COMP 6751 Sentence Sentiment Analysis Project 3 Report

Haochen Zou (Luke) (Student ID: 40158179)

0. Expectations of originality

I certify that this submission is my original work and meets the Faculty's Expectations of Originality.
Name: Haochen Zou; I.D: 40158179; Date: 2021.11.06

1. Design Justification

1.1. Sentiment feature

In the feature-based grammar, we have introduced a new notation that categories have **feature** called **SENTIMENT**. The value has three features: **positive**, **negative**, and **neutral**. We are using **?s** as a variable over values of **SENTIMENT**. It can be instantiated either to **positive**, **negative**, or **neutral**, within a given production. The example of sentiment feature in feature grammar is displayed as follows:

```
>>S[-INV, SENTIMENT=?s] -> NP[-WH, NUM=?n, PERSON=?p, SENTIMENT=?s] VP[TENSE=?t, NUM=?n, PERSON=?p,
   SENTIMENT=?s]
>>S[-INV, SENTIMENT=negative] -> S[-INV, SENTIMENT=positive] CC[CONJUNCTION=but] S[-INV,
   SENTIMENT=negative] | S[-INV, SENTIMENT=positive] CC[CONJUNCTION=but] S[-INV, SENTIMENT=neutral]
```

1.2. Conjunction of sentiment bearing adjective, nouns, and sentence

We have implemented grammar for adjective phrases, noun phrases, and sentences which contain the conjunction of sentiment: **and; or**; **but**. We improved the label **CC** to enable it show what specific conjunction it contains, for example: **CC[CONJUNCTION=and]**. The example of grammar for conjunction is displayed:

```
>>S[-INV, SENTIMENT=neutral] -> S[-INV, SENTIMENT=neutral] CC[CONJUNCTION=and] S[-INV,
   SENTIMENT=neutral]
>>NP[-WH, NUM=?n, PERSON=?p, SENTIMENT=?s] -> NP[-WH, NUM=?n, PERSON=?p, SENTIMENT=?s].
   CC[CONJUNCTION=and] NP[-WH, NUM=?n, PERSON=?p, SENTIMENT=?s] | NP[-WH, NUM=?n, PERSON=?p,
   SENTIMENT=neutral] CC[CONJUNCTION=and] NP[-WH, NUM=?n, PERSON=?p, SENTIMENT=?s] | NP[-WH,
   NUM=?n, PERSON=?p, SENTIMENT=?s] CC[CONJUNCTION=and] NP[-WH, NUM=?n, PERSON=?p,
   SENTIMENT=neutral]
>>ADJP[SENTIMENT=?s] -> ADJ[SENTIMENT=?s] CC[CONJUNCTION=and] ADJ[SENTIMENT=?s] |
   ADJ[SENTIMENT=?s] CC[CONJUNCTION=and] ADJP[SENTIMENT=?s]
```

1.3. Negation feature

We have introduced feature **negation** to adverbs. For those adverbs with negative meanings, we labeled: **ADV[+NEGATION]**. For the adverbs without negative meanings, we labeled **ADV[-NEGATION]**. Notice that some adverbs does not have negation feature but have negative sentiment value, for example:

```
>>ADV[-NEGATION, SENTIMENT=negative] -> 'shamelessly' | 'Shamelessly'
```

1.4. Wide coverage

We have developed a grammar with wide coverage and less acceptance of ungrammatical information. To ensure wide coverage, we have introduced new categories and features to the featured-based grammar:

```
>>S: Sentence; -INV: Not Inverted Clauses; +INV: Inverted Clauses
>>NP: Noun Phrase; -WH: Noun Phrase without Question Pronoun; +WH: Noun Phrase with Question Pronoun;
   NUM=?n/ sg/ pl: Noun Phrase is Singular, Plural or Either; PERSON=?p/1/2/3: Person 1, 2, 3 or Either
>>VP: Verb Phrase; TENSE=?t/pres/past: Verb Phrase Tense is Present, Past or Either
>>PP: Prepositional Phrase
>>SBAR: Subordinate Clause
>>ADJP: Adjective Phrase
>>DET: Determinator
>>N: Noun
>>V: Verb; +AUX: Auxiliary Verb; -AUX: Non-Auxiliary Verb; SUBCAT=intrans/trans/clause: Subcategorization:
   Transitive Verb, Intransitive Verb, and Clause Verb
>>PRP: Possessive Pronoun
>>ADJ: Adjective
>>ADV:  Adverb
>>PREP: Preposition
>>CC: Coordinating Conjunction
>>PUNCT: Punctuation
```

