Metaphors

Shutova 2010

Tassilo Barth

Saarland University

06. June 2011

Metaphor or not?

Metaphor

To understand one concept in terms of another.

1. I killed the program
2. Please don't hold back your ideas
3. The presentation stirred some excitement
4. He attacked my arguments
5. For I have neither wit, nor words, nor worth, Action or utt' range, nor the power of speech, *To stir men's blood*

Metaphor
To understand one concept in terms of another.

1. I killed the program
2. Please don't hold back your ideas
3. The presentation stirred some excitement
4. He attacked my arguments
5. For I have neither wit, nor words, nor worth, Action or utt' range, nor the power of speech, *To stir men's blood*

Make clear the distinction conventional vs. creative metaphors

Difficulty in assessing metaphoricality

Anatomy of Metaphors

I *killed* the program.

Living Entity	$\Rightarrow$	Computer Program
Source	$\Rightarrow$	Target
Vehicle	$\Rightarrow$	Tenor

Theories: Comparison, Interaction, Conceptual, Selectional Restriction Violation

I *killed* the program.

Living Entity	⇒	Computer Program
Source	⇒	Target
Vehicle	⇒	Tenor

Theories: Comparison, Interaction, Conceptual, Selectional Restriction Violation

Introduce key notions

Explain each theory shortly, 1 sentence or so

What about Salience Imbalance?

Selectional Preference Violation

Selectional Preference Violation

My car *drinks* gasoline.

My car *drinks* gasoline.

Quickly illustrate the concept of selectional preference violation
And how it can be used to detect metaphors

Mention issues with SelPrefs here or later? Currently it's at the end

Computational Approaches

- **Metaphor Recognition**: met* (Fass 1991), CorMet (Mason 2004)
- **Metaphor Interpretation**: MIDAS (Martin 1990), KARMA (Narayanan 1997), Shutova (2010)

- **Metaphor Recognition**: met* (Fass 1991), CorMet (Mason 2004)
- **Metaphor Interpretation**: MIDAS (Martin 1990), KARMA (Narayanan 1997), Shutova (2010)

Two tasks

Exemplary approaches: Knowledge-rich (met*, MIDAS, KARMA) vs. Knowledge-poor (CorMet, Shutova)

Theories they depend on: MIDAS (Conceptual metaphors, rich dependencies between metaphors) vs. met*, Shutova, CorMet (SelPrefViolation) - not sure about KARMA

Metaphor Knowledge

- Master Metaphor List
- MetaBank
- Hamburg Metaphor Database
- Automatic: TalkingPoints-Slipnet (Veale/Hao 2007)

- Master Metaphor List
- MetaBank
- Hamburg Metaphor Database
- Automatic: TalkingPoints-Slipnet (Veale/Hao 2007)

Some relevant resources - say a few words - most of them are based on the conceptual metaphor theory

TP/Slipnet is listed under “Interpretation” by Shutova 2010, but IMHO it’s rather a resource for Interpretation than an actual approach

Metaphor Annotation in Corpora

- .. is hard (see beginning)!
- Binary annotation vs. source-target tags
- Search for source + target vocabulary
- Search for linguistic markers ("metaphorically speaking")
- Manual search: Metaphor Interpretation Procedure (MIP)

- .. is hard (see beginning)!
- Binary annotation vs. source-target tags
- Search for source + target vocabulary
- Search for linguistic markers ("metaphorically speaking")
- Manual search: Metaphor Interpretation Procedure (MIP)

Refer to the slide with example metaphors from the beginning

MIP especially relevant, it is used by Shutova to tag their corpus

Shutova 2010:

Automatic Metaphor Interpretation as a Paraphrasing Task

Shutova 2010

The Ingredients:

Parsed corpus, annotated metaphorical *verbs* plus direct object or subject -

- the new idea *stirred* excitement$_{object}$
- the report$_{subject}$ *leaked* to the media

The Recipe:

