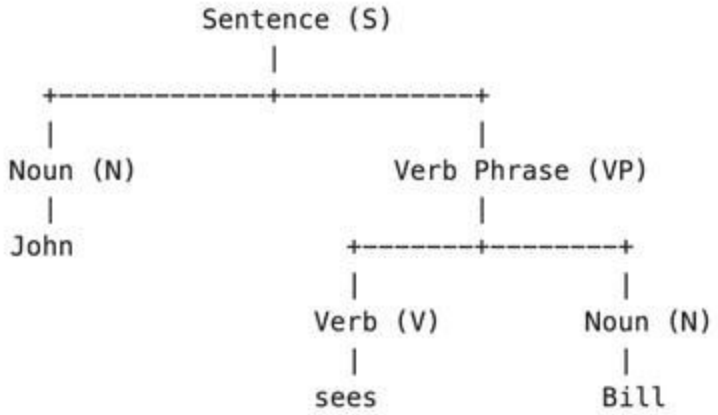
Constituency Parsing

Subtle Constraints in Descriptive Grammar

How do we explain these examples? (* indicates unacceptability)

- Bender, Sag, Wasow's examples
  - F— yourself!
  - Go f— yourself!
  - F— you!
  - *Go f— you!
- Wanna contraction (from Wikipedia)
  - Who does Vicky want to vote for?
    $\Rightarrow$ Who does Vicky wanna vote for?
  - Who does Vicky want to win?
    $\Rightarrow$ *Who does Vicky wanna win
- Gonna contraction
  - I am gonna get lunch
  - *I am gonna New York
- Gonna and wanna function like AUX verbs

Why do we need grammar in English?


Sentence (S)
Noun (N)
Verb Phrase (VP)
John
Verb (V)
Noun (N)
sees
Bill

Context-Free Grammar

- CFGs are needed to handle natural structure in human languages: think of matching parentheses
- Bender, Sag, Wasow's example:
  - That Sandy left bothered me
  - That that Sandy left bothered me bothered Kim
  - That that that Sandy left bothered me bothered Kim bothered Bo
- A grammar describes (and generates) all and only the valid finite strings over a given alphabet
- For NL, the alphabet is words or tokens in a lexicon (Jurafsky seems to use "lexicon" oddly in this setting)

How did we define NL in the first lecture?

CFG Example Sentence: I prefer a morning flight

- Initial grammar and lexicon to derive the above sentence

$$
\begin{array}{l}
S \longrightarrow NP\ VP \\
NP \longrightarrow Pronoun \mid Determiner\ Nominal \\
VP \longrightarrow Verb\ NP \\
Nominal \longrightarrow Nominal\ Noun \mid Noun
\end{array}
$$

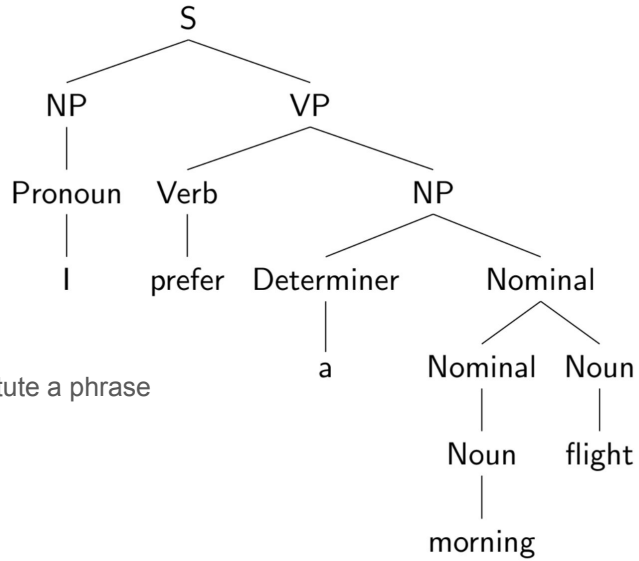
$$
\begin{array}{l}
Pronoun \longrightarrow I \\
Verb \longrightarrow prefer \\
Determiner \longrightarrow a \\
Noun \longrightarrow morning \mid flight
\end{array}
$$

CFG Example Sentence: I prefer a morning flight

- Initial grammar and lexicon to derive the above sentence

$S \longrightarrow NP\ VP$
$NP \longrightarrow Pronoun \mid Determiner\ Nominal$
$VP \longrightarrow Verb\ NP$
$Nominal \longrightarrow Nominal\ Noun \mid Noun$

$Pronoun \longrightarrow I$
$Verb \longrightarrow prefer$
$Determiner \longrightarrow a$
$Noun \longrightarrow morning \mid flight$



Multiple terminals or non terminals constitute a phrase

Draw a Parse Tree

I prefer leaving Boston in the morning

$S \longrightarrow NP\ VP$
$NP \longrightarrow Pronoun \mid Determiner\ Nominal$
$VP \longrightarrow Verb\ NP$
$Nominal \longrightarrow Nominal\ Noun \mid Noun$

$Pronoun \longrightarrow I$
$Verb \longrightarrow prefer$
$Determiner \longrightarrow a$
$Noun \longrightarrow morning \mid flight$

- For additional sentences, we could insert

$VP \longrightarrow Verb\ VP\ Nominal\ PP$ (leaving Boston in the morning)
$VP \longrightarrow VP\ PP$ (leaving in the morning)
$PP \longrightarrow Preposition\ NP$ (from Boston)

