

exp4

October 31, 2021

Neha Kulkarni

A060 BTech. IT SEM 7

Computational Linguistics and Natural Language Processing Lab Experiment 4

```
[2]: import spacy
     from spacy import displacy
```

```
[11]: nlp = spacy.load('en_core_web_sm')
      text = 'She drove the Greek piano'
      doc = nlp(text)

      for token in doc:
          print(token.text, token.dep_, token.head.text, token.pos_, [child for
            ↳child in token.children])

      displacy.render(doc, style = 'dep', jupyter = True, options ={'fine_grained' :
            ↳True})
```

```
She nsubj drove PRON []
drove ROOT drove VERB [She, piano]
the det piano DET []
Greek amod piano ADJ []
piano dobj drove NOUN [the, Greek]
<IPython.core.display.HTML object>
```

```
[12]: text = 'The queen of England drove the Greek piano'
```

```
[13]: doc = nlp(text)

      for token in doc:
          print(token.text, token.dep_, token.head.text, token.pos_, [child for
            ↳child in token.children])

      displacy.render(doc, style = 'dep', jupyter = True, options ={'fine_grained' :
            ↳True})
```

```
The det queen DET []
queen nsubj drove NOUN [The, of]
```

```

of prep queen ADP [England]
England pobj of PROPN []
drove ROOT drove VERB [queen, piano]
the det piano DET []
Greek amod piano ADJ []
piano dobj drove NOUN [the, Greek]
<IPython.core.display.HTML object>

```

```

[14]: import nltk
      nltk.download('punkt')
      nltk.download('averaged_perceptron_tagger')
      from nltk import pos_tag, word_tokenize, RegexpParser

      sample_text = "The quick brown fox jumps over the lazy dog"
      tagged = pos_tag(word_tokenize(sample_text))

      chunker = RegexpParser(r""" Chunk0: {((<NN|CD.?|RB>)<CD.?|VBD.?|VBP.?|VBN.?|NN.?
      ↳|RB.?|JJ>*)<NN|CD.??>}" """)

      output = chunker.parse(tagged)
      print("After Extracting\n", output)

```

After Extracting

```

(S
  The/DT
  quick/JJ
  (Chunk0 brown/NN fox/NN)
  jumps/VBZ
  over/IN
  the/DT
  lazy/JJ
  dog/NN)

```

```

[nltk_data] Downloading package punkt to
[nltk_data]   C:\Users\rosha\AppData\Roaming\nltk_data...
[nltk_data]   Package punkt is already up-to-date!
[nltk_data] Downloading package averaged_perceptron_tagger to
[nltk_data]   C:\Users\rosha\AppData\Roaming\nltk_data...
[nltk_data]   Package averaged_perceptron_tagger is already up-to-
[nltk_data]   date!

```

```

[15]: output.draw()

```

```

[2]: import nltk
      text = "The cat sat on the dirty mat"
      tokens = nltk.word_tokenize(text)
      print(tokens)
      tag = nltk.pos_tag(tokens)

```

```
print(tag)
grammar = "NP: {<DT>?<JJ>*<NN>}"
cp = nltk.RegexpParser(grammar)
result = cp.parse(tag)
print(result)
result.draw()
```

```
['The', 'cat', 'sat', 'on', 'the', 'dirty', 'mat']
[('The', 'DT'), ('cat', 'NN'), ('sat', 'VBD'), ('on', 'IN'), ('the', 'DT'),
('dirty', 'NN'), ('mat', 'NN')]
(S (NP The/DT cat/NN) sat/VBD on/IN (NP the/DT dirty/NN) (NP mat/NN))
```

[]: