Natural Language Processing CSCI 4152/6509 — Lecture 20 Syntax of Natural Languages; CYK Algorithm

Instructors: Vlado Keselj

Time and date: 16:05 – 17:25, 25-Nov-2024

Location: Carleton Tupper Building Theatre C

Previous Lecture

- Bracket representation of a parse tree
- Parsing NL in Prolog using Difference Lists
- Definite Clause Grammar (DCG)
  - Basic DCG example
  - Building a parse tree in DCG
  - Agreement example in DCG
  - Embedded code in DCG
- Probabilistic Context-Free Grammars (PCFG)
- PCFG definition
- PCFG as a probabilistic model
- Typical phrase structure rules in English (started): S

Noun Phrase (NP)

- typically: pronouns, proper nouns, or determiner-nominal construction
- some typical rules

```
NP -> PRP                            e.g.: you
NP -> NNP | NNPS                     e.g.: Halifax
NP -> PDT? DT JJ* NN PP*
NP -> NN NN                          e.g.: computer science
```

- in the last rule, we use regular expression notation to describe a set of different rules
- example: all the various flights from Halifax to Toronto
- determiners and nominals
- modifiers before head noun and after head noun
- postmodifier phrases `NP -> DT JJ* NN RelC`

Relative Clauses

- RelC — relative clause
- clause (sentence-like phrase) following a noun phrase
- example: gerundive relative clause:
  flights arriving after 5pm
- example: infinitive relative clause:
  flights to arrive tomorrow
- example: restrictive relative clause:
  flight that was canceled yesterday

Verb Phrase (VP)

- organizes arguments around the verb
- typical rules

```
VP -> Verb             intransitive verbs;
                       e.g.: disappear
VP -> Verb NP          transitive verbs:
                       e.g.: prefer a morning flight
VP -> Verb NP NP       ditransitive verbs:
                       e.g.: send me an email
VP -> Verb PP*         sentential complements
VP -> Verb NP PP*
VP -> Verb NP NP PP*
```

- sentential complements, e.g.:
  You said these were two flights that were the cheapest.

Prepositional Phrase (PP)

- Preposition (IN) relates a noun phrase to other word or phrase
- Prepositional Phrase (PP) consists of a preposition and the noun phrase which is an object of that preposition
- There is typically only one rule for the prepositional phrase: `PP -> IN NP`
- examples: from Halifax, before tomorrow, in the city
- PP-attachment ambiguity

Adjective Phrase (ADJP)

- less common
- examples:
  - She is *very sure of herself.*
  - . . . the *least expensive* fare . . .

Adverbial Phrase (ADVP)

- Example:
```
(S (NP preliminary findings)
    (VP were reported
        (ADVP (NP a year) ago)))
```
- another example: years ago

About Typical Rules

- Only some typical rules are presented
- For example: We see the cat, and you see a dog.
- The sentence could be described with: `S -> S CC S`
- Relative clauses are labeled in Penn treebank using SBAR ($\bar{S}$) non-terminal; e.g.: `(S (NP (NP Lorillard Inc.)`

```
   ,
   (NP (NP the unit)
       (PP of (NP (ADJP New York-based)
                  Loews Corp.)))
   (SBAR that
         (S (NP *gap*)
            (VP makes (NP Kent cigarettes))))
   ,)
(VP stopped (VP using (NP crocidolite))))
```

Heads and Dependency

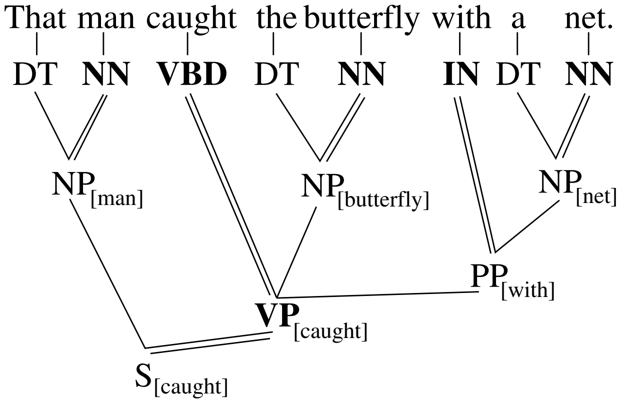
- a phrase typically has a central word called *head*, while other words are direct or indirect *dependents*
- a head is also called a *governor*, although sometimes these concepts are considered somewhat different
- phases are usually called by their head; e.g., the head of a noun phrase is a noun

Example with Heads and Dependencies

That man caught the butterfly with a net.

