Yale

Structural Semantic Fusion with Graph Transformers for Abstractive Summarization

Borui Wang[1], Alexander R. Fabbri[1], Ziming Mao[1], Michal Lewkowicz[1], Haoran Li[2], Yashar Mehdad[2], Dragomir Radev[1]

[1]Department of Computer Science, Yale University, [2]Facebook AI Research

LILY Lab

Introduction

We propose a new system for abstractive summarization of natural language text based on structural semantic fusion. Our new system is motivated by the popular need to improve the faithfulness of abstractive summarization models and the potential to better utilize the rich structural information contained in the source text. We first perform discourse parsing over the source text on the elementary discourse unit (EDU) level, and then conduct semantic parsing over the elementary discourse units to obtain their abstractive meaning representation (AMR) graphs. We then train a graph-transformer-based encoder-decoder model to perform recursive semantic fusion to merge intermediate AMR summary graphs over the discourse parsing tree of the source text in order to obtain a final summary AMR graph at the root of the tree. In the final step, we employ an AMR-to-text verbalizer to generate the final summarization result from the final summary AMR graph.

Discourse and Semantic Parsing

In order to let the summarization model gain deeper insights into the rich structural information contained in the source text, our system operates over the discourse-parsed and semantically parsed representations of the source text. More specifically, we first run a discourse parser to obtain a discourse parsing tree of elementary discourse units (EDUs), and then run a semantic parser to parse each EDU into their corresponding abstract meaning representation (AMR) graphs. Each EDU-AMR graph would then be further converted into their corresponding PENMAN notation for easier processing by downstream models.

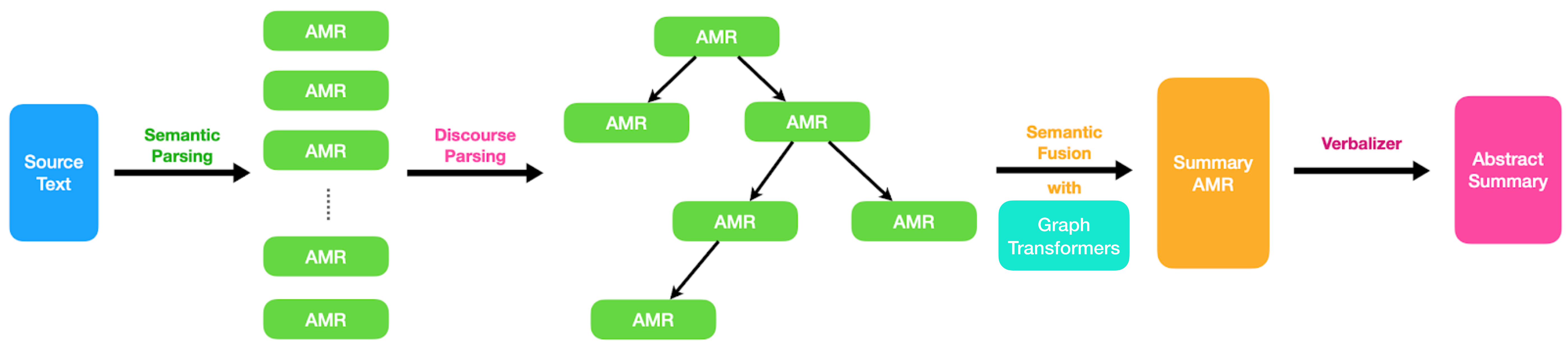

Figure 1. The System Pipeline.

