Name: Rajkumar Pillai

Note: The code for the solutions are present in q1.py file. For display of parse trees, the installation of Stanford CoreNLP library is necessary along with proper configurations. The code for generating the parse trees are commented in the q1.py file.

Question-1

Solution:

1. RESULTS FOR FILE Dracula.txt
Mean: 19.343680709534368
Standard Deviation: 13.544127056426982
Length of Longest sentence: 332

The Longest sentence:

You may copy it, give it away orre-use it under the terms of the Project Gutenberg License includedwith this eBook or online at www.gutenberg.org/licenseTitle: DraculaAuthor: Bram StokerRelease Date: August 16, 2013 [EBook #345]Language: English*** START OF THIS PROJECT GUTENBERG EBOOK DRACULA ***Produced by Chuck Greif and the Online DistributedProofreading Team at http://www.pgdp.net (This file wasproduced from images generously made available by TheInternet Archive)   DRACULA DRACULA   _by_   Bram Stoker   [Illustration: colophon] NEW YORK   GROSSET & DUNLAP   _Publishers_   Copyright, 1897, in the United States of America, according   to Act of Congress, by Bram Stoker   [_All rights reserved._]   PRINTED IN THE UNITED STATES   AT   THE COUNTRY LIFE PRESS, GARDEN CITY, N.Y.   TO   MY DEAR FRIEND   HOMMY-BEGCONTENTSCHAPTER I   PageJonathan Harker's Journal 1CHAPTER IIJonathan Harker's Journal   14CHAPTER IIIJonathan Harker's Journal 26CHAPTER IVJonathan Harker's Journal   38CHAPTER VLetters--Lucy and Mina 51CHAPTER VIMina Murray's Journal   59CHAPTER VIICutting from "The Dailygraph," 8 August   71CHAPTER VIIIMina Murray's Journal   84CHAPTER IXMina Murray's Journal   98CHAPTER XMina Murray's Journal 111CHAPTER XILucy Westenra's Diary   124CHAPTER XIIDr. Seward's Diary 136CHAPTER XIIIDr. Seward's Diary   152CHAPTER XIVMina Harker's Journal 167CHAPTER XVDr. Seward's Diary   181CHAPTER XVIDr. Seward's Diary 194CHAPTER XVIIDr. Seward's Diary   204CHAPTER XVIIIDr. Seward's Diary 216CHAPTER XIXJonathan Harker's Journal   231CHAPTER XXJonathan Harker's Journal   243CHAPTER XXIDr. Seward's Diary   256CHAPTER XXIIJonathan Harker's Journal   269CHAPTER XXIIIDr. Seward's Diary 281CHAPTER XXIVDr. Seward's Phonograph Diary, spoken by Van Helsing   294CHAPTER XXVDr. Seward's Diary   308CHAPTER XXVIDr. Seward's Diary 322CHAPTER XXVIIMina Harker's Journal   338DRACULACHAPTER IJONATHAN HARKER'S JOURNAL(_Kept in shorthand._)_3 May.

Length of shortest sentence: 1

The Shortest sentence: Oh!

The Selected Sentence for POS Tagging:

We left in pretty good time, and came after nightfall to Klausenburgh.

Parts of speech tag: [('We', 'PRP'), ('left', 'VBD'), ('in', 'IN'), ('pretty', 'RB'), ('good', 'JJ'), ('time', 'NN'), ('and', 'CC'), ('came', 'VBD'), ('after', 'IN'), ('nightfall', 'DT'), ('to', 'TO'), ('Klausenburgh', 'VB')]

The Selected Sentence for POS Tagging:

Do you know where you are going, and what you are going to?"

Parts of speech tag: [('Do', 'VB'), ('you', 'PRP'), ('know', 'VB'), ('where', 'WRB'), ('you', 'PRP'), ('are', 'VBP'), ('going', 'VBG'), ('and', 'CC'), ('what', 'WP'), ('you', 'PRP'), ('are', 'VBP'), ('going', 'VBG'), ('to', 'TO')]

2.RESULTS FOR FILE HoundBaskervilles.txt

Mean: 23.188818565400844
Standard Deviation: 15.594983775866266
Length of Longest sentence: 201
The Longest sentence:
The walk in the centre is about eight feet across."
    "Is there anything between the hedges and the walk?"
    "Yes, there is a strip of grass about six feet broad on either
    side."
    "I understand that the yew hedge is penetrated at one point by a
    gate?"
    "Yes, the wicket-gate which leads on to the moor."
    "Is there any other opening?"
    "None."
    "So that to reach the yew alley one either has to come down it
    from the house or else to enter it by the moor-gate?"
    "There is an exit through a summer-house at the far end."
    "Had Sir Charles reached this?"
    "No; he lay about fifty yards from it."
    "Now, tell me, Dr. Mortimer—and this is important—the marks which
    you saw were on the path and not on the grass?"
    "No marks could show on the grass."
    "Were they on the same side of the path as the moor-gate?"
    "Yes; they were on the edge of the path on the same side as the
    moor-gate."
    "You interest me exceedingly.

Length of shortest sentence: 1

The Shortest sentence: Halloa!

The Selected Sentence for POS Tagging:

Some people without possessing genius have a remarkable power of stimulating it.

Parts of speech tag: [('Some', 'DT'), ('people', 'NNS'), ('without', 'IN'), ('possessing', 'VBG'), ('genius', 'NN'), ('have', 'VBP'), ('a', 'DT'), ('remarkable', 'JJ'), ('power', 'NN'), ('of', 'IN'), ('stimulating', 'VBG'), ('it', 'PRP')]

The Selected Sentence for POS Tagging:

There were several Mortimers, but only one who could be our visitor.

