The Emergence of Symbolic Processes From the Subsymbolic Substrate*

Bruce MacLennan
Computer Science Department
University of Tennessee, Knoxville

March 13, 1991

Abstract

A central question for the success of neural network technology is the relation of symbolic processes (e.g., language and logic) to the underlying subsymbolic processes (e.g., parallel distributed implementations of pattern recognition, analogical reasoning and learning). This is not simply an issue of integrating neural networks with conventional expert system technology. Human symbolic cognition is flexible because it is not purely formal, and because it retains some of the "softness" of the subsymbolic processes. If we want our computers to be as flexible as people, then we need to understand the emergence of the discrete and symbolic from the continuous and subsymbolic.

1 Human Symbolic Cognition

Human symbolic cognition, by which we mean the use of language, logic and explicit reasoning, is much more flexible than symbolic AI systems (MacLennan, 1988). If an expert system has too few rules, then it exhibits *brittle* behavior, failing in catastrophic ways when faced with novel situations or minor exceptions to the rules. On the

*Text of invited panel presentation, *International Symposium on New Information Processing Technologies '91*, Tokyo, March 13–14, 1991.

other hand, trying to anticipate all the situations and exceptions that may occur leads to a proliferation of rules and an exponential explosion in the machine resources required. Human cognition does not have these limitations. When people reason explicitly, their use of categories is sensitive to the context of the problem, and their inferential processes are generated from and constrained by relevance to the situation. This seems to be a natural result of the implementation of these processes, rather than a result of special context rules or relevance rules. Human language use has similar characteristics: it is flexible, context-sensitive and controlled by relevance. Further, flexibility and "softness" is the natural state of both reason and language; "hard" logic and precise language use are special skills that are not easily acquired, and they are special tools used by the expert when they are called for, but not otherwise. How can we achieve similar flexibility in our computer systems?

It is not sufficient to replace categories with hard boundaries by categories with fuzzy boundaries (as is done in fuzzy set theory). Although fuzzy categories do eliminate some sources of brittleness, they do not address the complex processes by which categories may be sensitive to the global context. To achieve the flexibility of human cognition and true context-sensitive symbol use, a more radical reinterpretation of symbolic processing is necessary. We must see context-sensitive holistic processes as the normal mode of operation, and context-free discrete symbol manipulation as a specialized modification of this norm.

These observations have practical implications, for they mean human-like flexibility and competence is unlikely to result from a simple hybrid of neural network technology and expert systems technology. Wherever it appears in a hybrid system, discrete, formal, context-free symbol manipulation will be a source of brittleness and other forms of unskillful behavior. If we want implementations of symbolic cognition that exhibit the flexibility and competence of people's, then they must be built on a foundation that is fundamentally continuous, holistic and context-sensitive.

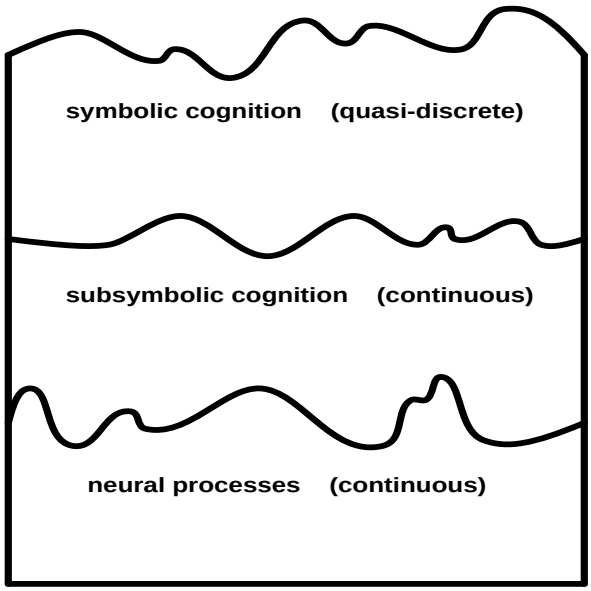

Figure 1: Emergence of Symbolic Cognition from Subsymbolic Substrate.

