Analysis of Antithetical Elements in English Literary Passages Using Stochastic Models

Chenjie Zeng

School of Humanities and Social Science

The Chinese University of Hong Kong, Shenzhen

E-mail: chenjiezeng@link.cuhk.edu.cn

Clement H. C. Leung

School of Science and Engineering

The Chinese University of Hong Kong, Shenzhen

E-mail: clementleung@cuhk.edu.cn

Content

Abstract

- Thesis and antithesis are rhetorical figures often employed by well-known authors in English literary passages. Antithesis refers to the juxtaposition of contrasting words or ideas, which often, although not always, in parallel structure. Such rhetorical figure is used to contrast opposing ideas, creating a sense of tension and urgency, as well as heightening the emotional impact of the speech. Contrasting the antithesis is the thesis, which refers to a statement, assertion, or tenet. A thesis also is a proposition laid down or stated, especially as a theme to be discussed and proved, or to be maintained against attack. It is observed that the points of occurrence of thesis and antithesis are random, yet they occur with remarkable regularity alternating with each other. In this study, we represent the lengths of thesis and antithesis by using an alternating renewal stochastic process. We find that the underlying parameters of the alternating renewal process are able to usefully characterize the writing style of individual authors and help us to quantify and understand their linguistic technique and intention. Using the present method in conjunction with other techniques, such as sentiment analysis, and word embedding models, it is possible to gain a deeper understanding of the literature and its underlying themes and structures.

I Thesis & Antithesis

- **Antithesis** refers to the juxtaposition of contrasting words or ideas, which often, although not always, in parallel structure. For example, Shakespeare's Sonnet says, 'Before, a joy proposed; behind, a dream'. This figure is used to contrast opposing ideas and create a sense of tension in "Henry V".

- The contrast of antithesis is **thesis**. According to Oxford English Dictionary, a "thesis" refers to a statement, assertion, or tenet. Thesis also is a proposition laid down or stated, especially as a theme to be discussed and proved, or to be maintained against attack.

I Thesis & Antithesis in Literary Passages

The formal connection between these figures demonstrates that despite their discrete, distinct, or potentially conflicting natures, they have the capacity and necessity to contribute to the larger entity of creation or society, existing alongside each other in a meaningful manner. For example, when Henry says (in stanza 1, Act 4 Scene 3, *Henry V*),

X   2   3   4

If we are marked to **die**, we are enough

5   6 7   8   9   10 11 12 -X

To do our country loss; and if to **live**,

X   2   3   -X

The **fewer** men, the **greater** share of honour

I Thesis & Antithesis in Literary Passages

- We use "X" to represent the first occurrence of rhetoric "die/fewer," and use "-X" to represent the follow-up antithesis "live/greater". Given the condition that an "X" appears in a poetic line, it is expected to have an "-X" sooner or later.

- In the above lines that contain antithesis in Henry's speech, when we come across "X," such as "die" in the third line, we expect to have an antithesis "-X" like "live" later. Likewise, when we have the word "fewer", which is an "X". We will have its antonym "greater", which can be symbolized as "-X".

- Another example would be in Stanza 7 of Henry's speech to his men before the Battle of Agincourt in *Henry V*.

II Modelling Thesis and Antithesis Using Alternating Renewal Processes

- The alternating recurrence of Thesis and Antithesis can be modelled as an alternating renewal stochastic process. Upon the occurrence of the thesis, a certain length of words (Xi) elapsed before the occurrence of the antithesis; for such words, we shall refer to them as being under the dominance of the thesis. By focusing on such rhetorical devices, we have the following model of a passage.

$$S = X_1 + Y_1 + X_2 + Y_2 + \cdots + \cdots X_N + Y_N$$

$$= \sum_{k=1}^{N} X_k + \sum_{k=1}^{N} Y_k$$

II Modelling Thesis and Antithesis Using Alternating Renewal Processes

- given a thesis occurs, then the thesis length has a probability distribution

$$\Pr[\, X_i = n\,] = qp^{n-1} \qquad n = 1, 2, \dots$$

- This can be seen as follows. Given the occurrence of a thesis, then it must consist of at least one positive word (the first word of the thesis), for otherwise, it would not be a thesis. For long passages, the number of words is large and may be mathematically approximated by infinity. Following the same reasoning, we also have

$$\Pr[\, Y_i = n\,] = pq^{n-1} \qquad n = 1, 2, \dots$$

II Modelling Thesis and Antithesis Using Alternating Renewal Processes

- The probability generating function F (z) of Xi is therefore given by

$$F(z) = \sum_{k=0}^{\infty} \Pr[X_i = k] z^k = \frac{q}{p} \sum_{k=0}^{\infty} p^k z^k$$
$$= \frac{q}{p(1-pz)}.$$

- The probability generating function G (z) of Yi is given by

$$G(z) = \sum_{k=0}^{\infty} \Pr[Y_i = k] z^k = \frac{p}{q} \sum_{k=0}^{\infty} q^k z^k$$
$$= \frac{p}{q(1-qz)}.$$

