Houston Holman
4/7/23

ACL Paper Summary

Title
Think Before You Speak: Explicitly Generating Implicit Commonsense Knowledge for Response Generation

Authors
Pei Zhou, Karthik Gopalakrishnan, Behnam Hedayatnia, Seokhwan Kim, Jay Pujara, Xiang Ren, Yang Liu, Dilek Hakkani-Tur

Affiliation
Department of Computer Science, University of Southern California

The Problem
Effective human communication relies on more than just what was previously said. Implicit, common knowledge is the foundation for many conversations. For example, if a person were to say “I need to buy some flowers for my wife”, then a likely response may be “Perhaps you’d be interested in red roses.” This is based on the implicit knowledge that a rose is both a type of flower and a symbol of love. Modern response generation models are trained to take the conversation history and produce a response. Because the history lacks implicit knowledge, the models never learn this knowledge, sometimes causing uninformative responses.

Prior Work
One solution that addresses this issue is a knowledge-grounded response generator (Ghazvininejad et al., 2018; Dinan et al., 2019; Gopalakrishnan et al., 2019). This is done through the use of a knowledge base from which knowledge is retrieved during response generation. The problem with this solution is that the model's knowledge is limited to the coverage of the knowledge bases.

Another solution uses knowledge as a hidden factor during generation (Tuan et al., 2020; Xu et al., 2021). The problem with this is that it is nearly impossible to judge the quality of knowledge generation and how it is being used by the response generator.

Paper Contributions

This paper proposes a new response generation framework, dubbed Think-Before-Speaking (TBS). With TBS, the response generation model explicitly generates the implicit knowledge that is used to generate a response. This comes with three benefits: 1) compared to tradition models, responses are augmented and/or constrained by the generated knowledge, resulting in more informative responses; 2) compared to knowledge-retrieval models, generating knowledge means that the responses can include knowledge that is not in the knowledge bases; 3) explicitly generating knowledge provides an explanation of the intent of each response.

TBS relies on ConceptNet, a free semantic network, to help understand the meaning of words while creating its implicit knowledge. Knowledge in ConceptNet is stored in (subject s, relation r, object o) triples, such as (rose, TypeOf, flower). This is converted to natural language using a question-answer template. For example, the previous triple would be converted into *"What is a type of flower? Rose is a type of flower."* In order to have the response generation model make use of this information, the model talks to itself to make itself aware of the background information when generating a response.

The following is a sample input:
"I need to buy some flowers for my wife."

The model then pulls from ConceptNet and tells itself:
"The following background knowledge is helpful for generating the response: rose is a type of flower. Grounded on the background knowledge, what does the speaker probably say in the next response?"

Using this knowledge and self-talk, the model responds:
"Perhaps you'd be interested in red roses."

In order to evaluate their work, the authors used human evaluation to rate each of the models on a variety of metrics. When comparing the results of the TBS model with DiabloGPT, a commonly used traditional response generation model, the models were about equal in grammar, coherence, and engagingness. However, TBS outperformed DiabloGPT when it came to informativeness, specificity, and common sense. Comparing TBS to other knowledge-augmented response generation models, TBS significantly improves specificity and common sense aspects while remaining about the same in the other metrics. MTurkers were used to check the quality of the knowledge generated by TBS. Results from 300 sampled test instances show that 85% of the knowledge generated makes sense and is relevant. When analyzing the knowledge and responses, it was found that 77% of the generated knowledge was used in responses, showing that most responses are knowledge grounded. To check how much knowledge influenced the response, the researchers introduced "noisy" knowledge from unrelated queries into their tests. This causes a statistically significant drop in response quality, showing that the knowledge plays a role in the quality of responses.

Author with Most Citations
Dilek Hakkani-Tur - 18250 Citations

Conclusion
I think that this work is important because it highlights an overlooked step in the field of response generation. So many chatbots are obsessed with the idea of mimicking the feel of human conversation. The problem with many of these bots is that they often do not have much to say, since they are only outputting the next most likely word. The added step of understanding the semantics of the input and creating knowledge based on that adds so much depth to the generation. I would much rather talk to TBS than a generic chatbot the same way I would rather have a deep conversation with someone rather than just small talk.