Measuring frame relatedness

Michael Wirth

Department of Computational Linguistics
Saarland University
miwirth@coli.uni-sb.de

(joint work with Marco Pennachiotti)

FEAST meeting
January 7th, 2009
Outline

- Introduction
  - Motivation
  - Related work
- Frame relatedness
  - FrameNet
  - Manually ranking Frames
- Frame relatedness measures
- Conclusions
What it's all about?

Defining and proving a notion of frame relatedness

Developing frame relatedness measures
Motivation

● Ontologies and taxonomies provide precise information about the relationship of words and concepts.

● Semantic similarity / relatedness is a important component in knowledge-driven applications.

● Support of advanced knowledge-driven NLP tasks

● Resources available
Related Work

Semantic relationships between words are intensively researched

- structure-based approaches
- information-based approaches

Word Similarity

Two words are *related* if any type of relation stands between them.

Two words are *similar* if they are connected through an “is-a” like relation.

Budanitsky and Hirst 2006
Frame relatedness

Definition:

1) Two frames are *similar* if they are linked via „is-a“ like relations.

2) Two frames are *related* if any relation stands between them.
Frame

| Frame name | Definition |
|------------|------------|
| **Frame: STATEMENT** | This frame contains verbs and nouns that communicate the act of a **SPEAKER** to address a **MESSAGE** to some **ADDRESSEE** using language. A number of the words can be used performatively, such as *declare* and *insist*. |

| Semantic roles | Frame Elements |
|----------------|----------------|
| **SPEAKER** | Evelyn said she wanted to leave. |
| **MESSAGE** | Evelyn announced **that she wanted to leave**. |
| **ADDRESSEE** | Evelyn spoke **to me** about her past. |
| **TOPIC** | Evelyn’s **statement** **about her past** |
| **MEDIUM** | Evelyn **preached** to me **over the phone**. |

| Words associated with the frame | Lexical Units |
|---------------------------------|---------------|
| **LUS**                         | acknowledge.v, acknowledgment.n, add.v, address.v, admission.n, admit.v, affirm.v, affirmation.n, allegation.n, allege.v, announce.v, ... |
FrameNet
(Ruppenhofer et al. 2005)

- Frames are concepts describing situations / events
- Frames are the basic units of FrameNet
- Structured semantic lexicon
- Words (lexical units) associated with frames
- ~ 10 000 lexical units
- ~ 800 Frames
- ~ 135 000 annotated sentences (BNC corpus)
Frame-to-Frame-Relations

- hierarchical
  - Inheritance
  - Using
  - Subframe
- non-hierarchical
  - Perspective_on
  - Causative_of
  - Inchoative_of
  - Precedes
  - See_also
FrameNet Hierarchy
(Release 1.3)

- Number of frames : 795
- Number of roots : 86
- Number of isolated nodes : 7
- Number of independent sub-graphs : 26
FrameNet Hierarchy
(Release 1.3)

- Number of frames reachable from more than 1 root: 559
- Number of incoming edges: 1136
- Average number of edges per frame: 2.86
- Maximum path length: 15
Manually ranking frames

Experiment

• 15 subjects -> 155 frame pairs
  – 15 frame pairs
    • 10 judged only by this subject
    • 5 judged by all subjects

• sort the pairs according to their similarity

• rate every pair on a scale from 0 to 4
Manually ranking frames
Experiment

• Data

– A set of controlled frame pairs
  → Controlled Set (155 pairs)

– A set of randomly selected frame pairs
  → Simple Set (155 pairs)
Manually ranking frames

Results

• high significant correlation among the annotators
  – Simple Set $\tau = 0.600 \quad \alpha < 0.005$
  – Controlled Set $\tau = 0.547 \quad \alpha < 0.005$

$\rightarrow$ The notion of “frame relatedness” is intuitive and principled for humans.
Manually ranking frames

