

Context Free Grammars

October 2008

1

Syntactic Grammaticality

Doesn't depend on

- Having heard the sentence before
- The sentence being true
 - I saw a unicorn yesterday
- The sentence being meaningful
 - Colorless green ideas sleep furiously
 - *Furiously sleep ideas green colorless
 - I sperred a couple of gurdy fipps.

Grammaticality is a formal property that we can investigate and describe

2

Syntax

By syntax, we mean various aspects of how words are strung together to form components of sentences and how those components are strung together to form sentences

- New Concept: Constituency
- Groups of words may behave as a single unit or constituent
- E.g., noun phrases
- Evidence
 - Whole group appears in similar syntactic environment
 - E.g., before a verb
 - Preposed/postposed constructions
 - Note: notions of meaning play no role in syntax (sort-of)

3

What is Syntax?

- Study of structure of language
- Specifically, goal is to relate surface form (e.g., interface to phonological component) to semantics (e.g., interface to semantic component)
- Morphology, phonology, semantics farmed out (mainly), issue is word order and structure
- Representational device is **tree structure**

4

Types of Linguistic Activity

- **Descriptive**: provide account of syntax of a language; often good enough for NLP engineering work
- **Explanatory**: provide principles-and-parameters style account of syntax of (preferably) several languages
- **Prescriptive**: "prescriptive linguistics" is an oxymoron

5

Structure in Strings

- Some words: *the a small nice big very boy girl sees likes*
- Some good sentences:
 - the boy likes a girl
 - the small girl likes the big girl
 - a very small nice boy sees a very nice boy
- Some bad sentences:
 - *the boy the girl
 - *small boy likes nice girl
- Can we find subsequences of words (**constituents**) which in some way behave alike?

6

Structure in Strings Proposal

- Some words: *the a small nice big very boy girl sees likes*
 - Some good sentences:
 - (the boy) likes (a girl)
 - (the small girl) likes (the big girl)
 - (a very small nice boy) sees (a very nice boy)
 - Some bad sentences:
 - *(the boy) (the girl)
 - *(small boy) likes (the nice girl)
- This is a reasonable proposal:
seems to capture some generality

7

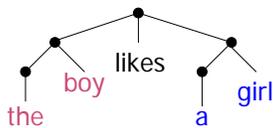
More Structure in Strings Proposal -- ctd

- Some words: *the a small nice big very boy girl sees likes*
- Some good sentences:
 - ((the) boy) likes ((a) girl)
 - ((the) (small) girl) likes ((the) (big) girl)
 - ((a) ((very) small) (nice) boy) sees ((a) ((very) nice) girl)
- Some bad sentences:
 - *((the) boy) ((the) girl)
 - *((small) boy) likes ((the) (nice) girl)

8

From Substrings to Trees

- ((the) boy) likes ((a) girl)



9

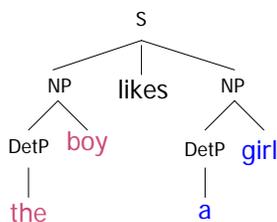
Node Labels?

- ((the) boy) likes ((a) girl)
- Group words by distribution of constituents they head (part-of-speech, POS):
 - Noun (N), verb (V), adjective (Adj), adverb (Adv), determiner (Det)
- Category of constituent: XP, where X is POS
 - NP, S, AdjP, AdvP, DetP

10

Node Labels

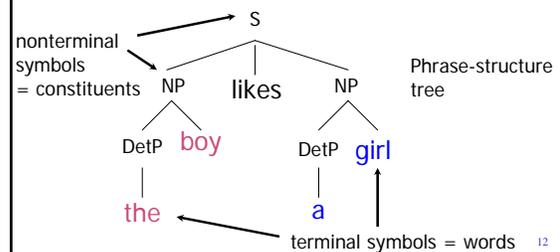
- ((the/det) boy/N) likes/v ((a/det) girl/N)



11

Types of Nodes

- ((the/det) boy/N) likes/v ((a/det) girl/N)



12

Context-Free Grammars

- Defined in formal language theory (comp sci)
- Terminals, nonterminals, start symbol, rules
- String-rewriting system
- Start with start symbol, rewrite using rules, done when only terminals left
- NOT A LINGUISTIC THEORY, just a formal device

13

CFG: Example

- Many possible CFGs for English, here is an example (fragment):
 - $S \rightarrow NP VP$
 - $VP \rightarrow V NP$
 - $NP \rightarrow DetP N \mid DetP AdjP N$
 - $AdjP \rightarrow Adj \mid Adv AdjP$
 - $N \rightarrow boy \mid girl$
 - $V \rightarrow sees \mid likes$
 - $Adj \rightarrow big \mid small$
 - $Adv \rightarrow very$
 - $DetP \rightarrow a \mid the$

the very small boy likes a girl

14

Derivations in a CFG

S

S

$S \rightarrow NP VP$
 $VP \rightarrow V NP$
 $NP \rightarrow DetP N \mid DetP AdjP N$
 $AdjP \rightarrow Adj \mid Adv AdjP$
 $N \rightarrow boy \mid girl$
 $V \rightarrow sees \mid likes$
 $Adj \rightarrow big \mid small$
 $Adv \rightarrow very$
 $DetP \rightarrow a \mid the$

