Natural Language Processing (CS5803)

Lecture 8
(Parsing - Part 1)

Parsing

- Natural language sentence: sequence of text – linear structure
- The **construction** of the sentence is governed by **grammatical rules**
- How the **words connect with each other** according to the rule – structural composition of the sentence
- Identifying the **structural composition** – **parsing**
  - **Constituency Parsing:** Phrase information
  - **Dependency Parsing:** Relationship between words
  - **Semantic Parsing:** Logical interpretation

Part 1 - Constituency Parsing

Constituency

- Fundamental notion underlying the idea of constituency is that of abstraction
  - groups of words behaving as single units, or constituents.
- How do words group together in English?
  - On July nineth, I'd like to fly from Atlanta to Denver.
  - I'd like to fly on July nineth from Atlanta to Denver.
  - I'd like to fly from Atlanta to Denver on July nineth .
- How to model constituent structure in English?
  - **Context-Free Grammar (CFG)**

CFG

- Also called Phrase-Structure Grammar (PSG)

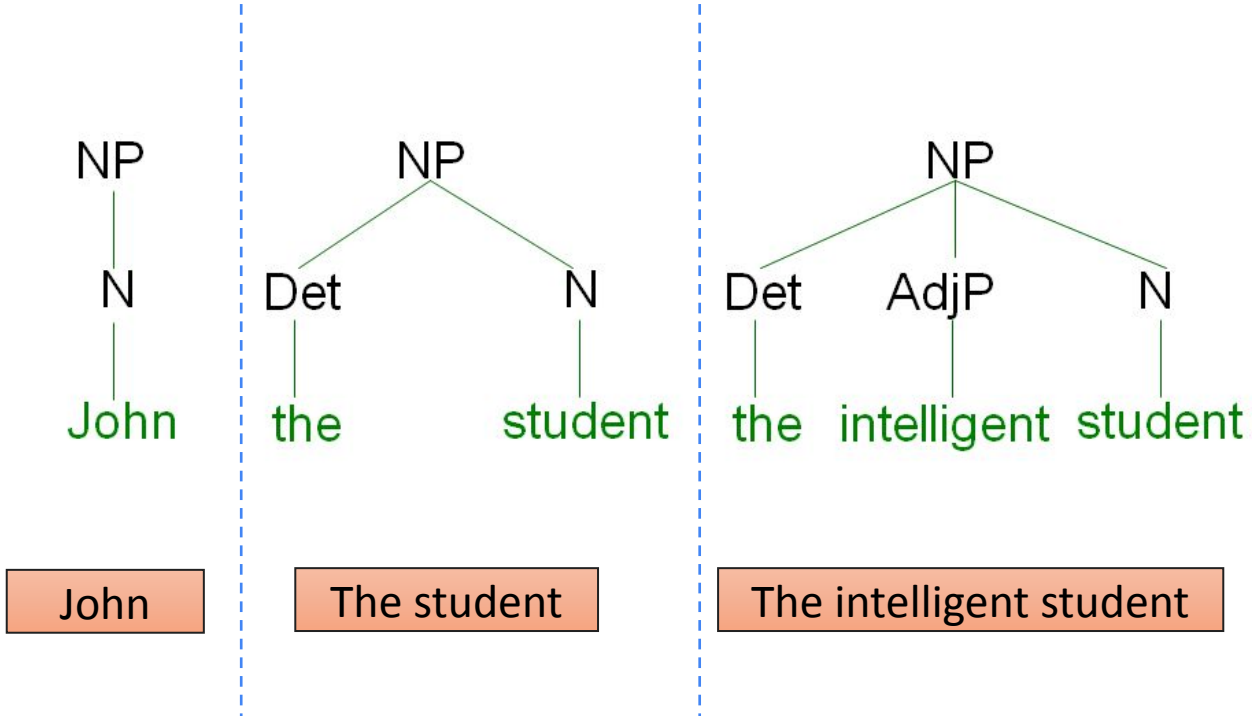
- CFG can be viewed as a device for
  - generating sentences
  - assigning a structure to a given sentence

- A context-free grammar G is defined by four parameters: (***N***,***Σ***,***R***,***S***)
  - ***N*** : a set of non-terminal symbols (or variables)
  - ***Σ*** : a set of terminal symbols (disjoint from *N*)
  - ***R*** : a set of rules or productions, each of the form $A \rightarrow \beta$, where *A* is a non-terminal, β is a string of symbols from the infinite set of strings $(S \cup N)^*$
  - ***S*** : a designated start symbol and a member of *N*

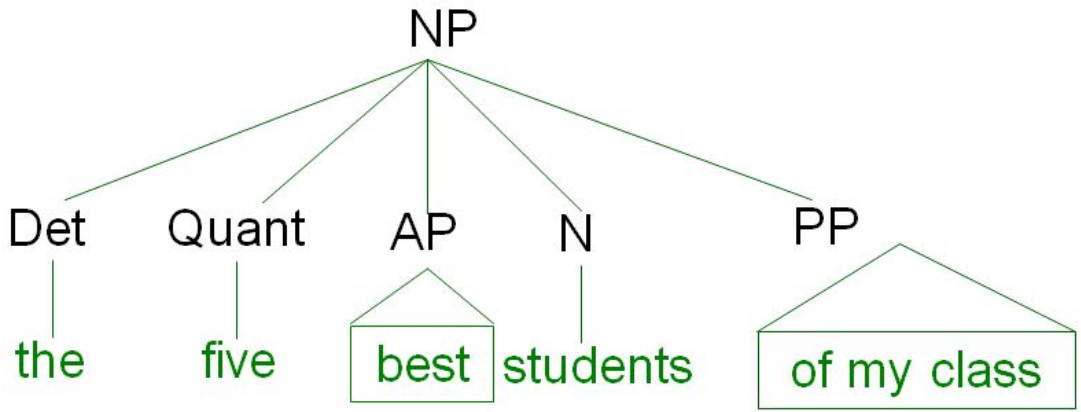
S -> AB
A -> to X
B -> in Y
X -> HYD
X -> BLR
Y -> May
Y -> Jun

S -> AB
A -> to X
B -> in Y
X -> HYD|BLR
Y -> May|Jun

Noun Phrase

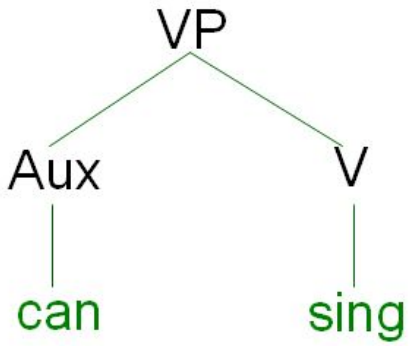

NP
N
John
NP
Det
N
the
student
NP
Det
AdjP
N
the
intelligent
student
John
The student
The intelligent student

Noun Phrase

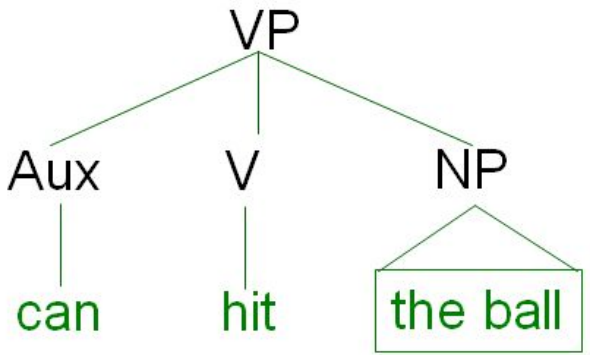


The five best students of my class

Verb Phrase

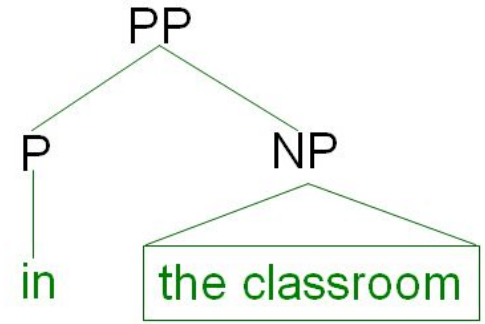


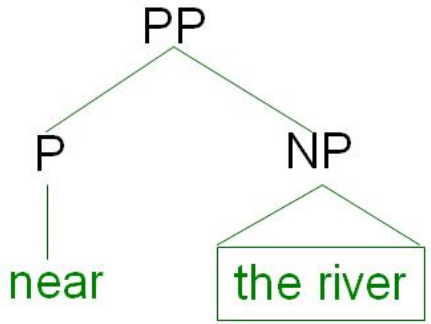
can sing



can hit the ball

Prepositional Phrase

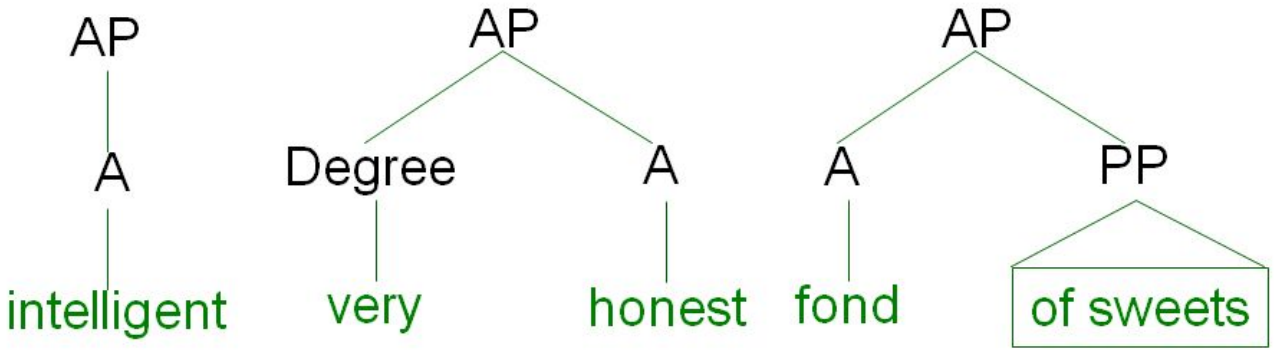




in the classroom

near the river

Adjective Phrase

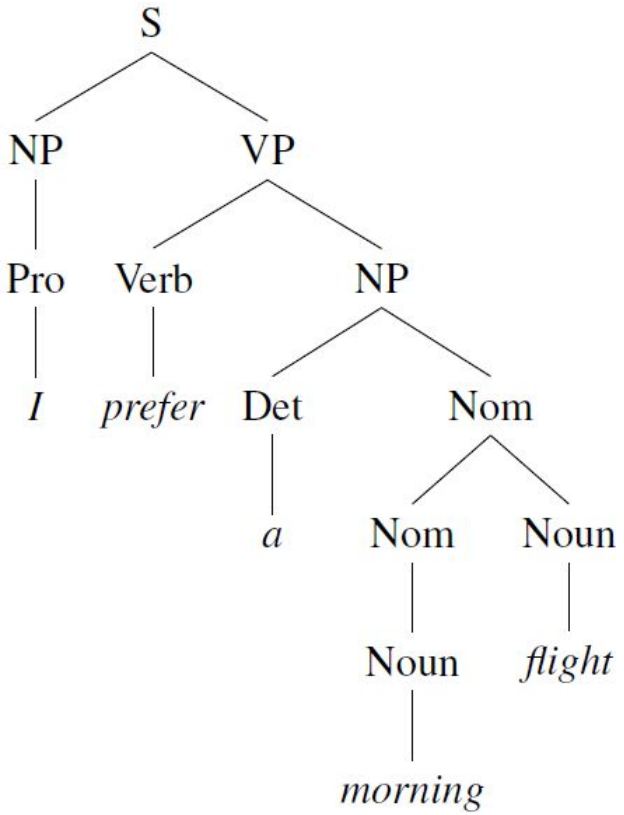


intelligent

very honest

fond of sweets

CFG/PSG Parse Tree



Noun	→	*flights* | *breeze* | *trip* | *morning*
Verb	→	*is* | *prefer* | *like* | *need* | *want* | *fly*
Adjective	→	*cheapest* | *non-stop* | *first* | *latest* | *other* | *direct*
Pronoun	→	*me* | *I* | *you* | *it*
Proper-Noun	→	*Alaska* | *Baltimore* | *Los Angeles* | *Chicago* | *United* | *American*
Determiner	→	*the* | *a* | *an* | *this* | *these* | *that*
Preposition	→	*from* | *to* | *on* | *near*
Conjunction	→	*and* | *or* | *but*

