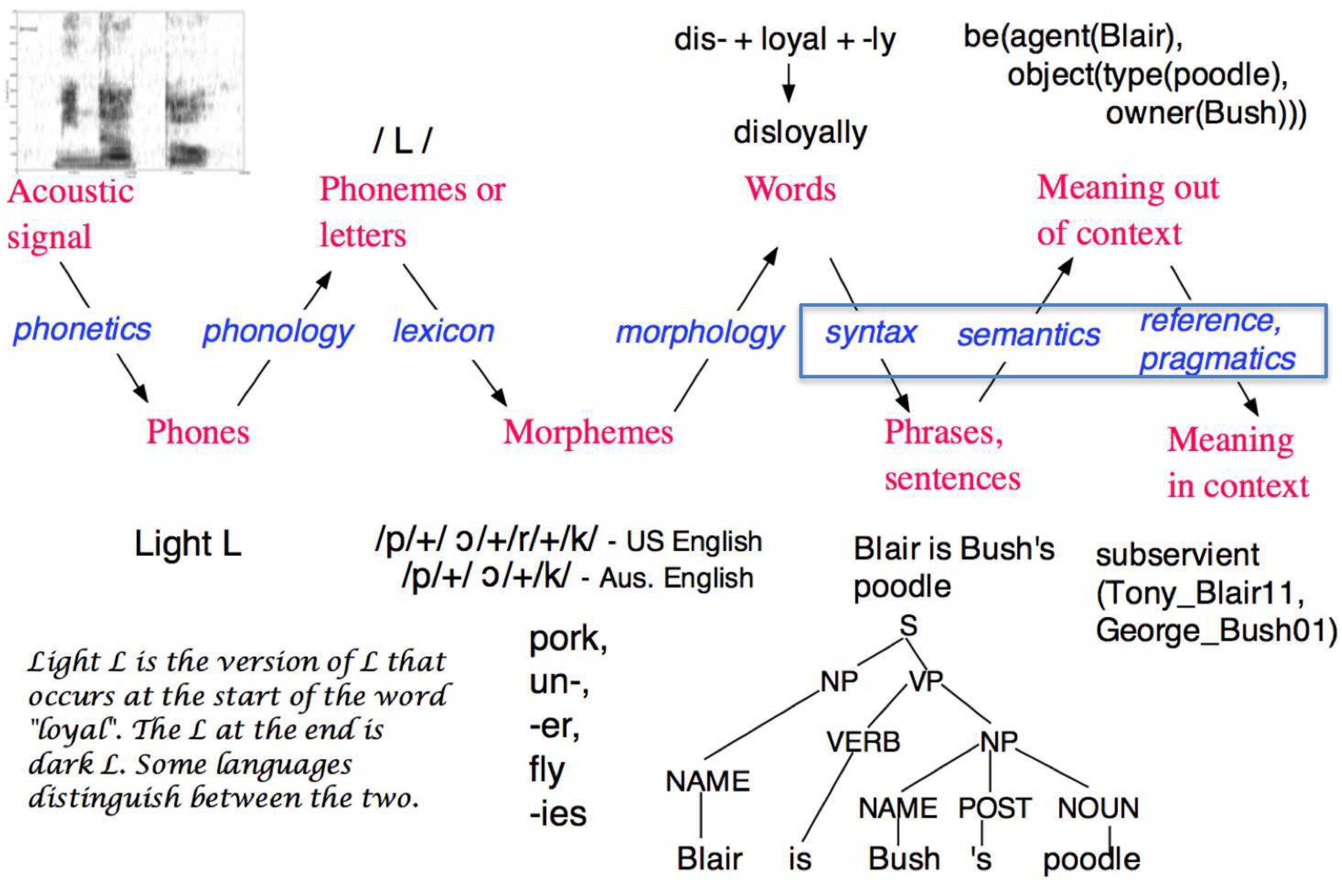
COMP3411: Artificial Intelligence

Grammars and Parsing

This Lecture

- Overview of Natural Languages
- Syntax and Grammar for Natural Languages

Linguistics Landscape



Natural Langue Processing

- Syntax
  - Linguistic Knowledge
  - Grammars and Parsing
  - Probabilistic Parsing
- Semantics
  - Semantic Interpretation and Logical Form
- Pragmatics
  - Discourse Processing
  - Speech Act Theory
  - (Spoken) Dialogue Systems