Sentences in English

- Declarative $\sim$ default form
  - S $\longrightarrow$ NP VP (here NP is the subject)
- Imperative, S $\longrightarrow$ VP
  - Usually, lack a subject "Go there"
  - But not always "You go there"
  - Subject *deletion* under a view that there is a subject
- Yes-no question, S $\longrightarrow$ Aux NP VP
  - Begin with auxiliary verb
  - Retain a main verb
- Wh-structures
  - In modern English, who, whose, when, where, what, which, how, why; also: whence, whereby, wherein
  - Contain a wh-phrase

Wh Structures

- Wh-subject question, S $\longrightarrow$ Wh-NP VP
  - What airlines fly from Burbank to Denver?
  - The wh-phrase yields the subject
  - Wh-NP $\longrightarrow$ Wh-Pronoun (who, whom, whose, which)
  - Wh-NP $\longrightarrow$ Wh-Determiner NP (what, which)
- Wh-non-subject question, S $\longrightarrow$ Wh-NP Aux NP VP
  - What flights do you have from Burbank to Denver?
  - The wh-phrase is not the subject of the sentence, which is something else
  - Long-distance dependencies

Penn Treebank Tagset

Grammar	Lexicon
$S \rightarrow NP\ VP$	$Det \rightarrow that \mid this \mid the \mid a$
$S \rightarrow Aux\ NP\ VP$	$Noun \rightarrow book \mid flight \mid meal \mid money$
$S \rightarrow VP$	$Verb \rightarrow book \mid include \mid prefer$
$NP \rightarrow Pronoun$	$Pronoun \rightarrow I \mid she \mid me$
$NP \rightarrow Proper\text{-}Noun$	$Proper\text{-}Noun \rightarrow Houston \mid NWA$
$NP \rightarrow Det\ Nominal$	$Aux \rightarrow does$
$Nominal \rightarrow Noun$	$Preposition \rightarrow from \mid to \mid on \mid near \mid through$
$Nominal \rightarrow Nominal\ Noun$	
$Nominal \rightarrow Nominal\ PP$	
$VP \rightarrow Verb$	
$VP \rightarrow Verb\ NP$	
$VP \rightarrow Verb\ NP\ PP$	
$VP \rightarrow Verb\ PP$	
$VP \rightarrow VP\ PP$	
$PP \rightarrow Preposition\ NP$	

How to check if two grammars are equivalent?

Grammar Equivalence and Normal Form

- Weak equivalence: generate the same strings
- Strong equivalence
  - Weak plus assign the same phrase structure (up to renaming of nonterminals)
- Chomsky Normal Form, in which productions are of these forms:
  - Two at a time: $A \longrightarrow B\ C$
  - Single terminal: $A \longrightarrow a$
  - Not generating the empty string: Exclude $A \longrightarrow \varepsilon$

Let's convert L1 to CNF?

$\mathscr{L}_1$ Grammar	$\mathscr{L}_1$ in CNF
$S \rightarrow NP\ VP$	$S \rightarrow NP\ VP$
$S \rightarrow Aux\ NP\ VP$	$S \rightarrow X1\ VP$
	$X1 \rightarrow Aux\ NP$
$S \rightarrow VP$	$S \rightarrow book \mid include \mid prefer$
	$S \rightarrow Verb\ NP$
	$S \rightarrow X2\ PP$
	$S \rightarrow Verb\ PP$
	$S \rightarrow VP\ PP$
$NP \rightarrow Pronoun$	$NP \rightarrow I \mid she \mid me$
$NP \rightarrow Proper\text{-}Noun$	$NP \rightarrow TWA \mid Houston$
$NP \rightarrow Det\ Nominal$	$NP \rightarrow Det\ Nominal$
$Nominal \rightarrow Noun$	$Nominal \rightarrow book \mid flight \mid meal \mid money$
$Nominal \rightarrow Nominal\ Noun$	$Nominal \rightarrow Nominal\ Noun$
$Nominal \rightarrow Nominal\ PP$	$Nominal \rightarrow Nominal\ PP$
$VP \rightarrow Verb$	$VP \rightarrow book \mid include \mid prefer$
$VP \rightarrow Verb\ NP$	$VP \rightarrow Verb\ NP$
$VP \rightarrow Verb\ NP\ PP$	$VP \rightarrow X2\ PP$
	$X2 \rightarrow Verb\ NP$
$VP \rightarrow Verb\ PP$	$VP \rightarrow Verb\ PP$
$VP \rightarrow VP\ PP$	$VP \rightarrow VP\ PP$
$PP \rightarrow Preposition\ NP$	$PP \rightarrow Preposition\ NP$

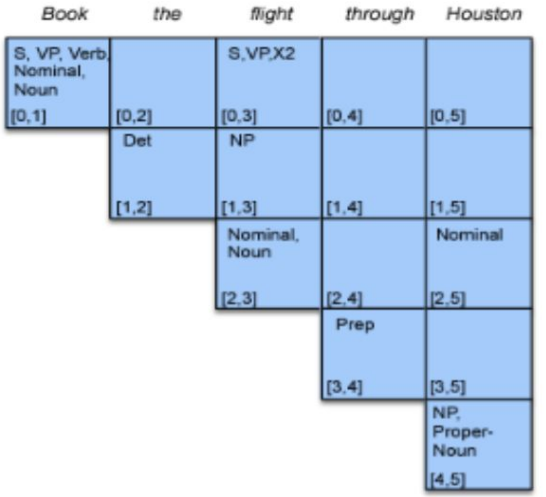
Figure 18.10 $\mathscr{L}_1$ Grammar and its conversion to CNF. Note that although they aren't shown here, all the original lexical entries from $\mathscr{L}_1$ carry over unchanged as well.

How to check if a sentence is syntactically correct?

