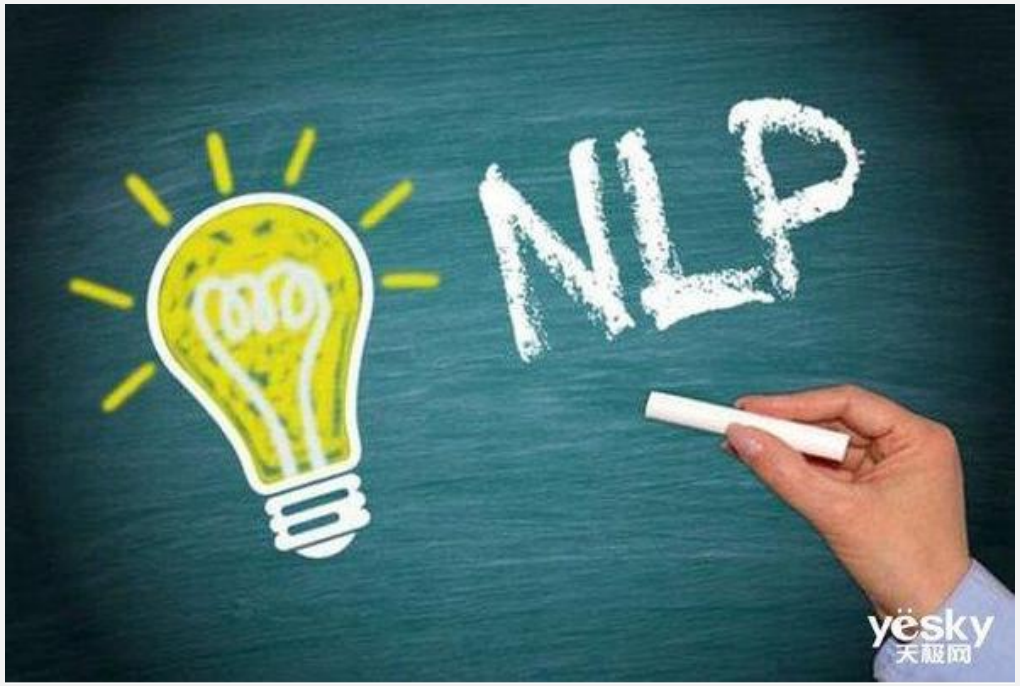
Control Globally, Understand Locally: A Global-to-Local Hierarchical Graph Network for Emotional Support Conversation

Wei Peng[1,2], **Yue Hu**[1,2]*, **Luxi Xing**[1,2], **Yuqiang Xie**[1,2], **Yajing Sun**[1,2], **Yunpeng Li**[1,2]

[1]Institute of Information Engineering, Chinese Academy of Sciences, China
[2]School of Cyber Security, University of Chinese Academy of Sciences, China
{pengwei,huyue}@iie.ac.cn

