

# A5: Syntax

---

Author: Nathan Schneider

This is a pen-and-paper assignment to give you practice with syntactic formalisms and parsing algorithms. The breakdown into 60 points is shown. (Your score will be scaled so that it is worth the same amount as other homework assignments.)

Warning: Allow plenty of time for the last 3 parts!

Questions will concern the sentence:

What will you purchase me for my birthday in July?

## Syntactic Description

---

1. [1 pt] What is the *sentence type*: declarative, imperative, interrogative, or interjection?
2. [2 pts] Identify two words in the sentence that are ambiguous with respect to part-of-speech (POS). Explain.
3. [2 pts] Assuming the contextually correct POS tags, identify a *structural ambiguity* in the sentence.
4. [5 pts] Assign the contextually correct POS to words (ignoring punctuation) from the following list:
  - ADJ : adjective
  - ADV : adverb
  - AUX : auxiliary
  - CONJ : conjunction
  - DET : determiner
  - NOUN : common noun
  - NUM : number
  - PREP : preposition
  - PROPN : proper noun
  - PRON:INDEF : indefinite pronoun
  - PRON:PERS : personal pronoun
  - PRON:POSS : possessive pronoun
  - PRON:WH : wh pronoun
  - VERB : verb
5. [10 pts] Draw two dependency trees: one for each way the structural ambiguity can be disambiguated. Use dependency relations from the following list (ignore punctuation), and use *content heads*:
  - adjmod : adjective modifier
  - advmod : adverb modifier
  - aux : auxiliary
  - det : determiner
  - dobj : direct object

- iobj : indirect object
- nmod : noun modifier
- nummod : number modifier
- pobj : object of preposition
- possmod : possessive modifier
- prep : preposition
- root : root
- subj : subject

6. [10 pts] For each dependency tree, draw a corresponding constituency (phrase structure) tree. Include parts of speech as preterminal categories. Use the following phrasal (ignore punctuation):

- s : sentence
- NP : noun phrase
- VP : verb phrase
- PP : prepositional phrase

Hint 1: Neither the subject nor anything before it should be in a VP.

Hint 2: Apart from the terminals, there should be 4 unary branching structures.

7. [4 pts] Give a small context-free grammar that allows the two constituency trees. You do not have to write out rules for the lexicon entries.

Hint: My solution has 9 rules.

8. [1 pt] Assuming a finite lexicon, does your CFG describe a *finite* or *infinite* string language?
9. [1 pt] Give one example where your CFG *overgenerates*, i.e., allows a sentence that is not really grammatical English.
10. [4 pts] Binarize the grammar so it is in Chomsky-Normal Form (CNF).

Hint: There are multiple valid ways to do this. My solution has 23 binary rules, which contain only nonterminals (phrasal categories and POS tags) on the RHS.

## Parsing Algorithms

11. [10 pts] Use the CNF CFG to fill out the CKY chart for constituency parsing the sentence. [A template](#) is provided. Assume the correct POS tags are already determined. (Ignore punctuation.) Ensure both valid parses are encoded in the chart. There should also be some possible constituents that do not lead to a full parse.
12. [10 pts] For the first dependency parse you drew, write out the steps for the transition-based algorithm to yield the structure. (Ignore the edge labels and punctuation.) Show the stack, buffer, and relations columns, and use the transition operations:
- s : shift
  - LA : left arc

◦ RA : right arc