Reasoning before Responding: Integrating Commonsense-based Causality Explanation for Empathetic Response Generation

Yahui Fu, Koji Inoue, Chenhui Chu, Tatsuya Kawahara

京都大学

KYOTO UNIVERSITY

SIGDIAL 2023, Prague, Czechia

Research Background

What's empathy?

Empathy is a desirable capacity of humans to place themselves in another's position to ***show understanding*** of his/her ***experience and feelings***.

Why empathy?

An empathetic dialogue system can serve as chit-chat friends for ***companion***, psychologists for ***health care***, etc.

Figure 1: Empathetic dialogue system by mutli-modality avatar Gene[1]

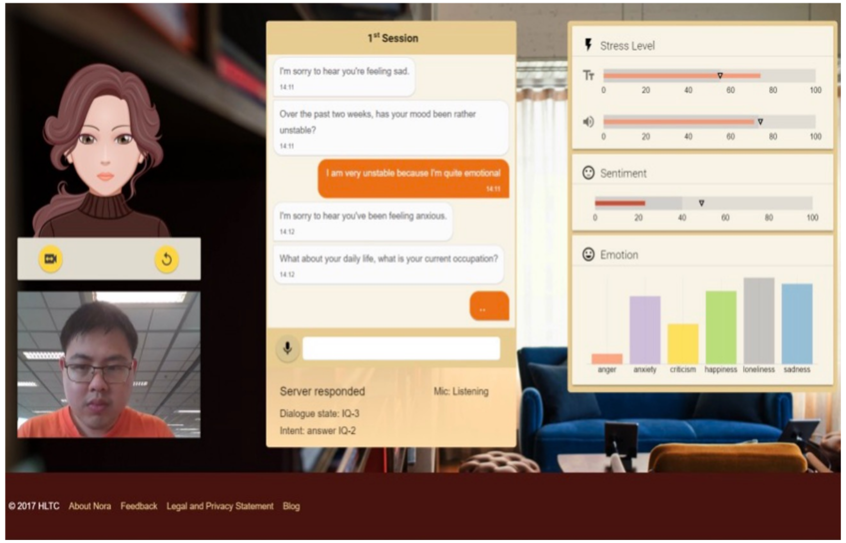


Figure 2: Nora, the empathetic psychologist[2]

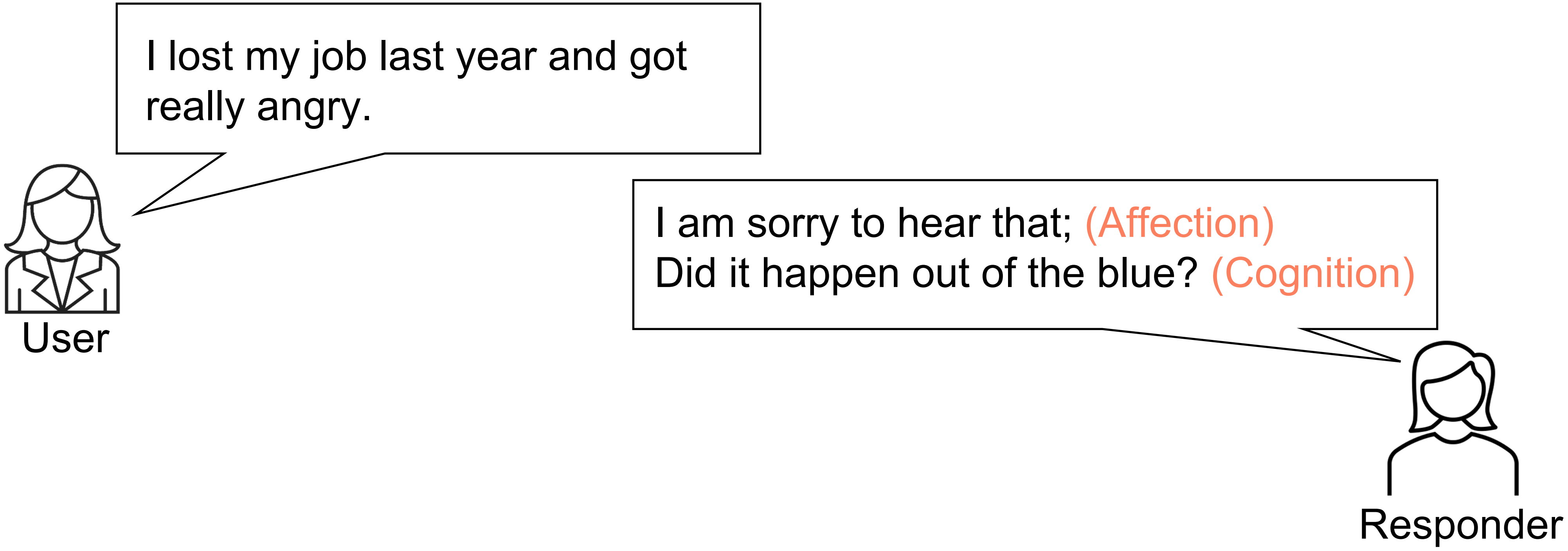
[1] Fu, Y et al, T. Improving Empathetic Response Generation with Retrieval based on Emotion Recognition. IWSDS 2023
[2] Winata, Genta Indra, et al. "Nora the Empathetic Psychologist." INTERSPEECH. 2017.

