A Structure of basic emotions: A review of basic emotion theories using an emotionally fine-tuned language model

Junho Lee[1] **Cheongtag Kim**[1,2]
[1]Interdisciplinary Program in Cognitive Science, [2]Department of Psychology
Seoul National University, Republic of Korea
{smbslt3, ctkim}@snu.ac.kr

Abstract

There is growing interest in emotions in textual data. Based on psychological theories, research has been conducted on assigning emotions as labels to text datasets or developing models to detect emotions present in text. However, little research has been done on the appropriateness of using these theories. In this study, we reviewed three commonly used basic emotion theories: Ekman's Basic Emotions, Plutchik's Wheel of Emotions, and GoEmotions. By leveraging a research finding that evaluated the emotional values of words, we were able to fine-tune a language model emotionally. Using it, we analyzed the emotional relationship between the names of basic emotions and evaluated the adequacy of the emotional structure each theory presents. Clear patterns of similarity emerged based on the emotional meaning of the words. Ekman's and GoEmotions were almost in line with our results, while Plutchik's had some differences. We discussed these matches and mismatches.

Keywords: basic emotion; emotion model; emotional fine-tuning; emotion embedding

Introduction

Researchers on emotion have attempted to find structures and processes underlying emotional phenomena and behaviors. However, emotions are context-dependent because they are influenced by cultures and languages in which they are expressed. To find a general theory of emotion, psychologists have attempted to find a universal "basic emotion" that all humans feel independently of cultures and languages. Two seminal theories were proposed about basic emotion. Ekman proposed a basic emotion theory based on facial expressions and Plutchik proposed a structural model of emotions based on human motivations. These theories, based on psychoevolutional assumptions, have not only become standards in psychological research on emotions but have also been used as standards for emotional classification in applied research fields such as HCI and data science. On the one hand, recently reported emotion classification studies do not assume psychological structures and processes. These studies investigated which emotions are observable from actual behaviors by analyzing large amounts of data with computational methods. They found that our emotional experiences were distributed non-linearly in high dimensional spaces, which is different from the assumptions of classic emotion psychology.

In this study, we critically analyzed their findings from a new perspective. By utilizing an emotionally fine-tuned language model to analyze the relationships between emotions used in previous studies, we gained insight into the structure of basic emotions they proposed. Furthermore, we present our perspective on existing emotion models. Our contribution is twofold: we presented a method for fine-tuning a language model to emotional meanings and analyzed the structure of basic emotions using state-of-the-art NLP methods.

Related Work

Basic Emotion Theories

Darwin stated in his book that emotions are the products of evolution and, therefore, exhibit universality (Darwin & Prodger, 1874/1998). Ekman disagreed with this claim, but after several experiments, he found evidence supporting it and accepted the existence of universal emotions (Ekman, 2003). Ekman proposed a method of classifying emotions through facial expressions to verify Darwin's claims, which can identify emotions unaffected by cultures and languages. As a result, he identified six basic universal emotions (happiness/enjoyment, anger, fear, disgust, sadness, surprise) that are pancultural and panlinguistic (Ekman & Friesen, 1971), and later expanded the basic emotions list to seven by adding contempt (Ekman & Friesen, 1986; Ekman & Heider, 1988). Russell (1991) argued that contempt is not distinguished from disgust or sadness, but Ekman counterargued that contempt is distinct from disgust or sadness and qualifies as a basic emotion (Ekman, O'Sullivan, & Matsumoto, 1991).

Plutchik also proposed the existence of universal emotions from a psychoevolutional perspective (Plutchik, 1980). He assumed that emotions play a role in aiding the survival of organisms. For example, when an organism encounters a dangerous situation, the emotion *"fear"* is triggered, which motivates the organism to escape, thereby enhancing its chances of survival. Therefore, from this perspective, emotions are innate and universal. Plutchik proposed a three-dimensional cone-shaped model based on eight basic emotions that is consistent with the results of various emotion studies (Figure 1). This model incorporates variations in the intensity of the basic emotions on each emotional axis (`mild-intense`), and arranges each axis in proximity to each other based on their emotional similarities. He suggested that new emotions can be explained by combining two emotions on different axes (Plutchik, 1982). He visualized this circumplex model known as the Wheel of Emotions or Emotion Wheel (Plutchik, 2001).