II Modelling Thesis and Antithesis Using Alternating Renewal Processes

- The mean thesis length is obtained by differentiation

$$E(X_i) = F'(1) = \frac{1}{q},$$

- and similarly, the mean antithesis length is

$$E(Y_i) = G'(1) = \frac{1}{p}.$$

- The variance of the thesis length and the variance of the antithesis length are as follows:

$$\mathrm{Var}(X_i) = F''(1) + F'(1) - F'(1)^2 = \frac{p}{q^2},$$

$$\mathrm{Var}(Y_i) = G''(1) + G'(1) - G'(1)^2 = \frac{q}{p^2}.$$

II Modelling Thesis and Antithesis Using Alternating Renewal Processes

- Another common way of characterizing a passage is determining the total length L of words in relation to a given length r of the thesis. Now, L=k (≥r), iff the kth word encountered coincides with the rth positive word, and this happens with probability

$$\Pr\left[L = k\right] = \binom{k-1}{r-1}p^r q^{k-r}$$

- for k = r, r+1, r+2, …. This can be seen by noting that in order for L= k to be true, we must have (r—1) positive words among the first (k—1) words encountered, and this has the binomial distribution

$$\binom{k-1}{r-1}p^{r-1}q^{k-r}$$

II Modelling Thesis and Antithesis Using Alternating Renewal Processes

- On multiplying this by the probability that the kth word is positive, we obtain Pr[L = k]. The mean value of L is given by

$$\mathrm{E}[L] = r\left(1 + \frac{q}{p}\right).$$

- On identifying 1/q as the mean thesis length and that, similarly, 1/p as the mean antithesis length, we obtain the following relation

$$\mathrm{E}[L] = r\left(1 + \frac{\text{mean length of antithesis}}{\text{mean length of thesis}}\right).$$

- Or in normalizing by r, we have

$$\frac{\mathrm{E}[L]}{r} = 1 + \frac{\text{mean length of antithesis}}{\text{mean length of thesis}},$$

- which can be utilized to characterize the style of a passage.

II Modelling Thesis and Antithesis Using Alternating Renewal Processes

- Another useful characterization is represented by the renewal function, which in the present context is the number of distinct thesis episodes N within r positive words. The r positive words have at most (r—1) gaps among them, since between any two successive positive words, there may or may not be intervening negative words. The probability of having no such intervening words is of course p and that of having one or more such words is q. Hence, the number of distinct episodes of the thesis equals k iff there are (k—1) interventions among the (r—1) gaps, i.e.,

$$\Pr\left[N = k\right] = \binom{r-1}{k-1} q^{k-1} p^{r-k}$$

II Modelling Thesis and Antithesis Using Alternating Renewal Processes

- where 1 ≤k ≤r. This has mean

$$E[N] = 1 + (r-1)q.$$

- If r is large and q small, then we can use the Poisson approximation

$$\Pr[N = k] \cong \frac{[(r-1)q]^{k-1}e^{-(r-1)q}}{(k-1)!}.$$

III Experiments & Illustrations

(1) Russian Novel

- The first chapter of Leo Tolstoy's *War and Peace*

- When Anna Pavlovna talked about family members, Prince Vasili said,

<pre>
                                    X   2
Hippolyte is at least a **quiet** fool,
  3   4   5 6   -X
but Anatole is an **active** one.

He said this smiling in a way
                          X   2   3
more natural and **animated** than usual,
  4   5   6   7   8   9   10
so that the wrinkles round his mouth
 11   12   13   14
very clearly revealed something
 15   16   17   -X
unexpectedly coarse and **unpleasant**.
</pre>

III Experiments & Illustrations

(1) Russian Novel

- The first "X" is "quiet" and its follow-up "-X" is "active". The length between "quiet" and "active" is 6 words. Then we have the second pair of "X" (animated) and "-X" (unpleasant), and the length is 17 words. The average length of these two sentences is (6+17)/2=11.5 words. Compared with Shakespeare's "Henry V," Leo Tolstoy's "War and Peace" has a shorter length of antithesis. To summarize, we have:

- $E(X_i) = 11.5$
- $E(Y_i) = 12$
- For $r = 10$, $E(L) = 21$

III Experiments & Illustrations

(2) Presidential Speech

- Barack Obama delivered his New Hampshire Primary Concession Speech entitled "Yes We Can" on January 8th, 2008, in Nashua, New Hampshire.

And whether we are **rich** or **poor**, **black** or **white**
(X 2 -X   X   2 -X)

We can bring **doctors** and **patients**,
(X 2 -X)

workers and **businesses**,
(X 2 -X)

Democrats and **Republicans** together
(X 2 -X)

- E(Xi) = 2
- E(Yi) = 1+4+1+1=1.75
- For r = 10, E(L) = 17

III Experiments & Illustrations

(3) Shakespearean Classic

X
What's he that wishes so?
-X   X   -X
My cousin, Westmoreland? No, my fair cousin;

$E(X_i) = (5+1)/2 = 3$
$E(Y_i) = 2$
For $r = 6$, $E(L) = 8$

We see that E(Xi), E(Yi) = 2, and E(L) for these three categories of literary styles are distinctly different.