Example with Heads and Dependencies

- the parse tree of “That man caught the butterfly with a net.”
- annotate dependencies, head words



Head-feature Principle

- Head Feature Principle:
  It is a principle that a set of characteristic features of a head word are transferred to the containing phrase.
- Examples of annotating head in a context-free rule:

$$NP \rightarrow DT\ NN_H$$

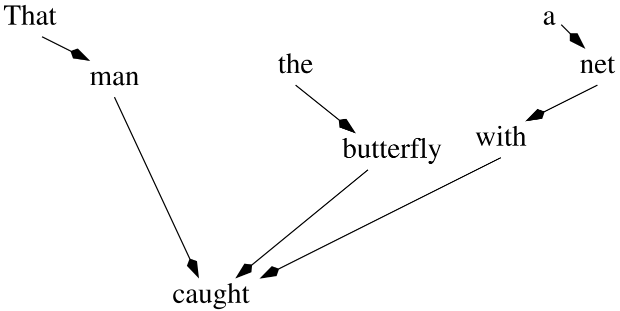
- or

$$[NP] \rightarrow [DT]\ H[NN]$$

- HPSG—Head-driven Phrase Structure Grammars

Dependency Tree

- dependency grammar
- example with “That man caught the butterfly with a net.”



Arguments and Adjuncts

- There ar two kinds of dependents:
  1. **arguments,** which are required dependents, e.g.,
     We deprived him of food.
  2. **adjuncts,** which are not required;
     - ⋆ they have a “less tight” link to the head, and
     - ⋆ can be moved around more easily

     Example:
     We deprived him of food yesterday in the restaurant.

Efficient Inference in PCFG Model

- Using backtracking is not efficient approach
- Chart parsing is an efficient approach
- We will take a look at the CYK chart parsing algorithm

CYK Chart Parsing Algorithm

- When parsing NLP, there are generally two approaches:
  1. Backtracking to find all parse trees
  2. Chart parsing
- CYK algorithm: a simple chart parsing algorithm
- CYK: Cocke-Younger-Kasami algorithm
- CYK can be applied only to a CNF grammar

Chomsky Normal Form

- all rules are in one of the forms:
  1. $A \rightarrow B\,C$, where A, B, and C are nonterminals, or
  2. $A \rightarrow w$, where A is a nonterminal and w is a terminal
- If a grammar is not in CNF, it can be converted to it

Is the following grammar in CNF?

S	→	NP VP	VP	→	V NP	N	→	time	V	→	like
NP	→	N	VP	→	V PP	N	→	arrow	V	→	flies
NP	→	N N	PP	→	P NP	N	→	flies	P	→	like
NP	→	D N				D	→	an			

How about this grammar? (Is it in CNF?)

$$\begin{array}{rcl@{\qquad}rcl@{\qquad}rcl@{\qquad}rcl}
\mathrm{S} & \rightarrow & \mathrm{NP\ VP} & \mathrm{VP} & \rightarrow & \mathrm{V\ NP} & \mathrm{N} & \rightarrow & \mathrm{time} & \mathrm{V} & \rightarrow & \mathrm{like} \\
\mathrm{NP} & \rightarrow & \mathrm{time} & \mathrm{VP} & \rightarrow & \mathrm{V\ PP} & \mathrm{N} & \rightarrow & \mathrm{arrow} & \mathrm{V} & \rightarrow & \mathrm{flies} \\
\mathrm{NP} & \rightarrow & \mathrm{N\ N} & \mathrm{PP} & \rightarrow & \mathrm{P\ NP} & \mathrm{N} & \rightarrow & \mathrm{flies} & \mathrm{P} & \rightarrow & \mathrm{like} \\
\mathrm{NP} & \rightarrow & \mathrm{D\ N} & & & & \mathrm{D} & \rightarrow & \mathrm{an} & & &
\end{array}$$

CYK Example: time flies like an arrow