In M. Goldwater, F. K. Anggoro, B. K. Hayes, & D. C. Ong (Eds.), *Proceedings of the 45th Annual Conference of the Cognitive Science Society.*

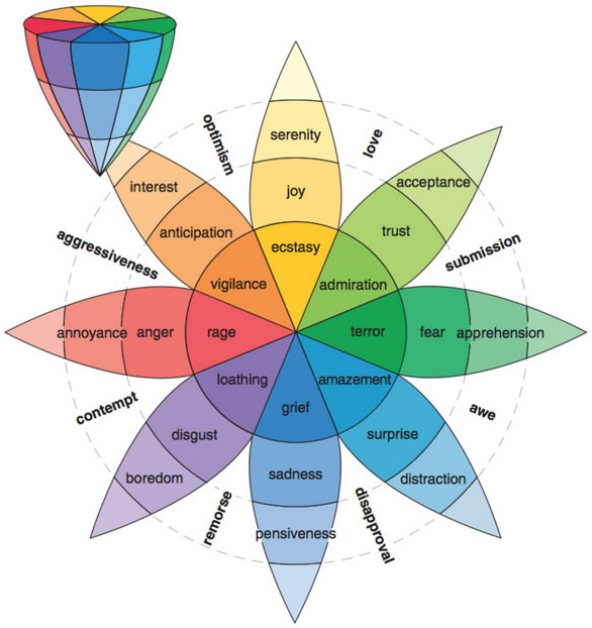

Figure 1: The structure of Plutchik's Wheel of Emotions (Plutchik, 2001). Reprinted by permission of American Scientist, magazine of Sigma Xi, The Scientific Research Honor Society.

Recent emotion classification studies are not based on such psychological hypotheses. Instead, emotions are assumed to be distributed in semantic space, and they classified emotions statistically based on large data (Cowen & Keltner, 2021). Cowen and colleagues reported that they could distinguish 27, 13, and 28 different emotional experiences in short videos, music, and facial expressions, respectively (Cowen & Keltner, 2017; Cowen, Fang, Sauter, & Keltner, 2020; Cowen & Keltner, 2020). Based on these studies, Demszky and colleagues selected 27 basic emotions that can cover emotions discovered in various domains, and released GoEmotions, a largescale text-emotion multi-label dataset through crowdsourcing (Demszky et al., 2020). The study found that the basic emotion list based on the semantic space assumption is still valid for classifying emotions in text. Finally, they analyzed the similarity between emotions and clustered the overall emotions into positive/ambiguous/negative, and reported the polarity of the emotions they used.

Emotion Classification in NLP Field

In the field of Natural Language Processing (NLP), with the advancement of methods for converting text into vectors, there has been growing efforts to analyze the emotions of words or sentences. With the rise of people's emotional expressions on social media and the ease of building large dataset through crowdsourcing, text-emotion dataset studies have been published. Most of these datasets provide labels based on Ekman's basic emotions, making it easier for labelers to distinguish each emotion (Strapparava & Mihalcea, 2007; Mohammad, 2012; Li et al., 2017).

In emotion detection/classification studies, Plutchik's model which can encompass more diverse emotions was preferred. Mondal and Gokhale (2020) built a classifier that categorizes which emotion pair in Wheel of Emotions corresponds to the given tweet, and what its polarity is. Kumar and Vardhan (2022) proposed a model that detects the emotions of tweets by utilizing a word embedding method and synonyms of the basic emotions of the wheel of emotions.

However, discussions on the emotional classification criteria themselves are limited in the NLP field. Bann and Bryson (2014) attempted to identify the emotions embedded in Twitter texts using clustering. The study reviewed the basic emotions in several basic emotion literatures and found that clustering based on Ekman's theory was most appropriate. However, adding a few emotions not included in Ekman's basic emotions resulted in a better clustering outcome. Jeon, Lee, and Kim (2022) attempted to identify emotions in Korean text based on the semantic space hypothesis proposed by Cowen and Keltner (2021). They analyzed the distribution of 1,787 emotion words in the vector space of a word embedding model using a clustering ensemble. As a result, they claimed that 43 emotions were classified in Korean text.

Based on these discussions, we suspect that it may not be appropriate to directly use Ekman's or Plutchik's theories as criteria for emotion classification. In fact, some studies that built a text-emotion dataset modified Ekman's basic emotions to successfully reflect the context of the data. (Alm, Roth, & Sproat, 2005; Lyu, Chen, Wu, & Wang, 2020). Thus, it is necessary to re-examine the basic emotions presented by prior studies.

Methods

According to basic emotion theory, emotions are not independent states but rather families of related states. Emotions within each family should share the same characteristics and be different from each other (Ekman et al., 1999). From this perspective, the adequacy of an emotion model can be examined by analyzing the similarity between the basic emotions proposed by each emotion theory. We applied sentence embedding methods to calculate the similarity between the words in the three emotion models.

