# Chapter 8

Associative Theories of Long-term Memory

## Nodes and Spreading Activation

- Knowledge represented in connections (associations) between nodes
- *Nodes* = individual concepts or ideas (Do not necessarily correspond to neurons. May be groups of strongly interconnected neurons.)
- Nodes = detectors in feature net
- Learning = making or strengthening connections between nodes
  - Attention needed to both active nodes

## Nodes and Spreading Activation- 2

- Retrieval = reactivating associations
- Not all associations have equal strength
- Associations analygous to 'highways'
- Activation (traffic) spreads along all connections from activated nodes (cities)

## Nodes and Spreading Activation - 3

- Nodes like detectors in that they receive activation until threshold is reached & node “fires” & sends activation to connected nodes
  - Stronger connections → spread more activation.
- *Subthreshold activation* – partial activation insufficient to cause node to fire can *summate* with slightly later input & produce firing

## Nodes and Spreading Activation - 4

- Activation accumulates: a few strong inputs can activate a node, or many weak inputs can summate to achieve activation
- Activation = attracting attention to node
- Get priming effect: prior input raises subthreshold activation level
  - Weak input can activate primed node
- Nodes are not neurons, more likely collections of connected neurons (Hebb cell assembly)
- Nodes in LTM include detectors

## Nodes and Spreading Activation - 5

- Connections between nodes vary in strength
- Spreading activation is divided among all connections according to connection strength. Highway analogy.
- Activation can spread from all active nodes → nodes can be activated from more than one pathway.
- Assumption: Response of node can be larger or smaller depending on recency, frequency etc.
  - Some nodes stronger = have more activation to spread

## Evidence for the Network Model

Hints
- In free recall giving categories (for categorized word list) or words associated with target words at input improves recall
- Hints must have been associated with target info. during learning.

State-dependent Learning, Context reinstatement
- Recall thoughts, feelings at time of learning. Not necessary to be in same mental state or same location
- Activation from multiple weak nodes (hints) may be sufficient to activate target node.

## Evidence for the Network Model - 2 Mnemonics

- Establish indirect but strong connections
  - One → bun → (visual image) →target word
- Few retrieval pathways, so activation does not spread out
- Retrieval slow but accurate if indirect connections used; e.g. retrieving from 4th word in list to 5th or recalling backwards.

## Experimental Evidence

- Lexical Decision Task
  - Show Ss real words or pseudowords (e.g. *lorse* or *clume*
  - S indicates 'yes' if words was word shown or 'no'
- Meyer et al. presented pairs of letter strings: 2 words vs. 2 pseudowords vs 1 word + 1 nonword
  - 2 words, associated or not

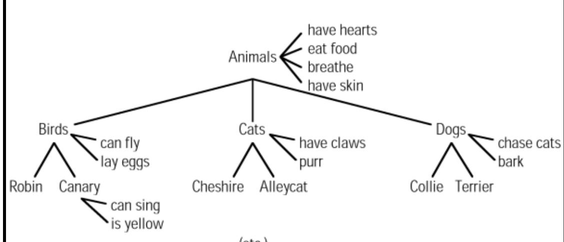
## Lexical Decision Task

- S reads first word, activates relevant node, activation spreads to related nodes
- If second word is associated to first word, it will be activated
- Get faster response to 2nd word if it is associated to first
- Faster reponses to 'white' in 'black -- white' than 'white' alone or 'clome– white'

## Sentence Verification

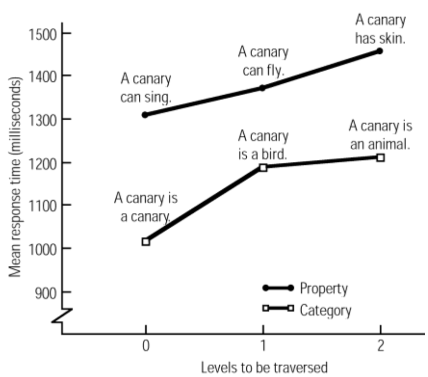
- Present sentences:
  - A robin is a bird. vs. A robin is an animal.
  - A cat has claws. vs. A cat can breathe.
  - A canary has fins. (False)
- Suppose memory is hierarchical with categories & subcategories
  - Properties stored at highest level possible
    - *Principle of Nonredundancy*
  - E.g. Animals breathe, have nervous systems, etc
  - Properties specific to cats, stored at 'cat' level, not at 'animal' level

## Sentence Verification - 2



## Sentence Verification - 3

- Prediction: faster responses if Ss have less "distance" to travel in order to verify or falsify sentence
- 'A robin is a bird,' should be faster than 'A robin is an animal.'
- Results: The more connections that need to be traveled, the slower the reaction time
- Note: Not all connections the same strength




## Typicality Effects

- Get faster response to 'A robin is a bird,' than to 'An emu is a bird.'
- → stronger connections for typical exemplars
- Get faster response to 'A peacock has feathers,' than to 'A robin has feathers."
- → Feathers more associated directly with peacocks than robins as well as being associated with category 'birds'.
  - Nonredundancy principle violated

Is LTM Storage Strictly Hierarchical?

