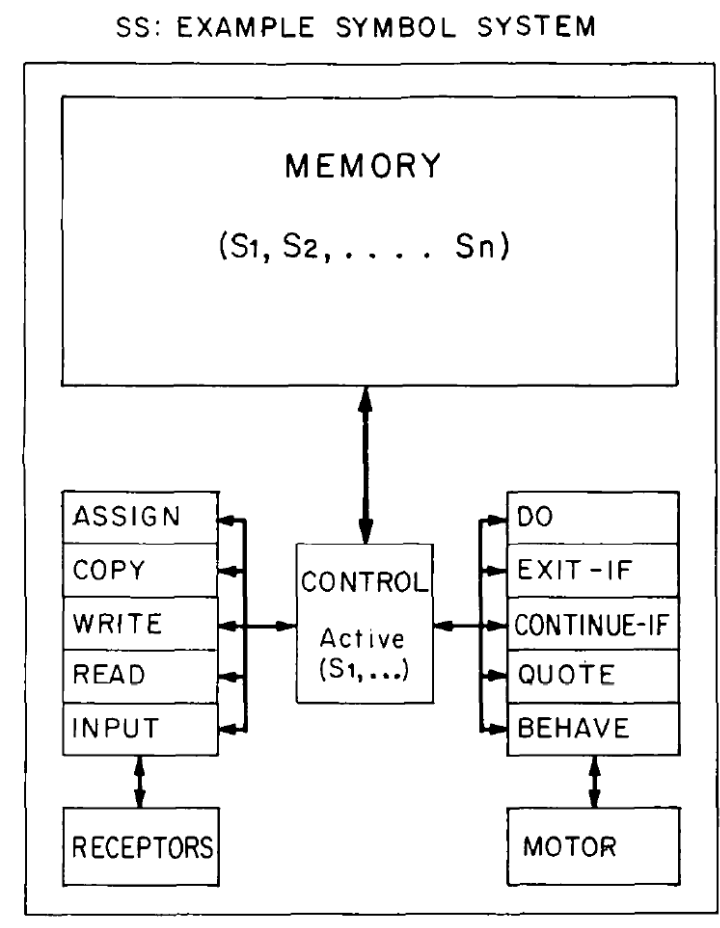


The PDP Approach to Understanding the Mind and Brain

Jay McClelland
Stanford University
January 21, 2014

Early Computational Models of Human Cognition (1950-1980)

- The digital computer instantiates a 'physical symbol system'
- Simon announces that he and Allan Newell have programmed a computer to 'think'.
- Symbol processing languages are introduced allowing success at theorem proving, problem solving, etc.
- Human subjects asked to give verbal reports while problem solving follow paths similar to those followed by N&S's programs.
- Psychologists investigate mental processes as sequences of discrete stages.
- Early neural network models fail to live up to expectations; Minsky and Pappert kill them off.
- Cognitive psychologists distinguish between algorithm and hardware; Neisser deems physiology to be only of 'peripheral interest'.