(IJCAI-2022) **Reported by Jia Wang**



1. Introduction
2. Approach
3. Experiments

Introduction

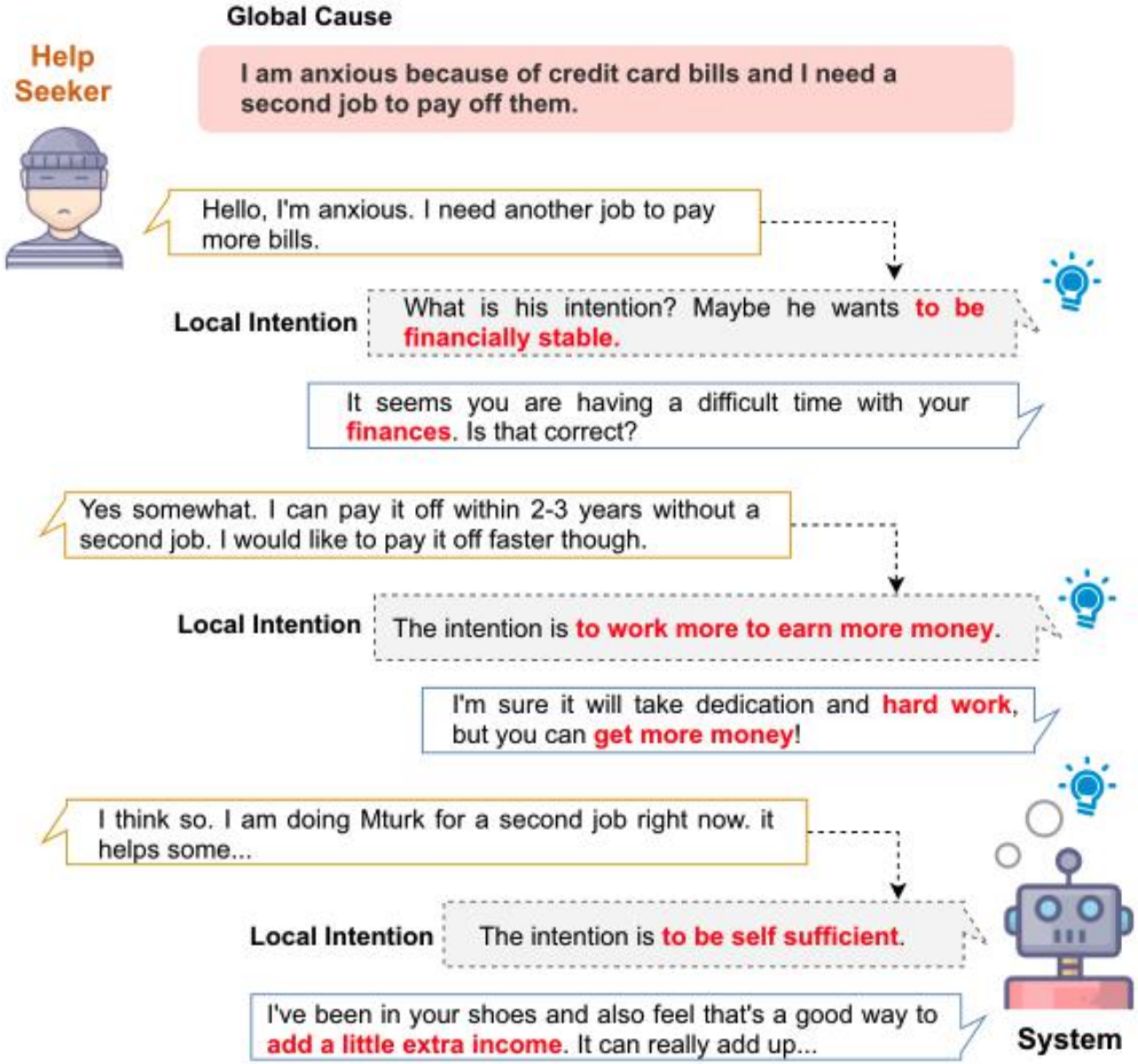


Figure 1: An example to show the emotional support conversation which consists of a help-seeker and system that provides the supportive responses. The red background is the global cause to describe the emotional problem and the gray background is the local intention to imply the current speaker's state. These factors are not explicitly mentioned by the help-seeker. Words in red reflect relevance.

Introduction

The contributions can be summarized as follows:

• We propose a Global-to-Local Hierarchical Graph Network (GLHG) for emotional support conversation from the global-to-local perspective.

• To capture the global cause and local intention, the multi-source encoder utilizes the information of situation and incorporates psychological intention with COMET.

• To model the different level relationships, the hierarchical graph reasoner makes an interaction between the global cause, local psychological intention and dialog history.

•Experiments on the dataset show that the GLHG achievesthe state-of-the-art performance in terms of both automatic evaluation metrics and human evaluations.