Results

• significant correlation on the gold standard ranking
  - Simple Set gold standard
    \[ \tau = 0.530 \quad \text{StdDev} = 0.146 \quad \alpha < 0.01 \]
  - Controlled Set gold standard
    \[ \tau = 0.566 \quad \text{StdDev} = 0.173 \quad \alpha < 0.01 \]

  gold standard ranking is reliable
Manually ranking frames
Gold Standards

| SIMPLE SET                        | CONTROLLED SET                                   |
|-----------------------------------|--------------------------------------------------|
| Measure volume - Measure mass (1) | Knot creation - Rope manipulation (1,5)         |
| Communication manner - Statement (2) | Shoot projectiles - Use firearm (1,5)         |
| Giving - Sent items (3)           | Scouring - Scrutiny (3)                         |
| Abundance - Measure linear extent (4) | Ambient temperature - Temperature (4)         |
| Remembering information - Reporting (5) | Fleeing - Escaping (5)                      |
| ...                               | ...                                             |
| Research - Immobilization (126)   | Reason - Taking time (142)                       |
| Resurrection - Strictness (126)   | Rejuvenation - Physical artworks (142)          |
| Social event - Word relations (126) | Revenge - Bungling (142)                     |
| Social event - Rope manipulation (126) | Security - Likelihood (142)             |
| Sole instance - Chatting (126)    | Sidereal appearance - Aggregate (142)           |

Human gold standard ranking: first and last 5 pairs (in brackets ranks allowing ties)
Frame relatedness measures

| Measure                  |                      |
|--------------------------|----------------------|
| wn_jcn                   |                      |
| wn_hso                   | WordNet-based measures|
| cr_occ_sent              |                      |
| cr_wgt_sent              | corpus-based measures|
| cr_occ_doc               |                      |
| cr_wgt_doc               | distributional measure|
| cr_dist_doc              |                      |
| hr_wu                    | FrameNet-based measures|
| hr_hso                   |                      |
| hr_fe                    |                      |
| `def overlap baseline`   |                      |
| `LU overlap baseline`    |                      |
| `human upper bound`      |                      |
Frame relatedness measures

- **WordNet-based measures**
  - map lexical units to WordNet senses
  - calculate the sense similarity using two different WordNet similarity measures:
    - *wn_jcn*: Jiang and Conrath (1997)
    - *wn_hso*: Hirst and St.Onge (1998)
Frame relatedness measures

- corpus-based measures
  - using the SemCor corpus
  - calculate the point wise mutual information of two frames
    - simple point wise mutual information
      - $cr_{occ}$
    - weighted point wise mutual information
      - $cr_{wgt}$
  - using two different types of context
    - sentences $\rightarrow cr_{occ\_sent}, cr_{wgt\_sent}$
    - documents $\rightarrow cr_{occ\_doc}, cr_{wgt\_doc}$
Frame relatedness measures

- distributional measure
  - using the TREC-2002 Vol.2 corpus
  - each frame is modelled by a distributional vector
  - documents are dimensions
  - the value of a dimension expresses the association between the document and the frame
  - relatedness computed by using cosine similarity
Frame relatedness measures

Results

| Measure                   | Simple Set | Controlled Set |
|---------------------------|------------|----------------|
| wn_jcn                    | 0.114      | 0.141          |
| wn_hso                    | 0.106      | 0.141          |
| cr_occ_sent               | 0.239      | 0.340          |
| cr_wgt_sent               | **0.281**  | **0.349**      |
| cr_occ_doc                | 0.143      | 0.227          |
| cr_wgt_doc                | 0.173      | 0.240          |
| cr_dist_doc               | 0.152      | 0.240          |
| hr_wu                     | 0.139      | 0.286          |
| hr_hso                    | 0.134      | 0.296          |
| hr_fe                     | 0.252      | 0.326          |
| *def overlap baseline*    | 0.056      | 0.210          |
| *LU overlap baseline*     | 0.080      | 0.253          |
| *human upper bound*       | 0.530      | 0.566          |

Correlation measure with gold standard using Kendall's Tau
Frame relatedness measures
Evaluation

- WordNet-based measures

| Measure                   | Simple Set | Controlled Set |
|---------------------------|------------|----------------|
| wn_jcn                    | 0.114      | 0.141          |
| wn_hso                    | 0.106      | 0.141          |
| def overlap baseline      | 0.056      | 0.210          |
| LU overlap baseline       | 0.080      | 0.253          |
| human upper bound         | 0.530      | 0.566          |