Parts of speech tag: [('There', 'EX'), ('were', 'VBD'), ('several', 'JJ'), ('Mortimers', 'NNS'), ('but', 'CC'), ('only', 'RB'), ('one', 'NN'), ('who', 'WP'), ('could', 'MD'), ('be', 'VB'), ('our', 'PRP$'), ('visitor', 'NN')]

3.RESULTS FOR FILE WarAndPeace.txt

Mean: 22.54235883527329

Standard Deviation: 17.80130410932107

Length of Longest sentence: 884

The Longest sentence:

   You may copy it, give it away or re-useit under the terms of the Project Gutenberg License included with thiseBook or online at www.gutenberg.orgTitle: War and PeaceAuthor: Leo TolstoyTranslators: Louise and Aylmer MaudePosting Date: January 10, 2009 [EBook #2600]Last Updated: January 21, 2019Language: EnglishCharacter set encoding: UTF-8*** START OF THIS PROJECT GUTENBERG EBOOK WAR AND PEACE ***An Anonymous Volunteer, and David WidgerWAR AND PEACEBy Leo Tolstoy/Tolstoi CONTENTS BOOK ONE: 1805   CHAPTER I   CHAPTER II   CHAPTER III   CHAPTER IV   CHAPTER V   CHAPTER VI   CHAPTER VII   CHAPTER VIII   CHAPTER IX   CHAPTER X   CHAPTER XI   CHAPTER XII   CHAPTER XIII   CHAPTER XIV   CHAPTER XV   CHAPTER XVI   CHAPTER XVII   CHAPTER XVIII   CHAPTER XIX   CHAPTER XX   CHAPTER XXI   CHAPTER XXII   CHAPTER XXIII   CHAPTER XXIV   CHAPTER XXV   CHAPTER XXVI   CHAPTER XXVII   CHAPTER XXVIII   BOOK TWO: 1805   CHAPTER I   CHAPTER II   CHAPTER III   CHAPTER IV   CHAPTER V   CHAPTER VI   CHAPTER VII   CHAPTER VIII   CHAPTER IX   CHAPTER X   CHAPTER XI   CHAPTER XII   CHAPTER XIII   CHAPTER XIV   CHAPTER XV   CHAPTER XVI   CHAPTER XVII   CHAPTER XVIII   CHAPTER XIX   CHAPTER XX   CHAPTER XXI   BOOK THREE: 1805   CHAPTER I   CHAPTER II   CHAPTER III   CHAPTER IV   CHAPTER V   CHAPTER VI   CHAPTER VII   CHAPTER VIII   CHAPTER IX   CHAPTER X   CHAPTER XI   CHAPTER XII   CHAPTER XIII   CHAPTER XIV   CHAPTER XV   CHAPTER XVI   CHAPTER XVII   CHAPTER XVIII   CHAPTER XIX   BOOK FOUR: 1806   CHAPTER I   CHAPTER II   CHAPTER III   CHAPTER IV   CHAPTER V   CHAPTER VI   CHAPTER VII   CHAPTER VIII   CHAPTER IX   CHAPTER X   CHAPTER XI   CHAPTER XII   CHAPTER XIII   CHAPTER XIV   CHAPTER XV   CHAPTER XVI   BOOK FIVE: 1806 - 07   CHAPTER I   CHAPTER II   CHAPTER III   CHAPTER IV   CHAPTER V   CHAPTER VI   CHAPTER VII   CHAPTER VIII   CHAPTER IX   CHAPTER X   CHAPTER XI   CHAPTER XII   CHAPTER XIII   CHAPTER XIV   CHAPTER XV   CHAPTER XVI   CHAPTER XVII   CHAPTER XVIII   CHAPTER XIX   CHAPTER XX   CHAPTER XXI   CHAPTER XXII   BOOK SIX: 1808 - 10   CHAPTER I   CHAPTER II   CHAPTER III   CHAPTER IV   CHAPTER V   CHAPTER VI   CHAPTER VII   CHAPTER VIII   CHAPTER IX   CHAPTER X   CHAPTER XI   CHAPTER XII   CHAPTER XIII   CHAPTER XIV   CHAPTER XV   CHAPTER XVI   CHAPTER XVII   CHAPTER XVIII   CHAPTER XIX   CHAPTER XX   CHAPTER XXI   CHAPTER XXII   CHAPTER XXIII   CHAPTER XXIV   CHAPTER XXV   CHAPTER XXVI   BOOK SEVEN: 1810 - 11   CHAPTER I   CHAPTER II   CHAPTER III   CHAPTER IV   CHAPTER V   CHAPTER VI   CHAPTER VII   CHAPTER VIII   CHAPTER IX   CHAPTER X   CHAPTER XI   CHAPTER XII   CHAPTER XIII   BOOK EIGHT: 1811 - 12   CHAPTER I   CHAPTER II   CHAPTER III   CHAPTER IV   CHAPTER V   CHAPTER VI   CHAPTER VII   CHAPTER VIII   CHAPTER IX   CHAPTER X   CHAPTER XI   CHAPTER XII   CHAPTER XIII   CHAPTER XIV   CHAPTER XV   CHAPTER XVI   CHAPTER XVII   CHAPTER XVIII   CHAPTER XIX   CHAPTER XX   CHAPTER XXI   CHAPTER XXII   BOOK NINE: 1812   CHAPTER I   CHAPTER II   CHAPTER III   CHAPTER IV   CHAPTER V   CHAPTER VI   CHAPTER VII   CHAPTER VIII   CHAPTER IX   CHAPTER X   CHAPTER XI   CHAPTER XII   CHAPTER XIII   CHAPTER XIV   CHAPTER XV   CHAPTER XVI   CHAPTER XVII   CHAPTER XVIII   CHAPTER XIX   CHAPTER XX   CHAPTER XXI   CHAPTER XXII   CHAPTER XXIII   BOOK TEN: 1812   CHAPTER I   CHAPTER II   CHAPTER III   CHAPTER IV   CHAPTER V   CHAPTER VI   CHAPTER VII   CHAPTER VIII   CHAPTER IX   CHAPTER X   CHAPTER XI   CHAPTER XII   CHAPTER XIII   CHAPTER XIV   CHAPTER XV

BOOK ONE: 1805CHAPTER I"Well, Prince, so Genoa and Lucca are now just family estates of theBuonapartes.