Research Background

How to express empathy?

Empathy includes two aspects: Cognition and Affection[3].

- Cognition: understand the other person’s perspective and situation.
- Affection: express suitable emotion

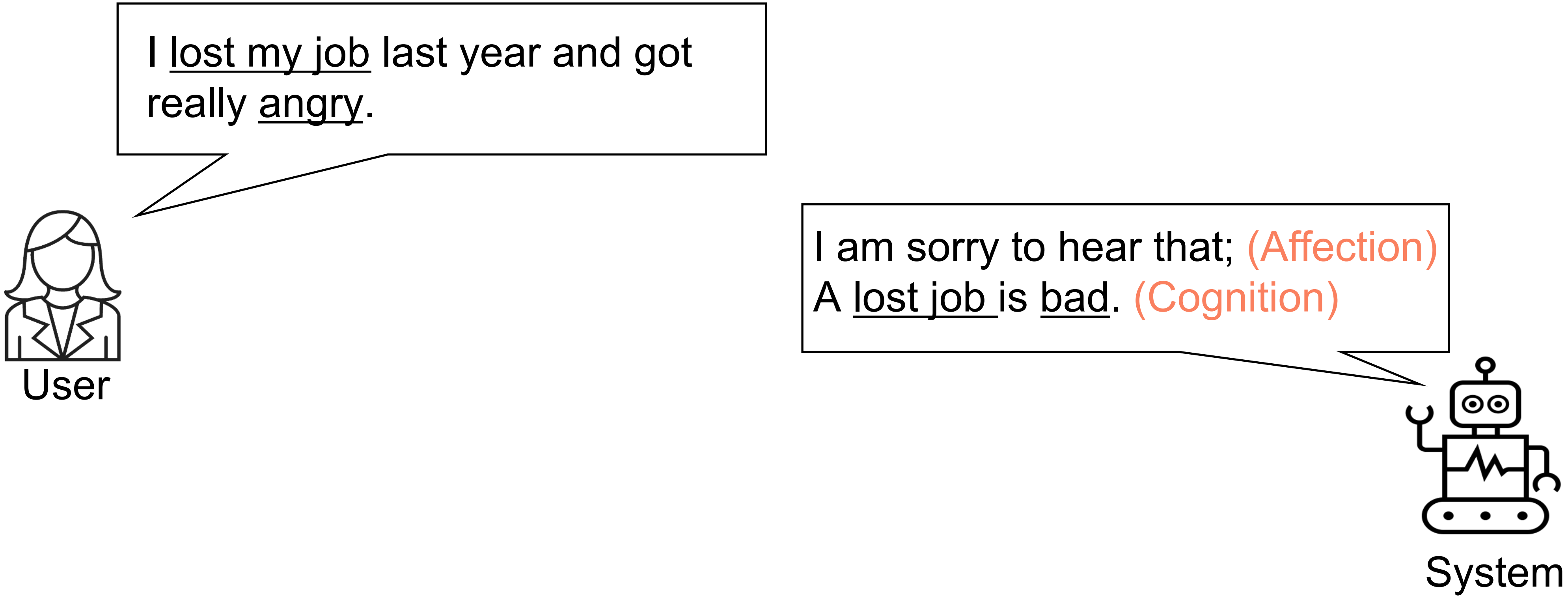


[3] Mark H Davis. 1983. Measuring individual differences in empathy: evidence for a multidimensional approach. Journal of personality and social psychology, 44(1):113.

