Does Shakespeare Pass the Bechdel Test?

Zhiyan Foo

November 11, 2016

1 Background and Significance

Despite the recent success of films such as *The Hunger Games* and *Star Wars: The Force Awakens*, most of our cultural output is disproportionally focused on male lives and their stories[1]. One popular measure of this gender bias is the Bechdel Test[2]. The test specifies 3 criteria for a piece to pass.

1. It must contain two named female characters
2. They must have a conversation
3. The conversation must not be about a man

While this is not a particularly high bar, a surprising amount of films—the original medium the test was applied to[1]—fail it. Gender disparities however are not a recent phenomena and this article looks at whether Shakespeare's plays in particular pass the Bechdel Test. This would not be done manually however but as a showcase for a new software package, **crunch-shake**[2].

2 Methodology

2.1 Modified Criteria

The criteria used by crunch-shake is slightly different from the original Bechdel Test. This is to both accommodate and utilize the computational nature of crunch-shake.

[1] see http://bechdeltest.com/ for a crowdsourced effort to classify movies by the Bechdel Test.

[2] https://github.com/zhiyanfoo/crunch-shake

Original Criteria	crunch-shake Criteria	Reason for Change
Contain two named female characters	Contain two female characters that are in the upper 50% of notable characters. See Gender and Notability.	The requirement that the two female characters be named is just a proxy for whether the character is significant to the piece, and using social network graphing algorithms and the number of lines the character has, this can be got at directly.
The two named female characters must have a conversation	In any scene, two notable female characters must speak in the presence of one another. See Presence.	It is hard to algorithmically determine when two characters are in a 'conversation' with one another. So the two females might be talking to a male, not each other, but unfortunately the scene will still 'pass'.
The conversation must not be about a man	In their 'conversation', the two notable females must not utter a word related to romantic relationships or mention a male. See Blacklist.	While there's a lot of subtext that an algorithm can miss out, a blacklist of words takes care of the more obvious cases.

2.2 Gender

The gender of each character is determined manually. The json files that include gender classifications can be found here [3]. While it is easy enough to be determine the gender of the named characters, the unnamed characters proved to be a bit harder. Some gender assignments were easy, such as

[3] https://github.com/zhiyanfoo/crunch-shake/tree/master/crunch-shake/gender

'groom' or 'maid'. Others such as 'Soldier', while could represent both genders in today's society, where for the most part male in Shakespeare's time and so classified as such. The more ambiguous designations such as 'Citizen' while were for the most part probably intended to be male, are left as 'N', and for the purposes of the algorithm, might as well not exist.

2.3 Notability

Whether or not a character is notable or not is dependent on how each character scores on 4 metrics: lines by character, out degree, page rank, betweenness. The last 3 metrics are network algorithms. crunch-shake uses the implementation found in the python package networkx [4].

Metric	Weight	Algorithm
lines by character	62.5%	Take the number of lines a character has and divide it by the lines of speech the character with the most lines has.
out degree	12.5%	networkx's implementation
page rank	12.5%	networkx's implementation
betweenness	12.5%	networkx's implementation

In this case, the network vertices here are speaking characters, with directed edges representing when a character speaks to another character, with more lines of speech indicating a stronger connection.

The use of network algorithms to classify importance of characters was taken from a paper on Game Of Thrones. Since Shakespeare's plays feature far fewer characters that GOT and since those characters are far more interconnected than in GOT, this methodology does not work as well for this paper, which is why the network algorithms are assigned such low weights.

2.4 Presence

2.4.1 Method

There are two ways the algorithm knows that a character is in a scene. The first is through stage directions. For example, if the algorithm sees the stage

[4] https://networkx.readthedocs.io/en/stable/reference/algorithms.centrality.html

direction

> Enter CAPULET in his gown, and LADY CAPULET.[5]

it would note that CAPULET and LADY CAPULET have entered. The second methods occurs if it sees that Romeo speaks a line, even if he never 'entered' the scene—so that scene started *in media res*—it would take it that Romeo was always there. Similarly, the algorithm will remove the character from the scene if it sees "Romeo Exit" or just when the scene ends.

2.4.2 Limitations

There are limitations to the algorithm however. For example if it sees,

> Exeunt all but MONTAGUE, LADY MONTAGUE, and BENVOLIO.[6]

it would erroneously note that MONTAGUE, LADY MONTAGUE, and BENVOLIO have all exited as it sees 'exeunt' while keeping the rest of the characters as it is not smart enough to interpret 'but all'. Also if a character enters, exit and re-enters, will note only first entrance and last exit. Finally sometimes the play directions do not refer to the characters by name. Take this example from Act IV, Scene III of *The Taming of the Shrew*, where the SERVANT is not mentioned directly.

> Enter four or five Serving-men¡/i¿¡/p¿

Yet is suppose to enter the scene.

Yet these errors are unlikely to skew the data in favor of males or females.

2.5 Blacklist

The list of forbidden words

romantic relationships marriage, matrimony, courting, love, wedlock, sex, sexual, intercourse.

male partners boyfriend, partner, husband, spouse, lover, admirer, fianc, amour, inamorato. (As well as the names of males in the piece).

[5] From Act I, Scene I in *Romeo and Juliet*

[6] From the same scene

The idea is not that it is 'wrong' for script writers to ever have their female characters use this words, but rather not ever scene involving female characters should have them focused on their relationships with male counterparts.

2.6 Avoiding Systematic Error

2.6.1 No Peeking

In order to not bias the data by, for example changing the blacklisted words, or changing the weights given to each metrics, I wrote all the specifications before running the crunch-shake on Shakespeare's plays, with the exception of *All's Well that Ends Well* and *Romeo and Juliet*, which I needed for debugging and checking feasibility purposes. As a result I will be omitting these two plays from the results.

2.6.2 Analysis

In order to avoid cutting up the data after running the experiment to support a potentially biased viewpoint, let me first state what results I wish to present.

1. Percentage of females to males overall.
2. Percentage of notable female characters to notable male characters (the rest would be ambiguous) per scene and overall.
3. A ranking of the plays using the percentages of scenes that past the Bechdel Test.
4. The percentages that fail the test because of a lack of females versus those that failed because the conversation was about males.
5. A summary of what the females talked about in 8 randomly chosen scenes that passed the test.
6. A summary of what it was they were talking about in 8 randomly chosen scenes that failed the test because they said a word on the blacklist.

3 Results

References

[1] S. Jia, T. Lansdall-Welfare, S. Sudhahar, C. Carter, and N. Cristianini, *Women are seen more than heard in online newspapers*, PLoS ONE, (2016).

[2] I. B. Lyle Friedman, Matt Daniels, *Hollywood's gender divide and its effect on films*, Polygraph.