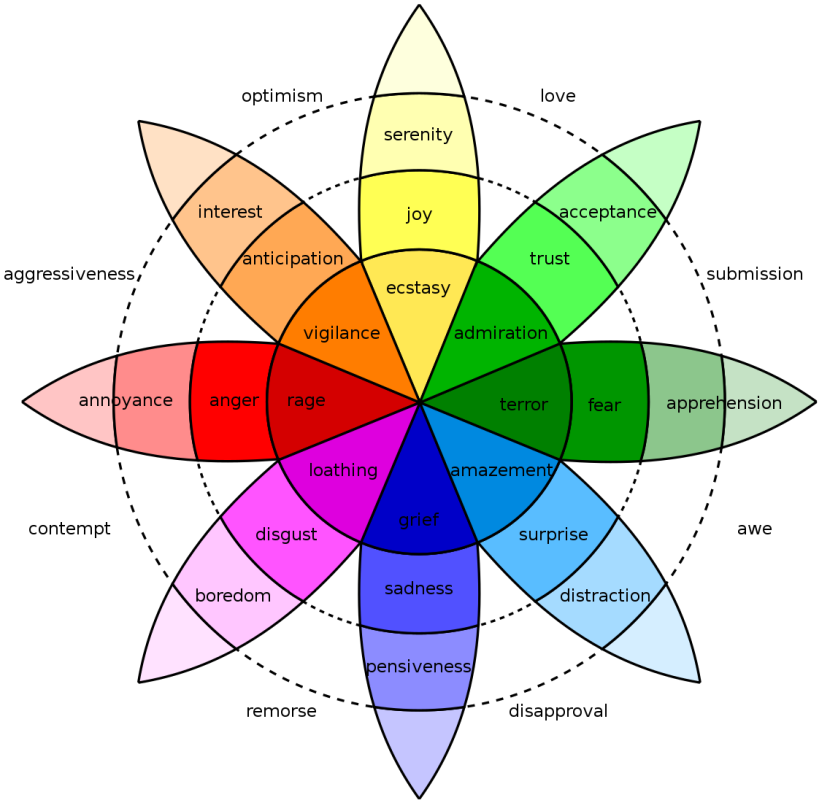
Ontology Based Emotion Recognition System

Paper:
An Ontology for Emotion Analysis (2016)
https://www.researchgate.net/publication/301016353_EmotiOn_An_ontology_for_emotion_analysis



Project:

A java application.

Input: raw text (e.g.: tweets, reviews, facebook posts, etc.). I used IMDB reviews from the imdb reviews dataset available on Kaggle.

Output: the predominant emotion(s) of the text. Focus switched from text polarity classification (positive/negative) to emotion recognition.

Technologies used:

- Ontosenticnet ontology: https://sentic.net/ontosenticnet.pdf
- OwlApi: http://owlapi.sourceforge.net/documentation.html
- OpenNlp: https://opennlp.apache.org/docs/1.8.0/manual/opennlp.html

- JSoup: https://jsoup.org/
- OpenCsv: http://opencsv.sourceforge.net/

Application Data Flow:
For a given input text, the application outputs one or more(if scores are equal) predominant emotion(s). The possible values are: admiration, anger, disgust, fear, interest, joy, sadness, surprise.

1. Initialization: build a cache of all ontology individuals which will be later used in the emotion recognition process. Also load in memory the input texts
2. Pre-processing: takes place for each review (html tags, emoticons, punctuation and numbers removal, tokenization, stop words removal, POS tagging, adjective retrieval)
3. For each adjective, query the ontology for its associated primitive emotion and polarity score and store the values.
4. Emotion Recognition: For each review count the number of occurrences of each primitive emotion given by the identified adjectives. If an adjective is negated, compute the opposite emotion (based on its position in the wheel of emotions) and increment the value of that emotion.
5. Compute the text polarity and tune the result according to it. E.g., if for an review the predominant emotions are disgust and admiration, if the text polarity is negative, remove admiration from the result.

Possible Improvements:
1. Use other parts of speech besides adjectives: adverbs, verbs, nouns
2. Use phrases instead of parts of speech
3. Elaborate the score computation formula