Research Background

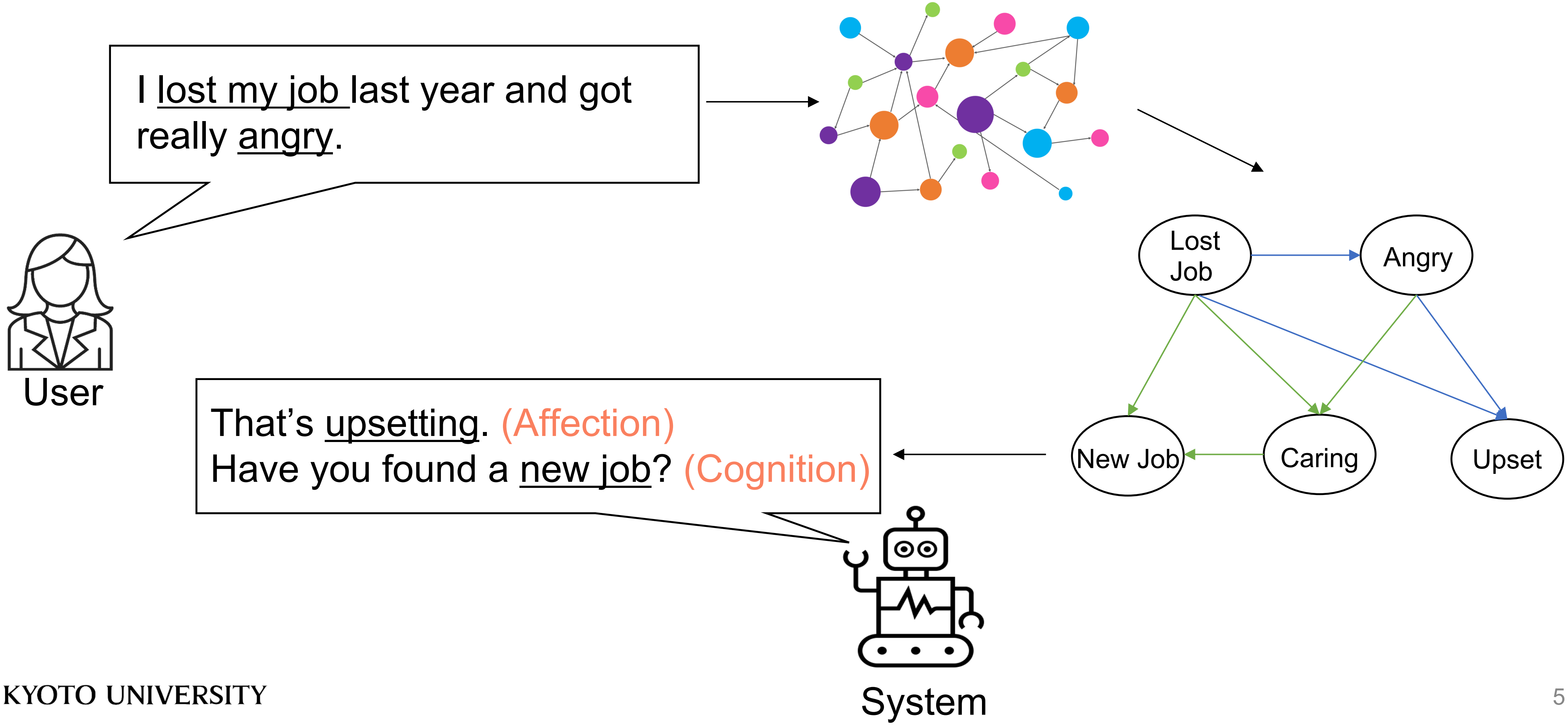
How to generate empathetic response?

A case with no causality explanation, generating an empathetic response based on context information.



