Analysis of Modifier Structure for Emotion Expressions

Liang-Chih Yu\textsuperscript{1,3} and K. Robert Lai\textsuperscript{2,3}
\textsuperscript{1}Department of Information Management
\textsuperscript{2}Department of Computer Science and Engineering
\textsuperscript{3}Innovation Center for Big Data and Digital Convergence
Yuan Ze University, 135 Yuan-Tung Road, Chung-Li, Taiwan 32003, R.O.C.
E-mail: lcyu@saturn.yzu.edu.tw Tel: +886-3-4638800

Abstract— Dimensional emotion representation such as valence and arousal (VA) space has been an emerging way to represent emotions. In this representation, emotion words can be projected to the VA space according to their valence and arousal values. Sentence and document-level emotions can then be projected based on the emotion words within them. However, emotion expressions in sentences and documents usually contain various modifier structure such as negation (e.g., not happy), degree (very happy) and emotion compounds. Such modifier structure can provide more precise information for measuring VA values in both sentence and document-levels. In this study, we analyze various types of modifier structure for emotion expressions. In addition, we also investigate the effect of different types of modifier structure on measuring VA values for emotion expressions.

I. INTRODUCTION

Automatic emotion recognition or sentiment analysis has been an emerging research field in recent years [1], [2], [3], [4], [5]. For emotion representation, categorical and dimensional are the two major representation schemes. Categorical representation represents emotions into several predefined emotion categories such as binary (positive and negative) and six basic emotions (e.g., anger, happiness, fear, sadness, disgust and surprise) [6]. An obvious limitation of the categorical representation is the problem of undefined categories because it is difficult to list all possible emotion categories for recognition. In addition, emotion categories are usually necessary to be refined when application domains have been changed. For dimensional representation, it represents emotion states into two or three dimensions such as valence-arousal (VA) space [7], [8]. The valence represents the degree of pleasant and unpleasant, and the arousal represents the degree of excited and calm. In this representation, all emotion states can be represented as points in the VA space. Moreover, this representation can be used as a platform for sharing emotion annotations across different research teams.

Both categorical and dimensional representations have been used in automatic emotion recognition or sentiment analysis. For categorical representation, automatic emotion recognition is to classify words/sentences/documents into a set of predefined emotion categories [9], [10], [11], [12], [13], [14]. Taboada et al. indicate that emotion lexicons are useful resources for sentiment analysis [9]. These emotion lexicons include ANEW (Affective norms for English words) [15], WordNet-Affect [16] and SENTIWORDNET [17], which are constructed manually or using (semi-)automatic methods. Devitt and Ahmad used emotion lexicons and lexical cohesion to identify positive and negative news articles [10]. Wu et al. constructed semantic dependency graphs to capture semantic structure within sentences to identify 17 depressive symptoms [11]. Wu et al. proposed the use of language patterns discovered using association rule mining to identify causal relations between sentences [12]. Husby and Barbosa used named entities as features to classify blog posts in eight classes (e.g., Government, Celebrities, Food & Drink and Religion) [13]. D'Mello et al. used several linguistic features and dialog features (e.g., dialog turns) for emotion recognition in a dialog system [14].

For dimensional representation, automatic emotion recognition is to predict valence and arousal values for words/sentences/documents. For example, Wei et al. proposed the use of linear regression to convert VA values of English emotion words to those of Chinese emotion words [18]. Paltoglou and Thelwall used the VA values defined in ANEW to predict the VA values of blog posts. They divided the VA dimensions into five different levels: very positive/high, positive/high, neutral/moderate, negative/low and very negative/low and the blog posts were then classified into these levels [19].

The above research has used various useful features for automatic emotion recognition such as bag-of-words (BOW), syntactic (e.g., language pattern), semantic features (e.g., named entity and semantic dependency) and dialog features. However, modifier structure is also a useful feature but seldom discussed in previous research. For example, not happy (negation), very happy (degree) and not very happy (negation + degree) represent different patterns of modifier structure. Such modifier structure can provide more precise information for measuring VA values in both sentence and document-levels. Accordingly, this study analyzes various types of modifier structure for emotion expressions. In addition, different combinations of modifiers and emotion words may have different effects on VA values (e.g., increase or decrease). Therefore, we also investigate several weighting schemes to calculate the VA values for the modifier structure.
### Table II
**Modifier Structure of Emotion Expressions**