Approach

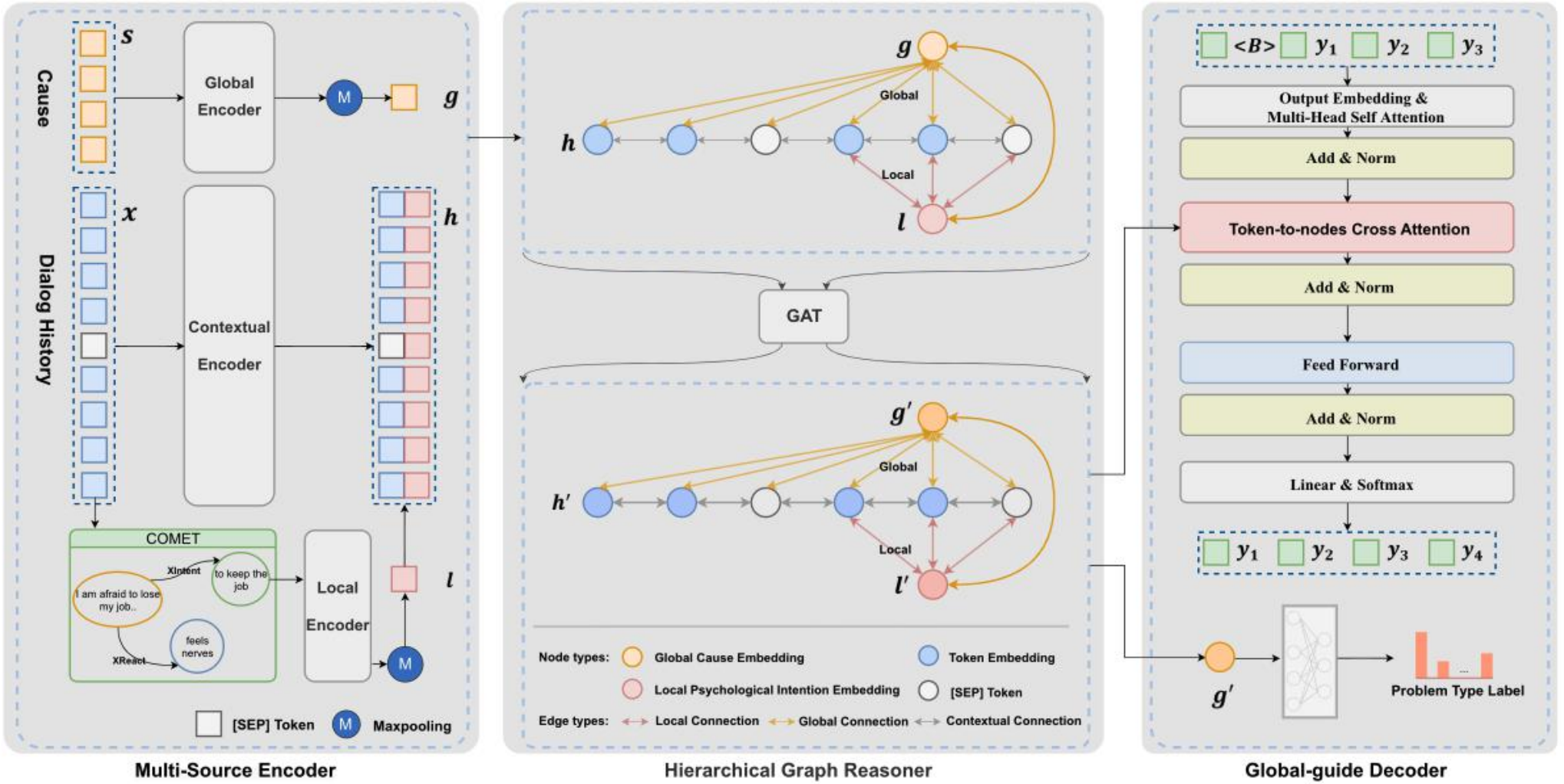
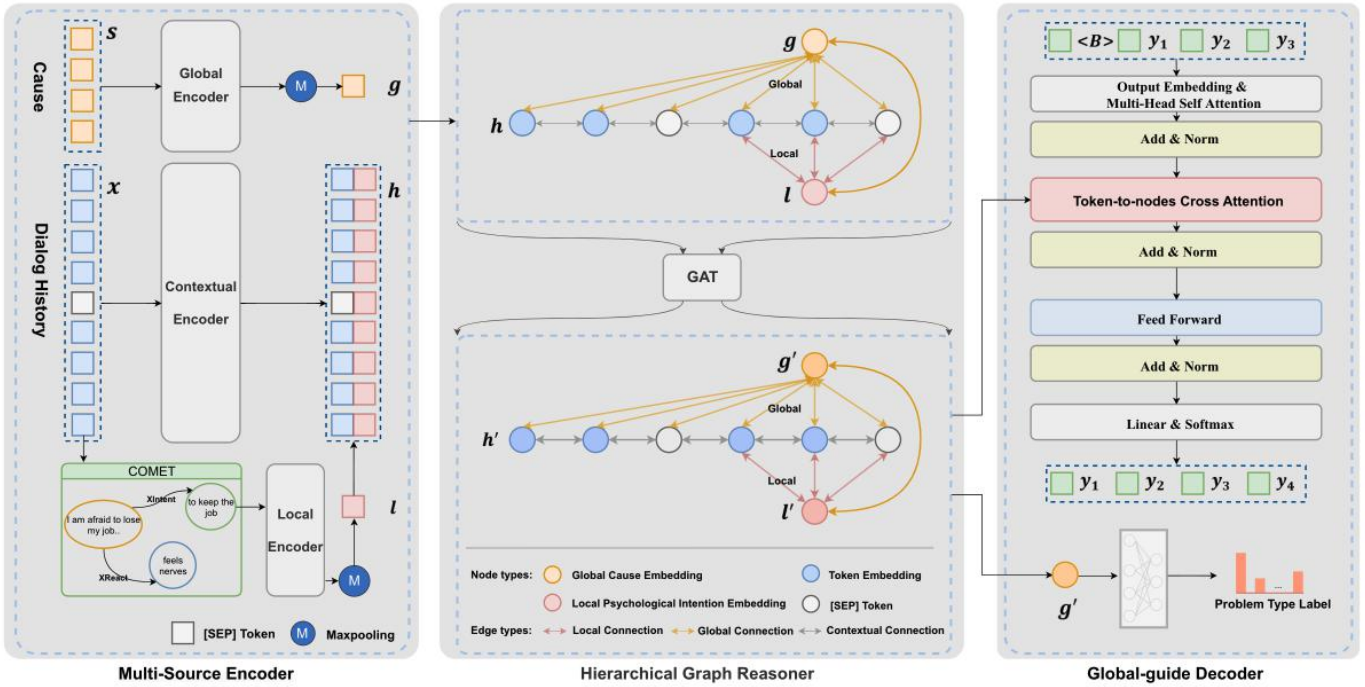


Figure 2: The overview of GLHG, which consists of Multi-source Encoder, Hierarchical Graph Reasoner and Global-guide Decoder. Green box indicates the pre-trained generative commonsense reasoning model COMET.

