

## **Part I**

For this assignment, my strategy was to work my way down the list of sentences in the sentences.txt file, and create rules to match up the text, line by line. I referenced the Stanford Parser and the Penn Treebank POS Tag List to assist me in defining new rules and correctly labeling “misc” items.

The first two sentences were already taken care of by the initial set of grammar rules. Both sentences included noun phrases followed by a verb phrase.

**['Arthur', 'is', 'the', 'king', '.']**

```
(START
(S1
  (NP (Proper Arthur))
  (VP (VerbT is) (NP (Det the) (NP (Noun king))))
(Eos .)))
```

For the first sentence, the noun phrase is “Arthur”, which is a proper noun (Proper). The verb phrase “is the king” was parsed with the initial rule: VerbT NP. Here, the third person singular present verb (VerbT) is “is” and the noun phrase is “the king”, (Det NP). The last noun phrase is the noun “king”.

**['Arthur', 'rides', 'the', 'horse', 'near', 'the', 'castle', '.']**

```
(START
(S1
  (NP (Proper Arthur))
  (VP
    (VerbT rides)
    (NP
      (Det the)
      (NP
        (Noun horse)
        (PP (Prep near) (NP (Det the) (NP (Noun castle))))))
    (Eos .)))
```

In the second sentence, “Arthur” was again the first noun phrase (NP -> Proper). The verb phrase was “rides the horse near the castle” using the rule VerbT NP. The VerbT is “rides” and the remaining noun phrase is “the horse near the castle”. Det NP is used next, where the determiner is “the” and the noun phrase is “horse near the castle”. This evokes the prepositional

phrase rule PP -> Prep NP where “near” is the preposition and “the castle” is the final noun phrase. “The” is once again the determiner and “castle” is the noun.

After looking at these two example sentences, I went on to create my own rules for the remaining sentences in the file.

**['Arthur', 'rides', 'the', 'plodding', 'horse', 'near', 'the', 'castle', '.']**

```
(START
(S1
  (NP (Proper Arthur))
  (VP
    (VerbT rides)
    (NP (Det the) (JJ plodding) (Noun horse))
    (PP (Prep near) (NP (Det the) (NP (Noun castle)))))
(Eos .)))
```

To parse this sentence, I added the grammar rule “NP -> Det JJ Noun”. This allowed for noun phrases to include determiners followed by an adjective and then a noun. I also defined adjectives as JJ. In this sentence, “plodding” is the adjective. The NP VP Eos rule now can be applied to this sentence.

**['the', 'Holy\_Grail', 'is', 'a', 'chalice', '.']**

```
(START
(S1
  (NP (Det the) (NP (NNP Holy_Grail)))
  (VP (VerbT is) (NP (Det a) (NP (Noun chalice))))
(Eos .)))
```

For this sentence, I tagged proper nouns that aren’t people (for example, Holy\_Grail) as NNP and created a new noun phrase rule: NP -> NNP. Now, noun phrases were expanded to include one proper noun. The standard S1 rule could now be used on this sentence.

**['the', 'sensational', 'Holy\_Grail', 'is', 'a', 'sacred', 'chalice', '.']**

```
(START
(S1
  (NP (Det the) (JJ sensational) (NNP Holy_Grail))
  (VP (VerbT is) (NP (Det a) (JJ sacred) (Noun chalice)))
(Eos .)))
```

I created a very similar rule to the one in sentence number three. Instead of having the only noun phrase with an adjective as “Det JJ Noun”, I added “Det JJ NNP” to the list. This follows the same structure, but allows for non-person proper nouns after the adjective.

**['every', 'coconut', 'was', 'carried', 'to', 'the', 'hottest', 'mountains', '.']**

```

(START
(S1
  (NP (Det every) (NP (Noun coconut)))
  (VP
    (VBD was)
    (VBD carried)
    (PP (TO to) (NP (Det the) (JJS hottest) (NNS mountains))))
(Eos .)))

```

In a similar fashion to the previous sentence, I added the noun phrase rule Det JJS NNS. This allowed for superlative adjectives, now tagged as JJS, like “hottest” and plural nouns, NNS, such as “mountains”. I also tagged past-tense verbs, including “was” and “carried”, as VBD. I created one new rule each for verb and prepositional phrases. VP -> VBD VBD PP accommodated the verb phrase after the initial NP (“every coconut”). I also added PP -> TO NP, which covered the word “to” through the end of the sentence.

**['sixty', 'strangers', 'are', 'at', 'the', 'Round\_Table', '.']**

```

(START
(S1
  (NP (CD sixty) (NNS strangers))
  (VP
    (VBZ are)
    (PP (Prep at) (NP (Det the) (NP (NNP Round_Table)))))
(Eos .)))

