L113 Word Meaning and Discourse Understanding
Session 5: Figurative Language and Sentiment

Simone Teufel
Natural Language and Information Processing (NLIP) Group
UNIVERSITY OF CAMBRIDGE
Simone.Teufel@cl.cam.ac.uk

2012/2013

22 different opinions on “line”

Number of senses:
5 12 15 16 17 18 19 20 23 25 26 32
1 2 1 3 2 2 4 1 3 1 1 1

By frequency

1. Phenomenology
   - Logical Metonymy
   - Regular Metonymy
   - Metaphor
   - Idioms

2. Automatic Approaches
   - Logical Metonymy
   - Regular Metonymy
   - Metaphor

3. Semantic Orientation of Adjectives
   - Antonymy
   - Linguistic tests for complementaries and antonymy type
   - Linguistic vs. natural polarity

4. Automatic Detection of Sem. Orientation
   - Hatzivassiloglou and McKeown
   - Turney (PMI Method)
Types of Figurative Language

- **Hyperbole** *(mile-high ice cream cone.)*
- **Simile** *(She is like a rose.)*
- **Metonymy**
  - **Creative** *(The ham sandwich is waiting for his check.)*
  - **Regular** *(All eyes were on Germany, but Berlin seemed unwilling to lead the Union.)*
  - **Logical** *(a fast plane)*
- **Metaphor** *(He shot down all my arguments.)*
- **Idiom** *(He has a bee in his bonnet.)*
- **Irony, Humour** *(Beauty is in the eye of the beer-holder)*

**Metonymy**

- Use one expression as placeholder for another
- Very frequent phenomenon in language
- Regular metonymy follows schemes:
  - *Press-men hoisted their notebooks and their Kodaks.* *(PRODUCT-FOR-PRODUCER)*
  - *After Lockerbie, people were more careful about saying that.* *(LOCATION-FOR-EVENT)*
- Creative metonymy is hard to recognise automatically, because it depends on the understanding of the entire situation (AI bottleneck).

**Logical Metonymy**

- Due to Pustejovsky (1991, 1995)
- Additional meaning arises for particular verb-noun and adjective-noun combinations in a systematic way
- Verb (or adjective) semantically selects for an event-type argument, but syntactically selects for a noun.
- The event is however predictable from the semantics of the noun.

Examples:
- *Mary finished her beer.*
  *Mary finished drinking her beer.*
- *easy problem
difficult language
good cook
good soup*

**Metaphor**

- Express one entire concept/situation in terms of another concept/situation (including all other participants, properties and events of that situation).
- Lakoff and Johnson (1980): Conceptual Metaphor Theory
- Mapping between two cognitive domains (source and target)
- Usually, source domain is more concrete/evocative

- **SOURCE DOMAIN: WAR**
  - war
  - shoot down
  - win
  - retreat
  - dig in
- **TARGET DOMAIN: ARGUMENT**
  - argument
  - win
  - respond
  - counter-argument
**Metaphor: ARGUMENT is WAR**

- Parties **go into battle** about how high to push the bar for skills
- Villagers **launch fight** to save their primary school from closure
- how to **defend** yourself against stupid arguments

**Metaphor: FEELINGS ARE LIQUIDS**

- A simple phone call had managed to **stir up** all these feelings.
- Now here I was, **seething** with anger
- is a kind of **pressure valve** for the release of **pent-up** nervous energy
- ... provide an **outlet** for creativity ... Just ignore the **turbulent** feelings and turn your attention towards ...

**Mixed Metaphor**

Combination of two incompatible metaphorical mappings:

- biting the hand that rocks the cradle
- it would somehow bring the public school system crumbling to its knees.
- She's been burning the midnight oil at both ends.
- He took it like a fish out of water.
- He wanted to get out from under his father's coat strings.
  (riding on coat tails + cling to mother's apron strings + hide behind your mother's skirts
- If we can hit that bullseye then the rest of the dominoes will fall like a house of cards... Checkmate.
  
