Word Processing

ELL 788

7/11

Words

The little star's beside the big star

- Grammar : a system of rules for producing meaningful sequences in language
- Lexicon: a large collection of symbols representing concepts
- Language users require both in order to make meaning
- Lexicon in particular is stored in long-term *memory*

Mental lexicon

Mental lexicon as a collection of words in memory, containing information about its sound (auditory form) and its meaning (semantic form)

- Lexical representation/organization
- Lexical access/retrieval

Words in the brain

- We are interested in the mental lexicon, its organization, its usage in real time.

- How do we mentally represent word forms?
- How are those representations organized?
- How are word meanings represented in the mind?
- How do we go about searching our memories for a matching form?
- . . .

Two types of studies

- Lexical decision task: used since the 1970s
- Targets our ability for word recognition from memory

- cat

- drom

- pillow

- straw

- fleg

- claw

- gorp

- tiger

- glor

- stripes

Lexical decision

Factors affecting lexical decision task

Frequency of the word
Semantic relatedness of the words (priming)

Lexical semantics

Organization of lexical items

- sound: gave vs. cave

Lexical semantics

Organization of lexical items

- sound: gave vs. cave
- visual cues: wow vs. mow

Lexical semantics

Organization of lexical items

- sound: gave vs. cave
- visual cues: wow vs. mow
- meaning: horse vs. donkey ✓

Word representation

- How are meanings represented? How to write a dictionary entry?
- → list of core characteristics
- eg. Bachelor → "human, "adult, "male and "unmarried
- there are many characteristics that can be included; what should remain? what should be left out?
- Core or essential properties: those things that make up the essence of an object
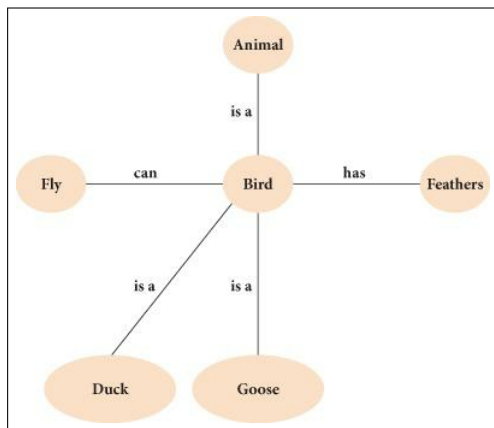- Discriminate between cats and other kinds of things; cat vs. everything else

Word representation

- What do we do for such words?
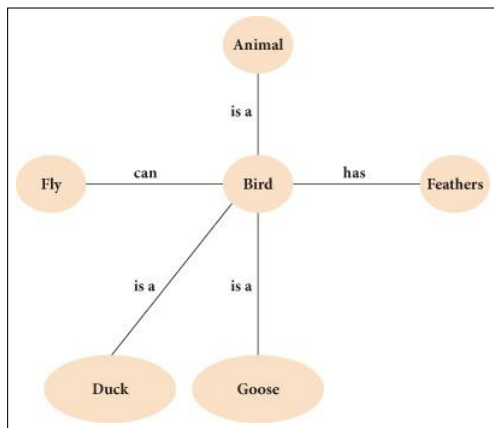- the "core feature" approach to lexical semantics is problematic

Exercise

- Think of words associated with the noun 'engineer'

Semantic network

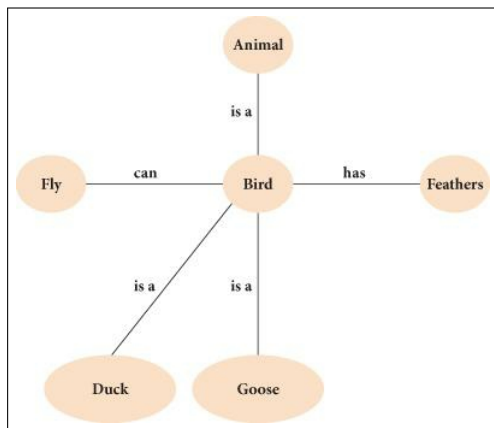- Organizing word meaning as collection of associated concepts

Semantic network

- Organizing word meaning as collection of associated concepts
- Semantic network: whatever comes to mind when someone says the word
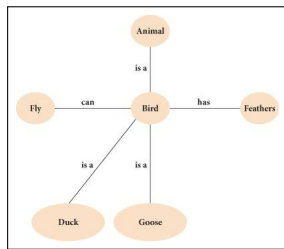
Semantic network

- Organizing word meaning as collection of associated concepts
- Semantic network: whatever comes to mind when someone says the word
- to explain certain behavioural pattern, and encoding of meaning

Semantic network
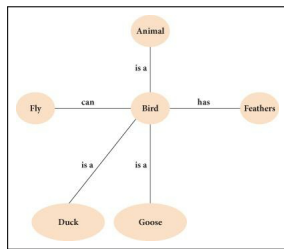
- Word's meaning represented via sets of nodes and the links between them (Quillian, 1968)

Semantic network

- Word's meaning represented via sets of nodes and the links between them (Quillian, 1968)
  - ▶ Node: concepts that a word represents,
  - ▶ links: relationship between concepts

Collins and Quillian (Sentence Verification Task)

