Quantitative Approaches to Metonymy

Yves Peirsman

KULeuven
Quantitative Lexicology and Variational Linguistics
Overview

1. Introduction

2. A corpus-based perspective on metonymy
   2.1 General perspective
   2.2 Contextual factors

3. Metonymy recognition
   3.1 Metonymy recognition
   3.2 Active Learning
   3.3 Learning on the basis of related words

4. Conclusions and outlook
1. Introduction

Metonymy

A figure of speech in which a word does not refer to its original referent A, but to a referent B that is contiguously related to A.

Metonymical patterns

- **PLACE FOR PEOPLE**: *Germany* opposed to the decision.
- **ORGANIZATION FOR PRODUCT**: He drives a *BMW*.
- **AUTHOR FOR WORK**: He really likes *Thomas Mann*.
1. Introduction

Theoretical purpose
A corpus-based perspective on metonymical proper nouns

- How often do metonymies occur?
- What contextual factors influence the reading of a possible metonymy?

Computational purpose
Use this statistical information in order to

- automatically recognize metonymical words.
- reduce the required amount of labelling.
Overview

1. Introduction

2. A corpus-based perspective on metonymy
   2.1 General perspective
   2.2 Contextual factors

3. Metonymy recognition
   3.1 Metonymy recognition
   3.2 Active Learning
   3.3 Learning on the basis of related words

4. Conclusions and outlook
2.1 General perspective

Starting point
Markert and Nissim’s corpus-based approach to metonymy recognition

- focus on country and organization names
- 1,000 examples of each from the BNC
- annotated with grammatical information
- used as training and evaluation corpora for a classification system that automatically recognizes metonymies
- but also useful for more linguistic purposes.
2.1 General perspective
2.2 Contextual factors: function

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*.

organizations

- BMW and Renault sign recycling pact.
- German firm’s export challenge CAR component maker Behr, which makes air conditioning for Mercedes and BMW . . .
2.2 Contextual factors: function
2.2 Contextual factors: function

![Bar chart showing contextual factors function]
2.2 Contextual factors: function
2.2 Contextual factors: determiner and number

**ORGANIZATION FOR PRODUCT**

- It was the largest *Fiat* anyone had ever seen
- Press-men hoisted their notebooks and their *Kodaks*.
- In the UK, more than one in 30 new cars is now either a *BMW* or a Mercedes.
2.2 Contextual factors: determiner and number

- Countries: 79.68%, 83.57%
- Organizations: 66.99%, 73.30%, 75.73%
2.2 Contextual factors: head

countries

- Or have you forgotten that *America* did once try to ban alcohol and look what happened!
- *Aruba* acquired separate status within the Kingdom of the Netherlands in 1986

organizations

- But in 1990 *Toyota*’s financial profit lengthened its lead over Honda and Nissan
- *Microsoft Corp*’s likely objections . . .
2.2 Contextual factors: head

![Bar chart showing percentages for countries and organizations]
2.2 Contextual factors

- Contextual factors like the function and head of a word captures
  - 85% of the variation in the country data, and
  - 78% of the variation in the organization data.

- Remaining variation?
  - Other variables: e.g., attachment information.
  - Data sparseness: semantic classes instead of words.

This statistical information can be used for the automatic recognition of metonymies in computational linguistics.
Overview

1. Introduction

2. A corpus-based perspective on metonymy
   2.1 General perspective
   2.2 Contextual factors

3. Metonymy recognition
   3.1 Metonymy recognition
   3.2 Active Learning
   3.3 Learning on the basis of related words

4. Conclusions and outlook
3.1 Metonymy recognition

Markert and Nissim

- Metonymy recognition as Word Sense Disambiguation
- Supervised recognition of metonymical country and organization names
- Grammatical and semantic information
- Successful approach: 87% for the country names, 76% for the organizations.

Problem
The supervised nature of the approaches hinders the development of a large-scale metonymy recognition system.
3.1 Metonymy recognition

Central question

- How can we reduce the number of manually labelled training examples?
- What data can we use in order to learn about metonymies?

Two solutions

- Active Learning
- Learning on the basis of words that are semantically related to one of the target senses
3.1 Metonymy recognition

Memory-Based Learning
solves a new problem by comparing it to related problems in its memory.

Learning phase
All labelled examples are stored in the memory.