  Zapp Brannigan (Futurama)

**Dead metaphor**

**Dead metaphor:** The image that the metaphor invokes has been established in the language, i.e., is now contained in the “lexicon”.

- I simply cannot **grasp** this idea.
- This really made an **impression** on me.

Often not perceived as metaphor. This is opposed to creative, situational metaphor.
Idioms

- Minimal semantic constituents which consist of more than one word.
  - pull somebody’s leg
  - be off one’s rocker
- Definition: the meaning of an idiom cannot be inferred as a compositional function of the meaning of its parts.

Syntactic Variability Tests:
- Arthur has a bee, apparently, in his bonnet. (insertion)
- Arthur kicked the large bucket. (modification)

Idioms: crosslingual issues

Level of translatability of idiom into another language is unpredictable.
- “donner sa langue au chat” (give your tongue to the cat)
- “appeler un chat un chat” (call a cat a cat)

Logical Metonymy: Lapata and Lascarides (2003)

- a fast { landing? } plane
- flying?
- I enjoyed { reading? } the book
- writing?
- eating?
- What is missing for full automatic recognition is the implicit verb (fly(ing)) and read(ing)).
- Cooccurrences of plane–fly and fly–fast and like-reading and read–book in corpus can give us the answer.
- Probabilistic model used collects counts for the two associations separately.

Rephrasing of a dead metaphor results in similar semantics:
- They tried to sweeten the pill. ≈ They tried to sugar the medicine.
- We shall leave no stone unturned in our search for the culprit.
  ≈ We shall look under every stone in our search for the culprit.

This is not the case for idioms (due to their non-compositional semantics):
- John pulled his sister’s leg ≠ John tugged at his sister’s leg
- Arthur kicked the bucket ≠ Arthur tipped over the water receptacle
Phenomenology
Automatic Approaches
Semantic Orientation of Adjectives
Automatic Detection of Sem. Orientation
Logical Metonymy
Regular Metonymy
Metaphor

Logical Metonymy: the adjective model

\[ P(a, e, n, rel) = \frac{f(a, e)f(\text{rel}, e, n)}{f(e)N} \]

e: verbal predicate \( e \) modified by adverb \( a \), bearing argument relation \( \text{rel} \) to head noun \( n \)

<table>
<thead>
<tr>
<th>Frequency: verbs modified by fast.</th>
<th>Frequency: verbs taking plane as argument.</th>
</tr>
</thead>
<tbody>
<tr>
<td>( f(\text{fast}, e) )</td>
<td>( f(\text{fast}, e) )</td>
</tr>
<tr>
<td>go</td>
<td>29</td>
</tr>
<tr>
<td>grow</td>
<td>28</td>
</tr>
<tr>
<td>beat</td>
<td>27</td>
</tr>
<tr>
<td>run</td>
<td>16</td>
</tr>
<tr>
<td>rise</td>
<td>14</td>
</tr>
<tr>
<td>travel</td>
<td>13</td>
</tr>
<tr>
<td>move</td>
<td>12</td>
</tr>
<tr>
<td>come</td>
<td>11</td>
</tr>
<tr>
<td>drive</td>
<td>8</td>
</tr>
<tr>
<td>get</td>
<td>7</td>
</tr>
</tbody>
</table>

Markert and Nissim (06):

- Country and organisation names are classified as metonymical or not
- Manually annotate large training corpus (1,000 examples of each from the BNC)
- Good human agreement
- Supervised learning problem: use grammatical information as features
- Roughly 20% of country names are used metonymically, and 33% of organisation names.

Metonymy: examples

Countries:
- Or have you forgotten that America did once try to ban alcohol and look what happened!
- At one time there were nine tenants there who went to America.

Organisations:
- How I bought my first BMW.
- BMW and Renault sign recycling pact.

