Who shot Mr.Burns?

William Biondi

October 2, 2023

Abstract

This project work aims to analyze emotions and interactions between characters in fictional data, it uses the script from a two part story of the series "The Simpsons"

1 Introduction

"Who Shot Mr. Burns?" is a two-part episode of "The Simpsons" TV show. The story is about Mr. Burns, the owner of the Springfield Nuclear Power Plant. In the first part of the episode, Mr. Burns is shot by an unknown. The rest of the town tries to find out who did it. The episode has a lot of surprises and jokes about American culture and politics. In the end, they find out who did it, and it's someone that nobody expected. The episode is widely regarded as one of the best in the series, and it has been chosen for this study for its structure and interesting lines in the script. In this project we will analyze emotions of the main characters, their emotional profile and its evolution thorough the story and finally, the interactions between the characters.

1.1 Emotions

An emotion is a complex psychological construct that involves a range of subjective experiences, including feelings, thoughts, and physiological responses. Emotions are often described as intense mental states that are triggered by specific events or stimuli and that motivate behavior. There are many different theories of emotion, but one of the most influential is the basic emotions theory by Paul Ekman [1], which proposes that there are six basic emotions that are universally recognized across cultures:

1. Joy
2. Sadness
3. Anger
4. Fear
5. Surprise
6. Disgust

The basic emotions are important in fiction because they help to create emotional engagement with the reader or viewer. By evoking these basic emotions, authors and filmmakers can create a sense of empathy and identification with the characters and their experiences. This can make the story more compelling and memorable, and can help to convey important themes and messages. In addition to the basic emotions mentioned before a neutral setting is included, this setting is used to create a sense of realism in the story to counterbalance more intense emotions, create suspense or anticipate an important episode: in this neutral setting will fall anything that does not fit into the basic categories of positive or negative emotions [2].

1.2 Dataset

The dataset for this project is derived from The Simpsons Dataset retreived from Kaggle [4] by filtering lines from the two episodes only, counting scenes by location changes and also assigning eventual character relate from the words spoken manually. After this preprocessing a dataset with 315 observations is obtained with the following features:

1. Episode: number of episode from the full scripts dataset, will be used to divide part one from part two
2. Line number:sequential number of the line inside the episode
3. Location id:identifier of a location
4. Scene: sequential number of the scene generated by location changes, mainly used to add finer granularity of the analysis
5. Character: name of the character
6. Spoken words: line said by the character
7. Refers: character related, can be none

2 Methodology

The methodology of this project consists of these macro steps:

1. Detecting emotions with a language model
2. Analyze the emotion profile of the main characters
3. Analyze the evolution of the emotions over the plot
4. Build a graph to analyze interactions and emotions between characters

2.1 Emotion detection

To perform emotion detection tasks transformers are broadly used, in particular BERT-based Large Language Models perform with good results. In this project i adopted a pre-trained RoBERTa architecture [3], which is a variant of the popular BERT model that has been optimized for improved performance on a range of natural language processing tasks: it is fine-tuned on a large dataset of English text that has been labeled with six basic emotions plus neutral.The model is designed to take a text as input and output a probability distribution over the seven emotions, indicating the likelihood that each emotion is present in the text. I performed this detection over all the lines assigning to the line the emotion with the highest likelihood.To enhance emotions over neutral i adopted a threshold system on the distribution: if neutral has the highest likelihood but the difference from the second highest likelihood does not exceed the threshold, then the second highest likelihood's emotion will be the label assigned to the line.

2.2 Network

Interactions between characters is analyzed by building a multi-directed graph G where each vertex v is a character and each edge u is a line said from one character related to another character (lines said without relating to another character were excluded)

3 Results

In this section, I will report the result of the analysis by setting a score difference threshold of 15% focusing on the main characters of the story which are:

1. Mr.Burns
2. Waylon Smithers
3. Lisa Simpson
4. Homer Simpson
5. Chief Wiggum
6. Seymour Skinner

3.1 Emotional Profiles

To study the emotional profiles of the main characters i compute the frequency of the emotions for each character and displayed radar plots in Figure 1
Besides neutral setting which was expected to be the most frequent it is possible to observe mostly