- Objects categorized in different ways
  - Tomato - vegetable
    - fruit
    - pizza topping
    - used in sandwiches
    - red object
    - nutritious
- We now think LTM storage is network, not hierarchy

## Fan Effects

- We know a lot about some objects (e.g. robins) and much less about others (e.g. mockingbirds, aardvarks)
  - Colour? Food? Where do they live? What do their vocalizations sound like?
- More associations to 'robin' than to 'mockingbird'
- *Fan* refers to the number of associations extending from a node

## Fan Effects - 2

- Total amount of activation is limited; activation must spread through all connecting nodes.
- The more nodes there are, the less strongly individual connecting nodes are activated
- Prediction: if two nodes differ only in the number of radiating connections (but not in the strength of the connections or the total activation from the node)
  - → faster responses for node with fewer connections
- Robin node activated frequently → connections strong
- → Two factors working in opposite directions

## Fan Effects: Anderson

- ☐ Anderson taught Ss sentences about people & locations.
- ☐ Actors appeared in 1 or 2 locations
- ☐ Locations had 1 or 2 people
- ☐ Amount of learning controlled. Number of connections to and from each node controlled.
- ☐ Ss memorized sentences and were given recognition test.
- ☐ Response latency on recognition test depended on number of sentences about a person or location.

## Fan Effects: Anderson - 2

- Anderson, 1974 – Response Times

| | No. sentences about person | |
|---|---|---|
| No. sentences about place | 1 | 2 |
| 1 | 1.11 | 1.17 |
| 2 | 1.17 | 1.22 |

## Searching Networks

- Compare searching neural network to
  - 1) encyclopedia
  - 2) Internet
- Encyclopedia has Table of Contents (items in serial order), index (items are alphabetical)
- Internet: Find sites with "key words"
  - Websites and encyclopedia entries often point to other useful sites (hyperlinks)
  - Need a way to rank relevance of various associations
- Importance of spread of activation from more than one starting point

## Entry Nodes

- Questions – words in questions activate nodes, activation spreads → retrieval of target information
- Free associations to environmental cues
- Can be perceptions – entry from feature detectors (See Chapt. 3)
  - Detectors are nodes in the network linked to sense organs
  - E. g. smell reminds you of a person or place or episode
- Can be other thoughts – free association
- Feature nets directly linked to memory networks

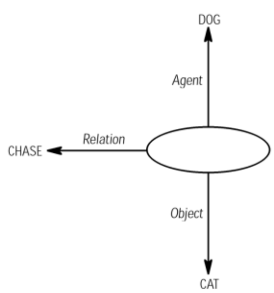
## Different Types of Associative Links: Anderson's ACT-R Model (1976, 1980, 1993)

- Nodes represent simple concepts (e.g. dog, cookie)
- Connections: *isa*, *hasa* represent different relationships (identity or equivalence, possession)
  - Connections represent syntactic role within the proposition
- Basic unit: proposition = smallest unit of knowledge that can be true or false
  - Proposition can be represented as a sentence

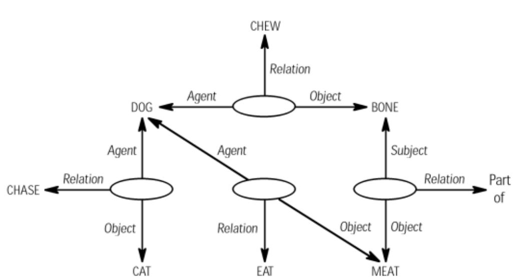
## Anderson's ACT Model - 3

- Ellipse represents the proposition
- Associations indicate syntactical relationships between elements
  - Agent or doer (noun)
  - Recipient of action, object (noun)
  - Relation or action (verb)

## Anderson's ACT Model - 2



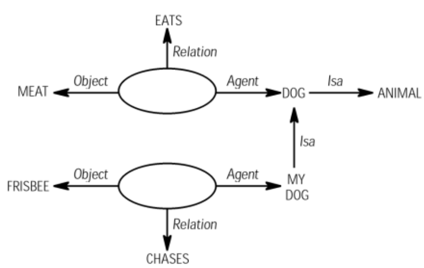

## Anderson's ACT Model - 4



## Anderson's ACT Model - 5

- Type vs. Token nodes
  - Type nodes: generic, general category (e.g. dogs)
  - Token nodes: specific instance of a category (e.g. my dog)