2. Design Critique

We have designed 17 sentences to test the program. The 17 sentences are guided by data provided in the project requirement. The 17 sentences will be first process sentiment analysis by the feature grammars and the opinion lexicon in NLTK. Specifically, we use **demo_liu_hu_lexicon (**Basic example of sentiment classification using Liu and Hu opinion lexicon. This function simply counts the number of positive, negative, and neutral words in the sentence and classifies it depending on which polarity is represented**)** to pre-analysis the sentiment value of the 17 sentences, and separate them into three files: **positive**, **negative**, and **neutral** according to the sentiment value results. The pre-analysis demo displayed below:

```
>>from nltk.sentiment.util import demo_liu_hu_lexicon
  demo_liu_hu_lexicon('This compelling story with gut-wrenching impact.')
>>Result: Neutral
```

Then we parse the 17 sentences by the project we designed with sentiment features, generate the sentiment value of the sentence, and compare the sentiment value with the results above. If the sentiment values of these two are the same, then we record the results and the parse tree in the **Good** file, else we write them in the **False** file. There are 2 sentences contain different sentiment values, the parse and sentiment values results shown below. So, in this part we mainly focus on analyze the different reasons.

```
【Sentence for Analysis】
It is a compelling story, but it has low impact.
【Initial Forecast Sentiment Value】
neutral
【Program Analysis Sentiment Value】
negative
【Earley Parse Result】
(S[-INV, SENTIMENT='negative']
  (S[-INV, SENTIMENT='positive']
    (NP[NUM='sg', PERSON=3, SENTIMENT=?s, -WH]
      (PRP[NUM='sg', PERSON=3] It))
    (VP[NUM='sg', PERSON=3, SENTIMENT='positive', TENSE='pres']
      (V[+AUX, NUM='sg', PERSON=3, TENSE='pres'] is)
      (NP[NUM='sg', PERSON=?p, SENTIMENT=?s, -WH] (DET[NUM='sg'] a))
      (NP[NUM='sg', PERSON=?p, SENTIMENT='positive', -WH]
        (ADJ[SENTIMENT='positive'] compelling)
        (NP[NUM='sg', PERSON=?p, SENTIMENT='neutral', -WH]
          (N[NUM='sg', SENTIMENT='neutral'] story)))))
  (PUNCT[] ,)
  (CC[CONJUNCTION='but'] but)
  (S[-INV, SENTIMENT='neutral']
    (NP[NUM='sg', PERSON=3, SENTIMENT=?s, -WH]
      (PRP[NUM='sg', PERSON=3] it))
    (VP[NUM='sg', PERSON=3, SENTIMENT='neutral', TENSE='pres']
      (V[-AUX, NUM='sg', PERSON=3, SUBCAT='trans', TENSE='pres'] has)
      (NP[NUM='sg', PERSON=?p, SENTIMENT='neutral', -WH]
        (ADJ[SENTIMENT='neutral'] low)
        (NP[NUM='sg', PERSON=?p, SENTIMENT='neutral', -WH]
          (N[NUM='sg', SENTIMENT='neutral'] impact))))))
```

```
【Sentence for Analysis】
It has gut-wrenching impact, and it is a compelling story.
【Initial Forecast Sentiment Value】
neutral
【Program Analysis Sentiment Value】
positive
【Earley Parse Result】
(S[-INV, SENTIMENT='positive']
  (S[-INV, SENTIMENT='neutral']
    (NP[NUM='sg', PERSON=3, SENTIMENT=?s, -WH]
      (PRP[NUM='sg', PERSON=3] It))
    (VP[NUM='sg', PERSON=3, SENTIMENT='neutral', TENSE='pres']
      (V[-AUX, NUM='sg', PERSON=3, SUBCAT='trans', TENSE='pres'] has)
      (NP[NUM='sg', PERSON=?p, SENTIMENT='neutral', -WH]
        (ADJ[SENTIMENT='neutral'] gut-wrenching)
        (NP[NUM='sg', PERSON=?p, SENTIMENT='neutral', -WH]
          (N[NUM='sg', SENTIMENT='neutral'] impact)))))
  (PUNCT[] ,)
  (CC[CONJUNCTION='and'] and)
  (S[-INV, SENTIMENT='positive']
    (NP[NUM='sg', PERSON=3, SENTIMENT=?s, -WH]
      (PRP[NUM='sg', PERSON=3] it))
    (VP[NUM='sg', PERSON=3, SENTIMENT='positive', TENSE='pres']
      (V[+AUX, NUM='sg', PERSON=3, TENSE='pres'] is)
      (NP[NUM='sg', PERSON=?p, SENTIMENT=?s, -WH] (DET[NUM='sg'] a))
      (NP[NUM='sg', PERSON=?p, SENTIMENT='positive', -WH]
        (ADJ[SENTIMENT='positive'] compelling)
        (NP[NUM='sg', PERSON=?p, SENTIMENT='neutral', -WH]
          (N[NUM='sg', SENTIMENT='neutral'] story))))))
```

