Manjunath Panduranga
SUID: 637425699

Homework – 3

Context-Free Grammar

Part 1*:* Writing grammar rules

1. Arthur is the king .

Solution:

```
(E:\Anaconda3) E:\fall 2017\NLP\PCFG>java -jar pcfg.jar parse -t dev.sen *.gr
[PCFGParser]    log prob = -27.68       sentence : Arthur is the king .
[PCFGParser]    best parse tree:
(START
   (S1
      (@NP-VP
         (NP
            (Proper Arthur))
         (VP
            (VerbT is)
            (NP
               (Det the)
               (Nbar
                  (Noun king)))))
      (last .)))
```

No rules were added to solve this sentence. It was already solved while I run through the command prompt.

- But as we can see in the sentence, the S1 is provided with Noun Phase (proper noun) and the Verb Phase.
- The VP has further subphases of transitive verb(VerbT) and Noun Phase, which is further divided into Determinant(Det) and Noun

2. Arthur rides the horse near the castle .

Solution:

```
[PCFGParser]    log prob = -45.36       sentence : Arthur rides the horse near the castle .
[PCFGParser]    best parse tree:
(START
   (S1
      (@NP-VP
         (NP
            (Proper Arthur))
         (VP
            (VerbT rides)
            (NP
               (Det the)
               (Nbar
                  (Nbar
                     (Noun horse))
                  (PP
                     (Prep near)
                     (NP
                        (Det the)
                        (Nbar
                           (Noun castle))))))))
      (last .)))
```

No rules were added to solve this sentence. It was already solved while I run through the command prompt.

- But as we can see in the sentence, the S1 is provided with Noun Phase (proper noun) and the Verb Phase.
- The VP has further subphases of transitive verb(VerbT) and Noun Phase, which is further divided into Determinant(Det) and Noun.
- The Preposition Phase (PP) is also having a subphase of preposition(prep) and Noun Phase (NP)

3. Arthur rides the plodding horse near the castle .

Solution:

```
[PCFGParser]    log prob = -54.93       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-Adjective
                  (Det the)
                  (Adjective plodding))
               (Nbar
                  (Nbar
                     (Noun horse))
                  (PP
                     (Prep near)
                     (NP
                        (Det the)
                        (Nbar
                           (Noun castle))))))))
      (last .)))
```

Vocab rules(Vocab.gr):

1 Adjective plodding

1 last .

Below rules are added for solving this sentence(S1.gr):

1 NP Det Adjective NP

1 NP Nbar PP

This sentence required us to write new rules for the ‘Plodding’ word i.e. Adjective

And also for a noun followed by PP i.e. (horse near).

- The word ‘plodding’ is an adjective. ‘the’ is a determinant. So there are no rules which includes the adjective and the det ‘the’ along with noun phase. Hence the new rule of [NP->Det Adjective NP] is added.
- The word ‘horse’ is an noun. So there are no rules which includes the Nbar and the preposition phase. Hence the new rule of [NP-> Nbar PP] is added.

4. the Holy Grail is a chalice .

Solution:

```
[PCFGParser]    log prob = -35.02       sentence : the Holy Grail is a chalice .
[PCFGParser]    best parse tree:
(START
  (S1
    (@NP-VP
      (NP
        (Det the)
        (NP
          (pn Holy) Grail)))
      (VP
        (VerbT is)
        (NP
          (Det a)
          (Nbar
            (Noun chalice)))))
    (last .)))
```

Vocab rules(Vocab.gr):

1 pn Holy Grail

Below rules are added for solving this sentence(S1.gr):

1 NP Det NP

There was no rule for Noun Phase that starts with a Det and not expects an Adj, so added one for the purpose of solving this sentence as it starts with ‘the’.

Hence the new rule of [NP-> Det NP] is added.

5. the sensational Holy Grail is a sacred chalice .