Collins and Quillian (Sentence Verification Task)

- A crow is a bird

Collins and Quillian (Sentence Verification Task)

- A crow is a bird
- A crow can fly

Collins and Quillian (Sentence Verification Task)

- A crow is a bird
- A crow can fly
- A crow has skin

Collins and Quillian (Sentence Verification Task)

- A crow is a bird
- A crow can fly
- A crow has skin
- A crow can sing

Collins and Quillian (Sentence Verification Task)

- A crow is a bird
- A crow can fly
- A crow has skin
- A crow can sing
- A crow lays eggs

Collins and Quillian (Sentence Verification Task)

- A crow is a bird
- A crow can fly
- A crow has skin
- A crow can sing
- A crow lays eggs
- A crow is an animal

Sentence Verification Task

- More time to retrieve the concept *A crow is an animal* or *A crow has skin*
- In contrast, *A crow can fly* or *A crow is a bird* is faster

Sentence Verification Task

- More time to retrieve the concept *A crow is an animal* or *A crow has skin*
- In contrast, *A crow can fly* or *A crow is a bird* is faster
- Each item in the tree for *crow* is connected to its superset and its subset
- Slower responses indicate searching through more links in the hierarchy

Hierarchy

- *goose*: an address in memory
- *is a*: encodes relationship between general categories and concepts that fall under the category
- subordinate categories inherit properties of superordinate nodes via transitive inference (a goose is a waterfowl, a waterfowl is a bird, therefore a goose is a bird)
- no need to connect goose to bird; helps conserve memory resources (cognitive economy)

Semantic network: Spreading activation

- A hypothetical mental process that takes place when one of the nodes in the semantic network is activated
- So, when we hear or read *goose*, the *goose* node is activated. The activation from this node then 'spreads' to other nodes that are connected to this node
- meaning of a word = activated nodes and links

Semantic network: Spreading activation

- Spreading activation helps explain how respond during *priming* tasks
- Priming: presentation of an item at time 1 helps in responding to another item at time 2

Spreading activation

- Mediated priming studies use word pairs like *lion-stripes*
  - lion → tiger → stripes
  - but experiments show *lion* does not prime *stripes*
  - showing that spreading activation is limited
- But why should this be so?
- Diminishes beyond one or two links
- Total amount of activation that can be spread across the network is limited

Lexical network: WordNet

- Hierarchically arranged semantic network
- In WordNet, the most frequently occurring synset will show up first

Noun

- (28)S: (n) **board** (a committee having supervisory powers) *"the board has seven members"*
- (18)S: (n) **board**, plank (a stout length of sawn timber; made in a wide variety of sizes and used for many purposes)
- (4)S: (n) **board** (a flat piece of material designed for a special purpose) *"he nailed boards across the windows"*

Word sense disambiguation

- S: (n) **bank** (sloping land (especially the slope beside a body of water)) *"they pulled the canoe up on the bank"; "he sat on the bank of the river and watched the currents"*
- S: (n) depository financial institution, **bank**, banking concern, banking company (a financial institution that accepts deposits and channels the money into lending activities) *"he cashed a check at the bank"; "that bank holds the mortgage on my home"*

- Lesk approach: take as input a word (+context) + WordNet senses, measure overlap in the senses
- Traversal based approach, using similarity measures based on path length, e.g. shorter the path, greater the similarity

Problems with semantic network

- Groundedness
- the meaning of a word can only be understood with reference to other words
- the meaning of these words is not grounded in a representation outside the system of symbols
- Chinese room problem
- Networks themselves may differ from one individual to another

Embodied semantics

- abstract symbols are tied to representations *outside* the linguistic system
- These representations are **sensory** in nature: vision, hearing, taste, touch and smell
- When we think "cat", we "perceive" a cat once again, using the same apparatus
- Perhaps these representations can solve the problem of symbol grounding

Embodied semantics

- Glenberg and Robertson (2000) refer to this as *indexical hypothesis*
- Words are indexed to actual objects in the world (or analog representations in the mind)
- Indexed objects are then used to derive affordances (interaction of our perceptual beliefs and physical properties of objects)
- Finally the index and affordance are meshed together to give us an utterance

Links between semantics, perception and motor control

- Glenberg and Kaschak (2002) on the link between semantics and the motor system
- Human subjects would be asked to rate sentences like *He opened the drawer* or *He closed the drawer*
- Subjects would answer with a gesture either towards the body or away from it
- If the gesture and sentence were congruent, they responded faster
- Suggests the role of mental simulation of the event

Links between semantics, perception and motor control

- Zwaan et al (2002): Named objects accessed with much visual detail
- subjects read sentences, then make judgements about an image of an object mentioned in the sentence
- Orientation of objects mentioned in the sentence affected this judgement task
- E.g. *He hammered the nail into the floor* vs *He hammered the nail into the wall*
- As the former example is incompatible with the action of hammering– judgement tasks took longer

Links between semantics, perception and motor control

- Tucker and Ellis (2001): action in the motor system facilitates response to a word
- Subjects judged whether an object was natural or man-made
- Some objects required a precision grip (pen, button, fork) or power grip (shovel,axe)
- Precision-grip or power-grip response device for a semantic categorization task
- When the object and response device was congruent, judgements were faster