Related Disciplines

- Linguistics
  - Study of language in the abstract and particular languages
- Psycholinguistics
  - Psychological models of human language processing
- Neurolinguistics
  - Neural models of human language processing
- Logic
  - Study of formal reasoning

NLP Applications

- Chatbots
  - Customer service, e.g. CBA, Amtrak, Lyft, Spotify, Whole Foods
- Personal Assistants
  - Siri, Alexa, Google Assistant
- Information Extraction
  - Financial reports, news articles
- Machine (Assisted) Translation
  - Weather reports, EU contracts, Canada Hansard
- Social Robotics
  - Home care robots

Central Problem – Ambiguity

- Natural languages exhibit ambiguity

  "The fisherman went to the bank" (lexical)

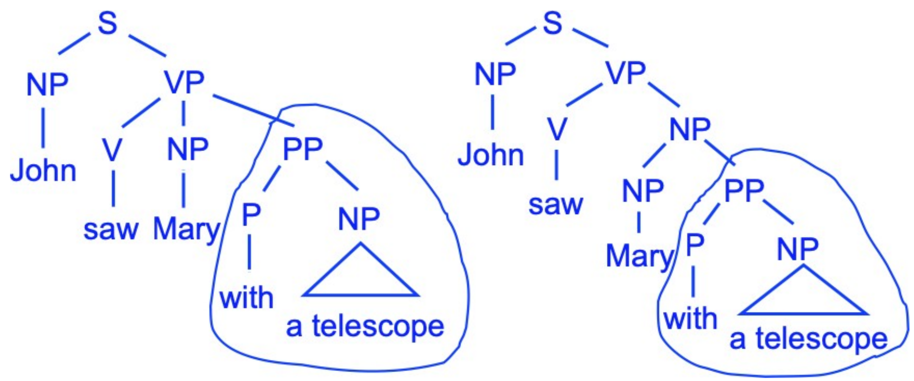
  "The boy saw a girl with a telescope" (structural) "Every student took an exam" (semantic)

  "The table won't fit through the doorway because it is too [wide/narrow]" (pragmatic)

- Ambiguity makes it difficult to interpret meaning of phrases/ sentences
  - But also makes inference harder to define and compute
- Resolve ambiguity by mapping to unambiguous representation

Structural Ambiguity

"John saw Mary with a telescope"



- Different interpretation →different representation

"John sold a car to Mary" and "Mary was sold a car by John"

- Same interpretation → same representation

Syntax

- Linguistic Knowledge and Grammars
- Context Free Grammars
- Parsing
  - Top Down Parsing
  - Bottom Up Parsing
  - Chart Parsing
  - Deterministic Parsing
  - Probabilistic Parsing

Sentence vs Utterance

- Sentence is a group of words that convey some meaning
- An utterance is a group of words between pauses in speech.

Lexical Items (Basic Words)

- Open class (can add new words)
  - Nouns: denote objects (e.g. cat, John, justice)
  - Verbs: denote actions, events (e.g. buy, break, believe)
  - Adjectives: denote properties of objects (e.g. red, large)
  - Adverbs: denote properties of events (e.g. quickly)
- Closed class (function words, mostly fixed in the language)
  - Prepositions: at, in, of, on, . . .
  - Articles: the, a, an
  - Conjunctions: and, or, if, then, than, . . .

Sentence Forms

- Declarative (indicative)
  - Bart is listening.
- Yes/No question (interrogative)
  - Is Bart listening?
- Wh-question (interrogative)
  - When is Bart listening?
- Imperative (command)
  - Listen, Bart!
- Subjunctive (conditional or imaginary situations)
  - If Bart were listening, he might hear something useful.