Solution:

```
[PCFGParser]    log prob = -49.84       sentence : the sensational Holy Grail is a sacred chalice .
[PCFGParser]    best parse tree:
(START
   (S1
      (@NP-VP
         (NP
            (@Det-Adjective
               (Det the)
               (Adjective sensational))
            (NP
               (pn Holy) Grail)))
         (VP
            (VerbT is)
            (NP
               (@Det-Adjective
                  (Det a)
                  (Adjective sacred))
               (Nbar
                  (Noun chalice)))))
      (last .)))
```

Vocab rules(Vocab.gr):

1 Adjective sensational

1 Adjective sacred

Below rules are added for solving this sentence(S1.gr):

1 NP pn

1 NP Det Adjective Nbar

- The word ‘Sensational’ is an adjective. So there are no rules which includes the adjective and the det ‘the’. Hence the new rule of [NP->Det Adjective Nbar] is added.
- ‘Holy Grail’ is a proper noun. There is no rule in Noun Phase which includes this pn. Hence [NP->pn] is added.

6. every coconut was carried to the hottest mountains .

```
[PCFGParser]    log prob = -52.66       sentence : every coconut was carried to the hottest mountains .
[PCFGParser]    best parse tree:
(START
   (S1
      (@NP-VP
         (NP
            (Det every)
            (Nbar
               (Noun coconut)))
         (VP
            (verbpast was)
            (VP
               (verbpastpar carried)
               (VP
                  (TO to)
                  (NP
                     (@Det-sAdjective
                        (Det the)
                        (sAdjective hottest))
                     (NP
                        (pnouns mountains)))))))
      (last .)))
```

Vocab rules(Vocab.gr):

1 verbpast was

1 verbpast carried

1 sAdjective hottest

1 pnouns mountains

Below rules are added for solving this sentence(S1.gr):

1 VP verbpast VP
1 VP TO NP
1 NP Det sAdjective NP
1 NP pnouns

- The word ‘was’ is a past tense verb. There are no rules which includes the past tense verb and the verb phase. Hence the new rule of [VP-> verbpast VP] is added.
- The word ‘to’ belongs to category TO in vocab. But there are no rules which includes TO and the noun phase. Hence the new rule of [VP-> TO NP] is added.

- The word ‘hottest’ is a Superlative adjective. So there are no rules which includes the Superlative adjectives and the det ‘the’ and noun phase. Hence the new rule of [NP->Det sAdjective NP] is added.
- ‘mountains’ is a Plural noun(pnouns). There is no rule in Noun Phase which includes this pnouns. Hence [NP->pnouns] is added.

7. sixty strangers are at the Round Table .

Solution:

```
[PCFGParser]    log prob = -48.47       sentence : sixty strangers are at the Round Table .
[PCFGParser]    best parse tree:
(START
   (S1
      (@NP-VP
         (NP
            (num sixty)
            (NP
               (pnouns strangers)))
         (VP
            (verb3 are)
            (PP
               (Prep at)
               (NP
                  (Det the)
                  (NP
                     (pn Round) Table))))))
      (last .)))
```

Vocab rules(Vocab.gr):

1 num sixty

1 pnouns strangers

1 verb3 are

1 pn Round Table

Below rules are added for solving this sentence(S1.gr):

1 NP num NP

1 VP verb3 PP

- ‘sixty’ belongs to category numbers(num) in vocab. But there are no rules which includes num and the noun phase. Hence the new rule of [NP-> num NP] is added.
- ‘are’ belongs to category of Verbs (present, plural, third person).There is no rule in Verb Phase which includes this verb3 and prepositions ‘at’. Hence [VP->verb3 PP] is added.