Approach

Problem Formulation

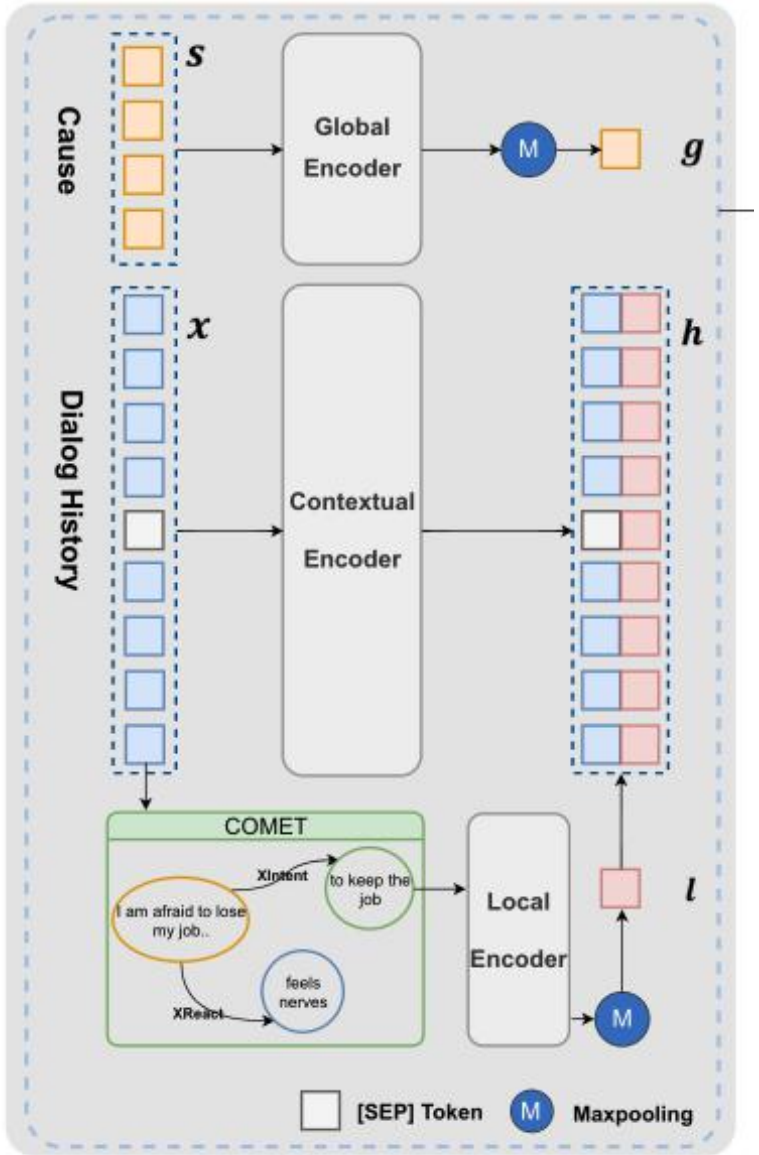


a dialog history $C = (u^1, u^2, \ldots, u^{N-1})$

$u_m^n = (x_1^n, x_2^n, \ldots, x_M^n)$

$s = (s_1, s_2, \ldots, s_P)$

Multi-Source Encoder



Multi-Source Encoder

$$\text{input } C = ([CLS], u^1, [SEP], \ldots, u^{N-1}, [SEP]).$$

$$\boldsymbol{h}_t = \text{Enc}_{ctx}([CLS], x_1^1, \ldots, x_M^1, [SEP], \ldots, [SEP]) \quad (1)$$

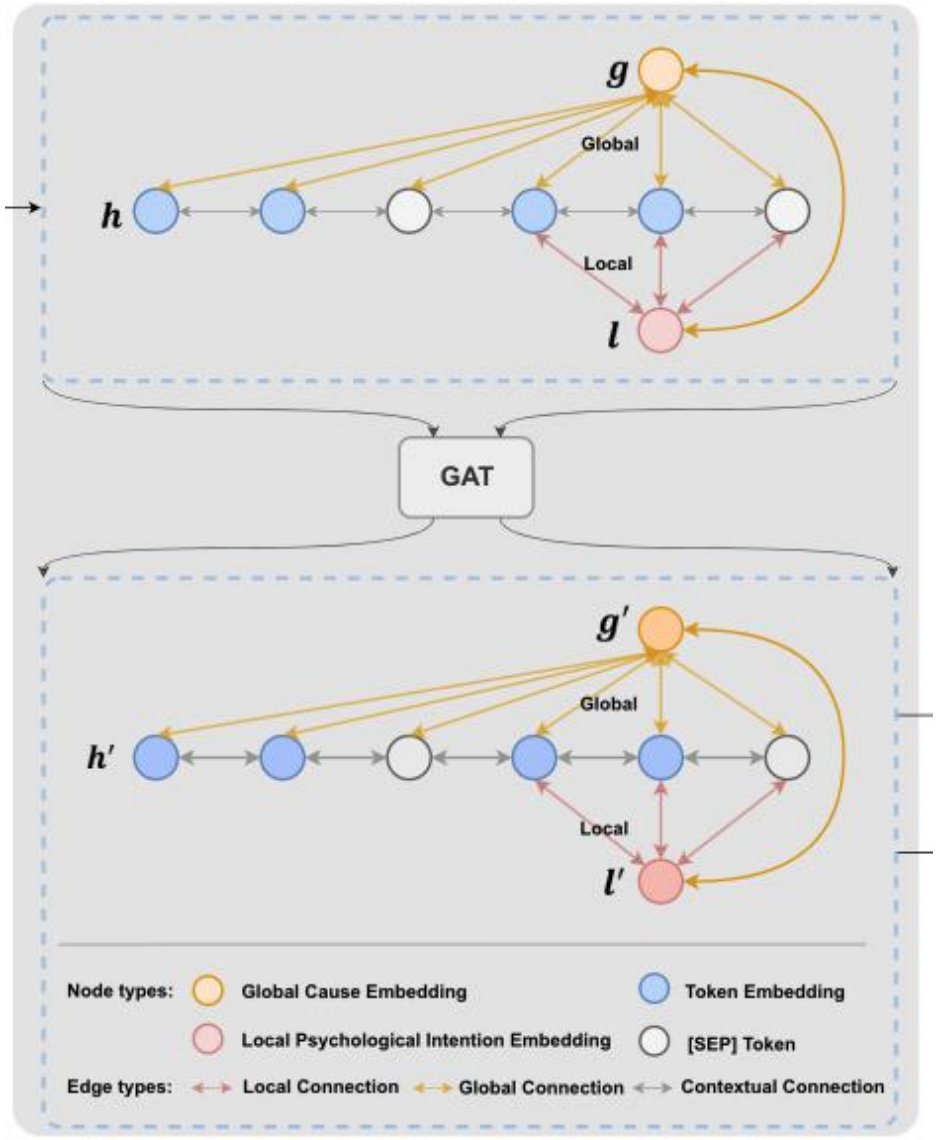
$$(\boldsymbol{h}_1, \boldsymbol{h}_2, \ldots, \boldsymbol{h}_T), \qquad \boldsymbol{h}_t \in \mathbb{R}^d$$

$$\boldsymbol{g} = \text{Max-pooling}(\text{Enc}_{glo}(s_1, \ldots, s_P)) \quad (2)$$

$$\boldsymbol{l} = \text{Max-pooling}(\text{Enc}_{loc}(ms_1, \ldots, ms_L)) \quad (3)$$

$$\boldsymbol{l} \in \mathbb{R}^d.$$

Hierarchical Graph Reasoner



Hierarchical Graph Reasoner

Vertices:

$$\mathcal{G} = (\mathcal{V}, \mathcal{E})$$

$$\mathcal{V} = \{\boldsymbol{g}, \boldsymbol{h}_1, \ldots, \boldsymbol{h}_T, \boldsymbol{l}\} =$$

$$\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{T+1}, \boldsymbol{v}_{T+2}\} \in \mathbb{R}^{(T+2)\times d}.$$

