LREC-COLING 2024

Beyond Linguistic Cues: Fine-grained Conversational Emotion Recognition via Belief-Desire Modelling

Bo Xu, Longjiao Li, Wei Luo, Mehdi Naseriparsa

Zhehuan Zhao, Hongfei Lin and Feng Xia

Dalian University of Technology, Dalian, China

Outline

1 Introduction

- Background
- Research status
- Our work

2 Method

- BDTE
- Framework

3 Experiments

- Dataset
- Performance comparison
- Ablation Study
- Effect of GT Layer and Channel Numbers
- Case Study

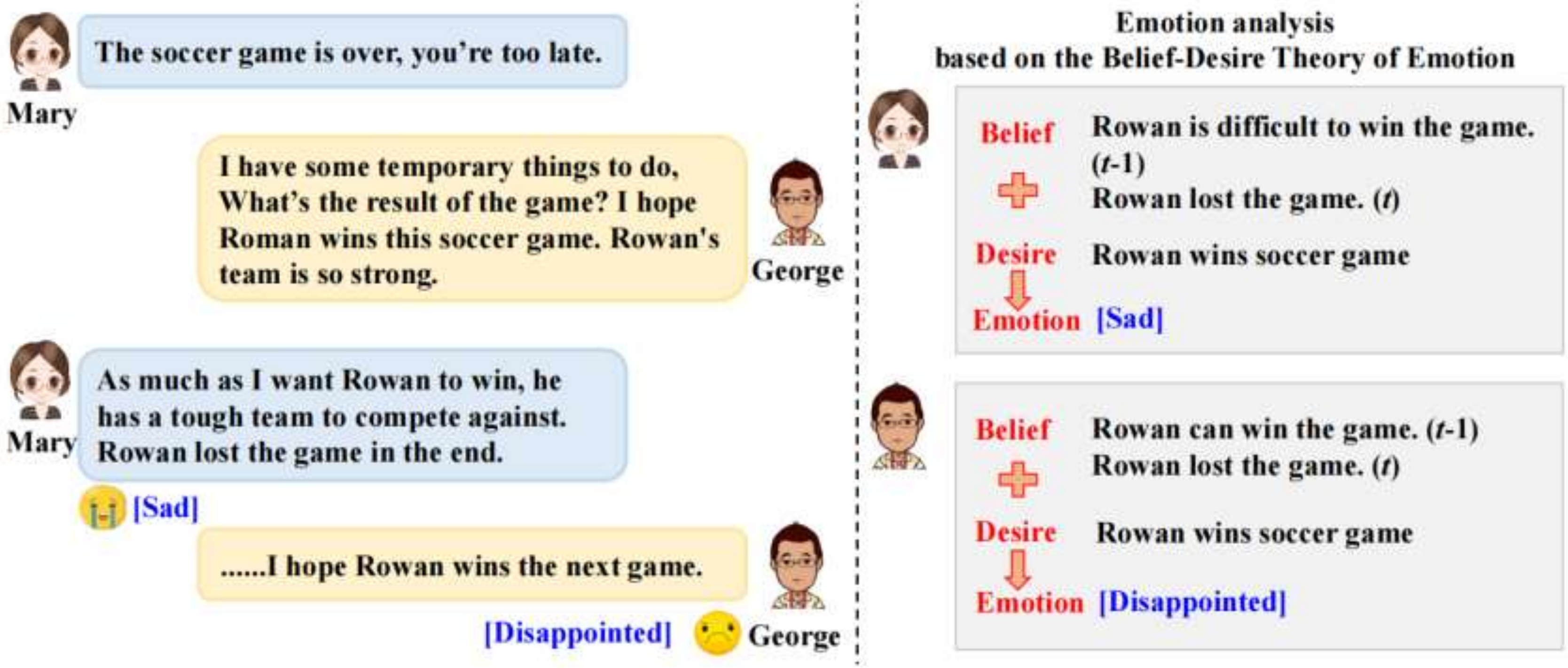
4 Conclusions

- Conclusions
- Future work

1 Introduction

Introduction

Background



- Emotion recognition in conversation (ERC) is essential for dialogue systems to identify the emotions expressed by speakers.
- The task of ERC requires not only the semantic information of the utterance itself but also the context modeling of each utterance.