8. Sir Lancelot might have spoken .

Solution:

```
[PCFGParser]    log prob = -35.90       sentence : Sir Lancelot might have spoken .
[PCFGParser]    best parse tree:
(START
   (S1
      (@NP-VP
         (NP
            (Proper Sir) Lancelot))
         (VP
            (Modals might)
            (VP
               (verbbase have)
               (VP
                  (verbpastpar spoken)))))
      (last .)))
```

Vocab rules(Vocab.gr):

1 Modals might

1 verbbase have

1 verbpastpar spoken

Below rules are added for solving this sentence(S1.gr):

1 VP Modals VP

1 VP verbbase VP

1 VP verbpastpar

- 'might' is a modal in vocab. But there is no rules which includes this modals with Verb Phase. Hence [VP-> Modals VP] rule is added.
- 'have' is a base form verb (verbbase). But there is no rules which includes this verbbase with Verb Phase. Hence [VP-> verbbase VP] rule is added
- 'spoken' is a past participle verb. But there is no rules which includes this verbpastpar with Verb Phase. Hence [VP-> verbpastpar] rule is added

9. Guinevere had been riding with Patsy for five weary nights .

Solution:

```
[PCFGParser]    log prob = -74.32       sentence : Guinevere had been riding with Patsy for five weary nights .
[PCFGParser]    best parse tree:
(START
   (S1
      (@NP-VP
         (NP
            (Proper Guinevere))
         (VP
            (verbpastpar had)
            (VP
               (verbpastpar been)
               (VP
                  (verbpre riding)
                  (PP
                     (Prep with)
                     (NP
                        (Proper Patsy)
                        (PP
                           (Prep for)
                           (NP
                              (num five)
                              (NP
                                 (Adjective weary)
                                 (pnouns nights)))))))))))
      (last .)))
```

Vocab rules(Vocab.gr):

1 verbpastpar had

1 verbpastpar been

1 verbpre riding

1 num five

1 Adjective weary

Below rules are added for solving this sentence(S1.gr):

1 VP verbpre PP

1 NP Proper PP

1 NP Adjective pnouns

- The word ‘riding’ is a present participle verb. But there is no rules which includes this verbpastpar with Prepositions. Hence [VP-> verbpre PP] rule is added.
- The ‘Patsy’ is a Proper Noun. But there is no rules which includes this proper noun with preposition in the sentence. Hence [NP-> Proper PP] rule is added to the sentence.
- ‘weary’ is an adjective. But there is no rules which includes this Adjective with pnoun ‘nights’. Hence [NP-> Adjective pnouns] rule is added.

10.Sir Bedevere might have been suggesting this quest .

Solution:

```
[PCFGParser]    log prob = -52.37       sentence : Sir Bedevere might have been suggesting this quest .
[PCFGParser]    best parse tree:
(START
   (S1
      (@NP-VP
         (NP
            (Proper Sir) Bedevere))
         (VP
            (Modals might)
            (VP
               (verbbase have)
               (VP
                  (verbpastpar been)
                  (VP
                     (verbpre suggesting)
                     (NP
                        (Det this)
                        (Nbar
                           (Noun quest)))))))
      (last .)))
```

Vocab rules(Vocab.gr):

1 verbpre suggesting

Below rules are added for solving this sentence(S1.gr):

1 VP verbpre NP

The ‘suggesting’ is a present participle verb. But there is no rules which includes this verbpre with Noun Phase in the sentence. Hence [VP-> verbpre NP] rule is added to the sentence.

11.the Britons migrate south frequently .

Solution:

```
[PCFGParser]    log prob = -41.46       sentence : the Britons migrate south frequently .
[PCFGParser]    best parse tree:
(START
  (S1
    (@NP-VP
      (NP
        (Det the)
        (NP
          (ppn Britons)))
      (VP
        (verbbase migrate)
        (VP
          (Adverbs south)
          (VP
            (Adverbs frequently)))))
    (last .)))
```

Vocab rules(Vocab.gr):

1 ppn Britons

1 verbbase migrate

1 Adverbs south

1 Adverbs frequently

Below rules are added for solving this sentence(S1.gr):

1 NP ppn

1 VP Adverbs VP

1 VP Adverbs

- The 'Britons' is a Plural proper noun. But there are no rules which includes this ppn with Noun Phase in the sentence. Hence [NP-> ppn] rule is added to the sentence.
- The 'south' is Adverb. But there are no rules which includes this adverb with Verb Phase in the sentence. Hence [VP-> Adverbs VP] rule is added to the sentence.
- 'frequently' is also a adverb. But there are no rules which includes this adverb. Hence [VP-> adverbs] rule is added to the sentence.