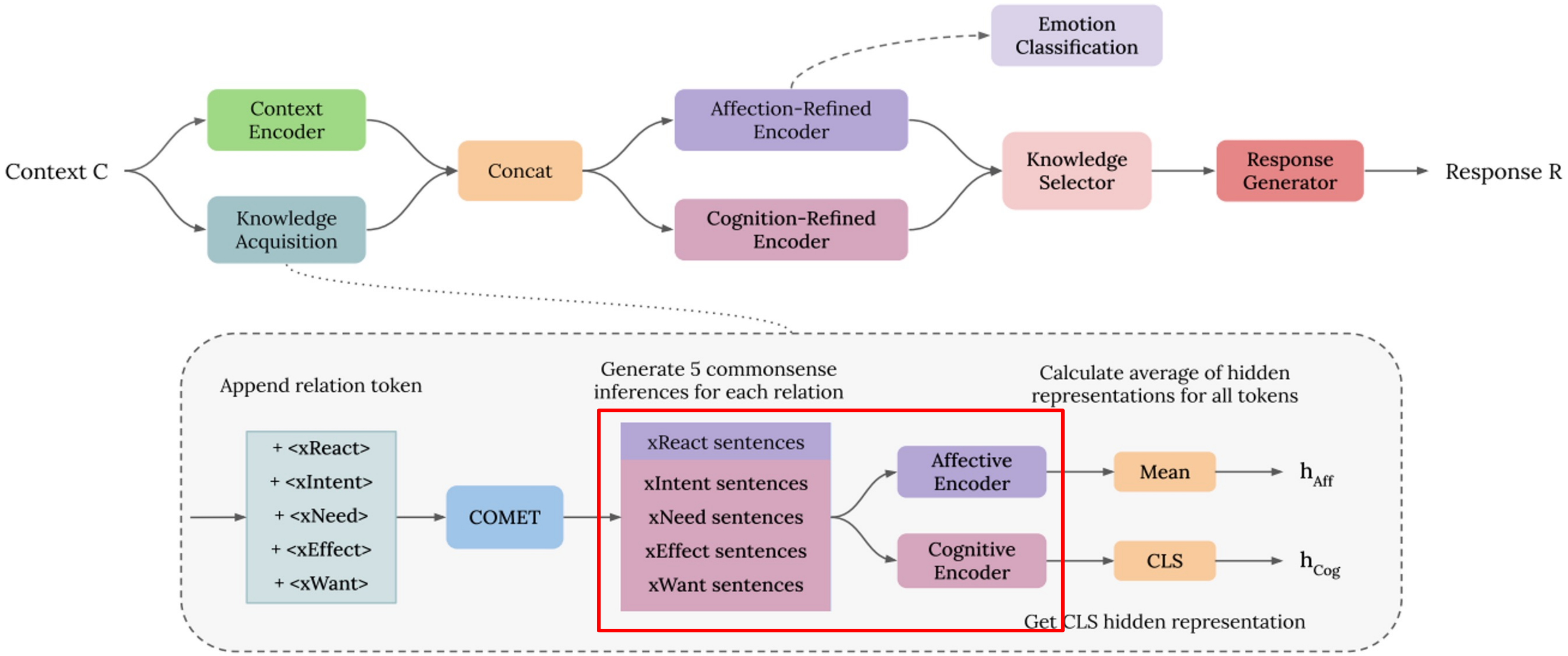
Research Background

A case with causality explanation, generating an empathetic response based on knowledge reasoning.



Related Work

[Sabour et al AAAI 2022] used a knowledge model COMET to obtain the user's ***react and situation*** for ***affective and cognitive*** encoding.

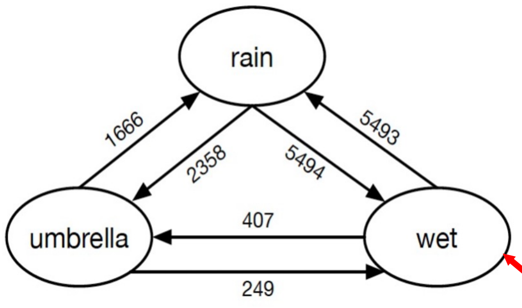


Weakness: concatenated related knowledges, no reasoning process.

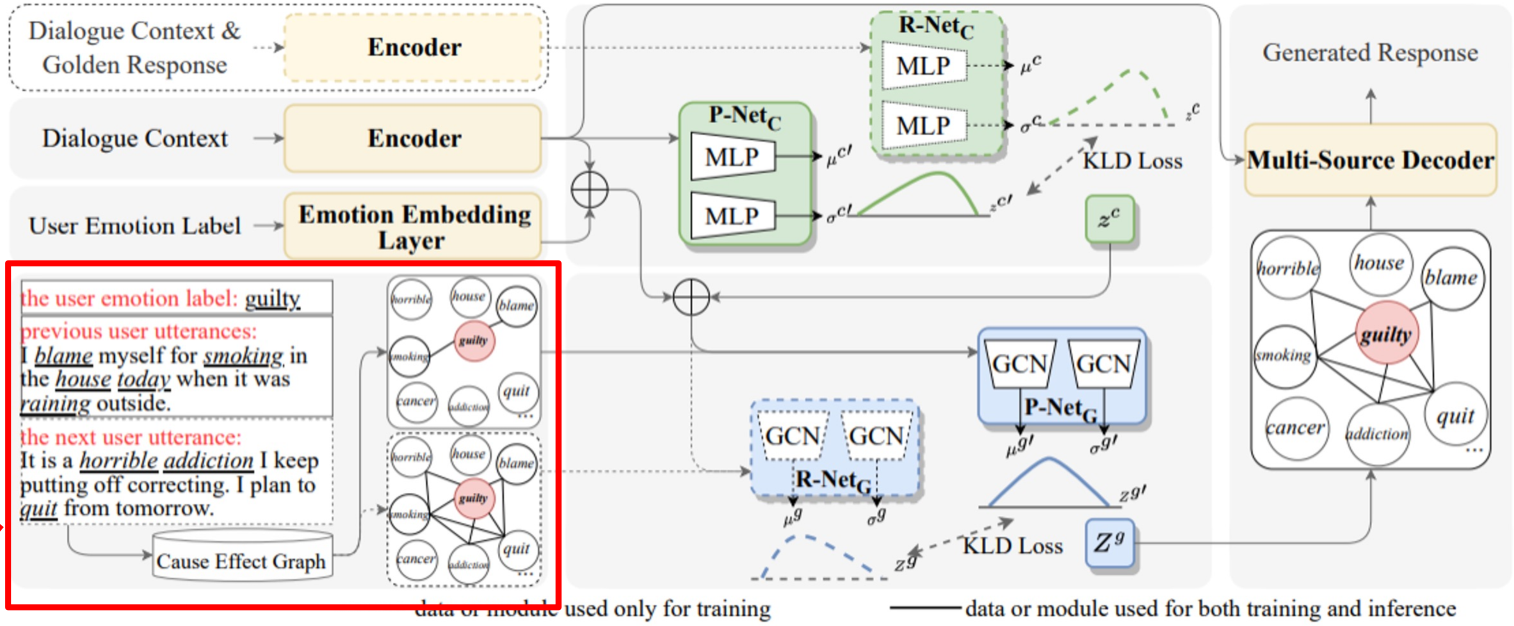
[4] Sabour, Sahand, Chujie Zheng, and Minlie Huang. "Cem: Commonsense-aware empathetic response generation." Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 36. No. 10. 2022.