Noun Phrases

- Noun phrases: occur as “subject” with a range of “predicates”
  - (noun phrase) ate the bone
  - (noun phrase) saw the bird in the sky
  - (noun phrase) believes that 2 + 2 = 4
- Examples
  - John, The dog, The big ugly dog, The man in the red car, The oldest man in the world with a beard, The oldest man who lives in China, . . .
- Sentences need not “make sense”

Verb Phrases

- Verb phrases: occur as “predicate” with a range of “subjects”

  John (verb phrase)
  The dog (verb phrase)
  Any noun phrase (verb phrase)

- Examples
  - ate the bone
  - saw the bird in the sky
  - believes that 2 + 2 = 4
- Verb phrase depends on noun phrase

Inside Noun Phrases

- Within noun phrase
  - Main item (the head of the phrase): noun
  - Optional specifiers
    - Determiners (articles, demonstratives, quantifiers)
    - Adjectives and other nouns
  - Mandatory arguments
    - Depend on head (e.g. capital (of France))
  - Optional modifiers
    - Adjectival phrases (e.g. larger than Spain)
    - Prepositional phrases (e.g. in the park)
    - Relative clauses (e.g. who likes beer)
  - Order specifiers, head, modifiers in English (e.g. firstly, …)

Inside Verb Phrases

- Within verb phrase
  - Main item (the head of the phrase): verb
  - Optional specifiers
    - Auxiliary verbs (e.g. do, does, will, might, . . .)
    - Adverbs (e.g. quickly)
  - Mandatory arguments
    - depend on head (e.g. bought (a book) (for Henry))
  - Optional modifiers
    - Adverbial phrases (e.g. more quickly than Henry)
  - Notice similar structure to noun phrases

Prepositional Phrases

- Within prepositional phrase
  - Main item (the head of the phrase): preposition
  - Mandatory arguments
    - (noun phrase) (e.g. in the park)
- Nouns, verbs, etc. are just the heads of phrases

Context Free Grammars

- Terminal symbols (lexical items)
- Nonterminal symbols (grammatical categories)
- Start symbol (a nonterminal) e.g. (sentence)
- Rewrite rules
  - nonterminal → sequence of nonterminals, terminals
  - e.g. (sentence) → (noun phrase) (verb phrase)
- Open question: is English context free?

Typical (Small) Grammar

S → NP VP

NP → [Det] Adj* N [AP | PP | Rel Clause]*

VP → V [NP] [NP] PP*

AP → Adj PP

PP → P NP

Det → a | an | the | . . .

N → John | Mary | park | telescope | . . .

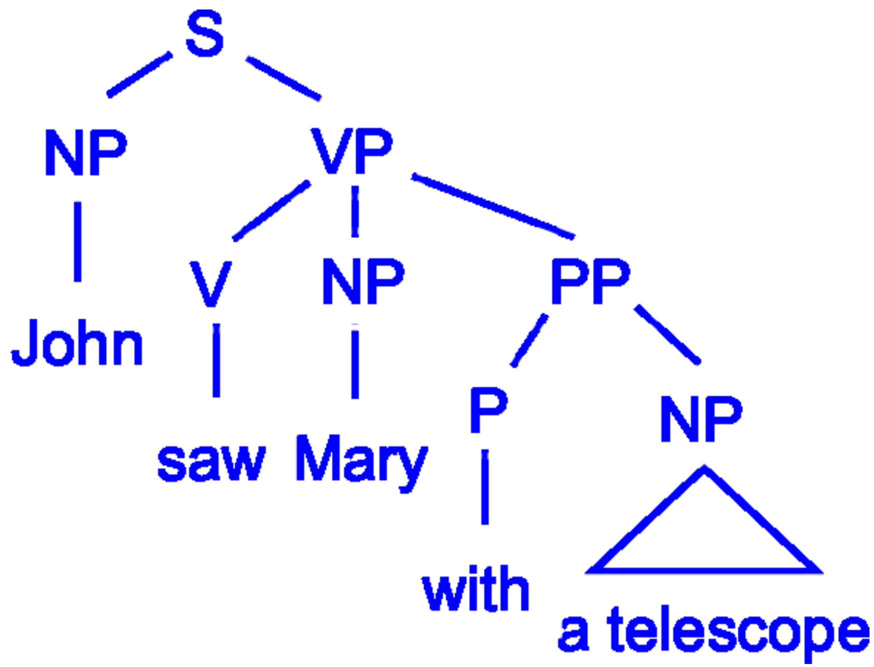
V → saw | likes | believes | . . .