Length of shortest sentence: 1

The Shortest sentence: None.

The Selected Sentence for POS Tagging:

Our gracious sovereign recognizes his high vocation and will be true to it.

Parts of speech tag: [('Our', 'PRP$'), ('gracious', 'JJ'), ('sovereign', 'NN'), ('recognizes', 'VBZ'), ('his', 'PRP$'), ('high', 'JJ'), ('vocation', 'NN'), ('and', 'CC'), ('will', 'MD'), ('be', 'VB'), ('true', 'JJ'), ('to', 'TO'), ('it', 'PRP')]

The Selected Sentence for POS Tagging:

She wanted to find, and still seeks, some secret motive in our actions.

Parts of speech tag: [('She', 'PRP'), ('wanted', 'VBD'), ('to', 'TO'), ('find', 'VB'), ('and', 'CC'), ('still', 'RB'), ('seeks', 'VBZ'), ('some', 'DT'), ('secret', 'JJ'), ('motive', 'NN'), ('in', 'IN'), ('our', 'PRP$'), ('actions', 'NNS')]

Question-2:

Solution:

Note: The parse trees are obtained by using Stanford CoreNLP library package the definition of these tags are present in this link: https://stanfordnlp.github.io/CoreNLP/pos.html

Parse Trees for Dracula.txt

1.We left in pretty good time, 89and came after nightfall to Klausenburgh.

```
                                      ROOT
                                       |
                                       S
                  _____|_____
                 |                          VP
                 |                           |
                 |        VP                 |_____
                 |      __|_____             |                           |
                 |     |       PP            |                           VP
                 |     |     __|____         |                      _____|_____
                 |     |    |      NP        |    |                PP              PP
                 |     |    |    __|___       |    |              __|____         __|_____
                NP     |    |   ADJP  |       |    |    |        |      NP        |      NP
                 |     |    |   _|_    |      |    |   |        |       |        |       |
               PRP   VBD   IN  RB    JJ     NN   CC  VBD  IN       NN      TO      NNP
                |     |    |    |      |      |    |    |    |        |       |        |
                We   left  in pretty  good  time and came after  nightfall  to   Klausenburgh
```

CFG:


```
S → NP | VP
NP → PRP | ADJP | NN | NNP
VP → VP | CC | VBD | PP
PRP → We
VBD → left | came
PP → IN | NP | TO
ADJP → RB | JJ
NN → time | nightfall
NNP → Klausenburgh
TO → to
CC → and
IN → in | after
RB → pretty
JJ → good
```