S	→	*NP VP*
NP	→	*Pronoun*
	|	*Proper-Noun*
	|	*Det Nominal*
Nominal	→	*Nominal Noun*
	|	*Noun*
VP	→	*Verb*
	|	*Verb NP*
	|	*Verb NP PP*
	|	*Verb PP*
PP	→	*Preposition NP*

Treebanks

- Syntactically annotated corpus
- Every sentence is paired with a corresponding parse tree
- **Penn Treebank**
  - "That cold, empty sky was full of fire and light."

```
((S
   (NP-SBJ (DT That)
     (JJ cold) (, ,)
     (JJ empty) (NN sky) )
   (VP (VBD was)
     (ADJP-PRD (JJ full)
       (PP (IN of)
         (NP (NN fire)
           (CC and)
           (NN light) ))))
   (. .) ))
```

Penn Treebank Syntactic Annotation Format

Penn POS Tags

Adjective:	JJ
Adverb:	RB
Cardinal Number:	CD
Determiner:	DT
Preposition:	IN
Coordinating Conjunction	CC
Subordinating Conjunction:	IN
Singular Noun:	NN
Plural Noun:	NNS
Personal Pronoun:	PP
Proper Noun:	NP
Verb base form:	VB
Modal verb:	MD
Verb (3sg Pres):	VBZ
Wh-determiner:	WDT
Wh-pronoun:	WP

Constituency Parsing Algorithm

- Top Down :
  - Start with S and generate the sentence.
- Bottom up :
  - Start with the words in the sentence and use the rewrite rules backwards to reduce the sequence of symbols to produce S

S → NP VP
VP → V NP
NP → NNP | ART N
NNP → Ram
V → ate | saw
ART → a | an | the
N → rice | apple | movie

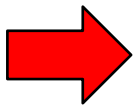
Top Down Parsing

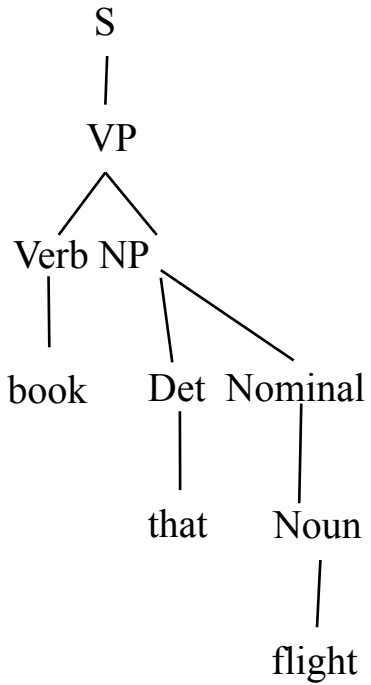
Ram ate the rice
=> NNP ate the rice (rewrite Ram)
=> NNP V the rice (rewrite ate)
=> NNP V ART rice (rewrite the)
=> NNP V ART N (rewrite rice)
=> NP V ART N (rewrite NNP)
=> NP V NP (rewrite ART N)
=> NP VP (rewrite V NP)
=> S

Bottom Up Parsing

Parsing Example

book that flight





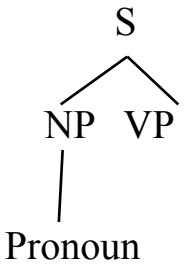
Grammar

S → NP VP
S → Aux NP VP
S → VP
NP → Pronoun
NP → Proper-Noun
NP → Det Nominal
Nominal → Noun
Nominal → Nominal Noun
Nominal → Nominal PP
VP → Verb
VP → Verb NP
VP → VP PP
PP → Prep NP

Lexicon

Det → the | that
Noun → book | flight
Verb → book
Pronoun → I | he | she | me
Proper-Noun → Houston | NWA
Aux → does
Prep → from | to

Top Down Parsing



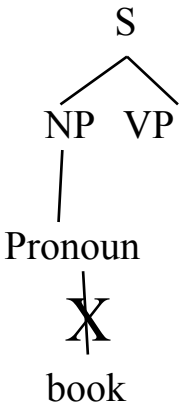
Grammar

S → NP VP
S → Aux NP VP
S → VP
NP → Pronoun
NP → Proper-Noun
NP → Det Nominal
Nominal → Noun
Nominal → Nominal Noun
Nominal → Nominal PP
VP → Verb
VP → Verb NP
VP → VP PP
PP → Prep NP

Lexicon

Det → the | that
Noun → book | flight
Verb → book
Pronoun → I | he | she | me
Proper-Noun → Houston | NWA
Aux → does
Prep → from | to

Top Down Parsing



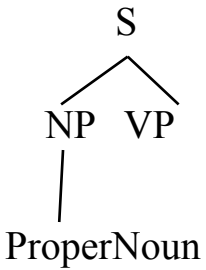
Grammar

S → NP VP
S → Aux NP VP
S → VP
NP → Pronoun
NP → Proper-Noun
NP → Det Nominal
Nominal → Noun
Nominal → Nominal Noun
Nominal → Nominal PP
VP → Verb
VP → Verb NP
VP → VP PP
PP → Prep NP

Lexicon

Det → the | that
Noun → book | flight
Verb → book
Pronoun → I | he | she | me
Proper-Noun → Houston | NWA
Aux → does
Prep → from | to

Top Down Parsing



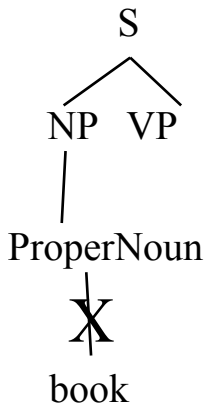
Grammar

S → NP VP
S → Aux NP VP
S → VP
NP → Pronoun
NP → Proper-Noun
NP → Det Nominal
Nominal → Noun
Nominal → Nominal Noun
Nominal → Nominal PP
VP → Verb
VP → Verb NP
VP → VP PP
PP → Prep NP

Lexicon

Det → the | that
Noun → book | flight
Verb → book
Pronoun → I | he | she | me
Proper-Noun → Houston | NWA
Aux → does
Prep → from | to

Top Down Parsing



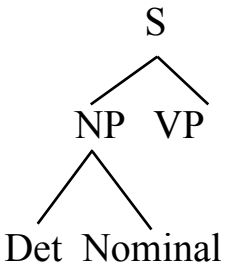
Grammar

S → NP VP
S → Aux NP VP
S → VP
NP → Pronoun
NP → Proper-Noun
NP → Det Nominal
Nominal → Noun
Nominal → Nominal Noun
Nominal → Nominal PP
VP → Verb
VP → Verb NP
VP → VP PP
PP → Prep NP

Lexicon

Det → the | that
Noun → book | flight
Verb → book
Pronoun → I | he | she | me
Proper-Noun → Houston | NWA
Aux → does
Prep → from | to

Top Down Parsing



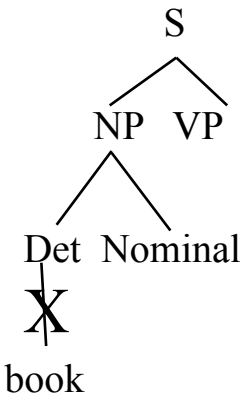
Grammar

S → NP VP
S → Aux NP VP
S → VP
NP → Pronoun
NP → Proper-Noun
NP → Det Nominal
Nominal → Noun
Nominal → Nominal Noun
Nominal → Nominal PP
VP → Verb
VP → Verb NP
VP → VP PP
PP → Prep NP

Lexicon

Det → the | that
Noun → book | flight
Verb → book
Pronoun → I | he | she | me
Proper-Noun → Houston | NWA
Aux → does
Prep → from | to

Top Down Parsing



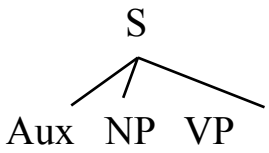
Grammar

S → NP VP
S → Aux NP VP
S → VP
NP → Pronoun
NP → Proper-Noun
NP → Det Nominal
Nominal → Noun
Nominal → Nominal Noun
Nominal → Nominal PP
VP → Verb
VP → Verb NP
VP → VP PP
PP → Prep NP

Lexicon

Det → the | that
Noun → book | flight
Verb → book
Pronoun → I | he | she | me
Proper-Noun → Houston | NWA
Aux → does
Prep → from | to

Top Down Parsing



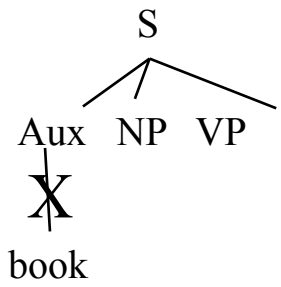
Grammar

S → NP VP
S → Aux NP VP
S → VP
NP → Pronoun
NP → Proper-Noun
NP → Det Nominal
Nominal → Noun
Nominal → Nominal Noun
Nominal → Nominal PP
VP → Verb
VP → Verb NP
VP → VP PP
PP → Prep NP

Lexicon

Det → the | that
Noun → book | flight
Verb → book
Pronoun → I | he | she | me
Proper-Noun → Houston | NWA
Aux → does
Prep → from | to

Top Down Parsing



Grammar

S → NP VP
S → Aux NP VP
S → VP
NP → Pronoun
NP → Proper-Noun
NP → Det Nominal
Nominal → Noun
Nominal → Nominal Noun
Nominal → Nominal PP
VP → Verb
VP → Verb NP
VP → VP PP
PP → Prep NP

Lexicon

Det → the | that
Noun → book | flight
Verb → book
Pronoun → I | he | she | me
Proper-Noun → Houston | NWA
Aux → does
Prep → from | to

Top Down Parsing

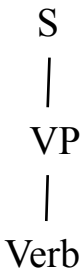
Grammar

S → NP VP
S → Aux NP VP
S → VP
NP → Pronoun
NP → Proper-Noun
NP → Det Nominal
Nominal → Noun
Nominal → Nominal Noun
Nominal → Nominal PP
VP → Verb
VP → Verb NP
VP → VP PP
PP → Prep NP

