A View-Clustering Approach to Abstractive Dialogue Summarization

Borui Wang, Alexander R. Fabbri, Faiaz Rahman and Dragomir Radev

Department of Computer Science, Yale University, New Haven, CT

LILY Lab

Introduction and Motivation

In real-world dialogue summarization tasks, many online dialogues involve multiple speakers and usually contain multiple major viewpoints that are quite different from each other. For example:

(1) Online community questions answering forums: Quora, Zhihu, Stackoverflow, Reddit;
(2) Social Networks and Online Microblogs: Facebook, Twitter, Weibo;
(3) Online Comments: The New York Times, CNN.

So instead of following the traditional corpus-coherent approach to dialogue summarization as used in previous work, we propose a new view-clustering approach to dialogue summarization that tries to discern the rich underlying view-structures hidden in the dialogue corpus.

Our Approach

In order to better detect and utilize the rich view-structures hidden in the dialogue corpus during our abstractive dialogue summarization task, we propose a new view-clustering approach to abstractive dialogue summarization. Our new summarization approach has three stages:

(1) Cluster utterances into major view clusters
(2) Summarize the utterances in each view cluster in a coherent manner
(3) Detect the general logical relationship among different view, and then generate the summarization, possibly by logically integrating individual view summarizations through logic templates
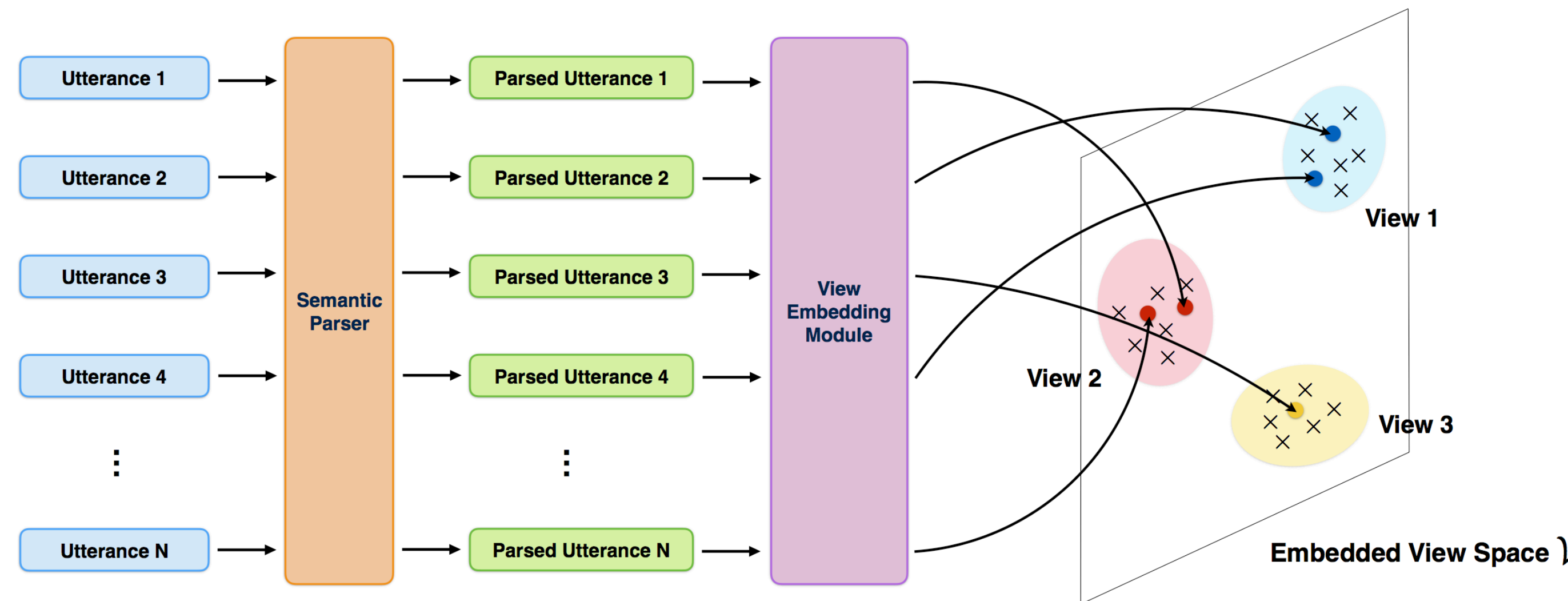