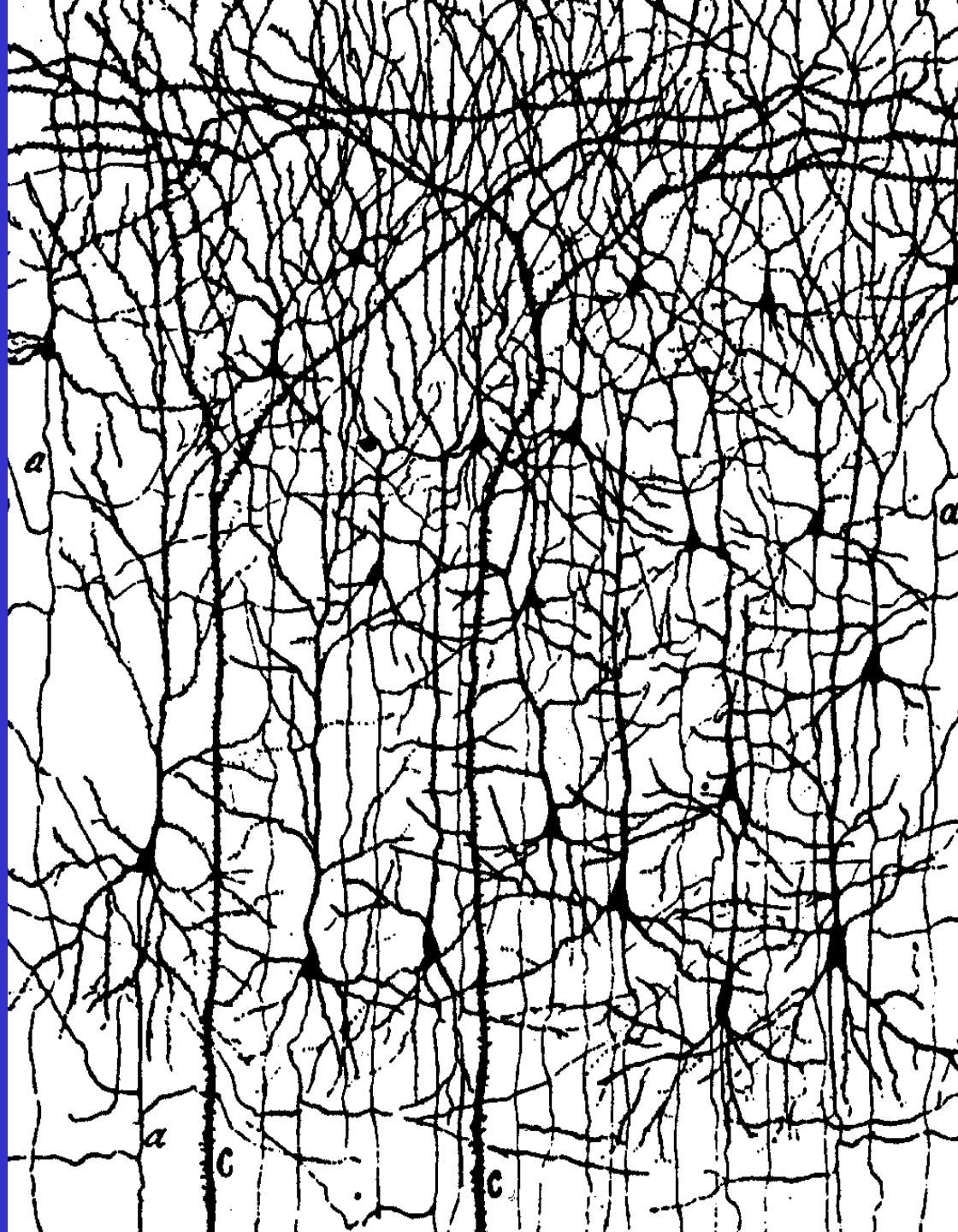


Ubiquity of the Constraint Satisfaction Problem

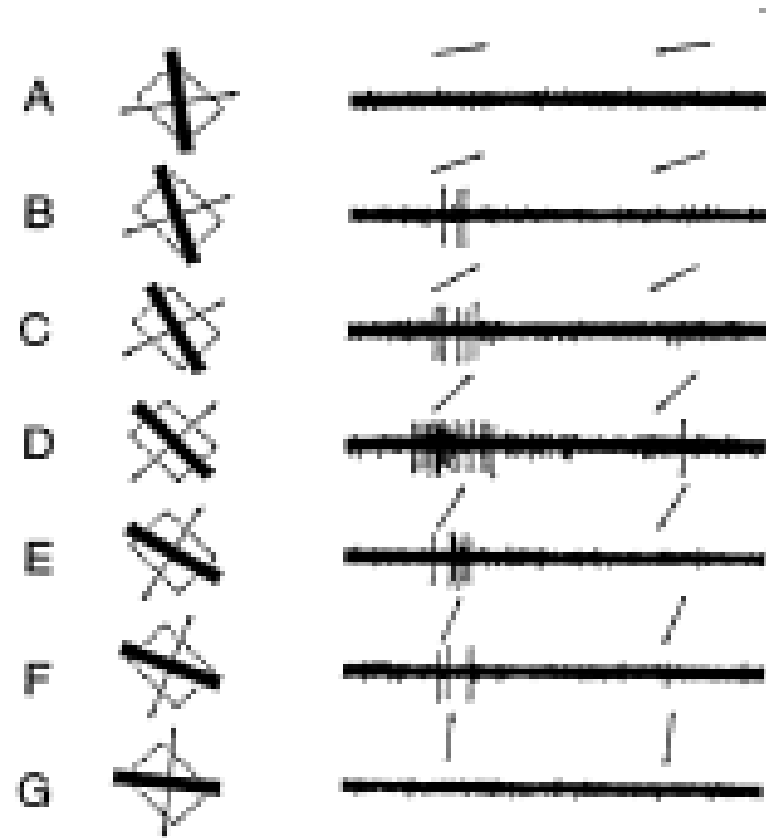


David E.
Rumelhart

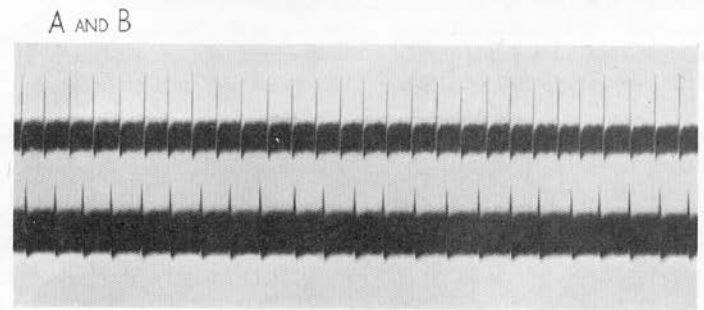
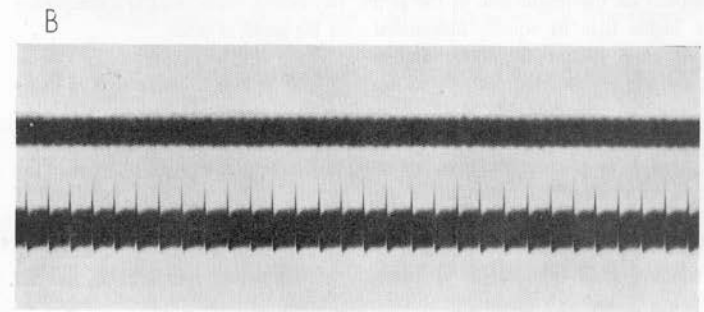
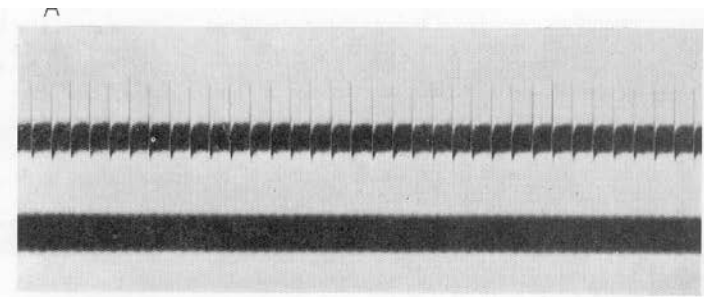
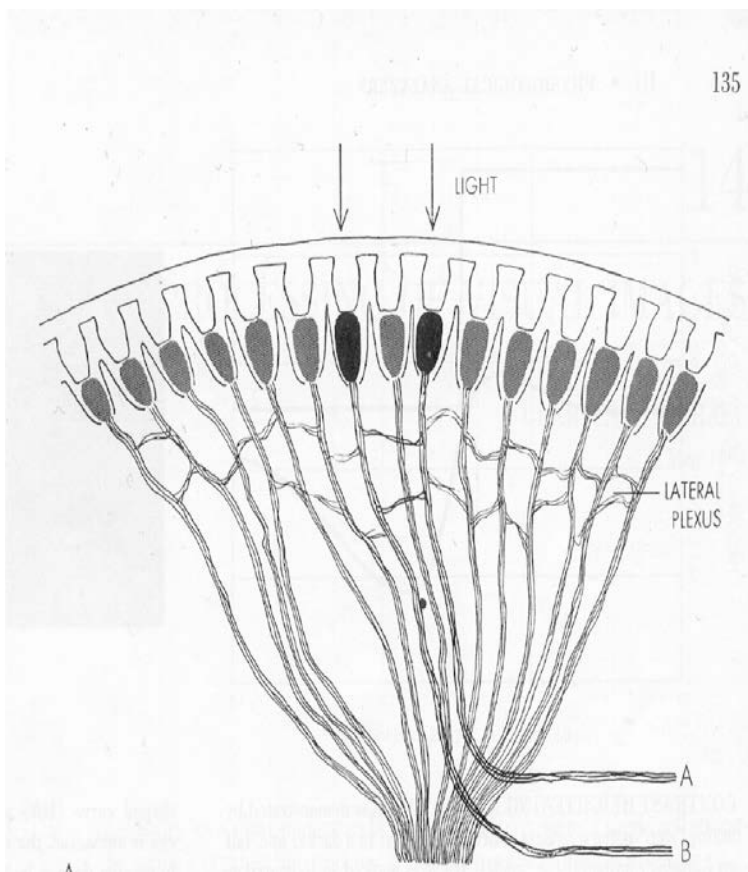
- In sentence processing
 - I saw the grand canyon flying to New York
 - I saw the sheep grazing in the field
- In comprehension
 - Margie was sitting on the front steps when she heard the familiar jingle of the “Good Humor” truck. She remembered her birthday money and ran into the house.
- In reaching, grasping, typing...



