

Metaphor Detection with Topic Transition, Emotion and Cognition in Context

Hyeju Jang, Yohan Jo, Qinlan Shen, Michael Miller,
Seungwhan Moon, Carolyn P. Rosé

総研大/NII 宮尾研究室 M2

宮澤 彰

2016 年 9 月 11 日

概要

以下の 2 つの仮説に基づいて、メタファーを検出するための新しい手法を提案している。

1. 周囲の文脈の話題から外れた文にメタファーが多い。
2. 感情の記述にメタファーが多い。

その提案手法を *Breast Cancer Discussion Forum* の投稿 (Jang et al., 2015) に適用して、既存手法を上回ることを示した。

背景

既存研究

提案手法

実験

主に2つのタスクがある.

1. 検出

メタファー的に使われている語や句を見つける. メタファーの定義は研究によって異なる.

(例) (...) these technical terms are not at all easy to **grasp**.

想定される応用

- ▶ テキストや演説などのメタファーを分析する際に補助的に使う.
- ▶ メタファー理解するタスクの前段階として使う.
- ▶ 情報検索, 語義曖昧性解消など他のタスクの改善に役立てる.

2. 理解

共通したタスクの定義があるわけではない。

- ▶ 言い換え (Shutova et al., 2013)
“(...) easy to **grasp**”  → “(...) easy to **understand**”
- ▶ 対応する概念写像の提示 (Shutova et al., 2013)
“That **kindled my ire**.”  → ANGER IS FIRE
- ▶ 性質の変化の理解 (Utsumi et al., 1998)
「林檎のような頬」  → 赤さや丸さの強調

背景

既存研究

提案手法

実験

既存研究におけるメタファー検出へのアプローチ |

メタファーの検出に関しては、多くの既存研究がある。それらについて詳しく知りたければ Veale et al. (2016) を参照するとよい。

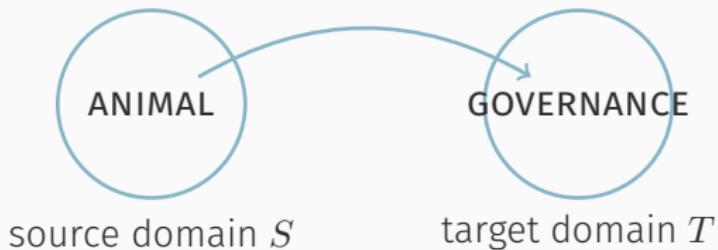
本論文のように、話題の転換を手がかりするものとして Broadwell et al. (2013) がある。この提案手法は、周囲から参照されたり、繰り返し使用されたりする名詞句はメタファーの可能性が低いという考察に基づいている。

These **qualities**¹ have helped **him**⁴ navigate the labyrinthine federal bureaucracy in his demanding \$191,300-a-year job as the top federal **official**³ responsible for bolstering airline, **border**², port and rail security against a second catastrophic terrorist attack. But those same personal **qualities**¹ also explain why the 55-year-old Cabinet **officer**³ has alienated so many Texans along the U.S.-Mexico **border**² with **his**⁴ relentless implementation (...)

既存研究におけるメタファー検出へのアプローチ II

また本論文のように、LDA のトピックを用いるものとしては Heintz et al. (2013) がある。ただし、これはトピックを概念写像のドメインに対応するものとして利用している。

- | Dems_T like rats_S sometimes attack when cornered (...)



背景

既存研究

提案手法

実験

メタファーに関する直感

例えば医療の話の中で “boat” が出たらメタファーの可能性が高い。

When my brain mets¹ were discovered last year, I had to see a neurosurgeon.

He asked if I understood that my treatment was palliative² care. Boy, did it **rock my boat**³ to hear that phrase!

I agree with Fitz, palliative treatment is to help with pain and alleviate symptoms.....but definitely different than hospice care.