Emotionally Fine-tuning a Language Model

Typically, a language model converts the contextual meaning of an input word or sentence into a vector. However, in order to compute the emotional meaning of a word, the vectors need to be rearranged according to their emotional meaning. Because all emotional words have an emotional meaning in common, thus they are located in a relatively similar vector space compared to those words that do not have emotional meaning. We solve this problem by fine-tuning the model in such a way that the similarity between the embedding vectors of two words is as close as the emotional similarity of the two words.

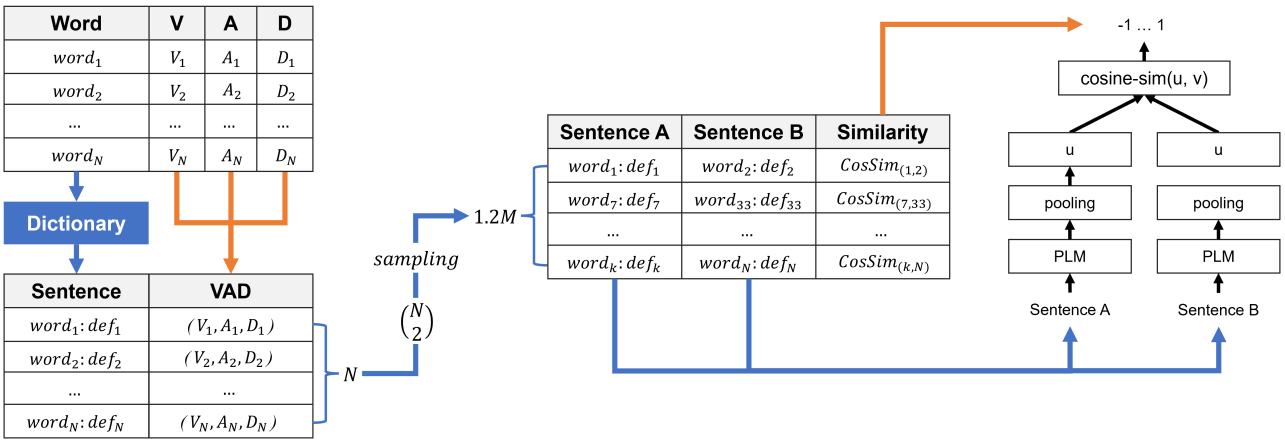

Figure 2: The procedures of dataset construction and model fine-tuning. ($N = 15,778$)

One of the effective methods of transforming a sentence or sequence of words into a vector representation is to use the Sentence-transformers model (Reimers & Gurevych, 2019). This model can be fine-tuned with NLI and STS tasks. The STS task is to predict the semantic similarity between two input sentences on a scale of $-1 \sim 1$. We adapted this task to create our task that predicts the emotional similarity between two input words. However, one limitation of this method is that it may not accurately represent words that were infrequent during pre-training. Some of the emotion names used in basic emotion theories were less frequently used than others in everyday life. To address this limitation and enrich the contextual meaning of words, we included dictionary definitions of emotion names as an input to the language model. By fine-tuning the model for this task, we ensure that the vectors of input words (and their dictionary definitions) is embedded according to their emotional meaning (Figure 2).

Dataset

In emotional psychology, the three-factor theory of emotions describes the emotional values of a given word as a three-dimensional vector (Russell, 1980). Based on this theory, several datasets have been built in various languages. The largest of these in English is the NRC-VAD Lexicon[1] built by Mohammad (2018). This dataset was built by aggregating English expressions used in previous affective lexicon studies and adding Valence, Arousal, and Dominance (V, A, D) annotations for 19,970 English expressions.

We added a dictionary definition to each word in this dataset. The dictionary we used was the Oxford Dictionary, and the search platform we used was the NAVER English Dictionary[2]. From the word list, we excluded grammatically incorrect words (`aaaaaaah`, `hateeee`, `yessss`), conjugations not found in the dictionary (`concurring`, `difficulties`, `lucky`), and phrases or sentences (`canthandleit`, `imissyou`, `somethingigetalot`). We also excluded emotion words from Ekman's Basic Emotions, Plutchik's Wheel of Emotions, and GoEmotions to ensure that the words to be evaluated were not directly learned during the fine-tuning process. After preprocessing, 15,778 words were left. To organize the data according to the format of the STS task, we selected two random words from the word list. The two sentences consisted of the selected words and their definitions (*word: definition*), and the similarity between the two sentences was defined by the cosine similarity of the VAD values of two words, $CosineSimilarity(\ (V_A, A_A, D_A),\ (V_B, A_B, D_B)\ )$. The total number of pairwise combinations in the 15,778 words was 124 million, so not all of the data could be used for fine-tuning. We randomly sampled from all possible word combinations to build 1 million training set and 0.2 million validation set.

