Knowledge-Bridged Causal Interaction Network for Causal Emotion Entailment

Weixiang Zhao, Yanyan Zhao *, Zhuojun Li, Bing Qin
Research Center for Social Computing and Information Retrieval
Harbin Insititute of Technology, China
{wxzhao, yyzhao, zjli, qinb}@ir.hit.edu.cn

Code：https://github.com/circle-hit/KBCIN.

2022.12.10 • ChongQing

—— AAAI 2023

Reported by Renhui Luo

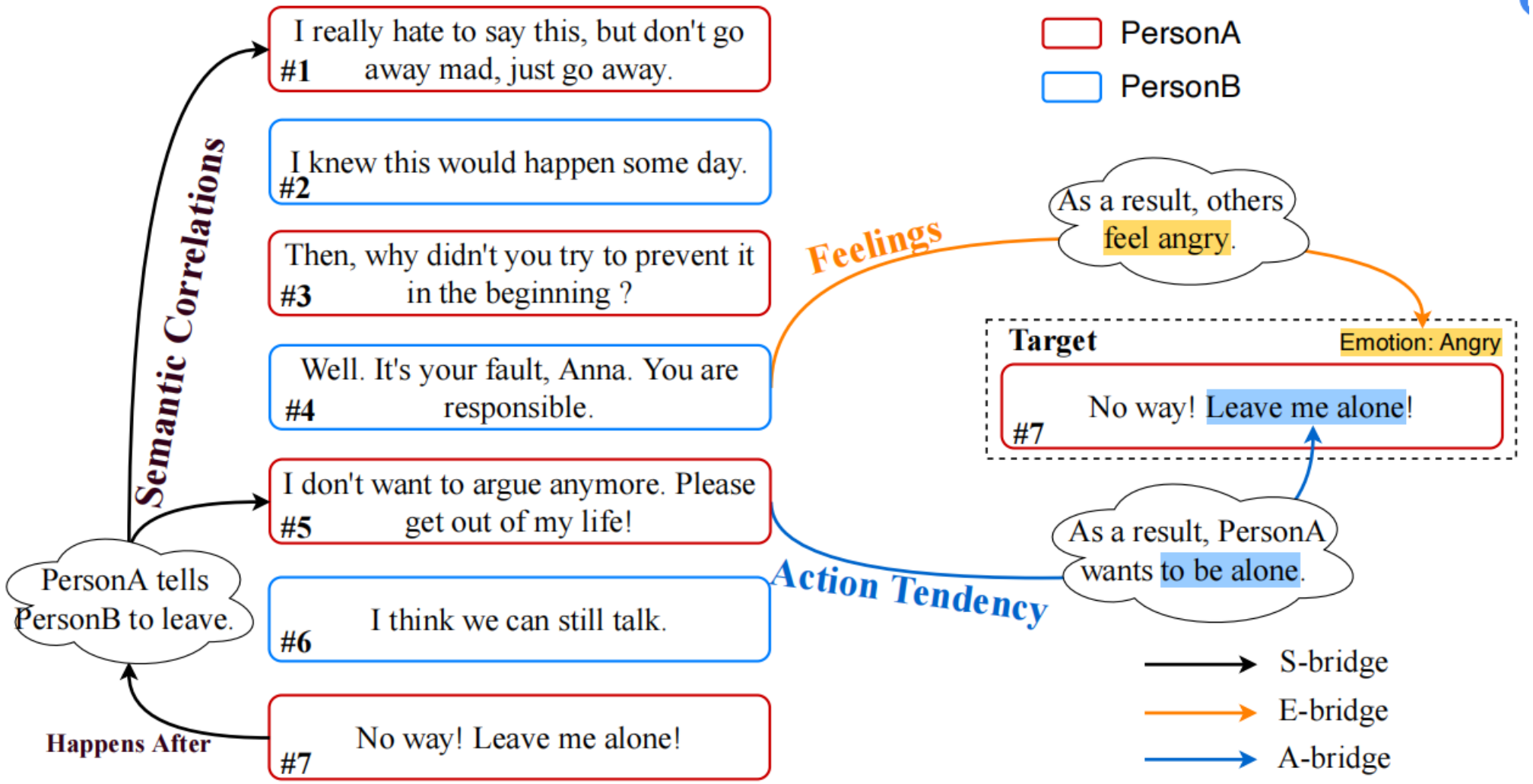
1.Introduction

2.Overview