Graded and variable nature of neuronal responses



Lateral Inhibition in Eye of Limulus (Horseshoe Crab)



MUTUAL INHIBITION results when two neighboring ommatidia are illuminated at the same time (*top*). The inhibition is exerted by cross connections among nerve fibers. When ommatidia attached to fiber *A* and fiber *B* were illuminated separately, 34 and 30 impulses were recorded respectively in one second. Illuminated together, they fired less often.

The Interactive Activation Model

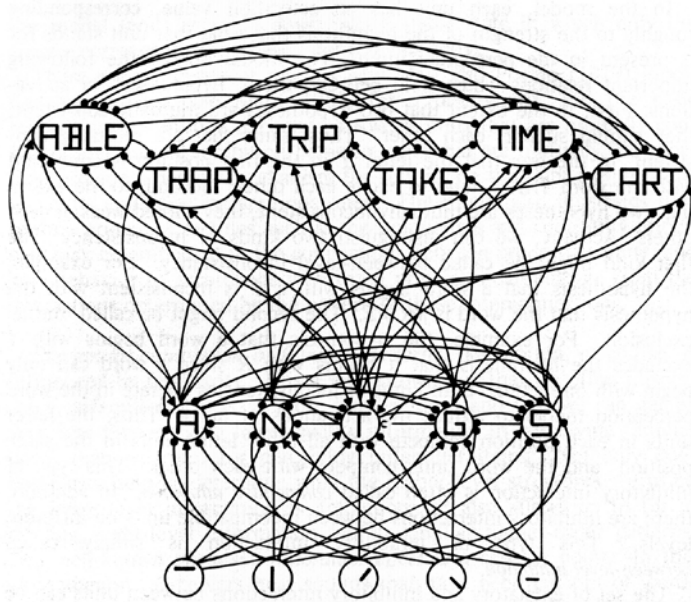


FIGURE 7. The unit for the letter *T* in the first position of a four-letter array and some of its neighbors. Note that the feature and letter units stand only for the first position; in a complete picture of the units needed from processing four-letter displays, there would be four full sets of feature detectors and four full sets of letter detectors. (From "An Interactive Activation Model of Context Effects in Letter Perception: Part I. An Account of Basic Findings" by J. L. McClelland and D. E. Rumelhart, 1981, *Psychological Review*, 88, p. 380. Copyright 1981 by the American Psychological Association. Reprinted by permission.)

