Models of Human Parsing

Informatics 2A: Lecture 22

Mirella Lapata (slides by Frank Keller)

School of Informatics
University of Edinburgh
keller@inf.ed.ac.uk

November 8, 2011

Reading: J&M, ch. 9 (pp. 350–352), ch. 12 (pp. 467–473), ch. 13 (pp. 491–496).

Human Parsing

So far, we looked at parsing from an engineering perspective. However, humans also do parsing to understand language.

The mathematical and algorithmic tools in this course can be used to analyze human parsing (human sentence processing). This is the domain of psycholinguistics.

Human Parsing

So far, we looked at parsing from an engineering perspective. However, humans also do parsing to understand language.

The mathematical and algorithmic tools in this course can be used to analyze human parsing (human sentence processing). This is the domain of psycholinguistics.

To study human parsing, we need:

- experimental data that tell us how humans parse;

Human Parsing

So far, we looked at parsing from an engineering perspective. However, humans also do parsing to understand language.

The mathematical and algorithmic tools in this course can be used to analyze human parsing (human sentence processing). This is the domain of psycholinguistics.

To study human parsing, we need:

- experimental data that tell us how humans parse;
- cognitive constraints derived from these data (e.g., incrementality, garden paths, memory limitations);

Human Parsing

So far, we looked at parsing from an engineering perspective. However, humans also do parsing to understand language.

The mathematical and algorithmic tools in this course can be used to analyze human parsing (human sentence processing). This is the domain of psycholinguistics.

To study human parsing, we need:

- experimental data that tell us how humans parse;
- cognitive constraints derived from these data (e.g., incrementality, garden paths, memory limitations);
- parsing models (and algorithms that implement them) that respect these constraints;

Human Parsing

So far, we looked at parsing from an engineering perspective. However, humans also do parsing to understand language.

The mathematical and algorithmic tools in this course can be used to analyze human parsing (human sentence processing). This is the domain of psycholinguistics.

To study human parsing, we need:

- experimental data that tell us how humans parse;
- cognitive constraints derived from these data (e.g., incrementality, garden paths, memory limitations);
- parsing models (and algorithms that implement them) that respect these constraints;
- an evaluation of the models against the data.

Incrementality

Parsing: extracting syntactic structure from a string; prerequisite for assigning a meaning to the string.

The human parser builds structures incrementally (word by word) as the input comes in.

This can lead to local ambiguity.

Example:

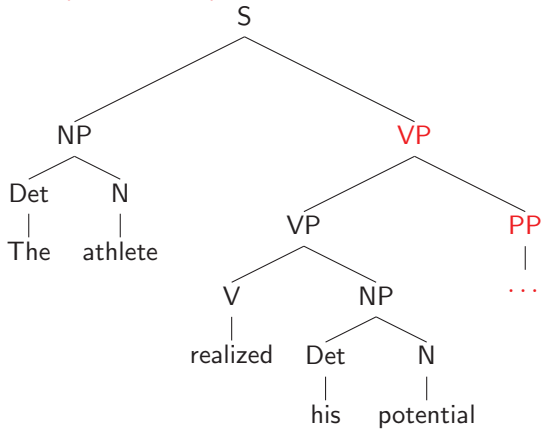
(1) The athlete realized his potential ...

a. ... at the competition.

b. ... would make him a world-class sprinter.

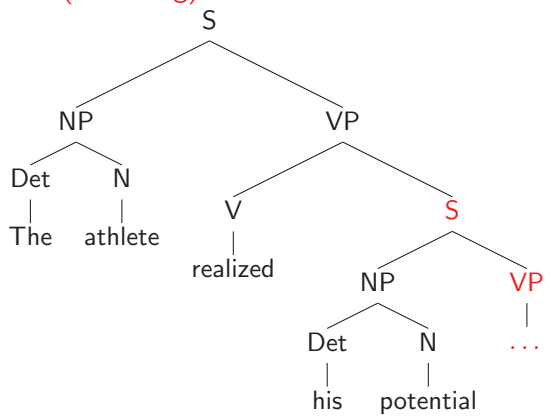
Incrementality

Structure 1 (NP reading):



Incrementality

Structure 2 (S reading):



Garden Paths

- Early commitment: when it reaches *potential*, the processor has to decide which structure to build.
- If the parser makes the wrong choice (e.g., NP reading for sentence (1-b)) it needs to backtrack and revise the structure.
- A garden path occurs, which typically results in longer reading times (and reverse eye-movements).
- Some garden paths are so strong that they parser fails to recover from them.

