Computer Science
Defense

Enhancing Document Understanding through the Incorporation of Structural Inference

Liyan Xu
Emory University

Abstract: Towards resolving a variety of Natural Language Processing (NLP) tasks, pretrained language models (PLMs) have been incredibly successful by simply modeling language sequences, backed by their powerful sequence encoding capabilities. However, for document understanding tasks involving multi-sentence or multi-paragraph inputs, the model still needs to overcome the inherent challenge of processing scattered information across the entire document context, such as resolving pronouns or recognizing relations among multiple sentences.
¡br¿¡br¿ To address the motivation of effectively understanding document context beyond sequence modeling, this dissertation presents an in-depth study on the incorporation of structural inference, utilizing intrinsic structures of languages and documents. Four research works are outlined within this dissertation. Particularly, the first work proposes to integrate syntactic dependency structures into the document encoding process, capturing inter-sentence dependencies through designed graph encoding for the task of machine reading comprehension, especially under the multilingual setting. The second work investigates different methods to perform inference on the discourse structure that concerns coreference relations, allowing for higher-order decision making. The third work presents a novel formulation of structural inference to facilitate joint information extraction, fusing multi-facet information of document entities in terms of both coreference and relations. The last work explores the potential of the sequence-to-sequence generation as an approach that performs implicit inference on linearized entity structures, motivated by its unified encoder-decoder architecture and inherent abilities to perform higher-order inference.
¡br¿¡br¿ Overall, this dissertation demonstrates that incorporating designed structural inference upon certain intrinsic structures of languages or documents can effectively enhance document understanding, and highlights that modeling dependencies among different parts of the context can lead to more accurate and robust encoding and decoding process, where auxiliary information can be provided that complements the sequence modeling of PLMs. ¡br¿¡br¿ Zoom Option: https://emory.zoom.us/j/95753738482

Tuesday, March 28, 2023, 1:45 pm
Carlos Hall Room 211

Computer Science
Emory University