- Accurate recognition and interpretation of similar fine-grained emotion properly accounting for individual variability remains a challenge.

Introduction

LREC-COLING 2024

Research status

Graph-based
Sequence-based
Commonsense knowledge-based

ignore

individual-level
cognitive processes
underlying emotions

To address this challenge, we turn to the Belief-Desire Theory of Emotion (BDTE) , which provides a psychological framework to understand emotion in conversation via complex interplay between individual beliefs and desires.

Introduction

Our work

individual-level cognitive processes underlying emotions

solve

Emotion-eliciting events + Heterogeneous conversation graph

- We introduce the Belief-Desire theory to the emotion recognition in conversation (ERC) task, for more nuanced and accurate modelling of individual emotion.
- We constructed an emotion-belief-desire conversation graph to support knowledge representation and inference based on key principles in BDTE. The graph captures the utterance context, speaker's global state, and supports the inference of beliefs and desires.
- We extensively evaluated our method on four commonly used ERC datasets, demonstrating its superiority and effectiveness.

2 Method

Method

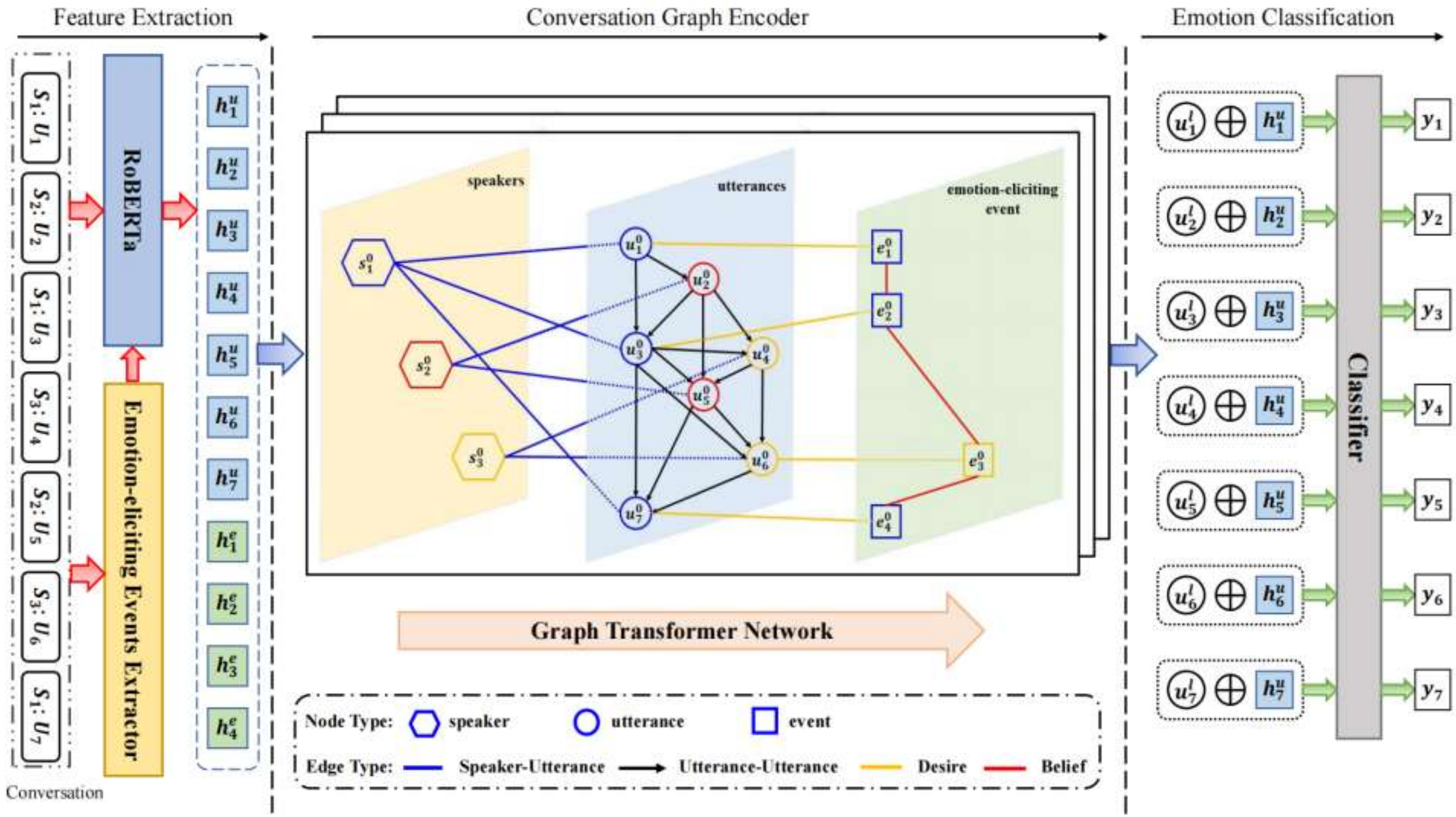
BDTE

Emotion	if	Belief at t	Desire at t	Belief at $t-1$
Happy/Joyful		Certain(e, t)	Des(e, t)	
Sad		Certain(e, t)	Des($\neg e$, t)	
Excited		Uncertain(e, t)	Des(e, t)	
Feared/Scared		Uncertain(e, t)	Des($\neg e$, t)	
Surprised		Certain(e, t)	–(irrelevant)	Bel($\neg e$, $t-1$)
Disappointed		Certain($\neg e$, t)	Des(e, t)	Bel(e, $t-1$)
Peaceful		Certain($\neg e$, t)	Des($\neg e$, t)	Bel(e, $t-1$)

The Belief-Desire Theory of Emotion emphasizes that emotion is not solely driven by desire/motivation, but also by how one perceives the situation.