Metonymy: Features and results

Features:
- Grammatical function (subj, premod, gen, obj, PP, pred, subjpassive, iobj, other)
- Number, definiteness of determiner
- Lexical head

Results:
- 87% correct for country names (EMNLP 2002 paper)
- 76% correct for organisations (IWCS 2005 paper)
Automatic Approaches to Metaphor Recognition

Selectional restrictions of metaphorically used word in literal interpretation are violated (Wilks 79)

- is-a metaphors violate WN-hyponymy relation: all the world is a stage (Krishnakumaran and Zhu, 2007)
- Or use manually created metaphor-specific knowledge bases (Martin 1980; Narayanan 1999; Barnden and Lee 2002).

Simone Teufel L113 Word Meaning and Discourse Understanding 21

Metaphor Recognition (Shutova et al. 2010)

- Start from seed set including a metaphorical verb (verb in source domain; e.g., stir excitement)
- Task: find other sourceVerb–targetNoun pairs (swallow anger)
- **Step 1**: Collect all subjects and arguments that occur with the seed sourceVerb.
  - Most of these are sourceNouns (soup; non-metaphors), but some are targetNouns (anger).
- **Step 2**: Clustering the nouns according to their semantics by verb association (cf. last lecture)
  - The targetNoun cluster is the most “abstract” cluster
  - Half the job done; we now need to find more sourceVerbs.
- **Step 3**: Go back from sourceNoun clusters
  - Now cluster the verbs they cooccur with
  - The cluster which has the seed verb in it is the sourceVerb cluster.

A Symbolic Approach to Metaphor Interpretation

SLIPNET (Veale and Hao 2008) relates two concepts via definitions, allowing for deletions, insertions and substitutions.

Goal: to find a connection between source and target concepts.

Example:
Make-up is a Western Burqa

- make-up =>
  typically worn by women
  expected to be worn by women
  must be worn by women
  must be worn by Muslim women

burqa <=

Simone Teufel L113 Word Meaning and Discourse Understanding 22

Metaphor Recognition – Examples

<table>
<thead>
<tr>
<th>Target domain N cluster</th>
<th>Source domain V cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td>desire hostility anxiety passion excitement doubt fear anger curiosity enthusiasm impulse instinct emotion feeling suspicion rage</td>
<td>gulp drain stir empty pour sip spill swallow drink pollute seep flow drip purify ooze pump bubble splash ripple simmer boil tread</td>
</tr>
</tbody>
</table>
Task 2: Metaphor Interpretation by literal paraphrase

Input: A *carelessly leaked* report
Output: A *carelessly disclosed* report

- Find lexically similar candidates for replacement (standard distributional semantics approach)
- Use a Resnik-type selectional restriction filter to filter out metaphorical expressions (those that have low selectional restriction strength), so that only literal ones are left over.

\[
AR(v, c) = \frac{1}{SR(v)} \frac{P(c|v)}{P(c)} \log \frac{P(c|v)}{P(c)}
\]

Shutova et al: Paraphrasing Example

<table>
<thead>
<tr>
<th>Initial ranking</th>
<th>SP reranking</th>
</tr>
</thead>
<tbody>
<tr>
<td>hold back truth</td>
<td>-13.09</td>
</tr>
<tr>
<td>-14.15 conceal</td>
<td>-14.15</td>
</tr>
<tr>
<td>-14.62 suppress</td>
<td>-14.62 suppress</td>
</tr>
<tr>
<td>-15.13 hold</td>
<td>-15.13 hold</td>
</tr>
<tr>
<td>-16.23 keep</td>
<td>-16.23 keep</td>
</tr>
<tr>
<td>-16.24 defend</td>
<td>-16.24 defend</td>
</tr>
</tbody>
</table>

| hold back truth | contain 0.1161 conceal |
| -13.09         | 0.0214 keep |
| -14.15 conceal | -14.15 conceal |
| -14.62 suppress| -14.62 suppress |
| -15.13 hold    | -15.13 hold |
| -16.23 keep    | -16.23 keep |
| -16.24 defend  | -16.24 defend |

