A Fuzzy Logic System to Encode Emotion-Related Words and Phrases

Author: Abe Kazemzadeh
Contact: kazemzad@usc.edu
class: EE590 Fuzzy Logic
professor: Prof. Mendel
Date: 2007-12-6

Abstract:
This project looks at using a fuzzy logic system (FLS) to represent the meaning in words that refer to emotions. To do this we take a computing with words (CWW) approach. The CWW approach applied to emotion words represents each emotional term as a vocabulary where each word in the vocabulary corresponds to a fuzzy set. This paper experiments using a FLS to relate the emotion terms to a dimensional representation of emotions. The two emotional words we consider are general emotional category terms (angry, disgusted, fearful, happy, neutral, sad, and surprised). These are presented as stimuli to 16 subjects, who use sliders to select ranges on the three scales of a dimensional representation of emotions: valence, activation, and dominance (VAD). In addition to the basic emotion words, we also present phrases using these words, such as “very X”, “sort of X”, and “not X”, to see how these modifying words can function as logical operators. The experimental data collected is analyzed using the Interval Approach (IA). We consider the results and come to the conclusion that the trends are promising. Finally we speculate on possible uses of such a FLS.

Introduction:
This paper presents a series of studies that look at using fuzzy logic systems (FLS) to represent emotions. Before going into the details, we should consider the rationale of the endeavor in general. Why would fuzzy logic be appropriate for representing emotions? One of the motivations is that emotions are difficult to measure precisely, yet still people can talk about being “slightly angry” or “very angry”, which leads one to infer some underlying scales. These types of judgments are a good fit for fuzzy sets, where an input can have a gradient membership in a set along some axis. Therefore, fuzzy logic can allow us to talk about a person, an animal, a document, or a song as having a degree of anger, happiness, sadness, and the like, while still allowing the notion of categories of emotions. Furthermore, different people may mean different things when they use such emotional descriptions. Fuzzy logic has methods for dealing with this phenomenon as well. The computing with words (CWW) framework treats these intra- and inter-personal differences by collecting data and creating a different person specific membership function (MF) [1]. Emotions present a further problem, namely that the vocabulary for describing emotions is vast. Attempts to make a list of emotional words is difficult for the following reasons. (1) Many words may have emotional connotations and can name very specific examples of emotions (e.g. the feeling of “abandonment”, a sad emotion in the context of being abandoned). Enumeration of these can be infeasible because exhaustively listing them would be too numerous and not permit generalization. (2)
Attempting to reduce the number of emotional terms will necessarily involve some implicit adoption of a specific theory of emotions. A number of these have been proposed by researchers, some dealing with physiological manifestations of emotions, some dealing with cognitive manifestations of emotions, some dealing with social manifestations of emotions. In this paper, we look at two systems of describing emotions. In the categorical description of emotions, we have a set of 7 different emotion categories (angry, fearful, happy, neutral, sad, and surprised). In the dimensional description of emotions, we have emotions represented by points in a space of valence, activation, and dominance (VAD), where valence is the positive/negative or pleasurable/not pleasurable axis, activation is the strength/weakness of the emotion, and dominance is the assertiveness/submissiveness of the emotion. It is possible to give a dimensional interpretation of a categorical description of an emotion. For example, sadness can be thought of as having a negative valence and a low activation, happiness can be thought of as having positive valence and high activation, and low dominance, and anger can be thought of as having negative valence, high activation, and high dominance.

Furthermore, if the categories can be taken as elements in a vocabulary, it should also be possible to construct operations with these words. As a test of this idea, we decided to consider “very X”, “sort of X”, and “not X”, where X is one of the seven emotion categories. Other operations such as adding two emotions are speculated about in the conclusion.

The proposal of this paper is to encode such natural language descriptions of emotions as a FLS. This will allow us to map from one way of describing emotions to another. The FLS will ideally allow us to model nonlinear mappings, uncertainty, and person-specific differences. It is hoped that establishing a relation between these two ways of describing emotions will facilitate the incorporation of more open vocabularies by allowing new words to be defined in terms of either of the two considered here.

