Finding Better Affective Meaning in Emojis: Mapping Emojis to an Affective Space

Nimesh Ghelani
University of Waterloo
nghelani@uwaterloo.ca

***Abstract*—Emojis have seen an increased popularity in recent years. However, most sentiment analysis and emotion research on text make limited or no use of them. Through this project, we attempt to provide a richer affective meaning to emojis; by devising a technique which maps them to the Arousal-Valence affective space.**

I. Introduction

Text is a primary way of communication over the internet. In recent years, the use of *emoticons* (or emojis) have become more common. Emojis are symbols ranging from smiley faces to road signs. Unlike images, they are encoded as part of the text. Almost all online platforms and smartphones have integrated support for emojis. Users often find it convenient to express complex emotions through them. It is so prominent that Unicode[1] maintains emoji specifications to ensure they are consistent across platforms. Unicode released Emoji 5.0[2] in March 2017, which supports 2623 different emojis.

Derks et al [1] performed a user study through which they concluded that emojis have a significant impact on user's understanding of text and can be equated to "non-verbal actions".

One of the key areas of interest in affective computing is understanding emotions in text. Most human-human and human-machine interactions on the internet is through text, and thus understanding emotions in text is important if we wish to build software which are more user oriented and human like. Many advancements have been made in the topic of textual *sentiment analysis* over the past two decades [2] [3]. The focus of these research are primarily on the words present in the text. Researchers do use emoticons as features when performing sentiment analysis of content in social media like twitter. However, the interpretation of emoticons in most of these techniques is very simple (for example, categorizing emoticons into small number of emotion classes or mapping them on a one dimensional sentiment scale).

Two popular models of emotions are the *EPA* (Evaluation-Potency-Activity) model [4] and the *AV* (Arousal-Valence) model [5]. These dimensional models of emotions assume emotion to be a point in a two or three dimensional space. These models work well with theories like the Affect Control Theory [6] which attempt to model complex human interactions. Words have been mapped into the AV or EPA model through past user studies [7] [8]. In this project, we aim to automatically map emojis to the Arousal Valence space using the AV values of words accompanying them. We experiment and analyse two methods for this task. The first method relied on combining AV values of words present in the description in the Unicode specification of an emoji. The second method relied on combining AV values of words present in tweets which also contained the emoji.

Rich affective definitions of emojis can be used for better emotion detection/generation. Conversational agents can also benefit from better use and understanding of emojis, making them seem more human like and intelligent. Emojis transcend language barriers making a better affective definition universally useful.

II. Related Work

Emojis are used as features when working with social media data like twitter. Kouloumpis et al [9] performed sentiment analysis on twitter data using emojis as a feature. They label emojis as positive and negative for this purpose. Agarwal et al [10] use emoji in a similar manner but instead hand annotate emoticons into five sentiment classes (extremely positive/negative, positive/negative, neutral).

Ptaszynski et al [11] built an emoticon dictionary for affect analysis from a kinesics point of view (using eyes/mouth/face information of emojis). The end result was a set of emotion classes automatically assigned to a list of emojis.

Healey and Ramaswamy's [12] work deal with mapping tweets into arousal valence scale by taking the mean of AV values of component words in the tweet. Our idea of using AV values of component words to derive AV values of an emoticon is roughly influenced from this work.

III. Dataset

Instead of performing our experiments on all available emojis, we clipped our emoji list to contain only the popular ones. *EmojiTracker*[3] tracks and logs real time usage of emojis in in twitter. We use all the top emojis tracked by this website, which total up to 837.

We fetched the emoticon descriptions of these 837 emojis from the unicode specifications. Table I lists some example emojis with their keyword descriptions from the specification. We used the twitter live stream api to collect tweets containing

[1] http://unicode.org
[2] http://unicode.org/emoji/charts/index.html
[3] http://www.emojitracker.com/

the desired emojis for approximately 24 hours. We observed a varying distribution of tweets per emoji which is summarized in Table II. There were many emojis with number of tweets in single digit.

For Arousal-Valence values of words, we relied on the affective norms collected by Warriner et al [8]. The AV dataset provides the mean and standard deviation of these values for around 13k words.

