Metaphor Identification Using Verb and Noun Clustering

Ekaterina Shutova

ICSI Open House
Berkeley, February 23, 2012
The main idea

- Start from a set of seed metaphorical expressions
- Implicitly model the underlying source – target domain mappings
- Identify new metaphors exemplifying those mappings
- by means of verb and noun clustering

..associated with MECHANISM

1. Our relationship is not really *working*.
2. Diana and Charles did not succeed in *mending* their marriage.
3. The *wheels* of Stalin’s regime were *well oiled* and already *turning*.
Clusters

Metaphor Identification Using Verb and Noun Clustering
System overview

- **seed set** (metaphorical expressions exemplifying a range of source-target domain mappings)
- **noun clustering** (to harvest various target concepts associated with the same source domain)
- **verb clustering** (to create a source domain verb lexicon)
- **search** the BNC for metaphorical expressions describing the target domain concepts using the verbs from the source domain lexicon
- **filter out** the candidates exposing weak selectional preference strength as non-metaphorical.
Seed expressions

- Single-word metaphors
  - expressed by a verb
  - subject and direct object relations
- annotated by 3 independent annotators
- using Metaphor Identification Procedure (MIP) (Pragglejaz Group, 2007)
- agreement of $\kappa = 0.64$
### Clustering Method

| Noun clustering |
|----------------|
| based on the method of Sun and Korhonen (2009) |
| **DATASET**: 2000 most frequent nouns in the BNC |
| **FEATURES**: subject, direct and indirect object relations; verb lemmas indexed by relation type |
| **ALGORITHM**: spectral clustering |

| Verb clustering |
|----------------|
| based on the method of Sun and Korhonen (2009) |
| **DATASET**: VerbNet |
| **FEATURES**: verb subcategorization frames parameterized by their selectional preferences |
| **ALGORITHM**: spectral clustering |
Example Clusters

| Target concepts cluster          |
|---------------------------------|
| desire hostility anxiety passion |
| excitement doubt fear anger     |
| curiosity enthusiasm impulse    |
| instinct emotion feeling suspicion rage |

| Source domain vocabulary         |
|---------------------------------|
| gulp drain stir empty pour sip  |
| spill swallow drink pollute     |
| seep flow drip purify ooze      |
| pump bubble splash ripple simmer boil tread |
Selectional Preference Strength Filter

- Metaphor as a violation of selectional preferences
- e.g. *remember* or *choose* are never in a metaphorical relation with their object
- quantified in terms of selectional preference strength (Resnik, 1993)

\[
S_R(v) = D(P(c|v)||P(c)) = \sum_c P(c|v) \log \frac{P(c|v)}{P(c)}
\]

*P(c)* is the prior probability of the noun class, *P(c|v)* its posterior probability given the verb; *R* is the GR

- candidates exposing weak selectional preferences are filtered out as non-metaphorical
Example Output

**Seed phrase expansion**

- stir excitement → swallow anger
- reflect concern → disguise interest
- throw remark → hurl comment
- cast doubt → spark enthusiasm etc.

**Output sentences from the BNC**

- K2W 1771 The committee heard today that gangs regularly *hurled* abusive *comments* at local people.
- CKM 391 Time and time again he would stare at the ground, hand on hip, [...] and then *swallow his anger* and play tennis.
- AD9 3205 He tried to *disguise the anxiety* he felt when he found the comms system down, [...] 
- ADK 634 *Catch their interest* and *spark their enthusiasm* so that they begin to see the product’s potential.
**Evaluation**

**Against WordNet baseline to evaluate coverage**

- baseline: synonyms of verbs and nouns in seed expressions represent source and target domain vocabularies
- baseline covers 13% of phrases identified using clustering
- clustering method goes far beyond synonymy

**Against human judgements to evaluate quality**

- randomly sampled sentences tagged by the system and by the baseline
- annotators were asked to mark tagged expressions they considered metaphorical as correct
- 5 annotators (agreement $\kappa = 0.63$)
- evaluation in terms of precision (P): system $P = 0.79$ vs. baseline $P = 0.44$
Governance example

**mobilize opinion**
mobilize will, alienate opinion, cultivate will, cultivate order, lend decision, lend conviction, liberate order, rally opinion, rally will, summon will, withdraw order, withdraw will, withdraw recommendation
G. Lakoff and M. Johnson.  
1980.  
*Metaphors We Live By.*  
University of Chicago Press, Chicago.

Pragglejaz Group.  
*MIP: A method for identifying metaphorically used words in discourse.*  
Metaphor and Symbol, 22:1–39, 2007.

P. Resnik.  
1993.  
*Selection and Information: A Class-based Approach to Lexical Relationships.*  
Ph.D. thesis, Philadelphia, PA, USA.

L. Sun and A. Korhonen.  
*Improving verb clustering with automatically acquired selectional preferences.*  
In Proceedings of EMNLP 2009, pages 638–647, Singapore, August 2009.
Pictures come from: I

open.salon.com