Document-level Sentiment Inference with Social, Faction, and Discourse Context

Eunsol Choi, Hannah Rashkin, Luke Zettlemoyer, Yejin Choi

ACL 2016
Targeted Sentiment Inference

Extract sentiment relation between real-world entities in news articles.

(Stoyanov and Claire 11, Yang and Cardie 14, Deng and Wiebe 14)

Russia criticizes Belarus

Opinion Holder

Russia

Belarus

Positive? Negative? Neutral?
This work: Document-level Sentiment Inference

Input: News article

Output: Sentiment Graph

Russia criticizes Belarus
Russia criticizes Belarus for …
The speaker of the Russian parliament Friday criticized Belarus…
Saakashvili announced on Belorussian television that he did not understand Russia’s claim

Challenge 1: Read the entire story

positive? negative? neutral?
Challenge 2: Inference across entities

- Entities form factions inside the document.
- Subset of sentiment relations can only inferred via relationships among entities.

- Taylor Swift
- Elie Goulding
- Katy Perry
- Rihanna
- Miley Cyrus

C.E.O
Tim Cook
Apple
**Challenge 3:**

**Sentiment Beyond Sentence Boundaries**

Russia heat, smog trigger health problems...........

...........................  “We never care to work with a future perspective in mind,” Alexei Skripkov of the Federal Agency said. “It’s a big systemic mistake.”

Russia

Alexei Skripkov
This work:
Document-level Sentiment Inference

Input: News article

Output: Sentiment Graph

Russia criticizes Belarus
Evidence: Explicit Sentiment Textual Cues

Belarus for permitting Georgian President Mikheil Saakashvili to appear on its television. Saakashvili announced Thursday that he did not understand Russia’s claims.
Evidence:
Sentiment Inference Through Factions

Entities in the same faction shares opinions
*Homophily (Lazarsfeld and Merton, 1954)*

- **Belarus**
- **Russia**
- **Georgia**
- **Saakashvilli**
- **President**
Evidence:

Sentiment Inference Through Relations

Enemy of an enemy is a friend.

*Social Balance Theory (Heider, 1946)*
This work: Document-level Sentiment Inference

Input: News article

Russia criticizes Belarus

Output: Sentiment Graph

Neg \rightarrow Pos
Related Work

Fine-grained Approach (MPQA)  
(Deng and Wiebe 15)

Opinion Frame
- Trigger: “hostility”
- Polarity: negative
- Intensity: high
- Source: “America”
- Target: “Chavez”

Corpus-level Approach (KBP)  
2013/4 Sentiment Task

Query Entity  
Conan O’Brien is positive towards

Answer
David Litterman, South Korea, Pokemon
Related Work

This Work:

- Document-level approach
- Considers all named entities
- Focuses on entity-entity interactions
Dataset

• Document-level sentiment data collection
• Dataset Statistics
1 Group urges engagement with authoritarian **Eritrea**.

2 A new report says the international community must engage more with the authoritarian leadership of **Eritrea** to prevent the tiny **Red Sea** nation from becoming another failed state in the **Horn of Africa**.

3 The **International Crisis Group** says the effects of **Eritrea**'s 1998-2000 war with **Ethiopia**, an economy in freefall and rising poverty are hemorrhaging the legitimacy of the authoritarian political system in the country.

4 The report -- "**Eritrea**: The Siege State" -- was released late Tuesday.

5 In late 2009, the **U.N. Security Council** imposed an arms embargo and other sanctions against **Eritrea** for supplying weapons to insurgents opposed to the **Somali** government and refusing to resolve a border dispute with neighboring **Djibouti**, a key **U.S.** ally.
Evaluation Dataset Collection

Mark the relationship between the pair as one of below.

<table>
<thead>
<tr>
<th>Positive</th>
<th>Not Negative</th>
<th>Unbiased</th>
<th>Not Positive</th>
<th>Negative</th>
</tr>
</thead>
</table>
Evaluation Dataset Collection

Mark the relationship between the pair as one of below.

