Support or Oppose?

Classifying Positions in Online Debates from Reply Activities and Opinion Expressions

Akiko Mukarami, Rudy Raymond
IBM Research, Tokyo
goal: identify author stance in online debate

premise: a local + global approach to the task is better than either approach alone
Approach

- Local - use textual content of a post to classify a post-reply pair into agreement, disagreement, or neutral

- Global - create a graph of authors where an edges represent overall agreement/disagreement level between authors

- Partition the author graph into 2 groups: support and oppose
intuition

- authors' stance remains same during debate
- authors often reply to previous post rather than topic post
- if A and B agree and B agrees with C, A agrees with C
the graph

- edge weight, aka reaction coefficient

\[ r(i, j) = \alpha N_{\text{disagree}}(i, j) + \beta N_{\text{neutral}}(i, j) + \gamma N_{\text{agree}}(i, j), \]

(1)

- first post is the topic post, so graph partition containing that author is support set

- reply-to structure of a debate
[weighted] max-cut

- a cut whose sum of edge weights is not smaller than the size of any other cut
- a cut is a partition of the vertices of a graph into two disjoint subsets
  - negative edges are allowed
  - NP-hard problem so need approx. method
  - BIQMAC is free online SAAS package

- Binary Quadratic and Max Cut solver
- [http://biqmac.uni-klu.ac.at/](http://biqmac.uni-klu.ac.at/)
text analysis

- their approach looks at opinionative and sentiment expressions in post-reply pair
  - opinionative (agree/disagree) expressions identified by matching from dictionary of such expressions in reply post
  - sentiment expressions identified from sentiment labeling tool which returns favorable or unfavorable expressions
  - post-reply score = (# of agrees + # of favorables) - (# of disagree + # of unfavorables)
    - > 0, agree
    - < 0, disagree
    - == 0, neutral
experiments

- 936 ideas (topics) on Japanese government site
- examined 17 with > 40 comments
- from those, selected 5 to experiment
- reply-to structure extracted from threaded organization

<table>
<thead>
<tr>
<th>Idea ID</th>
<th>Title</th>
<th># of Comments</th>
<th># of Participant</th>
<th># of Remarks per Participant</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Making “IT” Education as a Compulsory Subject in Schools</td>
<td>75</td>
<td>45</td>
<td>1.7</td>
</tr>
<tr>
<td>2</td>
<td>Making Personal-Computer Makers to Supplying Service Parts</td>
<td>130</td>
<td>21</td>
<td>6.2</td>
</tr>
<tr>
<td>3</td>
<td>Adoption of “Basic Income”</td>
<td>118</td>
<td>57</td>
<td>2.1</td>
</tr>
<tr>
<td>4</td>
<td>Votes in elections using Closed Networks</td>
<td>108</td>
<td>40</td>
<td>2.7</td>
</tr>
<tr>
<td>5</td>
<td>Computerized Books in Libraries</td>
<td>50</td>
<td>12</td>
<td>4.2</td>
</tr>
</tbody>
</table>
agree/disagree classifier

<table>
<thead>
<tr>
<th>Idea ID</th>
<th>Precision</th>
<th>Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.63</td>
<td>0.25</td>
</tr>
<tr>
<td>2</td>
<td>0.62</td>
<td>0.14</td>
</tr>
<tr>
<td>3</td>
<td>0.44</td>
<td>0.38</td>
</tr>
<tr>
<td>Ave.</td>
<td>0.56</td>
<td>0.26</td>
</tr>
</tbody>
</table>
test set

- manually label author stances (using an 'unclear' label)
- kappa value is 0.69 amongst 3 annotators
- used 1 of the 3 annotations (which one? chosen apriori?)
accuracy

- used custom accuracy to account for differing set size (oppose vs support)
  - average accuracies for support and oppose sets
- ignored 'unclear' label
- 'A' is actual, 'S' is system/predicted

\[
\text{accuracy} = \frac{1}{2} \left( \frac{|A_{\text{sup}} \cap S_{\text{sup}}|}{|A_{\text{sup}}|} + \frac{|A_{\text{opp}} \cap S_{\text{opp}}|}{|A_{\text{opp}}|} \right)
\]
results

- baseline (1,1,1): all replies indicate disagreement
- results show useful to distinguish agreement from disagreement in replies
- results greatly dependent on agree/disagree classifier accuracy

<table>
<thead>
<tr>
<th>ID</th>
<th>Baseline</th>
<th>(1,0,0)</th>
<th>(1,0.5,0)</th>
<th>(1,0.5,-1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>47.86</td>
<td><strong>66.67</strong></td>
<td>54.52</td>
<td>54.05</td>
</tr>
<tr>
<td>2</td>
<td>66.43</td>
<td>76.43</td>
<td>76.43</td>
<td><strong>89.29</strong></td>
</tr>
<tr>
<td>3</td>
<td>46.47</td>
<td>48.88</td>
<td>42.63</td>
<td><strong>55.45</strong></td>
</tr>
<tr>
<td>4</td>
<td>53.19</td>
<td>51.52</td>
<td>55.36</td>
<td><strong>77.60</strong></td>
</tr>
<tr>
<td>5</td>
<td>66.67</td>
<td>58.33</td>
<td>66.67</td>
<td><strong>75.00</strong></td>
</tr>
</tbody>
</table>

\[ r(i,j) = \alpha N_{\text{disagree}}(i,j) + \beta N_{\text{neutral}}(i,j) + \gamma N_{\text{agree}}(i,j), \]  
(1)
future work

- reply could be to both previous post(s) and main topic, rather than one or the other