Garden Paths

More examples of garden paths:

(2)
a. The horse raced past the barn fell.
b. I convinced her children are noisy.
c. Until the police arrest the drug dealers control the street.
d. The old man the boat.
e. We painted the wall with cracks.
f. Fat people eat accumulates.
g. The cotton clothing is usually made of grows in Mississippi.
h. The prime number few.

Dimensions of Parsing

In addition to incrementality, a number of properties are important when designing a model of the human parser:

- Directionality: the parser can process sentence bottom-up (from the words up) or top-down (from the non-terminals down). Evidence that the human parser combines both strategies.
- Parallelism: a serial parser maintains only one structure at a time; a parallel parser pursues all possible structures. Controversial issue; evidence for both serialism and limited parallelism.
- Interactivity: the parser can be encapsulated (only access to syntactic information) or interactive (access to semantic information, context). Evidence for limited interactivity.

Eye-tracking

An eye-tracker makes it possible to record the eye-movements of subjects while their are performing a cognitive task:

- looking at a scene;
- driving a vehicle;
- using a computer;
- reading a text.

Mind's Eye Hypothesis: where subjects are looking indicates what they are processing. How long they are looking at it indicates how much processing effort is needed.

Eye-tracking

A head-mounted, video-based eye-tracker.

Eye-movements and Reading

Let's look at eye-tracking data for reading in detail:

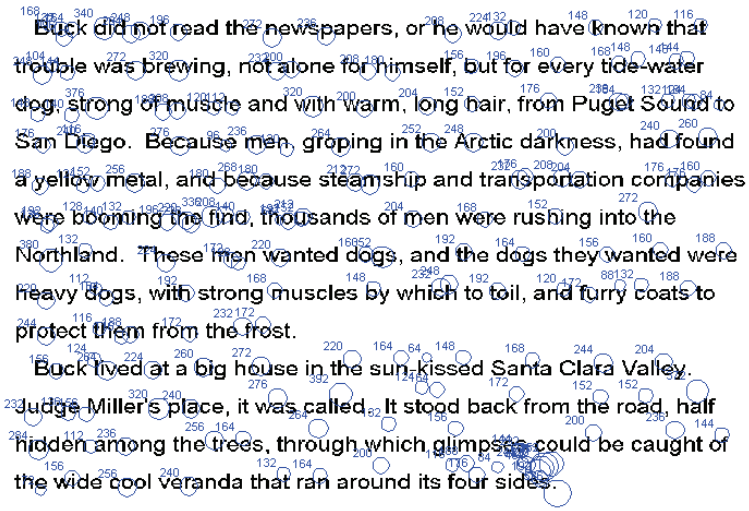
- eye-movements are recorded while subjects read texts;
- very high spatial (0.15° visual angle) and temporal (1 ms) accuracy;
- eye movements in reading are saccadic: a series of relatively stationary periods (fixations) between very fast movements (saccades);
- average fixation time is about 250 ms; can be longer or shorter, depending on ease or difficulty of processing;
- typically test a number of subjects, with a number of test sentences, and statistical analysis done on results.

Eye-movements and Reading

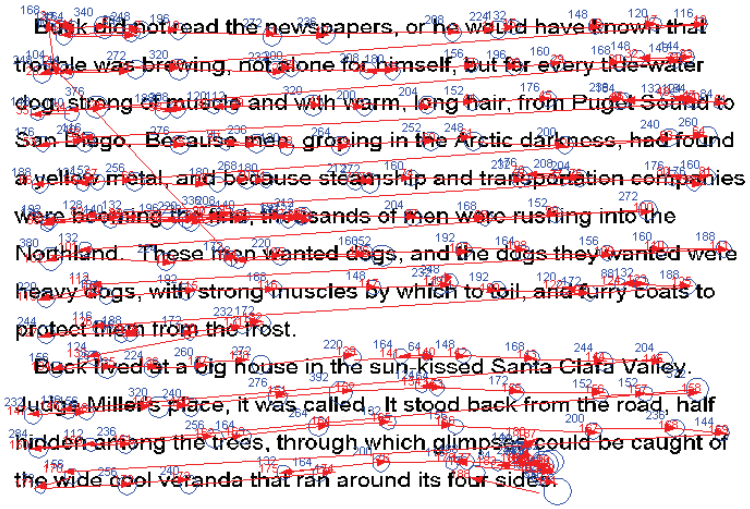
Buck did not read the newspapers, or he would have known that trouble was brewing, not alone for himself, but for every tide-water dog, strong of muscle and with warm, long hair, from Puget Sound to San Diego. Because men, groping in the Arctic darkness, had found a yellow metal, and because steamship and transportation companies were booming the find, thousands of men were rushing into the Northland. These men wanted dogs, and the dogs they wanted were heavy dogs, with strong muscles by which to toil, and furry coats to protect them from the frost.