2.Do you know where you are going, and what you are going to?"
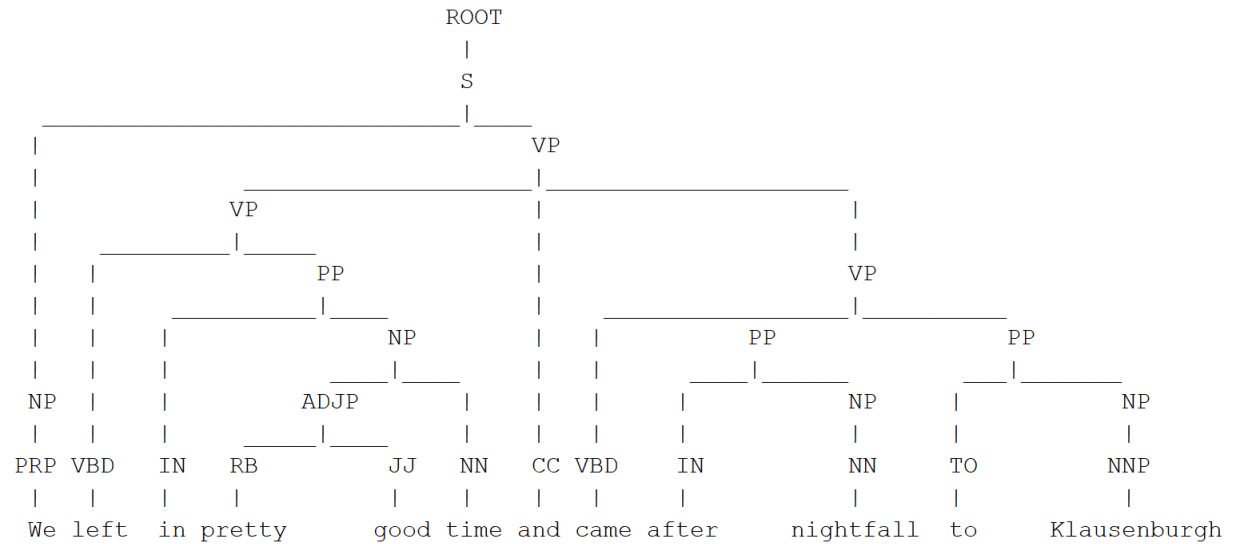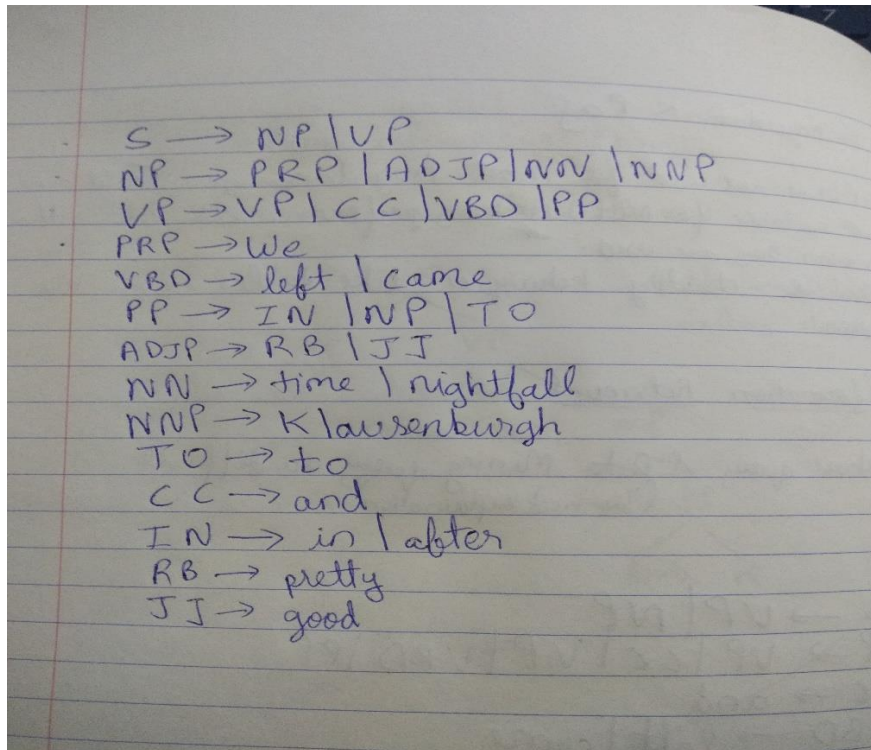
```
            ROOT
             |
             S
             |
             VP
          ___|____
         |        S
         |    ____|_____
         |   |                 VP
         |   |      _____|_____
         |   |     |                              SBAR
         |   |     |               _____|_____
         |   |     |              |                 |            SBAR
         |   |     |              |                 |         _____|____
         |   |     |            SBAR                |        |            S
         |   |     |       _____|____              |        |        ____|____
         |   |     |      |           S             |        |       |         VP
         |   |     |      |       ____|___          |        |       |     ____|____
         |   |     |      |      |        VP        |        |       |    |         VP
         |   |     |      |      |      __|__        |       |       |    |       __|__
         |  NP     |   WHADVP   NP     |     VP     |     WHNP   NP  |    |        PP
         |   |     |      |      |     |     |      |       |     |   |    |        |
        VB  PRP   VB    WRB    PRP   VBP    VBG    CC     WP   PRP  VBP  VBG       TO
         |   |     |      |      |     |     |      |       |     |   |    |        |
        Do  you  know  where   you   are  going  and   what  you  are  going     to
```
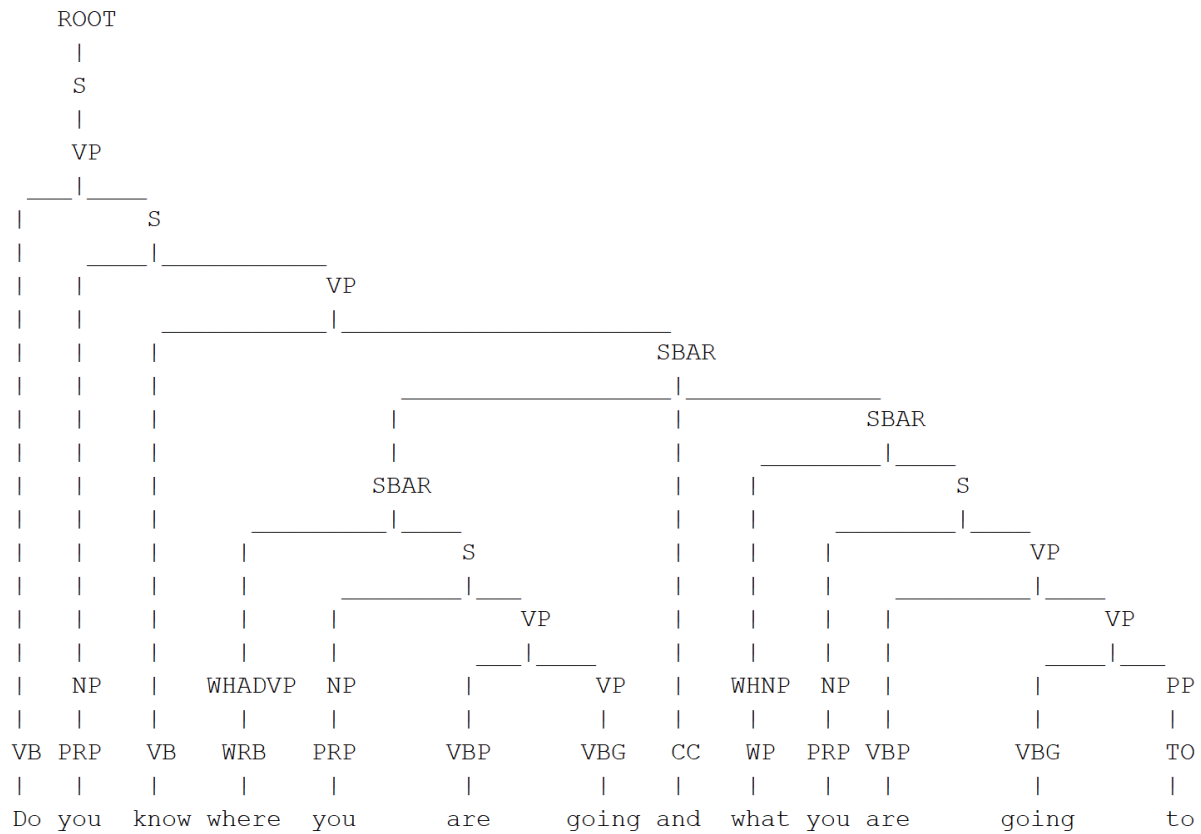
CFG:


```
S → VP | NP
VP → VB | S | SBAR | VBP | VP | VBG
SBAR → SBAR | CC | WHADVP | S | WHNP
NP → PRP
WHADVP → WRB
WHNP → WP
PP → TO
VB → Do
PRP → you
VB → know
WRB → where
PRP → you
VBP → are
VBG → going
CC → and
WP → what
PRP → you
VBP → are
VBG → going
TO → to
```