```

For this sentence, I tagged cardinal numbers, like “sixty”, as CD. To capture the word “are”, I used the VBZ tag for non-3rd person singular present verbs. Then, I added a new noun phrase rule, NP -> CD NNS, to cover cardinal numbers followed by plural nouns. Now, the first portion of the sentence “sixty strangers” was able to be parsed. To match the verb phrase, I added the rule VP -> VBZ PP. “Are” is the verb (VBZ) and “at the Round\_Table” is the preposition phrase.

**['Sir\_Lancelot', 'might', 'have', 'spoken', '.']**

```

(START
(S1
  (NP (Proper Sir_Lancelot))
  (VP (MD might) (VB have) (VBN spoken))
(Eos .)))

```

This short sentence required a new few tags and an additional verb phrase. I used MD to refer to modal words such as “might”. VB was used to tag base form verbs, such as “have”. Lastly, VBN covered past participle verbs; here it was “spoken”. The new verb phrase rule combining these three new tags was VP -> MD VB VBN (“might have spoken”). The noun phrase was the proper noun, “Sir\_Lancelot”.

**['Guinevere', 'had', 'been', 'riding', 'with', 'Patsy', 'for', 'five', 'weary', 'nights', '.']**

(START

(S1

(NP (Proper Guinevere))

(VP

(VBD had)

(VBN been)

(VBG riding)

(PP (Prep with) (NP (Proper Patsy)))

(PP (Prep for) (NP (CD five) (JJ weary) (NNS nights))))

(Eos .)))

I made this sentence rather targeted to avoid making recursive rules. I tagged present participle/gerund verbs as VBG. Here, the VBG was “riding”. Next, I added three new phrase rules. NP -> CD JJ NNS was used to parse the noun phrase “five weary nights”. PP -> Prep Proper matched “with Patsy”. The specific rule VP -> VBD VBN VBG PP PP brought it all together, with the other two new rules included inside or as the prepositional phrases.

**['Sir\_Bedeve', 'might', 'have', 'been', 'suggesting', 'this', 'quest', '.']**

(START

(S1

(NP (Proper Sir\_Bedeve))

(VP

(MD might)

(VB have)

(VBN been)

(VBG suggesting)

(NP (Det this) (NP (Noun quest))))

(Eos .)))

This sentence only needed one additional verb phrase rule, but it was also quite specific. VP -> MD VB VBN VBG NP connected “might have been suggesting this quest” with the noun phrase “Sir\_Bedeve”.

**['the', 'Britons', 'migrate', 'south', 'frequently', '.']**

(START

(S1

(NP (Det the) (NNPS Britons))

(VP (VB migrate) (RB south) (RB frequently))

(Eos .)))

Creating two new phrase rules and two new tag rules, I was able to parse through this sentence. RB was used to reference adverbs. Both “south” and “frequently” are adverbs. NNPS was used to tag plural proper nouns, like “Britons”. VP -> VB RB RB allowed for two adjectives

after a verb, “migrate south frequently”. NP -> Det NNPS made it so proper plural nouns could follow a determiner, “the Britons”.

**['Arthur', 'and', 'Guinevere', 'ride', 'frequently', 'near', 'the', 'castle', '.']**

```
(START
(S1
  (NP (Proper Arthur) (CC and) (Proper Guinevere))
  (VP
    (VB ride)
    (RB frequently)
    (PP (Prep near) (NP (Det the) (NP (Noun castle))))))
(Eos .)))
```

Still referencing the Penn Treebank POS tag list, I marked coordinating conjunctions, such as “and”, as CC. I also added the rules NP -> Proper CC Proper (“Arthur and Guinevere”) and VP -> VB RB PP (“ride frequently near the castle”).

**['he', 'suggests', 'to', 'grow', 'fruit', 'at', 'home', '.']**

```
(START
(S1
  (NP (PRP he))
  (VP
    (VBZ suggests)
    (TO to)
    (VB grow)
    (NP (Noun fruit) (PP (Prep at) (NP (Noun home))))))
(Eos .)))
```

To get the grammar to parse this sentence, I added two new tags and two more rules. I renamed personal pronouns as PRP and third person singular present verbs as VBZ. I set noun phrases to include NP -> PRP so they could consist only of personal pronouns. In this case, the noun phrase was “he”. I also added the rule VP -> VBZ TO VB NP. The VBZ was “suggests”, the TO was “to”, and the VB was “grown”. The final NP was using the rule NP -> Noun PP where “fruit” was the noun and “at home” was the prepositional phrase.

**['riding', 'to', 'Camelot', 'is', 'not', 'hard', '.']**

```
(START
(S2
  (VP (VBG riding) (PP (TO to) (NP (NNP Camelot))))
  (VP (VerbT is) (NOT not) (JJ hard))
(Eos .)))
```

This was the first sentence where I added a new type of sentence to the grammar. I added a new option after the start symbol START called S2. Then, I defined S2 as S2 -> VP VP Eos.