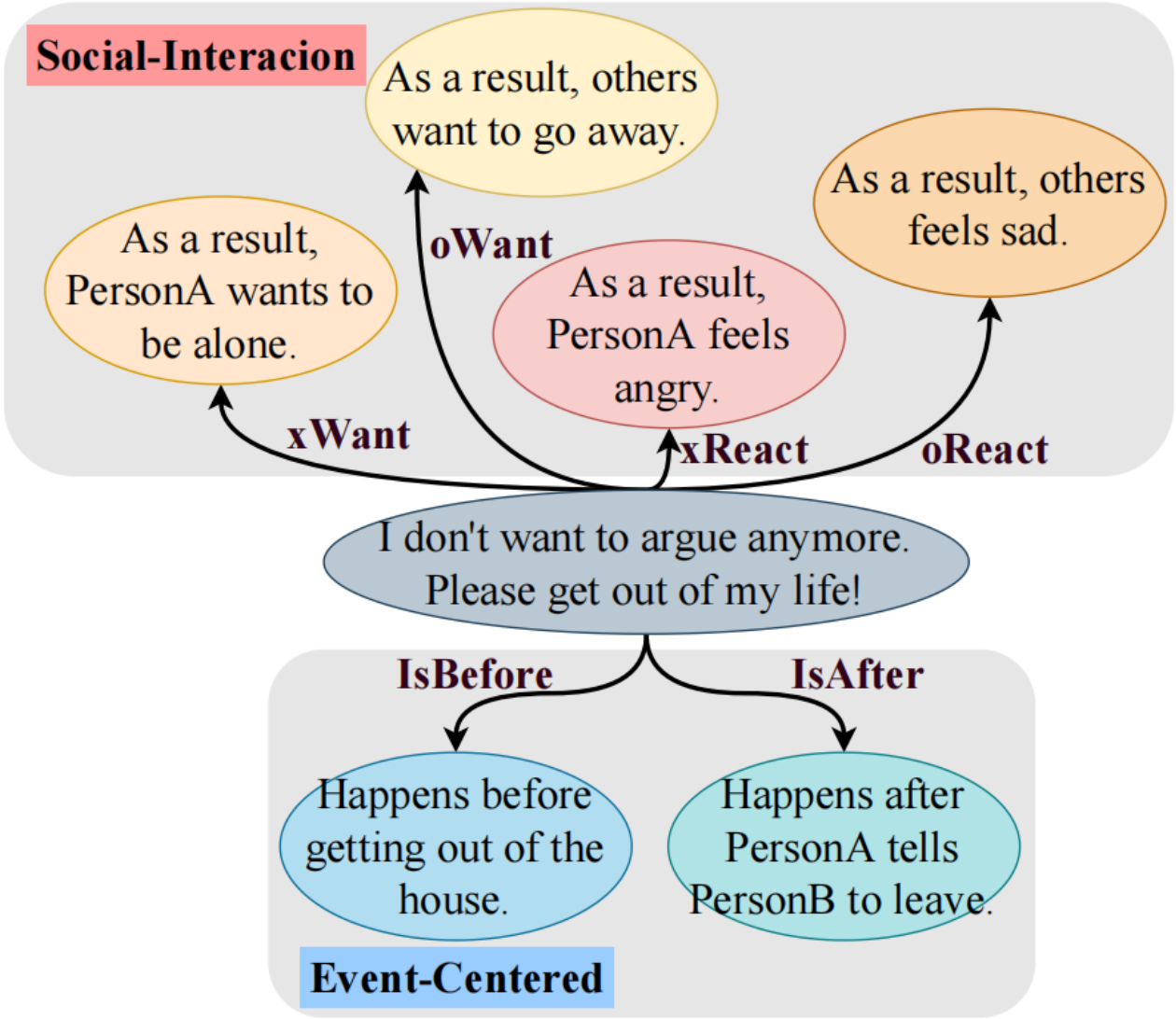
3.Methods

4.Experiments

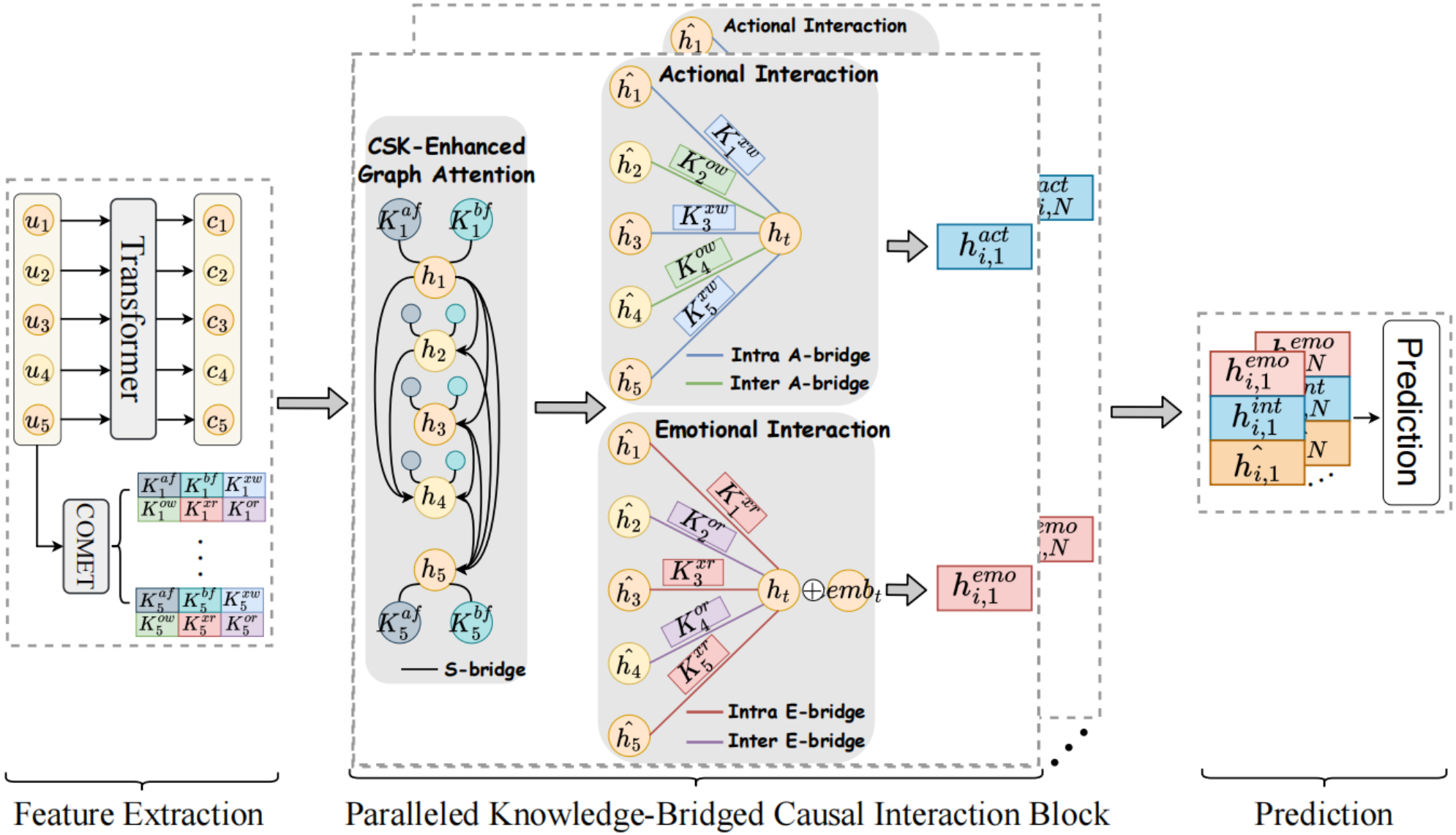
Introduction



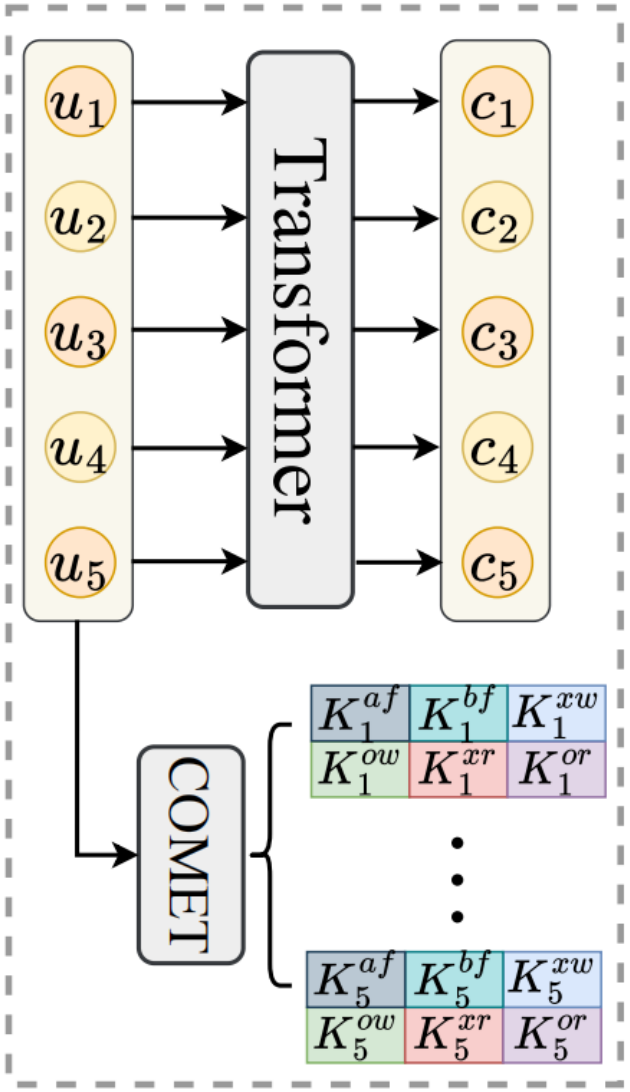
Introduction



Overview


Transformer
COMET
CSK-Enhanced Graph Attention
S-bridge
Actional Interaction
Intra A-bridge
Inter A-bridge
Emotional Interaction
Intra E-bridge
Inter E-bridge
Prediction
Feature Extraction
Paralleled Knowledge-Bridged Causal Interaction Block
Prediction