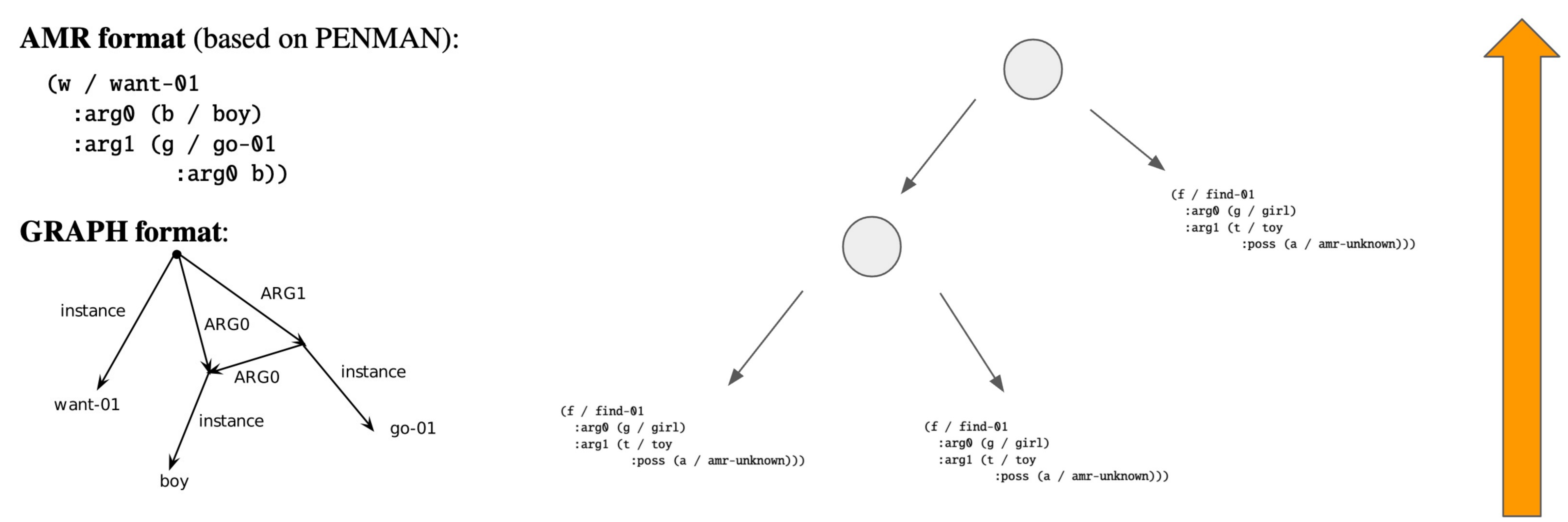

Figure 2. Left: an example AMR graph and its corresponding PENMAN notation; Right: recursive graph transformers applied over an example discourse parsing tree composed of AMR PENMAN scripts.

The Extractive Module

We observe that in many real-world abstractive summarization dataset, a large portion of the sentences in the source text are completely ignored in the human-annotated summary. Therefore, We first need to train a pre-selective extractive model to select K = 6 top sentences from each source text. Only these K = 6 top sentences will be fed into the semantic fusion module to generate the final summary.

To this end, we first process our training dataset to obtain its abridged version --- for each source text in the training set, we rank the sentences according to their ROUGE overlaps with the annotated summary, and only the top K with the highest ROUGE overlap will be kept as a representative for the full source text. We train the extractive model over this abridged version of the training dataset, which can then locate the set of top K sentences during testing.

Recursive Graph Transformers for Structural Semantic Fusion

The core component of our structural semantic fusion module is a recursive graph transformer model. This model recursively encodes pairs of PENMAN AMR-sequences and decode out a PENMAN AMR-sequence for intermediate summarization results along the discourse parsing tree, until we obtain the PENMAN notation for the final summary AMR graph at the top of the discourse parsing tree. Our recursive graph transformer model will recursively fuse the semantics of all the candidate sentences together in a structural manner.

AMR-to-Text Verbalizer

In the final step we train an AMR-to-text verbalizer to turn our final summary AMR into smooth and grammatically sound natural language summaries. For this part we can either take a good-performing off-the-shelve verbalizer or train our own verbalizer using pretrained models.

Experiment

In our experiment we are training our proposed structural semantic fusion model on the CNN/Dailymail dataset and the Xsum dataset. We also plan to test our model on our newly collected online dialogue summarization dataset as well as the Multi-News dataset.

Acknowledgement

We would like to thank Professor Dragomir Radev for his great guidance and support of our project.