<table>
<thead>
<tr>
<th>Positive</th>
<th>Not Negative</th>
<th>Unbiased</th>
<th>Not Positive</th>
<th>Negative</th>
</tr>
</thead>
</table>

When sentiment is not explicitly stated but can be inferred, mark it as Inferred.
1. Group urges engagement with authoritarian **Eritrea**.
2. A new report says the international community must engage more with the authoritarian leadership of **Eritrea** to prevent the tiny **Red Sea** nation from becoming another failed state in the **Horn of Africa**.
3. The **International Crisis Group** says the effects of **Eritrea**’s 1998-2000 war with **Ethiopia**, an economy in freefall and rising poverty are hemorrhaging the legitimacy of the authoritarian political system in the country.
4. The report -- "**Eritrea**: The Siege State" -- was released late Tuesday.
5. In late 2009, the **U.N. Security Council** imposed an arms embargo and other sanctions against **Eritrea** for supplying weapons to insurgents opposed to the **Somali** government and refusing to resolve a border dispute with neighboring **Djibouti**, a key **U.S.** ally.
Evaluation Dataset Collection

• Documents drawn from previous datasets (KBP, MPQA).
• All named entity pairs beyond sentence boundaries.
• Considers the first 15 sentences of document.
• Encouraged to capture implicit sentiment.
Dataset Statistics

<table>
<thead>
<tr>
<th></th>
<th>Positive</th>
<th>Unbiased</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>KBP</td>
<td>11.6</td>
<td>44.7</td>
<td>5</td>
</tr>
<tr>
<td>MPQA</td>
<td>9.5</td>
<td>91</td>
<td>11.6</td>
</tr>
</tbody>
</table>

# of Labeled Entity Pairs Per Doc

<table>
<thead>
<tr>
<th></th>
<th># Docs</th>
<th>Avg. # of entities</th>
</tr>
</thead>
<tbody>
<tr>
<td>MPQA</td>
<td>54</td>
<td>10.6</td>
</tr>
<tr>
<td>KBP</td>
<td>154</td>
<td>7.9</td>
</tr>
</tbody>
</table>

Total 15,185 entity pair labels

<table>
<thead>
<tr>
<th>Annotator Agreement (Cohen’s Kappa)</th>
<th>Same polarity</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.71</td>
<td></td>
</tr>
</tbody>
</table>
Data Characteristics

- High overlap (about 90%) with KBP and MPQA
- Denser (10x more) entity pair annotations than KBP/MPQA
- Cover all named entity pairs
- Capture inferred sentiment
The **FINRA** announced the fine, saying **Goldman** lacked adequate procedures to ensure the required disclosure.

<table>
<thead>
<tr>
<th>Inferred Ratio</th>
<th>Positive</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inferred Ratio</td>
<td>70%</td>
<td>58%</td>
</tr>
</tbody>
</table>
Sentiment Beyond Sentence Boundaries

~25% of entity pairs with sentiment occurs from pairs not appear in the same sentence.

Russia heat, smog trigger health problems............
............................................. “We never work with a future perspective,” Alexei Skripkov of the Federal Agency said. “It’s a big systemic mistake.”

Russia

Alexei Skripkov
Model

- Document-level ILP model framework
- Soft constraints capturing entity-entity interactions
- Individual entity pair model
Document level ILP model

Objective = $\psi_{faction} + \psi_{balance} + \psi_{reciprocity} + \psi_{pairwise}$
Document level ILP model

Objective = $\psi_{faction} + \psi_{balance} + \psi_{reciprocity} + \psi_{pairwise}$

Global soft constraints capturing entity-entity interactions
Document level ILP model

Objective = $\psi_{\text{faction}} + \psi_{\text{balance}} + \psi_{\text{reciprocity}} + \psi_{\text{pairwise}}$

Individual entity pair sentiment classifier model scores

- Russia criticizes Belarus for …
- The speaker of the Russian parliament Friday criticized Belarus…
- Saakashvili announced on Belorussian television that he did not understand Russia’s claim