Related Work

[Wang et al EMNLP 2022] used ***cause-effect graph*** to build the causality interdependence between user's emotion to user's context, and user's emotion to system's response.



Cause-effect graph



Weakness: It only reasoned casualties to the user's emotion, did not reason more fine-grained user's want and system's intent.

[5] Wang, Jiashuo, Yi Cheng, and Wenjie Li. "CARE: Causality Reasoning for Empathetic Responses by Conditional Graph Generation" EMNLP 2022 findings.

Motivation

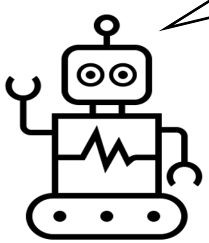
Exploring user’s perspective:
Affection: angry
Desire: to get a new job.

I lost my job last year and got really angry.

User

System’s intention is aligned with user’s desire:
Affection: sad
Intention: to give a new job.

I am sorry to hear that (Affection); I wish can give you a new job! (Cognition)



System

Motivation

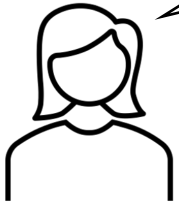
Reasoning user's perspective:
Affection: angry
Desire: want to complain.

I lost my job last year and got really angry.

User

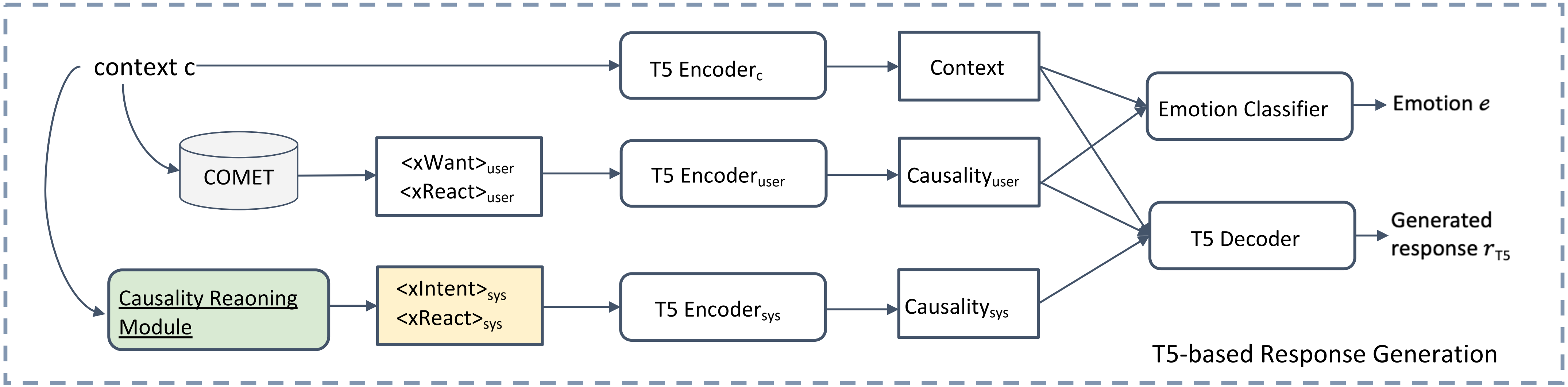
Reasoning responder's perspective to mimic humans:
Affection: sad
Intention: to know what happened.

I am sorry to hear that (Affection); Did it happen out of the blue? (Cognition)

