

NLP

Introduction to NLP

NLP Tasks

Part of Speech Tagging

The swimmer is getting ready to run in the final race.

Part of Speech Tagging

The swimmer is getting ready to **run** in the final race.

- Run – verb or noun?
- Final – noun or adjective?
- Race – verb or noun?

Part of Speech Tagging

The candidate is preparing for his **run** for the presidency.
The swimmer is getting ready to **run** in the final race.

Parsing

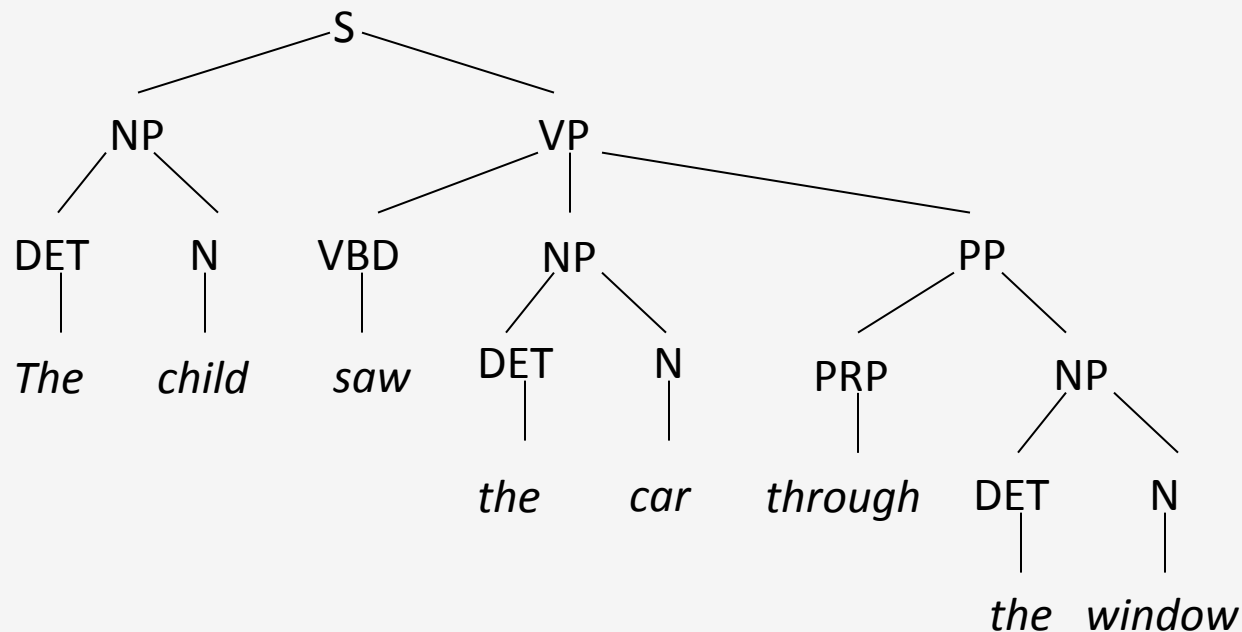
- Myriam slept.
- Myriam wrote a novel.
- Myriam gave Sally flowers.
- Myriam ate pizza with olives.
- Myriam ate pizza with Sally.
- Myriam ate pizza with a fork.
- Myriam ate pizza with remorse.

Phrase-Structure Grammar

S → NP VP
NP → DET N
NP → NP PP
VP → VBD
VP → VBD NP
VP → VBD NP NP
VP → VP PP
PP → PRP NP

DET → *the*
DET → *that*
DET → *a*
N → *child*
N → *window*
N → *car*
VBD → *found*
VBD → *ate*
VBD → *saw*
PRP → *in*
PRP → *of*
PRP → *through*

Parse Trees



Stanford Parser

Stanford Parser - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://nlp.stanford.edu:8080/parser/index.jsp opening line anna karenina

Most Visited Getting Started Latest Headlines

em demo - Go... Build and Pric... calculator.co... CM More states r... CM Parents, don'... Author Central Linux Load Re... NLP Stanfor...

Stanford Parser

Please enter a sentence to be parsed:

Housing starts, the number of new homes being built, rose 7.2% in March to an annual rate of 549,000 units, up from a revised 512,000 in February, the Commerce Department said.

Language: [Sample Sentence](#)

Your query

Housing starts, the number of new homes being built, rose 7.2% in March to an annual rate of 549,000 units, up from a revised 512,000 in February, the Commerce Department said.

Tagging

Housing/NN starts/NNS ,/, the/DT number/NN of/IN new/JJ homes/NNS being/VEG built/VBN ,/, rose/VBD 7.2/CD %/NN in/IN March/NNP to/TO an/DT annual/JJ rate/NN of/IN 549,000/CD units/NNS ,/, up/RB from/IN a/DT revised/VBN 512,000/CD in/IN February/NNP ,/, the/DT Commerce/NNP Department/NNP said/VBD ./.