Adj → hot | hotter | . . .

P → in | . . .

Extra notation: * is “0 or more”; [. .] is “optional”

Syntactic Structure



Syntactically ambiguous = more than one parse tree

(Leftmost) Derivation of Example

S
⇒ NP VP
⇒ N VP
⇒ John VP
⇒ John V NP PP
⇒ John saw NP PP
⇒ John saw N PP
⇒ John saw Mary PP
⇒ John saw Mary P NP
⇒ John saw Mary with NP
⇒ John saw Mary with Det N
⇒ John saw Mary with a N
⇒ John saw Mary with a telescope

S → NP VP
NP → [Det] Adj* N [AP | PP | Rel Clause]*
VP → V [NP] [NP] PP*
AP → Adj PP
PP → P NP
Det → a | an | the | …
N → John | Mary | park | telescope | . . .
V → saw | likes | believes | . . .
Adj → hot | hotter | . . .
P → in | . . .

⇒ means "rewrites as"

Rightmost Derivation

S

⇒ NP VP

⇒ NP V NP PP

⇒ NP V NP P NP

⇒ NP V NP P Det N

⇒ NP V NP P Det telescope

⇒ NP V NP P a telescope

⇒ . . .

⇒ . . .

⇒ . . .

S → NP VP
NP → [Det] Adj^* N [AP | PP | Rel Clause]*
VP → V [NP] [NP] PP^*
AP → Adj PP
PP → P NP
Det → a | an | the | ...
N → John | Mary | park | telescope | . . .
V → saw | likes | believes | . . .
Adj → hot | hotter | . . .
P → in | . . .

Parsing

- Aim is to compute a derivation of a sentence
  - produces parse tree
- Methods
  - Top down
    - Start with S, apply rewrite rules until sentence reached
  - Bottom up
    - Start with sentence, apply rewrite rules “in reverse” until S is reached
  - Chart parsing
    - Chart records parsed fragments and hypotheses
    - Can mix top down and bottom up strategies

Top-Down Parsing

- Use a stack to record working hypothesis
- Start with S as only symbol on stack
- At each step
  - Rewrite top of stack T using grammar rule T → RHS
- i.e. replace T by RHS (in reverse order), OR
- Match word on top of stack to next word in sentence
- Apply backtracking on failure
- Accept sentence when stack is empty and all words in sentence matched; reject sentence when no rules to try
- Produces leftmost derivation

Example

STACK	INPUT
S	John saw Mary with a telescope
NP VP	John saw Mary with a telescope
N VP	John saw Mary with a telescope
John VP	John saw Mary with a telescope
VP	saw Mary with a telescope
V NP PP	saw Mary with a telescope
Saw NP PP	saw Mary with a telescope
NP PP	Mary with a telescope
. . .	. . .
. . .	. . .

Bottom Up Parsing

- Use a stack to record parsed (left-right) fragment
- Start with stack empty
- At each step
  - Rewrite sequence at top of stack using rule T → RHS i.e. replace RHS (in reverse) by T, OR
  - Move word from input to stack
- Apply backtracking on failure
- Accept sentence when input empty and stack contains S; reject sentence when no more rules to try
- Produces rightmost derivation (in reverse)

Example

STACK	INPUT
	John saw Mary with a telescope
John	saw Mary with a telescope
N	saw Mary with a telescope
NP	saw Mary with a telescope
NP saw	Mary with a telescope
NP V	Mary with a telescope
NP V Mary	with a telescope
NP V N	with a telescope
. . .	. . .
. . .	. . .

