HW\_7

Gabriel Wang

The Sentiment Analysis via a Stanford Core NLP is based on an ‘enhanced lexical resources’ with high accuracy called the ‘SENTIWORDNET 3.0’ (Baccianella, et al). PC-ACE adpoted a scale level of 0 to 4, 0 is very negative, 4 is very positive, and 2 is neutral.

First, the pie chart output from PC-ACE gave a general sentiment tendency of the whole corpus of Song of Ice and Fire. The result is reported as follow:

Unfortunately buy not suprisingly, Song of Ice and Fire is definitely a tragedy story where there are only 197 very positve story out of this over-140,000-sentence masterpiece.

Sentiment analysis is also largely used to quantitively display the shape of story. According to Reagan and his colleagues’ study, there are six types of stories in total, and their shapes of story can be categorized from the change of sentiment of different segments. My method to analyze the shape of Song of Ice and Fire follows closely the route described in Reagan’s journal by dividing the whole corpus to segments of 1000 sentences. After averaging the every 1000 sentences, this miserable data showed that there isn’t one single segment even reach the ‘neutral’ lable. What a pity. So this plot is how realative this tragedy story ‘fall and rise’ (Reagan et al). The shape of the Song of Ice and Fire goes as follow:

The whole story is a really ‘bumpy’ ride. The low appeared around 77th segment, that is sentence 77,000 to 78,000, which is the story right after the notorious ‘Red Wedding’ where the whole Northern Army Alliance was slaughtered during the wedding of Edmure Tully and Roslin Frey. George R.R Martin surely superised all his readers and the viewers with this. The high was at the 97th segment, sentence 97,000 to 98,000, when Samwell Tarley, a chubby Night Watch lost his virginity to Gilly, a wildling girl. A conspiracy theorist of Song of Ice and Fire may actually agree with this as they believe that Samwell Tarley is the projection of George Martin himself in the story. However, even this high point only has a 1.762 sentiment number which is below neutral. Take a close look into the story, Sam sadly made love to Gilly in a tiny chamber on a boat and both of them were sick and vomitting. It was pretty sad.

Let’ examine more closely the local lowest point. Except for around the first 10 segments, almost all other low points experienced a sudden drop with a visbile downward gradient. It is well known that sometimes George Martin can get too dramatic with the plot and he is definitely cruel to his reader by using the ‘give you hope and make you desperate’ technique.

Despite the negaitve keynote, I was hopping to find some positive perspective from those roles that are commonly associated with postive words. So I picked several different charaters set, and did query on them.

* High-Born-Female: ‘Queen, Princes, Lady etc.’
* High-Born-Male: ‘King, Prince ,Lord, Knight, Master etc.’
* Other-Femal: ‘Girl, Woman, Mother, Sister, Daughter etc.’
* Other-Male: ‘Boy, Man, Father, Brother, Son etc.’
* Abnormal: ‘Eunuch, Cripple, Coward etc.’

Well, unfortunately, there is no differnece among these different set of characters. The result is reported as follow:

It seems that the author doesn’t like to mark a certain kind of characters as good or evil, he even used the highest percentage of ‘very positive’ words on the ‘eunuch and cripple.’

Trying to find which kind of characters the author favored was a dead end, how about which family. So, I repeated the same routine on the two major rival families, House Stark and House Lannister. The result was yet again disappointing as George Martin still showed no tendency.

The conclusion of the sentiment analysis of the Song of Ice and Fire revealed a lot of information. George Martin wrote the whole story on a tragic keytone, and he loves to make the characters suffer, regardless of gender, family or breed.

The end of the fifth book is experiencing a slide, so we can probably looking toward another major character’s death at the beginnign of the sixth book.