**Explanation of the Application**

This project mainly considers the encoding of a FLS. The application we consider is to take a word or phrase and encode it into an interval type 2 fuzzy set. In [2], a similar problem was considered, where the goal was to encode words referring to quantities in terms of a fuzzy set with a membership function over the domain of numbers between 0 and 10. Here, we extend the idea to emotions by positing a scale for 3 dimensions used to describe emotion: valence, activation, and dominance. As such, this can be thought of as three separate encoders, one for each scale. This conception may be more convenient when dealing with rule-based systems, so that each of the 3 dimensions can trigger rules separately. Like in [2], we also arbitrarily adopt the scale from 0 to 10 for each of the 3 dimensions.

One can visualize such a system like so:

Word -> [Encoder] -> VAD representation membership function

Or in terms of 3 separate encoders:

Word -> [Valence Encoder] -> Valence membership function
Word -> [Activation Encoder] -> Activation membership function
Word -> [Dominance Encoder] -> Dominance membership function

To begin studying operations on words, it can be hypothesized that an operation on a word will change the words membership function. To test this hypothesis, we look at modifying the seven basic emotion category words with “very”, “sort of”, and “not”. To make the hypothesis more precise, we can say that “not” will be have like the compliment operation, and that “very” and “sort of” might behave as a fuzzy set under a mapping function as described in [3] chapter 1.10.13, where the Extension Principle [4] is used to calculate such a mapping function.
Explanation of the Data

The data used in this experiment was obtained through an online survey taken by 14 subjects. Each subject was presented with a series of stimuli that consisted of 7 emotion category words (angry, disgusted, fearful, happy, neutral, sad, and surprised) and compound phrases made from these words and 3 modifiers (sort of X, very X, and not X, where X is one of the aforementioned 7 emotion category words). For each subject, the seven basic emotion category words were repeated 3 times while the phrases were only presented once. Below the presented word(s), were sliders for each of the 3 emotional dimensions (valence, activation, and dominance). The sliders had 2 handles, which allowed the subjects to select ranges. The range of the sliders was 0-10, which was the maximum allowable range that a user could select. The minimum range allowed was 1, since the steps were integer values and the implementation imposed a constraint that the upper and lower endpoint could not be the same. The instructions and a screen shot of the user interface are provided in the appendix.

The stimuli were presented in 3 groups. First the basic emotion category words were presented. Then a mixture of the basic emotion words and the compound words were presented. And finally the basic emotion category words were presented another time. The rationale for presenting the basic emotions first was to account for the fact that users would tend to normalize their responses over time. This also allowed for more data to be collected for the basic emotion category words. Within each of the 3 groups the stimuli were randomized using Fisher-Yates shuffling method.

The interface was implemented using a general LAMP methodology (Linux, Apache, MySQL, and Perl), with the slider functionality implemented using the Javascript library JQuery UI. The link to the experiment web page is [5].

FLS Description

To apply the CWW methodology to emotional words, I use the Interval Approach (IA) [2]. This primarily deals with the encoding component of a FLS. This approach uses interval end-point data for a word, collected from a group of subjects, to map each subjects data into a type-1 person membership function (MF) and then use the collection of all subject’s person MFs as embedded MFs in an interval type-2 fuzzy set.

The IA approach is divided into two main parts, the data part and the fuzzy set part. The data part preprocesses the data to remove outliers and compute statistics of the collected data. The fuzzy set part then uses the data statistics and instances remaining after preprocessing to create person MFs, aggregates them, establishes whether they are interior or shoulder functions, and then finally creates a mathematical model for the given word for which end-point data was collected. The Matlab files used to carry out this analysis is based on [6].

The IA approach offers several advantages over other methods, such as the person membership function and the interval endpoints approach [1]. It does not require experts who are knowledgeable in fuzzy logic as in the person membership function approach. Rather it collects data from subjects with minimal training. It does not require specifying whether the MF should be symmetrical or asymmetrical, shoulder or interior. It is a straightforward approach that goes sequentially from data to membership functions and if there is no uncertainty, the IA approach will result in type-1 membership functions.

Results

Basic Emotion Category Words

In the following figures we can see the results from the basic emotion category words. The figure 1 shows all the results of the 7 emotion category words together. Figure 2 shows that uncertainties in the footprints of uncertainty (FOUs) are smaller for some of the membership functions when the first set of stimuli were removed. This shows that it took some learning for the subjects to normalize their responses.
Another thing to notice is that the membership functions of “angry” and “disgusted” closely resemble each other, as do “happy” and “surprised”. This implies that a reverse mapping from the VAD representation to emotional category words would run into problems. This problem can be thought of as the decoding process or a classification into emotion categories based on VAD features.