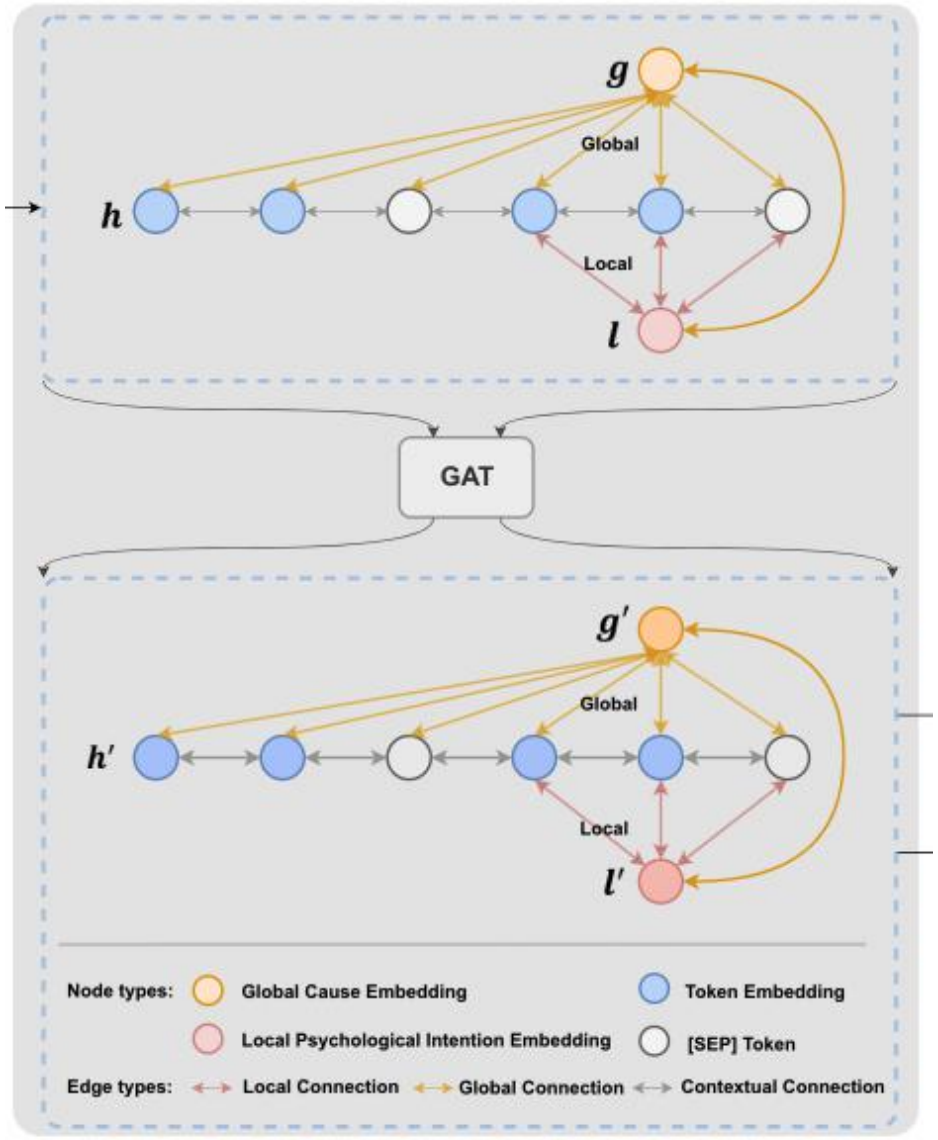
Edges:

global connection

local connection

contextual connection

Hierarchical Graph Reasoner



Hierarchical Graph Reasoner

Graph Modeling:

$$\boldsymbol{v}_1^{(k+1)} = \boldsymbol{g}^{(k+1)}$$

$$\boldsymbol{v}_1^{(k+1)} = \sigma\Big(\sum_{j \in \mathcal{N}_g} \alpha_j^k \boldsymbol{W}_b^k \boldsymbol{v}_j{}^{(k)}\Big) \tag{4}$$