Lexicon

Det → the | that
Noun → book | flight
Verb → book
Pronoun → I | he | she | me
Proper-Noun → Houston | NWA
Aux → does
Prep → from | to

Top Down Parsing



Grammar

S → NP VP
S → Aux NP VP
S → VP
NP → Pronoun
NP → Proper-Noun
NP → Det Nominal
Nominal → Noun
Nominal → Nominal Noun
Nominal → Nominal PP
VP → Verb
VP → Verb NP
VP → VP PP
PP → Prep NP

Lexicon

Det → the | that
Noun → book | flight
Verb → book
Pronoun → I | he | she | me
Proper-Noun → Houston | NWA
Aux → does
Prep → from | to

Top Down Parsing

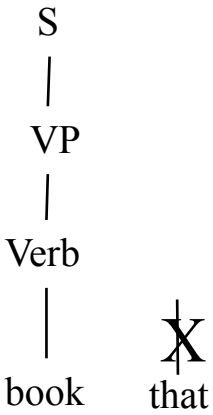
Grammar

S → NP VP
S → Aux NP VP
S → VP
NP → Pronoun
NP → Proper-Noun
NP → Det Nominal
Nominal → Noun
Nominal → Nominal Noun
Nominal → Nominal PP
VP → Verb
VP → Verb NP
VP → VP PP
PP → Prep NP

Lexicon

Det → the | that
Noun → book | flight
Verb → book
Pronoun → I | he | she | me
Proper-Noun → Houston | NWA
Aux → does
Prep → from | to

Top Down Parsing



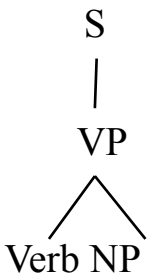
Grammar

S → NP VP
S → Aux NP VP
S → VP
NP → Pronoun
NP → Proper-Noun
NP → Det Nominal
Nominal → Noun
Nominal → Nominal Noun
Nominal → Nominal PP
VP → Verb
VP → Verb NP
VP → VP PP
PP → Prep NP

Lexicon

Det → the | that
Noun → book | flight
Verb → book
Pronoun → I | he | she | me
Proper-Noun → Houston | NWA
Aux → does
Prep → from | to

Top Down Parsing



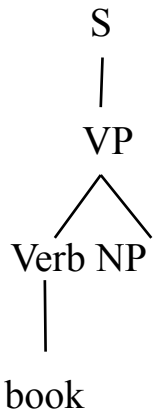
Grammar

S → NP VP
S → Aux NP VP
S → VP
NP → Pronoun
NP → Proper-Noun
NP → Det Nominal
Nominal → Noun
Nominal → Nominal Noun
Nominal → Nominal PP
VP → Verb
VP → Verb NP
VP → VP PP
PP → Prep NP

Lexicon

Det → the | that
Noun → book | flight
Verb → book
Pronoun → I | he | she | me
Proper-Noun → Houston | NWA
Aux → does
Prep → from | to

Top Down Parsing



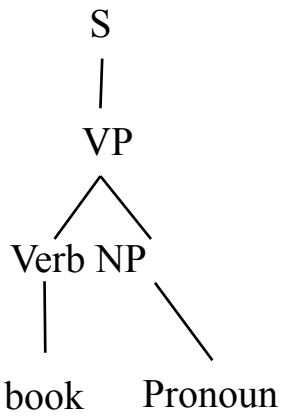
Grammar

S → NP VP
S → Aux NP VP
S → VP
NP → Pronoun
NP → Proper-Noun
NP → Det Nominal
Nominal → Noun
Nominal → Nominal Noun
Nominal → Nominal PP
VP → Verb
VP → Verb NP
VP → VP PP
PP → Prep NP

Lexicon

Det → the | that
Noun → book | flight
Verb → book
Pronoun → I | he | she | me
Proper-Noun → Houston | NWA
Aux → does
Prep → from | to

Top Down Parsing



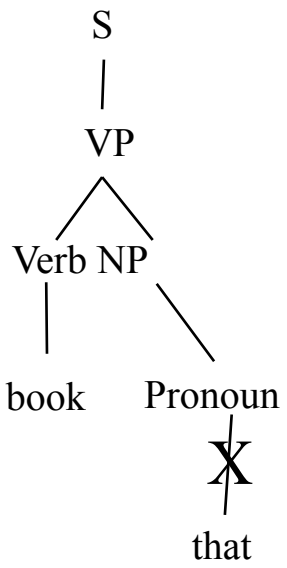
Grammar

S → NP VP
S → Aux NP VP
S → VP
NP → Pronoun
NP → Proper-Noun
NP → Det Nominal
Nominal → Noun
Nominal → Nominal Noun
Nominal → Nominal PP
VP → Verb
VP → Verb NP
VP → VP PP
PP → Prep NP

Lexicon

Det → the | that
Noun → book | flight
Verb → book
Pronoun → I | he | she | me
Proper-Noun → Houston | NWA
Aux → does
Prep → from | to

Top Down Parsing



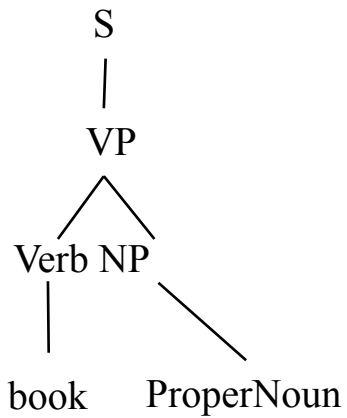
Grammar

S → NP VP
S → Aux NP VP
S → VP
NP → Pronoun
NP → Proper-Noun
NP → Det Nominal
Nominal → Noun
Nominal → Nominal Noun
Nominal → Nominal PP
VP → Verb
VP → Verb NP
VP → VP PP
PP → Prep NP

Lexicon

Det → the | that
Noun → book | flight
Verb → book
Pronoun → I | he | she | me
Proper-Noun → Houston | NWA
Aux → does
Prep → from | to

Top Down Parsing



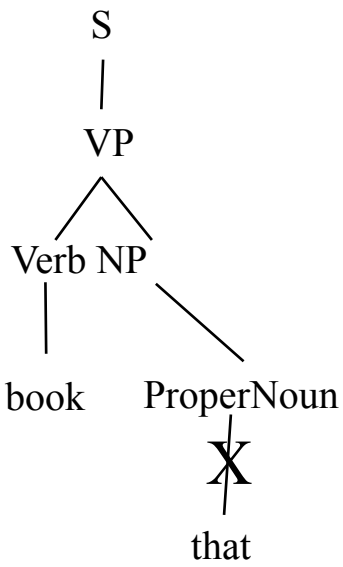
Grammar

S → NP VP
S → Aux NP VP
S → VP
NP → Pronoun
NP → Proper-Noun
NP → Det Nominal
Nominal → Noun
Nominal → Nominal Noun
Nominal → Nominal PP
VP → Verb
VP → Verb NP
VP → VP PP
PP → Prep NP

Lexicon

Det → the | that
Noun → book | flight
Verb → book
Pronoun → I | he | she | me
Proper-Noun → Houston | NWA
Aux → does
Prep → from | to

Top Down Parsing



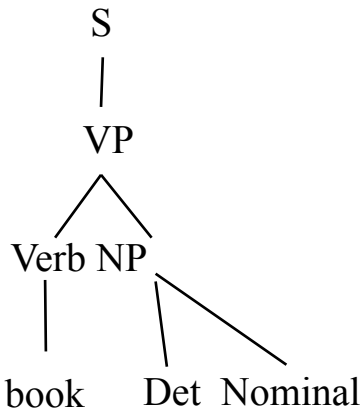
Grammar

S → NP VP
S → Aux NP VP
S → VP
NP → Pronoun
NP → Proper-Noun
NP → Det Nominal
Nominal → Noun
Nominal → Nominal Noun
Nominal → Nominal PP
VP → Verb
VP → Verb NP
VP → VP PP
PP → Prep NP

Lexicon

Det → the | that
Noun → book | flight
Verb → book
Pronoun → I | he | she | me
Proper-Noun → Houston | NWA
Aux → does
Prep → from | to

Top Down Parsing



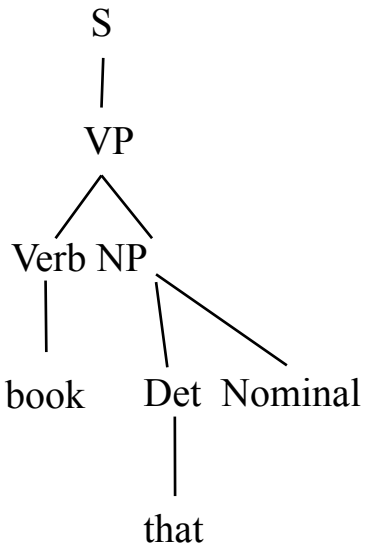
Grammar

S → NP VP
S → Aux NP VP
S → VP
NP → Pronoun
NP → Proper-Noun
NP → Det Nominal
Nominal → Noun
Nominal → Nominal Noun
Nominal → Nominal PP
VP → Verb
VP → Verb NP
VP → VP PP
PP → Prep NP

Lexicon

Det → the | that
Noun → book | flight
Verb → book
Pronoun → I | he | she | me
Proper-Noun → Houston | NWA
Aux → does
Prep → from | to

Top Down Parsing



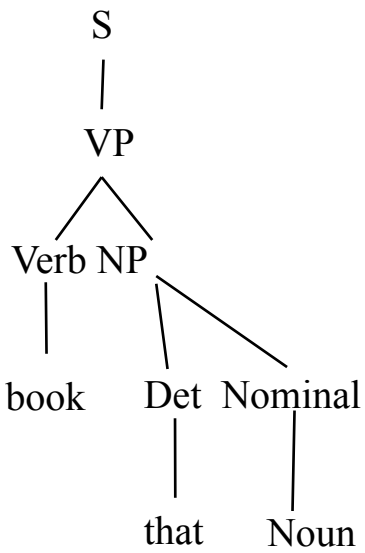
Grammar

S → NP VP
S → Aux NP VP
S → VP
NP → Pronoun
NP → Proper-Noun
NP → Det Nominal
Nominal → Noun
Nominal → Nominal Noun
Nominal → Nominal PP
VP → Verb
VP → Verb NP
VP → VP PP
PP → Prep NP

Lexicon

Det → the | that
Noun → book | flight
Verb → book
Pronoun → I | he | she | me
Proper-Noun → Houston | NWA
Aux → does
Prep → from | to

