Book Reviews

A Connectionist Approach to Word Sense Disambiguation

Garrison W. Cottrell
(University of California at San Diego)

London: Pitman, and San Mateo, CA: Morgan Kaufmann, 1989, xi + 220 pp.
(Research Notes in Artificial Intelligence)
Paperbound, ISBN 0-934613-61-3, $29.95

Reviewed by
Graeme Hirst
University of Toronto

The title of Cottrell's book mentions only two concepts: connectionism and lexical disambiguation. That's misleading, because the book has much more to offer than just that. Among the topics addressed are parsing, agrammatism, connectionist inheritance hierarchies, and structural ambiguity, and it is the integration of this wide-ranging set of topics that is one of the strengths of the work.

The book appears four years after the 1985 University of Rochester dissertation upon which it is based. Thus the flavor of connectionism that Cottrell uses is the coarse-grained localist representations used at Rochester in the early 1980s, in which each node in the network represents a concept. This is in contrast to the distributed representations ("PDP") that became popular in the latter part of the decade, in which many nodes may contribute to the representation of a concept (Rumelhart and McClelland 1986). Cottrell has taken advantage of the delay in publication to restructure the work substantially and to add discussions of the later research. He seems to suggest (p. 7) that distributed representations are generally preferable because they can learn, whereas localist networks like his own need to be individually hand designed. Nevertheless, this research shows that there is considerable appeal in hand-designed, localist networks.

Cottrell takes work in psycholinguistics as the starting point for his model of lexical access and disambiguation. In the early 1980s, it was discovered that in many circumstances, people subconsciously consider all meanings of an ambiguous word, even if the preceding context makes one alternative preferable *a priori*. For example, the floral sense of the word *rose* is activated even when one hears *The congregation rose.* Within a few hundred milliseconds all senses but the one chosen as correct become deactivated again. (While subsequent research has qualified these results somewhat—see Gorfein 1989—the basic principle has proven to be robust.) The usual explanation for these results is in terms of priming and spreading activation in a semantic network, so a localist model is very natural.

The input to Cottrell's networks is a string of words forming a syntactically simple sentence, such as *Bob threw a ball to the dog.* This is done by activating the nodes corresponding to the words. The activation of a node causes the activation of those other nodes in the network to which it is connected by excitory links and the deactivation of those to which it is connected by inhibitory links. A node can receive activation and inhibition at the same time; for example, an ambiguous word will send activation to all its senses, but the senses will be mutally inhibitory. Thus the network may be unstable for some time until it settles down into a pattern of activation that represents its "output"; the nodes representing the relevant concepts are activated and other nodes aren't. In the case of an ambiguous word, the correct meaning in context will presumably receive activation from more sources, or be pre-activated by the preceding context, and thus be able eventually to force its competitors into inhibition. This final pattern of activation may be construed as the interpretation of the sentence.

After the word-sense selection network, there are two more networks, running in parallel with one another: one for determining case roles and one for syntactic analysis. The case role network uses an "exploded" notion of cases; that is, rather than having one node representing, say, the agent role, Cottrell has one node for the agent of a *propel* action, one for the agent of a *vomit* action, and so on. (The topic area of Cottrell's example sentences ranges from baseball to emesis.) This seems counter-intuitive, or unparsimonious at the very least; but I must admit that, modern linguistic theory notwithstanding, I know of no particular psycholinguistic evidence for the reality of a single concept of, say, agency that is activated for any and every sentence that involves an agent.

A feature of the parsing network is that it need not be constructed by hand; rather, it is automatically generated from a grammar and lexicon by a Lisp program. It parses only the very simple one-clause sentences needed to test the other parts of the system. Unlike the other parts of the system, the parser has no special claim to psychological reality. However, the minimal-attachment strategy of structural ambiguity resolution (namely, to attach a new constituent in the way that creates the fewest new nodes) "falls out" as a natural consequence of the design.

Cottrell includes an interesting discussion of his system's predictions for aphasia. If the system has some psychological reality, then one would expect that "damage" to the network would result in behavior similar to that of aphasic patients. For example, if the connection between the case

and syntax networks is severed, they can no longer constrain each other. The result is a comprehension deficit rather like that of certain agrammatic aphasic patients.