Parse Trees for HoundBaskervilles.txt

1.Some people without possessing genius have a remarkable power of stimulating it.

```
                                        ROOT
                                         |
                                         S
                                         |____
                       |                         VP
                       |                     ____|_____
                       NP                    |                NP
             _____|_____             |            ____|_____
             |                  PP            |            |               PP
             |            _____|_____       |            |         _____|_____
             |            |            S       |            |         |            S
             |            |            |        |            |         |            |
             |            |            VP       |            |         |            VP
             |            |         ___|____    |            |         |         ___|___
             NP           |         |        NP  |     NP          |         |        NP
           __|____        |         |        |   |   __|___        |         |        |
           DT     NNS      IN       VBG       NN  VBP  DT    JJ      NN  IN    VBG      PRP
           |       |        |         |        |   |    |     |       |   |     |        |
          Some  people without possessing  genius have a remarkable power of stimulating it
```

CFG:


```
S → NP|VP
NP → NP|PP |DT|NNS|JJ|NN|PRP
VP → VBP|NP|VBG
PP → IN|S
DT → Some
NNS → People
IN → without
VBG → possessing
NN → genius
VBP → have
DT → a
JJ → remarkable
NN → power
IN → of
VBG → stimulating
PRP → it
```

2.There were several Mortimers, but only one who could be our visitor.
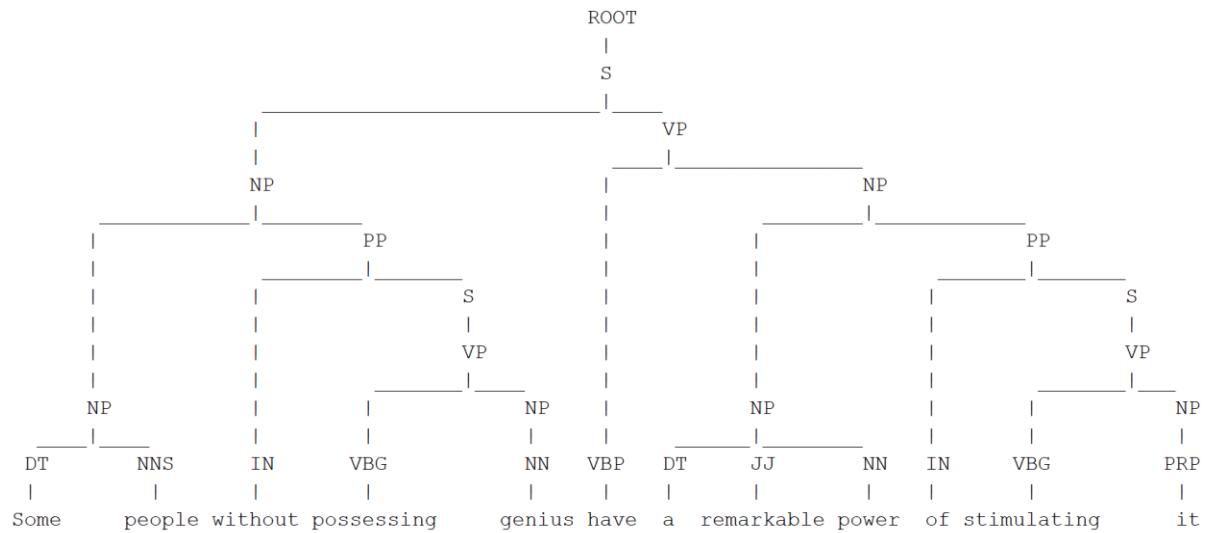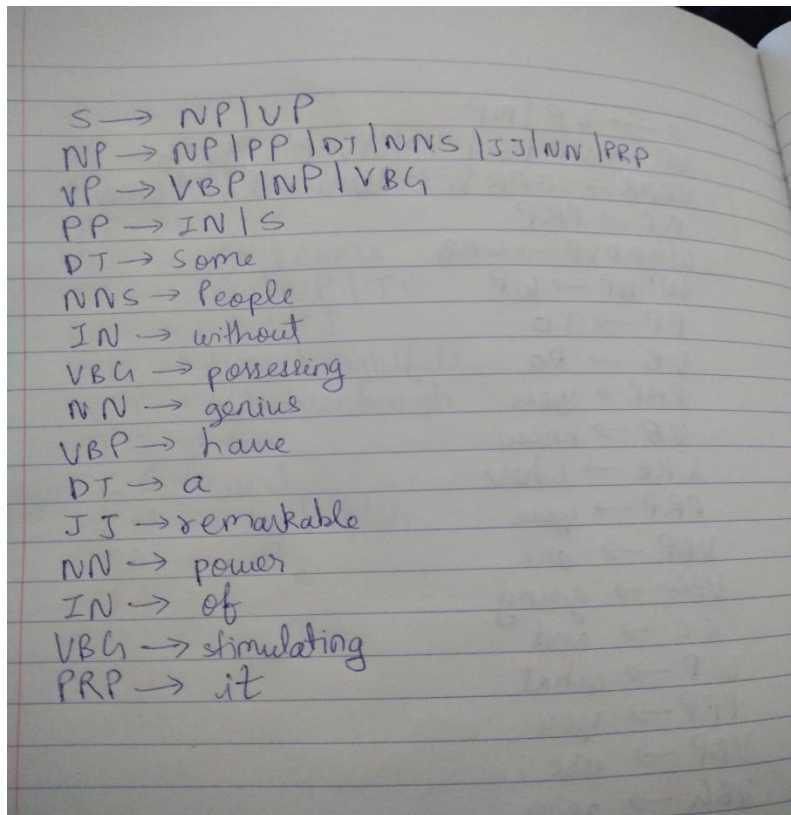
```
                ROOT
                 |
                 S
        _____|_____
       |         |                  VP
       |         |          _____|_____
       |         |         |                         NP
       |         |         |                 _____|_____
       |         |         |                |                NP
       |         |         |                |         _____|____
       |         |         |                |        |        SBAR
       |         |         |                |        |     ____|____
       |         |         |                |        |    |         S
       |         |         |                |        |    |         |
       |         |         |                |        |    |         VP
       |         |         |                |        |    |     ____|____
       |         |         |                |        |    |    |       VP
       |         |         |                |        |    |    |    ___|____
       NP        |         NP             CONJP     NP  WHNP   |   |      NP
       |         |      ___|___          ___|___     |    |    |   |    __|___
       EX       VBD    JJ     NNS       CC     RB    CD   WP   MD  VB  PRP$    NN
       |         |      |      |         |      |    |    |    |   |    |       |
     There     were  several Mortimers but    only  one  who could be  our   visitor
```
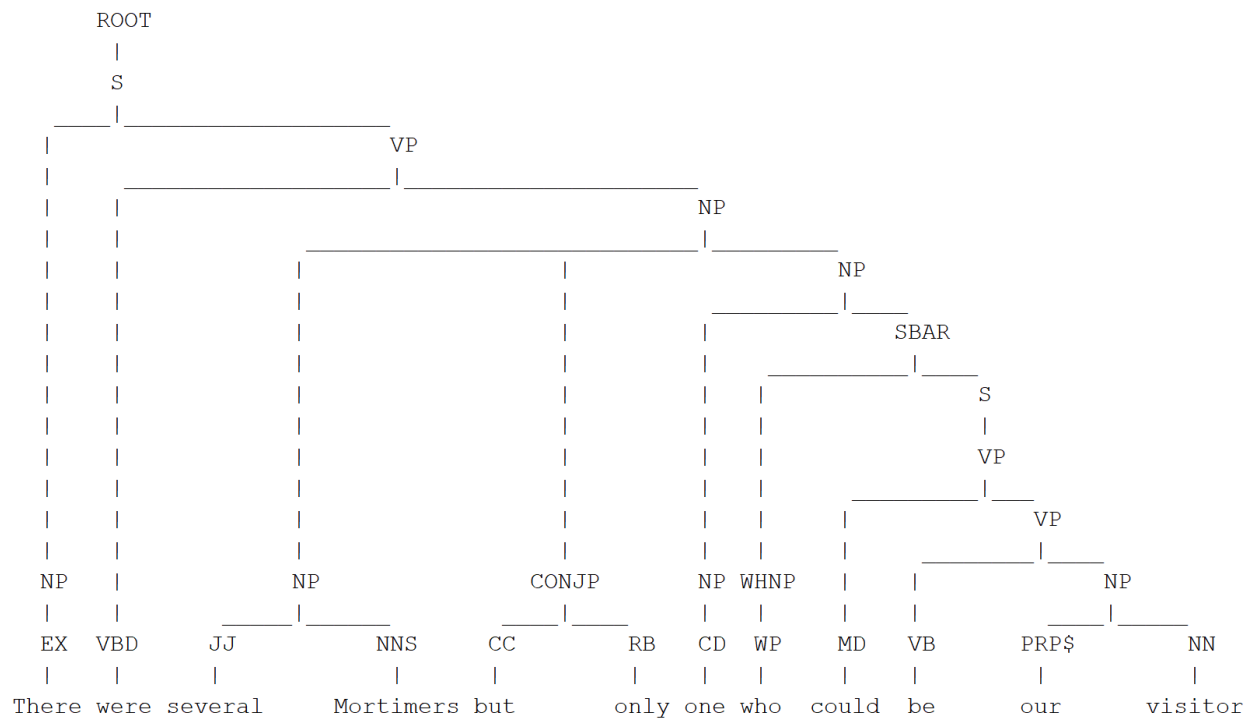
CFG:


```
S → NP | VP
VP → VBD | NP | MD | VP | VB
NP → N P | CONJP | SBAR | EX | JJ | NNS | CD | PRP$ | NN
SBAR → WHNP | S
CONJP → CC | RB
WHNP → WP
EX → There
VBD → were
JJ → several
NNS → Mortimers
CC → but
RB → only
CD → one
WP → who
MD → could
VB → be
PRP$ → our
NN → visitor
```