Testing phase
The algorithm . . .
- compares the test example to all training examples,
- singles out the most similar training examples,
- and assigns their most frequent label.
3.1 Metonymy recognition
3.1 Metonymy recognition
3.2 Active Learning

Underlying idea
Active Learning automatically selects those examples that are most interesting to the classifier.

Algorithm

- Select and label a number of seed instances;
- Train a classifier on those seeds and have it label the unlabelled pool;
- Select and label those instances whose classification the classifier is most uncertain of;
- Repeat.
3.2 Active Learning

Uncertainty as distance

- Uncertainty usually defined as entropy or other P-based measure.
- But memory-based classifiers only output distances.
- Hypothesis: uncertainty $\sim$ distance

Distance-based active learning

- Randomly choose seeds
- On each round, add 10 unlabelled instances based on their distance from the seeds.
3.2 Active Learning
3.2 Active Learning
3.2 Active Learning
3.2 Active Learning
3.2 Active Learning

Positive

- Active Learning gives a reduction in manual annotation of ± 30%.
- Reduction will increase when we take more contextual information into account.

Less positive

- Algorithms should be tested on other data sets.
- There is still manual semantic annotation involved.
3.3 Learning on the basis of related words

- Both the literal and metonymical meanings of a word have words that are semantically related to them.
  - country names
    - literal \(\approx\) country
    - metonymical \(\approx\) people, inhabitants, government
  - organization/company names
    - literal \(\approx\) company, organization
    - metonymical \(\approx\) people, president, representative
  - author names
    - literal \(\approx\) author, writer
    - metonymical \(\approx\) book

- The meaning of a possible metonymy can be found by comparing its context to the contexts of those related words.
3.3 Learning on the basis of related words

This approach combines the advantages of supervised and unsupervised learning:

- Semantic labelling can proceed automatically; no manual annotation is needed.
- Thanks to the semantic labels, we can use supervised algorithms.
3.3 Learning on the basis of related words

Algorithm

- Divide the target data in 10 folds: 1 as development test set, 9 as final test set.
- Choose 500 ‘literal’ and 100 ‘metonymical’ examples.
- On each round, add 10 ‘metonymical’ examples and evaluate on the development test set.
- Use the training set with the best result.
- Evaluate on the final test set.
- Repeat 10 times.
3.3 Learning on the basis of related words

Experiments

<table>
<thead>
<tr>
<th>ORGANIZATIONS</th>
<th>LITERAL</th>
<th>METONYMICAL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>company</td>
<td>people</td>
</tr>
<tr>
<td></td>
<td></td>
<td>car</td>
</tr>
<tr>
<td>COUNTRIES</td>
<td>country</td>
<td>people</td>
</tr>
<tr>
<td>AUTHORS</td>
<td>author</td>
<td>book</td>
</tr>
</tbody>
</table>
3.3 Learning on the basis of related words
Problem of noise

- Automatic labelling introduces noise into the training set.
- Some noise can be removed by *scrubbing* (cf. Birke): If a feature vector occurs both as a literal and a metonymical training example, remove it
  - either from the literal set,
  - or from both sets.
3.3 Learning on the basis of related words

- Country: 79.68% (79.68%), 81.03% (81.03%), 84.12% (84.12%)
- Company: 66.99% (66.99%), 69.33% (69.33%), 72.55% (72.55%)
- Author: 78.51% (78.51%), 80.82% (80.82%), 82.33% (82.33%)
Overview

1. Introduction

2. A corpus-based perspective on metonymy
   2.1 General perspective
   2.2 Contextual factors

3. Metonymy recognition
   3.1 Metonymy recognition
   3.2 Active Learning
   3.3 Learning on the basis of related words

4. Conclusions and outlook
4. Conclusions and outlook

Theoretical perspective
A closer look at the contextual variables that influence the reading of some proper noun classes.

Computational perspective
Possible ways of reducing the amount of manual semantic annotation for metonymy recognition.

- Active Learning
  - relatively successful
  - considerable reduction of annotation load

- Learning on the basis of related words
  - reduces manual semantic annotation to zero.
  - still achieves high results.
4. Conclusions and outlook

Theoretical perspective

- investigate more variables
- introduce semantic information

Computational perspective

- AL: more variables, use of probability distribution
- Related words: extension to more data sets
For more information:
http://wwwling.arts.kuleuven.be/qlvl
yves.peirsman@arts.kuleuven.be