$\mathscr{L}_1$ Grammar	$\mathscr{L}_1$ in CNF
$S \rightarrow NP\ VP$	$S \rightarrow NP\ VP$
$S \rightarrow Aux\ NP\ VP$	$S \rightarrow X1\ VP$
	$X1 \rightarrow Aux\ NP$
$S \rightarrow VP$	$S \rightarrow book \mid include \mid prefer$
	$S \rightarrow Verb\ NP$
	$S \rightarrow X2\ PP$
	$S \rightarrow Verb\ PP$
	$S \rightarrow VP\ PP$
$NP \rightarrow Pronoun$	$NP \rightarrow I \mid she \mid me$
$NP \rightarrow Proper\text{-}Noun$	$NP \rightarrow TWA \mid Houston$
$NP \rightarrow Det\ Nominal$	$NP \rightarrow Det\ Nominal$
$Nominal \rightarrow Noun$	$Nominal \rightarrow book \mid flight \mid meal \mid money$
$Nominal \rightarrow Nominal\ Noun$	$Nominal \rightarrow Nominal\ Noun$
$Nominal \rightarrow Nominal\ PP$	$Nominal \rightarrow Nominal\ PP$
$VP \rightarrow Verb$	$VP \rightarrow book \mid include \mid prefer$
$VP \rightarrow Verb\ NP$	$VP \rightarrow Verb\ NP$
$VP \rightarrow Verb\ NP\ PP$	$VP \rightarrow X2\ PP$
	$X2 \rightarrow Verb\ NP$
$VP \rightarrow Verb\ PP$	$VP \rightarrow Verb\ PP$
$VP \rightarrow VP\ PP$	$VP \rightarrow VP\ PP$
$PP \rightarrow Preposition\ NP$	$PP \rightarrow Preposition\ NP$

Figure 18.10 $\mathscr{L}_1$ Grammar and its conversion to CNF. Note that although they aren't shown here, all the original lexical entries from $\mathscr{L}_1$ carry over unchanged as well.

CKY (Cocke-Kasami-Younger) Algorithm



Which algorithmic paradigm is it?

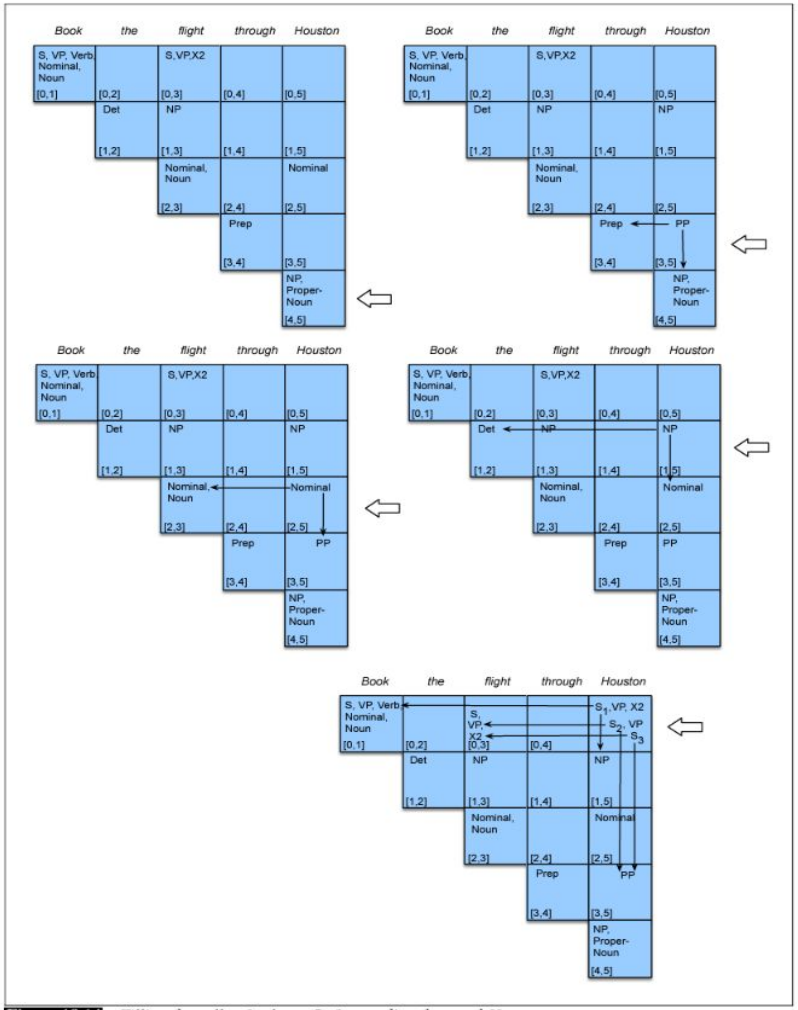


Figure 18.14 Filling the cells of column 5 after reading the word *Houston*.

Dynamic programing!

Time complexity?

$\mathscr{L}_1$ Grammar
$S \rightarrow NP\ VP$
$S \rightarrow Aux\ NP\ VP$
$S \rightarrow VP$
$NP \rightarrow Pronoun$
$NP \rightarrow Proper\text{-}Noun$
$NP \rightarrow Det\ Nominal$
$Nominal \rightarrow Noun$
$Nominal \rightarrow Nominal\ Noun$
$Nominal \rightarrow Nominal\ PP$
$VP \rightarrow Verb$
$VP \rightarrow Verb\ NP$
$VP \rightarrow Verb\ NP\ PP$
$VP \rightarrow Verb\ PP$
$VP \rightarrow VP\ PP$
$PP \rightarrow Preposition\ NP$

CKY - O(n^3 times |G|)

Let’s parse “I shot an elephant in my pajamas”

