STRUDEL: Structured Dialogue Summarization for Dialogue Comprehension

Borui Wang, Chengcheng Feng, Arjun Nair, Madelyn Mao, Jai Desai, Asli Celikyilmaz, Haoran Li, Yashar Mehdad, Dragomir Radev

Yale University, Meta AI

Introduction

Abstractive dialogue summarization has long been viewed as an important standalone task in natural language processing, but no previous work has explored the possibility of whether abstractive dialogue summarization can also be used as a means to boost an NLP system's performance on other important dialogue comprehension tasks. In this paper, we propose a novel type of dialogue summarization task - STRUctured DiaLoguE Summarization (STRUDEL) - that can help pre-trained language models to better understand dialogues and improve their performance on important dialogue comprehension tasks. In contrast to the holistic approach taken by the traditional free-form abstractive summarization task for dialogues, STRUDEL aims to decompose and imitate the hierarchical, systematic and structured mental process that we human beings usually go through when understanding and analyzing dialogues, and thus has the advantage of being more focused, specific and instructive for dialogue comprehension models to learn from. We further introduce a new STRUDEL dialogue comprehension modeling framework that integrates STRUDEL into a dialogue reasoning module over transformer encoder language models to improve their dialogue comprehension ability. In our empirical experiments on two important downstream dialogue comprehension tasks - dialogue question answering and dialogue response prediction - we demonstrate that our STRUDEL dialogue comprehension models can significantly improve the dialogue comprehension performance of transformer encoder language models.

Definition

We define Structured Dialogue Summarization (STRUDEL) as the task of generating a systematic and abstractive multi-entry dialogue summarization organized in a structured form that represents a comprehensive multi-aspect understanding and interpretation of a dialogue's content. It has the following key components:

- **Relationship**: the relationship between the two speakers of the dialogue.
- **Purpose/Theme**: the main purpose or theme for which the dialogue is made between the two speakers.
- **Task/Intention_1**: the main task or intention that the first speaker would like to achieve in the dialogue.
- **Task/Intention_2**: the main task or intention that the first speaker would like to achieve in the dialogue.
- **Problem/Disagreement**: the most important problem or disagreement that the two speakers need to solve in the dialogue.
- **Solution**: the solution that the two speakers reach for the most important problem or disagreement in the dialogue.
- **Conclusion/Agreement**: the final conclusion or agreement that the two speakers reach in the dialogue.

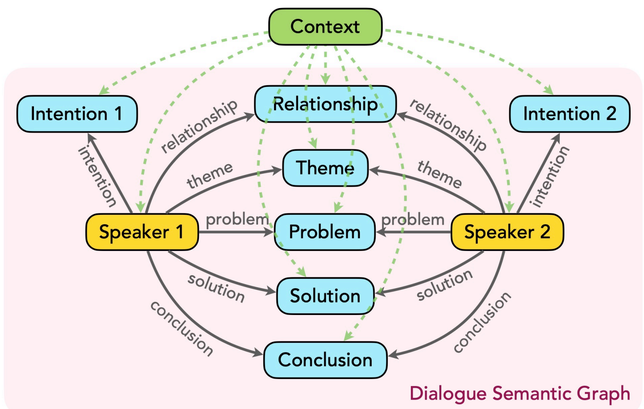

Figure 1. Dialogue Semantic Graph

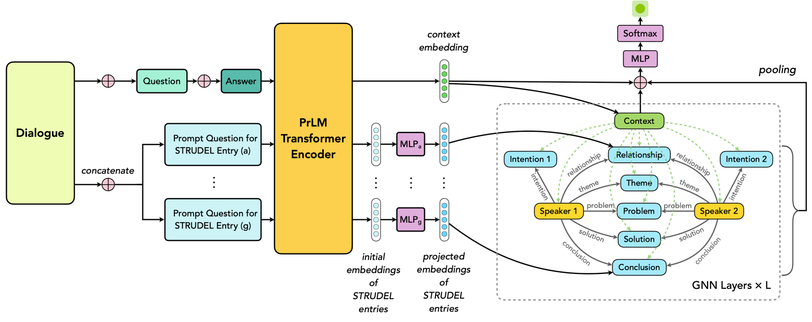

Figure 2. Model Architecture for STRUDEL Dialogue Comprehension

Model	MuTual			DREAM
	$R_4@1$	$R_4@2$	MRR	Accuracy
RoBERTa$_{large}$ (Liu et al., 2019)	0.695	0.878	0.824	0.821
RoBERTa$_{large}$ + STRUDEL	**0.869**	**0.947**	**0.919**	**0.838**
ALBERT$_{large}$ (Lan et al., 2020)	0.656	0.853	0.796	0.568
ALBERT$_{large}$ + STRUDEL	**0.673**	**0.872**	**0.812**	**0.596**

Table 1. Experiment Results on MuTual and DREAM

Model Training

During the training of our STRUDEL dialogue comprehension model, we first adopt a multi-task learning strategy to train the transformer model to learn to generate accurate STRUDEL embeddings and to infer the correct choices for dialogue question answering and response prediction tasks based on its generated STRUDEL embeddings at the same time.

Semantic Matching with Human Annotation:

$$\mathbb{L}_{\text{SM}} = -\sum_{\mathcal{E}\in\mathcal{S}} \text{Cos}\left(\text{MLP}^{\mathcal{E}}(h^{\mathcal{E}}_{\texttt{[CLS]}}),\ \tilde{h}^{\mathcal{E}}_{\texttt{[CLS]}}\right)$$

Using STRUDEL Reasoning to Infer Dialogue Comprehension Tasks:

$$\mathbb{L}_{\text{CE}} = -\log\left(\mathbb{P}^{\text{SDS}}(A = a^* \mid D, Q)\right)$$

Multi-Task Post-Training:

$$\mathbb{L} = \frac{1}{N}\sum_{i=1}^{N}\left(\alpha_1\mathbb{L}^i_{\text{SM}} + \alpha_2\mathbb{L}^i_{\text{CE}}\right)$$

Experiment Results

As we can see from the table, the accuracy results of our STRUDEL dialogue comprehension models on both the dialogue response prediction task (over the MuTual dataset) and the dialogue question answering task (over the DREAM dataset) are all consistently higher than their corresponding backbone transformer encoder models alone

Conclusion

In this paper, we presented STRUDEL (STRUctured DiaLoguE Summarization) - a novel type of dialogue summarization task that can help pre- trained language models to better understand dialogues and improve their performance on important dialogue comprehension tasks. In contrast to the traditional free-form abstractive summarization task for dialogues, STRUDEL provides a more comprehensive digest over multiple important aspects of a dialogue and has the advantage of being more focused, specific and instructive for dialogue comprehension models to learn from. In addition, we also introduced a new STRUDEL dialogue comprehension modeling framework that integrates STRUDEL into a dialogue reasoning module over transformer encoder language models to improve their dialogue comprehension ability.

Acknowledgement

We would like to thank Meta AI for their generous support of our work.