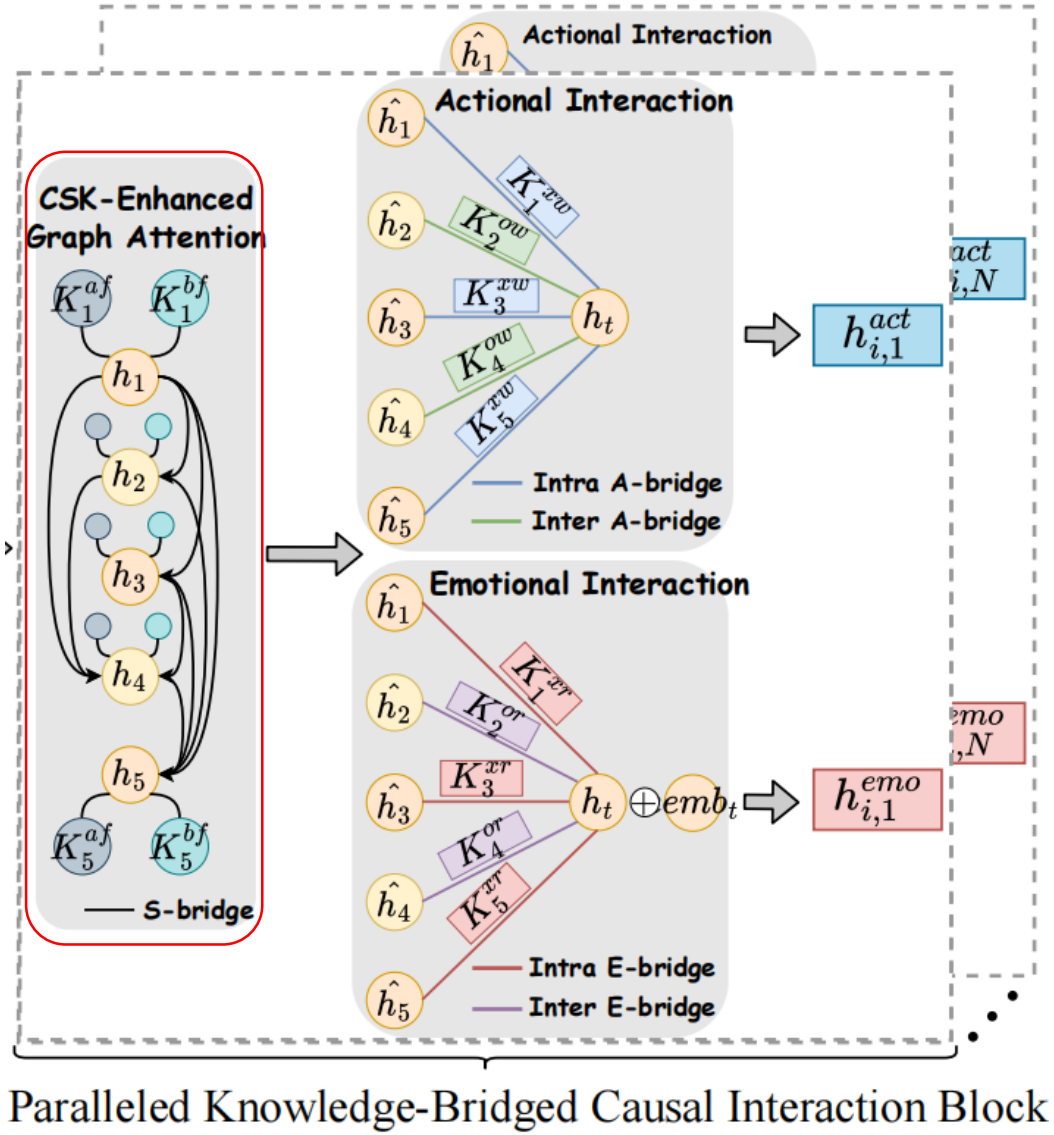
Method



$$c_i = \text{Maxpooling}(\text{Transformer}(\ [CLS], w_1, w_2, \cdots, w_L\ )) \quad (1)$$

$K_i^r, r \in \{af, bf, xr, or, xw, ow\}$ and af, bf, xr, or, xw, ow are short for the relation type *isAfter*, *isBefore*, *xReact*, *oReact*, *xWant* and *oWant*, respectively.