Top Down Parsing



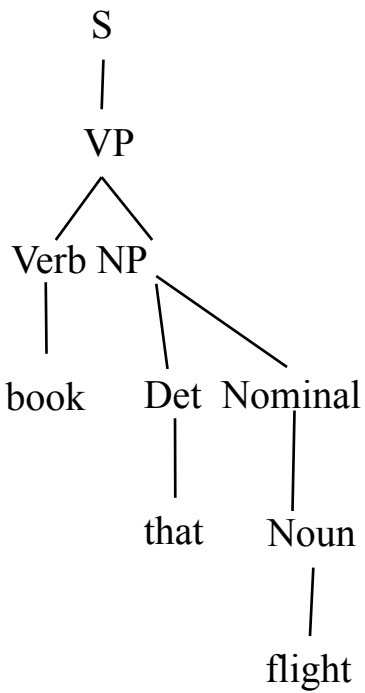
Grammar

S → NP VP
S → Aux NP VP
S → VP
NP → Pronoun
NP → Proper-Noun
NP → Det Nominal
Nominal → Noun
Nominal → Nominal Noun
Nominal → Nominal PP
VP → Verb
VP → Verb NP
VP → VP PP
PP → Prep NP

Lexicon

Det → the | that
Noun → book | flight
Verb → book
Pronoun → I | he | she | me
Proper-Noun → Houston | NWA
Aux → does
Prep → from | to

Top Down Parsing



Grammar

S → NP VP
S → Aux NP VP
S → VP
NP → Pronoun
NP → Proper-Noun
NP → Det Nominal
Nominal → Noun
Nominal → Nominal Noun
Nominal → Nominal PP
VP → Verb
VP → Verb NP
VP → VP PP
PP → Prep NP

Lexicon

Det → the | that
Noun → book | flight
Verb → book
Pronoun → I | he | she | me
Proper-Noun → Houston | NWA
Aux → does
Prep → from | to

Bottom Up Parsing

book that flight

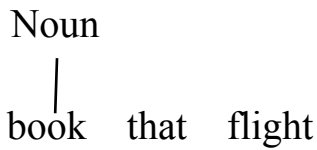
Grammar

S → NP VP
S → Aux NP VP
S → VP
NP → Pronoun
NP → Proper-Noun
NP → Det Nominal
Nominal → Noun
Nominal → Nominal Noun
Nominal → Nominal PP
VP → Verb
VP → Verb NP
VP → VP PP
PP → Prep NP

Lexicon

Det → the | that
Noun → book | flight
Verb → book
Pronoun → I | he | she | me
Proper-Noun → Houston | NWA
Aux → does
Prep → from | to

Bottom Up Parsing



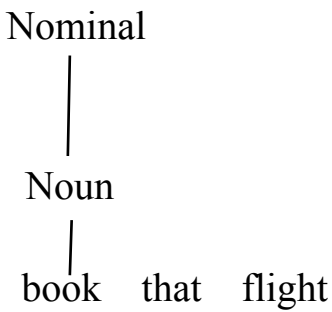
Grammar

S → NP VP
S → Aux NP VP
S → VP
NP → Pronoun
NP → Proper-Noun
NP → Det Nominal
Nominal → Noun
Nominal → Nominal Noun
Nominal → Nominal PP
VP → Verb
VP → Verb NP
VP → VP PP
PP → Prep NP

Lexicon

Det → the | that
Noun → book | flight
Verb → book
Pronoun → I | he | she | me
Proper-Noun → Houston | NWA
Aux → does
Prep → from | to

Bottom Up Parsing



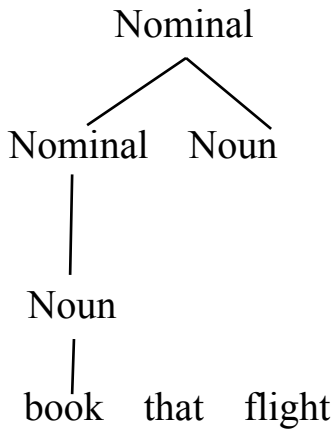
Grammar

S → NP VP
S → Aux NP VP
S → VP
NP → Pronoun
NP → Proper-Noun
NP → Det Nominal
Nominal → Noun
Nominal → Nominal Noun
Nominal → Nominal PP
VP → Verb
VP → Verb NP
VP → VP PP
PP → Prep NP

Lexicon

Det → the | that
Noun → book | flight
Verb → book
Pronoun → I | he | she | me
Proper-Noun → Houston | NWA
Aux → does
Prep → from | to

Bottom Up Parsing



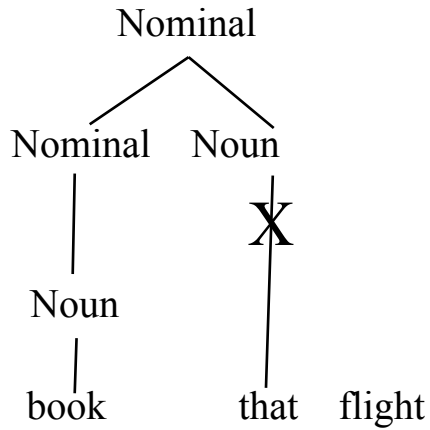
Grammar

S → NP VP
S → Aux NP VP
S → VP
NP → Pronoun
NP → Proper-Noun
NP → Det Nominal
Nominal → Noun
Nominal → Nominal Noun
Nominal → Nominal PP
VP → Verb
VP → Verb NP
VP → VP PP
PP → Prep NP

Lexicon

Det → the | that
Noun → book | flight
Verb → book
Pronoun → I | he | she | me
Proper-Noun → Houston | NWA
Aux → does
Prep → from | to

Bottom Up Parsing



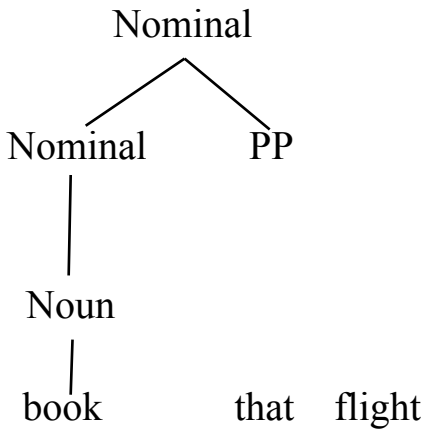
Grammar

S → NP VP
S → Aux NP VP
S → VP
NP → Pronoun
NP → Proper-Noun
NP → Det Nominal
Nominal → Noun
Nominal → Nominal Noun
Nominal → Nominal PP
VP → Verb
VP → Verb NP
VP → VP PP
PP → Prep NP

Lexicon

Det → the | that
Noun → book | flight
Verb → book
Pronoun → I | he | she | me
Proper-Noun → Houston | NWA
Aux → does
Prep → from | to

Bottom Up Parsing



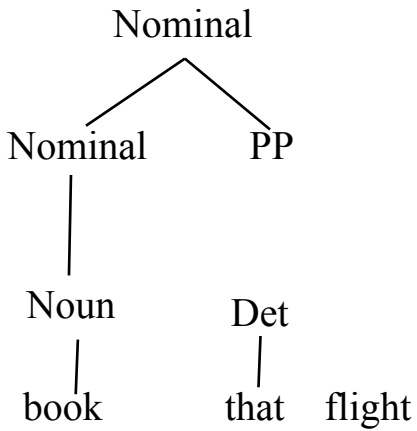
Grammar

S → NP VP
S → Aux NP VP
S → VP
NP → Pronoun
NP → Proper-Noun
NP → Det Nominal
Nominal → Noun
Nominal → Nominal Noun
Nominal → Nominal PP
VP → Verb
VP → Verb NP
VP → VP PP
PP → Prep NP

Lexicon

Det → the | that
Noun → book | flight
Verb → book
Pronoun → I | he | she | me
Proper-Noun → Houston | NWA
Aux → does
Prep → from | to

Bottom Up Parsing



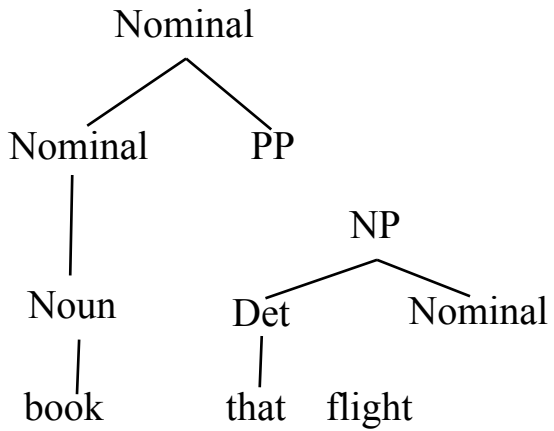
Grammar

S → NP VP
S → Aux NP VP
S → VP
NP → Pronoun
NP → Proper-Noun
NP → Det Nominal
Nominal → Noun
Nominal → Nominal Noun
Nominal → Nominal PP
VP → Verb
VP → Verb NP
VP → VP PP
PP → Prep NP

Lexicon

Det → the | that
Noun → book | flight
Verb → book
Pronoun → I | he | she | me
Proper-Noun → Houston | NWA
Aux → does
Prep → from | to

Bottom Up Parsing



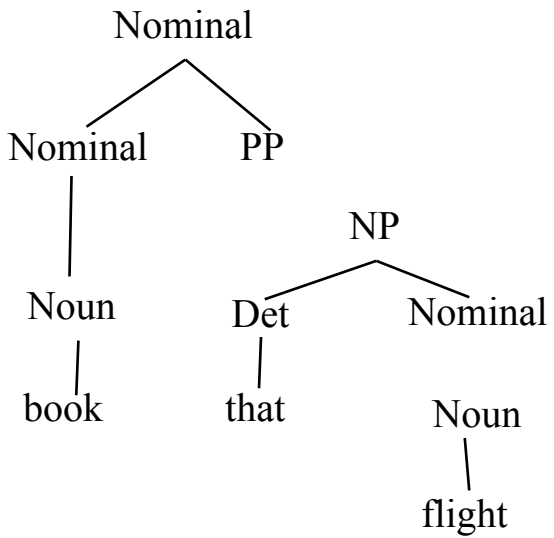
Grammar

S → NP VP
S → Aux NP VP
S → VP
NP → Pronoun
NP → Proper-Noun
NP → Det Nominal
Nominal → Noun
Nominal → Nominal Noun
Nominal → Nominal PP
VP → Verb
VP → Verb NP
VP → VP PP
PP → Prep NP

Lexicon

Det → the | that
Noun → book | flight
Verb → book
Pronoun → I | he | she | me
Proper-Noun → Houston | NWA
Aux → does
Prep → from | to

Bottom Up Parsing



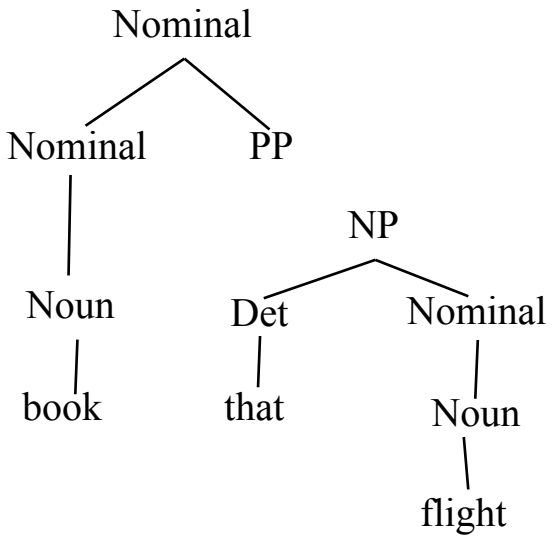
Grammar

S → NP VP
S → Aux NP VP
S → VP
NP → Pronoun
NP → Proper-Noun
NP → Det Nominal
Nominal → Noun
Nominal → Nominal Noun
Nominal → Nominal PP
VP → Verb
VP → Verb NP
VP → VP PP
PP → Prep NP

