

EmotionArcs

Emily Öhman

Waseda University,
Japan

Yuri Bizzoni

Aarhus University,
Denmark

Pascale Feldkamp

Moreira

Aarhus University,
Denmark

**Kristoffer L.
Nielbo**

Aarhus University,
Denmark

Jackie Daley

Introduction

Publish Date: March, 2024

Why is this research important?

- Utilizes recent technology methods towards literary analyses
- Develops a modern structure for measuring emotions in literature
- Provides a fresh set of emotion data and methodology to be expanded upon
- Provides a new perspective via correlational output

Background

Previous studies' topics...

Sentiment Arcs

Used to create visual arcs using clusters based on sentiment valence (positive/negative) at the word, sentence, and paragraph levels

Fractal Analysis

Used to study complex system dynamics to model the persistence, coherence, and predictability of sentiment arcs related to reader appreciation

NRC Lexicon

Tool that was used to determine the connection between the shape of emotion specific arcs and a book's success

Background cont.

Key Concepts:

What is EmotionArcs?

EmotionArcs is a dataset containing the emotion arcs derived from the Chicago Corpus (9,089 American novels, 1800s to 2000). It utilizes emotion intensity lexicons through word embeddings. This model covers the intensity of the following emotions: anger, anticipation, disgust, fear, joy, sadness, surprise, and trust.

What is Emotional Entropy?

Emotional Entropy is the variance of emotions within a novel. High emotional entropy means that there were an abundance of different emotions in the given text, while a low emotional entropy means that one emotion may have dominated.

Background cont.

Key Concepts:

What does the Slope of an Emotion Arc mean?

The slope represents the development of intensity for that particular emotion. A steep slope means that an emotion may be dragged out for a longer duration (positive slope) or absent for an extended period of time (negative).

What does the Skewness of an Emotion Arc mean?

The skewness is measured through the asymmetry of the arc. An arc with more outlier values is less symmetrical and therefore is positively skewed. A symmetrical arc with an even distribution will have a skewness nearing 0.

Research Questions

Does EmotionArcs accurately measure emotional content?

Can emotional arcs determine a novel's literary quality?

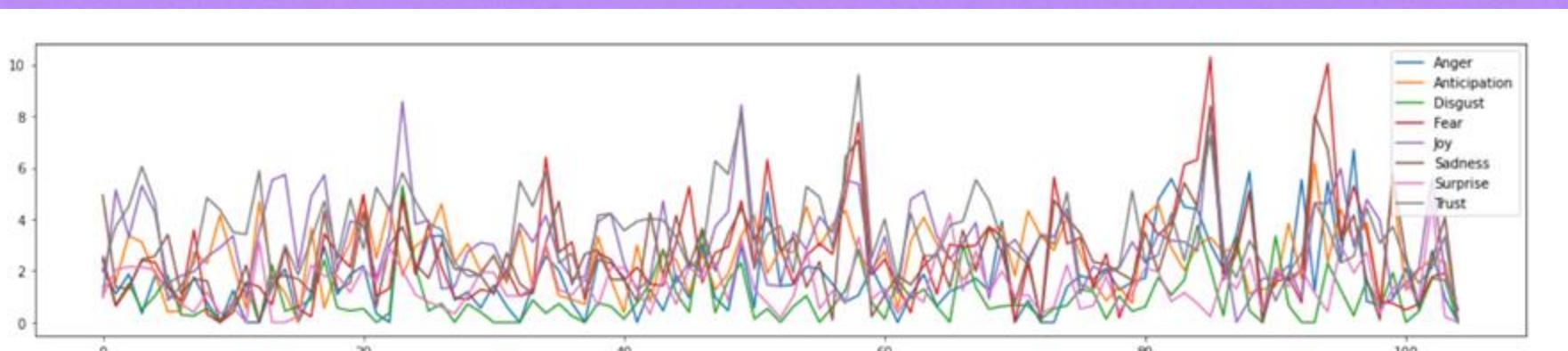
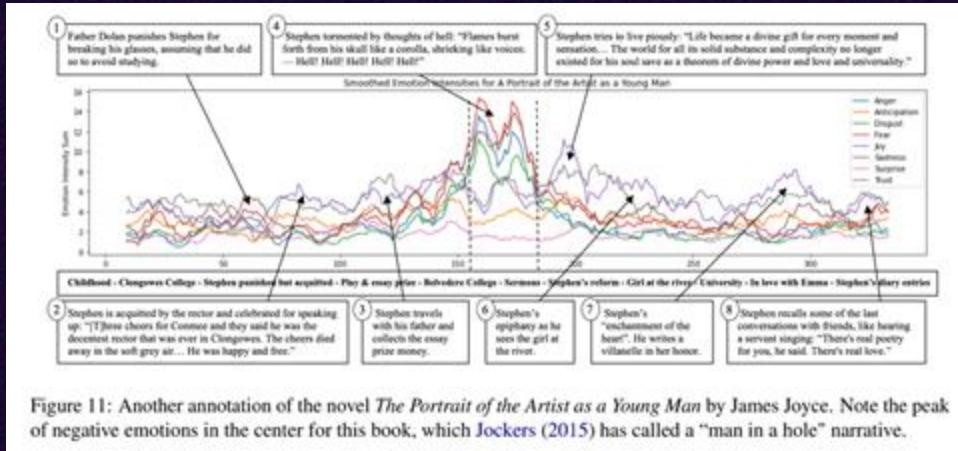


Figure 9: Unsmoothed emotion arcs for *The Old Man and the Sea*

Methodology

Validity:

- I. Two literary scholars related style and narrative events to the shapes of the emotion arcs with manual annotation.



