Towards a Better Semantic Role Labeling of Complex Predicates

Glorianna Jagfeld*, Lonneke van der Plas*
*IMS, University of Stuttgart, *IoL, University of Malta

Introduction

• Semantic Role Labeling (SRL) determines who did what to whom in sentences by identifying and associating predicates with their semantic arguments
• Complex predicates (CPs):
  – Multi-headed, composed of more than one grammatical element
  – Frequent phenomenon: 1 CP in every 5th sentence in the Wiki50 corpus
• Most frequent types:
  – VPCs: verb particle constructions, e.g. watch out, 50% of CPs in Wiki50
  – LVCs: light verb constructions, e.g. take a walk, 40% of CPs in Wiki50

• The PropBank corpus (PB): [Palmer et al., 2005]
  – One of the most widely used resources for training data for SRL systems
  – Poor coverage and treatment of CPs (e.g. take a hard line, take time and other take CPs annotated with the most frequent sense of take denoting acquire, come to have, choose, bring with you from somewhere)

Method

• Automatically finding aliases for CPs is related to finding (near-) synonymous predicates and their accompanying roleset for the CPs
• Apply the distributional hypothesis: assess the similarity of expressions by looking at their contexts
• Multilingual variant: multilingual contexts work better for synonym acquisition than monolingual syntactic contexts [Van der Plas and Tiedemann, 2006]

Parallel corpus
  + lemmas
  + POS
  + synt. dependencies
  + SRL
  + word alignments

Extract CPs and PB roleset groups

Populate matrix with translation counts (alignments of CPs and PB roleset groups)

For each CP vector calculate similarity with each PB roleset group vector

Alias: roleset with the highest similarity score

Overview of the alias finder

• The English portion of the Europarl corpus was processed with the MATE tools
• Word alignments (grow-diag-final-and) are taken from the OPUS project
• Pre-defined syntactic patterns used for finding occurrences of LVCs and VPCs in the corpus
• PB roleset group: all predicates that share a PB roleset (e.g. all verbs that were assigned to care/01 by the PB roleset system belong to the PB roleset group of care/01)

• Use the cosine similarity to measure the similarity between CPs and PB roleset groups

Alias PB roleset for the predicate take care:
  Roleset id: care.01 ‘to be concerned’
  Arg0: carer, agent
  Arg1: thing cared for/about

Representing CPs for SRL

• The current PB representation treats CPs as lexical usages of the verb

Frank takes care of business
take.01 care.01

Frank takes care of business
(take+care).01

• (Manually) creating new rolesets for CPs is cumbersome
• Aliasing: map CPs to existing PB entries with the same meaning and argument structure (take care → care.01 ‘to be concerned’) [Bonal et al., 2014]
• Our approach: Speed up the mapping process with an automatic system

Human Annotation

• 4 annotators independently annotated 100 CPs (50 VPCs and 50 LVCs) from the Wiki50 corpus in their original contexts
• Set was balanced for frequency
• All occurrences (197) of the CPs are included → account for polysemy
• Disagreements were discussed to obtain a consensus

IAA
A & B | A & C | A & D | Micro Average
Strict 87% 51% 14% 53%
Lenient 85% 74% 87% 75%

Decision # CP example
alias 96 take part
mw PB pred. 60 open up
compositional 18 obtain permission
no alias found 16 go into politics
7 taken control

• [aliased] PB roleset as true alias
• [mw. PB pred] PB roleset for this CP already exists

Evaluation Setup

• Evaluate on alias, and ‘mw. PB pred’ of the annotation effort (total 154, Wiki50 set) + 70 take CPs from Bonial et al. [2014] (take set)
• Strict accuracy: predicted alias corresponds to one of the gold aliases → Synonymous rolesets are counted as incorrect
• Lenient acc.: predicted alias belongs to same VerbNet class as gold alias
• Baseline selects for every CP first roleset (if available) of the (LVCs) verb predicate corresponding to the noun (take care → care.01) (VPCs) PB multwort predicate (open up → open.03)
• Very strong and informed baseline, only fails in lack of coverage

Future Work

• More parallel data from different domains
• Apply frequency weights to the matrix
• Include more divergent languages
• Retrain SRL system on the new annotations

Selected References


Selected References

Van der Plas and Tiedemann, 2006

Results and Discussion

<table>
<thead>
<tr>
<th>Set</th>
<th>Strict Cov</th>
<th>Strict Acc</th>
<th>Strict Hm</th>
<th>Len. Cov</th>
<th>Len. Acc</th>
<th>Len. Hm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wiki50</td>
<td>98.7</td>
<td>44.1</td>
<td>60.9</td>
<td>98.0</td>
<td>69.0</td>
<td>81.0</td>
</tr>
<tr>
<td></td>
<td>(65.6)</td>
<td>(54.5)</td>
<td>(59.5)</td>
<td>(59.5)</td>
<td>(85.9)</td>
<td>(70.3)</td>
</tr>
<tr>
<td>alias</td>
<td>98.9</td>
<td>36.6</td>
<td>53.4</td>
<td>98.4</td>
<td>60.0</td>
<td>74.5</td>
</tr>
<tr>
<td></td>
<td>(50.0)</td>
<td>(34.0)</td>
<td>(40.5)</td>
<td>(40.5)</td>
<td>(68.6)</td>
<td>(51.0)</td>
</tr>
<tr>
<td>mw PB pred.</td>
<td>98.3</td>
<td>55.9</td>
<td>71.3</td>
<td>97.6</td>
<td>62.5</td>
<td>89.4</td>
</tr>
<tr>
<td></td>
<td>(86.7)</td>
<td>(71.2)</td>
<td>(78.1)</td>
<td>(84.6)</td>
<td>(97.7)</td>
<td>(90.7)</td>
</tr>
<tr>
<td>take</td>
<td>67.1</td>
<td>25.5</td>
<td>37.0</td>
<td>56.6</td>
<td>60.0</td>
<td>58.3</td>
</tr>
<tr>
<td></td>
<td>(71.4)</td>
<td>(32.0)</td>
<td>(44.2)</td>
<td>(64.9)</td>
<td>(45.0)</td>
<td>(53.8)</td>
</tr>
</tbody>
</table>

Percentage coverage (Cov), accuracy (Acc) and the harmonic mean (Hm) of coverage and accuracy of the predicted aliases in the Wiki50 set (+ its two subsets) and the take set; the results of the baseline are in brackets

• The system always outperforms the baseline in terms of coverage
• Beats the baseline in terms of strict accuracy for the alias subset
• Performance is quite close to the IAA from human annotation

Error Analysis

1. Domain specificity of the corpus → low frequency of some gold aliases (e.g. melt down → fuse.01 ‘melt into lump’)