III Experiments & Illustrations

(3) Shakespearean Classic

- The antithesis is a literary device in which opposites are put close to one another in a sentence or phrase for contrasting effects. In the opening of Shakespeare's "Henry V" (before Act 1 Scene I), there are several examples of antithesis when Chorus enters and delivers the following speech.

Stanza 1:

O for a Muse of fire, that would ascend
           X   2   3
The brightest **heaven** of invention,
4   - X   X
A **kingdom** for a stage, princes to **act**
2   3   4   - X
And monarchs to **behold** the swelling scene!
Then should the warlike Harry, like himself,
Assume the port of Mars; and at his heels,
Leash'd in like hounds, should famine, sword and fire
Crouch for employment.
But pardon, and gentles all,
                     X
The flat unraised spirits that have **dared**
2 3   4   5   6 7   8
On this unworthy scaffold to bring forth
9 10   11 12   13 14   15   16
So great an object: can this cockpit hold
17   18   19   20 21   22 23 24 25
The vasty fields of France? or may we cram
26   27   28   29 30 31   32
Within this wooden O the very casques
33   34   -X
That did **affright** the air at Agincourt?

- First, the length between "heaven" and "kingdom" is 4. Regarding heaven and kingdom, the thesis in this contrast is the idea that the kingdom is like a stage, a place of grandeur and excitement worthy of the presence of princes and monarchs. The stage is an unworthy scaffold, a place that is limited in scope and unable to contain the vastness of the world outside like heaven. However, the antithesis of the kingdom is heaven, which is limitless because it usually refers to the expanse in which the celestial bodies, such as the sun, moon, and stars are observed, which was historically considered to resemble a large vault stretching over the earth, also known as the sky or the firmament.

- Second, in terms of "act" and "behold", the length between them is 4 as well. The thesis is that the stage is a place for actors to take on roles, an opportunity to "act" and perform for audiences. The antithesis is that the monarchs and audience members who watch the play are passive spectators who "behold" the action, emphasizing their potential power and leashed engagement.

Stanza 2:

O, pardon! since a crooked figure may
        X   2   3   4
Attest in **little** place a million;
5   6 7   8   9 10 - X
And let us, ciphers to this **great** accompt,
On your imaginary forces work.
Suppose within the girdle of these walls
Are now confined two mighty monarchies,
Whose high upreared and abutting fronts
              X   2   3
The perilous narrow **ocean** parts asunder:
4   5 6   7   8   9   10
Piece out our imperfections with your thoughts;
11 12 13   14   15   16 17
Into a thousand parts divide on man,
18 19   20   21
And make imaginary puissance;
22   23   24 25 26 27   28 29 30   31
Think when we talk of horses, that you see them
32   33   34   35 36 37   38   -X
Printing their proud hoofs i' the receiving **earth**;
For 'tis your thoughts that now must deck our kings,
         X 2   -X
Carry them **here** and there; jumping o'er times,
Turning the accomplishment of many years
Into an hour-glass: for the which supply,
Admit me Chorus to this history;
Who prologue-like your humble patience pray,
Gently to hear, kindly to judge, our play.

- Between "little" and "great", the length is 10 words. Here, the thesis is that a small place can be of great importance, while the antithesis is that something great can also be imperfect and have flaws. The line "a crooked figure may / Attest in little place a million" highlights the significance of small things in the grand scheme of things.

- Regarding "ocean" and "earth", the length is 38 words. The thesis is the idea that two mighty monarchies are separated by the narrow and perilous "ocean" that lies between them. It suggests a vastness and distance that capture the imagination, allowing the audience to visualize and join the drama. The antithesis of "earth" is that these monarchies are confined within the walls of the stage, reducing the scale of the conflict and the drama.

Conclusion

➢ **Antithesis and Thesis**

Antithetical elements are often used by authors in English literary writings and speeches. Antithesis refers to the juxtaposition of contrasting words or ideas, which often, although not always, in parallel structure. Such rhetorical figure is used to contrast opposing ideas, creating a sense of tension and urgency, as well as heightening the emotional impact of the speech. Contrasting the antithesis is the thesis, which refers to a statement, assertion, or tenet. It is observed that the points of occurrence of thesis and antithesis are random, yet they occur with remarkable regularity alternating with each other.

➢ **Renewal Stochastic Process**

In this study, we represent the lengths of the thesis and antithesis by using an alternating renewal stochastic process. We find that the underlying attributes, such as the length of the thesis, the length of the antithesis, and the renewal function, of the alternating renewal process are able to usefully characterize the writing style of individual authors and help us to quantify and understand their linguistic technique and intention.

➢ **Experiments**

Experiments are carried out to illustrate the present approach. Using the present method in conjunction with other techniques, such as sentiment analysis, and word embedding models, it is possible to gain a deeper understanding of the literature and its underlying themes and structures.

Thank you