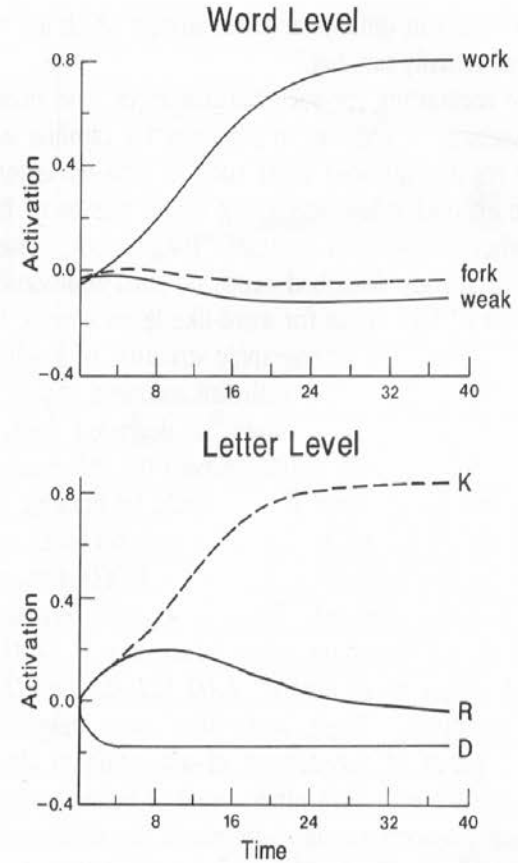


FIGURE 8. A possible display which might be presented to the interactive activation model of word recognition, and the resulting activations of selected letter and word units. The letter units are for the letters indicated in the fourth position of a four-letter display.

Input and activation of units in PDP models

- General form of unit update:

$$net_i = \sum_j w_{ij} a_j + bias_i + input_i + noise$$

if $net_i > 0$:

$$\Delta a_i = net_i(1 - a_i) - d(a_i - rest)$$

else

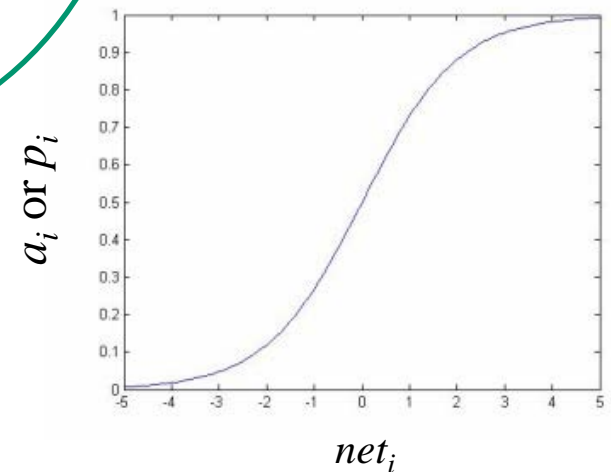
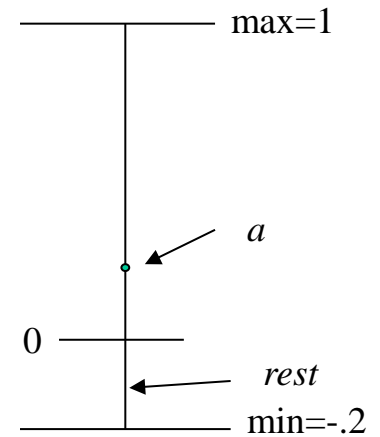
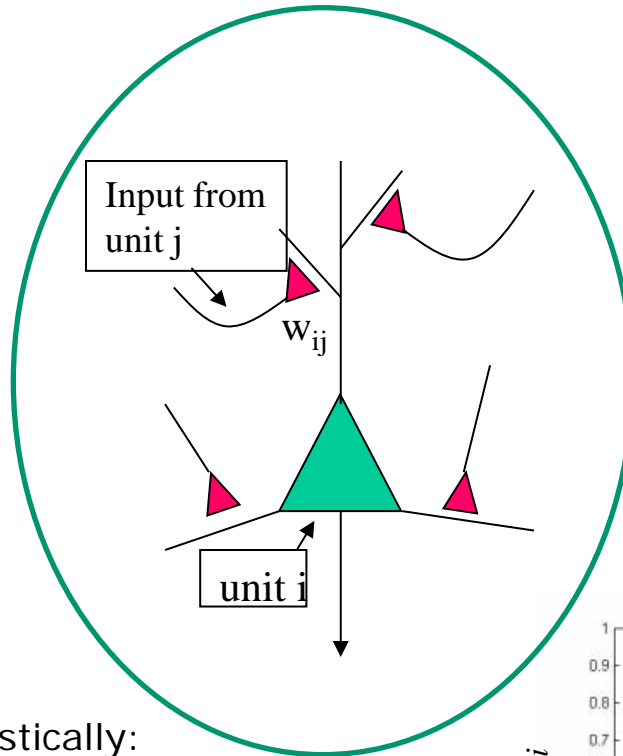
$$\Delta a_i = net_i(a_i - min) - d(a_i - rest)$$