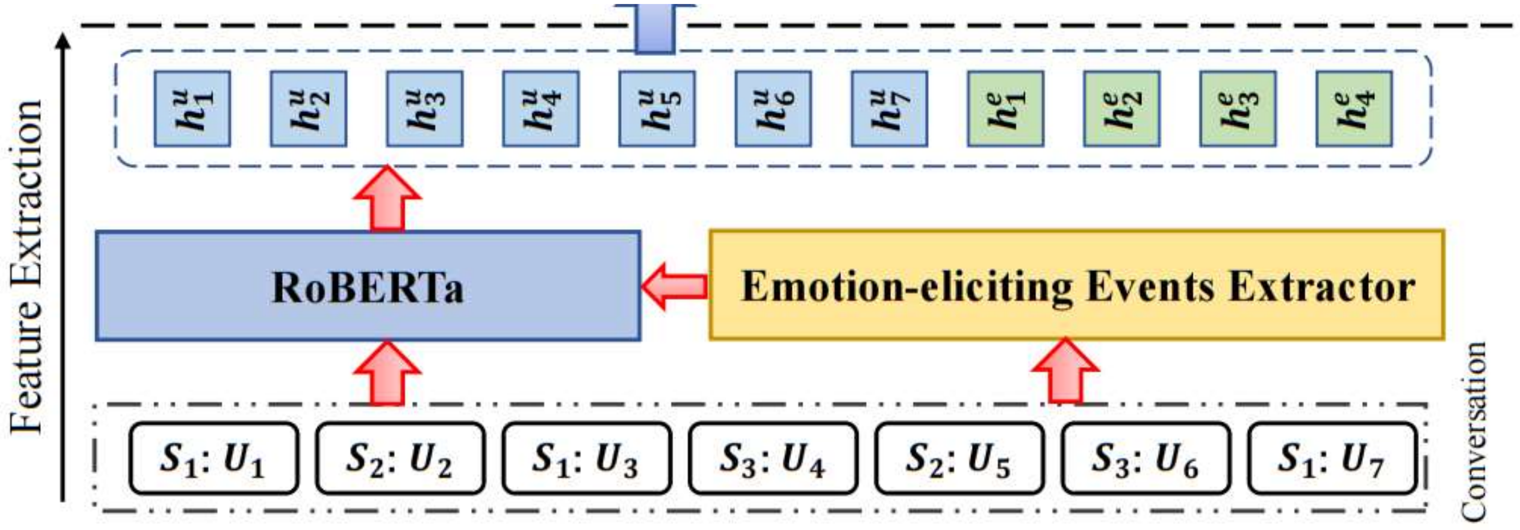
Method

Framework



Method

Feature Extraction



Utterance-level Feature Extraction

$$h_i^u = RoBerta([CLS], w_1, w_2, ..., w_{L_i}).$$