Fine-tuning and Evaluation

Following the training method of the Sentence-transformers model in the STS task, we set two words and their definitions as the input of the model and the emotional similarity $CosSim_{(A,B)}$ of the two words as the target of CosineSimilarityLoss. The base language model of Sentence-transformers was `all-mpnet-base-v2`, which is known to be the best performer for general purposes[3]. Our model was trained for 1 epoch with a batch size of 32 and a learning rate of 3e-7.

We used the fine-tuned model to perform analysis on the three basic emotion theories. The emotions named by each theory were annotated with dictionary definitions of the words using the Oxford Dictionary, as was the training dataset. The definition of `pensiveness` is not in the Oxford dictionary, so we used the definition from Wiktionary[4]. Plutchik's theory of emotions claims that additional emotions arise from the combinations of eight emotional axes (Plutchik, 2001), but the present study selected only the eight emotions proposed as basic emotions and their variation emotions (mild and intense). In the case of GoEmotions,

[1] `https://saifmohammad.com/WebPages/nrc-vad.html`
[2] `https://english.dict.naver.com/english-dictionary`
[3] `https://www.sbert.net/docs/pretrained_models.html`
[4] `https://en.wiktionary.org/wiki/Wiktionary:Main_Page`

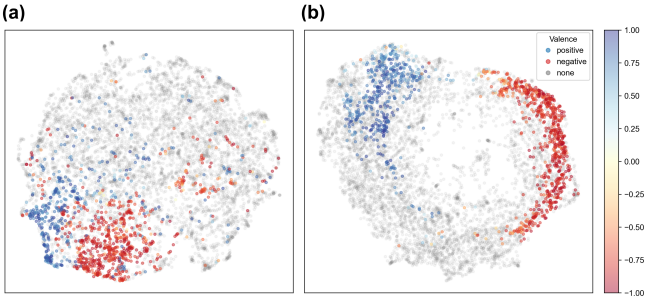

Figure 3: The distribution of embedding vectors of words. Distribution of embedding vectors before (a) and after (b) fine-tuning.

we did not use the dictionary definitions of the words, but the descriptions of the words presented to the labelers during the labeling process in the original paper (Demszky et al., 2020). From the dictionary definitions of each word, we excluded non-emotional connotations (interest: *interest rates*, ecstasy: *drugs*, etc.). For each theory, we constructed a similarity matrix by computing the cosine similarity of two words for each pair of combinations of emotion words. This similarity matrix was created before and after fine-tuning the language model to evaluate the performance of fine-tuning. And the values of the similarity matrix were converted into cosine distances in order to cluster the emotions of each theory by hierarchical clustering. The distances between the clusters were calculated based on the `ward`. The clusters containing *Joy (Enjoyment)* and *Sadness*, which are the best representatives of positive/negative emotions, were designated as positive/negative clusters, respectively, and the other clusters were treated as ambiguous clusters to classify the whole emotion word list into positive/ambiguous/negative.

Result and Discussion

Emotional Fine-tuning

The distribution of embedding vectors of words before and after fine-tuning is shown in Figure 3. Emotional words ($\sqrt{V^2+A^2+D^2} > 1$, $n = 1{,}245$) and non-emotional words ($\sqrt{V^2+A^2+D^2} < 0.5$, $n = 6{,}496$) were extracted from the dataset. Before fine-tuning, emotional words were densely distributed in the vector space, but after fine-tuning, they were found to be distributed widely in the space. Also, before fine-tuning, positive and negative words were in the same direction based on the center of the distribution. After fine-tuning, they were found to be distributed in opposite directions. Through this, It is shown that the emotional fine-tuning of the language model was carried out appropriately as we expected.

Ekman's Basic Emotions

The result of the Ekman's seven basic emotions is shown in Figure 4. Comparing the similarity matrix before and after fine-tuning, the cosine similarities between `enjoyment`, a positive emotion, and `fear`, `sadness` and other negative emotions has changed from positive to negative. This suggests that the fine-tuning performed in this study helped create a closer representation to Ekman's. `Anger`, `contempt`, `disgust`, and `fear` are highly similar to each other, and the similarity between `disgust` and `contempt` is the highest (0.96). The similarities between `sadness` and other emotions were relatively low, but showed a similar overall pattern to the aforementioned negative emotions. `Surprise` has a moderate similarity to most of the emotions. In the clustering result, `disgust` and `contempt` clustered first. Given that the two emotions were highly alike in the similarity matrix, this was consistent with the argument that contempt is not a distinct emotion from disgust (Russell, 1991; Ekman et al., 1991), and with recent research showing that disgust and contempt are very closely related (Miceli & Castelfranchi, 2018). These results are in line with extant emotion research. We categorized `enjoyment` and `surprise` as positive and ambiguous emotions, respectively, considering the polarity of each emotion.