| stir excitement | create 0.0696 provoke |
| -14.28 create   | -14.28 create |
| -14.84 provoke  | -14.84 provoke |
| -15.53 make     | -15.53 make |
| -15.53 elicit   | -15.53 elicit |
| -15.53 arouse   | -15.53 arouse |
| -16.23 stimulate| -16.23 stimulate |
| -16.23 raise    | -16.23 raise |
| -16.23 excite   | -16.23 excite |
| -16.23 conjure  | -16.23 conjure |

| stir excitement | create 0.0696 provoke |
| -14.28 create   | -14.28 create |
| -14.84 provoke  | -14.84 provoke |
| -15.53 make     | -15.53 make |
| -15.53 elicit   | -15.53 elicit |
| -15.53 arouse   | -15.53 arouse |
| -16.23 stimulate| -16.23 stimulate |
| -16.23 raise    | -16.23 raise |
| -16.23 excite   | -16.23 excite |
| -16.23 conjure  | -16.23 conjure |

Summary

- Logical Metonymy can be solved by individual associations of implicit verb with explicitly mentioned lexical items
- Problem with Lapata/Lascarides (2003): word senses all conflated
- Regular Metonymy can be solved by supervised classification with features similar to supervised WSD.
- Metaphors can be recognised by seed clustering and paraphrased by lexical similarity and selectional restrictions.
- Shutova et al.’s system: precision is high (~80%), but recall is very low (0.25%)
Opposites, Antonyms and Semantic Orientation

- There are different kinds of opposites: complementaries and antonyms.
- Antonyms are closely related to semantic orientation (degree of positiveness/negativeness).
- If we know that two adjectives relate to the same property (e.g., hot and cold) but have different semantic orientations they are typically antonyms.
- Exceptions: verbose—terse

Oppositeness and Antonymy

Complementaries

Between them exhaustively divide some conceptual domain into mutually exclusive compartments.

- Antonyms don’t.
- *neither–nor* test:
  - *? Mary is neither married nor is she single.* (→ opposites)
  - *It’s neither hot nor cold today.* (→ antonyms)

Gradability test:
- *? extremely true – extremely safe*
- *? more pregnant than most – longer than some*
- *? moderately female – moderately clean*

Antonyms 1: Pseudo comparatives and true comparatives

- *This box is light, but it’s heavier than that one.*
- *? Today it’s cold, but hotter than yesterday.*

What is going on?

- *light–heavy* and *hot–cold* do not behave in the same way:
  - *This box is light, but it’s heavier than that one.*
  - *? Today it’s cold, but hotter than yesterday.*
Antonyms 1: Pseudo comparatives and true comparatives

*light–heavy* and *hot–cold* do not behave in the same way:

- *This box is light, but it’s heavier than that one.*
- *Today it’s cold, but hotter than yesterday.*

What is going on? *Heavier* seems to mean *of greater weight* here (relative property), whereas *hot* seems to express a more absolute property.

- *hotter* is a true comparative of *hot*
- *heavier* is a pseudo-comparative of *heavy/1*, and a true comparative of *heavy/2*

Antonyms 2: How-adj questions

Are they possible for both antonyms?

Compare *long–short*:

- *How long is it?*
- *How short is it?*

with *hot–cold*:

- *How cold is it?*
- *How hot is it?*

Antonyms 3: Impartiality of how-adj questions

Does one of the questions imply something about your presuppositions? (this is about neutrality) *hot–cold*:

- *How cold is it?* → committed
- *How hot is it?* → committed

*clean–dirty*:

- *How clean was the room?* → impartial
- *How dirty was the room?* → committed

Three types of antonyms

- **good–bad** is an example of an **overlapping** antonym.
  - Overlapping antonyms are evaluative, and thus carry semantic orientation in our sense.
- **hot–cold** is an example of an **equipollent** antonym.
  - Equipollent antonyms are often correlated with sensory perceptions.
- **long–short** is an example of a **polar** antonym.
  - Polar antonyms show the greatest level of abstraction, but are neutral/descriptive.
Can we predict which one of the antonyms is more “salient” (that is typically the one that is positive)?