Emotion-eliciting Event Feature Extraction

$$E = Matching(Parsing(U_i)).$$

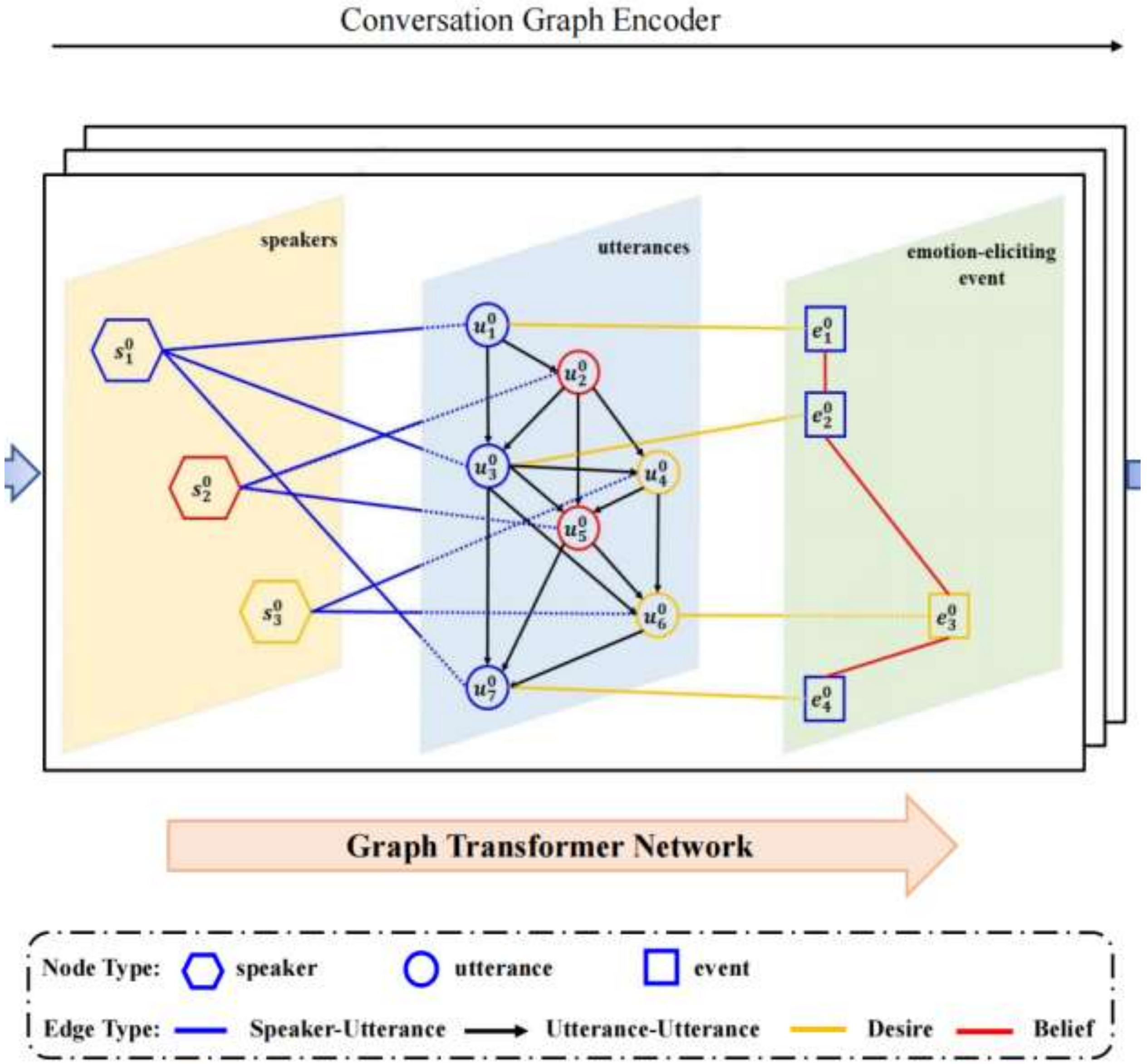
$$h_i^e = Maxpooling(RoBERTa(e_i)), e_i \in E.$$