2 Emergence of the Discrete from the Continuous

Human symbolic cognition is built upon a *subsymbolic substrate*, which is continuous in its principles of operation, and is the ultimate source of the flexibility of human symbol use (Fig. 1; see also Smolensky, 1988). To see this, observe that the basic neural processes are best described as continuous, since information is represented by continuous quantities such as spiking frequency and membrane potential. Most neurons seem to operate as low precision analog devices. Furthermore, the large number of neurons found in the brains of the higher animals means that functional areas can often be viewed as *spatially* continuous. That is, we can view such an area as a continuum of neurons, rather than as a large number of discrete neurons (MacLennan, 1987). Thus the basic neural processes are both temporally and spatially continuous. This is important from a theoretical perspective, because it means that powerful mathematical tools may be brought to bear on the problem of mental representation (e.g., see MacLennan, 1991b).

In addition, many of the most basic cognitive processes, such as

perception, association, sensory-motor coordination, and judgement of similarity, are by their nature continuous. We share these faculties with the lower animals, and we observe that they show the same flexibility in their use as we do. Discreteness is most apparent when we come to higher cognitive processes, such as language use and explicit reason, but they partake of the flexibility of the underlying continuous processes.

These considerations suggest that a critical research goal is to understand the processes by which discrete and approximately discrete symbols can emerge from continuous processes. Of course this is an inversion of the usual situation in computer science and even logic, where discrete symbols are assumed to be primitive, and continuous quantities are approximated by discrete structures. We must understand how the discrete structures found in logic and language can emerge from continuous representations and processes (MacLennan 1991a, in press).

A possible basis for the required theory can be found in the mathematics of Hilbert spaces, which establish a fundamental link between discrete and continuous phenomena (MacLennan, 1990b; Pribram, 1991). The simplest example is the Fourier series for a square wave; here we have a discrete object being approximated by a superposition of continuous objects. Of course Fourier analysis has been extensively applied to the analysis of the visual and auditory systems (MacLennan, 1991b), but it is perhaps surprising that it may inform our understanding of symbolic cognition.

A few more examples may be in order. Since Hilbert spaces also permit multiple representations in linear superposition, we can represent in the same functional area several interpretations of a retinal image, or several parses of a sentence. Further, these superpositions can be processed in parallel by a single neural network (MacLennan, 1990a, p. 674n); see Fig. 2. Linear superpositions can also be used to simultaneously represent several objects, thus avoiding problems with illusory conjunctions of features. Such superpositions suggest simple means for shifting objects in and out of the focus of attention, while retaining the contextual sensitivity of the foreground to the background (Fig. 3).

We are not suggesting that Hilbert spaces are the answer to every question about symbolic cognition; there are many other tools available. For example the topology of connected metric spaces seems a

