NLP

Introduction to NLP

Syntax

Syntax

- Is language more than just a “bag of words”?
- Grammatical rules apply to categories and groups of words, not individual words.
- Example – a sentence includes a subject and a predicate. The subject is a noun phrase and the predicate is a verb phrase.
  - Noun phrase: The cat, Samantha, She
  - Verb phrase: arrived, went away, had dinner
- When people learn a new word, they learn its syntactic usage.
  - Examples: wug (n), cluvious (adj) – use them in sentences
  - Hard to come up with made up words: forkle, vleer, etc. all taken.

Defining Parts of Speech

- What do nouns typically have in common?
  - E.g., *can* be preceded by "the".
- Verbs can be preceded by "can't".
- Adjectives can come between "the" and a noun.
- How is this different from grade school definitions?
- Determiners: a, the, many, no, five
- Prepositions: for, to, in, without, before

The Lexicon

- How do we think of words like cat, run, five?
  - pronunciation, part of speech, meaning
- Five: /faɪv/, numeral, “5”
- Ambiguity

Constituents

- Constituents are continuous
- Constituents are non-crossing
  - if two constituents share one word, then one of them must completely contain the other.
- Each word is a constituent

Constituent Tests

- "coordination" test
- "pronoun" test
  - A small dog is barking in the park.
  - It is barking in the park
- "question by repetition" test:
  - I have seen blue elephants
  - Blue elephants?
  - * Seen blue?
  - Seen blue elephants?
- "topicalization" test:
  - Blue elephants, I have seen.
- "question" test:
  - *What* have I seen?
- "deletion" test
  - Last year I saw a blue elephant in the zoo.
- "semantic" test
- "intuitition" test

How To Generate Sentences

- One way: tree structure
  - Generate the tree structure first
  - Then fill the leaf nodes with terminals

A Simple Syntactic Rule

- The simplest rule for a sentence, e.g. “Birds fly”

$$S \rightarrow N\ V$$

Simplest Grammar

S → N V

N → Samantha | Min | Jorge

V → left | sang | walked

Sample sentences:

Samantha sang

Jorge left

Syntax

- The verbs so far were intransitive (no direct object)
- What rules are needed next?
  - Transitive verbs and direct objects (“Jorge saw Samantha”)
  - Determiners (“the cats”)
- Combinatorial explosion (even for the simplest form of sentences)
- Need for noun phrases
- Ditto for verb phrases

Latest Grammar

S → NP VP

NP → DT N

VP → V NP

DT → the | a

N → child | cat | dog

V → took | saw | liked | scared | chased

Sample sentences:

a dog chased the cat

the child saw a dog

Alternatives

- Different expansions of a category are delineated with "|"
  - `NP → PN | DT CN`
- One rule for proper nouns and another for common nouns

Latest Grammar

```
S → NP VP
NP → DT CN
NP → PN
VP → V NP
DT → the | a
CN → child | cat | dog
PN → Samantha | Jorge | Min
V → took | saw | liked | scared | chased
```

Sample sentences:

a child scared Jorge

Min took the child

Optional Categories

- Wherever `N` is allowed in a sentence,
  - `DT N`
  - `JJ N`
  - `DT JJ N`

  are also allowed
- We can use the notation for alternatives
  - `NP → N | DT N | JJ N | DT JJ N`
- Optional categories can be also marked using parentheses:
  - `NP → (DT) (JJ) N`

Verb Phrases

- Samantha ran.
- Samantha ran to the park.
- Samantha ran away.
- Samantha bought a cookie.
- Samantha bought a cookie for John.
- Overall structure: `VP → V(NP)(P)(NP)`

Latest Grammar

S → NP VP
NP → DT CN
NP → PN
VP → V (NP) (P) (NP)
DT → the | a
CN → child | cat | dog
PN → Samantha | Jorge | Min
P → to | for | from | in
V → took | saw | liked | scared | chased | gave

Sample sentences:

Samantha saw the cat
Jorge gave the cat to Min

Prepositional Phrases

- Examples:
  - Mary bought a book for John **in a bookstore**.
  - The bookstore sells magazines.
  - The bookstore **on Main St.** sells magazines.
  - Mary ran away.
  - Mary ran **down the hill**.
- Changes are needed to both NP and VP to accommodate prepositional phrases
  - Wherever a preposition is allowed, it can be followed by a noun phrase.
  - Run up
  - NP can contain any number of PPs but only up to two NPs.
- How do we revise the grammar accordingly?

The Rules So Far

- S → NP VP
- NP → (DT) (JJ) N (PP)
- VP → V (NP) (PP)
- PP → P (NP)

PP Ambiguity

- The boy saw the woman with the telescope.

```
PP → PREP NP
VP → V NP PP
VP → V NP
NP → DT N
NP → DT N PP
```

Repetition (*)

- (JJ*) = a sequence of zero or more JJ
- Are all sequences of adjectives allowed?
  - a big red house
  - * a red big house
- Adjective ordering in English depends on semantics!

Exercise

- The Little Red Riding Hood
- Three Little Pigs
- The Three Musketeers
- The Steadfast Tin Soldier
- The French Connection
- Old Macdonald
- Five Golden Rings
- The Ancient Mariner

Adjective Ordering

- **Det**
- Number
- Strength
- Size
- Age
- Shape
- Color
- Origin
- Material
- Purpose
- **Noun**

- det < number < size < color < purpose < noun
- strength < material < noun
- origin < noun

Nested Sentences

- Examples:
  - I don't recall whether I took the dog out.
  - Do you know if the mall is still open?
- **VP → V (NP) (NP) (C S) (PP*)**
- Can (C S) appear inside an NP?
  - Whether he will win the elections remains to be seen.

Recursion

- S can generate VP, VP can generate S
- NP can generate PP, PP can generate NP
- What does recursion allow?
- Is there a longest sentence in English?
- Conjunction of NPs:

```
NP → NP and NP
```

- Conjunction of PPs:

```
PP → PP and PP
```

- Conjunction of VPs:

```
VP → VP and VP
```

Meta-patterns

- S → NP VP
  - NP → (DT) (JJ) N (PP)
  - VP → V (NP) (PP)
  - PP → P (NP)
- Is there a meta-pattern here?
  - XP → (specifier) X'
  - X' → X (complement)
- Example: NP → DT N'
- X-bar Theory
  - http://www.unlweb.net/wiki/X-bar_theory

Meta-rules for Conjunctions

- Conjunction
  - X → X and X
- This kind of rule even covers entire sentences
  - S → S and S

Auxiliaries

- Is “Aux V” a constituent?
  - I have seen blue elephants and will remember them forever.
- Recursion:
  - VP -> Aux VP
  - Raj may have been sleeping.
- Is such recursion unlimited?

Exercise

- Grammar:
  - S → NP VP | CP VP
  - NP → (DT) (JJ*) N (CP) (PP*)
  - VP → V (NP) (NP) (PP*) | V (NP) (CP) (PP*)
  - PP → P NP
  - CP → C S
- What rules are needed to generate these three sentences:
  - 1. The small dog of the neighbors brought me an old tennis ball.
  - 2. That wugs have three eyes is unproven by scientists.
  - 3. I saw the gift that the old man gave me at the meeting.

NLP