© 2009 Sebastian Dorn

¹met abbreviation metastasis /mɪ'tæstəsis/ (癌などの) 転移

²palliative adjective /'pæliətɪv/ 一時的に和らげる

³rock the 'boat (informal) to do something that upsets a situation and causes problems (OALD)

トピックを利用した手法 I

周囲の文脈の話題から外れた文にメタファーが多い。

→ Sentence LDA (Jo and Oh, 2011) を利用する。

素性として以下の 5 つを用いる。

1. Target Sentence Topic (TargetTopic in $\{0, 1\}^T$)
注目している語を含む文のトピック
2. Topic Difference (TopicDiff in $\{0, 1\}^2$)
注目している語を含む文のトピックと、直前/直後の文のトピックが異なるかどうか

トピックを利用した手法Ⅰ

3. Topic Similarity (TopicSim in $[0, 1]^2$)

注目している語を含む文のトピックと、直前/直後の文のトピックの類似度⁴

4. Topic Transition (TopicTrans in $\{0, 1\}^{2T}$)

注目している語を含む文のトピックと、直前/直後の最も近い位置にトピックが異なる文のトピック

5. Topic Transition Similarity (TopicTransSim in $[0, 1]^{2T}$)

注目している語を含む文のトピックと、直前/直後の最も近い位置にトピックが異なる文のトピックとの類似度

⁴各トピックの語の分布を表すベクトルの間のコサイン類似度

感情とメタファーに関する考察

私たちは抽象なことについて語るときメタファーを使うことが多い。

(例) 問題の {芽を摘む, 根が深い, 抱える, 背負う, 乗り越える, ...}

実際, Fainsilber and Ortony (1987)
は, 行動の説明より感情の説明のほうにより多くメタファーが多く含まれることを示している。

このため, メタファー検出において抽象度や想像しやすさがしばしば素性として使われている
(Broadwell et al., 2013; Tsvetkov et al., 2014).



© 2009 Steven Lilley

感情に関する素性

局所的な文脈（着目している語を含む文、及びその前後の文）と大域的な文脈（投稿全体）で LIWC (Tausczik and Pennebaker, 2010) の各カテゴリーに含まれる語の出現回数を数える。

Table 1: Selected LIWC categories.

| LIWC category | Example Terms |
|------------------|------------------------|
| affect | ache, like, sweet |
| positive emotion | passion, agree, giving |
| negative emotion | agony, annoy, miss |
| anxiety | embarrass, avoid |
| anger | assault, offend |
| sadness | despair, grim |

背景

既存研究

提案手法

実験

実験の設定

- ▶ タスク
メタファーの候補となる語を含む文章において、その語がメタファー的に使われているかそうでないか識別する。
- ▶ データ
Breast Cancer Discussion Forum の投稿のうち、以下の 7 種類の語を含むもの。
“boat”, “candle”, “light”, “ride”, “road”, “spice”, “train”
これらには Jang et al. (2015) によってアノテーションが付与されている。
- ▶ 識別器
次数 2 の多項式カーネルを用いた SVM

実験結果

提案手法がすべての項目で既存手法を上回った。

Table 2: Performance on metaphor identification task.

| Model | κ | F1 | P-L | R-L | P-M | R-M | A |
|------------------------|----------|------|------|------|------|------|------|
| Unigram | .435 | .714 | .701 | .434 | .845 | .943 | .824 |
| U + Topic + LIWC | .533 | .765 | .728 | .550 | .872 | .937 | .847 |
| U + MM Topic + MM LIWC | .543 | .770 | .754 | .546 | .872 | .946 | .852 |
| Jang et al. (2015) | .575 | .786 | .758 | .587 | .882 | .943 | .859 |
| J + Topic + LIWC | .609 | .804 | .772 | .626 | .892 | .943 | .869 |
| J + MM Topic | .619 | .809 | .784 | .630 | .893 | .947 | .873 |
| J + MM LIWC | .575 | .787 | .757 | .589 | .882 | .942 | .859 |
| J + MM Topic + MM LIWC | .631 | .815 | .792 | .642 | .896 | .948 | .876 |

