The Relative Clauses AMR Parsers Hate Most

Xiulin Yang    Nathan Schneider

Georgetown University

May 19, 2024
How effectively do AMR parsers handle different types of English relative clauses (RCs)?
I know the person who you like.
Relative Clauses

I know the person who you like.
Long Distance Dependency (LDD)

I know the person who you like.
Long Distance Dependency (LDD)

I know the person who you think that Tom mentioned once that ___ likes you.
Abstract Meaning Representation (AMR)

- AMR is a graph semantic representation that captures the core semantic roles and relations in a sentence.
  - Usually *who did what to whom, where and when*.
- Each AMR is a single rooted, directed graph, which can be represented with Penman Notation.
- *I know the person who likes you*
Abstract Meaning Representation (AMR)

- AMR is a graph semantic representation that captures the core semantic roles and relations in a sentence.
  - Usually *who did what to whom, where and when*.
- Each AMR is a single rooted, directed graph, which can be represented with Penman Notation.
- *I know the person who likes you*

\[(k / know-01
:ARG0 (i / i)
:ARG1 (p / person
 :ARG0-of (l / like-01
  :ARG1 (y / you)))\]
Relative Clause in AMRs

- *I know the person who likes you.*

**Figure:** Canonical AMR graph. The ARG0-of edge corresponds to the relative clause.
Relative Clause in AMRs

- *I know the person who likes you.*

**Figure:** Canonical AMR graph. The ARG0–of edge corresponds to the relative clause.

**Figure:** Normalized AMR graph. The ARG0 edge from like-01 to person corresponds to the relative clause.
Relative Clause Types

- **Subject RC:**
  He is the person who __ stole my book.

- **Object RC:**
  He is the person that you like __.

- **Oblique RC:**
  He is the person that I borrowed the book from __.

- **Passive Subject RC:**
  He is the person who __ is liked by you.
Relative Clause Types

- **Subject RC:**
  He is the person who __ stole my book.

- **Object RC:**
  He is the person that you like __.

- **Oblique RC:**
  He is the person that I borrowed the book from __.

- **Passive Subject RC:**
  He is the person who __ is liked by you.
Reduced Relative Clause Types

- Reduced Object RC:
  He is the person you like __.

- Reduced Oblique RC:
  He is the person I borrowed the book from __.
“AMR parsing is far from solved” (Groschwitz et al., 2023)

- SOTA AMR Parser (Lee et al., 2022) achieved over 0.85 in Smatch (Cai and Knight, 2013).
- Relying solely on overall F-scores does not fully reveal a parser’s performance across different linguistic phenomena (Groschwitz et al., 2023).
- Seq2seq models that simply take input as sequence string fail at structural generalization compared with models that explicitly encode structural information (Yao and Koller, 2022; Li et al., 2023; Shaw et al., 2021).
- Recovering reentrancy structures is a challenge for AMR parsers (Szubert et al., 2020; Damonte et al., 2017).
“AMR parsing is far from solved” (Groschwitz et al., 2023)

- SOTA AMR Parser (Lee et al., 2022) achieved over 0.85 in Smatch (Cai and Knight, 2013).
- Relying solely on overall F-scores does not fully reveal a parser’s performance across different linguistic phenomena (Groschwitz et al., 2023).
- Seq2seq models that simply take input as sequence string fail at structural generalization compared with models that explicitly encode structural information (Yao and Koller, 2022; Li et al., 2023; Shaw et al., 2021).
- Recovering reentrancy structures is a challenge for AMR parsers (Szubert et al., 2020; Damonte et al., 2017).
Research Questions

- How well can AMR parsers capture the long-distance predicate-argument dependencies in RCs?
  - Does structure-awareness help the models to parse?
  - Which types of RC are most challenging and why?
Relative clause

- **Subject RC:**
  He is the person who **__** stole my book.

- **Object RC:**
  He is the person that you **__** like.

- **Oblique RC:**
  He is the person that I **__** borrowed the book from.

- **Passive Subject RC:**
  He is the person who **__** is liked by you.

- **Reduced Object RC:**
  He is the person you **__** like.