Chart Parsing

- Top-down and bottom-up parsers may have to repeat parsing because they backtrack
- E.g. “The old man the boats”:
  - “man” can be a noun or a verb
  - Initially “old man” might be put together as a noun phrase,
  - but then we see that “man” is the verb
- A chart parser maintains a table or graph of all possible parsed fragments to avoid backtracking

“Garden Path” Sentence

Grammar:

1. $S \rightarrow NP\ VP$
2. $NP \rightarrow ART\ N$
3. $NP \rightarrow ART\ ADJ\ N$
4. $VP \rightarrow V\ NP$

Lexicon:

the: ART

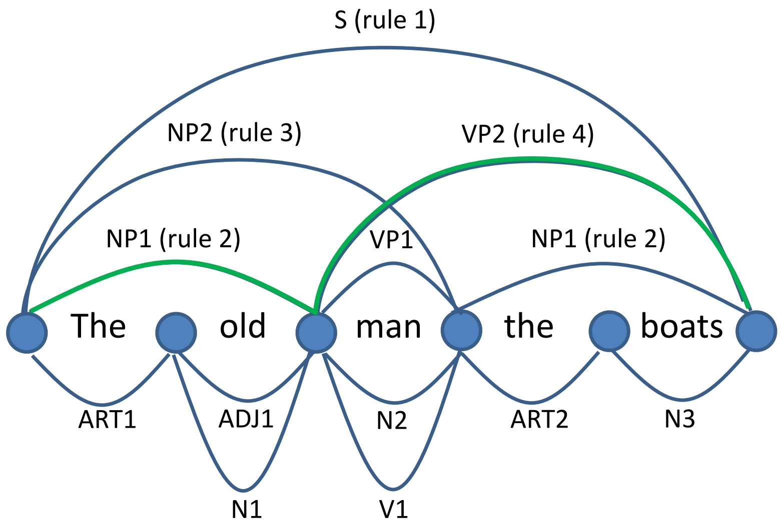
old: ADJ, N

man: N, V

boat: N

Sentence: $_1$ The $_2$ old $_3$ man $_4$ the $_5$ boat $_6$

Chart Parser Example



Grammar:

1. S → NP VP
2. NP → ART N
3. NP → ART ADJ N
4. VP → V NP

NP → ART • N

NP → ART • N

NP → ART • ADJ N

NP → ART • ADJ N

NP → ART ADJ • N

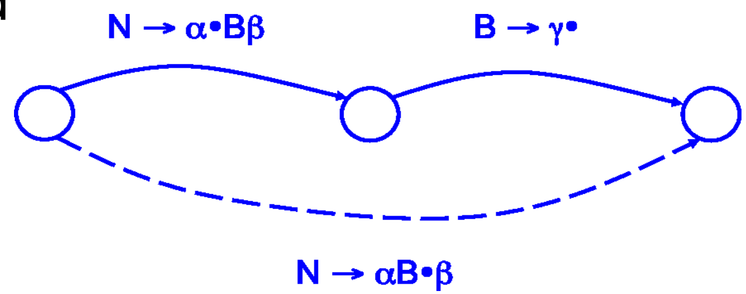
S → NP • VP

VP → V • NP

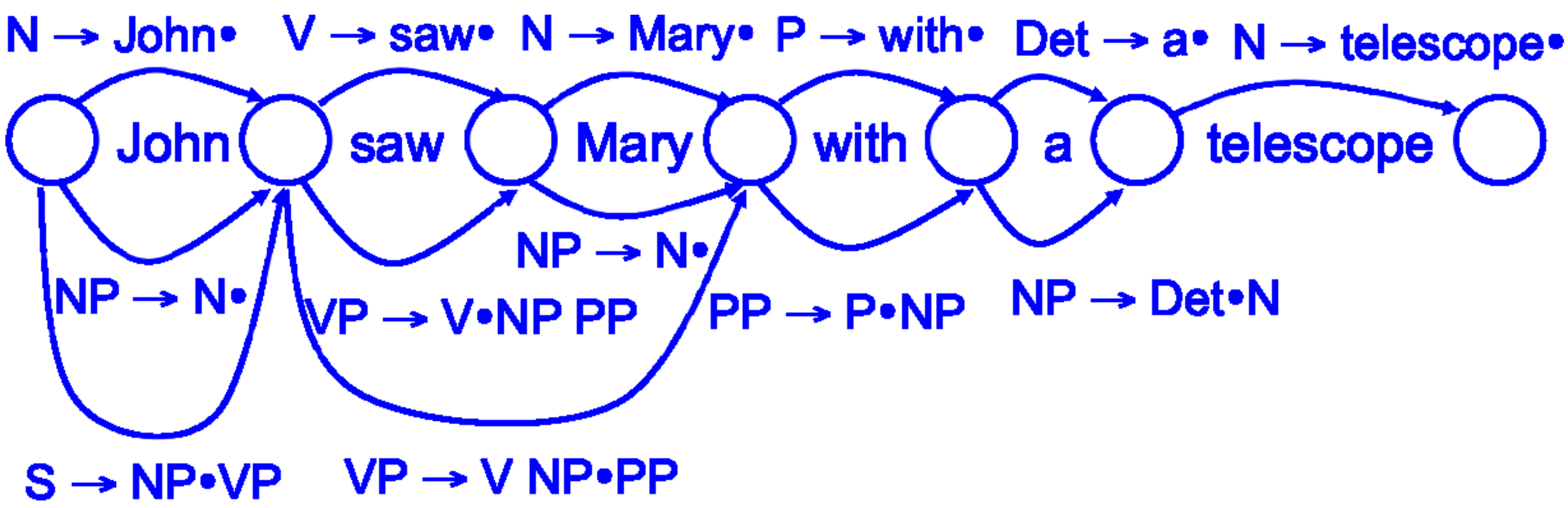
S → NP • VP

Chart Parsing

- Use a chart to record parsed fragments and hypotheses
- Hypotheses $N \to \alpha \cdot \beta$ where $N \to \alpha\beta$ is a grammar rule means "trying to parse N as $\alpha\beta$ and have so far parsed α"
- One node in chart for each word gap, start and end
- One arc in chart for each hypothesis
- At each step, apply fundamental rule
  - If chart has $N \to \alpha \cdot B\beta$ from n1 to n2 and $B \to \gamma\cdot$ from n2 to n3
  - add $N \to \alpha B \cdot \beta$ from n1 to n3
- Accept sentence when $S \to \alpha\cdot$ is added from start to end
- Can produce any sort of derivation



Example Chart



Comparing Parsing Methods

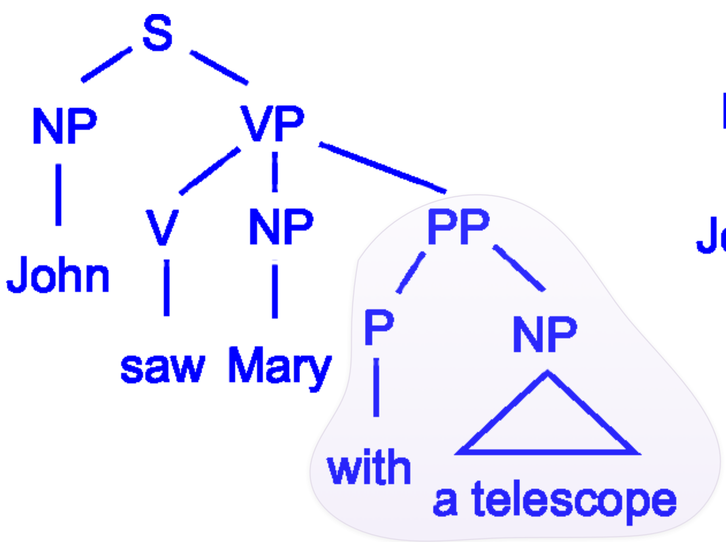
- Top Down Parsing
  - Simple, Memory efficient
  - Much repeated work, may loop infinitely
- Bottom Up Parsing
  - Less repeated work, harder to control
- Chart Parsing
  - Memory inefficient (especially with features)
  - No repeated work, difficult to control