<table>
<thead>
<tr>
<th>Modifier</th>
<th>Structure</th>
<th>Operator</th>
<th>Example</th>
<th>Operation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Degree</td>
<td>deg(+) + ew</td>
<td>Increase</td>
<td>(very) + (very) + (happy), 滿意(satisfactory)</td>
<td>$Val_{new} = (1 + w_{e}^{+}) \times Val_{ew}$; $Aro_{new} = (1 + w_{e}^{+}) \times Aro_{ew}$</td>
</tr>
<tr>
<td></td>
<td>deg(-) + ew</td>
<td>Decrease</td>
<td>(not) + (not) + (happy), 滿意(satisfactory)</td>
<td>$Val_{new} = (1 - w_{e}^{-}) \times Val_{ew}$; $Aro_{new} = (1 - w_{e}^{-}) \times Aro_{ew}$</td>
</tr>
<tr>
<td>Negation</td>
<td>neg_1 + ew ∈ I, IV</td>
<td>Reverse</td>
<td>(not) + (happy), 滿意(satisfactory)</td>
<td>$Val_{new} = -Val_{ew}$; $Aro_{new} = -Aro_{ew}$</td>
</tr>
<tr>
<td></td>
<td>neg_2 + ew</td>
<td>Neutral</td>
<td>(not) + (happy), 滿意(satisfactory)</td>
<td>Neutral</td>
</tr>
<tr>
<td>Combine</td>
<td>deg (-) + neg_1 + ew ∈ I, IV</td>
<td>Reverse + Decrease</td>
<td>(not) + (happy), 滿意(satisfactory)</td>
<td>$Val_{new} = (1 - w_{e}^{-}) \times (-Val_{ew})$; $Aro_{new} = (1 - w_{e}^{-}) \times (-Aro_{ew})$</td>
</tr>
<tr>
<td></td>
<td>deg (+) + neg_1 + ew ∈ I, IV</td>
<td>Reverse + Increase</td>
<td>(not) + (happy), 滿意(satisfactory)</td>
<td>$Val_{new} = (1 + w_{e}^{+}) \times (-Val_{ew})$; $Aro_{new} = (1 + w_{e}^{+}) \times (-Aro_{ew})$</td>
</tr>
<tr>
<td></td>
<td>neg + deg (+) + ew</td>
<td>Increase + Decrease</td>
<td>(not) + (happy), 滿意(satisfactory)</td>
<td>$Val_{new} = (1 - w_{e}^{-}) \times [(1 + w_{e}^{+}) \times Val_{ew}]$; $Aro_{new} = (1 - w_{e}^{-}) \times [(1 + w_{e}^{+}) \times Aro_{ew}]$</td>
</tr>
<tr>
<td></td>
<td>ew_i + ew_j</td>
<td>Average</td>
<td>既生氣又難過(sad and happy)</td>
<td>$Val_{new} = (Val_{ew_i} + Val_{ew_j}) / 2$</td>
</tr>
</tbody>
</table>

### Table I
**Modifier Types**

<table>
<thead>
<tr>
<th>Type</th>
<th>Example modifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Degree</td>
<td>- 有點, 有些 (a little)</td>
</tr>
<tr>
<td></td>
<td>+ 很, 好, 十分, 相當 (quite)</td>
</tr>
<tr>
<td></td>
<td>既, 超, 非常 (very)</td>
</tr>
<tr>
<td>Negation</td>
<td>1 不, 不是 (not)</td>
</tr>
<tr>
<td></td>
<td>2 沒, 没有 (not, no)</td>
</tr>
</tbody>
</table>

### II. Modifier Structure Analysis

Table I lists several example modifiers for the two types of modifiers: degree and negation. Table II shows various types of modifier structure and their effects on VA values. Each modifier structure is constituted by modifiers and emotion words, as shown in the column Structure. The column Operator explains how a modifier affects the VA values of a modified emotion word when they are combined to form a modifier structure. The operators include Increase, Decrease, Reverse, Neutral, Average and combinations of them. The column Example shows modifier structure examples using different modifiers combined with the emotion words in different quadrants. For example, the representative emotion words used herein for quadrants I, II, III and IV are 高興(happy), 生氣(angry), 難過(sad), 滿意(satisfactory), respectively. The column Operation shows the equations used to calculate the VA values of the modifier structure, where $Val_{new}$ and $Aro_{new}$ respectively represent valence and arousal values of a modifier structure; $Val_{ew}$ and $Aro_{ew}$ respectively represent valence and arousal values of an emotion word; $w_{e}$ represents a weight to increase/decrease the VA values of modifier structure from original emotion words.

For degree, modifiers may increase or decrease the VA values once combining with emotion words. For example, 有點高興(a little happy) is a combination of the modifier 有點(a little) and emotion word 高興(happy), and the modifier 有點(a little) may decrease the VA values of the emotion word...
The percentages of increase or decrease are determined based on the weight $w_{ea}$. For negation, different modifiers with emotion words in different quadrants may have different effect on VA values. For example, the modifier structure $neg_1 + ew \in I$, IV yields a reverse effect while $neg_1 + ew \in II$, III yields a neutral effect. Additionally, different types of modifiers can be further combined to modify emotion words, and thus produce additional effects such as Reverse +Decrease, Reverse+Increase and Increase+Decrease. In addition to modifiers, emotion words themselves can also be combined to form a compound structure such as $ew_i + ew_j$. For such emotion compounds, their VA values are calculated as the average of the VA values of individual emotion words.

III. CONCLUSIONS

This paper has analyzed various types of modifier structure for emotion expressions. In addition, we have also investigated the effect of different types of modifier structure on measuring VA values for emotion expressions. Future work will focus on using automatic or semi-automatic methods to discover more modifier structure from large corpora. The weights used to measure VA values of modifier structure will also be investigated based on human annotation. Once the weights for different types of modifier structure have been determined, the VA values of sentences and documents can then be estimated based on the modifier structure within them.

ACKNOWLEDGMENT

This work was partially supported by the Ministry of Science and Technology, Taiwan, R.O.C., under Grant No. NSC 102-2221-E-155-029-MY3.

REFERENCES