Human Language Parsing

- We understand language incrementally, word-by-word
- How precisely do people construct interpretations?
- We must resolve local and global ambiguity
- How do people decide upon a particular parse/interpretation?
- Why are some grammatical sentences so difficult to comprehend
- Ambiguity: the human parser has been misled, takes time to recover
- Resources: the human parse has exceeded available memory
Human Language Parsing

- We understand language **incrementally**, word-by-word
- How precisely do people construct interpretations?
- We must resolve local and global **ambiguity**
- How do people decide upon a particular parse/interpretation?
- Why are some grammatical sentences so difficult to comprehend?
  - Ambiguity: the human parser has been misled, takes **time** to recover
  - Resources: the human parse has exceeded available **memory**

The Problem

- How do people recover the meaning of an utterance in real-time?

“**The man held at the station was innocent**”

(Reduced) Relative Clauses

- One of the most famous (English) syntactic ambiguities!
- *The man delivered the junkmail* ... threw it away.

**Main clause analysis**

```
  S
   NP  VP
     The man  delivered NP
                the junkmail
```

**Reduced relative clause analysis**

```
  S
   NP  VP
     The man OP_i NP_i
     delivered  the junkmail
```

Reading time studies

- We can use controlled experiments of reading times to investigate local ambiguity resolution
  - (a) The man held at the station was innocent (LA)
  - (b) The man who was held at the station was innocent (UA)

- Compare the reading times of (b) where there is no ambiguity, with (a) to see if and when the ambiguity causes reading difficulty.

- Need a “linking hypothesis” from theory to measures

- Can then manipulate other linguistic factors to determine their influence on on RTs in a controlled manner
Experimental Design

• Decide on the Factor: e.g. Ambiguity

• Decide on the Levels: e.g. ambiguous & non-ambiguous

• One experimental Item includes all conditions (each Level of each Factor)
  • carefully match conditions within an item, minimal differences
  • create multiple items

• Multiple participants, each sees only one condition of each item

• Then average across items and participants, and do inferential statistics!

Reading Methods

Whole sentence reading times:

The man held at the station was innocent

Self-paced reading, central presentation:

in

Self-paced reading, moving window:

The man held at the station was innocent
But what do RTs tell us, really?

- Suppose you find an increased reading time for one condition, versus another. What does that mean?

- Various things can result in longer reading times:
  - lexical: longer words, infrequent words, implausible words
  - syntactic: memory load, ungrammaticality, disambiguation, revision

- We need our models to make clear the linking to empirical measures:
  - e.g. high RT when dispreferred structure must be recovered

- And, we need to make sure this is what is causing the RT in our experiment

Summary

- People construct an interpretations, word-by-word
  - People must resolve ambiguity
  - Sometimes we must revise our interpretation of the sentence so far

- Reading time measures can tell us about how/when this occurs
  - Different phenomena are found in different measure

- We can design experiments which exploit these methods (and others!) to investigate the underlying processing architectures and mechanisms
Parsing Mechanisms

• Syntactic processing requires a solution to the problem of:
  • How structures are incrementally constructed
  • How local and global ambiguity is resolved

• Incremental Parsing
  • Top-down; Bottom-up; Mixed strategies

• Ambiguity and parsing:
  • Serial (deterministic/non-deterministic)
  • Parallel (bounded/unbounded)

Parsing Algorithms for PSGs

• Algorithms to recover the parse tree for an utterance vary ...
  • left-to-right, head-driven, right to left
  • top-down, bottom-up, mixed
  • deterministic, serial, parallel

• Processing complexity:
  • Time: what time is required to parse a sentence as a function of sentence length, grammar size?
  • Space: how much memory does the parser require?
Bottom-up Parsing

- “The woman reads”

Shift-reduce Algorithm

1. Initialise $Stack = []$
2. loop: Either $shift$:
   
   2.1. Determine category, $C$, for next word in sentence;
   
   2.2. Push $C$ onto the stack;
3. Or $reduce$:
   
   3.1. If categories on the $Stack$ match the RHS of a rule: $LHS \rightarrow Stack$
      
      3.1.1. Remove those categories from the $Stack$;
      
      3.1.2. Push the LHS category onto the $Stack$;
4. No more words to process?
   
   4.1. If $Stack = [S]$, then done;
5. Goto loop
Simple example: Bottom-up

```
the man read  
Det  N  V  NP  VP  
Det  N       
```

Top-down Parsing

"The woman reads"

```
S [S]  S [NP,VP]  S [Det,N,VP]  
NP  VP  NP  VP  NP  VP  Det  N  
the  man  the  woman  the  woman  reads  
```
Top-down Algorithm

1. Initialise $Stack = [S]$
2. If $top(Stack)$ is a non-terminal, $N$:
   2.1. Select rule $N \rightarrow RHS$;
   2.2. pop($N$) off the stack and push($RHS$) on the stack;
3. If $top(Stack)$ is a pre-terminal, $P$:
   3.1. Get next word, $W$, from the input;
   3.2. If $P \ W$, then pop($P$) from the stack;
   3.3. Else fail;
4. No more words to process?
   4.1. If $Stack = []$, then done;
5. Goto $\Rightarrow$ $\Rightarrow$

Simple example: Top-down

```
S
   NP
  Det | N
the man
  VP
   V | NP
  read every
  NP
     Det | N
    every book
```

```
S → NP VP
NP → Det N
VP → V NP
NP → Det N
```
Evaluating top-down & bottom-up

- Are these parsers psychologically plausible?
  - Incrementality:
    - Bottom-up: no
    - Top-down: yes
  - Input-driven:
    - Bottom-up: yes
    - Top-down: no + problems with left-recursion

A Psychologically Plausible Parser

- Left-Corner Parsing
  - Rules are ‘activated’ by their ‘left-corner’

- Combines input-driven with top-down
- There is a ‘class’ of LC parsers
An example LC parse

“The woman read the book”

Simple example: Left-Corner
Evaluating the LC Parser

- Variations: **arc-standard** *versus* **arc-eager**

Affect on ambiguity resolution for arc-eager:

- Commitment to attachments is early, before daughters are completely built