Parse

```

(ROOT
  (S
    (S
      (NP
        (NP (NN Housing) (NNS starts))
        (, ,)
      )
    )
  )
)

```

Done

Parser Output

```
(ROOT
  (S
    (S
      (NP
        (NP (NN Housing) (NNS starts))
        (, ,)
        (NP
          (NP (DT the) (NN number))
          (PP (IN of)
            (NP
              (NP (JJ new) (NNS homes))
              (VP (VBG being)
                (VP (VBN built))))))
          (, ,))
```

```
(VP (VBD rose)
  (NP (CD 7.2) (NN %))
  (PP (IN in)
    (NP (NNP March)))
  (PP (TO to)
    (NP
      (NP (DT an) (JJ annual) (NN rate))
      (PP (IN of)
        (NP (CD 549,000) (NNS units))))
    (, ,)
    (ADVP (RB up)
      (PP (IN from)
        (NP
          (NP (DT a) (VBN revised) (CD 512,000))
          (PP (IN in)
            (NP (NNP February))))))
        (, ,)
        (NP (DT the) (NNP Commerce) (NNP Department))
        (VP (VBD said))
        (. .)))
```

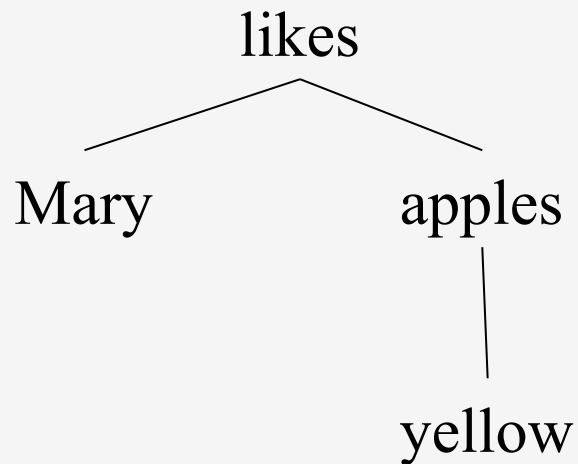
This Problem is Pretty // Easy

- Commercial for a phone company
- Garden path sentences
 - Don't bother coming
 - Don't bother coming early
 - Take the turkey out of the oven at five
 - Take the turkey out of the over at five to four
 - I got canned
 - I got canned peaches for dinner
 - All Americans need to buy a house
 - All Americans need to buy a house is a lot of money
- Can you think of more such examples?

Solution

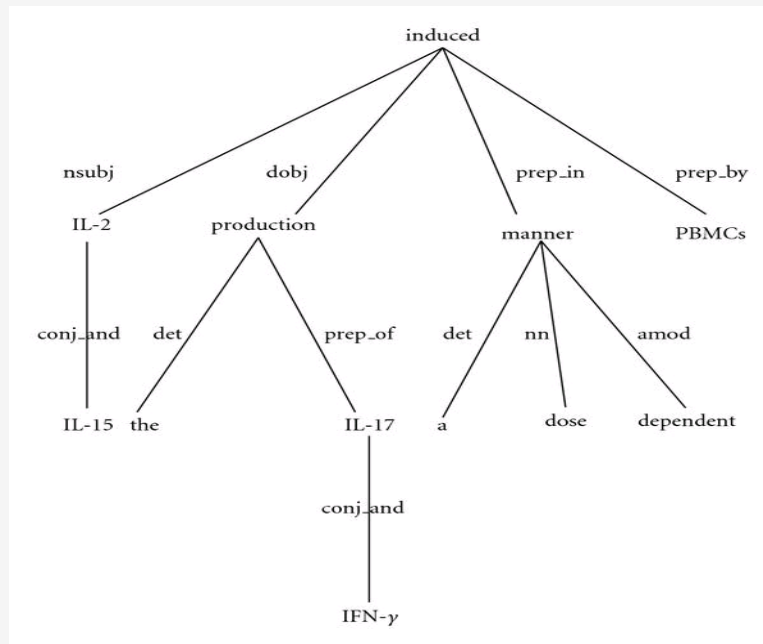
- This problem is pretty // easy
 - <http://www.naclo.cs.cmu.edu/problems2007/N2007-HS.pdf>
- Criteria
 - The part before // should be a complete sentence
 - The full sentence has a different meaning than the part before //
 - The part before // should not already be ambiguous

Dependency Parsing



Dependency Parsing

IL-2 and IL-15 induced the production of IL-17 and IFN- γ by PBMCs in a dose dependent manner.



Parser Output

nn(starts-2, Housing-1)
nsubj(rose-12, starts-2)
det(number-5, the-4)
appos(starts-2, number-5)
prep(number-5, of-6)
amod(homes-8, new-7)
pobj(of-6, homes-8)
auxpass(built-10, being-9)
partmod(homes-8, built-10)
ccomp(said-36, rose-12)
num(%-14, 7.2-13)
dobj(rose-12, %-14)
prep(rose-12, in-15)
pobj(in-15, March-16)
prep(rose-12, to-17)
det(rate-20, an-18)

amod(rate-20, annual-19)
pobj(to-17, rate-20)
prep(rate-20, of-21)
num(units-23, 549,000-22)
pobj(of-21, units-23)
advmod(rose-12, up-25)
dep(up-25, from-26)
det(512,000-29, a-27)
amod(512,000-29, revised-28)
pobj(from-26, 512,000-29)
prep(512,000-29, in-30)
pobj(in-30, February-31)
det(Department-35, the-33)
nn(Department-35, Commerce-34)
nsubj(said-36, Department-35)

NLP