Lexicon

Det → the | that
Noun → book | flight
Verb → book
Pronoun → I | he | she | me
Proper-Noun → Houston | NWA
Aux → does
Prep → from | to

Bottom Up Parsing



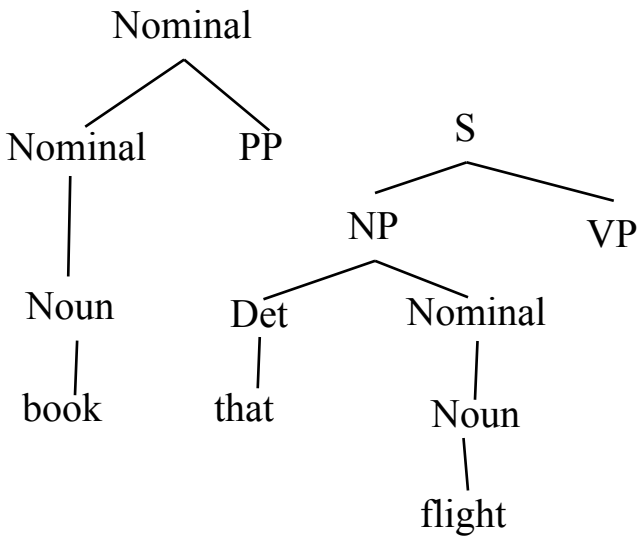
Grammar

S → NP VP
S → Aux NP VP
S → VP
NP → Pronoun
NP → Proper-Noun
NP → Det Nominal
Nominal → Noun
Nominal → Nominal Noun
Nominal → Nominal PP
VP → Verb
VP → Verb NP
VP → VP PP
PP → Prep NP

Lexicon

Det → the | that
Noun → book | flight
Verb → book
Pronoun → I | he | she | me
Proper-Noun → Houston | NWA
Aux → does
Prep → from | to

Bottom Up Parsing



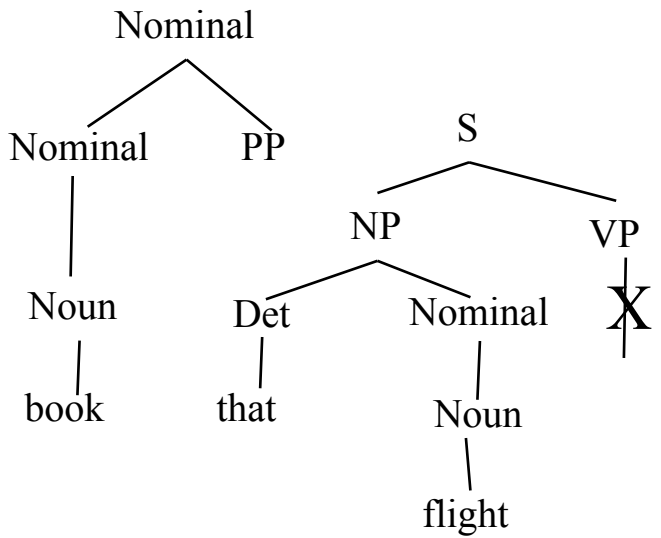
Grammar

S → NP VP
S → Aux NP VP
S → VP
NP → Pronoun
NP → Proper-Noun
NP → Det Nominal
Nominal → Noun
Nominal → Nominal Noun
Nominal → Nominal PP
VP → Verb
VP → Verb NP
VP → VP PP
PP → Prep NP

Lexicon

Det → the | that
Noun → book | flight
Verb → book
Pronoun → I | he | she | me
Proper-Noun → Houston | NWA
Aux → does
Prep → from | to

Bottom Up Parsing



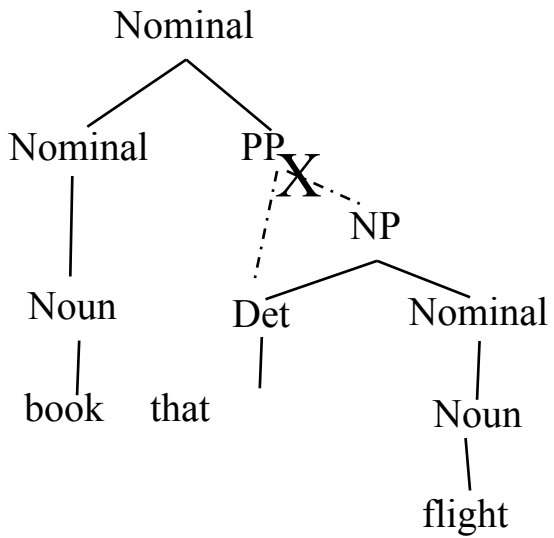
Grammar

S → NP VP
S → Aux NP VP
S → VP
NP → Pronoun
NP → Proper-Noun
NP → Det Nominal
Nominal → Noun
Nominal → Nominal Noun
Nominal → Nominal PP
VP → Verb
VP → Verb NP
VP → VP PP
PP → Prep NP

Lexicon

Det → the | that
Noun → book | flight
Verb → book
Pronoun → I | he | she | me
Proper-Noun → Houston | NWA
Aux → does
Prep → from | to

Bottom Up Parsing



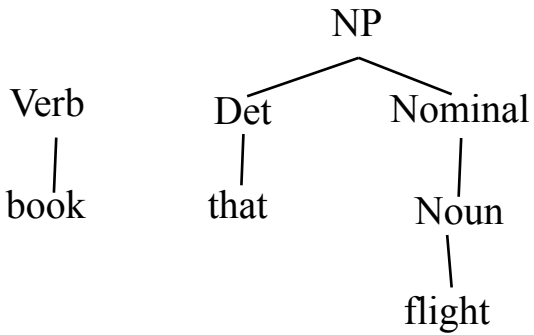
Grammar

S → NP VP
S → Aux NP VP
S → VP
NP → Pronoun
NP → Proper-Noun
NP → Det Nominal
Nominal → Noun
Nominal → Nominal Noun
Nominal → Nominal PP
VP → Verb
VP → Verb NP
VP → VP PP
PP → Prep NP

Lexicon

Det → the | that
Noun → book | flight
Verb → book
Pronoun → I | he | she | me
Proper-Noun → Houston | NWA
Aux → does
Prep → from | to

Bottom Up Parsing



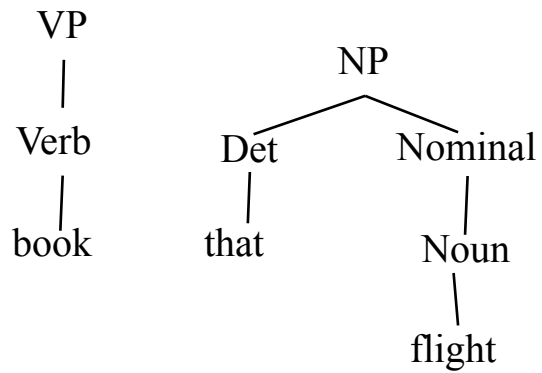
Grammar

S → NP VP
S → Aux NP VP
S → VP
NP → Pronoun
NP → Proper-Noun
NP → Det Nominal
Nominal → Noun
Nominal → Nominal Noun
Nominal → Nominal PP
VP → Verb
VP → Verb NP
VP → VP PP
PP → Prep NP

Lexicon

Det → the | that
Noun → book | flight
Verb → book
Pronoun → I | he | she | me
Proper-Noun → Houston | NWA
Aux → does
Prep → from | to

Bottom Up Parsing



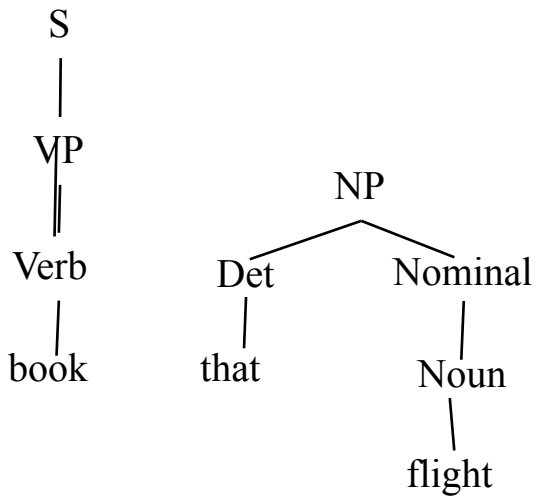
Grammar

S → NP VP
S → Aux NP VP
S → VP
NP → Pronoun
NP → Proper-Noun
NP → Det Nominal
Nominal → Noun
Nominal → Nominal Noun
Nominal → Nominal PP
VP → Verb
VP → Verb NP
VP → VP PP
PP → Prep NP

Lexicon

Det → the | that
Noun → book | flight
Verb → book
Pronoun → I | he | she | me
Proper-Noun → Houston | NWA
Aux → does
Prep → from | to

Bottom Up Parsing



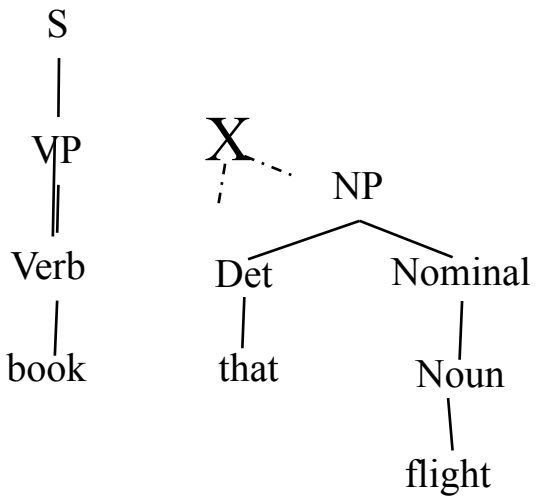
Grammar

S → NP VP
S → Aux NP VP
S → VP
NP → Pronoun
NP → Proper-Noun
NP → Det Nominal
Nominal → Noun
Nominal → Nominal Noun
Nominal → Nominal PP
VP → Verb
VP → Verb NP
VP → VP PP
PP → Prep NP

Lexicon

Det → the | that
Noun → book | flight
Verb → book
Pronoun → I | he | she | me
Proper-Noun → Houston | NWA
Aux → does
Prep → from | to

Bottom Up Parsing



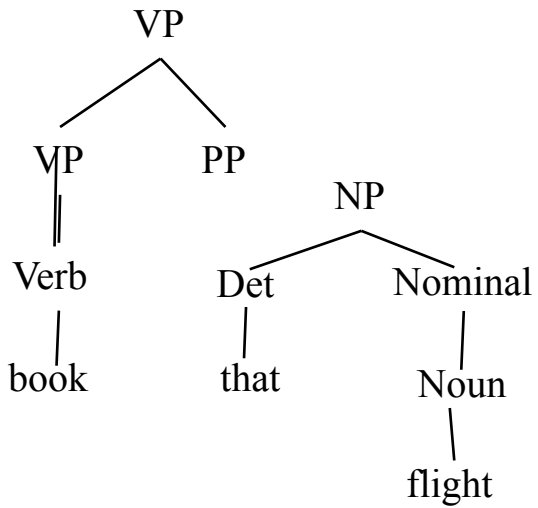
Grammar

S → NP VP
S → Aux NP VP
S → VP
NP → Pronoun
NP → Proper-Noun
NP → Det Nominal
Nominal → Noun
Nominal → Nominal Noun
Nominal → Nominal PP
VP → Verb
VP → Verb NP
VP → VP PP
PP → Prep NP

