Simple example: Left-Corner
Evaluating the LC Parser

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

![Diagram showing arcs with nodes labeled NP, VP, Det, N, and V]

- Affect on ambiguity resolution for arc-eager:
  - Commitment to attachments is early, before daughters are completely built

Quick experiment

- “The mouse died”

- “The mouse that the cat chased died”

- “The mouse that the cat that the dog bit died”

- “The mouse that the cat that the dog bit chased died”
Incrementality and Memory

• It wasn't incrementality that led to the LC algorithm, but memory load:
  • “The mouse died”
  • “The mouse the cat chased died”
  • “The mouse the cat the dog bit chased died”
    (Or: “The mouse that the cat that the dog bit chased died”)

• Grammatical, not ambiguous, what's the problem?

• Memory load: too high for centre embedding
  • “[The mouse [the cat [the dog bit] chased] died]”

Memory Load in Parsing

• Left-embedding is easy:
  • [[[John’s brother]’s car door]’s handle] broke off.

• Right-embedding too:
  • John believes [Bill knows [Mary said [she likes cats]]]

• Centre-embedding is hard:
  • [The mouse [the cat [the dog bit] chased] died]

• Memory load for parsers:
  • Top-down: LE: hard  CE: hard  RE: easy
  • Bottom-up: LE: easy  CE: hard  RE: hard
  • Left-corner: LE: easy  CE: hard  RE: easy
Evaluating the LC Parser

- Variations: *Arc-standard* versus *Arc-eager*

```
S
  NP
    the  man

V
S
  knew
  S
    NP
      VP
        the dog
```

Summary of Behaviour

<table>
<thead>
<tr>
<th>Node</th>
<th>Arcs</th>
<th>Left</th>
<th>Centre</th>
<th>Right</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top-down</td>
<td>Either</td>
<td>O(n)</td>
<td>O(n)</td>
<td>O(1)</td>
</tr>
<tr>
<td>Shift-reduce</td>
<td>Either</td>
<td>O(1)</td>
<td>O(n)</td>
<td>O(n)</td>
</tr>
<tr>
<td>Left-corner</td>
<td>Standard</td>
<td>O(1)</td>
<td>O(n)</td>
<td>O(n)</td>
</tr>
<tr>
<td>Left-corner</td>
<td>Eager</td>
<td>O(1)</td>
<td>O(n)</td>
<td>O(1)</td>
</tr>
<tr>
<td>People</td>
<td></td>
<td>O(1)</td>
<td>O(n)</td>
<td>O(1)</td>
</tr>
</tbody>
</table>
Comments on Left-Corner

- Mixed data-driven and hypothesis driven approaches
  - Eager corresponds to composition of partial structures
- **Arc Standard**: less ambiguity
  - attach when constituents are complete: safer
  - delayed attachment means more is kept on the stack
- **Arc Eager**: less memory
  - early composition reduces stack growth
  - eager attachments are less bottom-up

Ambiguity in Parsing

- Rule selection: *what if more than one rule can be selected?*
  - Local ambiguity: a parse derivation may fail later
  - Global ambiguity: multiple parses can succeed
- How can we handle local and global ambiguities during parsing:
  - Backtracking
  - Parallelism
  - Determinism
  - Underspecification
Ambiguity in Parsing

- Rule selection: *what if more than one rule can be selected?*
  - Local ambiguity: a parse derivation may fail later
  - Global ambiguity: multiple parses can succeed
- How can we handle local and global ambiguities during parsing:
  - **Backtracking**
  - **Parallelism**
  - Determinism
  - Underspecification

Backtracking Parsers

- Parsing is a sequence of rule selections
- If at one point, more than one rule can be applied, this is called a choice point
  - Make a decision, based on some selection rule
  - If subsequently parsing ‘blocks’, return to a choice point and re-parse from there
- Which choice point to return to?
  - usually the last, why?
  - what other choice point selection rules could be used
Backtracking: an example

Bill reads

Parallel Parsers

- Build parse trees through successive rule selections

- If more than one rule may be applied, create a new parse derivation for each possibility

- Pursue all parses in parallel

- If any of the parses 'blocks', discard it

- Because of multiple local ambiguities, the number of parallel derivation grows exponentially

- Bounded parallelism: pursue a fixed number

- How do we choose which ones to keep?
Parallel: an example

Bill reads

Theories of Sentence Processing

- Explanatory and descriptive goals

- Theories of parsing typically determine …
  - what architecture is assumed: modular? symbolic? …
  - what mechanism is used to construct interpretations?
  - which information sources are used by the mechanism?
  - which representation is preferred/constructed when ambiguity arises?

- Linking Hypothesis: Relate theory/model to observed measures
  - Preferred sentence structures should have faster reading times in the disambiguating region than dispreferred
Garden-Path Theory: Frazier

- What architecture is assumed?
  - Modular syntactic processor, with restricted lexical (category) and semantic knowledge

- What mechanisms is used to construct interpretations?
  - Incremental, serial parsing, with reanalysis

- What information is used to determine preferred structure?
  - General syntactic principles based on the current phrase structure

- Linking Hypothesis:
  - Parse complexity and reanalysis cause increased RTs

The Garden Path Theory (Frazier)

Which attachment do people initially prefer?

```
S
   NP
     PN
       John
   VP
     V
       saw
     NP
       Det
         the
       N
         man
     P
       with
     NP
       the
         telescope
```
First Strategy: Minimal Attachment

**Minimal Attachment**: Adopt the analysis which requires postulating the fewest nodes

NP/S Complement Ambiguity

**Minimal Attachment**: Adopt the analysis which requires postulating the fewest nodes
Second Strategy: Late Closure

**Late Closure:** Attach material into the most recently constructed phrase marker

Well-known local ambiguities

**NP/VP Attachment Ambiguity:**
- “The cop [saw [the burglar] [with the binoculars]]”
- “The cop saw [the burglar [with the gun]]”

**NP/S Complement Attachment Ambiguity:**
- “The athlete [realised [his goals]] last week”
- “The athlete realised [[his goals] were unattainable]”

**Clause-boundary Ambiguity:**
- “Since Jay always [jogs [a mile]] [the race doesn’t seem very long]”
- “Since Jay always jogs [[a mile] doesn’t seem very long]”

**Reduced Relative-Main Clause Ambiguity:**
- “[The woman [delivered the junkmail on Thursdays]]”
- “[[The woman [delivered the junkmail]] threw it away]”

**Relative/Complement Clause Ambiguity:**
- “The doctor [told [the woman] [that he was in love with her]]”
- “The doctor [told [the woman [that he was in love with]] [to leave]]”
Summary of Frazier

- Parsing preferences are guided by general principles:
  - Serial structure building
  - Reanalyze based on syntactic conflict
  - Reanalyze based on low plausibility ("thematic fit")

- Psychological assumptions:
  - Modularity: only syntactic (not lexical, not semantic) information used for initial structure building
  - Resources: emphasizes importance of memory limitations
  - Processing strategies are universal, innate