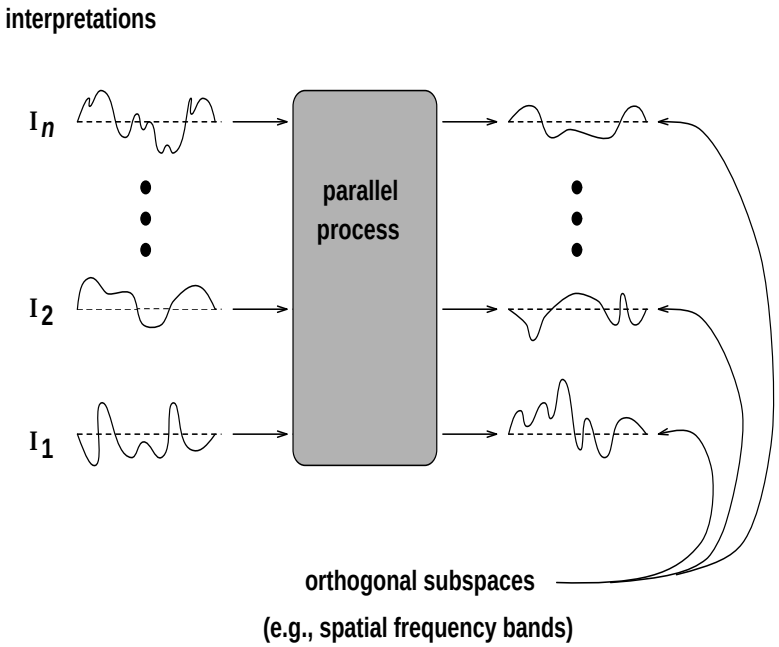

Figure 2: Parallel Processing by Linear Superposition in a Neural Network

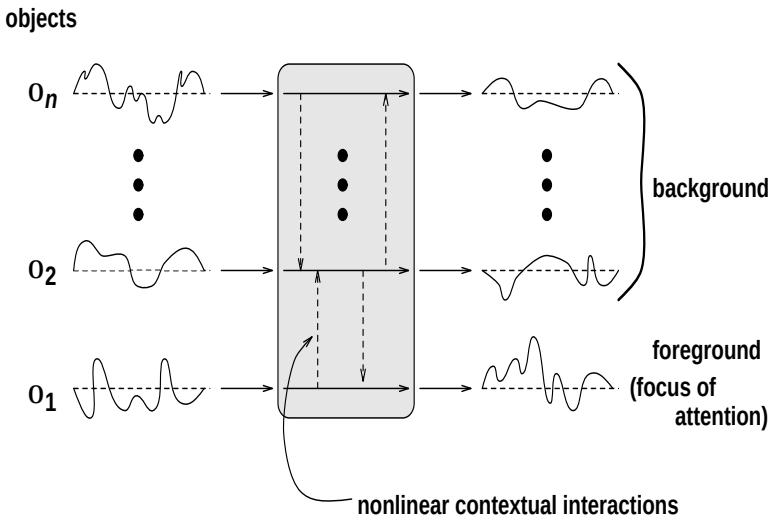

Figure 3: Contextual Effects Represented by Nonlinear Interactions.

useful basis for a qualitative theory of symbolic cognition (MacLennan 1991a, in press). In any case, much more research needs to be done.

3 Implications for Sixth Generation Technology

Neural networks and related technologies show great promise as efficient, flexible implementations of lower order cognitive processes, such as perception, motor control, recognition, association and judgement of similarity. Many applications, however, will depend on implementations of higher-order cognitive processes. Unfortunately, efficient, flexible implementations of higher-order cognitive faculties, such as natural language and explicit reasoning, will require an understanding of their relation to the lower-order processes. Although many connectionist researchers are working on this problem, it is still unsolved. Nevertheless, a theory of the emergence of symbolic processes from the subsymbolic substrate will be of great practical and theoretical importance. We conclude with a quotation:

> The disadvantage of regarding things in separate parts is that when one begins to cut up and analyze, each one tries to be exhaustive. The disadvantage of trying to be exhaustive is that it is consciously (mechanically) exhaustive. One goes on deeper and deeper, forgetting to return and sees a ghost. Or one goes on and imagines he has got it, and what he has got is only a carcass. For a thing which retains its substance but has lost the magic touch of life is but a ghost. *Only one who can imagine the formless in the formed can arrive at the truth.*
>
> — Soshi

4 References

MacLennan, B. J. (1987). Technology-independent design of neurocomputers: The universal field computer. In M. Caudill & C. Butler (Eds.), *Proceedings, IEEE First International Conference on Neural Networks* (Vol. 3, pp. 39–49).

MacLennan, B. J. (1988). Logic for the new AI. In J. H. Fetzer (Ed.), *Aspects of artificial intelligence* (pp. 163–192). Dordrecht, Holland: Kluwer Academic Publishers.

MacLennan, B. J. (1990a). The discomforts of dualism. *Behavioral and Brain Sciences*, **13**, 673–674.

MacLennan, B. J. (1990b). *Field computation: A theoretical framework for massively parallel analog computation, parts I–IV* (Report No. CS-90-100). Knoxville, TN: University of Tennessee, Department of Computer Science.

MacLennan, B. J. (1991a). *Characteristics of connectionist knowledge representation* (Report No. CS-91-147). Knoxville, TN: University of Tennessee, Computer Science Department. Also submitted for publication.

MacLennan, B. J. (1991b). *Gabor representations of spatiotemporal visual images* (Report No. CS-91-144). Knoxville, TN: University of Tennessee, Computer Science Department. Also submitted for publication.

MacLennan, B. (in press). Continuous symbol systems: The logic of connectionism. In Daniel S. Levine and Manuel Aparicio IV (Eds.), *Neural networks for knowledge representation and inference*. Hillsdale, NJ: Lawrence Erlbaum.

Pribram, K. H. (1991). *Brain and perception: Holonomy and structure in figural processing*. Hillsdale: Lawrence Erlbaum.

Smolensky, P. (1988). On the proper treatment of connectionism. *Behavioral and Brain Sciences*, **11**, 1–74.