Buck lived at a big house in the sun-kissed Santa Clara Valley. Judge Miller's place, it was called. It stood back from the road, half hidden among the trees, through which glimpses could be caught of the wide cool veranda that ran around its four sides.

Eye-movements and Reading



Eye-movements and Reading



Eye-movements and Reading

We can use the data generated by eye-tracking experiments to investigate how the human parser works. For example:

- evidence for garden paths comes from increased reading times, and more reverse saccades, when reading certain words;
- evidence for incrementality comes from studies where participants view visual scenes while listening to sentences;
- evidence for interactivity comes from the fact that semantic properties of words influence reading times in the same way as syntactic ones.

We will look at how to model these properties by building a parser that mimics human parsing behavior.

Clicker Questions

Which of the following sentences in not a garden path?

1. The man returned to his house was happy.
2. The complex houses married and single soldiers and their families.
3. The tomcat that curled up on the cushion seemed friendly.
4. The sour drink from the ocean.

Clicker Questions

Which of the following sentences in not a garden path?

1. The man returned to his house was happy.
2. The complex houses married and single soldiers and their families.
3. The tomcat that curled up on the cushion seemed friendly.
4. The sour drink from the ocean.

What does it mean if we skip words during reading?

1. We have dyslexia
2. We are non-native speakers
3. These are usually unknown words.
4. These are usually non-content bearing words.

A Small Grammar of English

We need a grammar of English to study our parsing models.

Phrasal categories:
S: sentence, NP: noun phrase, VP: verb phrase

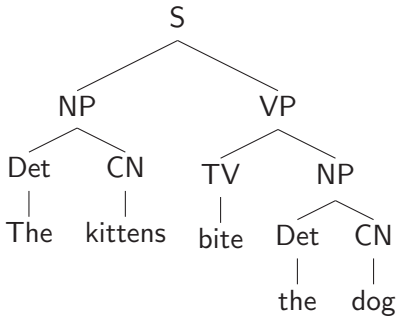
Syntactic categories (aka parts of speech):
Det: determiner, CN: common noun, TV: transitive verb

Phrase structure rules:

S	→	NP VP	Det	→	the
NP	→	Det CN	CN	→	kittens
VP	→	TV NP	TV	→	bite
			CN	→	dog

Syntax Tree

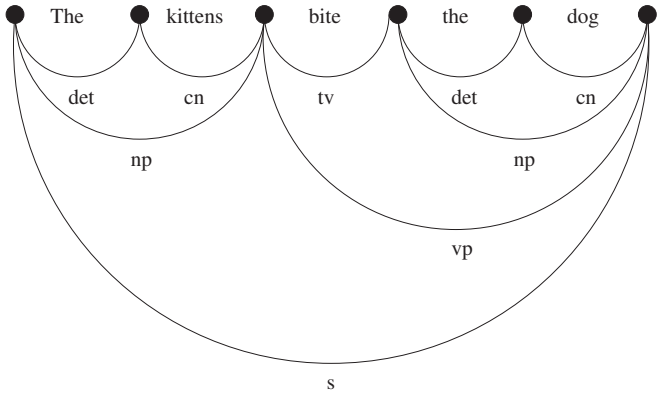
The syntax tree for the sentence *the kittens bite the dog:*



Let's assume that this tree is constructed a CYK parser as discussed in lecture 16.