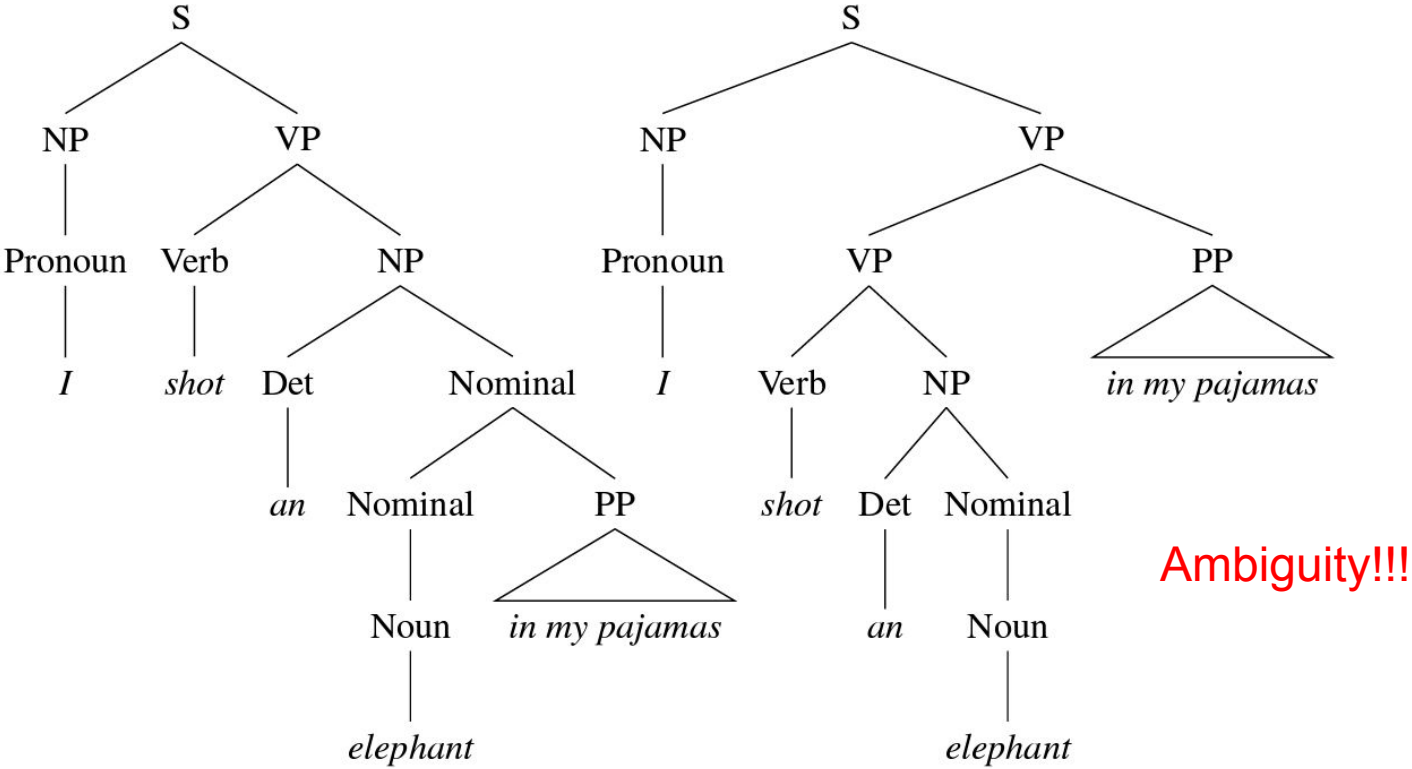
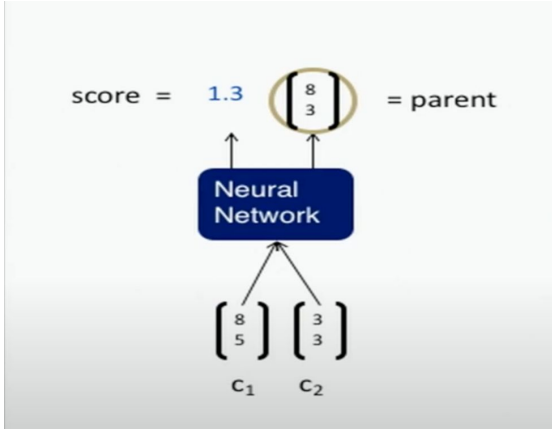


Figure 18.9 Two parse trees for an ambiguous sentence. The parse on the left corresponds to the humorous reading in which the elephant is in the pajamas, the parse on the right corresponds to the reading in which Captain Spaulding did the shooting in his pajamas.