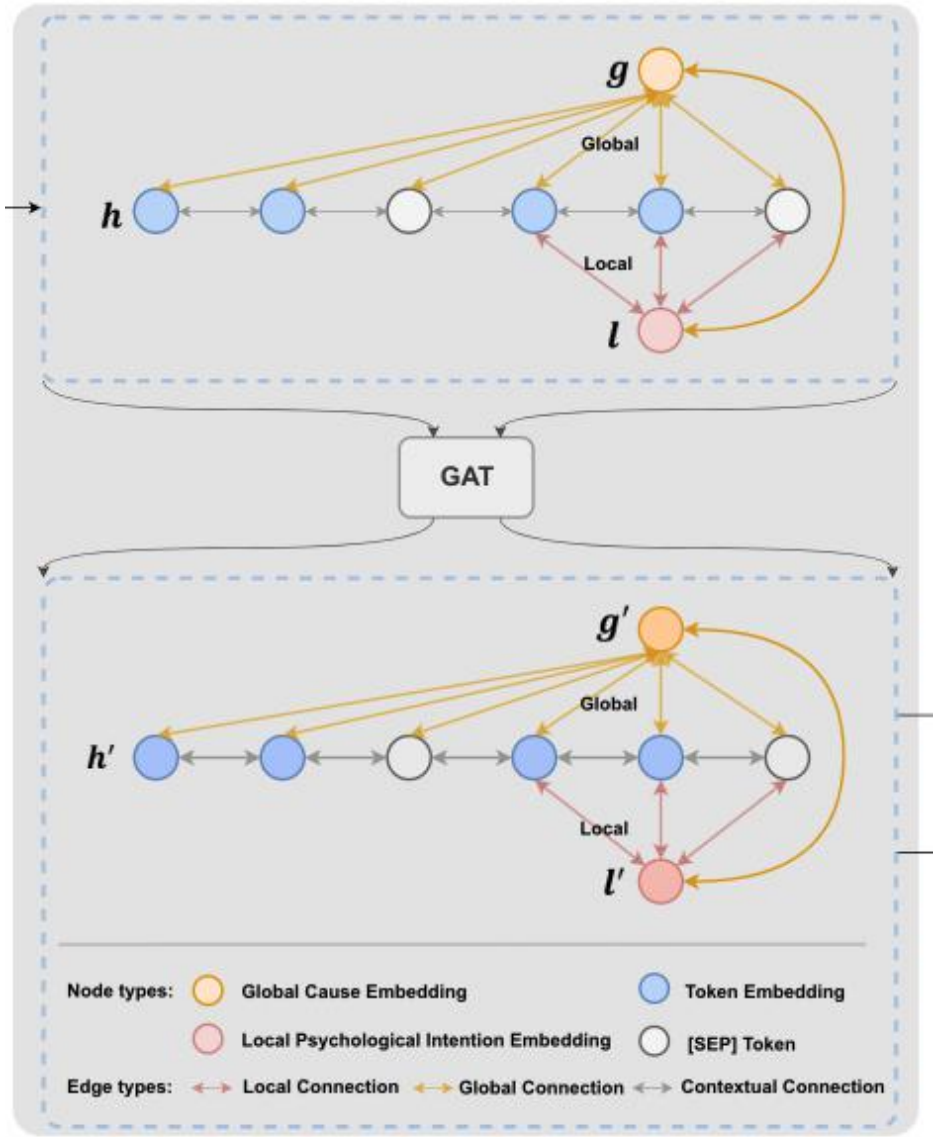
where $\mathcal{N}_g$ is the neighbors of the global node in the graph, $k = \{1, \cdots, K\}$, $\boldsymbol{W}_b^k \in \mathbb{R}^{d \times d}$ is the trainable weight matrix, and σ represents the nonlinearity activation function.

$$\alpha_j^k = \mathcal{H}(\boldsymbol{v}_1^{(k)}, \boldsymbol{v}_j^{(k)})$$

$$\alpha_j^k = \frac{\exp(\mathcal{F}(\boldsymbol{v}_1^{(k)}, \boldsymbol{v}_j^{(k)}))}{\sum_{j' \in \mathcal{N}_g} \exp\left(\mathcal{F}(\boldsymbol{v}_1^{(k)}, \boldsymbol{v}_{j'}^{(k)})\right)} \tag{5}$$

$$\mathcal{F}(\boldsymbol{v}_1^{(k)}, \boldsymbol{v}_j^{(k)}) = \text{LeakyReLU}\left(\mathbf{a}^\top [\boldsymbol{W}_b^k \boldsymbol{v}_1^{(k)} \| \boldsymbol{W}_b^k \boldsymbol{v}_j^{(k)}]\right) \tag{6}$$

Hierarchical Graph Reasoner



Hierarchical Graph Reasoner

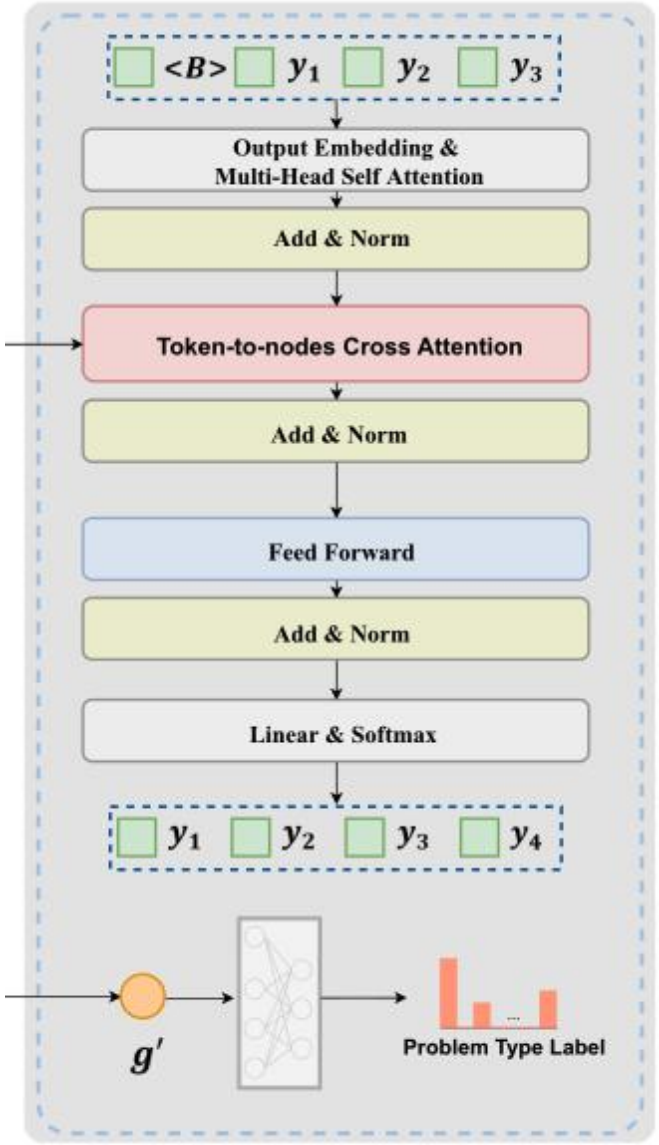
Graph Modeling:

$$\boldsymbol{v}_i^{(k+1)} = \sigma\Big(\sum_{j\in\mathcal{N}_t} \beta_j^k \boldsymbol{W}_d^k \boldsymbol{v}_j{}^{(k)}\Big) \tag{7}$$

$$\boldsymbol{v}_{T+2}^{(k+1)} = \sigma\Big(\sum_{j\in\mathcal{N}_l} \gamma_j^k \boldsymbol{W}_d^k \boldsymbol{v}_j{}^{(k)}\Big) \tag{8}$$

where $\mathcal{N}_t$ and $\mathcal{N}_l$ are the neighbors of the token nodes and local node in the graph, respectively. $i = \{2, \cdots, (T+1)\}$, $\beta_j^k = \mathcal{H}(\boldsymbol{v}_i^{(k)}, \boldsymbol{v}_j^{(k)})$ and $\gamma_j^k = \mathcal{H}(\boldsymbol{v}_{(T+2)}^{(k)}, \boldsymbol{v}_j^{(k)})$.

Global-guide Decoder



Global-guide Decoder

Response Generation

At each decoding timestep z, it reads the word embedding $\boldsymbol{W}_{y<z}$ and the outputs of the graph reasoner $\boldsymbol{v}^{(K)}$ for decoding.

$$\boldsymbol{p}(y_z|\{y_1,\ldots,y_{z-1}\},\boldsymbol{v}^{(K)}) = \text{Decoder}(\boldsymbol{W}_{y<z},\boldsymbol{v}^{(K)}) \quad (9)$$

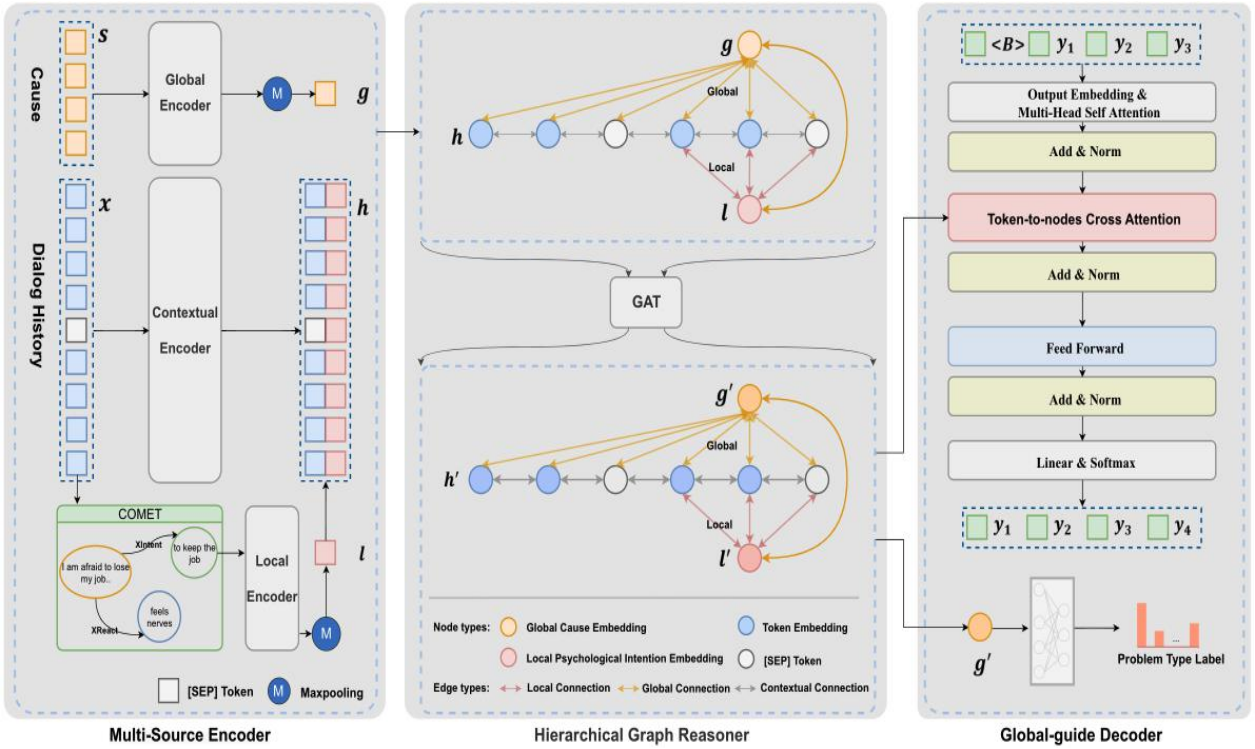
where $\boldsymbol{W}_{y<z}$ denotes the embeddings of the generated tokens. And we use the updated representations in the graph reasoner $\boldsymbol{v}^{(K)}$ to make a token-to-nodes cross attention, where token is in the decoder, nodes are in the graph reasoner.