Ekman's basic emotions consist of a relatively small number of emotions, which makes the results of the analysis clear. The theory also consists of only those emotions that can be distinguished by facial muscle movements. Nevertheless, the finding that our analysis using word meanings produced similar results suggests that Ekman's theory is quite robust.

Plutchik's Wheel of Emotions

The result of Plutchik's Wheel of Emotions is shown in Figure 5. The emotions are listed clockwise, starting with `joy` in the Wheel of Emotions (Figure 1), and within the same emotion group, the order is `mild-basic-intense` (e.g., `joy` group = `serenity-joy-ecstasy`). Comparing the similarity matrix before and after fine-tuning, it is shown that the similarity within each emotion group has increased and the similarity between groups has decreased. This is particularly evident in the `trust`, `fear`, and `anticipation` groups. On the other hand, some emotions did not show strong similarities within groups. `Ecstasy` in the `joy` group, `distraction` in the `surprise` group, and `pensiveness` in the `sadness` group showed low within-group similarity. According to Plutchik's theory, emotions in the same group should be highly similar because they are the same emotions with different intensities. However, emotions such as `distraction` and `pensiveness` do not meet this assumption.

The clustering result shows even more discrepancies between the assumptions of the theory and the result of our analysis. According to the assumptions of Plutchik's theory, emotions on each emotion axis should be clustered first, followed by emotions on adjacent axes, and finally symmetrical cross-axis emotions. However, the hierarchical clustering result in this study contradicts these assumptions entirely. The `interest` group and the `trust` group clustered relatively close to Plutchik's theory. However, although `fear` and `anger` are opposite emotions in Plutchik's theory, the clustering result showed that `rage-anger-terror` and `annoyance-fear` were close to each other. The three

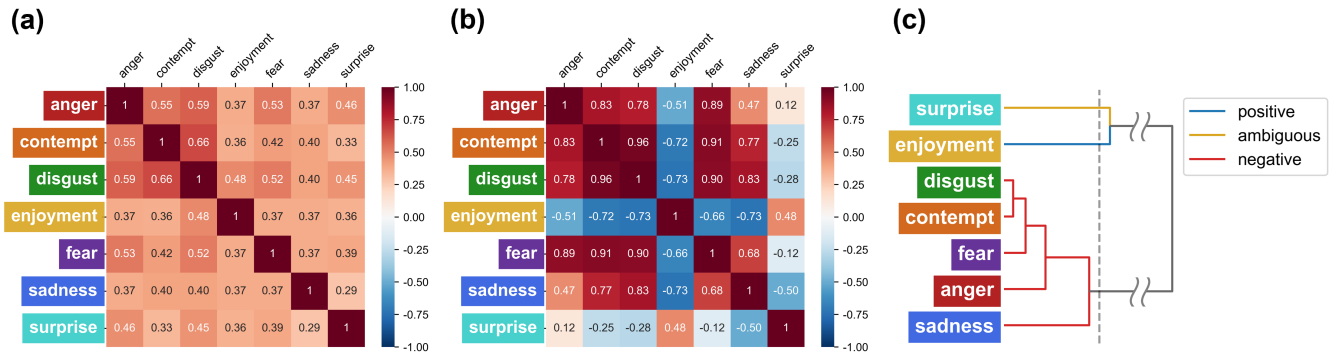

Figure 4: The result of Ekman's Basic Emotions. (a): The similarity matrix with the raw model. (b): The similarity matrix with the fine-tuned model. (c): The result of hierarchical clustering.