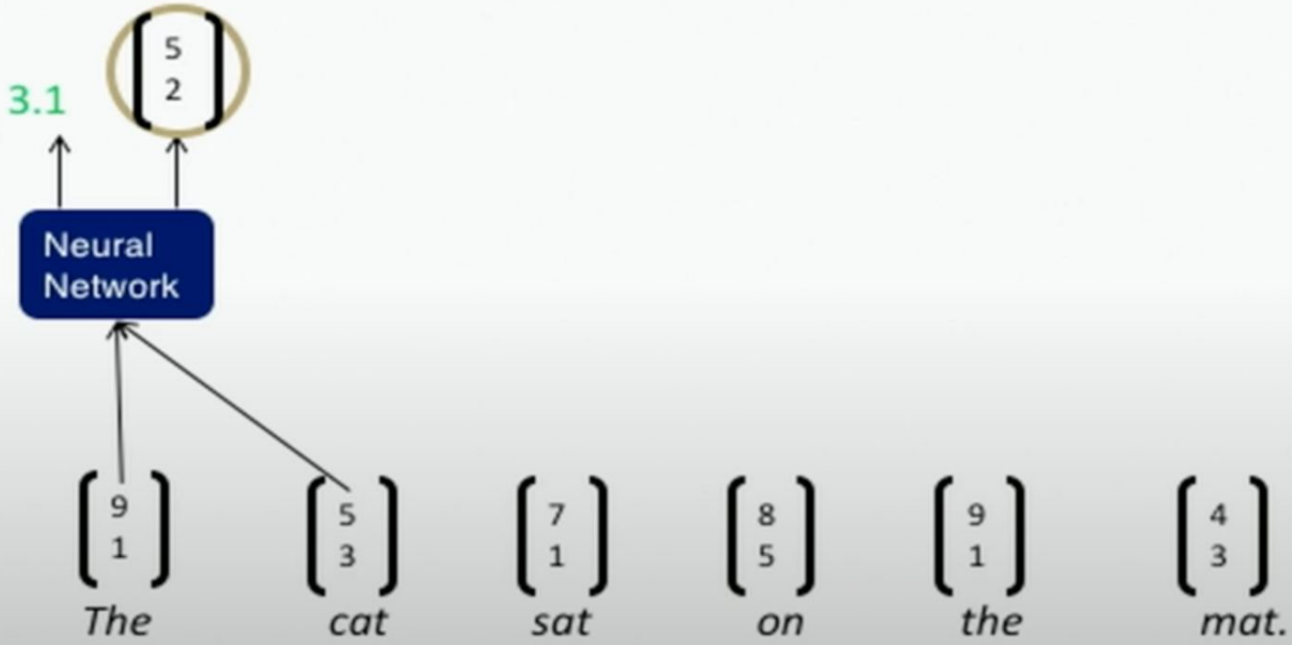
How to decide the constituency?

Input: Two children representation
Output: Semantic representation if the two nodes are merged
Score of how plausible the new node would be

