

Alekhya Pinnamaneni

CS 4395.001

### Portfolio 8: Reading ACL Papers

The ACL paper I chose to summarize for this assignment is titled “Multi-Party Empathetic Dialogue Generation: A New Task for Dialog Systems”. The authors of this paper are Ling.Yu Zhu, Zhengkun Zhang, Jun Wang, Hongbin Wang, Haiying Wu, and Zhenglu Yang.

Empathetic dialogue and conversations have been studied in the NLP field for years, but these solutions focus on 2-party dialogue generation. However, empathetic conversations often involve multiple parties in settings such as conferences, group chats, and seminars. The paper addresses this issue by presenting a solution for multi-party empathetic discussions. Another issue this paper addresses is the gap between emotional and sensible dialogue. Previous empathetic dialogue generation solutions focus only on the emotional side, failing to address the complexity of human emotions with empathy as well as sensibility. The solution presented in this paper addresses that issue as well.

In 2019, Rashkin created the first empathetic dialogue dataset called “EMPATHETICDIALOGUES” (ED), which contains 32 emotions collected from 25,000 dialogues. Another dataset called PEC also contains empathetic dialogue data, but the data comes from just 2 Reddit forums, which is very limited and not generalizable. Other data sets include BlendedSkillTalk, ConvAI2, and PersonaChat. However, none of these data sets have more than 2 parties participating in a conversation. Although there have been multiple studies on multi-party conversations in the past, none of these studies focused on text.

Instead, they focused on speech. A recent study focused on the Address and Response System (ARS) and addressed the multi-party conversation issue, but failed to consider the role of emotions/empathy in the conversations.

There are multiple factors that make his paper unique compared to other empathic dialogue studies. First, it is one of the rare studies to incorporate more than 2 parties in an empathetic conversation. It also takes into account not just emotion but also sensibility when generating a response in the multi-party conversations. Another unique factor is the used of a static-dynamic model called SDMPED.

The authors used 2 evaluation criteria to evaluate their work: automatic evaluation criteria and human evaluation criteria. For the automatic evaluation criteria, the response generated by the model is compared against human response by calculating the AVG BLEU and ROUGE-L scores. For the human evaluation criteria, the authors collected 100 random dialogues generated by the system and asked human annotators to rate the dialogues on a scale of 1 to 5 based on 3 criteria: empathy, relevance, and fluency.

This paper has received 0 citations on Google Scholar as of currently, probably due to its recent release. I think the authors' work was important because it explores a rarely touched area of emotional/empathetic dialogue generation study and has the potential to facilitate and improve empathetic dialogue generation and has many use applications. On Google Scholar Jun Wang has 4,010 citations, the most citations out of all the authors of this paper.