Two of the subjects completed the survey twice. The membership functions created from their data is shown in figures 3 and 4. From these, we can see that given enough responses, subjects will display intra-person variation. The amount of variation could be a person specific trait, and may also relate to the time in between experiments. It can also be seen that for these two subjects, the intra-uncertainty is much smaller than the inter-uncertainty seen in figures 1 and 2. In fact, some of the membership functions in figures 3 and 4 reduce to type 1.

Another difference between these subjects and the aggregate results is that the subject in figure 3 does not have similar membership functions for “angry” and “disgusted” or “happy” and “surprised”. Rather the membership functions for “disgusted” and “fearful” are confusable for this subject. The subject in figure 4 only shows the similarity of “happy” and “surprised”.

### Compound Emotion Phrases

In figures 5-11, the results from the compound emotion phrases are presented. There are a number of different effects that the modifying words have on the base forms. In general the FOU of the phrased is smaller compared to the base words. “Very” had the general tendency to make the membership function steeper and more peripheral, while “sort of” made the membership function less steep and more central. “Not” tended to shift the membership function to the opposite side of at least one of the scales.

“Not neutral” failed to come up with a membership function for valence and dominance. The fact that neutral can be thought of the absence of an emotion rather than an emotion per se seems to give some explanation for this difference in behavior.

Another noticeable trend is that for the basic emotion words, there seems to be more shoulder membership functions. This could be because the undifferentiated emotion category word represents
Figure 2: These are the membership functions generated by IA for the emotion category words excluding the first round of stimuli. This had an effect of slightly decreasing the uncertainty.

Figure 3: These are the membership functions for one subject, who did the survey multiple times. There is still a fair amount of uncertainty within the subject.
Figure 4: These are the membership functions for another subject who did the survey multiple times. There is less uncertainty for this subject.

Discussion

It is instructive to observe the sources of uncertainty in this methodology to see what the implications are for building upon these results. From the first two figures, we can see that there is some uncertainty among the subjects as they “warm up” and normalize their responses based on their previous ones. The interface lends itself to this uncertainty because most of the subjects were not familiar with interval selection, so they had to take a survey that was less familiar to them than typical Likert surveys. Also, subjects may not be used to thinking of emotions in such terms.

Another source of uncertainty is in the words and concepts themselves. Emotion words can be very general. Although they are coherently understandable, they are not very precise. The phrases tended to decrease the FOU and make the range of the membership function smaller. One interpretation of this is that the phrases act as operators that modify the basic emotions. Another interpretation is that the basic emotions are just aggregates of more specific emotion words.

Another source of uncertainty is in the fact that words mean different things to different people. It was seen that the membership functions of individuals were smaller and had less uncertainty than the membership functions for all the subjects combined.

One difference between this study and [2] is that here there was a step size of 1 enforced on the slider so that all values were integers. In the data of [2] (provided at [6]), any values were acceptable if they were between 0 and 10. Enforcing this and the 0-10 limit is easy to do with computer-based surveys, but a possible drawback is that the subjects who goof around are harder to spot when the range is constrained automatically.

One possible shortcoming of the interval approach is that there might be times when the membership
Figure 5: Membership functions for “very angry”, “sort of angry”, and “not angry” for all subjects.
Figure 6: Membership functions for “very disgusted”, “sort of disgusted”, and “not disgusted” for all subjects.
Figure 7: Membership functions for “very fearful”, “sort of fearful”, and “not fearful” for all subjects.
Figure 8: Membership functions for “very happy”, “sort of happy”, and “not happy” for all subjects.
Figure 9: Membership functions for “very neutral”, “sort of neutral”, and “not neutral” for all subjects.
Figure 10: Membership functions for “very sad”, “sort of sad”, and “not sad” for all subjects.
Figure 11: Membership functions for “very surprised”, “sort of surprised”, and “not surprised” for all subjects.
function should be everything outside a given range. An example of this is “not neutral”. For this, it would be most logical to think of it as the opposite of the range for neutral, i.e. everything outside of mid-range.

One implication of the results pertains to the idea of using a controlled vocabulary as a codebook for efficiently constructing a small rule set in a CWW FLS. The results show that the basic emotion category words do cover the VAD space. However, there is a lot of uncertainty in the membership functions. The more specific phrases have less uncertainty but, since there are more of them, they would lead to more rules. This trade-off between uncertainty and concise representation is a design decision that would have to be made later on.

