NLP model and tools for detecting and interpreting metaphors in domain-specific corpora

IsoMeta Project

P. Beust, S. Ferrari, V. Perlerin
GREYC UMR 6072 – University of Caen - France

Interdisciplinary Workshop on Corpus-Based Approaches to Figurative Language

Thursday 27th March 2003 - CORPUS LINGUISTICS 2003 - Lancaster
Talk overview

• Goals
• Framework
  – Metaphor
  – Lexical Representation
  – Isotopy
• Tools & Results
  – Browsing and Visualization
  – Examples
• Conclusion and further works
Goals

- Metaphor detection
- Computer-assisted metaphor interpretation

Metaphor study
- user-centered approach
- corpus-based approach

NLP model and Tools for detecting and interpreting metaphors in domain-specific corpora – P. Beust, S. Ferrari and V. Perlerin
A metaphor involves two concepts: a source concept and a target concept.

The Dow Jones is a stock exchange barometer.

NLP approaches for metaphor interpretation depend on how the relation is viewed:

» as an analogy
» as a novelty
Conventional metaphor

• Conventional conceptual metaphor

• The corpus
  – contains numerous examples of this conventional metaphor
  – also contains meteorological terms not used in a figurative way.

• In order to address the interpretation process of the metaphor, we focus on a conventional conceptual metaphor already observed in the corpus:
  « economics is meteorology »
Lexical representation

• The Anadia model:
  – is based on both componential and differential lexical categorization.
  – allows a user to produce his own descriptions of meaning components.

<table>
<thead>
<tr>
<th>Pressure zone</th>
<th>Zone</th>
<th>Pressure</th>
</tr>
</thead>
<tbody>
<tr>
<td>anticyclone</td>
<td>area</td>
<td>high</td>
</tr>
<tr>
<td>tropical wave</td>
<td>line</td>
<td>low</td>
</tr>
<tr>
<td>easterly wave</td>
<td>area</td>
<td>low</td>
</tr>
<tr>
<td>depression,</td>
<td>line</td>
<td>high</td>
</tr>
<tr>
<td>cyclone</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A *seme* with 2 opposite features

A *table* made from the combinatory of 2 semes features

A *set of tables* describing a domain

---

NLP model and Tools for detecting and interpreting metaphors in domain-specific corpora – P.Beust, S.Ferrari and V. Perlerin
Isotopy

- Isotopy is the recurrence of one *seme* in a linguistic unit.
- Isotopy is an easy and understandable way of expressing themes in linguistic units.
- Isotopy is the base of our interpretation and detection process.

*After starting off as a formidable tropical wave of the coast of Eastern Africa, it quickly developed into a depression* …  

The two recurring *semes* involve two isotopies that contribute to the meaning of the sentence.

The recurring features also show that the sentence deals with different kind of pressure zones.
Analysis

Building of sets of tables using **shared semes**

Results Analysis

Colouration of each set of tables

**Corpus tools**

NLP model and Tools for detecting and interpreting metaphors in domain-specific corpora – P. Beust, S. Ferrari and V. Perlerin
Corpus study tools (1)
Les Saints de glace BOURSE DE PARIS SEMAINE DU 11 AU 15 MAI

DESSOT ANDRÉ

Secteur rédaction : Economie Analytique
Sujets en France : BOURSE

SAINT Mamert, saint Pancrace et saint Servais, les fameux Saints de glace, doivent se frotter les mains. Leur passage annuel a, cette semaine, été couronné de succès. La température a très brutallement baissé, non seulement à l'extérieur, obligeant les plus friables à remettre une petite laine, mais aussi dans les affaires. La Bourse de Paris, qui depuis deux mois et demi patataugeait joyeusement dans une interminable consolidation, s'est repliée d'un bloc à toute allure vers la cote 430 de l'indice CAC, péniblement atteinte au début du mois de mars dernier. Le coup de froid a été sévère (-5, 5%) et d'autant plus inquiétant pour la "végétation mobilière" que le thermomètre a chuté de 4% au cours de la seule séance de vendredi.

Ce jour-là, pour se redonner du courage devant la débandade des cours, des professionnels laissent entendre: "Appelez la police pour arrêter le massacre!" Inutile de le préciser; aucune main n'a secouru du 15 mai. Cette semaine n'avait pourtant pas mal commencé. Avec l'assèchement de la tendance, s'était une alarme qui s'allumaït. Le thermomètre a chuté de 4% au cours de la seule séance de vendredi.

Outils de mesure
Axe : température
The Dow Jones, for instance, the thermometer of Wall Street, which had fallen...

thermometer is used in the same way as “graphics”, “ratio”... i.e. to suggest an object for analysis and study in the stock market domain.

 analogy point of view
This crash was due (...) to the vertiginous and uncontrolled fall of the dollar, sign.

storm is used to evoke a not only bad but also violent dynamic phenomenon in the stock market domain.

Isotopies at the storm will henceforth affect the financial markets.
Conclusion and further works

• IsoMeta Project is still in progress:
  - automatic detection
  - interpretation process
  - results evaluation

• Shared semes represent previous user’s knowledge on a conceptual metaphor
  ► "new metaphors" are not covered by our approach.
NLP model and tools for detecting and interpreting metaphors in domain-specific corpora

IsoMeta Project

P. Beust, V. Perlerin, S. Ferrari
GREYC UMR 6072 – University of Caen - France

http://www.info.unicaen.fr/~perlerin/recherche/Anadia.html

Interdisciplinary Workshop on Corpus-Based Approaches to Figurative Language

Thursday 27th March 2003 - CORPUS LINGUISTICS 2003 - Lancaster