We can discover the main reason for the difference of sentiment values is that these two sentences both contain **conjunction**, and the sentiment value before and after the conjunction in the different. For example, in the first sentence, the sentence is constructed by two small sentences with conjunction but. Before the conjunction but, the sentiment value of the small sentence is positive, and after the sentiment value is neutral. From the featured grammar we designed, we defined that the sentence with a positive sentence followed by punctuation followed by conjunction but followed by a neutral sentence show a negative sentiment value:

```
S[-INV, SENTIMENT=negative] -> S[-INV, SENTIMENT=positive] PUNCT CC[CONJUNCTION=but] S[-INV,
SENTIMENT=negative] | S[-INV, SENTIMENT=positive] PUNCT CC[CONJUNCTION=but] S[-INV, SENTIMENT=neutral]
```

For example: The house is amazing but way too example, which is contradiction to the initial forecast sentiment value from NLTK feature grammar and opinion lexicon.

In the second different sentiment values sentence, the reason for a diverse conclusion is the same as above: when a sentence has a form that a small sentence followed by conjunction followed by another small sentence and the two small sentences have different sentiment value, the total sentiment value result may be unable to reach an agreement. From the record, we can find that in the sentence, the front sentence has a neutral sentiment value, the back sentence has a positive sentiment value. In the grammar we developed, the whole sentence reveals a positive sentiment value, for example: The building is tall and beautiful, which is contradiction to the initial forecast sentiment value from NLTK feature grammar and opinion lexicon.

```
S[-INV, SENTIMENT=positive] -> S[-INV, SENTIMENT=positive] PUNCT CC[CONJUNCTION=and] S[-INV,
SENTIMENT=positive] | S[-INV, SENTIMENT=positive] PUNCT CC[CONJUNCTION=and] S[-INV, SENTIMENT=neutral]
| S[-INV, SENTIMENT=neutral] PUNCT CC[CONJUNCTION=and] S[-INV, SENTIMENT=positive]
S[-INV, SENTIMENT=negative] -> S[-INV, SENTIMENT=negative] PUNCT CC[CONJUNCTION=and] S[-INV,
SENTIMENT=negative] | S[-INV, SENTIMENT=negative] PUNCT CC[CONJUNCTION=and] S[-INV,
SENTIMENT=neutral] | S[-INV, SENTIMENT=neutral] PUNCT CC[CONJUNCTION=and] S[-INV, SENTIMENT=negative]
S[-INV, SENTIMENT=neutral] -> S[-INV, SENTIMENT=neutral] PUNCT CC[CONJUNCTION=and] S[-INV,
SENTIMENT=neutral]
```

3. Grammar Semantics

The feature-based sentiment analysis grammar we develop cover the semantics below:

(1) A compelling story. ***(Noun phrase)***

S[-INV, SENTIMENT=?s] -> NP[-WH, NUM=?n, PERSON=?p, SENTIMENT=?s]

(2) It is a compelling story. ***(Noun phrase followed by verb phrase)***

S[-INV, SENTIMENT=?s] -> NP[-WH, NUM=?n, PERSON=?p, SENTIMENT=?s] VP[TENSE=?t, NUM=?n, PERSON=?p, SENTIMENT=?s]

(3) Is a story compelling. ***(Inverted sentence: auxiliary verb followed by noun phrase followed by adjective phrase)***

S[+INV, SENTIMENT=?s] -> V[+AUX] NP[-WH, NUM=?n] ADJP[SENTIMENT=?s]

(4) AND Conjunction:

It has gut-wrenching impact and it is a compelling story.