15

Derivations in a CFG

NP VP

S

$S \rightarrow NP VP$
 $VP \rightarrow V NP$
 $NP \rightarrow DetP N \mid DetP AdjP N$
 $AdjP \rightarrow Adj \mid Adv AdjP$
 $N \rightarrow boy \mid girl$
 $V \rightarrow sees \mid likes$
 $Adj \rightarrow big \mid small$
 $Adv \rightarrow very$
 $DetP \rightarrow a \mid the$

16

Derivations in a CFG

DetP N VP

S

$S \rightarrow NP VP$
 $VP \rightarrow V NP$
 $NP \rightarrow DetP N \mid DetP AdjP N$
 $AdjP \rightarrow Adj \mid Adv AdjP$
 $N \rightarrow boy \mid girl$
 $V \rightarrow sees \mid likes$
 $Adj \rightarrow big \mid small$
 $Adv \rightarrow very$
 $DetP \rightarrow a \mid the$

17

Derivations in a CFG

the boy VP

S

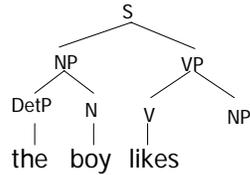
$S \rightarrow NP VP$
 $VP \rightarrow V NP$
 $NP \rightarrow DetP N \mid DetP AdjP N$
 $AdjP \rightarrow Adj \mid Adv AdjP$
 $N \rightarrow boy \mid girl$
 $V \rightarrow sees \mid likes$
 $Adj \rightarrow big \mid small$
 $Adv \rightarrow very$
 $DetP \rightarrow a \mid the$

18

Derivations in a CFG

the boy likes NP

S → NP VP
 VP → V NP
 NP → DetP N | DetP AdjP N
 AdjP → Adj | Adv AdjP
 N → boy | girl
 V → sees | likes
 Adj → big | small
 Adv → very
 DetP → a | the

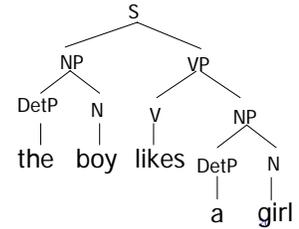


19

Derivations in a CFG

the boy likes a girl

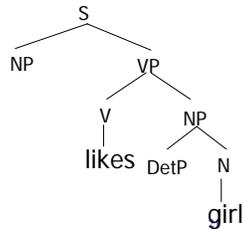
S → NP VP
 VP → V NP
 NP → DetP N | DetP AdjP N
 AdjP → Adj | Adv AdjP
 N → boy | girl
 V → sees | likes
 Adj → big | small
 Adv → very
 DetP → a | the



Derivations in a CFG; Order of Derivation Irrelevant

NP likes DetP girl

S → NP VP
 VP → V NP
 NP → DetP N | DetP AdjP N
 AdjP → Adj | Adv AdjP
 N → boy | girl
 V → sees | likes
 Adj → big | small
 Adv → very
 DetP → a | the



Key Constituents (English)

- Sentences
- Noun phrases
- Verb phrases
- Prepositional phrases

22

Recursion

- We'll have to deal with rules such as the following where the non-terminal on the left also appears somewhere on the right (directly).

NP → NP PP [[The flight] [to Boston]]
 VP → VP PP [[departed Miami] [at noon]]

23

Recursion

- Can make things interesting. Consider the rule:
- NP → NP PP
 flights from Denver
 flights from Denver to Miami
 flights from Denver to Miami in February
 flights from Denver to Miami in February on a Friday
 flights from Denver to Miami in February on a Friday under \$300
 flights from Denver to Miami in February on a Friday under \$300 with lunch

24

Recursion

[[flights] [from Denver]]
[[[[flights] [from Denver]] [to Miami]]
[[[[[flights] [from Denver]] [to Miami]] [in February]]
[[[[[[flights] [from Denver]] [to Miami]] [in February]] [on a Friday]]
Etc.

25

The Point

- If you have a rule like
 - VP → V NP
 - It only cares that the thing after the verb is an NP. It doesn't have to know about the internal affairs of that NP

26

The Point

- VP → V NP
- I hate
 - flights from Denver
 - flights from Denver to Miami
 - flights from Denver to Miami in February
 - flights from Denver to Miami in February on a Friday
 - flights from Denver to Miami in February on a Friday under \$300
 - flights from Denver to Miami in February on a Friday under \$300 with lunch

27