トピック

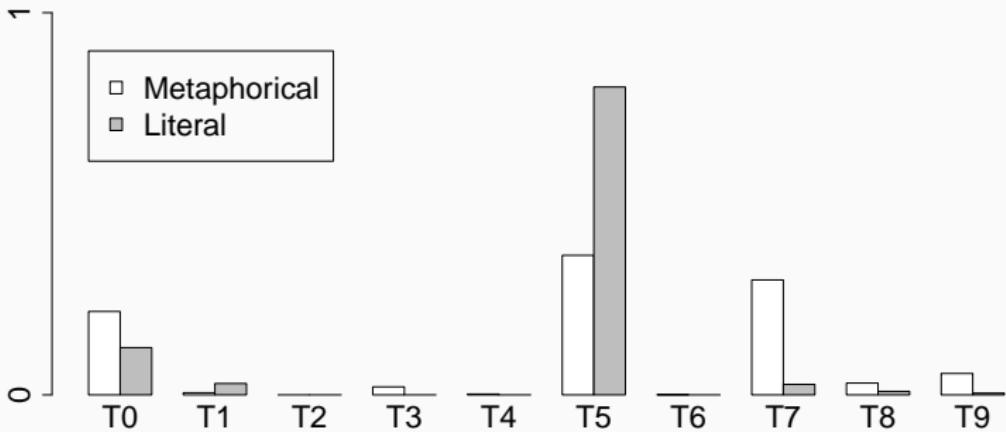
トピックの数は 10 が最適で、以下のようなトピックが得られた。

Table 3: Topics learned by Sentence LDA.

| | Topic | Example Sentences |
|---|-------------------|--|
| 0 | Disease/Treatment | I'm scared of chemo and ct scans ... |
| 1 | Food | *Martha's Way* Stuff a miniature marshmallow ... |
| 2 | Emotions | Too funny. / You're so cute! |
| 3 | Time | I am now 45, and just had my ONE year ... |
| 4 | Greetings/Thanks | Thank you so much for the story!! |
| 5 | People | She has three children and her twin sister ... |
| 6 | Support | YAY! / lol. / I wish you all good luck and peace. |
| 7 | Relation | I just read your message and I wondered about you. |
| 8 | Religion | Dear Lord, I come to you with a friend ... |
| 9 | Diagnosis | I was 64 when diagnosed with pure DCIS ... |

トピック別に見た文の分布 (TargetTopic)

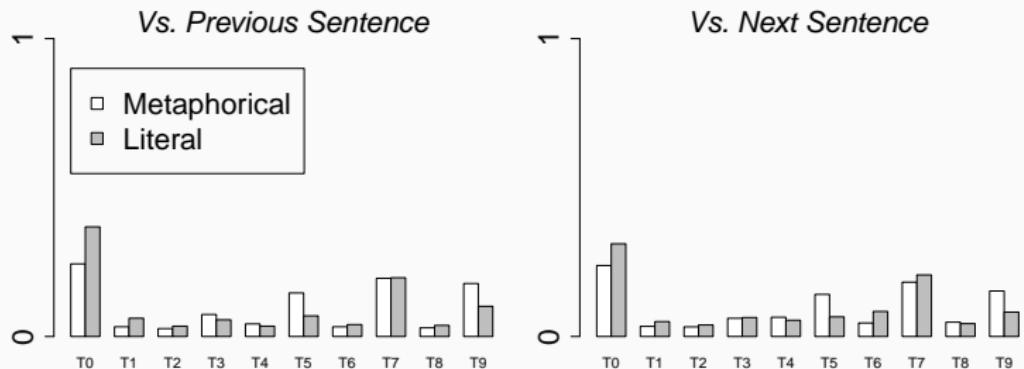
Figure 1: Topic Distribution of Target Sentences



T0 (Disease/Treatment) や T7 (Relation) でメタファーの割合が大きくなっている。一方 T5 (People) はメタファーの割合が小さい。