- fail to predict relatedness for many pairs
  - wn_hso assigns zero to 137 (Simple Set) pairs and 119 (Controlled Set) pairs
- WordNet misses situational relations
- 18% of FrameNet LUs are adjectives or adverbs
- 7% of verbal FrameNet LUs don't have a WordNet mapping
Frame relatedness measures

Evaluation

- corpus-based measures

| Measure                | Value 1 | Value 2 |
|------------------------|---------|---------|
| cr_occ_sent            | 0.239   | 0.340   |
| cr_wgt_sent            | 0.281   | 0.349   |
| cr_occ_doc             | 0.143   | 0.227   |
| cr_wgt_doc             | 0.173   | 0.240   |
| cr_dist_doc            | 0.152   | 0.240   |
| def overlap baseline   | 0.056   | 0.210   |
| LU overlap baseline    | 0.080   | 0.253   |
| human upper bound      | 0.530   | 0.566   |

- correlation decreases using documents as context
- corpus-based measures promote frame pairs in non-hierarchical relations
- the distributional measure promotes frame pairs in hierarchical relations
Frame relatedness measures

Evaluation

- FrameNet hierarchy is a good indicator for frame relatedness
- hierarchy-based measures promote pairs related by diverse relations
- measures slightly penalized by low coverage
Conclusions

- the notion of frame relatedness is cognitively principled
- introduce a variety of measures for automatically estimating frame relatedness
- measures offer good performance (significant at the 99% level)
Bibliography

• Marco Pennachiotti and Michael Wirth. 2009. Measuring frame relatedness. In: Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2009). Athens, Greece. to appear.
Frame relatedness measure

WordNet-based measures

\[
wn(F_1, F_2) = \sum_{s_1 \in S_{F_1}} \sum_{s_2 \in S_{F_2}} wn\_rel(s_1, s_2)
\]

\[
S_F : \text{the set of WordNet senses mapping to lexical units of frame } F
\]

\[
wn\_rel(s_1, s_2) : \text{a function estimating the relatedness of two WordNet senses}
\]

used functions: Jiang and Conrath (1997) wn_jcn
Hirst and St.Onge (1998) wn_hso
Frame relatedness measure

corpus-based measures

\[ cr_{occ}(F_1, F_2) = \log_2 \frac{|C_{F_1,F_2}|}{|C_{F_1}||C_{F_2}|} \]

\[ C_{F_i} = \{c \in C : \exists l_{F_i} \text{ in } c\} \]

\[ C_{F_1,F_2} = \{c \in C : \exists l_{F_1} \text{ and } \exists l_{F_2} \text{ in } c\} \]

\( C \) : the corpus
\( c \) : context in Corpus \( C \)
\( l_{F_i} \) : a lexical unit of frame \( F \)
Frame relatedness measure

corpus-based measures

$$cr_{wgt}(F_1, F_2) = \log_2 \frac{\sum_{c \in C_{F_1,F_2}} w_{F_1}(c) \cdot w_{F_2}(c)}{\sum_{c \in C_{F_1}} w_{F_1}(c) \cdot \sum_{c \in C_{F_2}} w_{F_2}(c)}$$

$$w_F(c) = \arg \max_{l_F \in L_F} P(S_{l_F} | l_F)$$

$$P(S_{l_F} | l_F) = \frac{|S_{l_F}|}{|S_l|}$$

$L_F$: the set of LUs of frame $F$
Frame relatedness measure

FrameNet-based measures

\[ hr \_wu(F_1, F_2) = \frac{2 \cdot dp(LCS)}{ln(F_1, LCS) + ln(F_2, LCS) + 2 \cdot dp(LCS)} \]

\( dp(F) \) : the depth of frame \( F \) in the FrameNet hierarchy

\( ln(F_1,F_2) \): the path length between the frames \( F_1 \) and \( F_2 \)

\( LCS \) : least common subsumer
Frame relatedness measure

FrameNet-based measures

\[ hr_{hso}(F_1, F_2) = M - \text{path length} - k \cdot d \]

\( M, k : \text{constants} \)
\( d : \text{number of changes of direction} \)
Frame relatedness measure

FrameNet-based measures

\[ hr_{fe}(F_1, F_2) = \frac{|FE_1 \cap FE_2|}{\max(|FE_1|, |FE_2|)} \]

FE : the set of frame elements