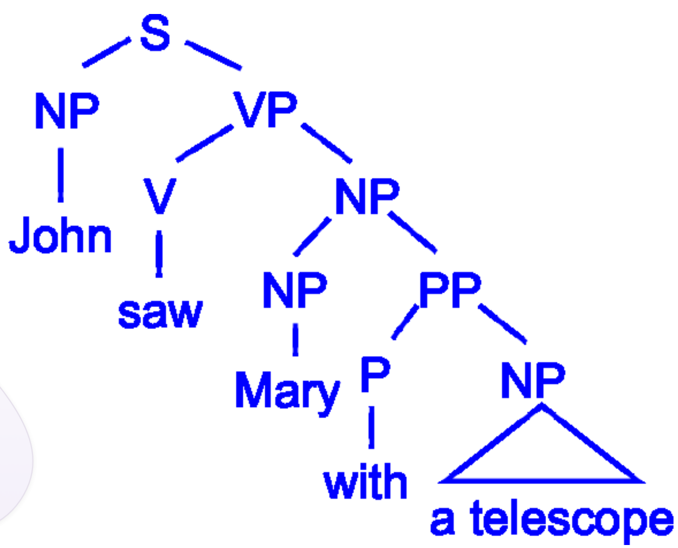
Deterministic Parsing

- Motivation
  - People don’t notice ambiguity . . .
  - But sometimes have trouble
- “The horse raced past the barn fell”
  “We painted all the walls with cracks”
  “The man kept the dog in the house”
- Can we do what the “human parser” does?