トピックの変化がメタファーの生じ方に与える影響 (TopicTrans)

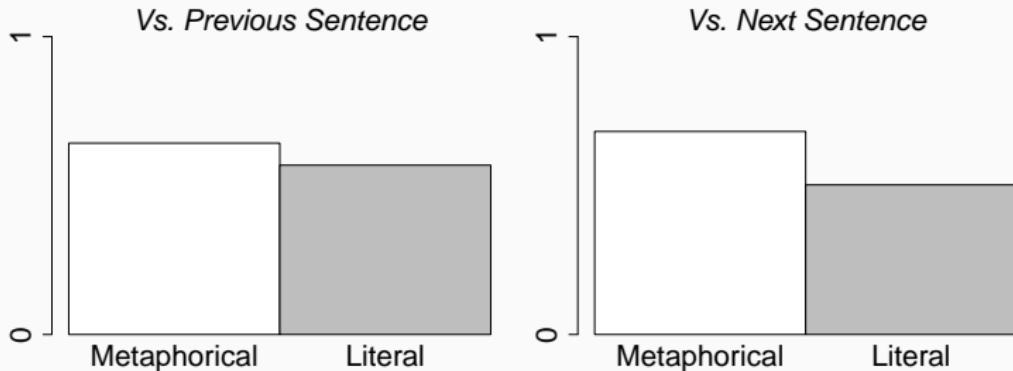
Figure 2: Topic Distribution of the Sentences Nearest to the Target Sentence and with a Different Topic



T5 (People) の前後はメタファー的になりやすい。一方 T0 (Disease/Treatment) の前後はメタファー的になりにくい。

トピックの変化が与える影響 (TopicTrans)

Figure 3: Proportions of Target Sentences with a Different Topic from Context

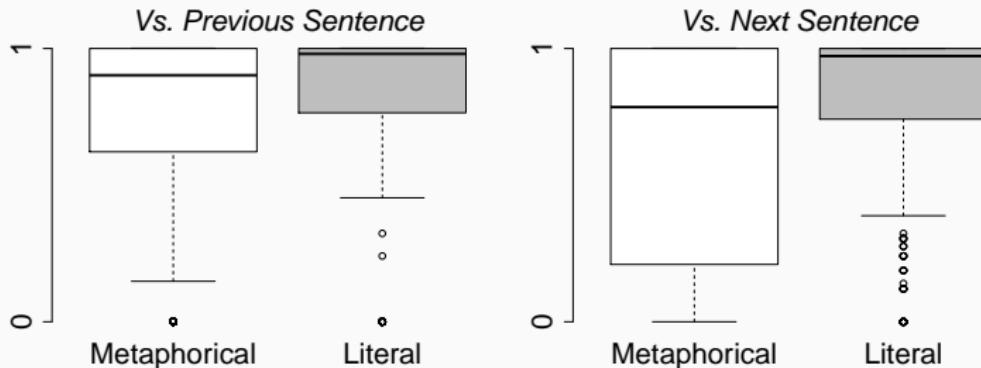


前後と異なるトピックの文はメタファーの割合が大きい⁵.

⁵前後で同じトピックの場合がどうなのが記述がない？

トピックの変化が与える影響 (TopicTransSim)

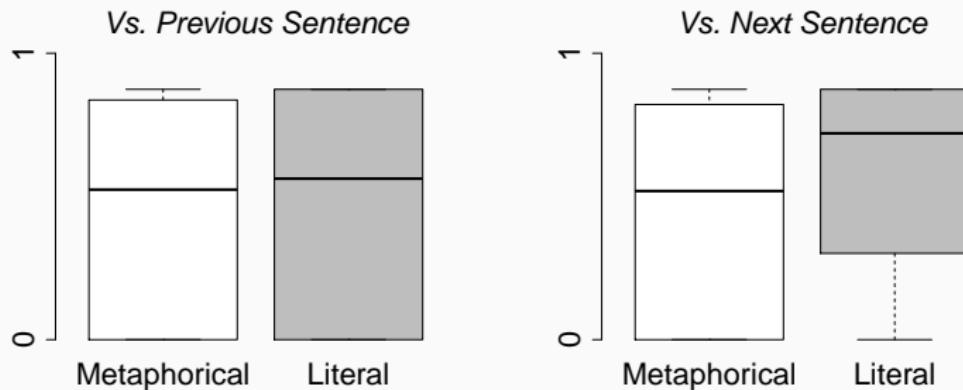
Figure 4: Topic Similarity Between Target Sentence and Context