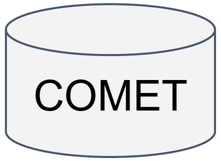


Responder

Proposed Method

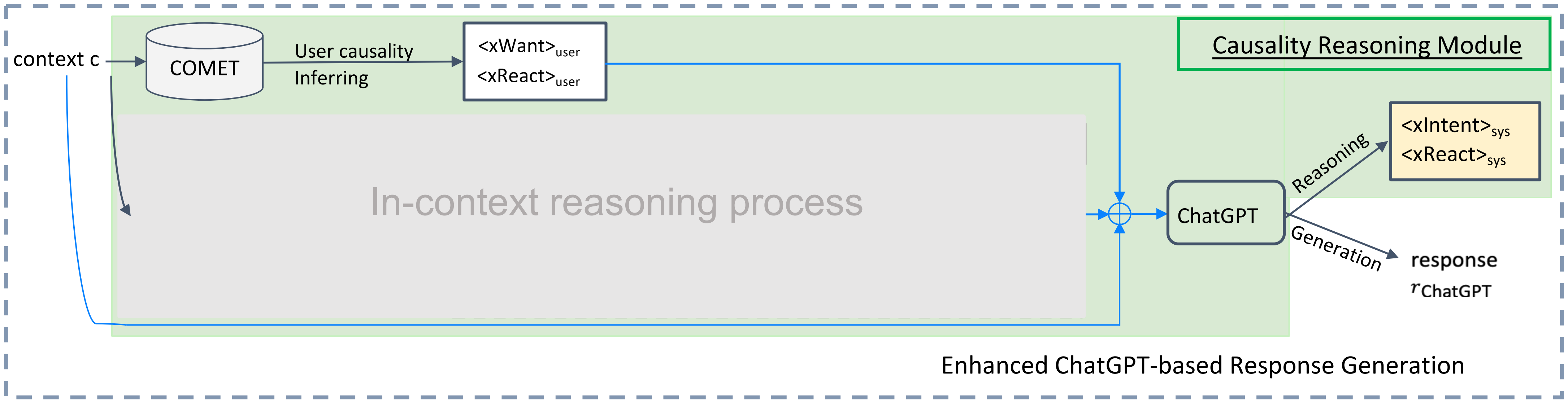


- For the context c, using COMET to predict user's want/react, and causality resoning module to predict system's intent/react.



- It is a BART-based model which is fine-tuned on the ***cause-effect*** graph from ATOMIC-2020 dataset.

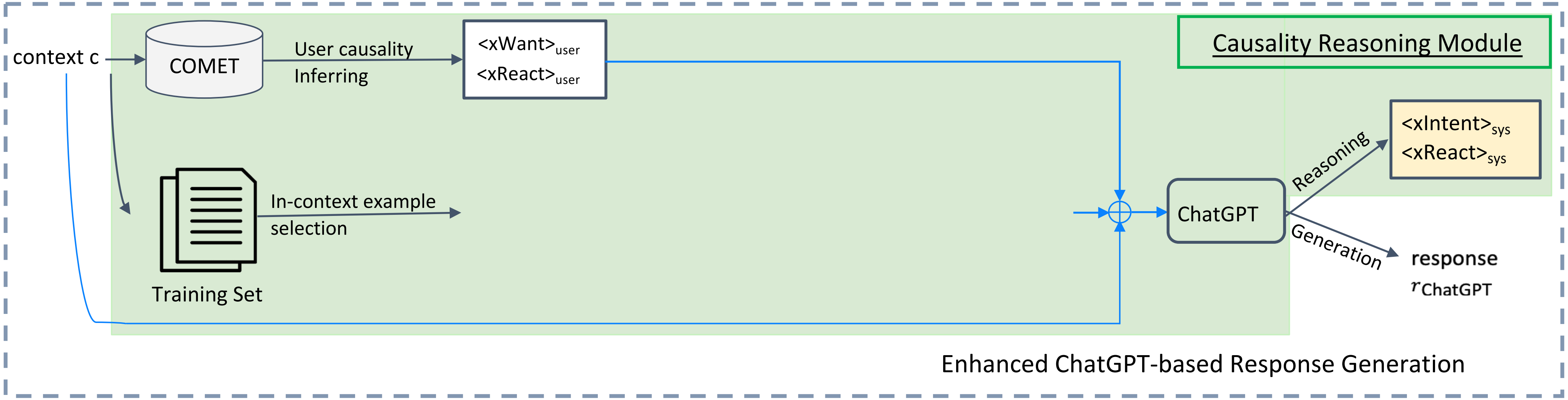
Proposed Method



Input: *context c*; *user's want/react*, *outputs of in-context reasoning process*