Emoji	Keywords
😍	*eye* | *face* | *love* | *smile*
😎	*bright* | *cool* | *eye* | *eyewear ...*
😭	*cry* | *face* | *sad* | *sob* | *tear*
😠	*angry* | *face* | *mad*

TABLE I
EXAMPLE KEYWORDS IN THE UNICODE SPECIFICATION

Emoji	Count	Emoji	Count
😂	*43790*	😥	*917*
😊	*12221*	💩	*620*
😎	*4978*	👷	*36*
😩	*9072*	🔏	*3*

TABLE II
EXAMPLE EMOJIS AND THEIR RESPECTIVE TWEET COUNTS

IV. METHODOLOGY

A. Using unicode specification data

Given an emoji and a list of keywords describing it, we wish to compute its AV value. We fetched the AV values for a subset of these keywords (some of them might not be present in the dataset). We defined the AV value of the emoji as a combination of AV values of these keywords. One simple way to combine them is to take a simple average. However, in our experiments, this led to the final values being closer to neutral AV values (5, 5) due to the wide presence of neutral words like "face", "man", "mouth", etc in the specification. To mitigate this and make our AV values more expressive, we propose doing a weighted average of component AV values.

The combined arousal and valence $A_{combined}$ and $V_{combined}$ is computed from the component arousal and valence values A_i and V_i as

$$A_{combined} = \frac{\sum_i A_i \times wa_i}{\sum_i wa_i}$$

$$V_{combined} = \frac{\sum_i V_i \times wv_i}{\sum_i wv_i}$$

The weights wa_i and wv_i are computed as

$$wa_i = 1 + (A_i - 5)^2$$

$$wv_i = 1 + (V_i - 5)^2$$

The 1 is added to avoid zero weights. To further reduce the noise, we ignored the top ten frequent keywords (*"face"*, *"woman"*, and so on).

B. Using tweets

Given an emoji and a list of tweets containing it, we wish to compute its AV value. We followed the similar strategy of weighted average from the previous section, but twice:

- Compute AV value of an emoji by doing a weighted average of AV values of all the tweets that contain that emoji. The weighting scheme described in the previous section is used.
- Compute AV value of a tweet by doing a weighted average of AV values of the component words in the tweet text. The weighting scheme is slightly different, using mean AV values of words across the entire tweet collection.

$$wa_i = 1 + (A_i - A_{mean})^2$$

$$wv_i = 1 + (V_i - V_{mean})^2$$

This was done based on the observation that words used in tweets are generally biased towards high valence and low arousal. The A_{mean} and V_{mean} in our experiments were found out to be 4.29 and 5.69, respectively.

V. RESULTS AND DISCUSSION

Figure 1 and 2 show a sample of emojis plotted according to their AV values. The emojis were hand picked to optimize presentation and exclude obscure emojis. Compared to the AV values computed using unicode specs (lets call it *AV-unicode*), the AV values computed using tweets (*AV-tweets*) lack expressiveness. Figures 3 and 4 demonstrate this difference in expressiveness.

Is the difference between the results of two methods merely due to poor calibration? To find out, we measured the stability of relative AV values between *AV-unicode* and *AV-tweets* using Kendall's τ. Given two ranklists, Kendall's τ can be used to find correlation between two ranklist:

$$\tau = \frac{(\#\ of\ stable\ pairs) - (\#\ of\ unstable\ pairs)}{total\ \#\ of\ pairs}$$

A pair (x, y) is stable if their relative positions are same in both the ranklists (i.e. if $(position\ of\ x) < (position\ of\ y)$ in the first ranklist, same is true in the second ranklist, and vice versa). An unstable pair is the opposite of a stable pair. $\tau = -1$ indicate that the ranklists are reverse of each other. $\tau = 1$ indicate that the ranklists are exactly the same. $\tau = 0$ indicates there is no correlation. Positive and negative values of τ indicate a positive and negative correlation, respectively.

In our case, we had ranked lists of emojis sorted by arousal or valence values. For emojis ranked on arousal values, we found $\tau_a = 0.52$ between $AV - unicode$ and $AV - tweets$. For emojis ranked on valence values, we found $\tau_v = 0.70$. High value of τ_v confirm the similarity of relative valence

