implicit	Explicit
Non decomposable	Decomposable
Automatic	Controlled
Unconscious	Conscious
Fast	Slow

Table 2 Some features of typical synthetic and analytical operations

Synthetic operations are mostly concerned with low-level, non decomposable, unconscious, perceptual tasks. Some typical examples are vision and speech recognition: given a low-level input, the goal is to extract some higher-level information, such as objects present in the image, or pronounced words. The knowledge involved in these operations is non-propositional, that is hard to express in words, but rather pictorial for example.

In contrast, analytical operations are mostly concerned with high-level, conscious, decomposable, reasoning tasks. Some typical examples are problem solving, decision-making, or game-playing: in each case, there is a conscious thought process and some conscious decisions, made from explicit high-level information. The knowledge involved in these operations is symbolic and propositional, that is, related to language. Alan Newell [23] has classified these operations according to their duration.

In the case of this supposed dichotomy, it is once again clear that some operations can be mostly synthetic or mostly analytical rather than purely-synthetic or purely-analytical. A typical case is mental calculus: if the computation of 4×3 is clearly synthetic, the computation of 34×13 clearly analytical, what about the computation of 12×7 or 5×15? Certainly somewhere in-between. Another interesting case is musical practice: a beginner at guitar or piano needs to think of what he has to do; his knowledge about the chords and the piece to play is explicit, etc. A more experienced player is less explicitely conscious of his acts, and his playing improves. The playing of a professional musician is fully automatic.

Some complex cognitive operations can use synthetic operations as well as analytical operations. For example, driving a car uses synthetic operations (perception of the environment, of the motor noise...) as well as analytical operations (decision between turning left or right to reach the destination as soon as possible). Therefore we propose to draw a new mental continuum leading from synthetic to analytical operations, namely the *cognitive spectrum* (see figure 3) on which we can situate simple and complex cognitive operations. The complex ones will be located on a wider segment of the spectrum, indicating that they imply several types of processes (see examples on figure 4).

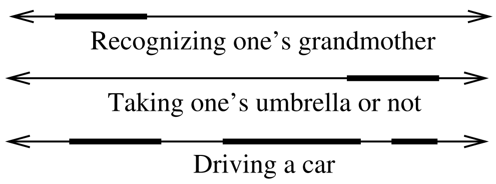

Figure 4 Some cognitive operations

Our two spectra propose a simplified, but accurate view of the two domains. In the AI spectrum, we take into account only the dichotomy between symbolic and connectionist models, ignoring other models such as statistical pattern recognition or genetic learning. But between the two extremities of our spectrum lies the vast majority of the various AI models—which do not belong at either pole. Thus our spectrum covers a wide portion of the field of AI today. In the cognitive spectrum, we do the same kind of simplification, taking into account only one dimension to classify the different kinds of cognitive operations. In this case again, it appears that a single axis covers the variations of several important characteristics (more or less conscious, more or less decomposable, etc.) and thus is not an over-simplification.

3 THE CASE FOR NEUROSYMBOLIC INTEGRATION

The cognitive spectrum and the AI spectrum can be quite easily related to one another: the connectionist side of the AI spectrum mirrors the synthetic side of the cognitive spectrum, and the same is true for the symbolic and analytical sides. The same kind of vocabulary is used to define the synthetic and connectionist extremities on the one hand, and the analytical and symbolic extremities on the other hand. Connectionist models and synthetic operations share many characteristics (e.g, parallelism, non decomposability, non propositional knowledge). This, together with the pattern matching and classification abilities of connectionist models, enables them to perform synthetic perceptual operations.

At the other ends of the spectra, symbolic models and analytical operations also share many features (e.g., sequentiality, decomposability, propositional knowledge). This, together with the symbol manipulation abilities of symbolic models enables them to perform analytical reasoning operations. Moreover, knowledge representation along the AI spectrum is akin to knowledge representation along the cognitive spectrum: from implicit, pictorial to explicit, symbolic. To summarize, connectionist models are best suited to perform synthetic operations, whereas symbolic models are best suited to perform analytical operations.

Thus, a cognitive operation, as defined above, can be situated on the cognitive spectrum, but also on the AI spectrum, according to the kind of model that will best simulate it. For example, a purely perceptual, discrimination operation will be placed on the connectionist side of the AI spectrum. In the same way, an AI model, according to the kind of operation it simulates, can be situated on the cognitive spectrum too. Typically, a pure symbolic model will be placed on the analytical side of the cognitive spectrum.