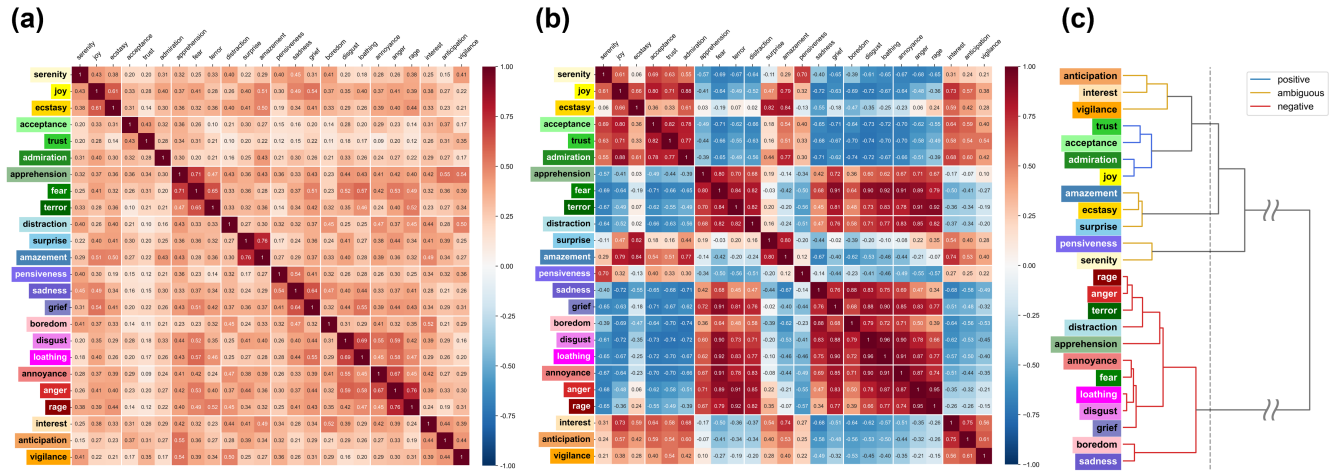

Figure 5: The result of Plutchik's Wheel of emotions. (a): The similarity matrix with the raw model. (b): The similarity matrix with the fine-tuned model. (c): The result of hierarchical clustering.

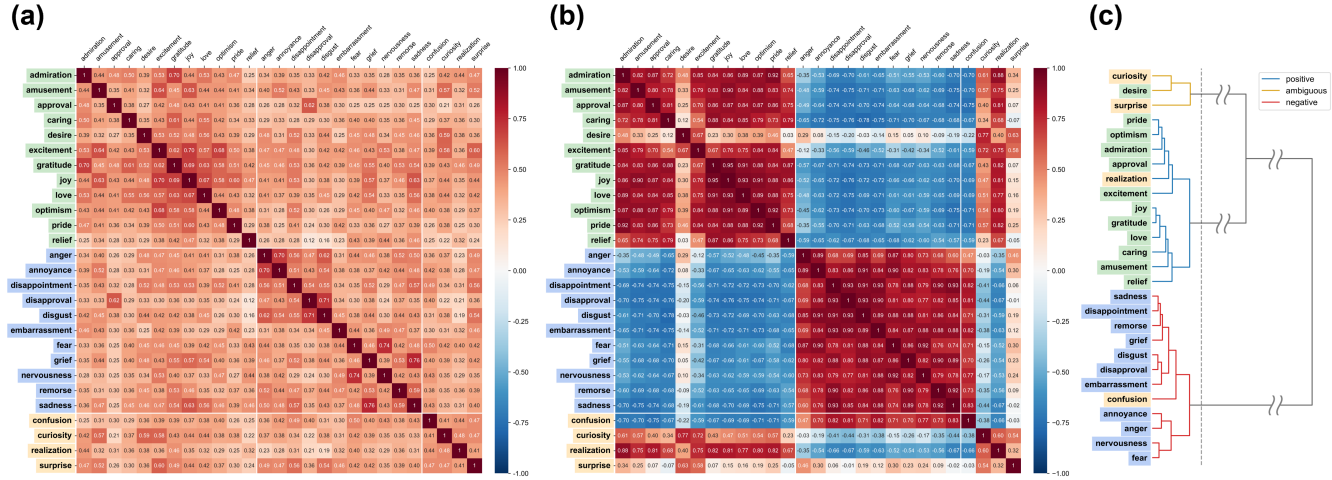

Figure 6: The result of GoEmotions. (a): The similarity matrix with the raw model. (b): The similarity matrix with the fine-tuned model. (c): The result of hierarchical clustering.

emotions in the joy group were all assigned to different clusters. Among the clusters categorized as ambiguous, two clusters, amazement-ecstasy-surprise and pensiveness-serenity, seem to be clustered according to the emotional intensity of the words. The dictionary used in the analysis provided the following definitions: surprise as *"an unexpected or **astonishing** event, ... to feel mild **astonishment** or **shock**"*, ecstasy as *"an **overwhelming** feeling of **great** happiness or joyful **excitement**; an emotional or religious **frenzy** or ..."*, and amazement as *"a feeling of **great** surprise or wonder"*. These words had expressions that mean high arousal in their definitions. On the other hand, serenity was defined as *"the state of being **calm**, **peaceful** and untroubled"* and pensiveness was defined as *"A thoughtful or reflective state, especially if sad or **melancholic**"* with low arousal. Ecstasy is an intense emotion of joy, a typical positive emotion, while pensiveness is a mild emotion of sadness and even has *sad* and *melancholic* in its definition. Despite this, the clustering results suggest that these emotions cannot be clearly categorized as either positive or negative. Therefore, we consider them to be ambiguous emotions.