Prediction: the more salient antonym often has a positive polarity.

Test 1: The antonym that can be paraphrased as the other one plus a negative prefix is the less salient one.

Test 2: The more salient antonym is associated with “more” properties:
- Something is dead when there is no life present.
- Something is alive when there is no deadness present.

Test 3: The more salient antonym yields the impartial interpretation in the how-adj question.

Algorithm

Extract all coordinated adjectives from 21 million word WSJ corpus (and, or, but, either-or, neither-nor)
- 15048 adj pairs (token), 9296 (type)
- Classify each extracted adjective pair as same or different orientation (82% accuracy)
  - features used: number of modified noun; type of coordination; type of modification (attributive, predicative, appositive, resultative ("Bill laughed himself hoarse")); and is most reliable same-orientation predictor, particularly in predicative position (85%), this drops to 70% in appositive position.
  - but has 31% same-orientation.
  - This information comes from an independently annotated gold standard (1336 most frequent adjectives; 657 positive, 679 negative)
  - Additional different orientations comes from simple morphological analysis. Out of the labelled adjectives, 93% of

Hatzivassiloglou and McKeown’s (1997) algorithm classifies adjectives into those with positive or negative semantic orientation.

In coordinations, antonymy results in constraints on the semantic orientation:

1. The tax proposal was simple and well-received by the public.
2. The tax proposal was simplistic but well-received by the public.
3. The tax proposal was simplistic and well-received by the public.

but combines adjectives of opposite orientation; and adjectives of the same orientation

This indicates information that is not provided.

Clustering adjectives with same orientation

Now cluster adjectives into two orientations, placing as many words of the same orientation as possible into the same subset
- Interpret classifier’s P(same-orientation) as dissimilarity value.
- Perform non-hierarchical clustering via Exchange Method
- Start from random partition, locate the adjective which reduces the cost c most if moved.

\[
c = \sum_{i=1}^{2} \left( \frac{1}{|C_i|} \sum_{x,y \in C_i, x \neq y} d(x, y) \right)
\]

Repeat until no movements can improve the cost; overall dissimilarity cost is now minimised.
Phenomenology
Automatic Approaches
Semantic Orientation of Adjectives
Automatic Detection of Sem. Orientation

Exchange method, final step

- At final iteration, move any adjective which violates the following constraint:

\[
\frac{1}{|C|-1} \sum_{y \in C, x \neq y} d(x, y) < \frac{1}{|\bar{C}|} \sum_{y \in \bar{C}} d(x, y)
\]

Labelling Clusters as Positive or Negative

- Hatzivassiloglou empirically find that the cluster with overall higher frequency tends to be the positive one; so this is the final step in their algorithm
- Possible reason: In overlapping antonym pairs, the positive adjective tends to be semantically unmarked (as we heard earlier today).
- Semantically unmarked adjectives should occur more frequently in language (neutral questions etc)

Results

- Dependent on how sparse the test set is, results between 78% and 92% correct
- Baselines: MFC 51% negative
- Classified as positive: **bold, decisive, disturbing, generous, good, honest, important, large, mature, patient, peaceful, positive, proud, sound, stimulating, straightforward, strange, talented, vigorous, witty.**
- Classified as negative: **ambiguous, cautious, cynical, evasive, harmful, hypocritical, inefficient, insecure, irrational, irresponsible, minor, outspoken, pleasant, reckless, risky, selfish, tedious, unsupported, vulnerable, wasteful.**

Discussion

Strengths:
- Algorithm only needs gold standard list
- Convincing results

Weaknesses:
- Analysis of isolated adjectives, not phrases
- Needs large corpus in order to contain enough coordinated adjectives
- Clustering algorithm is not optimal (problem is NP-hard); it is a steepest-descending hill climbing method, which is at least guaranteed to converge (but might run algorithm repeatedly with different start partitions)
**Turney’s 2002 method**