Method

Emotion-belief-desire Conversation Graph Construction

- Utterance node
- Speaker node
- Emotion-eliciting event node

- Utterance-utterance edge
- Speaker-utterance edge
- Desire edge
- Belief edge



Method

LREC-COLING 2024

Graph Transformer Network

We learn the node representation of the graph by using the graph transformer network.

$$H = \mathop{\Big\|}_{i=1}^{C} \sigma(\widetilde{D}_i^{-1} \widetilde{A}_i^{(l)} X W).$$

$$u_i^l = H_{u_i}.$$

Emotion Classification

We connect the utterance-level embedding representation of an utterance with the final node-embedding representation of an utterance node and feed it to a feedforward neural network for emotion classification:

$$z_i = h_i^u \| u_i^l.$$

$$p_{x,i} = Softmax(W_z z_i + b_z).$$

$$y_{x,i} = Argmax(p_{x,i}).$$

3 Experiments

Experiments

LREC-COLING 2024

Dataset

	Dataset	IEMOCAP	DailyDialog	MELD	EmoryNLP
#Dia	Train	100	11118	1038	713
	Val	20	1000	114	99
	Test	31	1000	280	85
#Utt	Train	5236	87170	9989	9934
	Val	574	8069	1109	1344
	Test	1623	7740	2610	1328