Heuristics

Minimal Attachment

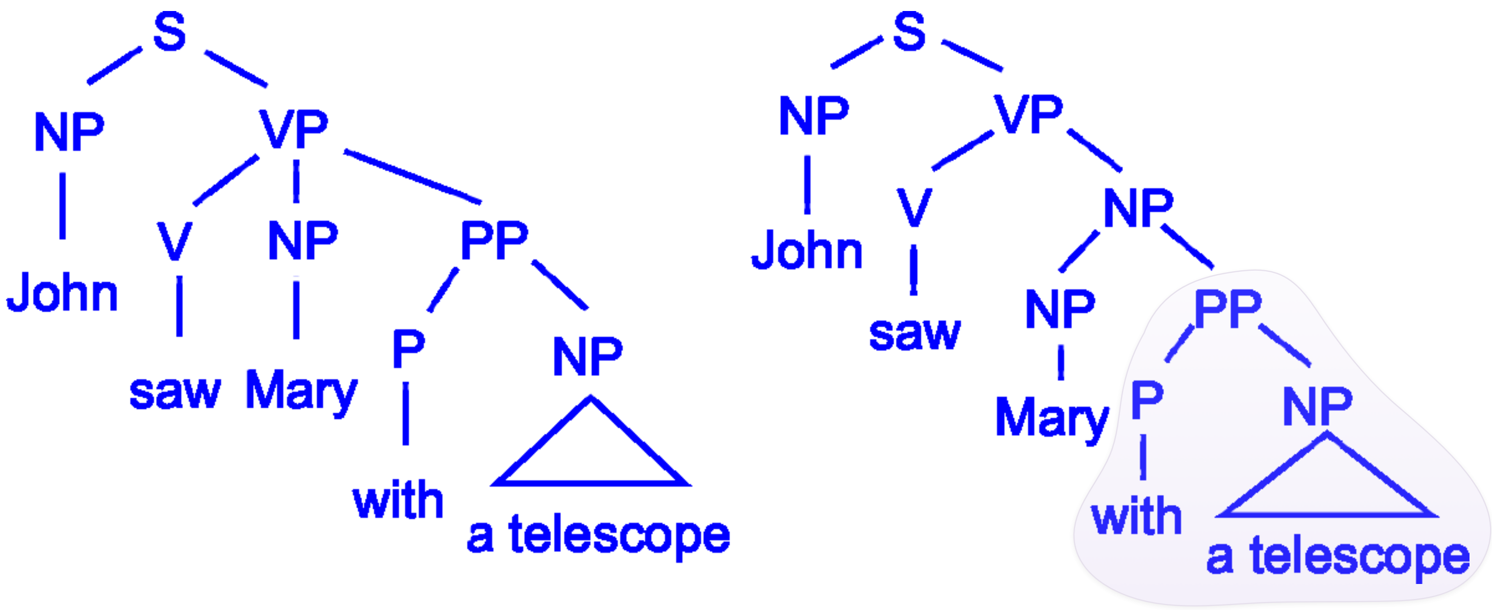




Minimise size of parse tree

Heuristics

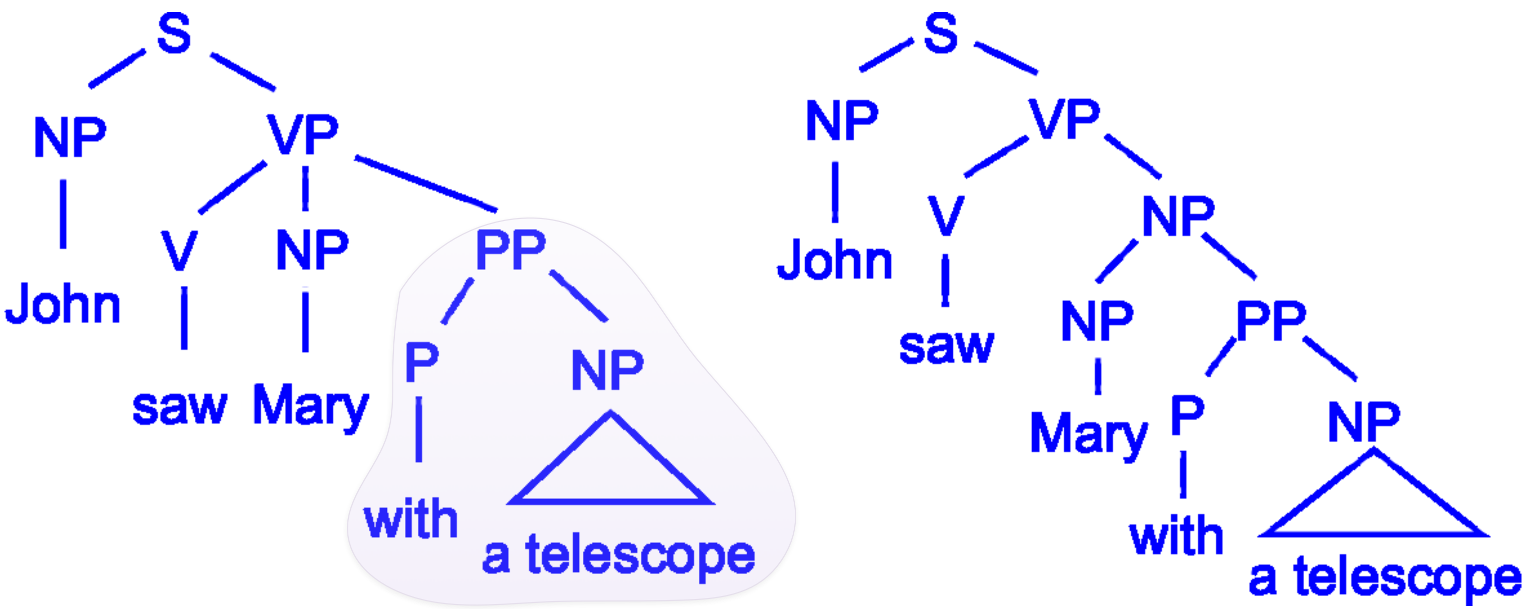
Right Association



Always attach to rightmost (lower) nodes

Heuristics

Lexical Preference



Try to fill most common sub-categorisation frame

Probabilistic Context Free Grammars

- Associate probabilities with grammar rules
  - Requires parsed corpus (e.g. Penn Treebank)
  - Count number of times rule used in parsing corpus sentences
- Probability of parse tree
  - $\prod_r$ probability rule $r \times \prod_w$ probability of word w given category
  - Assuming independence

Probabilistic Chart Parsing

- Start with probabilities calculated by part of speech tagger
- Multiply probabilities when applying fundamental rule
- Best-First Chart Parsing
  - Examine most likely constituents first (priority queue)
  - Never constructs constituents with lower probability than parse

Summary

- Syntactic Knowledge
  - Grammatical categories defined by distribution
  - Much determined by properties of lexical items
- Context Free Grammars
  - Useful and powerful formalism
  - Relatively efficient parsers
  - Limited when dealing with complex phenomena
- Parsing
  - Top down method is easy to understand, but not efficient
  - Bottom up method is more efficient