Cottrell's work in some ways resembles my own (Hirst and Charniak 1982; Hirst 1987) and that of Waltz and Pollack (1985). The most important difference is that this other work tried to mix conventional symbolic approaches together with connectionist-like spreading activation for disambiguation. Waltz and Pollack, for example, use a chart parser to build a network that represents the alternative parses of the input sentence. Activation is then spread through the network, causing one of the parses and one meaning of each ambiguous word to be chosen. My own work started from the same psycholinguistic data as Cottrell's. However, lexical disambiguation was performed by a set of parallel cooperating processes, one per word, which drew on the results of spreading activation in a semantic network as just one of several sources of knowledge for disambiguation. Parsing and semantic interpretation were purely symbolic.

As NLU systems go, Cottrell's is pretty dinky; it doesn't do anything new. What's different and important about it is *how* it does what it does. By using localist connectionist networks for everything, Cottrell shows the potential of the approach, and lays a foundation for the development of non-dinky systems. However, the price paid for this is the need to reinvent, almost from scratch, everything that computational linguistics has done in the least 20 years. It seems a little perverse to be slaving away, for example, on a connectionist parser for simple sentences like *Bob barfed badly* when highly sophisticated parsers and grammars are already available.

The reply, of course, is that one day the connectionist systems will outstrip anything that we have now; they'll be faster and more elegant, and so natural that all known principles of parsing and interpretation (and maybe a few more) will be "emergent properties" of the systems. In particular, symbolic systems have had great difficulty with some of the fuzzier aspects of language understanding, such as trading off conflicting preferences in the interpretation of an utterance, and such trade-offs are clearly a strength of connectionism. But while recent research in connectionist NLU suggests that useful systems may indeed be possible, it will remain for quite some time an article of faith rather than science that such a research program can be carried through to completion. It is books like Cottrell's that help to sustain that faith.

Cottrell is excellent at analyzing the strengths and weaknesses of various approaches—his own and those of other researchers—and his discussions of other research are a valuable part of the book. It is also nice to see a book in which the author can so honestly present the good and bad points of his own work. Cottrell has an easy and breezy writing style (with a whimsical canine leitmotif) that is always clear and a pleasure to read. His book is an impressive integration of AI, psycholinguistics, and neurolinguistics, in the best traditions of cognitive science.

References

Gorfein, David, editor (1989). *Resolving semantic ambiguity.* NY: Springer.

Hirst, Graeme (1987). *Semantic interpretation and the resolution of ambiguity.* Cambridge: Cambridge University Press.

Hirst, Graeme and Charniak, Eugene (1982). "Word sense and case slot disambiguation." In *Proceedings of the Second National Conference on Artificial Intelligence* (AAAI-82), Pittsburgh, August 1982, 95–98.

Rumelhart, David and McClelland, James (1986). *Parallel distributed Processing: Explorations in the microstructure of cognition.* Cambridge, MA: MIT Press.

Waltz, David and Pollack, Jordan (1985). "Massively parallel parsing: A strongly interactive model of natural language interpretation." *Cognitive Science,* 9(1), 51–74.

Graeme Hirst once thought about becoming a connectionist, but he's better now, thank you. Hirst's address is: Department of Computer Science, University of Toronto, Toronto, Canada M5S 1A4. E-mail: gh@cs.toronto.edu

---

Artificial Intelligence Techniques in Language Learning

Rex W. Last
(Department of Modern Languages, University of Dundee)

Chichester, England: Ellis Horwood, 1989, 173 pp.
(Ellis Horwood series in computers and their applications)
Hardbound, ISBN 0-7458-0177-3 and 0-470-21503-8, $64.95

Reviewed by
Camilla Schwind
Centre National de la Recherche Scientifique

This book is a state-of-the-art review of the techniques of artificial intelligence in computer-assisted language learning (CALL). This is an extremely interesting subject, which has up to now not been treated extensively in AI, nor more especially in natural language understanding. The book's objectives are:

- to examine the current developmental level of computer-assisted language learning (from the point of view of the informed modern language teacher and researcher);
- to disentangle the present state of the art of artificial intelligence as it relates to CALL;
- to establish the extent to which artificial intelligence applications can be applied to the future development of CALL.

First, a survey of CALL is given, explaining the how and why of the evolution of the field up to the present. The next chapter, entitled "What is AI?" tries to "consider the whole question of the nature of AI." The rest of the book is devoted to the presentation and discussion of several areas of AI that the author considers relevant to CALL, such as human/computer interfaces, knowledge representation, and