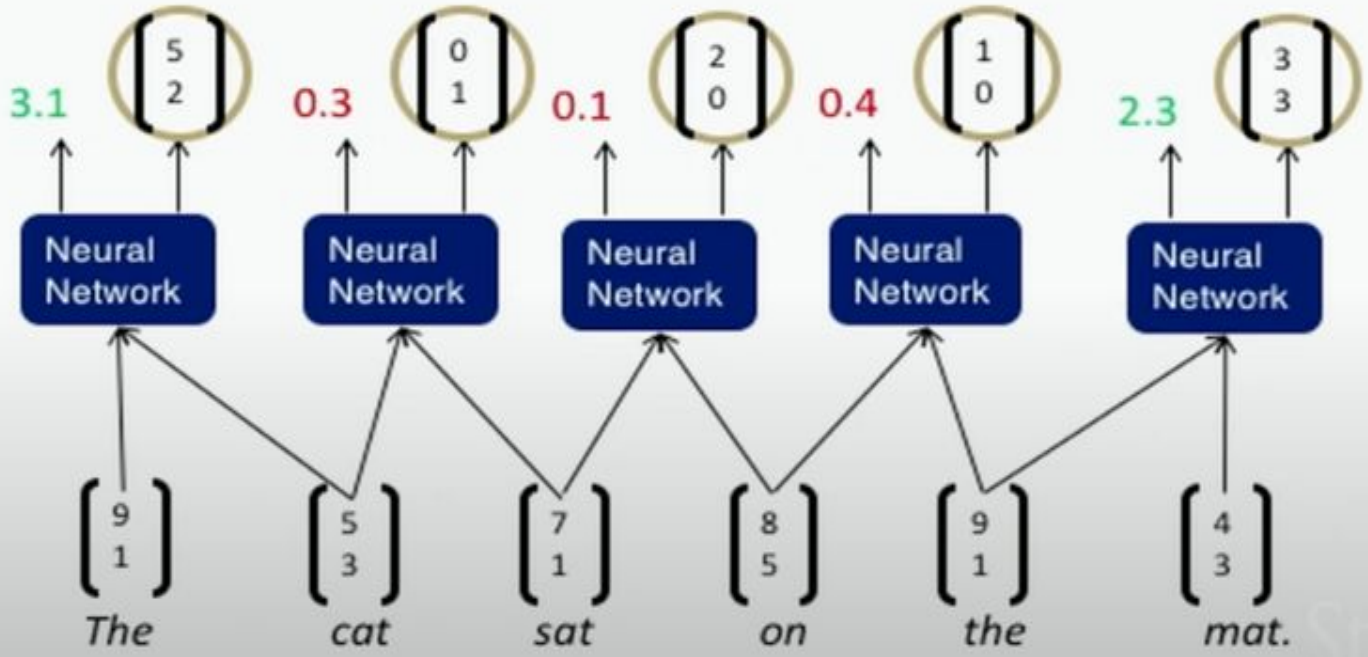


Neural Constituency Parsing

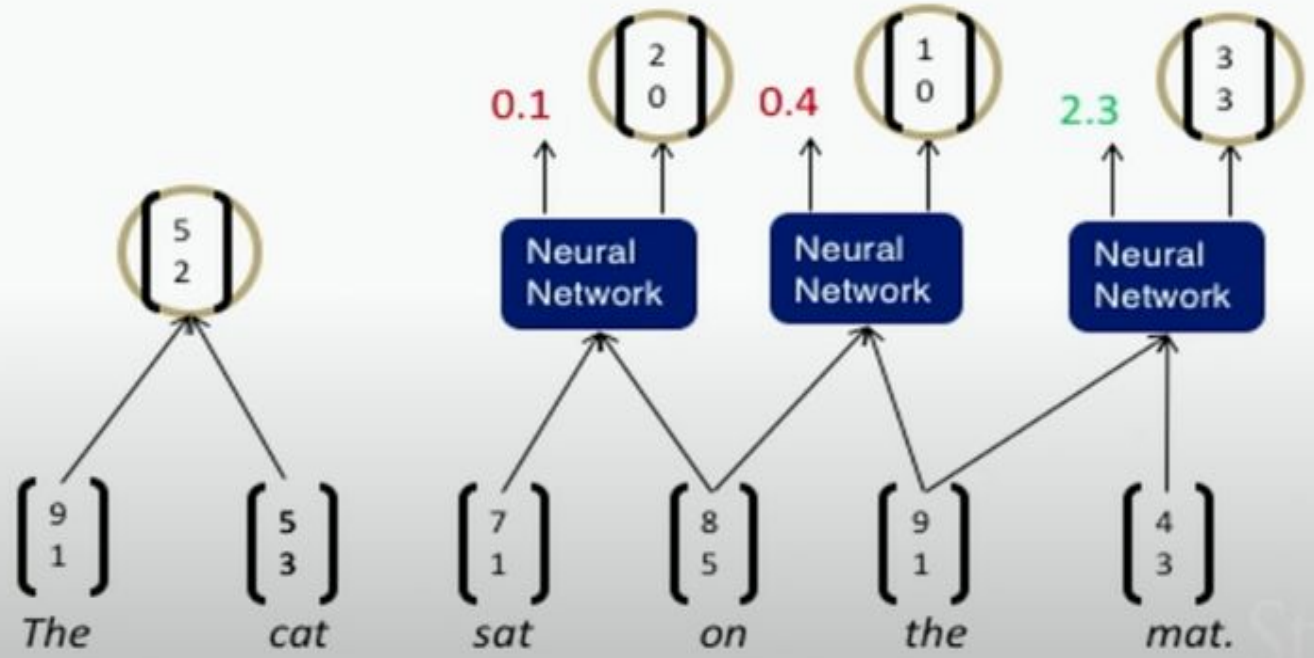
Used for disambiguation