Lexicon

Det → the | that
Noun → book | flight
Verb → book
Pronoun → I | he | she | me
Proper-Noun → Houston | NWA
Aux → does
Prep → from | to

Bottom Up Parsing



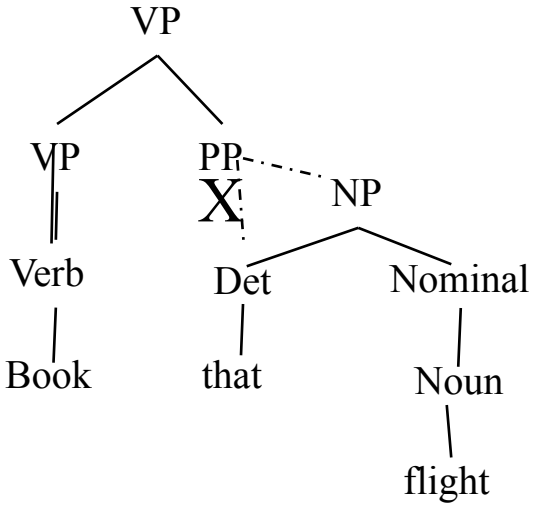
Grammar

S → NP VP
S → Aux NP VP
S → VP
NP → Pronoun
NP → Proper-Noun
NP → Det Nominal
Nominal → Noun
Nominal → Nominal Noun
Nominal → Nominal PP
VP → Verb
VP → Verb NP
VP → VP PP
PP → Prep NP

Lexicon

Det → the | that
Noun → book | flight
Verb → book
Pronoun → I | he | she | me
Proper-Noun → Houston | NWA
Aux → does
Prep → from | to

Bottom Up Parsing



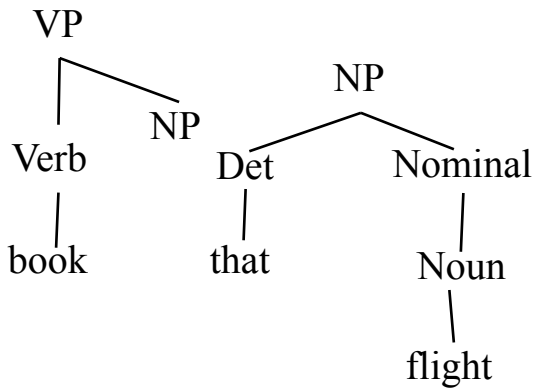
Grammar

S → NP VP
S → Aux NP VP
S → VP
NP → Pronoun
NP → Proper-Noun
NP → Det Nominal
Nominal → Noun
Nominal → Nominal Noun
Nominal → Nominal PP
VP → Verb
VP → Verb NP
VP → VP PP
PP → Prep NP

Lexicon

Det → the | that
Noun → book | flight
Verb → book
Pronoun → I | he | she | me
Proper-Noun → Houston | NWA
Aux → does
Prep → from | to

Bottom Up Parsing



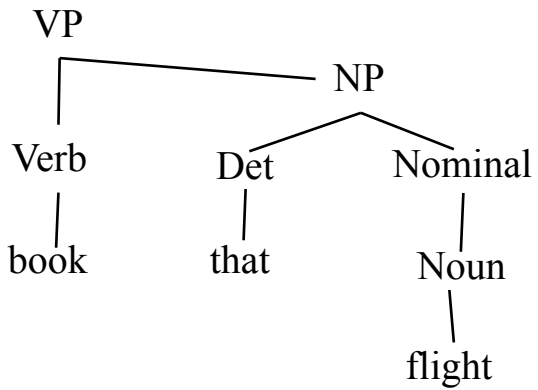
Grammar

S → NP VP
S → Aux NP VP
S → VP
NP → Pronoun
NP → Proper-Noun
NP → Det Nominal
Nominal → Noun
Nominal → Nominal Noun
Nominal → Nominal PP
VP → Verb
VP → Verb NP
VP → VP PP
PP → Prep NP

Lexicon

Det → the | that
Noun → book | flight
Verb → book
Pronoun → I | he | she | me
Proper-Noun → Houston | NWA
Aux → does
Prep → from | to

Bottom Up Parsing



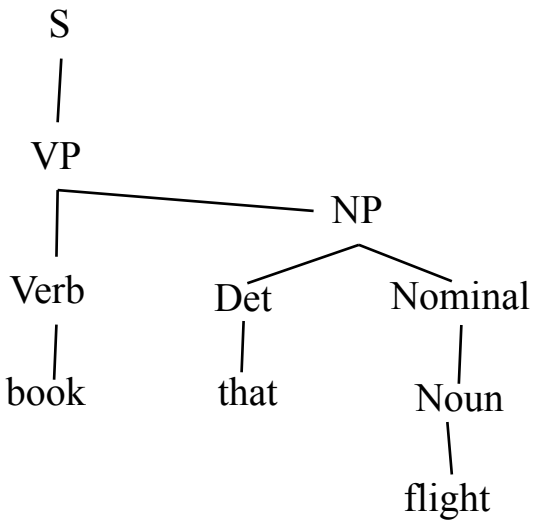
Grammar

S → NP VP
S → Aux NP VP
S → VP
NP → Pronoun
NP → Proper-Noun
NP → Det Nominal
Nominal → Noun
Nominal → Nominal Noun
Nominal → Nominal PP
VP → Verb
VP → Verb NP
VP → VP PP
PP → Prep NP

Lexicon

Det → the | that
Noun → book | flight
Verb → book
Pronoun → I | he | she | me
Proper-Noun → Houston | NWA
Aux → does
Prep → from | to

Bottom Up Parsing



Grammar

S → NP VP
S → Aux NP VP
S → VP
NP → Pronoun
NP → Proper-Noun
NP → Det Nominal
Nominal → Noun
Nominal → Nominal Noun
Nominal → Nominal PP
VP → Verb
VP → Verb NP
VP → VP PP
PP → Prep NP

Lexicon

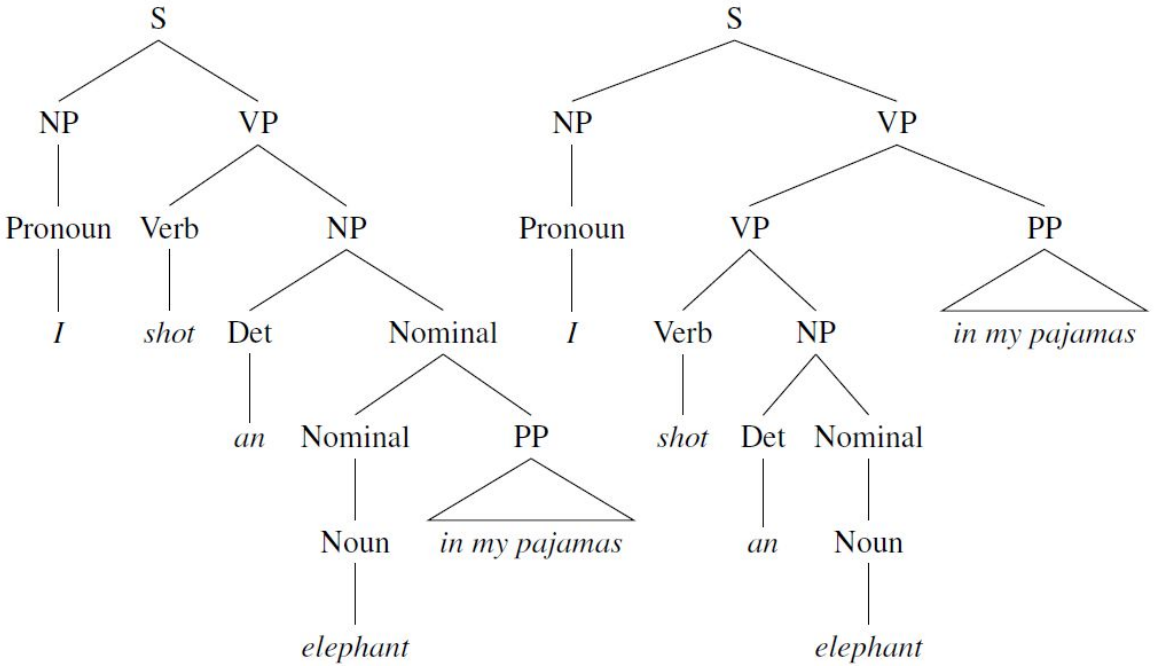
Det → the | that
Noun → book | flight
Verb → book
Pronoun → I | he | she | me
Proper-Noun → Houston | NWA
Aux → does
Prep → from | to

Top Down vs. Bottom Up

- Top down never explores options that will not lead to a full parse for the current sentence, but can explore many options that never connect to the actual sentence.
- Bottom up never explores options that do not connect to the actual sentence but can explore options that can never lead to a full parse.
- Relative amounts of wasted search depend on how much the grammar branches in each direction.

Structural Ambiguity

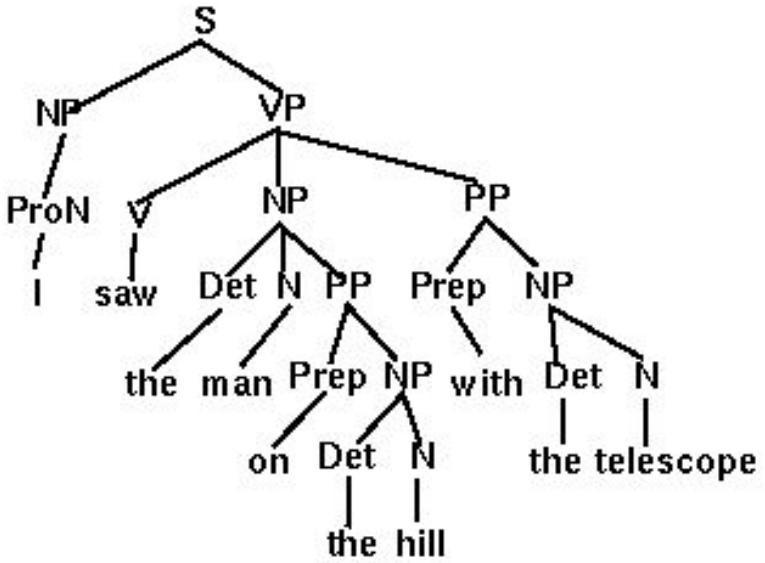
- "I shot an elephant in my pyjamas"