12. Arthur and Guinevere ride frequently near the castle .

Solution:

```
[PCFGParser]    log prob = -51.86       sentence : Arthur and Guinevere ride frequently near the castle .
[PCFGParser]    best parse tree:
(START
   (S1
      (@NP-VP
         (NP
            (@Proper-cc
               (Proper Arthur)
               (cc and))
            (NP
               (Proper Guinevere)))
         (VP
            (verbbase ride)
            (VP
               (Adverbs frequently)
               (PP
                  (Prep near)
                  (NP
                     (Det the)
                     (Nbar
                        (Noun castle)))))))
      (last .)))
```

Vocab rules(Vocab.gr):

1 cc and

1 verbbase ride

Below rules are added for solving this sentence(S1.gr):

1 NP Proper cc NP

1 VP Adverbs PP

- The ‘Arthur’ is a proper noun. Word ‘and’ is a Coordinating conjunction. But there are no rules which includes this proper noun with Coordinating conjunction along with Noun Phase in the sentence. Hence [NP-> Proper cc NP] rule is added to the sentence.
- The ‘frequently’ is an adverb. But there are no rules which includes this adverb with preposition in the sentence. Hence [VP-> Adverbs PP] rule is added to the sentence.

13.he suggests to grow fruit at home .

Solution:

```
[PCFGParser]    log prob = -55.78       sentence : he suggests to grow fruit at home .
[PCFGParser]    best parse tree:
(START
   (S1
      (@NP-VP
         (NP
            (pr he))
         (VP
            (sverb suggests)
            (VP
               (TO to)
               (VP
                  (verbbase grow)
                  (NP
                     (Nbar
                        (Noun fruit))
                     (PP
                        (Prep at)
                        (NP
                           (Nbar
                              (Noun home)))))))))
      (last .)))
```

Vocab rules(Vocab.gr):

1 pr he

1 sverb suggests

1 TO to

1 verbbase grow

Below rules are added for solving this sentence(S1.gr):

1 NP pr

1 NP Nbar

1 VP sverb VP

1 VP TO VP

1 VP verbbase NP

- The ‘he’ is a Personal pronoun. But there are no rules which includes this Personal pronoun to the Noun phase in the sentence. Hence [NP-> pr] rule is added to the sentence.

- The ‘suggests’ is a third person singular verb. But there are no rules which includes this third person singular verb with the Verb phase in the sentence. Hence [VP-> sverb VP] rule is added to the sentence.
- The ‘to’ belongs to the TO category in vocab. But there are no rules in S1 which includes this TO with the Verb phase in the sentence. Hence [VP-> TO VP] rule is added to the sentence.
- The term ‘grow’ is a verb base. But there are no rules in S1 which includes this ‘grow’ with the Noun phase in the sentence. Hence [VP-> verbbase NP] rule is added to the sentence.
- The term ‘home’ is a singular noun. But there are no rules in S1 which includes only singular noun with the Noun phase in the sentence. Hence [NP-> Nbar] rule is added to the sentence.

14.riding to Camelot is not hard .

