Parsing

PSY 200

Greg Francis

Lecture 28

This is a valid sentence:

"Buffalo buffalo buffalo buffalo buffalo buffalo buffalo buffalo."

Linguistics

- So far we have looked at
  - grammar
  - words
- From the point of view of *generating* a sentence
- Today we look more closely at *interpreting* a sentence

Eliza / Emacs

- Weizenbaum (1966)
- *I am the psychotherapist. Please, describe your problems. Each time you are finished talking, type RET twice.*
- I am sad.
- *Why do you say you are sad?*
- My parents won't give me a new car.
- *Tell me something about your family.*
- I have five brothers and two sisters.
- *You seem to dwell on your family.*
- They are important.
- *Is it because of your plans that you say they are important?*

Eliza / Emacs

- *I am the psychotherapist. Please, describe your problems. Each time you are finished talking, type RET twice.*
- Hi, my name is Greg.
- *How do you do? What brings you to see me?*
- I am showing a demonstration for class.
- *Why do you say you are showing a demonstration for class?*
- Because I am.
- *Can you elaborate on that?*
- Well, I am a professor.
- *Tell me about your professor.*
- I am the professor!

Eliza just picks up on key phrases and generates some stock responses.

Parsing

- Eliza does *not* really *understand* anything
  - Try it yourself at
  - http://www.manifestation.com/neurotoys/eliza.php3
- The difference between
  - *Dog bites man.*
  - *Man bites dog.*
- Requires identifying the subject, object, and verb
- The system that does this is called a *parser*

Who does what?

- You can learn quite a bit about a sentence's meaning by knowing the phrase tree *structure* of the sentence
  - indicates some aspects of *meaning*

The green idea eats the girl's candy.

- We know the sentence is about an idea rather than a girl
  - we also know the idea is doing the eating

Building phrase trees

- You *create* a sentence with ideas in your head
- Those ideas are converted into appropriate words and phrase trees to convey those ideas
- Sometimes two different ideas can give rise to the same sentence
  - leads to ambiguous sentences
  - the parser does not work in the same way as the creator

Ambiguous sentences

- Consider the following

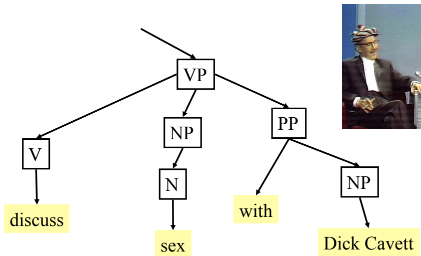
I saw a man on a hill with a telescope.

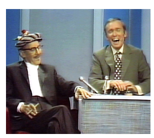
Two cars were reported stolen by the Purdue police yesterday.

Tonight's program discusses stress, exercise, nutrition, and sex with former Celtic forward Scott Wedman, Dr. Ruth Westheimer, and Dick Cavett.

Ambiguous sentences

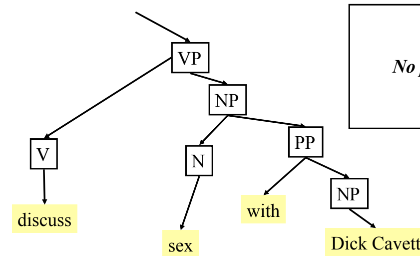
- The writer had in mind a phrase tree like





Ambiguous sentences

- But a reader/listener could interpret it like



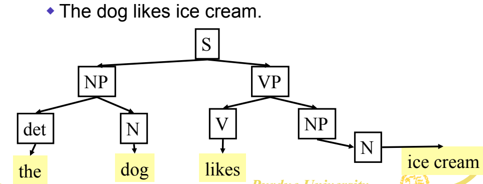
No photo

Mentalese

- That two different internal thoughts can give rise to the same language statement is interesting
  - it suggests that we *think* in some way that is different from language
  - a *mentalese*, if you will

Parsing

- Parsing is something like building a phrase tree in reverse
- Let's parse through a simple sentence word by word
  - The dog likes ice cream.



Parsing

- Once every slot is filled, the sentence is parsed
  - a mental "click" of understanding
- Each word has its role defined
  - and the order of the phrases identifies the meaning (usually)

Purdue University

Two problems

- Parsing is complicated in two ways
  - (1) Phrases are not always consistent with word order
  - (2) The same spoken sounds are sometimes used for words with different meanings (noun vs verb vs adjective)

Purdue University

Word order

- This sentence is relatively easy to parse, even though it is a complicated sentence

Remarkable is the rapidity of the motion of the wing of the hummingbird.

Purdue University

Word order

- This sentence is not as easy
- One type of phrase is embedded in another

The rapidity that the motion has is remarkable.

Purdue University

Word order

- This sentence is nearly impossible

The rapidity that the motion that the wing that the hummingbird has has has is remarkable.