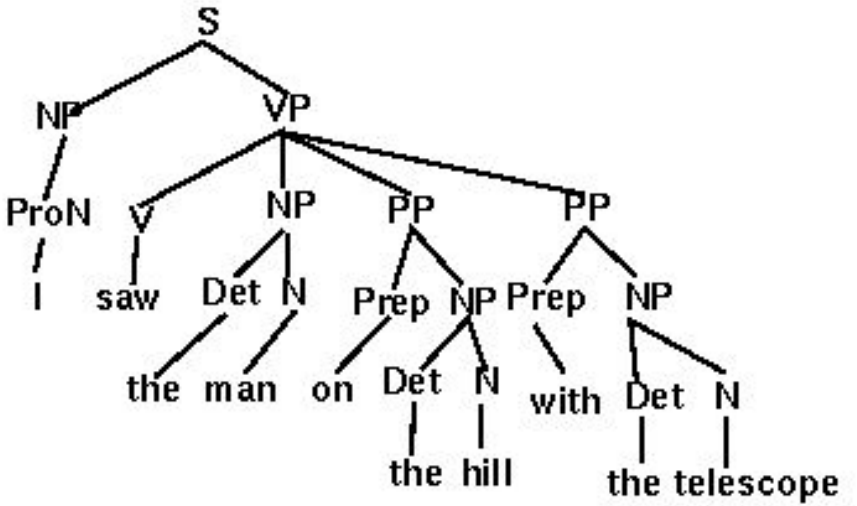


Grammar
$S \rightarrow NP\ VP$
$S \rightarrow Aux\ NP\ VP$
$S \rightarrow VP$
$NP \rightarrow Pronoun$
$NP \rightarrow Proper\text{-}Noun$
$NP \rightarrow Det\ Nominal$
$Nominal \rightarrow Noun$
$Nominal \rightarrow Nominal\ Noun$
$Nominal \rightarrow Nominal\ PP$
$VP \rightarrow Verb$
$VP \rightarrow Verb\ NP$
$VP \rightarrow Verb\ NP\ PP$
$VP \rightarrow Verb\ PP$
$VP \rightarrow VP\ PP$
$PP \rightarrow Preposition\ NP$

Structural Ambiguity

- "I saw the man on the hill with telescope"





Syntactic Disambiguation

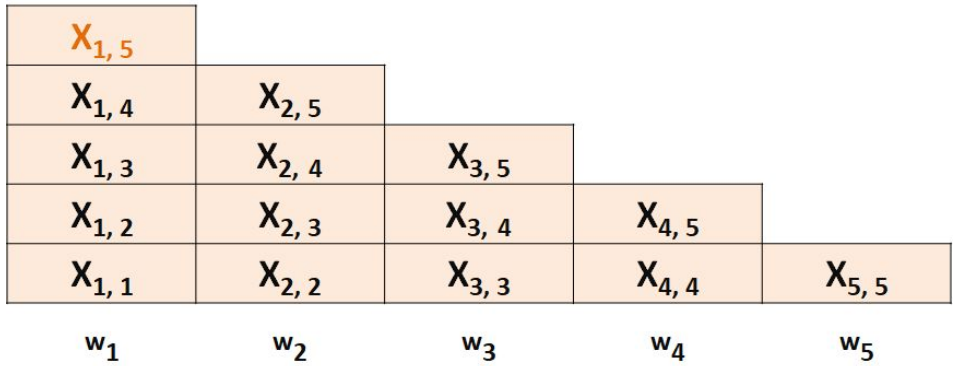
- The process of choosing a single correct parse from the multitude of possible parses is called syntactic disambiguation
- Effective disambiguation algorithms require statistical, semantic, and contextual knowledge sources
- How to efficiently handle structural ambiguities?

CYK Parsing : A Dynamic Programming Approach

- Requires the grammar to be in Chomsky Normal Form (CNF)
  - restricted to rules of the form $A \rightarrow B\ C$ or $A \rightarrow w$

- CYK(Cocke–Younger–Kasami) Algorithm
  - Convert CFG to CNF
  - Construct a triangular table
  - Check if $X_{1,n}$ contains start symbol S

- Time Complexity
  - **$O(n^3.|G|)$**
  - **n** : length of string
  - **|G|** : size of the CNF Grammar



CKY parsing example

1. S -> NP VP
2. NP -> Det N
3. NP -> NP PP
4. VP -> V NP
5. VP -> VP PP
6. PP -> P NP
7. NP -> Tom
8. N -> Ball
9. N -> Net
10. V -> Net
11. V -> Sent
12. P -> to
13. Det -> The
14. Det -> a

Tom sent the ball to the net

Tom (1)	Sent (2)	The (3)	Ball (4)	To (5)	The (6)	Net (7)

Statistical Constituency Parsing

- Can CYK Algorithm resolve ambiguity? NO
  - It can only represent ambiguities
- How to resolve ambiguity?
  - Compute the probability of each interpretation and choose the most probable interpretation
  - Need a probabilistic parser

Pseudocode

```
function CKY-PARSE(words, grammar) returns table

  for j ← from 1 to LENGTH(words) do
      for all {A | A → words[j] ∈ grammar}
              table[j − 1, j] ← table[j − 1, j] ∪ A
      for i ← from j − 2 downto 0 do
          for k ← i + 1 to j − 1 do
              for all {A | A → BC ∈ grammar and B ∈ table[i,k] and C ∈ table[k, j]}
                      table[i,j] ← table[i,j] ∪ A
```

```
function PROBABILISTIC-CKY(words,grammar) returns most probable parse
                                                and its probability
  for j ← from 1 to LENGTH(words) do
    for all { A | A → words[j] ∈ grammar}
      table[j − 1, j,A] ← P(A → words[j])
    for i ← from j − 2 downto 0 do
      for k ← i + 1 to j − 1 do
          for all { A | A → BC ∈ grammar,
                          and table[i,k,B] > 0 and table[k, j,C] > 0 }
              if (table[i,j,A] < P(A → BC) × table[i,k,B] × table[k,j,C]) then
                  table[i,j,A] ← P(A → BC) × table[i,k,B] × table[k,j,C]
                  back[i,j,A] ← {k,B,C}
    return BUILD_TREE(back[1, LENGTH(words), S]), table[1, LENGTH(words), S]
```

Probabilistic Context-Free Grammars (PCFG)

- Probabilistic augmentation of context-free grammars
- Differs from CFG by augmenting each rule in *R* with a conditional probability
  - $\boldsymbol{\alpha \rightarrow \beta}$ [*p*]
  - *p* is the conditional probability of $\boldsymbol{\beta}$ given $\boldsymbol{\alpha}$
  - Defined as $P(\alpha \rightarrow \beta)$ or $P(\alpha \rightarrow \beta | \alpha)$
  - $\sum_i P(\alpha \rightarrow \beta_i) = 1$

- How to learn the probabilities?

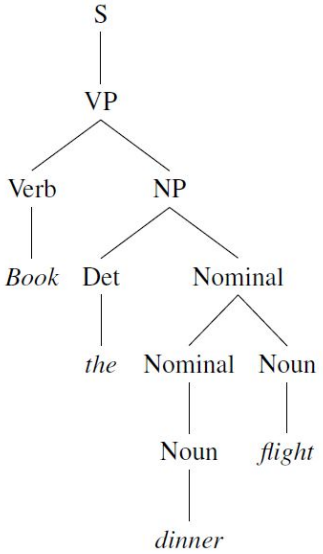
$$P(\alpha \rightarrow \beta | \alpha) = \frac{\text{Count}(\alpha \rightarrow \beta)}{\sum_\gamma \text{Count}(\alpha \rightarrow \gamma)} = \frac{\text{Count}(\alpha \rightarrow \beta)}{\text{Count}(\alpha)}$$

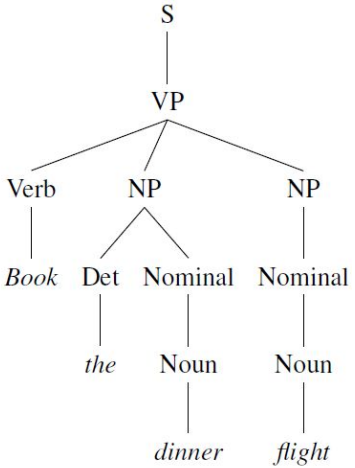
Grammar		Lexicon
S → *NP VP*	[.80]	*Det* → *that* [.10] | *a* [.30] | *the* [.60]
S → *Aux NP VP*	[.15]	*Noun* → *book* [.10] | *flight* [.30]
S → *VP*	[.05]	| *meal* [.05] | *money* [.05]
NP → *Pronoun*	[.35]	| *flight* [.40] | *dinner* [.10]
NP → *Proper-Noun*	[.30]	*Verb* → *book* [.30] | *include* [.30]
NP → *Det Nominal*	[.20]	| *prefer* [.40]
NP → *Nominal*	[.15]	*Pronoun* → *I* [.40] | *she* [.05]
Nominal → *Noun*	[.75]	| *me* [.15] | *you* [.40]
Nominal → *Nominal Noun*	[.20]	*Proper-Noun* → *Houston* [.60]
Nominal → *Nominal PP*	[.05]	| *NWA* [.40]
VP → *Verb*	[.35]	*Aux* → *does* [.60] | *can* [.40]
VP → *Verb NP*	[.20]	*Preposition* → *from* [.30] | *to* [.30]
VP → *Verb NP PP*	[.10]	| *on* [.20] | *near* [.15]
VP → *Verb PP*	[.15]	| *through* [.05]
VP → *Verb NP NP*	[.05]	
VP → *VP PP*	[.15]	
PP → *Preposition NP*	[1.0]	

PCFGs for disambiguation

Choose the parse tree with highest probability

Rules			P	Rules			P
S	→	VP	.05	S	→	VP	.05
VP	→	Verb NP	.20	VP	→	Verb NP NP	.10
NP	→	Det Nominal	.20	NP	→	Det Nominal	.20
Nominal	→	Nominal Noun	.20	NP	→	Nominal	.15
Nominal	→	Noun	.75	Nominal	→	Noun	.75
				Nominal	→	Noun	.75
Verb	→	book	.30	Verb	→	book	.30
Det	→	the	.60	Det	→	the	.60
Noun	→	dinner	.10	Noun	→	dinner	.10
Noun	→	flight	.40	Noun	→	flight	.40



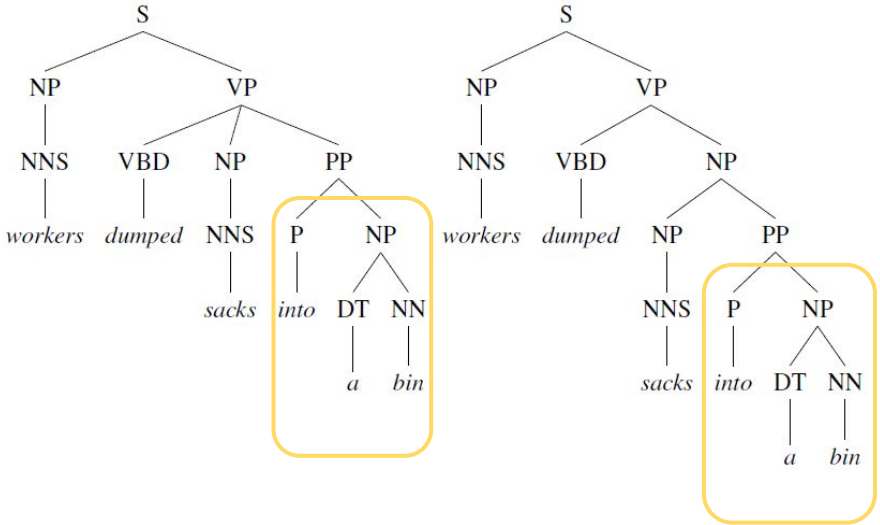


$$P(T_{left}) = .05 * .20 * .20 * .20 * .75 * .30 * .60 * .10 * .40 = \mathbf{2.2 \times 10^{-6}}$$

$$P(T_{right}) = .05 * .10 * .20 * .15 * .75 * .75 * .30 * .60 * .10 * .40 = \mathbf{6.1 \times 10^{-7}}$$