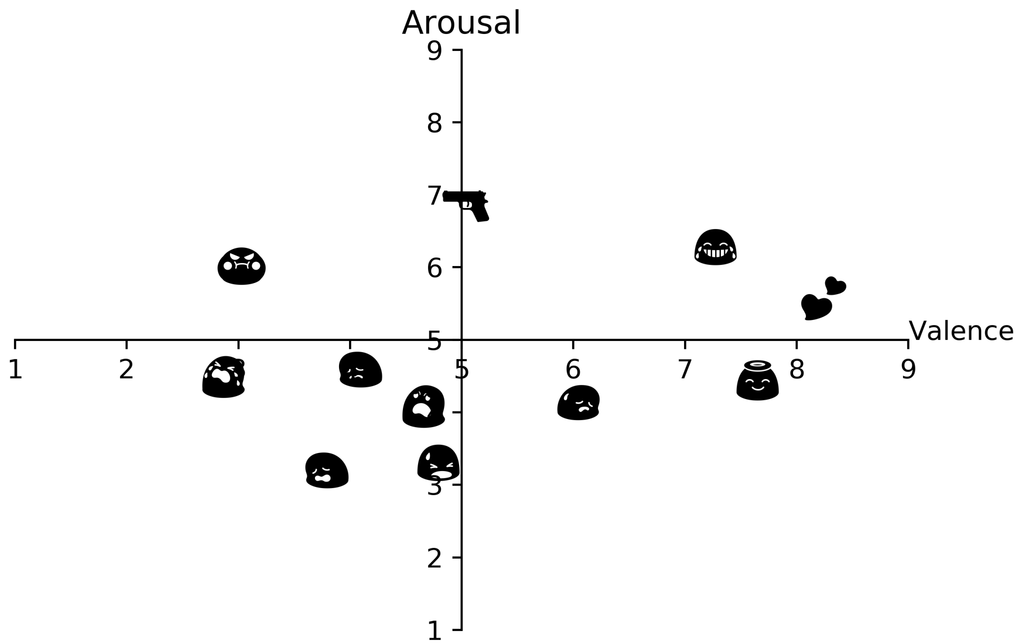

Fig. 1. Emojis plotted on Arousal Valence scale (values computed from unicode specs)

score. However, the same cannot be said for relative arousal scores.

As an example application of our new AV mappings, we wrote a small program which predicts emojis based on the input text. The AV values of the input text is computed the same way as tweets in this project. Using euclidean distance, we fetched three emojis closest to the AV value of the input text. To exclude obscure emojis, we only considered popular emojis (emojis with at least 1000 tweets in our collection). When the input text is *"I am feeling great"*, the program predicted the following emojis: 🌸, 😌, 😄. When the input text is *"My dog died"*, the program predicted the following emojis: 😱, 😡, 💀.

VI. Conclusion and Future Work

In this paper, we described two automatic methods to map emojis to the Arousal-Valence model of emotion. The method using data from the unicode specifications produced promising results. While the method using tweets produced AV values which are relatively consistent (in the valence scale) with the other method, it lacked expressiveness. Based on our observations, we list few reasons for poor quality of *AV-tweets*

- Large number of terms from a tweet contributed to the combined AV value. These terms usually had a high valence and low arousal. These terms shadowed the terms which actually reflect the emotion of the text. Improving the weighting technique proposed in this paper can help mitigate this issue to some extent.
- Significant number of terms which are crucial for computing the *true* AV values of the tweet were absent from the AV dataset. For example, many slangs and short forms like *"omg"* and *"wtf"* were not present in the AV dictionary. Expanding the AV dictionary or translating these slangs (for example, to their formal definition) can solve this issue.
- Averaging of AV values of component words can fail in many cases, specially for tweets where the words are part of a sentence. For example, in the tweet *"A man killing a 5year old child and drinking his blood without feeling impenitent, tf is wrong with these people"*, averaging individual words (even when weighted) dilute the overall AV value towards neutral due to words like "drinking", "blood", "man", "child" considered without any context. Incorporating sophisticated techniques like Affect Control Theory can be a potential fix to this problem [13].

In the previous section we demonstrated a sample application which predicted emojis for a text using our new AV mappings. There is potential future work in these lines. Emoji prediction can be a useful feature for users of smartphones and social media. Our own sample application can be improved by using techniques like affect control theory to better understand input text and predict suitable emojis. Conversational agents can become more interactive by meaningful use of emojis in conversations. Many more creative ideas can be perceived using the data computed by the work mentioned in this paper.

References

[1] D. Derks, A. E. Bos, and J. Von Grumbkow, "Emoticons and online message interpretation," *Social Science Computer Review*, 2007.

[2] B. Pang, L. Lee *et al.*, "Opinion mining and sentiment analysis," *Foundations and Trends® in Information Retrieval*, vol. 2, no. 1–2, pp. 1–135, 2008.

