Confirming the Non-compositionality of Idioms for Sentiment Analysis

Alyssa Hwang | Columbia University | a.hwang@columbia.edu
Christopher Hidey | Columbia University | chidey@columbia.edu

Background

A multiword expression (MWE) is defined as a phrase that can be decomposed into multiple lexemes and shows lexical, syntactic, semantic, pragmatic, or statistic idiosyncrasy (Baldwin and Kim, 2010), where a lexeme is a linguistic unit that constitutes the basic block of a language (Ramisch, 2015). Idioms, a subset of MWEs, are particularly hard to analyze because they are non-compositional: the meaning of the entire idiom cannot be deduced from the definitions of the individual component words (Jochim et al., 2018). Our goal is to confirm or deny the non-compositionality of idiom sentiment, which is not explicitly delineated in the definition of an idiom. We test idiom sentiment non-compositionality by comparing compositional and phrase-level scores for idioms in the Sentiment Lexicon of IDiomatic Expressions (Jochim et al., 2018).

Methods

SLIDE Positive Percent Index

$$POS - NEG - NEU = PPI$$

The percentages of negative and neutral votes were subtracted from the percentage of positive votes to compute the PPI.

Component-Wise Idiom Scoring

Based on Agarwal et al.'s (2009) method of measuring phrase-level polarity.

Example: "No better than evil"

1. Assign each token a pleasantness score from the DAL; use synonyms/negated antonyms from WordNet for missing words.

No	better	than	evil
1.2308	2.5000	2.000	1.8750

2. Z-Normalize and boost the scores of each word.

No	better	than	evil
-0.135	30.078	3.650	-51.661

3. Handle local negations with a finite state machine.

No	better	than	evil
0.135	30.078	3.650	-51.661

4. Final component-wise score: sum of component scores normalized by length of idiom.

-17.798

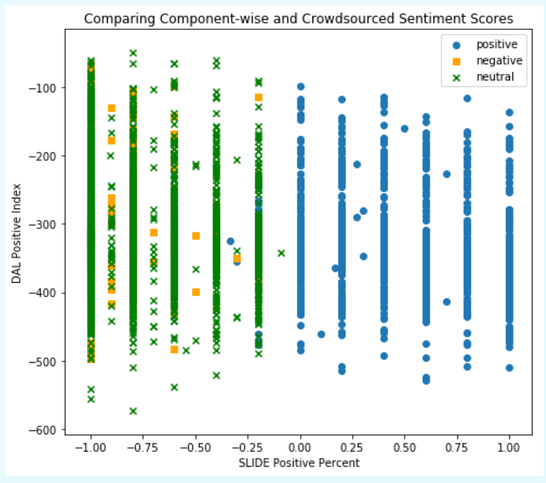
Conclusion and Future Work

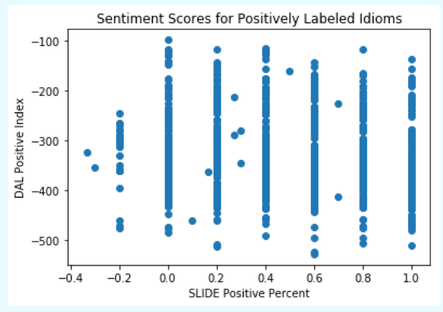
Our analysis shows that there is no consistent correlation between component-wise sentiment scores and crowdsourced phrase-level labels, which supports the hypothesis that idioms are non-compositional for sentiment as well as meaning. Possible future work in the sentiment analysis of MWEs include learning domain-specific sentiment without manual annotation, like predicting a negative sentiment for the phrase "high blood pressure" in the context of a poor health condition. Work must also be done in recognizing new MWEs as language evolves, as well as associating new meanings to already existing words and phrases. Learning to recognize and associate proper sentiment scores to MWEs is an important step in improving overall sentiment classification.

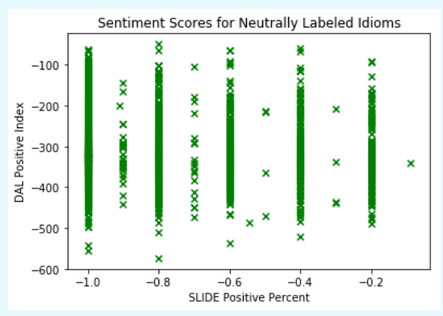
Results

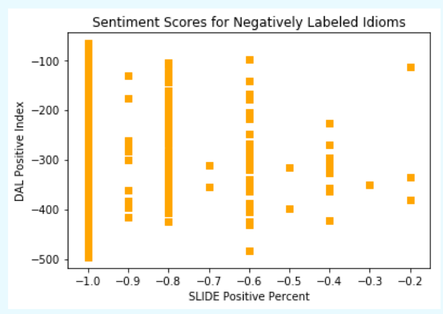
	Spearman corr.	p-value
Positive	-0.144	9.35×10^{-6}
Neutral	0.012	0.503
Negative	0.007	0.813

Spearman correlation and p-value for each sentiment class. The Spearman correlation of each sentiment class is close to 0, which implies no correlation, and we fail to reject the null hypothesis for idioms labeled neutral and negative.









Component-wise and crowdsourced sentiment scores in aggregate (top left) and separated by sentiment class (all right). If idioms were compositional for sentiment, we would expect SLIDE positive percent and DAL positive index to be directly related, but we can see from the top left figure that idioms with the highest SLIDE positive percent rating do not strictly correspond to a higher DAL positive index.

Acknowledgments

We would like to thank Kathy McKeown from Columbia University for her generous support.

Literature

Apoorv Agarwal, Fadi Biadsy, and Kathleen McKeown. 2009. Contextual Phrase-Level Polarity Analysis Using Lexical Affect Scoring and Syntactic N-Grams.

Timothy Baldwin and Su Nam Kim. 2010. *Multiword Expressions,* p. 267-292.

Charles Jochim, Francesca Bonin, Roy Bar-Haim, and Noam Slonim. 2018. SLIDE – A Sentimental Lexicon of Common Idioms.

Carlos Ramisch. 2015. Definitions and Characteristics, p. 23-51.

COLUMBIA | ENGINEERING
The Fu Foundation School of Engineering and Applied Science