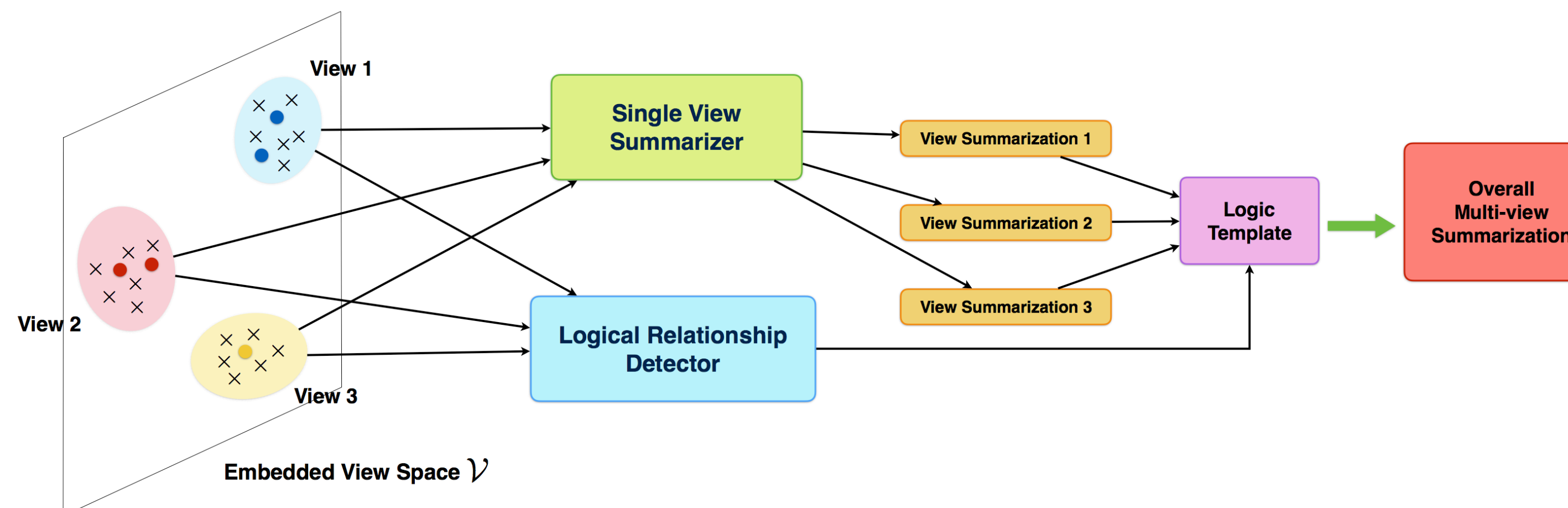Figure 1. Model Architecture for View-Clustering of different utterances in a dialogue.

Figure 2. Model Architecture for Summarizing Multiple Clustered Views with Detected Logic Template.

The Learning Model

Our proposed new learning model for multi-view abstractive dialogue summarization has four key components:

(1) **Summarization View Parser**: to parse out the major viewpoints contained in human-annotated dialogue summarizations in the training dataset;
(2) **Utterance-View Matching Mechanism**: to match each individual utterance in the dialogue with a specific parsed viewpoint that is semantically closest to it;
(3) **Single-View Summarizer**: to summarize coherent collections of utterances that are matched into same viewpoints.
(4) **View-Embedding Model for Semantically Parsed Utterances**: a mapping module to embed the semantically parsed structural form of utterances into vector embeddings in an embedded view space V. During the training of this mapping module, the training objective pushes that utterances matched into the same view cluster to be mapped into points that are closer to each other in the embedded view space V and that utterances matched into the different view clusters to be mapped into points that are further away from each other in V.

Experiment

In our experiment we will train our proposed model on our newly collected online dialogue summarization dataset that includes a large collection of human-annotated online dialogues from the New York Times, Reddit, Stackoverflow etc.

Acknowledgement