Align-smatch: A Novel Evaluation Method for Chinese Abstract Meaning Representation Parsing based on Alignment of Concept and Relation

1. From AMR to Chinese AMR

Abstract Meaning Representation (AMR): 
- Sentence’s semantic structure
- Rooted, directed, acyclic graph with nodes (concepts) and labels on arcs (relations)
- Lack of explicit alignments required for AMR parsing

He wants to see the show in Beijing. 他想在北京看演出。

Chinese AMR (CAMR): 
- Concept Alignment: number after “x” in node’s ID is the index of the aligned word in sentence
- Relation Alignment: the function word on arcs, whose label has the same function as the function word

2. Smatch

Smatch\(^{(1)}\) is a popular evaluation method for AMR parsing.

**Process:**
- Reset node’s ID and transform two AMRs into two triple sets, each of which consists of node triple, arc triple and node property triple

Num(Triples of Parsed)=4
Num(Triples of Gold)=9

- Use hill-climbing method to find the optimal matching number of two triple sets

Num(Match Triples)=2

- Calculate Precision(P), Recall(R) and F1 score

**Shortcoming:**
- Concepts are not considered when comparing two arcs and two top nodes
- Alignments cannot be evaluated

3. Align-smatch

**Modify Smatch:**
- Two arc triples can match only if the concepts at both ends are same
- Transform top node into arc triple instead of node property triple

**Add tuples for alignments:**
- Concept alignment: anchor(Node, ID)
- Relation alignment: (Word, ID, Source Node, Target Node)

<table>
<thead>
<tr>
<th>Align-smatch Tuples</th>
<th>Node</th>
<th>Arc</th>
<th>Prop.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Match Tuples**

- Smatch scores
  - P=0/13, R=0/7, F1=0

4. Smatch VS Align-smatch

**Data:**
- 100 sentences/CAMR pairs from CAMR1.0 corpus (LDC2019T07) as data G
- Two annotators re-annotate 100 sentences as data A/B

**Metric:**
- Smatch, Align-smatch
- Concept-smatch: only put triple of concept alignment into Smatch
- MOD: the version that perfects the shortcoming of Smatch

**Results:**
- Performance under MOD version is worse, but MOD version is more tough
- With the help of annotation platform, performance under Concept-smatch is better than Smatch
- Performance under Align-smatch is worse, which indicates the difficulty of annotating relation alignment

**Conditions Inducing Errors:**
- More than one function words align one relation
- One function word can align many kinds of relations

**References**