CSC: Classic Paper Review/Analysis

Title and Author

Title The Appeal of Parallel Distributed Processing

Author J. L. McCLELLAND, D. E. RUMELHART, and G. E. HINTON

Summary/Hook

Introductory section to Parallel Distributed Processing (PDP) models in The Appeal of Parallel Distributed Processing. The authors compile many network theory models they have developed with examples like motor function, visual processing, and knowledge representations. The grounding is within their biological inspiration from the function of neurons. This book collected many well cited papers and contributions to the area, leading to the artificial neural networks we have today.

Knowledge Relating to the Cognitive Science Program Learning Outcomes

1. 6.Psychological Investigations

   As Feldman and Ballard (1982) have pointed out , the biological hardware is just too sluggish for sequential models of the microstructure to provide a plausible account , at least of the microstructure' of human thought. And the time limitation only gets worse, not better, when sequential mechanisms try to take large numbers of constraints into account. Each additional constraint requires more time in a sequential machine, and, if the constraints are imprecise, the constraints can lead to a computational explosion. Yet people get faster, not slower, when they are able to exploit additional constraints

2. 2. Symbol Systems

   Graceful degradation. A few of the desirable properties of this kind of model are visible from considering what happens as we vary the set of features we use to probe the memory in an attempt to retrieve a particular individual's name. Any set of features which is sufficient to uniquely characterize a particular item will activate the instance node for that item .more strongly than any other instance node. A probe which contains misleading features will most strongly activate the node that it matches best. This will clearly be a poorer cue than one which contains no misleading information-but it will still be sufficient to activate the "right answer " more strongly than any other, as long as the introduction qf misleading information does not make the probe closer to some other item. In general , though the degree of activation of a particular instance node and of the corresponding name nodes varies in this model as a function of the exact content of the probe, errors in the probe will not be fatal unless they make the probe point to the wrong memory. This kind of model's handling of incomplete or partial probes also requires no special error-recovery scheme to work-it is a

natural by-product of the nature of the retrieval mechanism that it is capable of graceful degradation.

3.3. Neural Networking

In most models, knowledge is stored as a static copy of a pattern. Retrieval amounts to finding the pattern in long-term memory and copying it into a buffer or working memory. There is no real difference between the stored representation in long-term memory and the active representation in working memory. In PDP models, though, this is not the case. In these models, the patterns themselves are not stored. Rather, what is stored is the connection strengths between units that allow these patterns to be re-created. In the Jets and Sharks model , there is an instance unit assigned to each individual , but that unit does not contain a copy of the representation of that individual. Instead , it is simply the case that the connections between it and the other units in the system are such that activation of the unit will cause the pattern for the individual to be reinstated on the property units.

4. 3. Neural Networking

The approach that we take in developing PDP models is completely different. First , we do not assume that the goal of learning is the formulation of explicit rules. Rather, we assume it is the acquisition of connection strengths which allow a network of simple units to act though it knew the rules. Second, we do not attribute powerful computational capabilities to the learning mechanism. Rather, we assume very simple connection strength modulation mechanisms which adjust the strength of connections between units based on information locally available at the connection.

5. 3. Neural Networking

The strengths of the links they share in common with other patterns. Thus, if we present the same pair of patterns over and over, but each time we add a little random noise to each element of each member of the pair, the system will automatically learn to associate the central tendency of the . two patterns and will learn to ignore the noise. . What will be stored will be an average of the similar patterns with the slight variations removed. On the other hand, when we present the system with completely uncorrelated patterns, they will not interact with each other in this way. Thus, the same pool of units can extract the central tendency of each of a number of pairs of unrelated patterns