Solution:

```
[PCFGParser]    log prob = -34.95       sentence : riding to Camelot is not hard .
[PCFGParser]    best parse tree:
(START
   (S1
      (VP
         (verbpre riding)
         (VP
            (TO to)
            (NP
               (pn Camelot)
               (VP
                  (@VerbT-NOT
                     (VerbT is)
                     (NOT not))
                  (Adjective hard)))))
      (last .)))
```

Vocab rules(Vocab.gr):

1 pn Camelot

1 NOT not

1 Adjective hard

Below rules are added for solving this sentence(S1.gr):

1 S1 VP last

1 NP pn VP

1 VP VerbT NOT Adjective

- The sentence S1 does not include the rule which includes only VP and Last. Since NP is separated by using 'to'
- The term 'Camelot' is a proper noun. But there are no rules in which includes this proper noun with the Verb phase in the sentence. Hence [NP-> pn VP] rule is added to the sentence.
- The term 'is' is a verb tag. The term 'not' belongs to NOT category tag. 'hard' belongs to Adjective. But there are no rules in which includes this these three in the sentence. Hence [VP-> VerbT NOT Adjective] rule is added to the sentence.

15.do coconuts speak ?

Solution:

```
[PCFGParser]    log prob = -23.72       sentence : do coconuts speak ?
[PCFGParser]    best parse tree:
(START
   (S1
      (verb1 do)
      (S1
         (@NP-VP
            (NP
               (pnouns coconuts))
            (VP
               (verb3 speak)))
         (last ?))))
```

Vocab rules(Vocab.gr):

1 verb1 do

1 pnouns coconuts

1 verb3 speak

Below rules are added for solving this sentence(S1.gr):

1 S1 verb1 S1

1 VP verb3

Below rules are added for solving this sentence. This was a tricky sentence that required us to introduce new Vocab tag ‘verb3’ for ‘speak’ word. Later the sentence parsing was completed by implementing the LOOP mechanism i.e. by resolving to S1 since S1 has already had rules that could solve the rest of the sentence

16.why does England have a king ?

Solution:

```
[PCFGParser]    log prob = -35.20       sentence : why does England have a king ?
[PCFGParser]    best parse tree:
(START
   (S1
      (whadv why)
      (S1
         (verb1 does)
         (S1
            (@NP-VP
               (NP
                  (pn England))
               (VP
                  (verbbase have)
                  (NP
                     (Det a)
                     (Nbar
                        (Noun king)))))
            (last ?)))))
```

Vocab rules(Vocab.gr):

1 whadv why

1 verb1 does

1 pn England

1 last ?

Below rules are added for solving this sentence(S1.gr):

1 S1 whadv S1

The sentence parsing was completed by implementing the LOOP mechanism i.e. by resolving to S1 since S1 has already had rules that could solve the rest of the sentence

Hence the rule [S1->whadv S1] is included, which is a loop mechanism.

Challenge sentences

17.the king drank to the castle that was his home .

Solution:

```
[PCFGParser]    log prob = -66.34       sentence : the king drank to the castle that was his home .
[PCFGParser]    best parse tree:
(START
   (S1
      (@NP-VP
         (NP
            (Det the)
            (Nbar
               (Noun king)))
         (VP
            (verbpast drank)
            (VP
               (TO to)
               (NP
                  (Det the)
                  (Nbar
                     (Nbar
                        (Noun castle))
                     (PP
                        (Det that)
                        (VP
                           (verbpast was)
                           (VP
                              (ppr his)
                              (NP
                                 (Nbar
                                    (Noun home)))))))))))
      (last .)))
```

Vocab rules(Vocab.gr):

1 verbpast drank

1 ppr his

Below rules are added for solving this sentence(S1.gr):

1 PP Det VP

1 VP ppr NP

- The term 'that' is a Determiners. But there are no rules in which includes this Determiners with the Verb Phase in the sentence. Hence [PP-> Det VP] rule is added to the sentence.
- The term 'his' is a Possessive personal pronoun. But there are no rules in which includes this ppr with the Noun Phase in the sentence. Hence [VP-> ppr NP] rule is added to the sentence.

18.do not speak again !