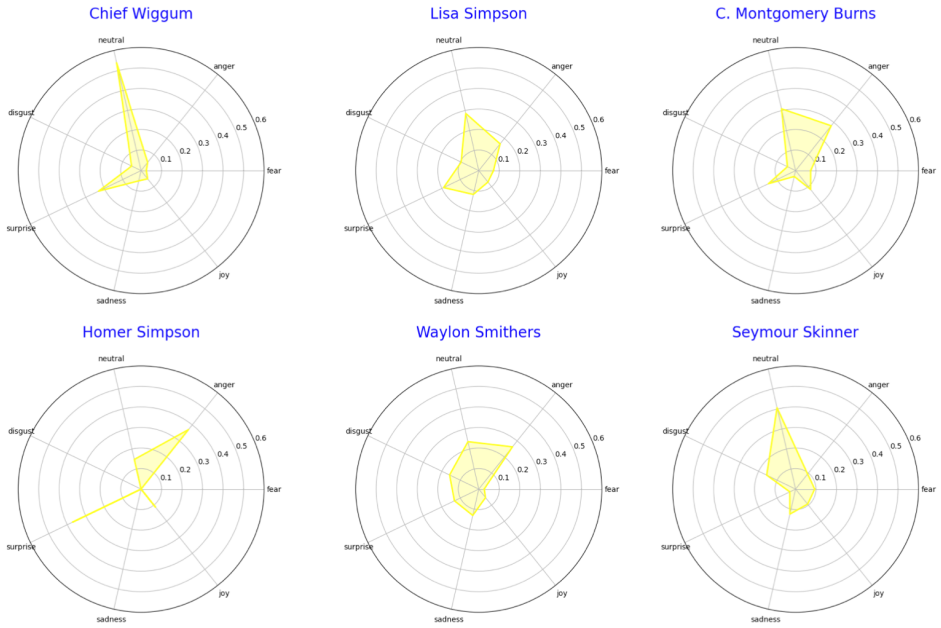

Figure 1: Radar plot of the emotions profiles

anger lines from Mr.Burns and Homer Simpson. Lisa Simpson and Waylon Smithers seems to have a more evenly distributed emotions whilst Seymour Skinner and Chief Wiggum being mostly neutral.

3.2 Evolution over the plot

By using the scene's structure i derived the most dominant emotion in each scene for each present character displayed in 2 and blank if the character is not present. As expected anger, surprise and neutral are the most dominant emotions over the plot: we can observe a similar emotional pattern from Mr.Burns and Waylon Smithers in the first 30 scenes and surprise lines at the same scenes from Burns, Homer and Lisa Simpson

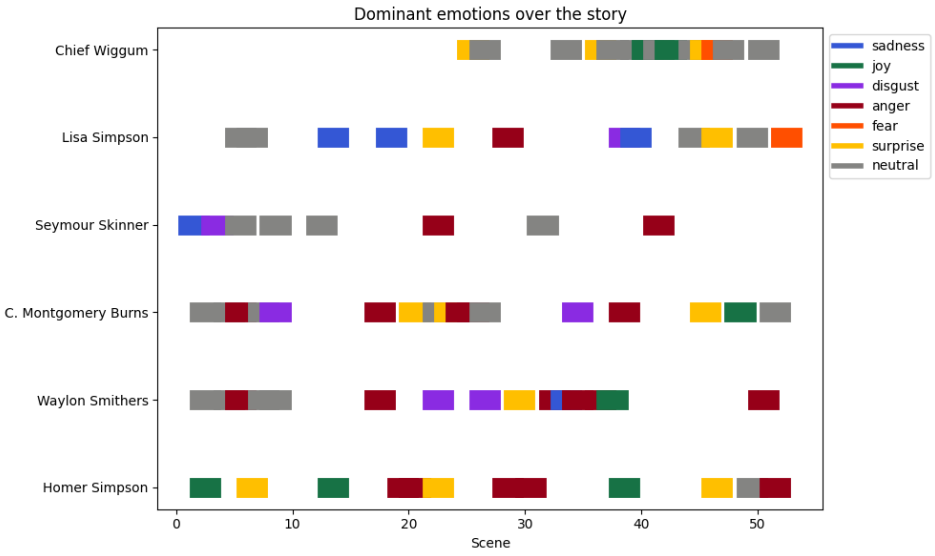

Figure 2: Evolution of characters over the plot

3.3 Emotional Interactions between the characters

The whole story is divided into two parts so to study the interaction i preferred to build two separate graphs to study the emotions before Mr.Burns got shot and after:In the first part's graph in figure 3 :as expected there are relevant anger and disgust interactions against Mr.Burns but surprisingly most anger interactions are against Seymour Skinner, especially from Mr.Burns, Smithers and Willie. By studying part 2 displayed in figure 4: as expected there is an intense anger interaction between