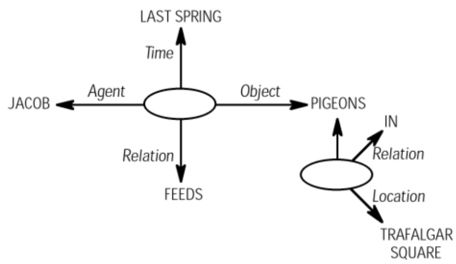
## Anderson's ACT Model - 6




## Anderson's ACT Model - 7

- Time and Location nodes (adjectival phrases)
  - Distinguish between timeless truths (generic knowledge) and specific episodes (episodic memory) or knowledge about specific objects rather than all objects in a class

## Anderson's ACT Model - 8




## Logic of Simulation

- Implement model on a computer.
- Have model learn some information.
- Give model questions to answer.
- Compare model's performance to performance of humans.
  - Same type of errors?
  - Variables produce same effects?
- Change assumptions to see when model works best.

## Tip of the Tongue (TOT)

- ☐ Instrument used to calculate positions of celestial bodies.  Invented before the sextant.
- ☐ Candelabrum used in Jewish worship
- ☐ Legendary warrior & hero of English poem
- ☐ Another name for camel.
- ☐ Russian sled pulled by 3 horses

## Tip of the Tongue - 2

- Can't recall target word, but often know first letter, number of syllables, stress pattern, some of the sounds
- We are in general memory area & there is lots of activation.
- Why can't we retrieve the word? Why doesn't activation spread to target node?
- Evidence against network model ???

## Finding Distant Connections – Weaknesses of Network Models

- Many connections to and from individual concept
- What happens when there are many connections from a node?
  - Activation of any single node very weak.
- How far does activation spread from starting node?
  - If activation spreads beyond first node → too many nodes activated weakly

## Inhibitory Effects – Winner Takes All

- Need mechanism for selection of desired nodes. Want to inhibit irrelevant nodes.
- Node 1 connected to Nodes A to Z. Connections vary in strength.
- Assume inhibitory connections as well as excitatory→ the more active a node is the more it will inhibit a competing node node (Recall Lateral Inhibition)

## Inhibitory Effects – Winner Takes All - 2

- Node M strongest → inhibits weaker nodes
- Inhibitory connections will decrease activation in weakly associated nodes.
- Ultimately only most strongly connected node will be activated.
- Relevant node(s) receive activation from other sources (e.g. context)

## Winner Takes All – Retrieval Blocks & Incubation Effects

- In TOT task, can retrieve related word which you know is not the target.
- Activation of non-target is high & blocks retrieval of target.
- Need to let activation decline (work on something else) & try to retrieve target later
- Relevant cues help
- Get reminiscence, incubation effect

## Connectionism: Parallel Distributed Processing

- Anderson's ACT model – local representations (nodes) represent concepts or propositions
- PDP models: nodes have no meaning
  - Meaning is in *pattern* of connections and active nodes → distributed representation
- Widespread activation pattern can evoke another widespread activation pattern
- → parallel activation spreading throughout brain

## Connectionism: Parallel Distributed Processing - 2

- Learning = changing connection weights
- Thinking = nodes currently active
- Knowledge = how activation would flow
- After learning, activation flow is different than before. (new responses or associations)
- No central executive, but activation patterns satisfy *simultaneous multiple constraints*

## Connectionism: Parallel Distributed Processing - 3

- Simultaneous Multiple Constraints
- Retrieval cue activates certain nodes
- Activation spreads to related nodes
- Target information – what best satisfies all of the currently active nodes
- Review discussion of feature nets in Chapter 3
- Word *clock* activated because component bigrams activated.

- Demo: Ladle rat rotten hut

## Connectionism: Learning

- Small amount of learning → adjustment in many connection weights
- Adjustments must be made locally & in parallel throughout brain.
- No Central Executive to coordinate
- Learning algorithms depend on Contiguity
  - Iif two 'nodes' often active or inactive together → get increase in connection weights → activity in one 'node' increases activity in the other

## Back Progagation

- Learning depends on Feedback
  - Error signal from external source sends activation backwards to nodes activating the incorrect response
  - Activation spreads backwards through network to weaken inappropriate links
- Is there a parallel process in the brain? Does negative feedback decrease connection weights?

## Limits to Connectionist Models

- Make biological sense:
  - neurons & nodes fire in all-or-none manner
  - excitatory & inhibitory connections
  - Summation of activation, thresholds
  - bidirectional connections (top down & bottom up)
  - Learning = adjustment of connection strength = changes in synaptic transmission
  - Brain is a parallel processor
  - Brain uses "divide-and-conquer" strategy
    - Different parts of brain deal with different tasks

Limits to Connectionist Models - 2

- Connectionist models can learn skills & generalize appropriately.
  - Learn English grammar, learn to read, can play games etc.
- Learning very slow in connectionistic models; learning often very fast in humans & animals.
- Connectionist models need help to learn. The right examples must be used. Not true for humans (I disagree)
- Existing models very specific & limited.