Solution:

```
[PCFGParser]    log prob = -20.55       sentence : do not speak again !
[PCFGParser]    best parse tree:
(START
   (S1
      (verb1 do)
      (S1
         (NOT not)
         (S1
            (VP
               (verb3 speak)
               (Adverbs again))
            (last !)))))
[PCFGParser]    cross-entropy=5.844 perplexity=5.742e+01
```

Vocab rules(Vocab.gr):

1 Adverbs again

1 last !

Below rules are added for solving this sentence(S1.gr):

1 S1 NOT S1

Manjunath Panduranga
SUID: 637425699

1 VP verb3 Adverbs

- The term ‘speak’ is a Verb3 category. The term ‘again’ is a Advebrb3 category. But there are no rules in which includes this category of verb3 with adverbs in the Verb Phase in the sentence. Hence [VP-> verb3 Adverbs] rule is added to the sentence.
- the sentence parsing was completed by implementing the LOOP mechanism i.e. by resolving to S1 since S1 has already had almost all rules that could solve the rest of the sentence.
- hence the rule [S1->whadv S1] is included, which is a loop mechanism.

Part 2: Exemplar Sentences

make up a sentence that uses some of the same words in the sentences that you already parsed and is an actual English sentence, but cannot be parsed by those rules.

1. why are snakes frequently migrating to the south ?

Solution:

```
[PCFGParser]    log prob = -81.30       sentence : why are snakes frequently migrating to the south ?
[PCFGParser]    best parse tree:
(START
   (S2
      (_whadv
         (whadv why)
         (_verb3
            (verb3 are)
            (_pnouns
               (pnouns snakes)
               (_Adverbs
                  (Adverbs frequently)
                  (_verbpre
                     (verbpre migrating)
                     (_TO
                        (TO to)
                        (_Det
                           (Det the)
                           (_Adverbs
                              (Adverbs south)
                              (_last
                                 (last ?))))))))))))
```

The above sentence is made up of the words that are taken from the sentences that are already parsed.

The sentence is also a meaningful, hence it can be called as an actual English sentence.

But then it cannot be parsed through those rules, the reason is:

1. the word ‘why’ is Wh-adverbs, for which rules are not added with the Verb Phase.
2. The word ‘are’ belongs to the Verbs (present, plural, third person). But there is no rules which includes this verb3 with noun phase(NP).
3. The word ‘snakes’ is a plural noun. But no rule is added for pnouns with verb phase in the sentence
   Hence based on the 3 rules which are not added to the sentence, could say that the sentence cannot be parsed.

make up a string of words that should not be an actual English sentence, but your grammar will parse it.

2. not horse growing through Uther Pendragon neither another coconut .

Solution:

```
[PCFGParser]    log prob = -48.12       sentence : not horse growing through Uther Pendragon neither another coconut .
[PCFGParser]    best parse tree:
(START
   (S1
      (NOT not)
      (S1
         (@NP-VP
            (NP
               (Nbar
                  (Noun horse)))
            (VP
               (verbpre growing)
               (PP
                  (Prep through)
                  (NP
                     (@Proper-cc
                        (Proper Uther) Pendragon)
                        (cc neither))
                     (NP
                        (Det another)
                        (Nbar
                           (Noun coconut)))))))
         (last .))))
[PCFGParser]    cross-entropy=5.945 perplexity=6.163e+01
```

Normally a sentence which is meaningful is an English sentence.

The above sentence is not an English sentence since it has no meaning. But it is able to parse through the normal grammar rules. The reason for declaring this sentence as non-English sentence is,

1. there is no meaning involved in this sentence
2. the word neither which is a coordinating conjunction is being used. But ‘nor’ is not included in this sentence.

If the grammar parses obvious non-English sentences, it is “over- generalization”. Here the grammar includes rules like{ [S1->NP VP], [S1->NOT S1], [NP->Nbar], [VP->verbpre PP], [PP->Prep NP], [NP->Proper cc NP], [NP->Det Nbar]}. These rules are already mentioned, which are allowing this sentence to parse through.

The parse tree output is shown in the screenshot shared above.