(Sentence followed by and conjunction followed by sentence)

S[-INV, SENTIMENT=**positive**] -> S[-INV, SENTIMENT=**positive**] CC[CONJUNCTION=and] S[-INV, SENTIMENT=**positive**] | S[-INV, SENTIMENT=**positive**] CC[CONJUNCTION=and] S[-INV, SENTIMENT=**neutral**] | S[-INV, SENTIMENT=**neutral**] CC[CONJUNCTION=and] S[-INV, SENTIMENT=**positive**]

S[-INV, SENTIMENT=**negative**] -> S[-INV, SENTIMENT=**negative**] CC[CONJUNCTION=and] S[-INV, SENTIMENT=**negative**] | S[-INV, SENTIMENT=**negative**] CC[CONJUNCTION=and] S[-INV, SENTIMENT=**neutral**] | S[-INV, SENTIMENT=**neutral**] CC[CONJUNCTION=and] S[-INV, SENTIMENT=**negative**]

S[-INV, SENTIMENT=**neutral**] -> S[-INV, SENTIMENT=**neutral**] CC[CONJUNCTION=and] S[-INV, SENTIMENT=**neutral**]

It has gut-wrenching impact, and it is a compelling story.

(Sentence followed by punctuation followed by and conjunction followed by sentence)

S[-INV, SENTIMENT=**positive**] -> S[-INV, SENTIMENT=**positive**] PUNCT CC[CONJUNCTION=and] S[-INV, SENTIMENT=**positive**] | S[-INV, SENTIMENT=**positive**] PUNCT CC[CONJUNCTION=and] S[-INV, SENTIMENT=**neutral**] | S[-INV, SENTIMENT=**neutral**] PUNCT CC[CONJUNCTION=and] S[-INV, SENTIMENT=**positive**]

S[-INV, SENTIMENT=**negative**] -> S[-INV, SENTIMENT=**negative**] PUNCT CC[CONJUNCTION=and] S[-INV, SENTIMENT=**negative**] | S[-INV, SENTIMENT=**negative**] PUNCT CC[CONJUNCTION=and] S[-INV, SENTIMENT=**neutral**] | S[-INV, SENTIMENT=**neutral**] PUNCT CC[CONJUNCTION=and] S[-INV, SENTIMENT=**negative**]

S[-INV, SENTIMENT=**neutral**] -> S[-INV, SENTIMENT=**neutral**] PUNCT CC[CONJUNCTION=and] S[-INV, SENTIMENT=**neutral**]

(5) OR Conjunction

It has gut-wrenching impact or it is a compelling story.

S[-INV, SENTIMENT=?s] -> S[-INV, SENTIMENT=?s] CC[CONJUNCTION=or] S[-INV, SENTIMENT=?s]

S[-INV, SENTIMENT=**neutral**] -> S[-INV, SENTIMENT=**positive**] CC[CONJUNCTION=or] S[-INV, SENTIMENT=**negative**] | S[-INV, SENTIMENT=**negative**] CC[CONJUNCTION=or] S[-INV, SENTIMENT=**positive**]

(6) **BUT Conjunction**

It does not have gut-wrenching impact but it is a compelling story.

S[-INV, SENTIMENT=?s] -> S[-INV, SENTIMENT=**neutral**] CC[CONJUNCTION=but] S[-INV, SENTIMENT=?s]

It has low impact but it is a compelling story.

S[-INV, SENTIMENT=**negative**] -> S[-INV, SENTIMENT=**positive**] CC[CONJUNCTION=but] S[-INV, SENTIMENT=**negative**] | S[-INV, SENTIMENT=**positive**] CC[CONJUNCTION=but] S[-INV, SENTIMENT=**neutral**]

S[-INV, SENTIMENT=**positive**] -> S[-INV, SENTIMENT=**negative**] CC[CONJUNCTION=but] S[-INV, SENTIMENT=**positive**] | S[-INV, SENTIMENT=**negative**] CC[CONJUNCTION=but] S[-INV, SENTIMENT=**neutral**]

It is a compelling story, but it has low impact.

S[-INV, SENTIMENT=**negative**] -> S[-INV, SENTIMENT=**positive**] PUNCT CC[CONJUNCTION=but] S[-INV, SENTIMENT=**negative**] | S[-INV, SENTIMENT=**positive**] PUNCT CC[CONJUNCTION=but] S[-INV, SENTIMENT=**neutral**]

S[-INV, SENTIMENT=**positive**] -> S[-INV, SENTIMENT=**negative**] PUNCT CC[CONJUNCTION=but] S[-INV, SENTIMENT=**positive**] | S[-INV, SENTIMENT=**negative**] PUNCT CC[CONJUNCTION=but] S[-INV, SENTIMENT=**neutral**]