HW3

Rashmi K Shivanna(329589544)

Part 1

I have generated new rules and listed the words in vocabulary with meaningful tags instead of misc and used them in the rules. I have explained how each sentence was parsed using grammar in S1 and the words whose tags were changed from misc to something specific in the rest of this document. Also, I picked 2 sentences from the challenge sentences list and I have explained about them too.

Step 1: I picked up the sentence "Arthur rides the plodding horse near the castle .". All the words were tagged meaningfully in the vocal.gr file except for plodding. Plodding being an adjective, it was listed under Misc. So I created a new tag with name JJ and put plodding under this new tag. Also, I created a new tag EOF for the end of sentence symbols !, . and ?

When I looked at the structure of this sentence, it mainly had 2 parts -
1. A noun phrase which is just a proper noun "Arthur"
2. A verb phrase which is a combination of verb "rides" and noun phrases. The Noun phrases were "the plodding horse near the castle" which was further divided into a noun phrase "the plodding horse" and a prepositional phrase "near the castle"

We had rules for a noun phrase to have a noun phrase and a prepositional phrase, a noun phrase to be just a proper noun and a verb phrase to be just a verb. The only terminal word which wasn't fitting any of these rules was plodding. So, I added a new rule for a noun phrase. A noun phrase could be noun preceded by an adjective. So I added the rule Nbar -> JJ Noun. Now, I had a tree generated for this sentence using the grammar S1:-

```
sentence : Arthur rides the plodding horse near the castle .
[PCFGParser]   best parse tree:
(START
  (S1
    (@NP-VP
      (NP
        (Proper Arthur))
      (VP
        (VerbT rides)
        (NP
          (Det the)
          (Nbar
            (Nbar
              (JJ plodding)
              (Noun horse))
            (PP
              (Prep near)
              (NP
                (Det the)
                (Nbar
                  (Noun castle))))))))
    (EOF .)))
```

Step 2: The next sentence was "the Holy Grail is a chalice ." which had a noun phrase(the holy grail) and a verb phrase(is a chalice). The Noun phrase "the holy grail" had a determinant "the" and a proper noun "holy grail". "holy grail" was named as misc in the vocabulary file. It is a proper noun but not referred to

people. So, I moved it to a new tag PrNN(proper noun). When I looked at the rules in S1, a verb phrase was a combination of singular present verb(is) and a noun phrase which is further a combination of a determinant(a) and a noun(chalice). However, the noun phrase "the holy grail" wasn't getting parsed by any rule since it had a proper noun preceded by a determinant. So I added a new rule under noun phrases as NP->Det Nbar and Nbar-> PrNN. Now, the sentence was parsed using grammar in S1 as below:

```
sentence : the Holy Grail is a chalice .
[PCFGParser]   best parse tree:
(START
  (S1
    (@NP-VP
      (NP
        (Det the)
        (Nbar
          (PrNN Holy) Grail)))
      (VP
        (VerbT is)
        (NP
          (Det a)
          (Nbar
            (Noun chalice)))))
    (EOF .)))
```

Step 3: The next sentence "the sensational Holy Grail is a sacred chalice ." was very similar to the previous sentence except that the proper noun "holy grail" was prefixed with an adjective(sensational) and the noun "chalice" was also prefixed with an adjective. So I added rules NP->Nbar, Nbar-> Noun and Nbar-> JJ Nbar. I moved noun to Nbar so that we can generalize a noun phrase to have just a pronoun or a noun.

The adjectives Sensational and sacred were listed as misc in vocabulary. So I moved them to the adjectives section which was tagged as JJ. Now, the sentence was parsed with grammar in S1 as:

```
sentence : the sensational Holy Grail is a sacred chalice .
[PCFGParser]   best parse tree:
(START
  (S1
    (@NP-VP
      (NP
        (Det the)
        (Nbar
          (JJ sensational)
          (Nbar
            (PrNN Holy) Grail))))
      (VP
        (VerbT is)
        (NP
          (Det a)
          (Nbar
            (JJ sacred)
            (Nbar
              (Noun chalice))))))
    (EOF .)))
```

Step 4: The next sentence "every coconut was carried to the hottest mountains ." had words "was", "carried", "to,""hottest" and "mountains" which were listed as misc in vocabulary. So I created a new tag VerbD(verbs past tense) and moved "was" under it, created tag VerbPP(verb past participle) and moved "carried" under it. I moved "to" to newly created tag TO, "hottest" to a newly created tag JJS(superlative adjectives) and "mountains" under a new tag NNP(plural nouns)

