Connectionist Models of Language

James L. McClelland

Traditional models of language processing process language by rule. This approach faces two problems. First, there are difficulties in using the rules during processing, since often one rule must be pitted against another. In this case traditional approaches face the difficult problem of deciding which rule should win in such cases. Second, there are difficulties in acquiring rules, since it is often hard to know when a rule should be proposed, or when a sentence should be handled as one of many special cases.

In the connectionist approach my colleagues and I have been taking, language processing is viewed as a constraint satisfaction process. Each constituent of a sentence is viewed as imposing constraints on the representation of the state or event described by the sentence. During processing, as each constituent is encountered, it constrains the evolving representation of the sentence.

The knowledge that governs this constraint satisfaction is stored in the strengths of the connections among the units in a connectionist network. These connection strengths encode the knowledge that is traditionally encoded in the form of rules, but have the advantage that they are naturally capable of capturing constraints that differ in magnitude or degree. The acquisition of these connection strengths occurs through a connection adjustment process based on the back-propagation learning algorithm. The algorithm performs gradient descent in a measure of the extent to which the answers that the network gives to questions about the event described by a sentence actually match the probability that those answers are correct given the sentence. This algorithm is able to learn to assign the correct interpretations even when there are conflicting cues to the correct interpretation of a sentence.

To date this approach has been applied successfully to the processing of one-clause sentences. We have shown that it can learn to assign meanings to sentences containing vague and ambiguous words; that it fills in implicit arguments, and that it can use both word meaning and word order information correctly in making assignments of constituents to roles.

Current extensions focus on improving the rate of learning and on extending the approach to sentences of arbitrary complexity. In this regard we have recently established that a simpler variant of the model used for the comprehension of one-clause sentences is capable of learning, from a finite set of examples, to process all of the infinite corpus of sentences generated by a Finite State Automaton.

References

The following two Technical Reports give details of the research described above:

St. John, M. & McClelland, J. L. *Learning and applying contextual constraints in sentence comprehension.* AIP Technical Report, Departments of Psychology and Computer Science, Carnegie Mellon University.

Servan-Schreiber, D., Cleeremans, A., and McClelland, J. L. *Encoding sequential structure in simple recurrent networks.* Technical Report CMU-CS-88-183, Department of Computer Science, Carnegie Mellon University.