Neural Constituency Parsing

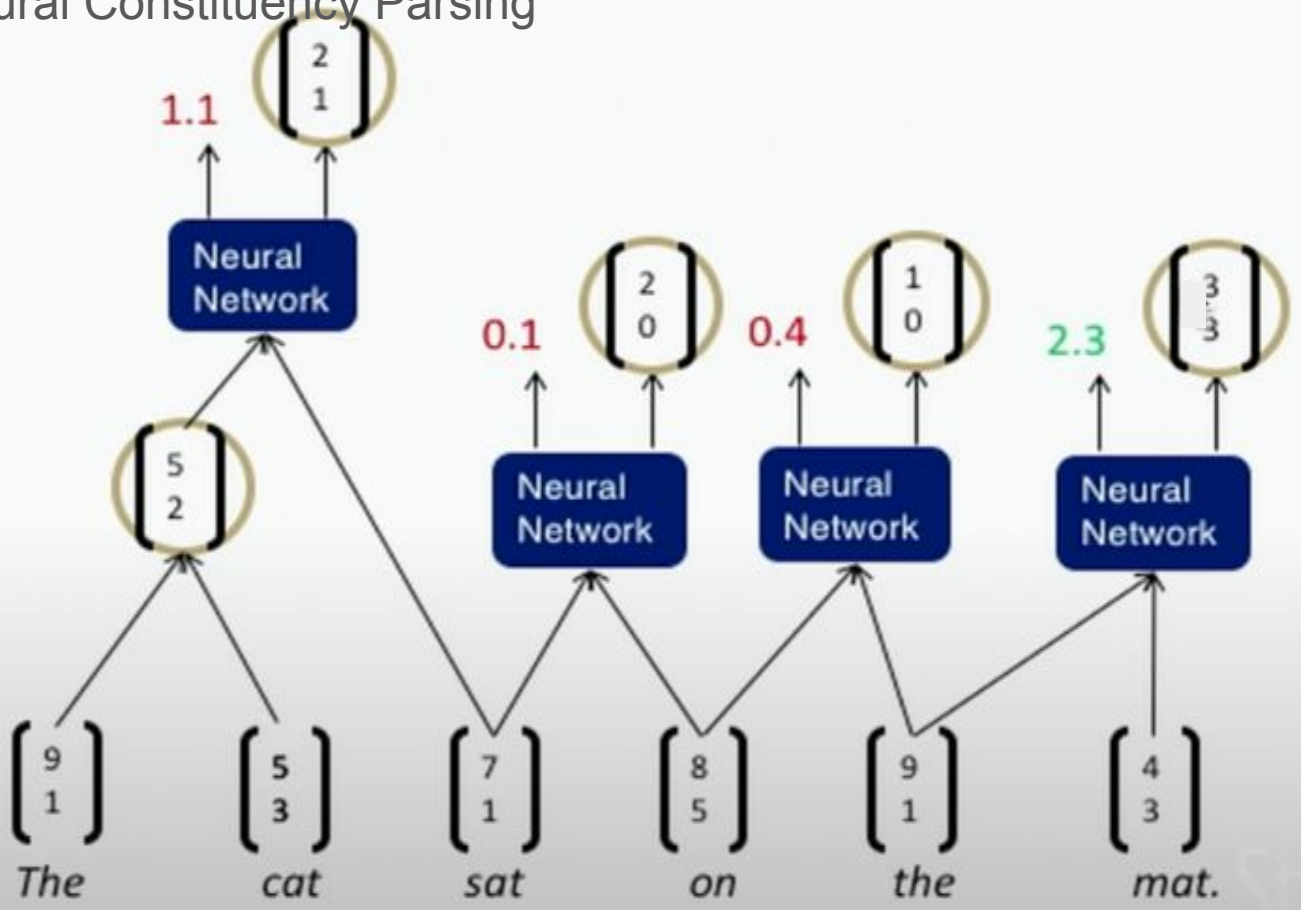


Neural Constituency Parsing

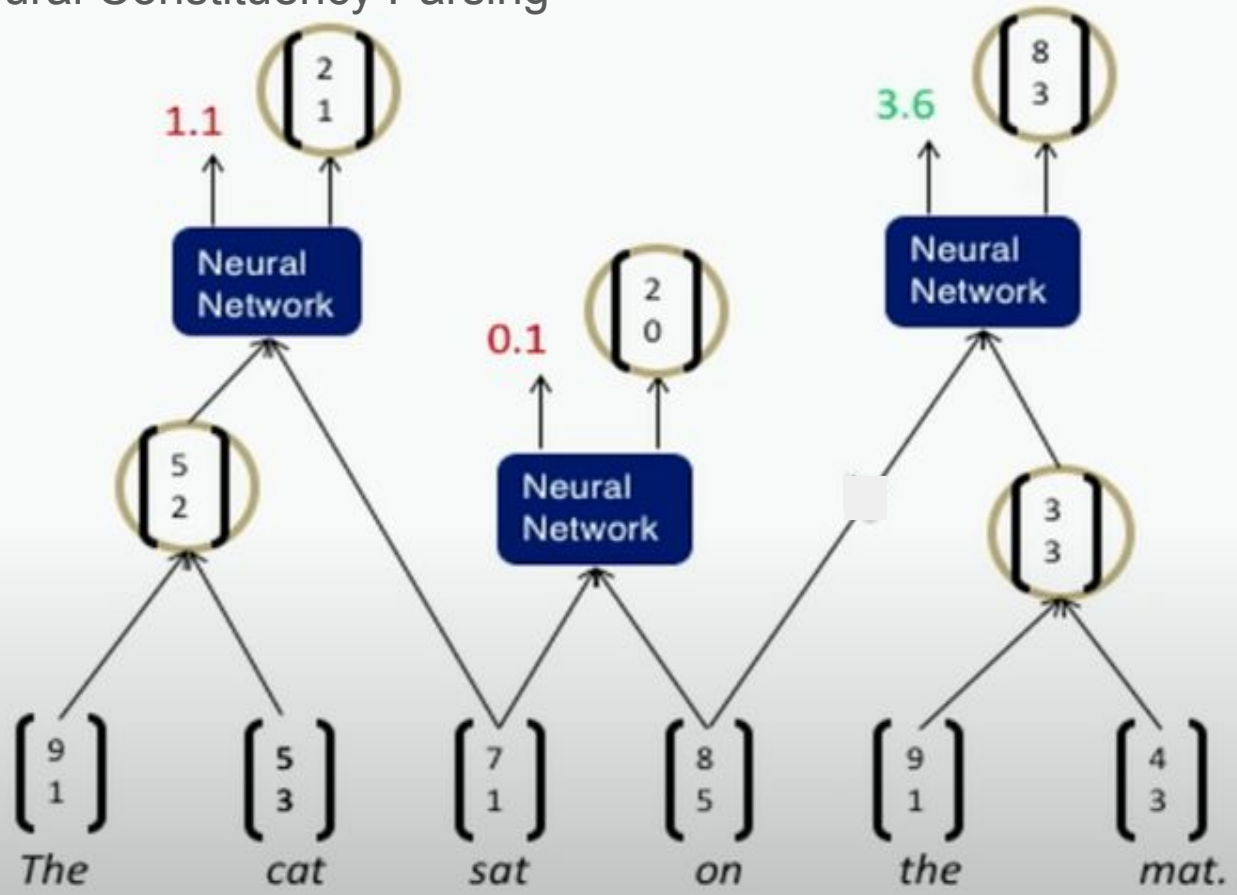


Which one?