Now the sentence was a noun phrase(every coconut) followed by a verb phrase(was carried to the hottest mountains). I already had the rule NP-> Det Nbar which could parse "every coconut". But the verb phrase "was carried to the hottest mountains" needed a new rule since it started with a past tense verb "was" followed by another verb phrase "carried to the hottest mountains" which is a past participle verb "carried" followed by "to" and a noun phrase "the hottest mountain" which is a determinant "the" followed by superlative adjective "hottest" and a noun "mountains". So, I added the rules VP -> VBD VP, VP-> VerbPP PP, Nbar-> NNP , Nbar-> JJS Nbar, PP-> TO NP which generated the below parse tree for this sentence using S1.

sentence : every coconut was carried to the hottest mountains .
[PCFGParser]   best parse tree:
(START
  (S1
    (@NP-VP
      (NP
        (Det every)
        (Nbar
          (Noun coconut)))
      (VP
        (Vbar
          (VerbD was))
        (VP
          (Vbar
            (VerbPP carried))
          (PP
            (TO to)
            (NP
              (Det the)
              (Nbar
                (JJS hottest)
                (Nbar
                  (NNP mountains)))))))))
    (EOF .)))

Step 5: In the sentence "sixty strangers are at the Round Table ." none of the words were tagged meaningfully in the vocabulary file except "at". So I created a tag Num and moved sixty under it, created VerbTP for a plural present verb and moved "are" under it. I moved "Round Table" to the proper noun(PrNN) section and "strangers" to the NNP(plural nouns) section.

The sentence had noun phrase "Sixty strangers" and a verb phrase "are at the round table". But since I added new tags, I had to add new rules to use these newly added tags. The noun phrases needed a rule NP-> Num Nbar since noun phrases could be a noun phrase prefixed with a number, Nbar-> NNP denoting a noun phrase to be a plural noun. Verb phrase needed rule VP-> VerbTP PP since a verb phrase could start with a plural present verb and a prepositional phrase. After adding all these rules, below tree was generated:

sentence : sixty strangers are at the Round Table .
[PCFGParser]   best parse tree:

```
(START
  (S1
    (@NP-VP
      (NP
        (Num sixty)
        (Nbar
          (NNP strangers)))
      (VP
        (VerbTP are)
        (PP
          (Prep at)
          (NP
            (Det the)
            (Nbar
              (PrNN Round) Table))))))
    (EOF .)))
```

Step 6: In the next sentence "Sir Lancelot might have spoken .", the words "might", "have" and "spoken" were tagged as misc. "might" being a modal, I created a new tag MD and moved it under it. "have" being a verb in base form, I created a new tag VerbB and moved have under it and moved "spoken" under the section VerbPP since it is a verb past participle.

The sentence was divided into a noun phrase "Sir Lancelot" and a verb phrase "might have spoken". Here noun phrases "NP" was just a noun phrase (Nbar). Verb phrase was a base verb(have) prefixed with a modal. Since the verb phrases were too specific with many terminals, I created a new non-terminal symbol Vbar which could be any form of the verb. I added rules VP->MD VP, VP-> Vbar, Vbar-> VerbB, Vbar->VerbPP

(I modified the rules which were initially written for verb phrases to use Vbar instead of direct verbs in the verb phrase. So rules became VP->Vbar NP, VP->Vbar VP, VP->Vbar PP, Vbar->VerbT, Vbar-> VerbTP, Vbar->VerbD)

sentence : Sir Lancelot might have spoken .
[PCFGParser]   best parse tree:
```
(START
  (S1
    (@NP-VP
      (NP
        (Nbar
          (Proper Sir) Lancelot)))
      (VP
        (MD might)
        (VP
          (VerbB have)
          (VP
            (Vbar
              (VerbPP spoken))))))
    (EOF .)))
```