Opinion Holder: Russia

Opinion Target: Belarus

positive? negative? neutral?
Inference from Faction Relation

Objective = $\psi_{faction} + \psi_{balance} + \psi_{reciprocity} + \psi_{pairwise}$

Entities in the same faction shares opinion, and is positive toward each other (Lazarsfeld and Merton, 1954)
Inference from Faction Relation

\[ \text{Objective} = \psi_{\text{faction}} + \psi_{\text{balance}} + \psi_{\text{reciprocity}} + \psi_{\text{pairwise}} \]

Entities in the same faction shares opinion, and is positive toward each other (Lazarsfeld and Merton, 1954)
Inference from Faction Relation

Objective = $\psi_{faction} + \psi_{balance} + \psi_{reciprocity} + \psi_{pairwise}$

Entities in the same faction shares opinion, and is positive toward each other (Lazarsfeld and Merton, 1954)
Inference from Faction Relation

Objective = $\psi_{faction} + \psi_{balance} + \psi_{reciprocity} + \psi_{pairwise}$

Entities in the same faction shares opinion, and is positive toward each other (Lazarsfeld and Merton, 1954)
Inference from Faction Relation

Objective = $\psi_{\text{faction}} + \psi_{\text{balance}} + \psi_{\text{reciprocity}} + \psi_{\text{pairwise}}$

$$\psi_{\text{faction}} = \sum_{i=1}^{n} \sum_{j=1}^{n} (\alpha_{\text{itself}} \cdot \text{itself}_{ij})$$

$$+ \sum_{k=1}^{n} (\alpha_{\text{fact}} \cdot (\text{tie\_same}_{ijk} - \text{tie\_diff}_{ijk}))$$

$$\text{tie\_same}_{ijk} = \text{tie}_{ij} \land \text{pos}_{ik} \land \text{pos}_{jk} + \text{tie}_{ij} \land \text{neg}_{ik} \land \text{neg}_{jk}$$

$$\text{tie\_diff}_{ijk} = \text{tie}_{ij} \land \text{pos}_{ik} \land \text{neg}_{jk} + \text{tie}_{ij} \land \text{neg}_{ik} \land \text{pos}_{jk}$$

$$\text{itself}_{ij} = \text{tie}_{ij} \land \text{pos}_{ij} - \text{tie}_{ij} \land \text{neg}_{ij}$$
Faction Detector

Heuristics to decide whether an entity belongs to the other.

- modifier, compound, possessive or appositive on dependency path

Saakhfs hvilli  

president of  

Georgia
Faction Detector

On small annotated study, ~30% recall, ~60% precision
Inference from Sentiment Relation

Objective = \( \psi_{faction} + \psi_{balance} + \psi_{reciprocity} + \psi_{pairwise} \)

Social Balance Theory (Heider, 1946):
Models balance or imbalance of sentiment relation in triadic relations.
Inference from Sentiment Relation

\[
\text{Objective} = \psi_{\text{faction}} + \boxed{\psi_{\text{balance}}} + \psi_{\text{reciprocity}} + \psi_{\text{pairwise}}
\]

Social Balance Theory (Heider, 1946):
Models balance or imbalance of sentiment relation in triadic relations.
Inference from Sentiment Relation

Objective = \psi_{action} + \psi_{balance} + \psi_{reciprocity} + \psi_{pairwise}

\psi_{balance} = \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} (\alpha_{balance} \cdot (\text{pos\_same}_{ijk} + \text{neg\_diff}_{ijk}))
+ \alpha_{bad\_balance} \cdot (\text{pos\_diff}_{ijk} + \text{neg\_same}_{ijk}))

\text{pos\_same}_{ijk} = \text{pos}_{ij} \wedge \text{pos}_{ik} \wedge \text{pos}_{jk} + \text{pos}_{ij} \wedge \text{neg}_{ik} \wedge \text{neg}_{jk}
\text{pos\_diff}_{ijk} = \text{pos}_{ij} \wedge \text{neg}_{ik} \wedge \text{pos}_{ik} + \text{pos}_{ij} \wedge \text{pos}_{ik} \wedge \text{neg}_{ik}
Inference from Sentiment Relation

