Can AMR Assist Legal and Logical Reasoning? Nikolaus Schrack, Ruixiang Cui, Hugo A. López, Daniel Hershcovich

08.12.2022

Semantic Parsing

- Transforming natural text into a formal meaning representation e.g. Abstract Meaning Representation (Banarescu et al., 2013) • Represents the meaning of a sentence as a graph



```
(s / see-01
    :ARGO (h / he)
    :ARG1 (b / book
        :poss h))
```

Idea & Approach

Understanding the logic in law is a major challenge in legal NLP

Can AMR help capture logical relationships?

- Theoretical analysis → might help with some logical relationship but not others
- Create model architectures utilizing AMR and run experiments on legal and logical reasoning tasks

Tasks

Multiple Choice

CaseHOLD (Zheng et al., 2021)

- Court decision statement
- Find the correct holding

LogiQA (Liu et al., 2020)

- Context & Question
- Find correct answer

violence of 18 usc 16

Court decision statement

Drapeau's cohorts, the cohort would be a "victim" of making the bomb. Further, firebombs are inherently dangerous. There is no peaceful purpose for making a bomb. Felony offenses that involve explosives qualify as "violent crimes" for purposes of enhancing the sentences of career offenders. See 18 U.S.C. $\S924(e)(2)(B)(ii)$ (defining a "violent felony" as: "any crime punishable by imprisonment for a term exceeding one year ... that ... involves use of explosives"). Courts have found possession of a bomb to be a crime of violence based on the lack of a nonviolent purpose for a bomb and the fact that, by its very nature, there is a substantial risk that the bomb would be used against the person or property of another. See United States v. Newman, 125 F.3d 863 (10th Cir.1997) (unpublished) (<HOLDING>); United States v. Dodge, 846 F.Supp. 181,

Holding Statement 1 (correct)

holding that possession of a pipe bomb is a crime of violence for purposes of 18 usc 3142f1

Holding Statement 2 (not correct)

holding that bank robbery by force and violence or intimidation under 18 usc 2113a is a crime of violence

Holding Statement 3 (not correct)

holding that sexual assault of a child qualified as crime of violence under 18 usc 16

Holding Statement 4 (not correct)

holding for the purposes of 18 usc 924e that being a felon in possession of a firearm is not a violent felony as defined in 18 usc 924e2b

Holding Statement 5 (not correct)

holding that a court must only look to the statutory definition not the underlying circumstances of the crime to determine whether a given offense is by its nature a crime of violence for purposes of 18 usc 16

Example of CaseHOLD

Model trained on AMR graphs AMRBART (Bai et al., ACL 2022)

- Based on BART (Lewis et al., 2020): a Seq2Seq denoising auto-encoder
- Further pre-trained on linearized AMR graphs
- Uses only AMR input

Fusion Model

- Combine Text and AMR input
- Extract embeddings from PLMs
- Use concatenation



Results **CaseHOLD – Legal Reasoning**

• Fusion model performs similar to base-line

Model	Input	Model Size	Accuracy
LegalBERT _{SMALL}	Text	35M	0.72 ¹²
LegalBERT _{SMALL} + adapter	Text	35M	0.73
BARTBASE	Text	139M	0.74
LegalBERT _{SMALL} + adapter	AMR (linearised and simplified)	35M	0.53
Smatch Model	AMR (Penman)	-	0.34
AMRBARTBASE	AMR (Spring prepr.)	142M	0.51
Fusion Model	Text and AMR (Spring prepr.)	252M	0.74

Results LogiQA – Logical Reasoning

AMR models underperform compared to base-line

Model	Input	Model Size	Accuracy
Random			0.25
BERTBASE	Text	1 39M	0.28
$\mathbf{AMRBART}_{\text{BASE}}$	AMR (Spring prepr.)	142M	0.27
Fusion Model	Text and AMR (Spring prepr.)	252M	0.27

Error Analysis Parser Quality

- Missing sentences in parsed AMRs
- lacksquare



Correlation between "missing information" in parsed AMR & accuracy



Discussion

Performance

- Cross-domain AMR parsing is a known problem (Bai et al. 2022)
 domain specific pretraining
 fine tuning on in-domain silver AMRs
- Fusion Model: fusion mechanism too simple

 use co-attention layer (Siriwardhana et al., 2020)