These results suggest that Plutchik's proposed structure of emotions is not consistent with the structure of actual emotions. Plutchik's assumptions of eight emotional axes may not be consistent with the human emotional structure, especially those opposite to each other. In the case of rage-anger-terror and annoyance-fear, emotions that should have opposite emotional meanings according to Plutchik's theory were clustered in the same clusters. It is also possible that each emotion was misnamed. Since our analysis relies on the names of the emotions used in each theory and their dictionary definitions to estimate their affective value, it is possible that they misnamed actual emotions and did not match the emotion to its name. The negative connotation of distraction compared to surprise and amazement makes it difficult to consider distraction-surprise-amazement as different intensity emotions on the same emotional axis. The three emotions that were assigned to different clusters, serenity-joy-ecstasy, are also examples of emotion-word matching failures. In addition, the distance between the three emotions separated by mild-basic-intense may be different for each emotion axis. For interest-anticipation-vigilance, the distances among the three emotions seem to be reasonable ($0.56 \sim 0.75$), but for disgust-loathing (0.96) and anger-rage (0.95), the distance between the basic-intense emotions seems to be closer than for the anticipation group. In conclusion, Plutchik's emotion model assumes that emotions are equally spaced on eight axes with constant angles, but actual emotions could be placed at unevenly spaced distances on uneven dimensional axes.

GoEmotions

We analyzed GoEmotions with our methodology and got the following results: Figure 6. First, the similarity matrix was consistent with the theory. GoEmotions categorizes 27 emotions into positive, negative and ambiguous. Our result was almost the same. Among the positive emotions, desire showed a different similarity pattern than other positive emotions, while among the ambiguous emotions, confusion showed a similar pattern to negative emotions and realization showed a similar pattern to positive emotions. These patterns were reflected in the clustering result. Desire, which was classified as a positive emotion in GoEmotions, was clustered with curiosity and surprise as an ambiguous emotion. Realization and confusion, which were classified as ambiguous emotions in GoEmotions, were clustered as positive and negative emotions, respectively. In the case of realization, it was found to be the closest to approval in the GoEmotions' analysis, but it was classified as ambiguous rather than positive by the researchers. In our analysis, realization clustered as a positive emotion close to approval, once again. Therefore, it seems appropriate to consider realization as a positive emotion rather than an ambiguous emotion. They also reported that emotions with similar meanings and different intensities appeared in close proximity to each other. In our results, joy-excitement clustered far away from each other, which was different from their results, but sadness-disappointment, fear-nervousness were the same as their results, and disappointment-sadness-grief, disgust-annoyance-anger were similar to their results. These findings may reflect that GoEmotions is a data-driven emotion model. The proposed structure of this model, which is built by human labeling of the emotions expressed in sentences, was quite consistent with our results using a language model. This is probably due to the similarity of people's emotion representations and the embedding vectors of the emotionally fine-tuned language model we used.

General Discussion

In the case of Ekman's basic emotions and GoEmotions, we found good agreement between the proposed theories and our analysis. However, in the case of Plutchik's Wheel of Emotions, we found a large gap between the theory's prediction and our results. This may be due to the fact that Plutchik tried to categorize emotions in terms of human instincts and motivations, and assign each emotion to an eight axes structure, creating a gap between theory and practice. In contrast, in the case of GoEmotions, which is based on data, our results closely match the theory. While many studies have used Ekman's or Plutchik's theories as a basis for emotion classification or as a label for data, our results suggest that using the emotion structure of semantic space theory may be a more appropriate choice in the field of data science.

Acknowledgments

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(NRF-2022R1A6A3A13071841)

References

Alm, C. O., Roth, D., & Sproat, R. (2005). Emotions from text: machine learning for text-based emotion prediction. In *Proceedings of human language technology conference and conference on empirical methods in natural language processing* (pp. 579–586).

Bann, E. Y., & Bryson, J. J. (2014). The conceptualisation of emotion qualia: Semantic clustering of emotional tweets. In *Computational models of cognitive processes: Proceedings of the 13th neural computation and psychology workshop* (pp. 249–263).

Cowen, A. S., Fang, X., Sauter, D., & Keltner, D. (2020). What music makes us feel: At least 13 dimensions organize subjective experiences associated with music across different cultures. *Proceedings of the National Academy of Sciences*, *117*(4), 1924–1934.

Cowen, A. S., & Keltner, D. (2017). Self-report captures 27 distinct categories of emotion bridged by continuous gradients. *Proceedings of the national academy of sciences*, *114*(38), E7900–E7909.

Cowen, A. S., & Keltner, D. (2020). What the face displays: Mapping 28 emotions conveyed by naturalistic expression. *American Psychologist*, *75*(3), 349.