Output: *system's intention and reaction; response*

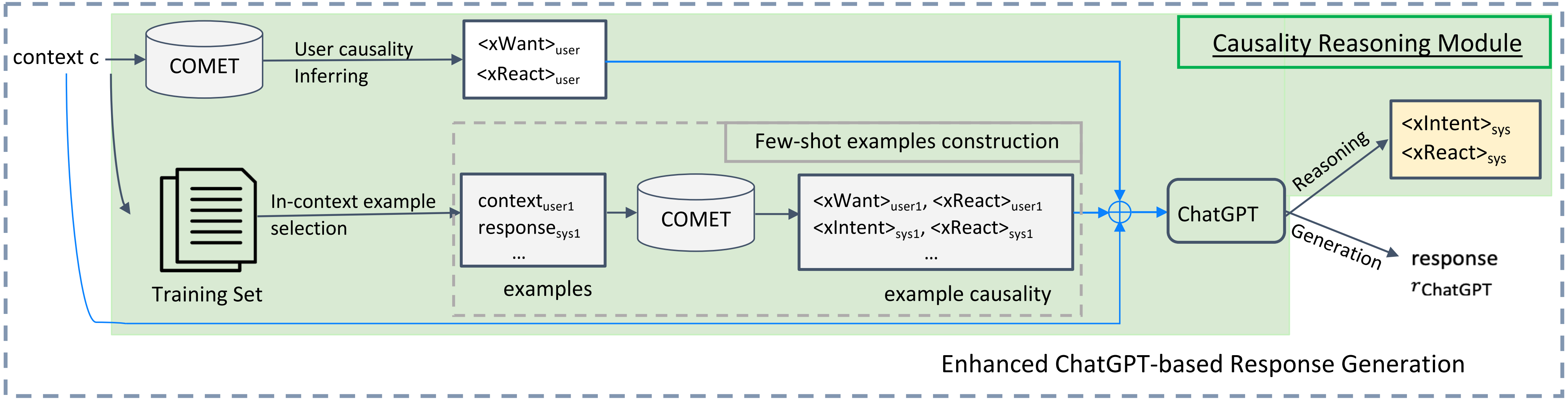
Proposed Method



- Select top-*k* conversations from training set based on cosine similariy as the in-context example.

京都大学

Proposed Method



- ➢ For the <context, reasponse> in each selected example, predict user’s want/react and system’s intention/reaction as the example causality.

In-context reasoning

Test input	user: I'm so excited because I'm finally going to visit my parents next month! I didn't see them for 3 years.	
Predictions	user wants: to spend time with family; to have fun with them; to see them again. user reacts to: excited; happy; nostalgic; anxious; joyful.	
Few-shot1	context1	user1: Someone is visiting me soon and I can't wait! sys1: Who is it? user1: My mom, she is amazing.
	example causality	user1 wants: to have a good time; to talk to their mom; to have fun with Mom. user1 reacts to: excited; happy; satisfied; good; loved. sys1's intent: to be with her; to be loved; to be nice; happy. sys1 reacts to: happy; excited; proud; good; loving.
	response1	sys1: I bet she is! I am so glad you get to see her. Mom's are awesome!
Few-shot2	context2	user2: My family is coming to visit! sys2: Awesome. When are they coming and for how long? user2: They are coming next year from Africa!
	example causality	user2 wants: to have a good time; to go to the airport; to have fun with the family. user2 reacts to: happy; excited; happy; excited; loved. sys2's intent: to see the sights; to be with family; to be with them; to have fun. sys2 reacts to: happy; excited; satisfied; tired; relieved.
	response2	sys2: That's a long trip. I hope they have a good time.
Reasoning	sys's intent: to be supportive; to be happy for them; to ask about the visit. sys reacts to: happy; excited; curious; supportive; interested. response: That's great news! I'm so happy for you. What are you planning to do when you visit them?	

User causality referring

In-context reasoning process

Resoned results

Experiment: Dataset

Dataset: EmpatheticDialogue [6]

25k empathetic conversations with 32 emotion labels.

The ratio for training/validation/test is 8:1:1.

[6] Rashkin, Hannah, et al. "Towards empathetic open-domain conversation models: A new benchmark and dataset." arXiv preprint arXiv:1811.00207 (2018).

Experiment: Number of few-shots

	EMOACC	IP	EX	ER
k=2	0.24	0.08	**0.57**	**1.10**
k=3	0.25	0.09	0.48	1.05
k=4	**0.27**	0.09	0.40	1.04
k=5	0.25	**0.10**	0.33	1.00
k=6	0.25	0.08	0.32	1.01

- EMOACC = Emotion accuracy, measured by a fine-tuned BERT-base model on the EmpatheticDialogue dataset.
- IP, EX, ER is measured by separately fine-tune pre-trained empathy identification models for each metric[7].
- IP = Interpretation
- EX= Exploration
- ER= Emotion reaction

[7] https://github.com/behavioral-data/Empathy-Mental-Health

Experiment: Results on ChatGPT

Results of automatic evaluations for single-turn.

	Method	Empathy				Coherence		
		EMOACC	ER	IP	EX	PBERT	RBERT	FBERT
k=2	ChatGPT	0.060	0.923	0.073	0.341	0.877	0.872	0.875
	ChatGPT+Causality$_{user,sys}$	**0.280**	**1.116**	**0.104**	**0.768**	**0.886**	**0.878**	**0.882**

Results of automatic evaluations for multi-turn.