Step 7: The next sentence "Guinevere had been riding with Patsy for five weary nights ." had words "had", "been", "riding", "five", "weary" and "nights" which were listed as misc. Since "had" and "been" are past tense verbs, I added them under the section VerbD in vocabulary file, "five" under Num, "weary" under adjectives JJ and "nights" under noun plural NNP. Riding is a verb present participle. So I created a new tag, VerbPrP and added riding under it and since it is a new tag, I added a new rule Vbar->VerbPrp which says that an inner level verb phrase could just be a present participle verb. Since there are 2

prepositional phrases "with Patsy" and "for five weary nights", followed by a verb, I added rule Vbar->
Vbar PP. This generated the below parse tree using grammar S1:

sentence : Guinevere had been riding with Patsy for five weary nights .
[PCFGParser]   best parse tree:
(START
  (S1
    (@NP-VP
      (NP
        (Nbar
          (Proper Guinevere)))
      (VP
        (Vbar
          (VerbPP had))
        (VP
          (Vbar
            (VerbPP been))
          (VP
            (Vbar
              (Vbar
                (VerbPrP riding))
              (PP
                (Prep with)
                (NP
                  (Nbar
                    (Proper Patsy)))))
            (PP
              (Prep for)
              (NP
                (Num five)
                (Nbar
                  (JJ weary)
                  (Nbar
                    (NNP nights)))))))))
    (EOF .)))

Step 8: The next sentence was "Sir Bedevere might have been suggesting this quest .". This was a noun
phrase "Sir Bedevere" which is a proper noun followed by a verb phrase which is a combination of
multiple verb phrases "might VP", "have VP", "been VP", "Suggesting NP", NP in the end was "this
quest". Since it had a base form verb followed by a noun phrase, I added a new rule Vbar -> VerbB and
this helped in picking up the rule VP -> Vbar VP to parse the other verb phrases in the sentence

sentence : Sir Bedevere might have been suggesting this quest .
[PCFGParser]   best parse tree:
(START
  (S1
    (@NP-VP
      (NP
        (Nbar
          (Proper Sir) Bedevere)))
      (VP
        (MD might)
        (VP
          (VerbB have)

```
        (VP
          (Vbar
            (VerbPP been))
          (VP
            (Vbar
              (VerbPrP suggesting))
            (NP
              (Det this)
              (Nbar
                (Noun quest))))))))))
    (EOF .)))
```

Step 9: The next sentence "the Britons migrate south frequently ." had a noun phrase "the Britons" and a verb phrase "migrate south frequently". The Noun phrase in this sentence had new vocabularies "south" and "frequently" which are adverbs. So, I created a new tag AD for adverbs and added "south" and "frequently" to them. Since "migrate" is a base form verb, I added it under VerbB tag. Now, the existing rules were able to make the parser generate the tree using S1, since the verb phrase was formed by combination of verb phrases VerbB(migrate) and adverb phrase "south frequently" and we needed a new non-terminal symbol for adverb phrase. So I created ADVP for adverb phrase. I added the rule VP-> VP ADVP and ADVP-> AD AD. Also, "Britons" was listed as misc. But since it a plural proper noun, I created a new tag PropP and added new rule Nbar->PropP since a noun phrase could just a plural proper noun. This generated below parse tree using S1:

```
sentence : the Britons migrate south frequently .
[PCFGParser]   best parse tree:
(START
  (S1
    (@NP-VP
      (NP
        (Det the)
        (Nbar
          (PropP Britons)))
      (VP
        (VP
          (Vbar
            (VerbB migrate)))
        (ADVP
          (AD south)
          (AD frequently))))
    (EOF .)))
```

Step 10: The next sentence "Arthur and Guinevere ride frequently near the castle ." had all the words which were tagged meaningfully in the vocabulary file except the word "ride" which is base form of verb. So I added it under VerbB. The word "and" is a conjunction. So created a new tag CC and moved "and" from misc to CC. Noun phrase "Arthur and Guinevere" is a combination of a noun phrase(Proper noun Arthur) , a conjunction(and) and noun phrase(proper noun Guinevere). So I added the rule NP -> NP CC NP and since the verb ride was followed by just one adverb, I added ADVP-> AD which generated the following parse tree using S1

```
sentence : Arthur and Guinevere ride frequently near the castle .
[PCFGParser]   best parse tree:
(START
  (S1
    (@NP-VP
```

```
   (NP
     (@NP-CC
       (NP
         (Nbar
           (Proper Arthur)))
       (CC and))
     (NP
       (Nbar
         (Proper Guinevere))))
   (VP
     (@VP-ADVP
       (VP
         (Vbar
           (VerbB ride)))
       (ADVP
         (AD frequently)))
     (PP
       (Prep near)
       (NP
         (Det the)
         (Nbar
           (Noun castle))))))
   (EOF .)))
```

