+/-EffectWordNet: Sense-level Lexicon Acquisition for Opinion Inference

Yoonjung Choi and Janyce Wiebe
Department of Computer Science, University of Pittsburgh

Introduction

- Most previous research in opinion mining and sentiment analysis tends to focus on explicit opinion expression.
- Deng and Wiebe (2014) show how sentiments toward one entity may be propagated to other entities via opinion inference rules.

The bill would lower the skyrocketing health care costs.

The writer expresses an **explicit** negative sentiment (*skyrocketing*).

This event has a negative effect on object, *costs*
The writer is **implicitly** positive toward this event.

The writer expresses an **implicit** positive toward *the bill.*

- These inference rules involve events that have positive or negative effect on entities. → ***+/-effect*** events

Background

- Examples of +/-effect events (Verb classes by Anand and Reschke, 2010)
  - Creation/Destruction (changes in states involving existence)
    - e.g., bake a cake → **+effect** on the cake
    - e.g., destroy the building → **-effect** on the building
  - Gain/Loss (changes in states involving possession)
    - e.g., increase the tax rate → **+effect** on the tax rate
    - e.g., decrease the tax rate → **-effect** on the tax rate
  - Benefit/Injury (changes in states involving affectedness)
    - e.g., comfort the child → **+effect** on the child
    - e.g., kill Bill → **-effect** on the Bill
- Sentiment vs. Connotation vs. +/-Effect

Perpetrate:
S: (v) perpetrate, commit, pull (perform an act, usually with a negative connotation) "perpetrate a crime"; "pull a bank robbery"

→ This is an **objective** term in SentiWordNet.
→ It has a **negative connotation** by Feng et al. (2011).
→ However, It has a **positive effect** on the object, *a crime*.

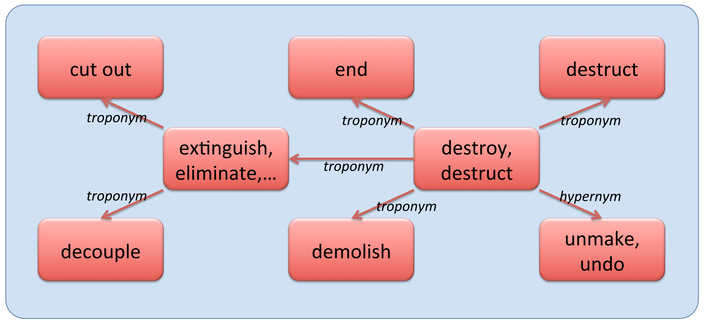
Sense-level +/-Effect Ambiguity

- Since a single word has one or more meaning, it may have both +effect and -effect meanings.

Carry:
S: (v) carry (win in an election) "The senator carried his home state"
→ **+effect** on the agent, *the senator*
S: (v) carry (keep up with financial support) "The Federal Government carried the province for many years"
→ **+effect** on the object, *the province*
S: (v) carry (capture after a fight) "The troops carried the town after a brief fight"
→ **-effect** on the object, *the town*

Sense-level Lexicon Acquisition

- Graph-based Learning for WordNet Relations
  - WordNet semantic relations can be used to build a graph.



  - For the label inference, the local and global consistency method (suggested by Zhou et al. 2004) is applied.
    - The cost function considers two parts;
      - It should not change too much between nearby points, and
      - It should not change too much from the initial label assignment.
- Gloss Classifier
  - A word vector representation is appropriate to utilize gloss information.
  - SVM classifier is adopted.
    - Features: Bag-of-words and sentiment features
- Hybrid Method
  - To use both WordNet relations and gloss information, we combine two methods by a linear combination.

Experiment

- +/-Effect Seed Lexicon
  - Word-level lexicon
    - FrameNet is utilized.
    - An annotator identified promising frames as +/-effect and extracted all lexical units (LUs). Then, he judges to be +effect or -effect.
    - 736 +effect LUs and 601 -effect LUs are selected from 463 semantic frames.
  - Sense-level lexicon
    - The senses of the words in the word-level lexicon are annotated.
    - 258 +effect senses, 487 -effect senses, and 880 null senses are annotated.
    - Half of each set is used as seed set in the graph model and training data in gloss classifier, and the other half is used for evaluation.
- Experimental Result

		Graph Model	Gloss Classifier	Hybrid
+Effect	Precision	0.642	0.588	0.610
	Recall	0.680	0.393	0.735
	F-score	0.660	0.472	0.667
-Effect	Precision	0.779	0.672	0.717
	Recall	0.612	0.511	0.669
	F-score	0.686	0.580	0.692
Null	Precision	0.583	0.451	0.556
	Recall	0.695	0.657	0.520
	F-score	0.634	0.535	0.538

Guided Annotation

- Our method can guide annotation efforts to find other words that have +/-effect senses.
  - Four steps: 1) rank all unlabeled data, 2) choose the top 5% and manually annotated them, 3) add them to the seed set, and 4) rerun the system.
  - The accuracy on the newly labeled annotated data in Step 2 stays good. (Right graph)