1. Find other verbs in same context
2. Rank by likelihood
3. Throw out junk verbs
   ———Put aside as BASELINE———-
4. Re-rank by selectional association
5. Choose top rank

Shutova 2010

The Ingredients:

Parsed corpus, annotated metaphorical *verbs* plus direct object or subject -

- the new idea *stirred* excitement$_{object}$
- the report$_{subject}$ *leaked* to the media

The Recipe:

1. Find other verbs in same context
2. Rank by likelihood
3. Throw out junk verbs
   ——Put aside as BASELINE——
4. Re-rank by selectional association
5. Choose top rank

Make clear that they were concerned with conventional metaphors
Will say more about concrete data set later

.. and SelAssociation ranking completely ignores likelihood ranking - is like a second approach

Example: . . . *stirred* excitement$_{object}$

1. Find other verbs in same context

provoked excitement
created excitement
made excitement
demand excitement

Example: . . . *stirred* $\underline{\text{excitement}}_{object}$

2. Rank by likelihood

$P(verb) * P((context - word, syntactic - rel)|verb)$
(P given by relative frequencies in corpus)

LogLh	Paraphrase
-14.28	create
-14.84	provoke
-15.53	make
-15.82	demand

Example: . . . *stirred* excitement$_{object}$

3. Throw out junk verbs

Requirement: New verb and metaphorical verb have common hypernym in WordNet (maximum 3 levels)

LogLh	Paraphrase
-14.28	create
-14.84	provoke
-15.53	make
~~-15.82~~	~~demand~~

Example: . . . *stirred* excitement$_{object}$

3. Throw out junk verbs

Requirement: New verb and metaphorical verb have common hypernym in WordNet (maximum 3 levels)

LogLh	Paraphrase
-14.28	create Baseline!
-14.84	provoke
-15.53	make
~~-15.82~~	~~demand~~

Example: . . . *stirred* excitement$_{object}$

4. Re-rank by selectional association

v = the verb.

c = one of 200 noun classes given by a noun clustering algorithm:

$$A'_R(v,c) = P(c|v) * log \frac{P(c|v)}{P(c)} \qquad A_R(v,c) = \frac{A'_R(v,c)}{\sum_{c'} A'_R(v,c')}$$

$\mathbf{A_R(v,c)}$	**Paraphrase**
0.069	provoke
0.003	create
0.000	make

Example: . . . *stirred* excitement$_{object}$

4. Re-rank by selectional association

v = the verb.
c = one of 200 noun classes given by a noun clustering algorithm:

$A'_R(v,c) = P(c|v) * log\frac{P(c|v)}{P(c)}$ $\quad A_R(v,c) = \frac{A'_R(v,c)}{\sum_{c'} A'_R(v,c')}$

$\mathbf{A_R(v,c)}$	**Paraphrase**
0.069	provoke
0.003	create
0.000	make

How much to say about the noun clustering?

Example: . . . *stirred* excitement$_{object}$

5. Choose top rank

This is the literal interpretation of the metaphor.

$\mathbf{A_R(v, c)}$	**Paraphrase**
0.069	provoke
0.003	create
0.000	make

Example: . . . *stirred* excitement$_{object}$

5. Choose top rank

This is the literal interpretation of the metaphor.

$\mathbf{A_R(v, c)}$	**Paraphrase**
0.069	provoke
0.003	create
0.000	make

Next: stir well and evaluate!

Evaluation

- Annotators tagged verb occurrences in subset of BNC as +/- metaphorical
- Filter out noisy cases (named entities and pronouns at subject/object position etc.)
- 62 metaphorical expressions in total
- Find paraphrases for all of them

First question: How good are top paraphrases chosen by system ("Precision")?
Second question: How good and exhaustive is overall ranking ("Recall")?

Evaluation

First question: How good are top paraphrases chosen by system (Precision)?
Second question: How good and exhaustive is overall ranking ("Recall")?