Experiments

Performance comparison with different methods

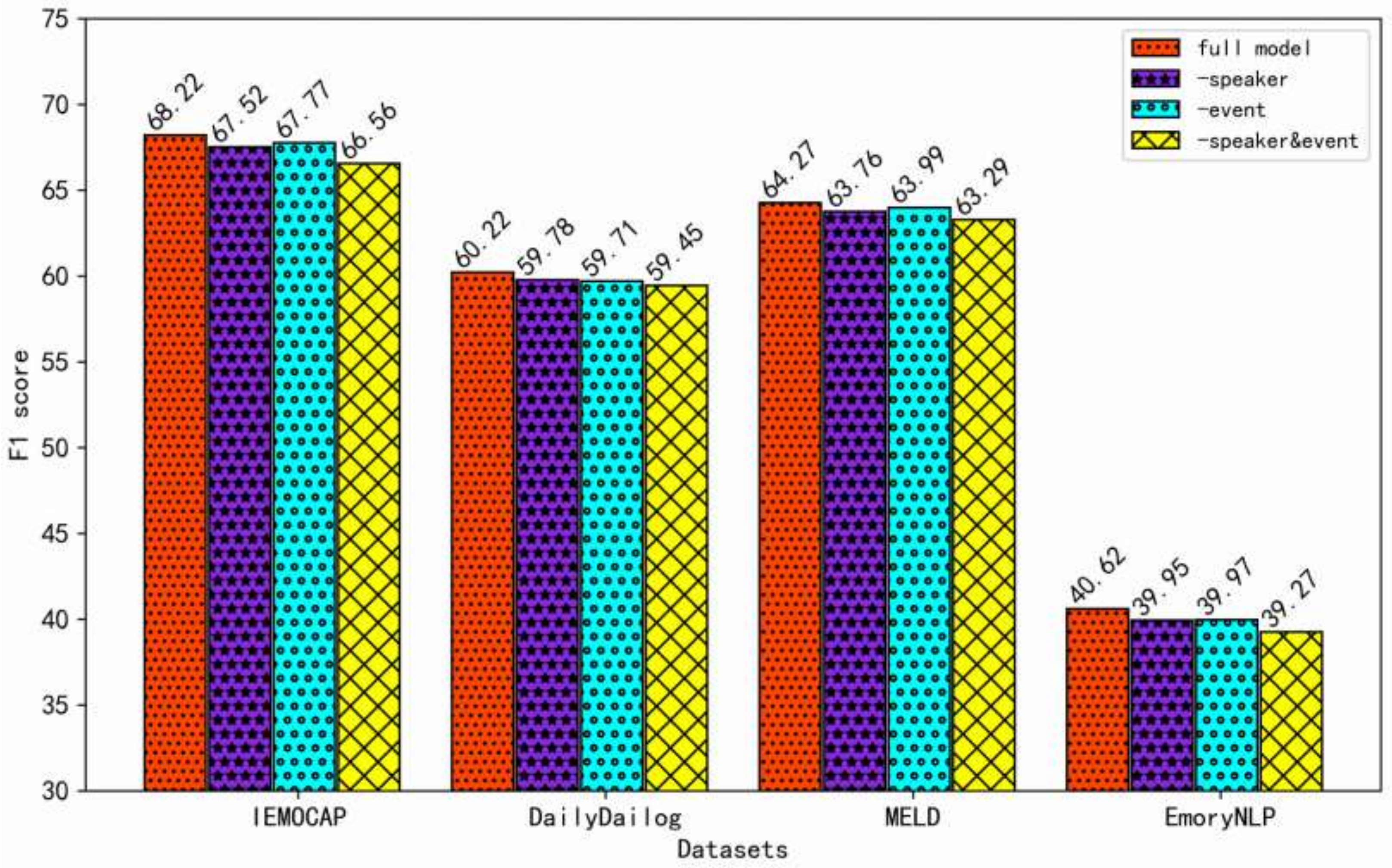
Model	IEMOCAP	DailyDialog	MELD	EmoryNLP
Metric	weighted-F1	micro-F1	weighted-F1	weighted-F1
CNN(Kim, 2014)	52.04	50.32	55.02	32.59
KET(Zhong et al., 2019)	59.56	53.37	58.18	34.39
DialogueRNN(Majumder et al., 2019)	62.57	57.03	57.03	31.70
DialogueGCN(Ghosal et al., 2019)	64.18	-	58.10	-
RGAT-POS(Ishiwatari et al., 2020)	65.22	54.31	60.91	34.42
DialogXL(Shen et al., 2021a)	65.94	54.93	-	34.73
DialogueCRN (Hu et al., 2021)	66.20	-	58.39	-
Cosmic (Ghosal et al., 2020)	65.28	58.48	65.21	38.11
SKAIG-ERC (Li et al., 2021)	66.96	59.75	65.18	38.88
CauAIN(Zhao et al., 2022)	67.61	58.21	65.46	-
CISPER (Yi et al., 2022)	-	-	66.10	39.86
MPLP (Zhang et al., 2023)	66.65	59.92	**66.51**	-
RoBERTa (Liu et al., 2019)	63.38	58.08	62.88	37.78
Ours	**68.22**	**60.22**	64.27	**40.62**