Parse Trees for WarAndPeace.txt

1.Our gracious sovereign recognizes his high vocation and will be true to it.

```
                                                    ROOT
                                                     |
                                                    FRAG
                                                     |
                                                     NP
                                           _____|_____
          |                                          |      SBAR
          |                                          |       |
          |                                          |       S
          |                                          |       |
          |                                          |       VP
          |                                  _____|_____
          |                                |         |              VP
          |                                |         |         _____|____
          |                                |         |        |         VP
          |                                |         |        |    _____|___
          |                                |         |        |   |       ADJP
          |                                |         |        |   |    ____|____
          |                          VP    |         |        |   |   |        PP
          |                         _|___   |        |        |   |   |     ___|__
          NP                       |    NP  |        |        |   |   |    |      NP
        __|_____          _____|____   |    |    |    |   |   |    |      |
     PRP$    JJ      JJ    VBZ   PRP$  JJ   NN  CC   MD   VB  JJ   TO     PRP
       |     |       |      |     |    |    |    |    |    |   |    |       |
      Our gracious sovereign recognizes his high vocation and will be true to      it
```

CFG:


```
FRAG → NP
NP → NP|SBAR |PRP$ |JJ      NN |PRP
SBAR → S
S → VP
VP → VP|CC |ADJP |VBZ |NP|MD |VB
ADJP → JJ|PP
PP → TO|NP
PRP$ → Our |his
JJ → gracious |sovereign |high |true
VBZ → recognizes
NN → vocation
CC → and
MD → will
VB → be
TO → to
PRP → it
```