- **Reduced Oblique RC:**
  He is the person I **__** borrowed the book from.
Method

- Datasets
- Models
- Evaluation Metric
## Datasets

<table>
<thead>
<tr>
<th>Dataset</th>
<th># sents</th>
<th># tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>EWT (Silveira et al., 2014)</td>
<td>1,449</td>
<td>26.5</td>
</tr>
<tr>
<td>CRC (Prasad et al., 2019)</td>
<td>1,400</td>
<td>13.7</td>
</tr>
<tr>
<td>AMR 3.0 (Knight et al., 2021)</td>
<td>259</td>
<td>29.1</td>
</tr>
</tbody>
</table>

**Table:** Number of sentences containing RCs in the datasets and the mean sentence length
## Datasets

<table>
<thead>
<tr>
<th>Dataset</th>
<th># sents</th>
<th># tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>EWT (Silveira et al., 2014)</td>
<td>1,449</td>
<td>26.5</td>
</tr>
<tr>
<td>CRC (Prasad et al., 2019)</td>
<td>1,400</td>
<td>13.7</td>
</tr>
<tr>
<td>AMR 3.0 (Knight et al., 2021)</td>
<td>259</td>
<td>29.1</td>
</tr>
</tbody>
</table>

**Table:** Number of sentences containing RCs in the datasets and the mean sentence length
Models

▶ Structure-aware models
  ▶ AM-Parser (Groschwitz et al., 2018): compositional parser composed of a supertagger + dependency parser
  ▶ AMRBART (Bai et al., 2022): structural pretraining + fine-tuning

▶ Structure-unaware models
  ▶ Spring (Spring et al., 2021)
  ▶ amrlib-BART\(^1\)
  ▶ amrlib-T5

▶ All models are fine-tuned on AMR 3.0.

\(^1\)https://github.com/bjascob/amrlib
Evaluation Metric - Reentrancy recall

- Our evaluation assesses whether the relativized noun in a sentence is **reentrant**, with two incoming edges—one originating from the main clause’s predicate verb and another from the predicate within the RC.

- To do so, we use LEAMR (Blodgett and Schneider, 2021), a probabilistic, fine-grained aligner optimized for English AMR.


![Figure: Normalized AMR graph for the sentence I know the person who likes you.](GNORMXG graph)

**Figure:** Normalized AMR graph for the sentence *I know the person who likes you.*
I know the person who likes you.

Figure: Correct prediction ✓
I know the person who likes you.

**Figure:** Correct prediction ✓

**Figure:** Incorrect prediction ×
Figure: RC reentrancy recall of AM-Parser, amrlib-BART and AMRBART, by RC subtype and overall.
Relative Clause Types

Figure: RC reentrancy recall of all parsers by RC subtype and overall.
Why contributes to such discrepancies?

**Figure:** Average Dependency Distance vs Mean Recall across RC Types.
Why contributes to such discrepancies?

Figure: Average Dependency Distance vs Mean Recall across RC Types.
Why contributes to such discrepancies?

**Figure**: Average Dependency Distance vs Mean Recall across RC Types.
Takeaways

▶ Does structure-awareness help the models to parse?
▶ Which types of RC are most challenging and why?
Takeaways

- **Does structure-awareness help the models to parse?**
  - Seq2seq models, on the whole, outperform the compositional model.
  - There is little difference in performance between seq2seq models that are aware of structure and those that are not.

- **Which types of RC are most challenging and why?**
  - Relative clauses are challenging for current parsers.
  - Reduced RCs are the most challenging RC types.
  - The full RCs with shorter dependency distances are easier to parse.
  - Linguistic cues?
Thank you for your attention!


Other RC Types

- Free relatives (e.g., *I heard what you said*)
- Possessive RCs (e.g., *I like the girl whose dress is blue*)
- Reduced subject RCs (e.g., *I met the person you mentioned finished all the work this week*)
- Adnominal participial clauses (e.g., *the sheep eaten by wolves*)
Figure: RC reentrancy recall (solid lines) and attainability rate (dashed) of all parsers, by RC subtype and overall.
## Dependency Distances and Counts across RC Types

<table>
<thead>
<tr>
<th>RC Category</th>
<th>Dep Dist</th>
<th>Count</th>
<th>Mean Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reduced oblique RC</td>
<td>3.06</td>
<td>1,092</td>
<td>41.9</td>
</tr>
<tr>
<td>Reduced object RC</td>
<td>3.13</td>
<td>1,371</td>
<td>56.6</td>
</tr>
<tr>
<td>Subject RC</td>
<td>4.30</td>
<td>4,226</td>
<td>64.5</td>
</tr>
<tr>
<td>Passive Subject RC</td>
<td>5.78</td>
<td>534</td>
<td>78.0</td>
</tr>
<tr>
<td>Object RC</td>
<td>5.21</td>
<td>516</td>
<td>61.2</td>
</tr>
<tr>
<td>Oblique RC</td>
<td>6.98</td>
<td>729</td>
<td>55.2</td>
</tr>
</tbody>
</table>

**Table:** Mean dependency distance of 6 types of RCs in our experiments