Step 11: The sentence "he suggests to grow fruit at home ." had a personal pronoun "he" which was listed as misc in vocabulary. So I created a new tag PProNN for personal pronouns and added "he" under it. Rest of the words in the sentence were added under meaningful tags in the vocabulary file. The sentence was divided into noun phrase "he" which is a personal pronoun and a verb phrase "suggests to grow fruit at home". Verb phrase had a singular present verb "suggests" which is denoted by non-terminal word Vbar, followed by TO and a verb phrase "grow fruit" which was followed by a prepositional phrase "at home". This prepositional phrase had a preposition "to" followed by a verb phrase "grow fruit at home" So I added a new rule VP-> TO VP and got the following parse tree using S1:

```
sentence : he suggests to grow fruit at home .
[PCFGParser]   best parse tree:
(START
  (S1
    (@NP-VP
      (NP
        (Nbar
          (PProNN he)))
      (VP
        (Vbar
          (VerbT suggests))
        (VP
          (TO to)
          (VP
            (VerbB grow)
            (NP
              (NP
                (Nbar
                  (Noun fruit)))
              (PP
                (Prep at)
```

```
                        (NP
                          (Nbar
                            (Noun home)))))))))
      (EOF .)))
```

Step 12: "riding to Camelot is not hard ." had word "camelot" which is a proper noun(place). It was listed as misc in vocabulary file. So I moved it under PrNN(non-people proper noun). Sentence begins with verb "riding". Even though "riding" is a verb, in this sentence "is" is the main verb. So I treated "riding to camelot" as a noun phrase and added "riding" under noun. This noun phrase is followed by verb phrase "is not hard" which has a verb "is" followed by an adjective "not hard". "not" was listed as misc. However, since it is a negating word, we didn't have any tag in the vocabulary. So I created a tag "NT" for "not". And added "hard" to JJ section as it is an adjective. We needed a rule for using an NT with an adjective forming an adjective phrase. So I added new rule VP-> Vbar ADJP and added a new non-terminal symbol ADJP-> NT JJ. This generated the following parse tree using S1:

```
sentence : riding to Camelot is not hard .
[PCFGParser]   best parse tree:
(START
  (S1
    (@NP-VP
      (NP
        (NP
          (Nbar
            (Noun riding)))
        (PP
          (TO to)
          (NP
            (Nbar
              (PrNN Camelot)))))
      (VP
        (Vbar
          (VerbT is))
        (ADJP
          (NT not)
          (JJ hard))))
    (EOF .)))
```

Step 13: The sentence "do coconuts speak ?" started with "do" and then a noun phrase "coconuts" which is a plural noun and a verb phrase "speak" which is just a base form of verb. I added "coconuts" under NNP(noun plural). I created a non-terminal symbol QP to represent a question phrase "do coconuts" and tagged "do" as QW in vocabulary file. Since we didn't have any rule for a sentence to begin with a question phrase, I added the rules S1->QP VP and QP-> QW NP. Now, the following parse tree was generated using S1.

```
sentence : do coconuts speak ?
[PCFGParser]   best parse tree:
(START
  (S1
    (@QP-VP
      (QP
        (QW do)
        (NP
          (Nbar
            (NNP coconuts))))
```

```
    (VP
      (Vbar
        (VerbB speak))))
  (EOF ?)))
```

Step 14: The sentence "why does England have a king ?" had WH phrase "why does England" and a verb phrase "have a king?". "why" was a new WH-adverb. So, I created a tag WAD for the wh-adverbs and a new non-terminal symbol WHP to represent a WH phrase. "why" was followed by "does England" which is a question phrase. So, the sentence was divided into a wh-phrase, question phrase and a verb phrase. So I added the rules S1-> WHP QP EOF, WHP->WAD and QP-> QW NP VP. This gave the following parse tree using S1.

```
sentence : why does England have a king ?
[PCFGParser]   best parse tree:
(START
  (S1
    (@WHP-QP
      (WHP
        (WAD why))
      (QP
        (@QW-NP
          (QW does)
          (NP
            (Nbar
              (PrNN England))))
        (VP
          (VerbB have)
          (NP
            (Det a)
            (Nbar
              (Noun king))))))
    (EOF ?)))
```

Sentences from challenge.Sentences.txt: I picked the 2 sentences "what horse does Arthur ride ?" and "who does Arthur suggest she carry ?" to solve in this homework.