A Bottom-Up Parallel Parser

Example of a CYK chart (in graph notation):



Properties of the Model

This offers a simple model of the human parser with the following properties:

- bottom-up: parsing is driven by the addition of words to the chart; chart is expended upwards from lexical to phrasal categories;
- limited incrementality: when a new word appears, all possible edges are added to the chart; then the system waits for the next word;
- parallelism: all chart edges are added at the same time; multiple analyses are pursued.

Properties of the Model

This offers a simple model of the human parser with the following properties:

- bottom-up: parsing is driven by the addition of words to the chart; chart is expended upwards from lexical to phrasal categories;
- limited incrementality: when a new word appears, all possible edges are added to the chart; then the system waits for the next word;
- parallelism: all chart edges are added at the same time; multiple analyses are pursued.

This model is therefore not cognitively plausible. Let's look at an alternative that is more promising.

Left Corner Parsing

Cognitively plausible incrementality: each word is integrated into the structure as it appears (no unconnected words).

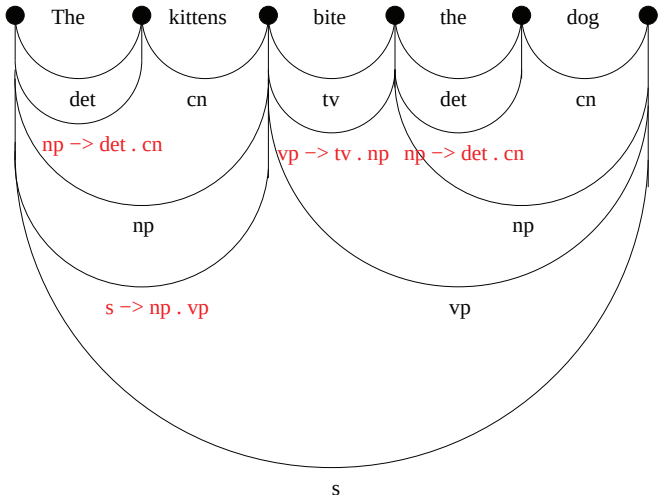
This can be achieved using left corner parsing. The chart of a left-corner parser contains edges that represent incomplete constituents.

This is similar to the Earley parsing, but uses a different Predictor:

PREDICTOR($(A \rightarrow \alpha\ .\ , [i, k])$)
1 **for each** $(B \rightarrow A\ \beta)$ **in** GRAMMAR_RULES_FOR(B, *grammar*)
2 **do** ENQUEUE($(B \rightarrow A\ .\ \beta, [i, k])$, *chart*$[i]$)

The Completer and Scanner as the same as in Earley.

Example of a Left Corner Chart



Serial Parsing

If parsing was fully parallel, all analyses of a sentence would be equally available; there would be no garden paths.

In the literature, two types of models have been assumed:

- ranked parallel: multiple structures are pursued in parallel; they are ranked in order of preference; garden paths occur if a low-ranked structure turns out to be correct;
- serial: only one structure is pursued; if it turns out to be incorrect, then a garden path occurs.

Let's assume a serial left-corner parser with backtracking.

Serial Parsing

Serial left-corner parser with backtracking:

- At each point of ambiguity, the parser has to chose one structure, instead of adding all possible structures to the chart;
- if this structure turns out to be incorrect; the parser has to backtrack;
- at the last point of ambiguity, the incorrect structure is removed from the chart, and an alternative one is added to the chart instead;
- this requires additional data structures to keep track of the choice points and the choices made.

Properties of the Model

Properties of the left corner model:

- this model will parse garden path sentences such as *the horse raced past the barn fell;*
- extensive backtracking will occur for such sentences; only possible if the stack size of the choice point stack is sufficient;

Potential problems:

- backtracking requires that parse failure is detected; requires that the parser knows where the sentence boundaries are;
- processing order is fixed; context or experience is not taken into accounts; no attempt to minimize backtracking;
- still not fully incremental; would need to add Composition operation to the algorithm.

Summary

- The human parser builds syntactic structure in response to strings of words;
- parsing models have to capture the incrementality of human parsing and account for ambiguity resolution (garden paths);
- parsing models can be implemented using a chart (representing partial syntactic structure);
- a simple bottom-up parser assumes limited incrementality, full parallelism: not cognitively plausible;
- left-corner parsing models achieves a higher degree of incrementality;
- to model garden paths, we can assume serial parsing with backtracking.