- II. Randomly selected 11 passages from a novel and let 20 non-expert volunteers choose the present emotion and its intensity. Then later view EmotionArc's emotions and whether or not they agree with the emotion and intensity

- III. Compared EmotionArc's valence scores to both a RoBerta that was fine-tuned for sentiment analysis on tweets and human annotations for valence scores.

Findings

Validity:

- I. Emotions can be expressed simultaneously within the context of the story's key scenes. Some emotions are observed to overlap more frequently than others.
- II. The majority of volunteers agreed with EmotionArcs, with mixed reviews on the emotional intensity level.
- III. All three methods for measuring valence followed similar sentiment trends.

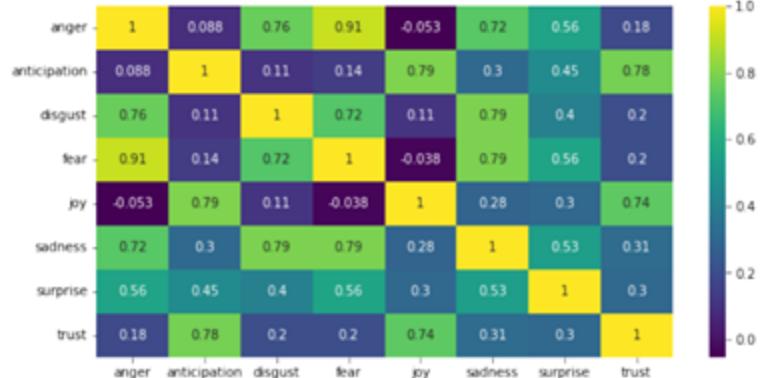


Figure 3: Correlation between emotions in all emotion arcs

So EmotionArcs' results are coherent!

Methodology

Literary Quality:

First Proxy: Number of ratings in GoodReads (crowd-based proxy)

Second Proxy: Library Holdings of the title (crowd and expert-based proxy)

Both proxy values will be compared to the EmotionArc's values for emotional entropy, arc slope, and skewness in order to identify any significant relationships.

The novels will be separated by if they were successful or unsuccessful, to observe if there is a connection to the EmotionArcs values.

Findings

Literary Quality:

Overall Novel Emotion:

Weak positive correlation - intensity of most emotions & number of library holdings

Weak negative correlation - overall emotional entropy & number of library holdings

When novels are separated by successfulness:

Strong positive correlation - higher entropy & higher number of ratings and library holdings

Slope and Skewness:

Positive correlation - more sloping arcs & less ratings and library holdings

Negative correlation - more skewness & less ratings and library holdings

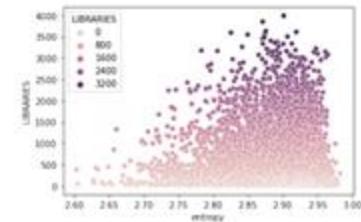


Figure 5: Distribution of library holdings with respect to emotion entropy.

Conclusion:

Overall, EmotionArcs finds that highly rated novels are most likely to have a high entropy (but not too high - there is a cutoff point), less sloping arcs, and minimal skewness. Novels with only one emotion intensity are too predictable, while having an excessive amount of emotion intensities leads to unpredictable or incoherent story arcs.

My Commentary

I wondered if EmotionArcs could expand its palette of emotions to select from. Although it might lead to each emotion intensity being more level and spaced out (which would make results harder to find)

I wondered if they could also apply this to dialogue from movies and shows to see if they come up with similar results

I thought that some of the diagrams were unintuitive and vague without context, which beats the purpose of being a visual aid to express data

Quiz Time!

Quiz Question 1

What are the main qualities of an emotion arc?

- A) Roundness, radius, and size
- B) Depth, valence, and peak
- C) Length, amplitude, and frequency
- D) Entropy, slope, and skewness

Quiz Question 2

What is EmotionArcs?

- A) An emotion lexicon designed specifically for poetry
- B) A new model for classifying literary genres by emotions
- C) A dataset of emotional arcs from over 9,000 English novels
- C) A website to rate the emotional flow within novels

Quiz Question 3

What is high emotional entropy and low emotional entropy?

- A) High emotional entropy is the most used emotions in a text, while low emotional entropy is the least used emotion in a text
- B) High emotional entropy is an abundant amount of complex vocabulary in a text, while low emotional entropy is not complex
- C) High emotional entropy is one intense emotion dominating in a text, while low emotional entropy is several less intense emotions in a text
- D) High emotional entropy is an diverse amount of emotions in a text, while a low emotional entropy is one emotion dominating

Quiz Question 4

Based on the Emotion Arcs, which two emotions correlate the least?

A) Anger and Fear

B) Joy and Fear

C) Joy and Anger

D) Trust and Surprise

Any Questions?

The End