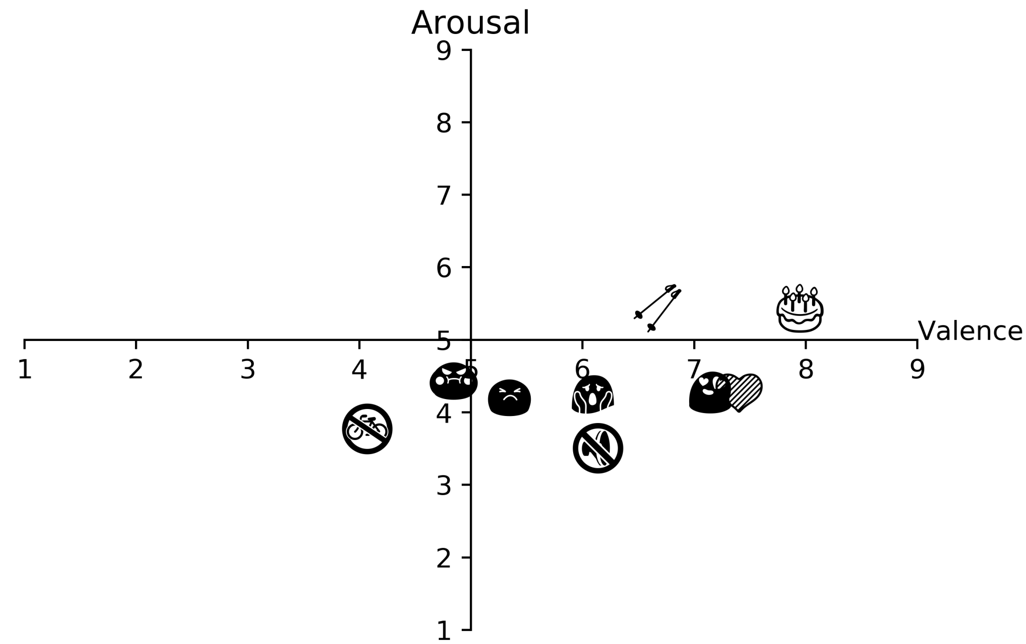

Fig. 2. Emojis plotted on Arousal Valence scale (values computed from tweets)

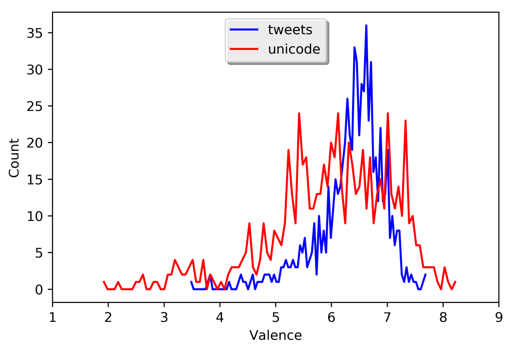

Fig. 3. Distribution of emojis by valence values

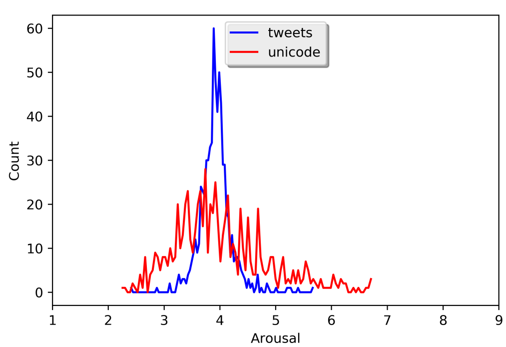

Fig. 4. Distribution of emojis by arousal values

[3] S. Baccianella, A. Esuli, and F. Sebastiani, "Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining." in *LREC*, vol. 10, 2010, pp. 2200–2204.

[4] C. E. Osgood, W. H. May, and M. S. Miron, *Cross-cultural universals of affective meaning*. University of Illinois Press, 1975.

[5] J. A. Russell, "Affective space is bipolar." *Journal of personality and social psychology*, vol. 37, no. 3, p. 345, 1979.

[6] D. R. Heise, "Affect control theory: Concepts and model," *Journal of Mathematical Sociology*, vol. 13, no. 1-2, pp. 1–33, 1987.

[7] M. M. Bradley and P. J. Lang, "Affective norms for english words (anew): Instruction manual and affective ratings," Technical report C-1, the center for research in psychophysiology, University of Florida, Tech. Rep., 1999.

[8] A. B. Warriner, V. Kuperman, and M. Brysbaert, "Norms of valence, arousal, and dominance for 13,915 english lemmas," *Behavior research methods*, vol. 45, no. 4, pp. 1191–1207, 2013.

[9] E. Kouloumpis, T. Wilson, and J. D. Moore, "Twitter sentiment analysis: The good the bad and the omg!" *Icwsm*, vol. 11, no. 538-541, p. 164, 2011.

[10] A. Agarwal, B. Xie, I. Vovsha, O. Rambow, and R. Passonneau, "Sentiment analysis of twitter data," in *Proceedings of the workshop on languages in social media*. Association for Computational Linguistics, 2011, pp. 30–38.

[11] M. Ptaszynski, J. Maciejewski, P. Dybala, R. Rzepka, and K. Araki, "Cao: A fully automatic emoticon analysis system." in *AAAI*, 2010.

[12] R. S. Healey, C., "Visualizing twitter sentiment (2010)," http://www.csc.ncsu.edu/faculty/healey/tweet_viz/.

[13] A. Alhothali and J. Hoey, "Good news or bad news: Using affect control theory to analyze readers' reaction towards news articles." in *HLT-NAACL*, 2015, pp. 1548–1558.