Syntax intro

CS 585, Fall 2015

Introduction to Natural Language Processing

http://people.cs.umass.edu/~brenocon/inlp2015/

Brendan O'Connor

College of Information and Computer Sciences

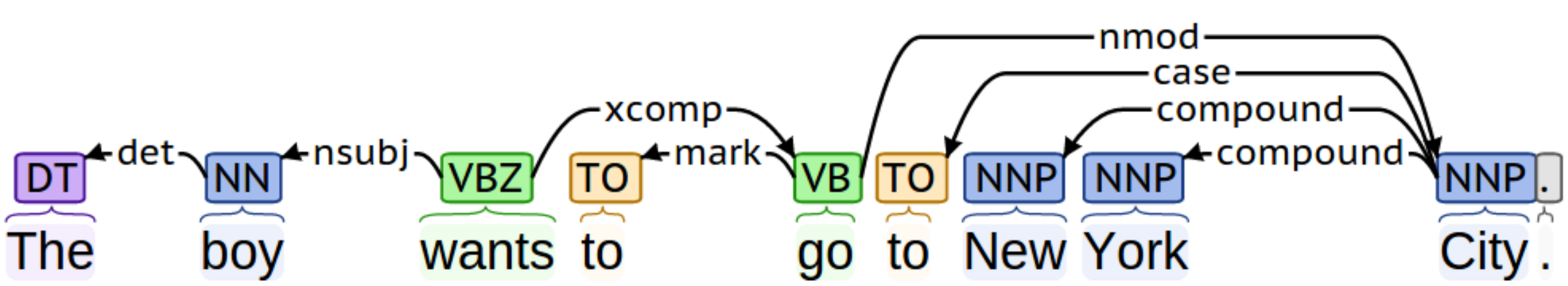
University of Massachusetts Amherst

[some slides borrowed from Percy Liang]

- What linguistic representations do we want?

Dependency parse trees?

The boy wants to go to New York City.



Frames?

Cynthia	*sold*	*the bike*	*to*	*Bob*	*for*	*$200*
SELLER	**PREDICATE**	GOODS		BUYER		PRICE

Logical forms?

What is the largest city in California?

$\texttt{argmax}(\lambda x.\texttt{city}(x) \wedge \texttt{loc}(x, \texttt{CA}), \lambda x.\texttt{population}(x))$

Levels of linguistic analyses

natural language utterance

Levels of linguistic analyses

Syntax: what is grammatical?

natural language utterance

Levels of linguistic analyses

Semantics: what does it mean?

Syntax: what is grammatical?

natural language utterance

Levels of linguistic analyses

Pragmatics: what does it do?

Semantics: what does it mean?

Syntax: what is grammatical?

natural language utterance

Analogy with programming languages

Syntax: no compiler errors

Semantics: no implementation bugs

Pragmatics: implemented the right algorithm

Analogy with programming languages

Syntax: no compiler errors

Semantics: no implementation bugs

Pragmatics: implemented the right algorithm

Different syntax, same semantics (5):

$$2 + 3 \quad \Leftrightarrow \quad 3 + 2$$

Analogy with programming languages

Syntax: no compiler errors

Semantics: no implementation bugs

Pragmatics: implemented the right algorithm

Different syntax, same semantics (5):

$$2 + 3 \Leftrightarrow 3 + 2$$

Same syntax, different semantics (1 and 1.5):

$$3 \ / \ 2 \text{ (Python 2.7)} \not\Leftrightarrow 3 \ / \ 2 \text{ (Python 3)}$$

Analogy with programming languages

Syntax: no compiler errors

Semantics: no implementation bugs

Pragmatics: implemented the right algorithm

Different syntax, same semantics (5):

$$2 + 3 \quad \Leftrightarrow \quad 3 + 2$$

Same syntax, different semantics (1 and 1.5):

$$3 \;/\; 2 \text{ (Python 2.7)} \quad \not\Leftrightarrow \quad 3 \;/\; 2 \text{ (Python 3)}$$

Good semantics, bad pragmatics:

correct implementation of deep neural network
for estimating coin flip prob.

Constituency

- Group of words that behaves like a single unit
- e.g. "Noun Phrase" could be any of

Harry the Horse
the Broadway coppers
they

a high-class spot such as Mindy's
the reason he comes into the Hot Box
three parties from Brooklyn

- Commonalities in behavior (allowable left/right contexts?)

Context Free Grammars

- A formal grammar
  - 1. defines ("generates") a set of strings
  - 2. can be used to analyze a string: is it in the set or not?
- Theoretical goal: What is the grammar for all of English?
- CFG for noun phrases

$$
\begin{aligned}
NP &\rightarrow Det\ Nominal \\
NP &\rightarrow ProperNoun \\
Nominal &\rightarrow Noun \mid Nominal\ Noun
\end{aligned}
$$

$$
\begin{aligned}
Det &\rightarrow a \\
Det &\rightarrow the \\
Noun &\rightarrow flight
\end{aligned}
$$

- Was J&K a CFG? How to add adjectives?

Sentence CFG

Rules

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

Figure 13.1 The $\mathscr{L}_1$ miniature English grammar and lexicon.

Example: *Book that flight*