- Determine semantic orientation of phrases, not just single adjectives
- Single adjectives do not always carry full orientation; context is needed. *unpredictable plot vs. unpredictable steering*
- Unsupervised method based on distributional semantics
- Assign a numerical ranking indicating strength of orientation
- Use search engine hits to estimate semantic orientation of a phrase

**Idea**

- If an adjectival phrase has a positive semantic orientation, it will appear more frequently in the intermediate vicinity of known positive adjectives, and vice versa.
- Measure an adjective’s tendency to appear in positive or negative vicinity via PMI-IR
  - Pointwise mutual information determines similarity of a pair of phrases
  - Use IR to quantify effect
- Measure success indirectly via classification of entire reviews

**PMI and SO**

\[
PMI(\text{word}_1, \text{word}_2) = \log_2 \left( \frac{P(\text{word}_1, \text{word}_2)}{P(\text{word}_1)P(\text{word}_2)} \right)
\]

- Semantic Orientation:
  \[
  SO(\text{phrase}) = PMI(\text{phrase}, \text{excellent}) - PMI(\text{phrase}, \text{poor})
  \]
- Counts are calculated via search engine hits
- Altavista’s NEAR operator – window of 10 words

Therefore:

\[
SO(\text{phrase}) = \log_2 \left( \frac{\text{hits}(\text{phrase NEAR excellent})\text{hits}(\text{poor})}{\text{hits}(\text{phrase NEAR poor})\text{hits}(\text{excellent})} \right)
\]

**Results: indirectly via classification of documents**

- 74% accuracy on classifying 410 reviews from Epinions
- 66% accuracy on movie reviews

An example:

<table>
<thead>
<tr>
<th>Phrase</th>
<th>PMI</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>little difference</td>
<td>-1.615</td>
<td></td>
</tr>
<tr>
<td>clever tricks</td>
<td>-0.040</td>
<td></td>
</tr>
<tr>
<td>programs such</td>
<td>0.117</td>
<td></td>
</tr>
<tr>
<td>possible moment</td>
<td>-0.668</td>
<td></td>
</tr>
<tr>
<td>unethical practices</td>
<td>-8.484</td>
<td></td>
</tr>
<tr>
<td>old man</td>
<td>-2.566</td>
<td></td>
</tr>
<tr>
<td>other problems</td>
<td>-2.748</td>
<td></td>
</tr>
<tr>
<td>probably wondering</td>
<td>-1.830</td>
<td></td>
</tr>
<tr>
<td>virtual monopoly</td>
<td>-2.050</td>
<td></td>
</tr>
<tr>
<td>other bank</td>
<td>-0.850</td>
<td></td>
</tr>
<tr>
<td>extra day</td>
<td>-0.286</td>
<td></td>
</tr>
<tr>
<td>direct deposits</td>
<td>5.771</td>
<td></td>
</tr>
<tr>
<td>online web</td>
<td>1.936</td>
<td></td>
</tr>
<tr>
<td>cool thing</td>
<td>0.395</td>
<td></td>
</tr>
<tr>
<td>very handy</td>
<td>1.349</td>
<td></td>
</tr>
<tr>
<td>lesser evil</td>
<td>-2.288</td>
<td></td>
</tr>
</tbody>
</table>

Total: -1.218. Rating: Not recommended.
Discussion

Strengths:
- Fully unsupervised
- Nominal context makes adjective semantics more interpretable

Weaknesses:
- No direct evaluation of SO provided
- Very simple model
- Requires many searches (too many without API)
- NEAR no longer supported
- Results depend substantially on lexical items chosen, but choice largely unmotivated

Reading

- Cruse (1986), chapters 9 and 11.3;

Coursework 2!!

- Frame semantics of “risk”
- Core task: identify the participants in risk-type situations
- Generalise to other words that are semantically related