Now, two verb phrases followed by an end of sentence tagged symbol could make up a sentence. I also added the phrase rules VP -> VBG PP, PP -> TO NNP, and VerbT NOT JJ. I didn't need to rename any tags to parse this sentence. The first verb phrase, following the first new phrase rule, was "riding to Camelot" where "to Camelot" was the new kind of prepositional phrase. The second verb phrase was "is not hard" where "is" was the VerbT, "not" was the NOT, and "hard" was the adjective.

**['do', 'coconuts', 'speak', '?']**

(START (S3 (DO do) (NP (NNS coconuts)) (VP (VB speak)) (Eos ?)))

For this, I created a third type of sentence structure. Now the start rule read START -> S1 | S2 | S3. The new sentence type was S3 -> DO NP VP Eos. I made it so "do" could precede the noun phrase in a sentence similar to a S1 type. I could have also created a new noun phrase type with DO as the first symbol. I also added the rule VP -> VB so verb phrases could consist of just one base form verb.

**['why', 'does', 'England', 'have', 'a', 'king', '?']**

(START

(S4

(WRB why)

(DO does)

(NP (NNP England))

(VP (VB have) (NP (Det a) (NP (Noun king))))

(Eos ?)))

Another new sentence rule was used to parse this sentence. I added START -> S1 | S2 | S3 | S4 where S4 -> WRB DO NP VP Eos. This tacked on an additional wh-adverb symbol (WRB) at the beginning of the third type of sentence. Of course, I also tagged adverbs like "why" as WRB. I also needed the verb phrase rule VP -> VB NP. This let verb phrases include base verbs followed by noun phrases. So, "why" was the WRB, "does" was the DO, the NNP "England" was the noun phrase, and the verb phrase was "have a king" where "have" was the VB, and "a king" was the noun phrase using the rule NP -> Det NP.

### **Challenging Sentences**

**['do', 'not', 'speak', 'again', '!']**

(START (S5 (DO do) (NOT not) (VP (VB speak) (RB again)) (Eos !)))

I created a fifth and final sentence type to parse the challenging sentence "do not speak again!". The start rule now became START -> S1 | S2 | S3 | S4 | S5. Sentence type five was defined as S5 -> DO NOT VP Eos. I adjusted S3 to allow an additional NOT word, like "not", between the DO word and the verb phrase. Here, the verb phrase was another new rule, VP -> VB RB, where "speak" was the base verb and "again" was the adverb.

**['what', 'horse', 'does', 'Arthur', 'ride', '?']**

(START

(S1

(NP (WDT what) (Noun horse))

(VP (DO does) (Proper Arthur) (VB ride))

(Eos ?)))

The second challenging sentence I added was “What horse does Arthur ride?”. I tagged wh-determiners like “what” as WDT. I added the rule NP -> WDT Noun so a noun phrase could be a wh-determiner and then a noun. I also added the rule VP -> DO Proper VB so a verb phrase could consist of a DO word, a proper noun, and then a base verb. The noun phrase was “what horse” and the verb phrase was “does Arthur ride”.

## **PART II**

After making my way through the list of sentences and additional challenge sentences, I created two sentences of my own to run through the grammar.

**['Sir\_Lancelot', 'will', 'ride', 'with', 'Arthur', 'below', 'the', 'yellow', 'mountains', '.']**

I was not able to parse the sentence above with my current grammar rules, even though it is a logical English sentence. Some of this works with the grammar, like the noun phrase “Sir\_Lancelot”, which uses the rule NP -> Proper. However, one portion of this sentence not represented by my grammar is “the yellow mountains”. Even though I have two rules for noun phrases with determiners and adjectives (NP -> Det JJ [adjective] NNP [proper nouns] and Det JJS [superlative adjective] NNS [plural nouns]), I am missing a rule for adjectives and plural nouns, NP -> Det JJ NNS.

**['one', 'ants', 'migrates', 'to', 'migrate', 'Zoot', 'and', 'Dingo', '.']**

(START

(S1

(NP (CD one) (NNS ants))

(VP

(VBZ migrates)

(TO to)

(VB migrate)

(NP (Proper Zoot) (CC and) (Proper Dingo)))

(Eos .)))

I also added this nonsensical sentence that is actually able to be parsed by the grammar. This works because my grammar doesn’t specify which noun phrases can go with which verb phrases. I have rules which allow for both singular and plural phrases and because the noun and verb phrases are valid and the grammar doesn’t care about plurality, it works. This

sentence uses the grammar rules  $S1 \rightarrow NP VP Eos$ ,  $NP = CD NNS$ ,  $NP \rightarrow Proper CC Proper$ , and  $VP = VBZ TO VB NP$ .