（続紙　１　）

京都大学	博士（情報学）	氏名	傅　雅慧(Fu Yahui)
論文題目	Dialogue Comprehension and Personalization for Empathetic Response Generation （共感的な応答生成のための対話の理解と個性化）		

（論文内容の要旨）

Incorporating empathy into spoken dialogue systems is crucial for improving interaction with conversational robots and virtual agents, as empathy is the emotional bonding among humans; Conversational robots and virtual agents expressing empathy would give humans a feeling of being understood and satisfied with the conversation. This thesis addresses empathetic response generation for text-based dialogue systems from the perspective of appropriate dialogue comprehension and personalization.

Generally, empathy is embodied in the aspects of both contextual understanding and affective expression. However, previous studies often focus on either aspect. We first address this problem by generating appropriate empathetic responses with both aspects via modeling emotion and content consistency between the user's input and empathetic response. Moreover, it is necessary to comprehend the cause-and-effect relationships in response generation. The end-to-end generation model operates as a black box, making it unclear what factors lead to a particular response in a given context. To address this issue, we further explore causal reasoning to make the generated empathetic responses explainable. An appropriate empathetic response also depends on personality traits. Recognition of the user's personality and the development of systems that accordingly express a consistent personality are important for enhancing rapport and engagement in the interactions. To achieve this, we enhance the personality recognition in dialogue and then stylize the system to generate responses that are both empathetic and reflective of a distinct personality.

Chapter 2 provides an overview of dialogue systems, specifically emphasizing techniques for empathetic response generation (ERG).

In Chapter 3, a dual variational generative model (DVG) is proposed for empathetic response generation based on both contextual understanding and affective expression. Specifically, an emotion classifier and a variational model are incorporated into a dual response and context generative model to learn the emotion and content consistencies efficiently. DVG also uses the reconstruction loss used in VAE for both contexts and responses. Evaluations on both Japanese and English EmpatheticDialogues datasets demonstrate DVG's superiority in generating empathetic responses with contextual and emotional appropriateness. In addition to the DVG model, we propose an auxiliary retrieval system to improve empathetic response generation. Furthermore, the proposed model's ability is extended to general response generation, which is not specific to empathetic but also chitchatting dialogue systems. We evaluated our system's effectiveness in enhancing dialogue by a virtual agent. Subsequently, we integrated the system into a humanoid robot for practical application.

Chapter 4 describes the empathetic response generation based on causal reasoning. Recent approaches mainly focus on understanding the causalities of context from the user's perspective, ignoring the system's perspective. We propose a commonsense-based causality reasoning for diverse empathetic response generation that considers both the user's perspective (user's desires and reactions) and the system's perspective (system's intentions and reactions). Enhances ChatGPT's ability to reason for the system's perspective by integrating in-context learning with common-sense knowledge. Then, the commonsense-based causality explanation

is integrated into both ChatGPT and a T5-based model. Integration of T5 with ChatGPT's reasoning capability realizes more empathetic responses that result in better evaluations. ChatGPT with the causality explanation can generate more empathetic and accurate responses.

Chapter 5 addresses the personality recognition of the user in dialog, which is useful for enhancing the ability of conversational robots and virtual agents to tailor user-adaptive responses. To address the challenge of the limited number of speakers in existing dialogue corpora, we introduce personality trait interpolation for speaker data augmentation. Moreover, a heterogeneous conversational graph neural network (HC-GNN) is incorporated to independently capture the interdependencies among interlocutors and the intradependencies within the speaker. Experimental results on the RealPersonaChat corpus demonstrate that increasing speaker diversity by data augmentation significantly improves personality recognition in both monologue and dialogue settings. The proposed HC-GNN outperforms baseline models, showcasing its effectiveness in dialogue setting.

Chapter 6 focuses on stylizing the empathetic response generation considering the system's personality. Specifically, a multi-grained prefix mechanism is designed to capture the intricate relationship between a system's personality and its empathetic expressions. Furthermore, a personality reinforcement module is designed to leverage contrastive learning to calibrate the generation model, ensuring that responses are both empathetic and reflective of a distinct personality. Automatic and human evaluations show the effectiveness of the proposed method in generating responses with enhanced empathy and personality expression.

Chapter 7 summarizes the findings of this thesis and discusses future work on adapting the system's empathetic style and personality to the user's personality in dialogue.

（続紙 ２ ）

（論文審査の結果の要旨）

　対話システムは大規模言語モデル(LLM)により大きな進展を遂げているが、ユーザに対して共感的な応答を生成することが大きな課題となっている。共感的な応答を生成するには、ユーザの状況や感情さらには個性を理解した上で、システムの適切な意図を推論したり、発話スタイルを調整するのが望ましいが、LLMなどのブラックボックスなモデルでは困難である。本研究は、このような対話の理解と個性化に基づいて共感的な応答生成を行う方法を複数提案し、その実験的評価をまとめたもので、主な成果は以下の通りである。

1．エンコーダ・デコーダモデルにおいて、文脈の理解と応答の生成に一貫性・双対性があることに着目し、双対型で変分オートエンコーダ型のモデルを定式化し、さらに感情認識を導入することで、効果的に共感的応答を生成するシステムを学習する方法を提案した。本手法により、応答の多様性が大きく改善し、共感度や関連度も向上することを示した。実際に会話エージェントに実装し、傾聴感やエンゲージメントが改善されることも確認した。

2．対話の文脈からユーザの欲求や感情を明示的に常識推論するとともに、システムの適切な意図や感情も文脈内学習を用いて推論する機構をend-to-end型の対話モデルに導入することを提案した。本手法をChatGPTとT5の２つのモデルに実装し、より共感的で適切な応答が生成されることを示した。

3．共感的な応答は個性に依存することに着目し、テキスト発話から外向性などの個性を予測するモデルを効果的に学習する方法を提案するとともに、共感的意図に加えて、個性を明示的に中間表現に埋め込んで共感的な応答生成を行う方法を提案した。本手法により、共感度とともに個性の一貫性に関する評価が向上することを示した。

　以上のように本論文は、対話システムにおいて共感的な応答を生成する方法を複数提示しており、学術上・実用上寄与するところが少なくない。よって、本論文は博士（情報学）の学位論文として価値あるものと認める。また、令和6年8月28日に論文とそれに関連した内容に関する口頭試問を行った結果、合格と認めた。また、本論文のインターネットでの全文公開についても支障がないことを確認した。