Learning Character Graphs from Literature

Sumit Gogia, Min Zhang, Tommy Zhang

November 6, 2015

1 Introduction

Literary scholars, when comparing different works of literature, frequently use character roles and relationships to make their comparisons. These roles include major single-character literary labels such as "protagonist", as well as two-character relationship labels such as "parent-child". These character relationships for a novel can be visualized as a graph, an example of which is shown below:

[insert graph example]

We propose, for this project, to build a system that automatically produces these graphs, which we call *character graphs*, given the text for a work of literature. While literary scholars frequently hypothesize trends in literature by performing a small scale analysis of different works manageable through close reading and manual annotation, having a system that can automatically visualize these character relationships can help with exploring trends across a large number of novels.

2 Methods

2.1 Data Collection

2.2 Baseline

2.3 Our Approach

Our proposed system takes the form of a pipeline of subsystems that extract character references from the text, determine semantic representations of them, and then classify their literary roles and relationships from this semantic representation. The pipeline is visualized below:

[Insert Pipeline Graphic]

The neural networks we train, for extracting semantic character representations (*character vectors*) and classifying literary relationships, are also visualized separately for clarity:

[Begin figure of neural networks]

Each of these networks is trained across the entire dataset of novels obtained from data collection, though we train the character vector network before the other two, as they require semantic character representations as input.

Our hypothesis is that a semantic character vector can be obtained for each character given context for the character's appearance in a novel, using the first network. The exact context that we should use is unclear - however, for a start, we believe that using the dependent verbs and direct objects in sentences of character appearance will be helpful. If this proves difficult to work with, either due to issues with training to target a high-dimensional sparse output, or beause we are unable to extract these features, we can fall back to features used in [insert McKeowan citation] for their manual method.

Given output from the first network, the other two networks are similar to many current classification networks using word vectors as input. We feed in a singlet or pair of characters, and then obtain an output label corresponding to their literary role or relationship. To keep learning feasible, we restrict to a small subset of major literary role and prominent relationship labels. Since the training is done over all novels, we believe we will have enough data to perform adequate learning.

After we have passed the characters through the last two networks, we then have labels for each character and a list of labeled 2-character relationships. As described in the introduction, these labels then induce a graph structure that we can visualize.

2.4 Result Measuring

3 Milestones