The isomorphism between our two spectra shows that current AI models are most often able to perform a kind of cognitive operation or another, but rarely both. It also shows that each kind of cognitive operation is related to some AI model which will best simulate it.

AI models striving to achieve the ultimate goal of AI must cover the full gamut of human cognitive processes. This can be easily illustrated in terms of our two spectra: such models have to cover the full cognitive spectrum, and thus the full AI spectrum. This implies realizing neurosymbolic models, cumulating connectionist and symbolic capabilities. In the next section, we examine two ways of developing such models.

4 NEUROSYMBOLIC MODELS: UNIFIED *VS* HYBRID APPROACHES

In order to cover the full AI spectrum, two methods can be identified: adding several segments from various places on the spectrum, or starting the coverage at one of the extremities and leading to the other one. We call the first method *hybrid*, and the second one *unified*: see figures 5 and 6.

Figure 5 Hybrid approach

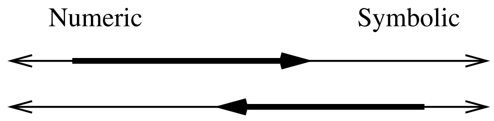

Figure 6 Unified approaches

4.1 The hybrid approach

The hybrid approach tries to combine several models from different origins. Its basic assumption is that covering the AI spectrum is beyond the reach of a single paradigm. The research in hybrid models is currently very varied, and applications include natural language processing [16], diagnosis [14], industrial application [18]...A brief overview of several hybrid models can be found in [17]. This approach has the benefit of a psychological plausibility: the faculty psychology [7] argues that human cognition consists of of several specialized subsystems. Each of these subsystems could be implemented by a particular kind of model. Of course, the synthetic subsystems will naturally implemented by connectionist models, and the analytical subsystems will naturally implemented by symbolic models.

4.2 The unified approaches

In contrast, the unified approaches start with a single paradigm and attempt to enhance it to get both the connectionist and symbolic capabilities. The two original paradigms can be seen as special cases of unified approaches, since, at the beginning, both of them claimed to be able to reach the final goal of AI. However, the current research in unified models is mostly concerned with neural networks. The starting point of a unified model can be either on the connectionist, or on the symbolic side of the spectrum. In the first case, the approach is bottom-up, neuron-to-symbol oriented. The point is to make symbols emerge from neural networks. Such models are often inspired by neuro-biology: see for example [2, 21, 22]. In the latter case, the approach is top-down, symbol-to-neuron oriented. The point is to build connectionist architectures for symbol processing, and thus to ground symbols in a connectionist substrate. This approach is widely known as *connectionist symbol processing*: see for example [1, 11, 25]. In [17] can be found a brief overview of several unified models.

5 DISCUSSION

Given our definitions—mono- or multi-paradigms—the two approaches seem to exhaust the current research in neuro-symbolic integration. It is worth noting that hybrid models may include models which stem from the unified approach, but, even in that case, they remain hybrid since they are made up of several kinds of models.

A question naturally comes to mind: which is the most promising approach? Our definitions say nothing about this: there is no reason why one way to cover the cognitive spectrum would be better than the other. The hybrid approach is sometimes viewed as a short-term, engineering expedient, whereas, in the long run, the unified approach would be the key to human cognition. Indeed, the hybrid approach has so far given more convincing results than the unified approach: for example, some models are already used in industrial applications. However, the two approaches are still in their infancy, and have to solve a number of research issues before any real assessment can be made.

The research issues concerning the hybrid approach are mainly those of communication and task sharing between different tools. In the case of the unified approach, the main issue is to continue the bottom-up or top-down strategies that have just begun, to date, to reach the *sub-symbolic* level that seems to be the most relevant in human cognition [12]. It is definitely not clear at the moment which approach will turn out to be the most successful.

6 SUMMARY

In this paper, we defined some representations (the AI and the cognitive spectra) that we used to underline the necessity for neurosymbolic integration and to show how this integration can be achieved: through a hybrid or a unified approach. We suggest that these two approaches have a significant role to play in the future of AI, and that both of them are complementary rather than exclusive. The complexity and the diversity of human cognition and the human brain leads us to think that a single unified model may not be sufficient to achieve the final goal of AI: as Marvin Minsky [20] writes:

> *To solve really hard problems, we'll have to use several different representations.*

And indeed, considering the variety of problems yet to be solved, a single type of representation seems to be very far away. An interesting next step in the domain of neurosymbolic integration would certainly be to integrate classical models (connectionist or symbolic) together with unified models into new multi-hybrid models. This approach may well lead to some fundamental breakthroughs in AI: such models would present both biological and psychological plausibility, which is not the case in the current AI models.