Cowen, A. S., & Keltner, D. (2021). Semantic space theory: A computational approach to emotion. *Trends in Cognitive Sciences*, *25*(2), 124–136.

Darwin, C., & Prodger, P. (1998). *The expression of the emotions in man and animals* (3rd ed.). Oxford University Press, USA. (Original work published 1874)

Demszky, D., Movshovitz-Attias, D., Ko, J., Cowen, A., Nemade, G., & Ravi, S. (2020). Goemotions: A dataset of fine-grained emotions. *arXiv preprint arXiv:2005.00547*.

Ekman, P. (2003). *Emotions revealed : recognizing faces and feelings to improve communication and emotional life*. New York: Times Books.

Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of personality and social psychology*, *17*(2), 124.

Ekman, P., & Friesen, W. V. (1986). A new pan-cultural facial expression of emotion. *Motivation and emotion*, *10*(2), 159–168.

Ekman, P., & Heider, K. G. (1988). The universality of a contempt expression: A replication. *Motivation and emotion*, *12*(3), 303–308.

Ekman, P., O'Sullivan, M., & Matsumoto, D. (1991). Contradictions in the study of contempt: What's it all about? reply to russell. *Motivation and Emotion*, *15*(4), 293–296.

Ekman, P., et al. (1999). Basic emotions. *Handbook of cognition and emotion*, *98*(45-60), 16.

Jeon, D., Lee, J., & Kim, C. (2022). User guide for kote: Korean online comments emotions dataset. *arXiv preprint arXiv:2205.05300*.

Kumar, P., & Vardhan, M. (2022). Pwebsa: Twitter sentiment analysis by combining plutchik wheel of emotion and word embedding. *International Journal of Information Technology*, 1–9.

Li, Y., Su, H., Shen, X., Li, W., Cao, Z., & Niu, S. (2017, November). DailyDialog: A manually labelled multi-turn dialogue dataset. In *Proceedings of the eighth international joint conference on natural language processing (volume 1: Long papers)* (pp. 986–995). Taipei, Taiwan: Asian Federation of Natural Language Processing. Retrieved from https://aclanthology.org/I17-1099

Lyu, X., Chen, Z., Wu, D., & Wang, W. (2020). Sentiment analysis on chinese weibo regarding covid-19. In *Natural language processing and chinese computing: 9th ccf international conference, nlpcc 2020, zhengzhou, china, october 14–18, 2020, proceedings, part i 9* (pp. 710–721).

Miceli, M., & Castelfranchi, C. (2018). Contempt and disgust: The emotions of disrespect. *Journal for the Theory of Social Behaviour*, *48*(2), 205–229.

Mohammad, S. (2012, 7-8 June). #emotional tweets. In **SEM 2012: The first joint conference on lexical and computational semantics – volume 1: Proceedings of the main conference and the shared task, and volume 2: Proceedings of the sixth international workshop on semantic evaluation (SemEval 2012)* (pp. 246–255). Montréal, Canada: Association for Computational Linguistics. Retrieved from https://aclanthology.org/S12-1033

Mohammad, S. (2018). Obtaining reliable human ratings of valence, arousal, and dominance for 20,000 english words. In *Proceedings of the 56th annual meeting of the association for computational linguistics (volume 1: Long papers)* (pp. 174–184).

Mondal, A., & Gokhale, S. S. (2020). Mining emotions on plutchik's wheel. In *2020 seventh international conference on social networks analysis, management and security (snams)* (pp. 1–6).

Plutchik, R. (1980). A general psychoevolutionary theory of emotion. In *Theories of emotion* (pp. 3–33). Elsevier.

Plutchik, R. (1982). A psychoevolutionary theory of emotions. *Social Science Information/sur les sciences sociales*.

Plutchik, R. (2001). The nature of emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice. *American Scientist*, *89*(4), 344–350. Retrieved 2023-01-27, from http://www.jstor.org/stable/27857503

Reimers, N., & Gurevych, I. (2019). Sentence-bert: Sentence embeddings using siamese bert-networks. *arXiv preprint arXiv:1908.10084*.

Russell, J. A. (1980). A circumplex model of affect. *Journal of personality and social psychology*, *39*(6), 1161.

Russell, J. A. (1991). The contempt expression and the relativity thesis. *Motivation and emotion*, *15*(2), 149–168.

Strapparava, C., & Mihalcea, R. (2007, June). SemEval-2007 task 14: Affective text. In *Proceedings of the fourth international workshop on semantic evaluations (SemEval-2007)* (pp. 70–74). Prague, Czech Republic: Association for Computational Linguistics. Retrieved from

https://aclanthology.org/S07-1013