- An activation function that links PDP models to Bayesian computation:

$$a_i = \frac{e^{net_i}}{e^{net_i} + 1}$$

- Or set activation to 1 probabilistically:

$$p_i = \frac{e^{net_i}}{e^{net_i} + 1}$$



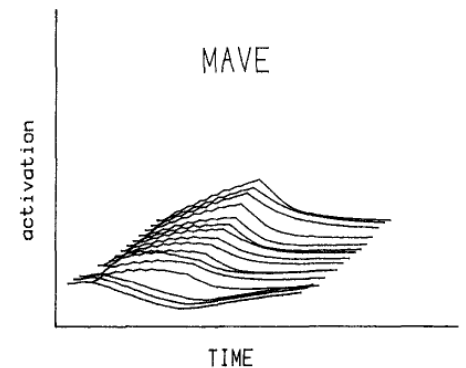
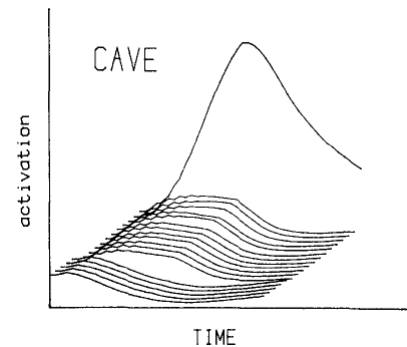
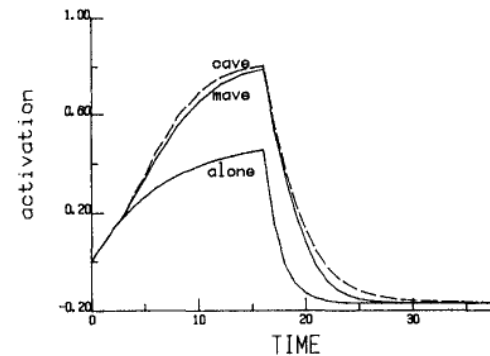
Rules or Connections?

- The IA model only knows rules, but human perceivers show perceptual facilitation when they perceive letters in non-words as well.
- Does our perceptual system follow rules based on a 'grammar' or legal forms?

Syl \rightarrow {Ons} + Body
Body \rightarrow Vwl + {Coda}

- The IA model simulates perceptual facilitation in pseudowords as well as words
- The knowledge is in the connections

a in different contexts



IA Model as a Bridge to a new Framework

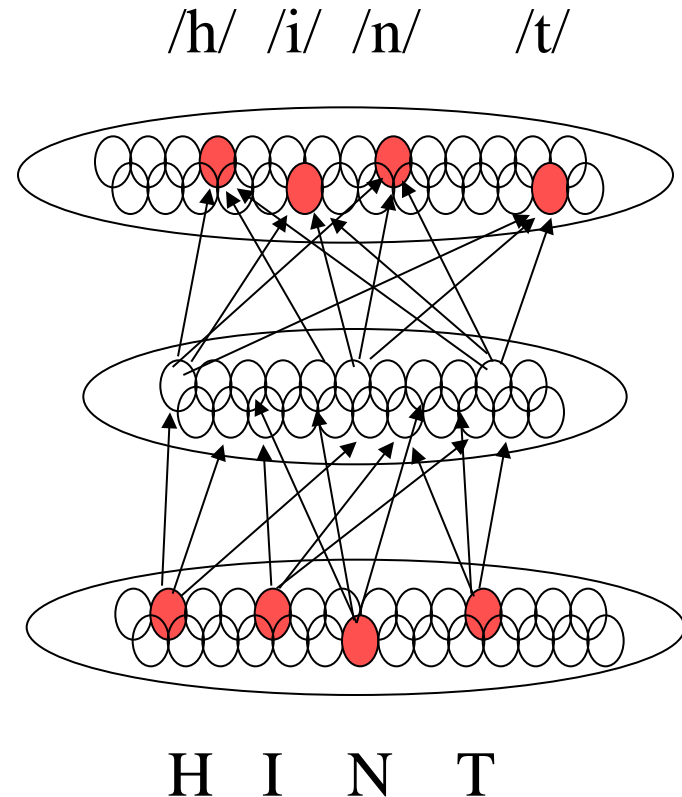
- It is different from the PSS framework in that:
 - Knowledge is in the connections, hence
 - Directly wired into the processing machinery rather than stored as such
 - Patterns are not retrieved by *constructed*
 - Intrinsically inaccessible to inspection
- But it is similar in that:
 - Programmed by its designer
 - Embodies designer's choices about how to represent knowledge
 - Units correspond directly to cognitive entities