Experimental results

LREC-COLING 2024

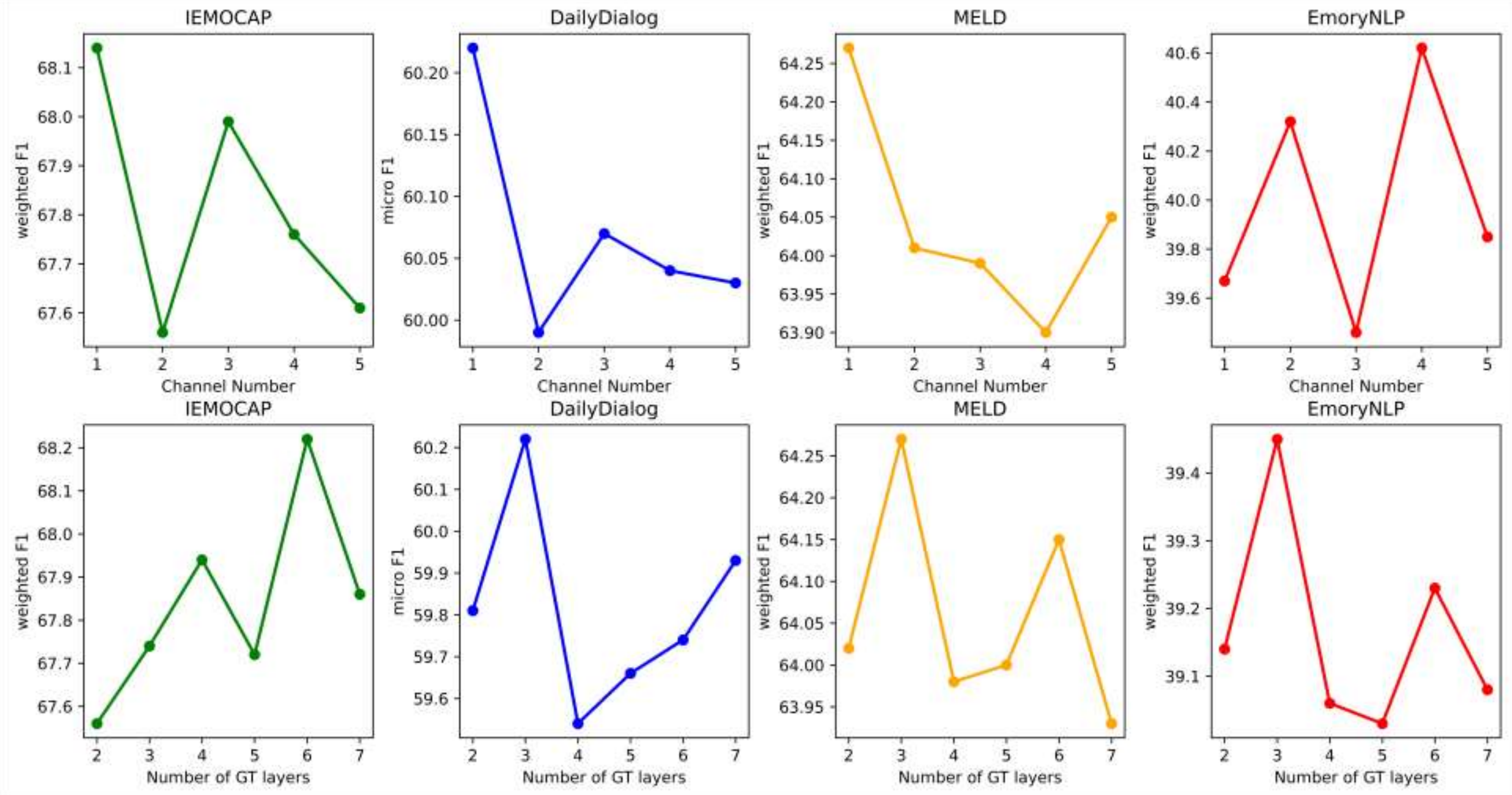
ELRA ICCL

Ablation Study



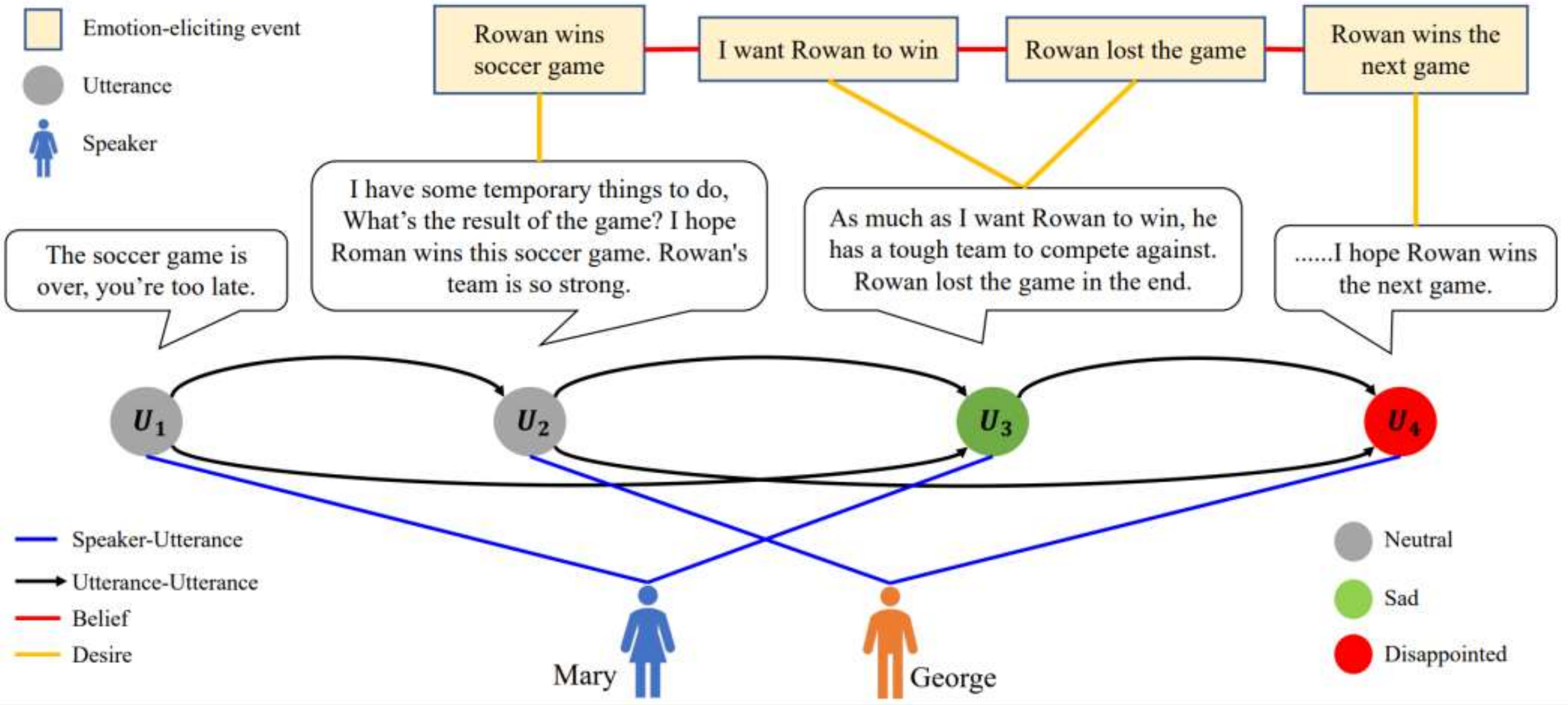
Experimental results

Effect of GT Layer and Channel Numbers



Experimental results

Case Study



4 Conclusions

Conclusions

Conclusions

- We proposed an approach based on the Belief-Desire Theory of Emotion for conversational emotion recognition.
- We construct an emotion-belief-desire heterogeneous conversation graph to model the context modeling, speakers' global state, and belief and desire inference between emotion-eliciting events and utterances.
- We conducted extensive experiments, the results verify that our proposed method for ERC based on belief and desire inference achieves superior performance than multiple state-of-the-art methods.

Future work

- How to better model the belief-desire representation of emotions is worth further research.

Thanks for your attention!