1. The sentence "what horse does Arthur ride ?" starts with a WH phrase "what horse" followed by a question phrase "does Arthur" and a verb phrase "ride" which is just a verb. "what" was listed as misc in vocabulary file. So I created a new tag "WHD" to represent WH-determiners. I moved the word "does" from misc to "QW" as it is a question word in this sentence. Since a WH phrase starts with a WH word and a noun phrase, I added the rule WHP-> WHD NP and since sentence was mainly divided into WH phrase "what horse" and a question phrase "does Arthur ride?" S1-> WHP QP EOF. This generated the following parse tree using the grammar in S1.

```
sentence : what horse does Arthur ride ?
[PCFGParser]   best parse tree:
(START
  (S1
    (@WHP-QP
      (WHP
        (WHD what)
        (NP
          (Nbar
```

```
              (Noun horse))))
       (QP
         (@QW-NP
           (QW does)
           (NP
             (Nbar
               (Proper Arthur))))
       (VP
         (Vbar
           (VerbB ride)))))
    (EOF ?)))
```

2. The 2nd sentence "who does Arthur suggest she carry ?" had a few more tricky phrases in it. It starts with a WH phrase "who" followed by a question phrase "does Arthur" and a verb phrase "suggest". The last part of the sentence had a clause "she carry" which was formed by a noun phrase "she" and a verb phrase "carry". The two verbs "suggest" and "carry" are in their base forms. So, I added them under VerbB tag . "she" being a personal pronoun, I added it under the tag PProNN. The existing rules were not enough to parse this sentence but since it had a clause in the end, I added the rule S1-> WHP QP CL EOF and created a new non-terminal symbol CL-> NP VP for a clause.

```
sentence : who does Arthur suggest she carry ?
[PCFGParser]    best parse tree:
(START
  (S1
    (@@WHP-QP-CL
      (@WHP-QP
        (WHP
          (WHD who))
        (QP
          (@QW-NP
            (QW does)
            (NP
              (Nbar
                (Proper Arthur))))
          (VP
            (Vbar
              (VerbB suggest)))))
      (CL
        (NP
          (Nbar
            (PProNN she)))
        (VP
          (Vbar
            (VerbB carry)))))
    (EOF ?)))
```

Part 2: Exemplar Sentences

1. I formed a sentence "was Arthur carried to the castle ?" which is a proper english sentence as it was parsed successfully by the Stanford parser. Also, it gives a meaning which is a question which was asked to find out if the person Arthur was carried to the castle. This sentences was not parsed by our parser using S1. It was parsed using the grammar in S2 generating the below parse tree.

sentence : was Arthur carried to the castle ?
[PCFGParser]   best parse tree:
(START
  (S2
    (_VerbD
      (VerbD was)
      (_Proper
        (Proper Arthur)
        (_VerbPP
          (VerbPP carried)
          (_Prep
            (Prep to)
            (_Det
              (Det the)
              (_Noun
                (Noun castle)
                (_EOF
                  (EOF ?)))))))))))

The reason is, this sentence had a question phrase which was divided into a verb phrase "was", noun phrase "Arthur" and a verb phrase "carried to the castle". However, rules which we had only allowed a question phrase to start with question words do or does. And also, a question phrase was restricted to be a question word followed by a noun phrase and a verb phrase or just a noun phrase. But in this sentence a question phrase was starting with a verb. Since none of the rules written were able to parse this, a parse tree using S2 was generated instead of S1.

2. I formed a sentence using the words which were parsed before using S1 grammar as "why is to horse". This sentence by normal english rules, doesn't imply anything meaningful. However, since it has a WH word "why" followed by a verb phrase "is to horse" and since the verb phrase already has rule VP-> Vbar PP which parsed "is to horse" and "to horse" is matching with PP using the rule PP-> Prep Nbar, the whole sentence was parsed using the rule S1-> WHP VP. So a parse tree was generated using grammar in S1, even though the sentence wasn't meaningful.

sentence : why is to horse ?
[PCFGParser]   best parse tree:
(START
  (S1
    (@WHP-VP
      (WHP
        (WAD why))
      (VP
        (Vbar
          (VerbT is))
        (PP
          (TO to)
          (NP
            (Nbar
              (Noun horse))))))
    (EOF ?)))