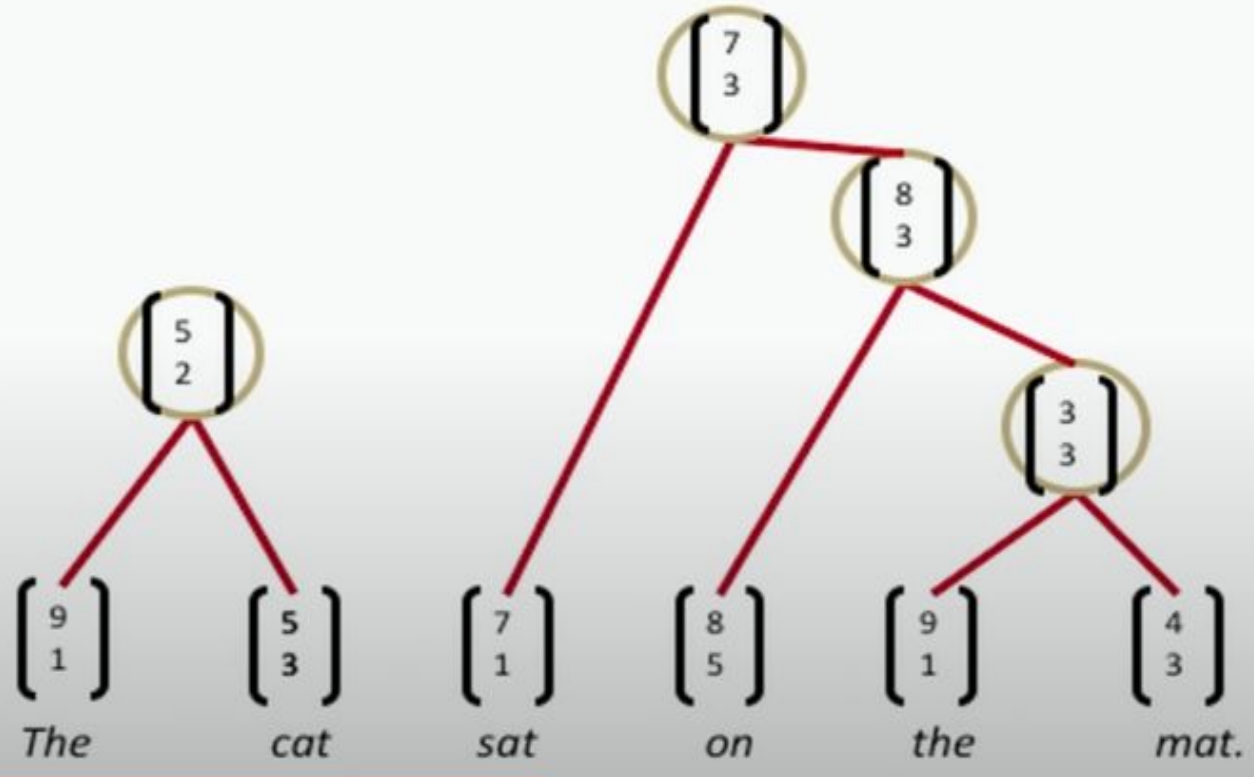
Neural Constituency Parsing



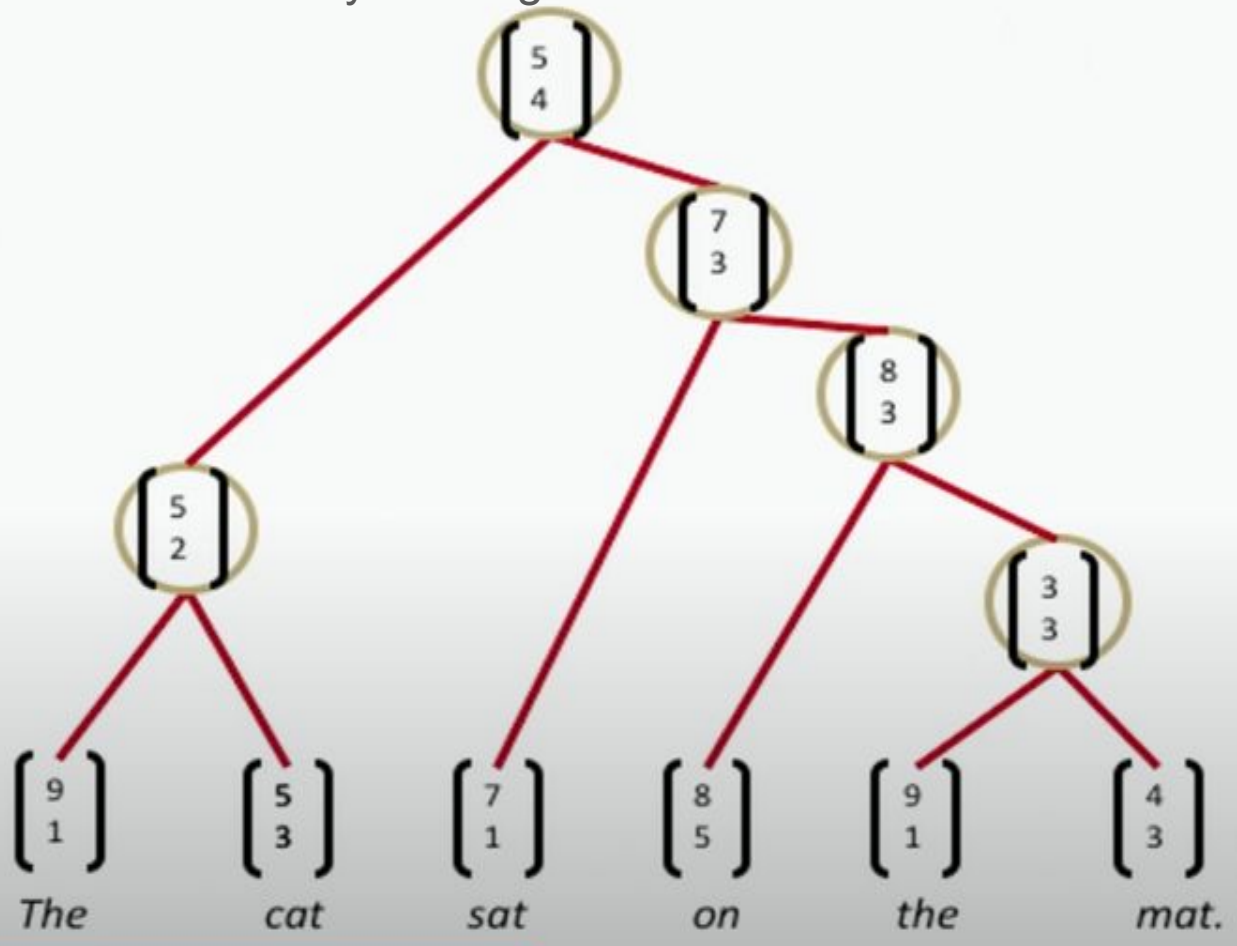
Neural Constituency Parsing



Neural Constituency Parsing



Neural Constituency Parsing



Neural Constituency Parsing

Greedy algorithm

95% times generates valid trees

Evaluating Constituency Parsers

Precision, Recall, F1

Acknowledgments

Dr. Munindar Singh

Dr. Christopher Manning

“Speech and Language Processing” by Dan Jurafsky and James H. Martin