Another implication of the results is that could be to use different scales than the VAD dimensions. It was seen in the results for the basic emotion category words that the membership functions for “angry” and “disgusted” closely resembled each other as did “happy” and “surprised”.

Conclusion

This project presents a way to encode emotional words into a fuzzy logic system using the Interval Approach. The results show general trends that could be of use in the Computing With Words framework. In particular, though the words by themselves resulted in large FOUs, the phrases exhibited less uncertainty. These results can help design codebook-like vocabularies for using emotional terminology in FLSs.

Some further directions for this research project could be to use a more open vocabulary as an input. The results showed that more precise phrases showed less uncertainty and sharper membership functions. The larger vocabulary could be mapped to a more controlled codebook. This could offer subjects of the experiment more challenge, and if such a FLS is implemented in an application then users would have more freedom in expressing themselves.

Another direction would be to show stimuli that combines emotion words, asking users to imagine an emotion that is, say, both angry and sad (“bitter” perhaps), both angry and happy (maybe “sarcastic”), or both happy and sad (which could be “nostalgic”). It would be interesting to see if such “addition” operations behave regularly considering the fuzzy set representation of the components.

Another direction would be to use other media as stimuli. This could enable CWW applications that use words as descriptions or tags of objects, events, and documents so that these media could be components of a CWW system. The original project I proposed was to use blog data in such a system, but first it was necessary to validate some preliminary ideas, which is what this project ended up doing. For future work, I can look at the blog data, which was labeled using an open set of words for the author’s mood of the blog posting. This larger vocabulary can be mapped to a more concise representation, like the category words or VAD scales, and also text statistics can be mined to find salient features of blog postings that have certain emotional characteristics.
Appendix 1: Survey Instructions

Emotion Survey For Fuzzy Logic

In many surveys, the person taking the survey is presented with some item and asked to rate it on a scale (0-10, 1-7, etc.). Here, we do the same except that we will use two points on the scale, one to represent the lower bound of a range of possible values and the other for the upper bound. In this way, it is possible to make your response more or less specific.

In this experiment, at the top of the page you will be presented with a word or words that describe a certain emotion. Below this, there will be 3 slider scales. You must express the emotion word(s) in terms of a RANGE using BOTH handles of each slider. This way the range will help you be able to express when the emotion word is more or less precise. Word meanings, especially for emotion words, can be vague and hard to define precisely, so by giving a range, it is possible to represent such uncertainty (As a side note, this is where fuzzy logic fits in, though you don’t need to know about fuzzy logic to take the survey).

The slider scales will be the same for each emotion word presented. They are 3 characteristics that psychologists use to classify emotions:

Valence is the first characteristic. It refers to whether the emotion is negative/unpleasant or positive/pleasurable.

Activation is the second characteristic. It refers to how strong the emotion is: dispassionate/calm or passionate/excited.

Dominance is the third characteristic. It refers to how assertive the emotion is: submissive/retreating or dominant/aggressive.

Use the sliders to define a range for each of the 3 characteristics that best fits your idea of the presented word(s) mean. The left slider represents the lower boundary of the range and the right represents the upper boundary.

Don’t worry if you are unclear at first. If you are unclear about something keep in mind the following three things: (1) The first several words that are presented will help you get your bearings. (2) There are cartoon pictures above the scales (like those above) that will help you visualize what they mean, so if the concepts of valence, activation, or dominance seem unclear, use the pictures as a reference. (3) Finally, words -- especially emotion words -- can mean different things to different people. It is natural to feel uncertain about a ‘right’ answer. Use the range between the sliders to specify the region that you can feel certain about.
If you wish to see the definitions of the valence, activation, and dominance scales, you can put your mouse over the picture. However, it is often more instructive to just consider the pictures.

If you have any comments, please write these down and email them to me. My email address will be shown when the experiment is over.

Thanks for your help. This should take between 20 and 40 minutes.

Please do the experiment in one session (without logging out). If you need to take a break, DON’T log out. Use the log out button only if you make a mistake and would like to start over.

Please log on using your first name and email address.
Appendix 2: Survey Screen Shot

Figure 12: This screen shot shows the word “fearful” being shown as a stimulus. Below it are the 3 sliders for valence, activation, and dominance. On the each side of the sliders, there are buttons that move the sliders to the extremes for ease of use.

References