2. She wanted to find, and still seeks, some secret motive in our actions.
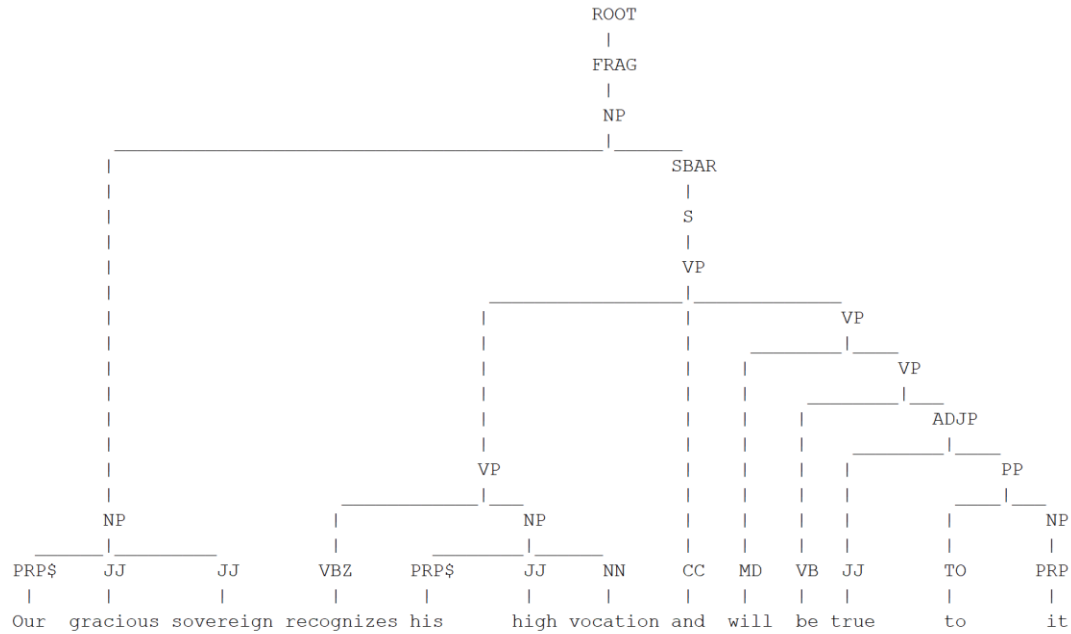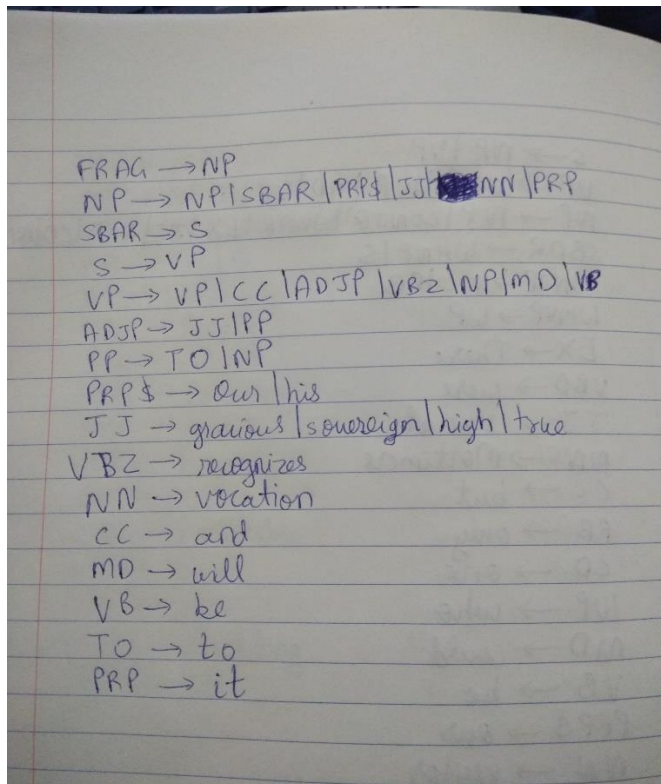
```
                                      ROOT
                                       |
                                       S
                 _____|_____
                |                               VP
                |                                |
                |           VP                   |_____
                |          __|___                |                VP
                |         |      S                |              __|_____
                |         |      |                |   |   |            NP
                |         |      VP                |   |   |          __|_____
                NP        |    __|___              |   |   |         |            PP
                |   |     |   |     VP  |  ADVP     |          NP    |          __|___
               PRP  VBD   TO  |     VB  CC   RB   VBZ   DT     JJ     NN   IN  PRP$        NP
                |    |     |   |      |   |    |     |    |      |     |     |    |          |
               She wanted to      find and still seeks some secret motive  in  our      actions
```
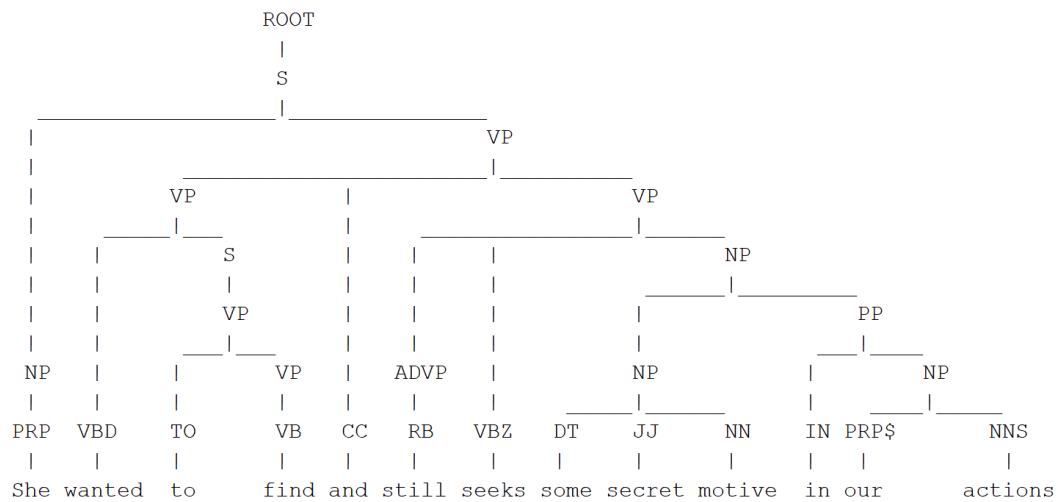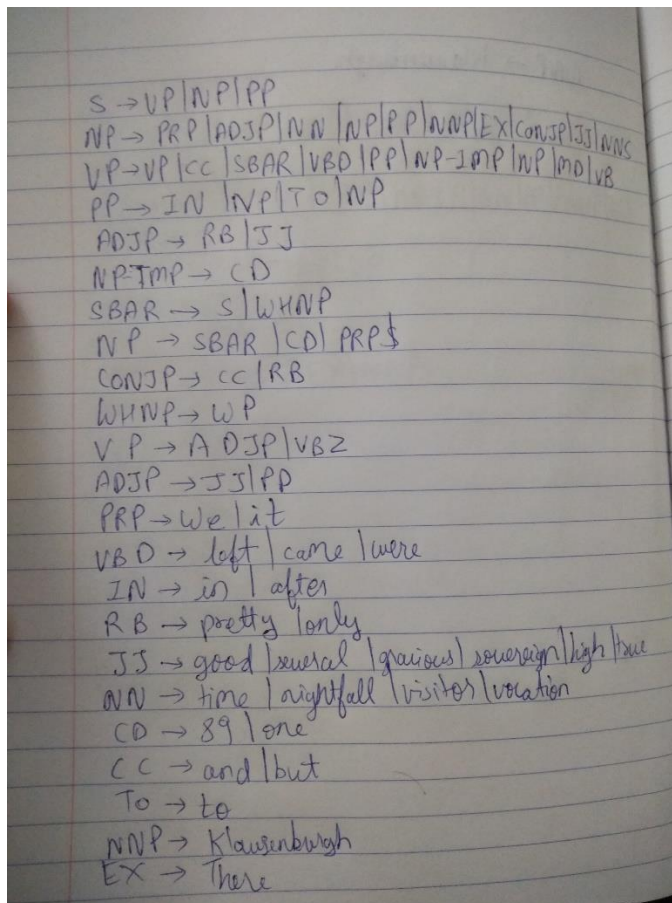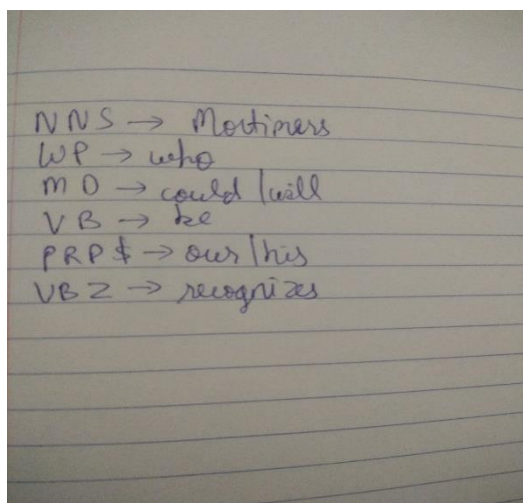
CFG:

S ⟶ NP | VP
VP ⟶ VP | VBD | S | CC | ADVP | VBZ | NP | TO | VB
NP ⟶ NP | PP | PRP | DT | JJ | NN | PRP$ | NNS
PP ⟶ IN | NP
ADVP ⟶ RB
PRP ⟶ she
VBD ⟶ wanted
TO ⟶ to
VB ⟶ find
CC ⟶ and
RB ⟶ still
VBZ ⟶ seeks
DT ⟶ some
JJ ⟶ secret
NN ⟶ motive
IN ⟶ in
PRP$ ⟶ our
NNS ⟶ actions

The corpus of three sentences from each .txt file:

We left in pretty good time, 89and came after nightfall to Klausenburgh. There were several Mortimers, but only one who could be our   visitor. Our gracious sovereign recognizes his high vocation and will be true to it.

(Note: The POS tag parse tree for this corpus is present is Parse_tree.txt file and can be seen in notepad++)

```
S → VP | NP | PP
NP → PRP | ADJP | NN | NP PP | NNP | EX | CONJP | JJ | NNS
VP → VP | CC | SBAR | VBD | PP | NP-TMP | NP | MD | VB
PP → IN | NP | TO | NP
ADJP → RB | JJ
NP-TMP → CD
SBAR → S | WHNP
NP → SBAR | CD | PRP$
CONJP → CC | RB
WHNP → WP
VP → ADJP | VBZ
ADJP → JJ | PP
PRP → We | it
VBD → left | came | were
IN → in | after
RB → pretty | only
JJ → good | several | gracious | sovereign | high | true
NN → time | nightfall | visitor | vocation
CD → 89 | one
CC → and | but
TO → to
NNP → Klausenburgh
EX → There
```

```
NNS → Mortimers
WP → who
MD → could |will
VB → be
PRP$ → our |his
VBZ → recognizes
```

In all the parse trees above it can be seen that in most cases Verb phrase is producing verb phrase, noun phrase produces possessive pronoun, noun and adjective, start variable produces noun phrase or verb phrase and determiner, adjective and noun are in a sequence.

Question-3:

Solution:

1. Text: We left in pretty good time, 89 and came after nightfall to Klausenburgh.

   Hypothesis1: We went to Klausenburgh

   Hypothesis2: We left at the right time.

   Hypothesis3: We came in the evening

2. Text: Do you know where you are going, and what you are going to?"

   Hypothesis1: Where are you going to?

   Hypothesis2: What are you going to?

   Hypothesis3: Do you know where and what you are going to?

3. Text: Some people without possessing genius have a remarkable power of stimulating it.
   Hypothesis1: Some people are not genius

   Hypothesis2: Some people can stimulate power to become genius

   Hypothesis3: Some people have stimulating power

4. Text: There were several Mortimers, but only one who could be our   visitor.

   Hypothesis1: There are many Mortimers

   Hypothesis2: There is only one visitor

   Hypothesis3: A visitor is a mortminer

5. Text: Our gracious sovereign recognizes his high vocation and will be true to it.

   Hypothesis1: The sovereign is gracious

   Hypothesis2: The sovereign will be true to vocation

   Hypothesis3: The sovereign has high vocation

6. Text: She wanted to find, and still seeks, some secret motive in our actions.

   Hypothesis1: She wants to find secret motives

   Hypothesis2: She wants to seek secret motives

   Hypothesis3: She wants to find and seek secret motives