Purdue University

Difficult sentences

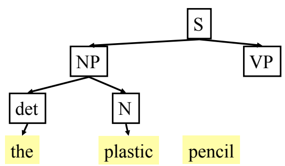
- These sentences are difficult for humans because of limited memory
  - when a phrase tree includes many unfilled branches of the same type (PP)
  - the parser becomes confused as to which phrase is associated with a new word
  - ends up backtracking to sort out the phrases
  - sometimes falls apart ("has has has")
- The grammar generator and the parser are different things in your language system
  - these are grammatically correct sentences
  - they are not *good* sentences
  - *you* make sentences like these

Don't make me show you your exams!

Purdue University

Word ambiguity

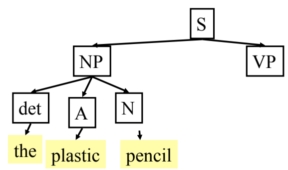
- A word by itself is often ambiguous
- Consider a parser trying to follow the phrase
  - The plastic pencil marks...



Word pencil is inconsistent with structure created!

Word ambiguity

- A differently designed phrase tree handles the new word
  - The plastic pencil marks...



Word pencil is consistent with structure created!

Word ambiguity

- But you run into the same problem with the word "marks" (noun or verb?)
  - The plastic pencil marks were ugly. (noun)
  - The plastic pencil marks easily (verb)
- Parsers build phrase trees on the fly, so backtracking is often required
  - many times it is so fast that we do not notice
  - seems effortless

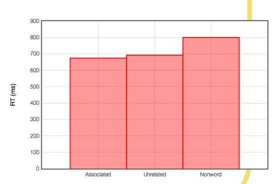
Lexical decision

- It is not effortless and it can be shown with an experiment
- The experiment is a variation of the lexical decision task, which you did in CogLab
- In the lexical decision experiment, you see a sequential pair of words/non-words, and we measure the reaction time for you to decide if the second "word" is a word
  - RT is faster if the second word is semantically related to the first word

Lexical decision

- The data find that RT is faster if the second word is *meaningfully* related to the first word
- CogLab Data (69 participants)

Condition	Reaction time (ms)
Associated words	675
Unassociated words	693
Nonwords	801



Evidence of ambiguity

- We can apply the lexical decision task to the ambiguity of parsing (Swinney, 1979)
- Consider the following paragraph, which subjects listened to
  - Rumor had it that, for years, the government had been plagued with problems. The man was not surprised when he found several spiders, roaches, and other bugs in the corner of his room.
- The word *bugs* is ambiguous
  - insects vs surveillance devices
  - Although the context makes one interpretation more reasonable

Evidence of ambiguity

- No one notices the ambiguity
- But, give a lexical decision test for words verses non-words
  - Flashed visually on a screen just after the word was spoken
  - Subjects respond faster for words related to *either* definition of bug

ant	sew	spy
fastest	slowest	in between

Sentence ambiguity

- Interestingly, people often miss ambiguities in sentences
  - Time flies like an arrow.
- Humans recognize only one interpretation
- Computer algorithms can find 5 interpretations
  - all grammatically correct!

Sentence ambiguity

- Time flies like an arrow.
- (1) Time proceeds as quickly as an arrow proceeds.
- (2) Measure the speed of flies in the same way that you measure the speed of an arrow.
- (3) Measure the speed of flies in the same way that an arrow measures the speed of flies.
- (4) Measure the speed of flies that resemble an arrow.
- (5) Flies of a particular kind, time-flies, are fond of an arrow. (*Fruit flies like a banana.)*

Ambiguity and computers

- Or consider the following (valid) sentence that computer algorithms can correctly interpret
  - Buffalo buffalo buffalo buffalo buffalo buffalo buffalo buffalo.
- Here's a hint to make it understandable in principle

Significance

- These types of results suggest that words and grammar are *not* enough to insure communication
- In a certain sense a speaker and listener must already be *agreeing* about the topic before anything can be communicated
- Thus, we can understand the following discourse
  - Woman: I'm leaving you.
  - Man: Who is he?

Schemas / scripts

- Cognitive devices
  - describe stereotypical properties of a situation
  - e.g., restaurant scene involves table, waiter, drinks, tips,…
- Fill-in the missing information that is critical for understanding language (and events in general)
  - explains why it is difficult to communicate across cultures, even with a common language
- Schemas provide the context to remove the almost constant ambiguities of language

Schemas / scripts

- Giving computers the general "knowledge of life" needed to create something like schemas is very difficult
- This is why computers do not carry on conversations with you
- Lots of work going on in artificial intelligence to address this problem

Conclusions

- Understanding language
- Parsing
- Phrase trees (in reverse)
- Ambiguities
- Computer generated interpretations
- Missing information / schemas

Next time

- Speech
- Phonemes
- Articulation / coarticulation
- CogLab on Categorical Perception – Discrimination due
- *Why do we say "razzle-dazzle" instead of "dazzle-razzel"?*