Acknowledgements

The authors wish to thank Mélanie Hilario and Bruno Orsier for numerous discussions which contributed to the propositions developed in this paper, in a noteworthy way. They also wish to thank Mark Walmsley for his helpful comments.

REFERENCES

[1] V. Ajjanagadde. Unclear distinctions lead to unnecessary shortcomings: Examining the rule vs fact, role vs filler, and type vs predicate distinctions from a connectionist representation and reasoning perspective. In *American Association Artificial Intelligence*, 1994.

[2] F. Alexandre, Y. Burnod, F. Guyot, and J-P. Haton. The cortical column: a new processing unit for multilayered networks. *Neural Networks*, 4(1), 1991.

[3] T. Cornu. *Machine Cellulaire Virtuelle : Définition, Implantation et Exploitation.* PhD thesis, Université de Nancy I and Supélec Metz, 1992. Available at ftp://ftp.loria.fr/ pub/loria/rfia/ publications/cortex/ cornu.machine-cellulaire.thesis92.tar.Z.

[4] D. Crevier. *AI: the tumultuous history of the search for artificial intelligence.* BasicBooks, 1993.

[5] J. Dinsmore, editor. *The Symbolic and connectionist paradigms: closing the Gap.* Lawrence Erlbaum, 1992.

[6] H. Dreyfus and S. Dreyfus. Making a mind versus modeling the brain: Artificial intelligence at a branchpoint. *Daedalus*, Winter, 1988.

[7] J. A. Fodor. *The Modularity of Mind: An essay on faculty psychology.* MIT Press, 1983.

[8] S. I. Gallant. Connectionist expert systems. *Communications of the ACM*, 31(2), February 1988.

[9] M. Gutknecht. The postmodern mind: hybrid models of cognition. *Connection Science*, 4(3 & 4), 1992.

[10] J. Hawthorne. On the compatibility of connectionist and classical models. *Philosophical Psychology*, 2(1), 1989.

[11] G. E. Hinton, editor. *Artificial Intelligence, special issue on connectionist symbol processing*, volume 46. Elsevier, November 1990.

[12] D. Hofstadter. *Metamagical themas: questing for the essence of mind and pattern*. Basic Books, 1985.

[13] V. Honavar and L. Uhr, editors. *Artificial intelligence and neural networks: steps toward principled integration*. Academic Press, 1994.

[14] D. L. Hudson, M. E. Cohen, P. W. Banda, and M. S. Blois. Medical diagnosis and treatment plans derived from a hybrid expert system. In Kandel and Langholz [15].

[15] A. Kandel and G. Langholz, editors. *Hybrid architectures for intelligent systems*. CRC Press, 1992.

[16] S. C. Kwasny and K. A. Faisal. Symbolic parsing via subsymbolic rules. In Dinsmore [5].

[17] Y. Lallement, M. Hilario, and F. Alexandre. Neurosymbolic integration: Cognitive grounds and computational strategies. In M. DeGlas and Z. Pawlak, editors, *World Conference on the Fundamentals of Artificial Intelligence*, Paris, July 1995. Available at ftp://ftp.loria.fr/pub/loria/rfia/publications/cortex/lallement.comp.neurosymbolic95.ps.Z.

[18] G. R. Madey, J. Weinroth, and V. Shah. Hybrid intelligent systems: Tools for decision making in intelligent manufacturing. In C. H. Dagli, editor, *Artificial neural networks for intelligent manufacturing*. Chapman & Hall, 1994.

[19] J.L. McClellan, D.E. Rumelhart, and the PDP Research Group. *Parallel Distributed Processing*, volume 1 and 2. MIT press, 1986.

[20] M. Minsky. Logical versus analogical or symbolic versus connectionist or neat versus scruffy. *AI Magazine*, Summer 1991.

[21] V. Nenov and M. Dyer. Perceptually grounded language learning: Part 1 - a neural network architecture for robust sequence association. *Connection Science*, 5(2), 1993.

[22] V. Nenov and M. Dyer. Perceptually grounded language learning: Part 2 - dete: a neural/procedural model. *Connection Science*, 6(1), 1994.

[23] Alan Newell. *Unified theories of cognition*. Harvard University Press, Cambridge, Mass., 1990.

[24] Jean-François Richard. *Les activités mentales.* Armand Colin, 1990.

[25] R. Sun. On variable binding in connectionist networks. *Connection Science*, 4(2), 1992.

[26] P. H. Winston. *Artificial Intelligence.* Addison Wesley, 1977.