First question: How good are top paraphrases chosen by system (Precision)?
Second question: How good and exhaustive is overall ranking ("Recall")?

Actual recall hard to determine, since gold standard not exhaustive

Evaluation - First question

How good are top paraphrases chosen by system?

Answered by human annotators. E.g.

Is “provoke excitement” a good literal paraphrase for “stir excitement”?

Evaluation - First question

How good are top paraphrases chosen by system?

Answered by human annotators. E.g.

Is "provoke excitement" a good literal paraphrase for "stir excitement"?

81% accuracy for system vs. **55%** for baseline.

Evaluation - Second question

How good and exhaustive is ranking?

- Gold standard: For each metaphorical expression, human annotators give paraphrases
- For each paraphrase ranking given by system: Calculate Reciprocal Rank
- Calculate Mean Reciprocal Rank over all metaphorical expressions

Evaluation - Second question

How good and exhaustive is ranking?

Mean Reciprocal Rank (MRR)

For each paraphrase ranking: R = rank of first gold paraphrase

$$RR = \begin{cases} R^{-1} & R \leq 5 \\ 0 & else \end{cases}$$

MRR = Mean over RR of all expressions

Selectional Association Ranking for *stir excitement*

$\mathbf{A_R}$	**Paraphrase**
0.069	provoke
0.003	**create:** $RR = \frac{1}{2}$
0.000	make

Evaluation - Second question

How good and exhaustive is ranking?

0.63 MRR vs. baseline MRR of **0.55**

How good and exhaustive is ranking?

0.63 MRR vs. baseline MRR of **0.55**

Strange: Why is Baseline MRR = Baseline Accuracy? Coincidence?

Conclusion

Metaphors + CoLi in general

- Most theories to metaphor w/o enough formal strictness
- Selectional Preference Violations not unproblematic, but at least helpful for computational approaches:
  - SPV $\neq$ Metaphor: Metonymies, Anomalies, metaphors which don't violate SelPref..
  - Very general verbs, like "improve"
  - Frequent conventional metaphors
- Metaphors are still a unsolved problem for NLP.

Metaphors + CoLi in general

- Most theories to metaphor w/o enough formal strictness
- Selectional Preference Violations not unproblematic, but at least helpful for computational approaches:
  - SPV $\neq$ Metaphor: Metonymies, Anomalies, metaphors which don't violate SelPref..
  - Very general verbs, like "improve"
  - Frequent conventional metaphors
- Metaphors are still a unsolved problem for NLP.

Maybe give intermediate conclusion after first part!

Problem with other theories? Are all rather vague and hard to formalize

Might be too much, need to summarize

Shutova 2010

- Unlike previous approaches, Shutova 2010 works without predefined knowledge (apart from WN)
- Interpretation result directly usable as input to other NLP modules
- But: Very restricted wrt kind of metaphor
- Issues with Selectional Preference Violation apply as well
- Evaluation: Only 62 metaphorical expressions, which seem to be (judging by the examples) rather strongly lexicalized (in other words, is it really necessary to paraphrase them?)

- Unlike previous approaches, Shutova 2010 works without predefined knowledge (apart from WN)
- Interpretation result directly usable as input to other NLP modules
- But: Very restricted wrt kind of metaphor
- Issues with Selectional Preference Violation apply as well
- Evaluation: Only 62 metaphorical expressions, which seem to be (judging by the examples) rather strongly lexicalized (in other words, is it really necessary to paraphrase them?)

Why no WordNet baseline (Replacing metaphorical senses by other verbs in Synset, or the hypernym)? Actually, the three examples in her paper are all contained in WN


[Shutova, 2010a] E. Shutova.
Automatic Metaphor Interpretation as a Paraphrasing Task
Proceedings of NAACL 2010, 2010.


[Shutova, 2010] E. Shutova.
Models of Metaphor in NLP
Proceedings of ACL 2010, 2010.