Distributed Connectionist Models

- What if we could *learn* from experience, without making prior commitments to the way cognitive entities are represented
 - Do there have to be units corresponding to such entities in our minds?
 - Do we need separate subsystems for items that follow the rules and items that do not?
- Two prominent application areas:
 - Past tense inflection
 - Pay – paid, lay – laid, tay – taid;
See-saw, Say – said, Have – had...
 - Spelling to sound
 - HINT, MINT, PINT

Core Principles of Parallel Distributed Processing Models using Learned Distributed Representations

- Processing occurs via interactions among neuron-like processing units via weighted connections.
- A representation is a pattern of activation.
- The knowledge is in the connections.
- Learning occurs through gradual connection adjustment, driven by experience.
- Learning affects both representation and processing.



Learning in a Feedforward PDP Network

- Propagate activation 'forward' producing a_r for all units using the logistic activation function.

$$a_r = \frac{e^{net_r}}{e^{net_r} + 1}$$

- Calculate error at the output layer:

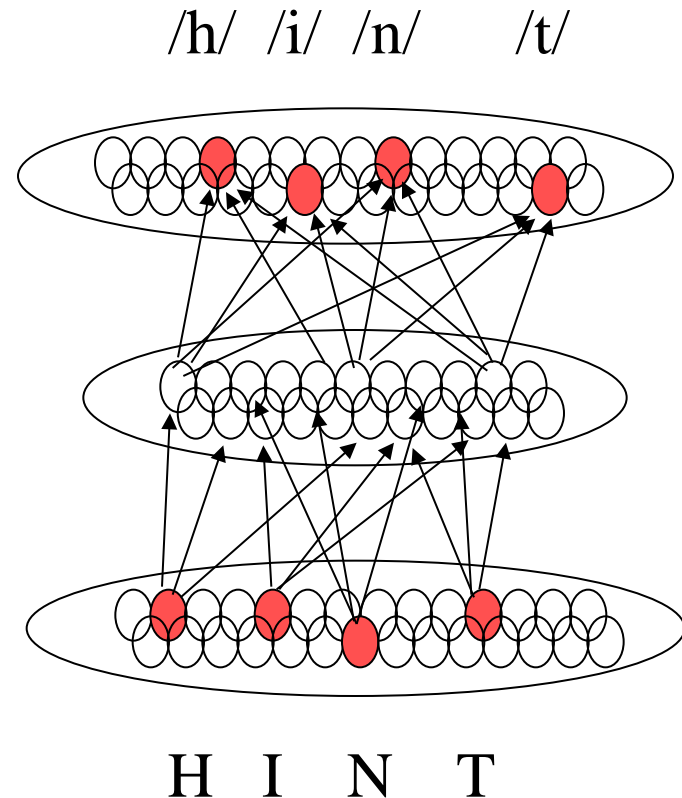
$$\delta_r = f'(t_r - a_r)$$

- Propagate error backward to calculate error information at the 'hidden' layer:

$$\delta_s = f'(\sum_r w_{rs} \delta_r)$$

- Change weights:

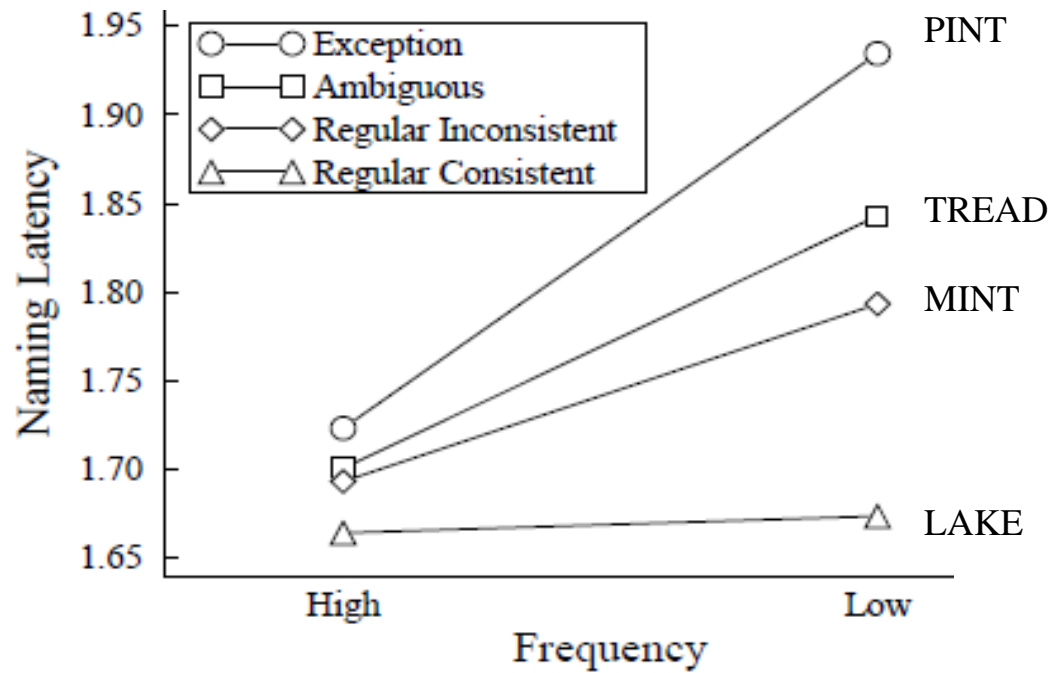
$$\Delta w_{rs} = \delta_r a_s$$



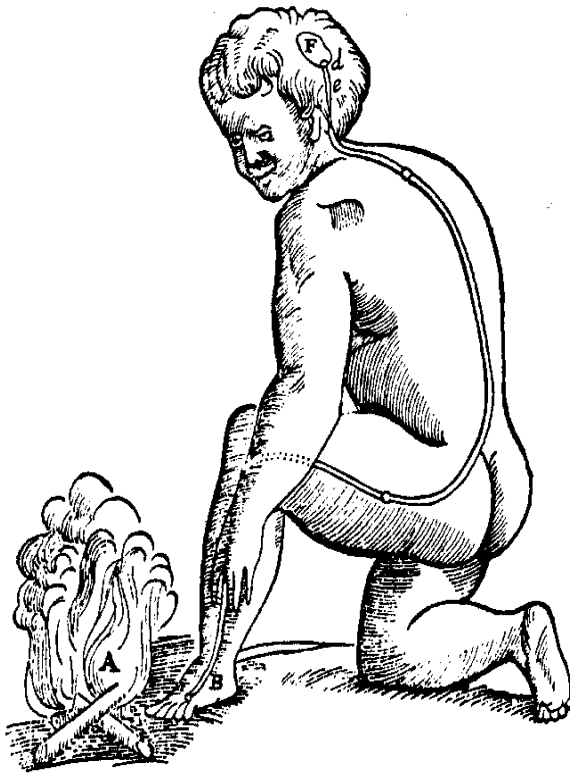
Characteristics of Past Tense and Spelling-sound models

- They use a single system of connections to correctly capture performance with regular, exceptional, and novel items
 - MINT, PINT, VINT
 - LIKE, TAKE, FIKE
- Tend to over-regularize exceptions early in learning as if they have 'discovered' a rule.
- The knowledge in the connections that informs processing of regular items also informs processing of the regular aspects of exceptions
 - Quasi-regularity: The tendency for exceptions to exhibit characteristics of fully regular items
 - PINT, YACHT – said, thought
- Exhibit graded sensitivity to frequency and regularity and a frequency by regularity interaction.

Frequency by Regularity Interaction



Descartes' Legacy



- Descartes' idea that the 'mind' and higher cognition is distinct from the physiological mechanisms of sensation and action persists to a degree to this day.
- Everyone accepts that neural networks underlie sensory and motor processes but are they really relevant to understanding thought?
- Recent and future work with neural networks addressing semantic and mathematical cognition addresses this question.