Objective = \psi_{fraction} + \psi_{balance} + \psi_{reciprocity} + \psi_{pairwise}

Reciprocity (Gouldner, 1960): Opinions are often reciprocal.
Inference from Sentiment Relation

Objective = $\psi_{faction} + \psi_{balance} + \boxed{\psi_{reciprocity}} + \psi_{pairwise}$

Reciprocity (Gouldner, 1960): Opinions are often reciprocal.
Inference from Sentiment Relation

Objective = $\psi_{\text{faction}} + \psi_{\text{balance}} + \psi_{\text{reciprocity}} + \psi_{\text{pairwise}}$

Reciprocity (Gouldner, 1960): Opinions are often reciprocal.

$$\psi_r = \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_r(r_{\text{same}_{ij}}) + \alpha_{\text{bad}_r}(r_{\text{diff}_{ij}})$$

$$r_{\text{same}_{ij}} = \text{pos}_{ij} \land \text{pos}_{ji} + \text{neg}_{ij} \land \text{neg}_{ji}$$

$$r_{\text{diff}_{ij}} = \text{pos}_{ij} \land \text{neg}_{ji} + \text{neg}_{ij} \land \text{pos}_{ji}$$
Data shows that these constraints are often satisfied.
Individual Entity Pair Model

- Incorporates the scores from a pairwise classifier model

\[
\text{Objective} = \psi_{\text{faction}} + \psi_{\text{balance}} + \psi_{\text{reciprocity}} + \psi_{\text{pairwise}}
\]

- Factors from linear SVM classifier model

\[
\psi_{ij} = \phi_{\text{pos}_{ij}} \cdot \text{pos}_{ij} + \phi_{\text{neg}_{ij}} \cdot \text{neg}_{ij} + \phi_{\text{neu}_{ij}} \cdot \text{neu}_{ij}
\]
Individual Entity Pair Model

- Russia criticizes Belarus for …
- The speaker of the Russian parliament Friday criticized Belarus…
- Saakashvili announced on Belorussian television that he did not understand Russia’s claim

- Document Features
- Dependency Path Features
- Quotation Features
Document Features

- Capture the entity salience.
  - Do they appear in headline?
  - Do they occur together frequently?

The most frequently mentioned entity is 3.4 times more likely to have sentiment.
Dependency Path Features

- The polarity of path by MPQA sentiment lexicon (Wilson et al. 05)
  - Path containing \([dobj, rev\_subj]\)

Olympic hero Skah accuses Norway over custody battle.
Quotation Features

Russia heat, smog trigger health problems ...........

“We never care to work with a future perspective in mind,” Alexei Skripkov of the Federal Agency said. “It’s a big systemic mistake.”

- The polarity of quotes by MPQA sentiment lexicon
Evaluation
Experimental Set-up

- Precision, Recall, and F1 score
- Positive and negative label sets
- Half as development set, half as a test set
Comparison Systems

- Random (following the data distribution of the dev set)
- Sentence-level RNN classifier:
  - Sentiment model on movie domain (Socher 2013)
  - Sentiment labels from sentences of the pair co-occurring
- Pairwise classifier (without global inference)
Results (KBP)

![Bar Chart]

- **Random**
- **Sentence**
- **Pairwise**
- **Global**

<table>
<thead>
<tr>
<th>F1 Score</th>
<th>Positive</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Legend:
- **Random**
- **Sentence**
- **Pairwise**
- **Global**
Ablation Study

F1 Score

Base +Reciprocity +Balance Theory +Faction

Positive Negative

F1 Score

Base +Reciprocity +Balance Theory +Faction

Positive Negative
Contributions

• Document-level sentiment inference among pairs of entities

• Intuitions from social balance theory and reciprocity

• Models factual relationship affecting sentiment relation
Future work

• Incorporating additional types of factual relationships

• Refine global constraints based on entity types

Data is available!

http://homes.cs.washington.edu/~eunsol/project_page/acl16/
Thanks!

Questions?