メタファーを含む文のほうが前後と大きくトピックが異なる。

トピックの変化が与える影響 (TopicTransSim)

Figure 5: Topic Similarity Between Target Sentence and Context



まとめ

メタファーを発見するときには、話題の転換や感情の表現がよい手がかりになるという考えに基づき、前者を Sentence LDA で、後者を LIWC でエンコードする検出の仕組みを示した。

Breast Cancer Discussion Forum のデータで行った実験では、Jang et al. (2015) を大きく上回る性能を示した。

他のドメインへの適用が今後の課題である。

参考文献 |

- Broadwell, George Aaron, Umit Boz, Ignacio Cases, Tomek Strzalkowski, Laurie Feldman, Sarah Taylor, Samira Shaikh, Ting Liu, Kit Cho, and Nick Webb (2013) "Using imageability and topic chaining to locate metaphors in linguistic corpora," in *International Conference on Social Computing, Behavioral-Cultural Modeling, and Prediction*, pp. 102–110, Springer.
- Fainsilber, Lynn and Andrew Ortony (1987) "Metaphorical uses of language in the expression of emotions," *Metaphor and Symbol*, Vol. 2, No. 4, pp. 239–250.
- Heintz, Ilana, Ryan Gabbard, Mahesh Srivastava, Dave Barner, Donald Black, Majorie Friedman, and Ralph Weischedel (2013) *Proceedings of the First Workshop on Metaphor in NLP*, Chap. Automatic Extraction of Linguistic Metaphors with LDA Topic Modeling, pp. 58–66: ACL LINK.
- Jang, Hyeju, Seungwhan Moon, Yohan Jo, and Carolyn Rose (2015) *Proceedings of the 16th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, Chap. Metaphor Detection in Discourse, pp. 384–392: ACL LINK.

参考文献 II

- Jang, Hyeju, Yohan Jo, Qinlan Shen, Michael Miller, Seungwhan Moon, and Carolyn Rose (2016) "Metaphor Detection with Topic Transition, Emotion and Cognition in Context," in *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, pp. 216–225: ACL LINK.
- Jo, Yohan and Alice H Oh (2011) "Aspect and sentiment unification model for online review analysis," in *Proceedings of the fourth ACM international conference on Web search and data mining*, pp. 815–824, ACM.
- Shutova, Ekaterina, Simone Teufel, and Anna Korhonen (2013) "Statistical metaphor processing," *Computational Linguistics*, Vol. 39, No. 2, pp. 301–353 LINK.
- Tausczik, Yla R and James W Pennebaker (2010) "The psychological meaning of words: LIWC and computerized text analysis methods," *Journal of language and social psychology*, Vol. 29, No. 1, pp. 24–54.
- Tsvetkov, Yulia, Leonid Boytsov, Anatole Gershman, Eric Nyberg, and Chris Dyer (2014) "Metaphor Detection with Cross-Lingual Model Transfer," in *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*, pp. 248–258 LINK.

参考文献 III

- Utsumi, Akira, Koichi Hori, and Setsuo Ohsuga (1998) "An Affective-Similarity-Based Method for Comprehending Attributional Metaphors," *Journal of Natural Language Processing*, Vol. 5, No. 3, pp. 3-31.
- Veale, Tony, Ekaterina Shutova, and Beata Beigman Klebanov (2016) *Metaphor: A Computational Perspective*: Morgan & Claypool Publishers [LINK](#).