	Method	Empathy				Coherence		
		EMOACC	ER	IP	EX	PBERT	RBERT	FBERT
k=2	ChatGPT	0.083	0.917	**0.065**	0.318	0.891	0.902	0.894
	ChatGPT+Causality$_{user,sys}$	**0.199**	**1.094**	0.058	**0.397**	**0.899**	**0.907**	**0.901**

Emotion expression

Cognition

- Compared with ChatGPT, ChatGPT with causality explanation can generate response with appropriate emotion and contents.

Experiment: Results on ChatGPT

Results of ***human A/B test*** evaluations.
Emp., Coh., Inf. refer to ***Empathy***, ***Coherence***, and ***Informativeness***

Comparisons	Aspects	Win	Loss	Tie
ChatGPT+Causality$_{user,sys}$ vs. ChatGPT (k=2)	Emp.	**50.7**	36.0	13.3
	Coh.	**42.7**	42.0	15.3
	Inf.	**51.3**	37.3	11.3

Experiment: Results on T5

Results of automatic evaluations

	Methods	PPL ↓	BLEU-2	BLEU-3	BLEU-4	D1	D2	PBERT	RBERT	FBERT
Baselines	MOEL	37.63	8.63	4.25	2.43	0.38	1.74	86.19	85.67	85.91
	MIME	36.84	8.37	4.31	2.51	0.28	0.95	86.27	85.59	85.92
	EmpDG	38.08	7.74	4.09	2.49	0.46	1.90	86.09	85.49	85.78
	CEM	36.36	6.35	3.55	2.26	0.54	2.38	86.61	85.39	85.98
	LEMPEx	30.42	2.1	0.8	0.35	1.02	**10.81**	83.60	83.09	83.34
Ours	T5	46.13	3.59	1.94	1.15	0.49	2.82	86.69	84.07	85.35
	T5+Causality$_{user}$	15.26	4.84	1.97	0.89	**1.08**	10.75	90.16	89.48	89.80
	T5+Causality$_{user,sys}$	**13.07**	**10.53**	**6.34**	**4.06**	0.75	5.52	**92.24**	**90.76**	**91.48**

Experiment: Results on T5

Results of ***human A/B test*** evaluations.
Emp., Coh., Inf. refer to ***Empathy***, ***Coherence***, and ***Informativeness***

Comparisons	Aspects	Win	Loss	Tie
T5+Causality$_{user,sys}$ vs. CEM	Emp.	**42.0**	40.0	18.0
	Coh.	**38.7**	33.3	28.0
	Inf.	38.3	**44.3**	17.3
T5+Causality$_{user,sys}$ vs. LEMPEx	Emp.	**53.0**	35.0	12.0
	Coh.	**39.0**	33.3	27.7
	Inf.	**50.0**	38.0	12.0

Comparation between ChatGPT and T5

Model	Empathy				Diversity		BLEU		
	EMOACC	IP	EX	ER	D1	D2	BLEU-2	BLEU-3	BLEU-4
T5+Causality$_{user,sys}$	0.125	**0.271**	0.498	0.751	0.75	5.52	**10.53**	**6.34**	**4.06**
ChatGPT+Causality$_{user,sys}$	**0.235**	0.046	**0.668**	**1.109**	**2.91**	**16.44**	3.95	2.17	1.32

- Stronger interpretation (IP), which involves *understanding and empathizing through shared experiences* is more frequently observed in the trained T5 model, while ChatGPT tends to respond from the view of a machine.
- Indicating a potential need of balancing the response diversity and the accuracy in generating empathetic response.

Case studies

	Context	user: I'm so sad because I've read an article about a newborn girl who died because her parents didn't believe in medication and doctors.
Baselines	MOEL	sys: I am so sorry to hear that. I hope you are able to find her.
	MIME	sys: That is sad. I am sorry to hear that.
	EmpDG	sys: I am sorry to hear that. I hope you find out of it.
	CEM	sys: Oh no, I am so sorry to hear that.
	LEMPEx	sys: I'm sorry to hear that. I'm so sorry to hear that.
	ChatGPT	sys: That's heartbreaking. It's important to trust medical professionals and seek help when needed.
Ours	Causality_{user}	user wants: to tell their friends; to cry; to find out what happened. user reacts to: sad; sorrowful; depressed; upset.
	Causality_{sys}	sys's intent: to comfort user; to show empathy; to be supportive. sys reacts to: sad; upset; empathetic; sympathetic; concerned.
	$\text{T5+Causality}_{user,sys}$	sys: I know what you mean, I would have been so mad!
	$\text{ChatGPT+Causality}_{user,sys}$	sys: That's really heartbreaking. I'm here for you if you need to talk or just need someone to listen.

Conclusion

- A *commonsense-based causality explanation* approach that reasons not only the user's desires/reaction but also the system's proper intention/reaction.
- Integration of T5 with ChatGPT's reasoning capability realizes more empathetic responses that result in better evaluations.
- They are *more accurate and empathetic* than the responses by ChatGPT while not so diverse.

Thanks for you attention!

Q&A