Acquiring a Dictionary of Emotion-Provoking Events

Hoa Trong Vu, Graham Neubig, Sakriani Sakti, Tomoki Toda, Satoshi Nakamura
Graduate School of Information Science, Nara Institute of Science and Technology
EACL 2014

Natural Language Processing Seminar - Fall 2014
PaUlô GômeS and Elahe Rahimtoroghi
Emotion-Provoking Events

● “Given a particular emotion, what are the most prevalent events (situation, context) that provoke it?”

Motivation:
● Recognizing emotion based on events
● Generate response to an emotion-related question
● Social science (discovering events that affect the emotion of a particular population)
What is Emotion?

- Basic emotions according to Ekman (1992)
- Ekman’s other basic emotions in the Basic Emotions chapter (1999).
- What other ways could we use to structure emotions?
Affect typings presented by David Elson from Maslow and Max-Neef

<table>
<thead>
<tr>
<th>#</th>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Life</td>
<td>Continuation of basic life functions; existence vs. non-existence</td>
</tr>
<tr>
<td>2</td>
<td>Health</td>
<td>Freedom from pain, disease, malnutrition, and other physical/mental ailments. (If a loss permits the character to live in greater pain, it is a Health matter; if life and death are immediately at stake, it is a Life matter.)</td>
</tr>
<tr>
<td>3</td>
<td>Ego</td>
<td>A positive perception of one's qualities by oneself and by others.</td>
</tr>
<tr>
<td>4</td>
<td>Wealth</td>
<td>Material possessions or currency, above that needed for basic sustenance (those for Health).</td>
</tr>
<tr>
<td>5</td>
<td>Love</td>
<td>Feelings of fondness, warmth, and romance for and from another person; familial companionship; compassion or a desire to heal the world.</td>
</tr>
<tr>
<td>6</td>
<td>Leisure</td>
<td>Entertainment and enjoyment, whether from peaceful solitude, active socializing, or another form of recreation</td>
</tr>
<tr>
<td>7</td>
<td>Actualization</td>
<td>Fulfillment of ones artistic, athletic, spiritual, professional or other aspirational potential in an elective endeavor. Freedom</td>
</tr>
<tr>
<td>8</td>
<td>Freedom</td>
<td>The state of being unrestricted in movement, action and behavior, whether the restricting force is other characters, natural forces, or an internal struggle.</td>
</tr>
</tbody>
</table>
Previous Work


- Japanese language/ Japanese Web corpus

- Emotion lexicon and patterns:
  - automatically obtain a huge collection of emotion-provoking event instances from the Web (1.3 M)

- Emotion classification
Manual Gold-Standard Data

- Survey on 30 subjects
- Write 5 events that provoke 5 emotions
- Manually merge events with similar meaning
- Extract events shared by more than 1 person
- Any comprehensive dictionary of emotion-provoking events should at least cover the pairs in this collection

Describe the population: Gender, Age, Education, Culture

Provide annotation guidelines

More systematic approaches instead of arbitrary/random ideas
Automatic Event Extraction

● seed pattern
  ○ I am <EMOTION> that <EVENT>
● Syntactic constraint on <EVENT>: at least one noun phrase and one verb phrase
● What about sarcasm?
Pattern Expansion

● **Pattern expansion with Espresso** (Pantel and Pennacchiotti, 2006)
  ○ expansion and ranking iterations

● **Normalization**
  ○ remove function words
  ○ replace proper nouns with special symbol
  ○ lemmatizing words
Clustering

- Hierarchical agglomerative single-linkage with cosine similarity
  - Aggregate similar events
  - Reduce sparsity

  ➔ Group synonyms using a lexical resource
    - WordNet: {meet, see, run across, encounter} a {friend, buddy}

- Heuristically choose number of clusters
  - average #events in each cluster = 4

  ➔ Threshold on cluster similarity
Evaluation Measures

- Adapted Mean Reciprocal Rank
- Recall on seed dictionary
  - manual correspondence between seed and tweets
- Other metrics?
Dataset

- Twitter
  - 30 M English tweets
  - 6 weeks June-July 2012

- Preprocessing
  - Remove tags and emoticons → losing information
Experiment

- Methods: only seed, clustering, pattern expansion, and both.
- Expresso: 10 iterations and top-20 patterns at each.
- Only consider 20 most frequent events.
- 3 annotators for EVENT/EMOTION pairs.
- How would you replicate the experiment?
Results

- Clear MRR difference between *only seed* (46.3 +/- 5.0) and *both* (71.7 +/- 2.9)
- Recall still relatively low in *both* (15.4 +/- 0.8)
- Why different results for different emotions?
Conclusions

- Notion of emotion
- Using lexical resources
- Reliability of seed pattern
- Biased dataset (tweets)
- Evaluation measures