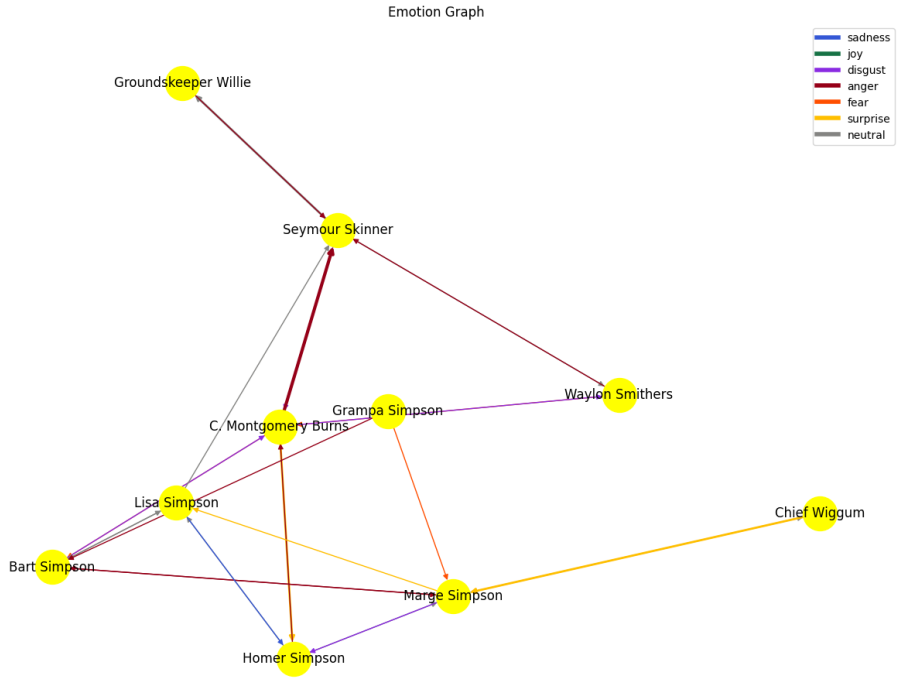

Figure 3: Interactions network in part 1: edges width indicate the intensity of the interaction

Mr.Burns and Maggie Simpson which is the true guilty of the story and in the same measure a sadness interaction derived from the delusion of Mr.Burns towards Chief Wiggum caused by the impossibility

to arrest a child. An unexpected surprise interaction is noted between Lisa and Bart Simpson, by checking however this interaction is derived from a gag which is not relevant to the main story.

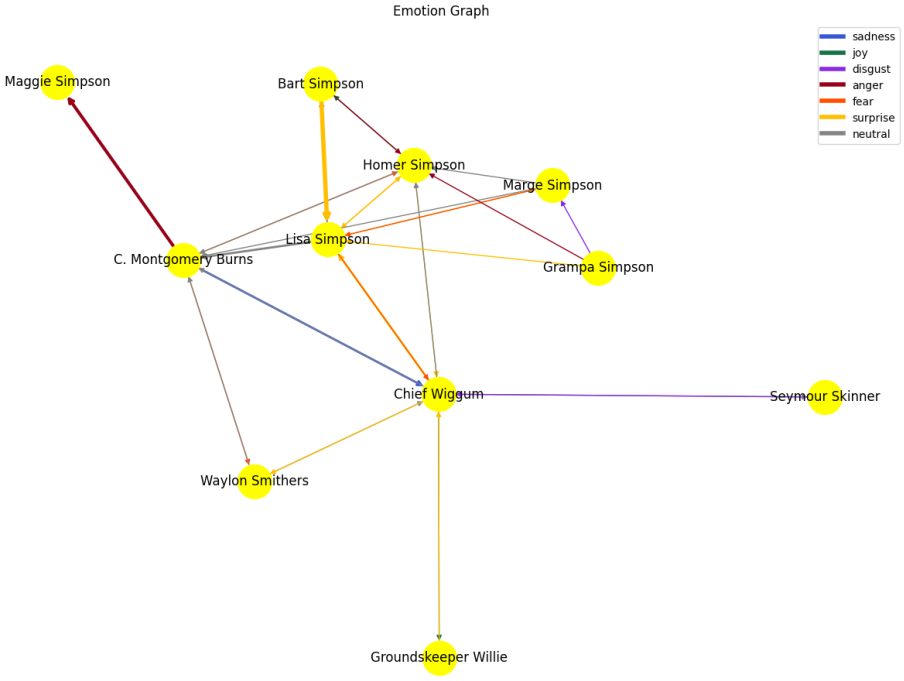

Figure 4: Interactions network in part 1: edges width indicate the intensity of the interaction

4 Conclusions and further work

The emotion detector with some adjustments with neutral settings brought results that align well with the anticipated plot of the story. However, a series like "The Simpsons" is full of side gags, flashbacks and other side components and it is essential to note that not all detected emotions may directly correlate with the main story. Some emotions might be tied to these side gags, which are often humorous and may not carry the same emotional weight as the primary plot events. Possible studies to make in the future are about refining the language model on this particular kind of fictional data in order to capture better the possible patterns in an episode (side gags, flashbacks, etc.) and also work with viewers point with an annotated script in order to check if the emotion detected corresponds to the audience perception.

Declaration

I declare that this material, which I now submit for assessment, is entirely my own work and has not been taken from the work of others, save and to the extent that such work has been cited and acknowledged within the text of my work. I understand that plagiarism, collusion, and copying are grave and serious offenses in the university and accept the penalties that would be imposed should I engage in plagiarism, collusion or copying. This assignment, or any part of it, has not been previously submitted by me or any other person for assessment on this or any other course of study.

References

[1] Paul Ekman. *Emotions Revealed: Recognizing Faces and Feelings to Improve Communication and Emotional Life*. Holt Paperbacks, 2003.

[2] Nico H Frijda. "The concept of basic emotions: A historical review". In: *Emotions: Essays on emotion theory* 1 (1988), pp. 345–368.

[3] Jochen Hartmann. *Emotion English DistilRoBERTa-base*. https://huggingface.co/j-hartmann/emotion-english-distilroberta-base/. 2022.

[4] Kaggle. *The Simpsons Dataset*. https://www.kaggle.com/datasets/prashant111/the-simpsons-dataset. Retrieved October 4, 2021. 2020.