

Natural Language Processing

Lecture 12: Syntax

Noun Phrases

- the elephant arrived
- it arrived
- elephants arrived
- the big ugly elephant arrived
- the elephant I love to hate arrived

Prepositional Phrases

- on Tuesday
- in March
- under the leaking roof

Sentences (a.k.a. Clauses)

- John loves Mary
- John loves the woman he thinks is Mary
- sometimes John thinks he is Mary
- it is patently false that
- sometimes John thinks he is Mary

Context-Free Grammars

- Vocabulary of terminal symbols, Σ
- Set of nonterminal symbols (a.k.a. variables), N
- Special start symbol $S \in N$
- Production rules of the form $X \rightarrow \alpha$

where

$$X \in N$$

$$\alpha \in (N \cup \Sigma)^*$$

λ . $S \rightarrow NP VP$

λ . $NP \rightarrow Det Noun$

λ . $VP \rightarrow Verb NP$

λ . $Det \rightarrow the, a$

λ . $Noun \rightarrow boy, girl, hotdogs$

λ . $Verb \rightarrow likes, hates, eats$

Building Noun Phrases

NP \rightarrow Determiner NounBar

NP \rightarrow ProperNoun

NounBar \rightarrow Noun

NounBar \rightarrow AP NounBar

NounBar \rightarrow NounBar PP

AP \rightarrow Adj AP

AP \rightarrow Adj

PP \rightarrow Preposition NP

Terminology

Grammatical: said of a sentence in the language

Ungrammatical: said of a sentence **not** in the language

Derivation: sequence of top-down production steps

Parse tree: graphical representation of the derivation

A string is grammatical iff there exists a derivation for it.

Ambiguity

λ $S \rightarrow NP VP$

λ $NP \rightarrow Det Noun$

λ $VP \rightarrow Verb NP$

λ $VP \rightarrow VP PP$

λ $PP \rightarrow Prep NP$

λ $Det \rightarrow the, a$

λ $Noun \rightarrow boy, girl, hotdogs, park$

λ $Verb \rightarrow likes, hates, eats, sees$

λ $Prep \rightarrow in, with$

Grammatically (varies sometimes)

λ I'll write the company

λ I'll write to the company

λ It needs to be washed

λ It needs washed

λ They met Friday to discuss it

λ The met on Friday to discuss it

Why are we building grammars?

Consider:

- .Oswald shot Kennedy
- .Oswald, who had visited Russia recently, shot Kennedy
- .Oswald assassinated Kennedy

- .Who shot Kennedy?

Why are we building grammars?

Consider:

- .Oswald shot Kennedy
- .Kennedy was shot by Oswald
- .Oswald was shot by Ruby
- .Who shot Oswald?

Why are we building grammars?

λ Active/Passive

- Oswald shot Kennedy

- Kennedy was shot by Oswald

λ Relative clauses

- Oswald who shot Kennedy was shot by Ruby

- Kennedy who Oswald shot didn't shoot anybody

Subject and Object

- Syntactic (not semantic)

- .The batter hit the ball [subject is semantic agent]

- .The ball was hit by the batter [subject is semantic *patient*]

- .The ball was given a whack by the batter
[subject is semantic *recipient*]

- {George, the key, the wind} opened the door

- Subject ≠ topic

- .I just married the most beautiful woman in the world

- .Now beans, I like

- .As for democracy, I think it's the best form of

credit: Lori Levin, Archana Bhatia

Subject and Object

- English subjects
 - .agree with the verb
 - .when pronouns, in nominative case
 - .(I/she/he vs. me/her/him)
 - .omitted from infinitive clauses
 - .(I tried _ to read the book, I hoped _ to be chosen)
- English objects
 - .when pronouns, in accusative case
 - .become subjects in passive sentences