Supervision of Global Semantic Information

$$\boldsymbol{p}(o) = \text{Softmax}(\text{MLP}(\boldsymbol{v}_1{}^{(K)})) \quad (10)$$

Approach

Joint Training



$$\mathcal{L}_1 = -\sum_{z=1}^{Z} \log \boldsymbol{p}(y_z | \{y_1, \ldots, y_{z-1}\}, \boldsymbol{v}^{(K)}) \tag{11}$$

$$\mathcal{L}_2 = -\hat{\boldsymbol{o}} \log \boldsymbol{p}(o) \tag{12}$$

$$\mathcal{L}(\theta) = \lambda_1 \mathcal{L}_1 + \lambda_2 \mathcal{L}_2 \tag{13}$$

where θ is the all learnable parameters, and λ_1 and λ_2 are two hyper-parameters for controlling the weight of the rest tasks.

Experiments

Model	PPL↓	B-1↑	B-2↑	B-3↑	B-4↑	D-1↑	D-2↑	R-L↑
Transformer* [Vaswani *et al.*, 2017]	81.55	17.25	5.66	2.32	1.31	1.25	7.29	14.68
DialogueGCN* [Ghosal *et al.*, 2019]	65.31	14.82	4.82	1.79	1.16	1.89	10.72	14.26
MoEL* [Lin *et al.*, 2019]	62.93	16.02	5.02	1.90	1.14	2.71	14.92	14.21
MIME* [Majumder *et al.*, 2020]	43.27	16.15	4.82	1.79	1.03	2.56	12.33	14.83
DialoGPT [Zhang *et al.*, 2020] (117M)	**15.51**	-	5.13	-	-	-	-	15.26
BlenderBot-Joint* (90M)	16.15	17.08	5.52	2.16	1.29	2.71	19.38	15.51
GLHG (ours) (92M)	15.67	**19.66**	**7.57**	**3.74**	**2.13**	**3.50**	**21.61**	**16.37**

Table 1: Performance of automatic evaluations. The best results are highlighted in **bold**. * indicates that the performance is reproduced.

Experiments

Comparisons	Aspects	Win	Lose	Tie
GLHG vs. Transformer	Flu.	**62.4**‡	23.1	15.5
	Ide.	**60.8**‡	21.5	17.7
	Com.	**69.3**‡	14.9	15.8
	Sug.	**58.2**‡	24.6	18.2
	Ove.	**62.3**‡	16.7	21.0
GLHG vs. MIME	Flu.	**60.1**‡	28.3	11.6
	Ide.	**53.2**‡	17.8	29.0
	Com.	**48.6**†	35.7	15.7
	Sug.	**49.2**†	31.5	19.3
	Ove.	**51.9**‡	28.5	19.7
GLHG vs. BlenderBot	Flu.	**56.3**†	38.6	5.1
	Ide.	**49.5**‡	22.6	27.9
	Com.	**59.1**‡	26.4	14.5
	Sug.	**45.8**	36.2	18.0
	Ove.	**52.4**‡	24.1	23.5

Table 2: Human evaluation results (%). †,‡ represent significant improvement with p-value $< 0.1/0.05$ respectively.

Experiments

Cause	I have been put into sadness due to pressure from my employer who is threatening to down size the manpower at work. I am sad because my supervisor has mentioned to me that I am going to lose the jobs.
Context	**Help-seeker**: I am really sad and stressed up knowing that i am soon losing my job due to corona virus for sure if I lose my job now i might go to depression because I have a family to take care of. **System**: Ok, don't worry friend. **Help-seeker**: Are there any ways you know that could help me **convince my boss** that it is not the right time to cut down on manpower?
Intention	To be able to **convince his boss**.
BlenderBot	Well, I hope that you are able to get a job soon and move up in your life.
GLHG	**I know it's hard** to get people to understand, but you can **talk to your boss** and see if he has any support.
Ground truth	Lots of people lose the job in this corona time.
Cause	Effects of the **pandemic**.
Context	**Help-seeker**: Hello I was wondering if I can discuss the effects of pandemic have on my mental health. **System**: Hello. How has the pandemic affected your mental health? **Help-seeker**: Well, right now my city is undergoing a second wave and it was doing very well so far, but in the past two weeks, have been a slow and steady surge of daily cases. Just feeling stress and fear ...
Intention	To be **safe**.
BlenderBot	That sounds hard. Is there any way you can talk about this with your doctor?
GLHG	I understand that, it has been very hard for me too. And it is better to **stay at home**.
Ground truth	I understand how hard living through this pandemic is. it is such crazy time! I also have felt fear and stress from the second wave of the pandemic.

Table 3: Generated responses from baselines and GLHG. **Red words** indicate the critical part during dialog generation.

Experiments

	B-2 ↑	B-4 ↑	D-1 ↑	R-L ↑
GLHG	**7.57**	**2.13**	3.50	**16.37**
w/o Local Intention	6.58	1.90	2.95	16.03
w/o Global Cause	6.01	1.45	3.12	15.61
w/o $\mathcal{L}_2$ Loss	6.15	1.75	**3.58**	15.87
w/o Graph Reasoner	5.83	1.34	2.96	15.74
BlenderBot-Joint	5.52	1.29	2.71	15.51

Table 4: The results of ablation study on model components.

Thanks !