Problems with PCFG

- Poor Independence assumption
  - Expansion of a non-terminal is independent of the context
  - Expansion can depend on the location of the node in the parse tree
- Lack of sensitivity to Lexical dependencies
  - Example : Prepositional phrase(PP) attachment
  - Solution : Probabilistic Lexicalized CFGs

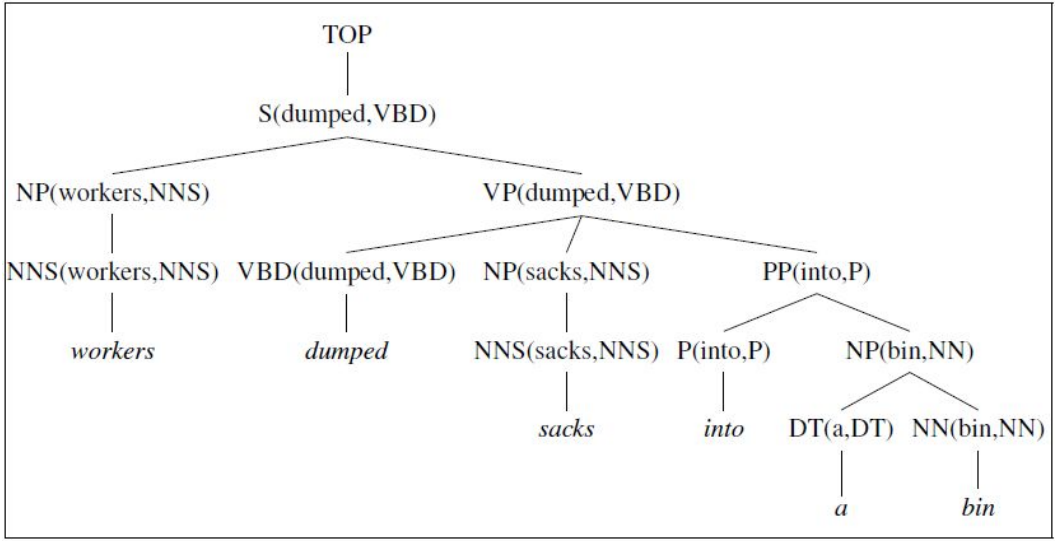


Two possible parse trees for a **PP attachment ambiguity**

- The left parse is the sensible one where "into a bin" describes the resulting location of the sacks
- In the right incorrect parse, the sacks to be dumped are the ones which are already "into a bin"

Probabilistic Lexicalised CFG

- Each non-terminal grammar in the tree is annotated
  - With Lexical head and Part-of-speech tags of the head word
- Parsing Method
  - Collins Parser



Internal Rules				Lexical Rules		
TOP	→	S(dumped,VBD)		NNS(workers,NNS)	→	workers
S(dumped,VBD)	→	NP(workers,NNS)	VP(dumped,VBD)	VBD(dumped,VBD)	→	dumped
NP(workers,NNS)	→	NNS(workers,NNS)		NNS(sacks,NNS)	→	sacks
VP(dumped,VBD)	→	VBD(dumped, VBD)	NP(sacks,NNS) PP(into,P)	P(into,P)	→	into
PP(into,P)	→	P(into,P)	NP(bin,NN)	DT(a,DT)	→	a
NP(bin,NN)	→	DT(a,DT)	NN(bin,NN)	NN(bin,NN)	→	bin

Parser Evaluation

- Precision

$$\textbf{labeled precision:} = \frac{\text{\# of correct constituents in hypothesis parse of } s}{\text{\# of total constituents in hypothesis parse of } s}$$

- Recall

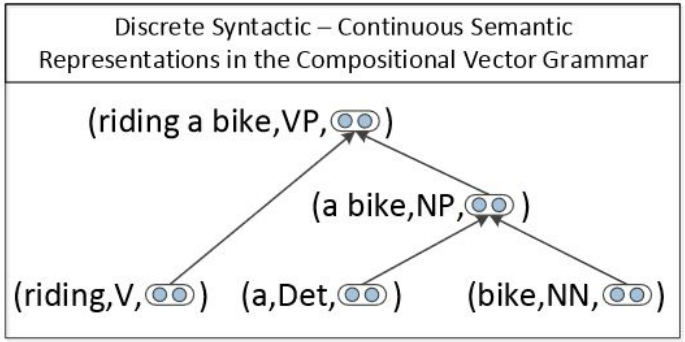
$$\textbf{labeled recall:} = \frac{\text{\# of correct constituents in hypothesis parse of } s}{\text{\# of correct constituents in reference parse of } s}$$

- F-Measure

$$F_\beta = \frac{(\beta^2 + 1)PR}{\beta^2 P + R}$$

  - $\boldsymbol{\beta} > 1$ favor recall
  - $\boldsymbol{\beta} < 1$ favor precision

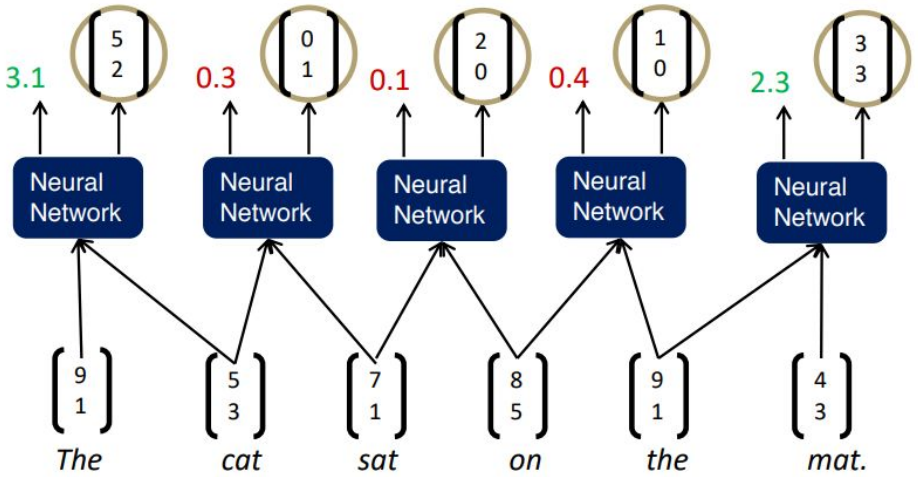
Scoring and selecting the constituents



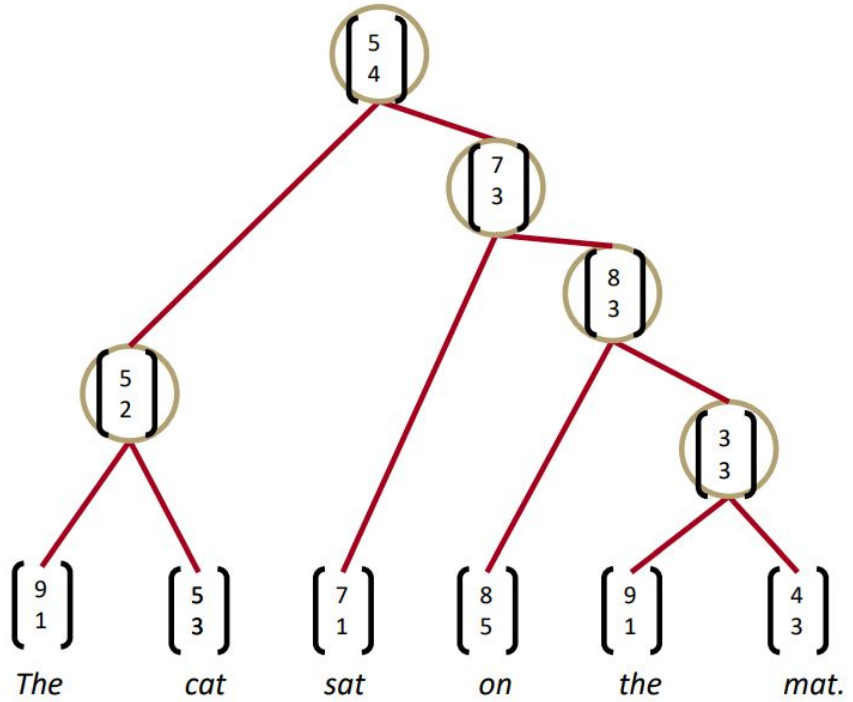
Dimensionality of $W_{A,B}$?

- Parse tree construction in recursive manner
- The next two nodes to be combined is decided based on greedy choice
  - For each pair of consecutive nodes, find a score
  - Pick the pair with highest score
  - Merge them and replace the two constituent nodes with this merged nodes
- Scoring
  - Based on the representations of the nodes, their labels and PCFG probability
- Score
  - $p_{X,a,b} = f\left(W_{A,B}\begin{bmatrix}a\\b\end{bmatrix} + bias_{A,B}\right) + P(X \rightarrow AB)$

Parsing with Recursive NN (greedily)



Parsing with Recursive NN (greedily)



Partial Parsing : Chunking

- Not all tasks require complete parse trees
  - Information extraction

- Chunking is the process of identifying and classifying the flat, non-overlapping segments of a sentence

- Objective of Chunking
  - **Segmentation**
    - Finding the non-overlapping extents of the chunks
    - Some input words may not be part of any chunk
  - **Labeling**
    - Assigning the correct tag to the discovered chunks

- Example
  - [$_{\text{NP}}$The old gentleman] from [$_{\text{NP}}$the neighboring village] is an amazing [$_{\text{NP}}$ artist].

More examples of chunking

- The morning flight from Denver has arrived

NP PP NP VP

- The morning flight from Denver has arrived

B_NP I_NP I_NP B_PP B_NP B_VP I_VP

- Let the equation of the circle be $ax^2 + by^2 = c.$

M

- The $sigmoid$ activation is applied on the output $y = Wx + b$ of the final layer

O_M B_M O_M O_M O_M O_M O_M O_M B_M I_M I_M I_M I_M O_M O_M O_M O_M

Chunking Model

- IOB Tagging
  - **B** : Beginning of a chunk
  - **I** : Inside of a chunk
  - **O** : Outside of any chunk
- Chunking reduces to tagging task
  - Can use any supervised classification model from ML
- Example features:
  - w_i, w_{i-1}, w_{i+1}, POS(w_i), POS(w_{i-1}), POS(w_{i+1}), punctuations, beginning-of-sentence?, end-of-sentence? Etc.