Learning the English past tense
Connectionism and Neural Networks

Lesson VII: Language
module 34

Connectionist models
Parallel distributed processing (PDP)

- Motivated by neural information processing
  - Huge # of "simple" information processing units
  - Information processing by activation flow between units
    and integration of activation within units
  - Units can be active (above threshold) or inactive

![Diagram of neural network]

Distributed representation

- Representation of concepts
  - Unlike traditional semantic network models, one unit
    doesn’t represent a concept by itself
  - Representations are distributed across many units
    (activation patterns within the network)
  - Allows for subtle, "sub-symbolic" interactions and fuzzy
    representations

![Image of distributed representation]
Learning in connectionist models

- Learning by statistical association
  - Units that are active at similar times will tend to strengthen their connections between each other
  - Units that are not correlated in their activity will decrease their connections

- Teaching
  - In many models, a network of units will be given an input activation pattern and a target output pattern
  - Changes in connection weights depend on how different the current activation is from target value
  - Effects can be back-propagated (passed on) to other units

Learning of English past tense

- Phenomenon
  - Children first learn correct forms of past tense of irregular verbs
  - Later children overgeneralize the rule “-ed” and make mistakes for irregular verbs (taked, tooked)
  - At the end, past tense for irregular verbs will return to correct forms

- Rumelhart & McClelland’s (1986) model
  - Connectionist network representing the phonological representations of present and past tense verbs as input and output “layers”
  - Training of the model with verbs at frequencies of normal occurrence of these verbs in language
One problem with phonology

- The model doesn’t assume that there are separate processes for regular and irregular verbs
- Traditional models assume that children have to memorize irregular forms and separately learn the general rules

Interesting aspects of the model

- Single route model
  - Shows qualitively similar behavior as children’s learning curves
  - Shows examples of overgeneralization in the middle of the learning process
  - Returns to almost perfect behavior after a large number of learning episodes