Method



Paralleled Knowledge-Bridged Causal Interaction Block

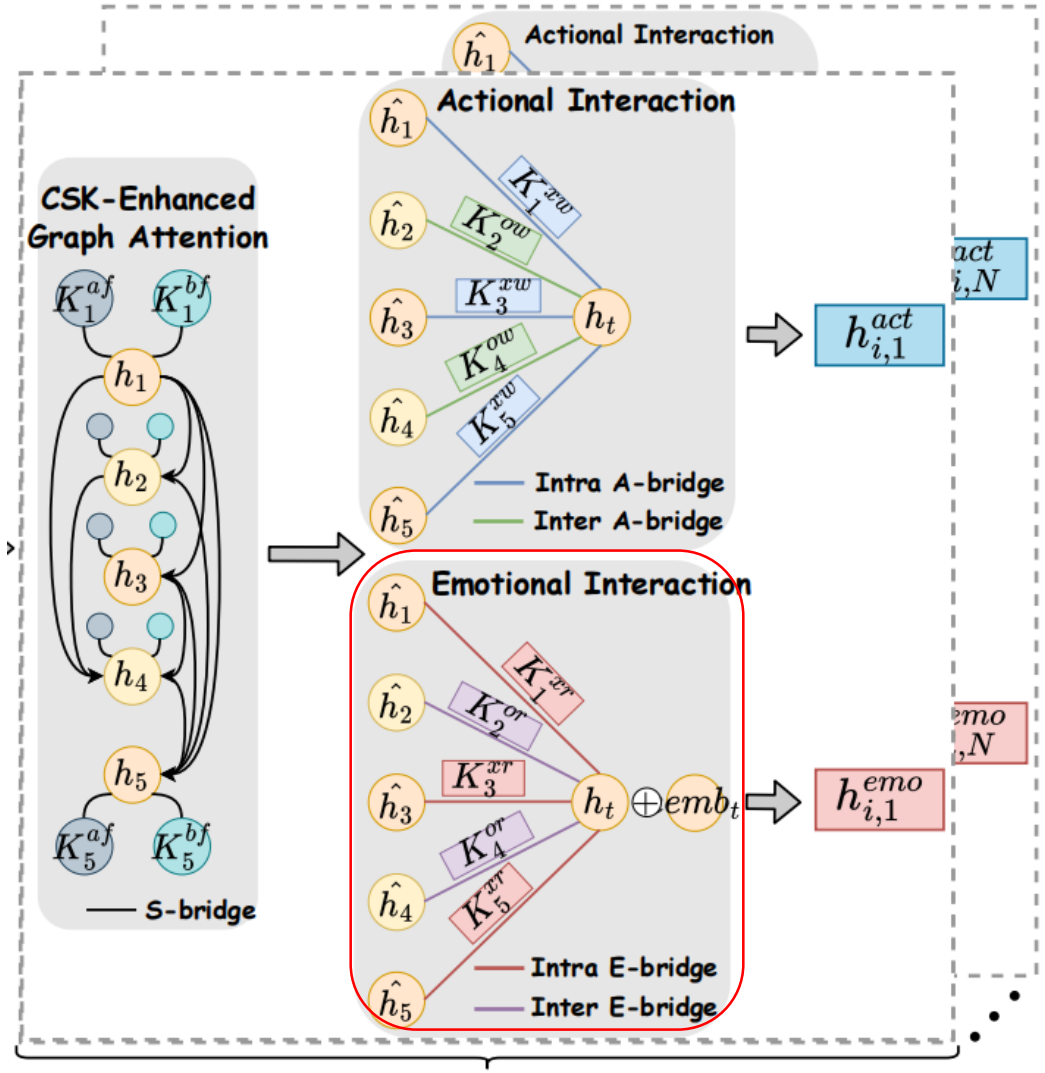
$$h_i = c_i \oplus pemb_i \oplus eemb_i \tag{2}$$

$$\hat{h}_i = \sigma\Big(\sum_{j \in N_i} \alpha_{ij} W_h h_j\Big) \tag{3}$$

$$\alpha_{ij} = \frac{\exp(\mathcal{F}(h_i, h_j, K_j^{af}, K_j^{bf}))}{\sum_{j' \in N_i} \exp(\mathcal{F}(h_i, h_{j'}, K_{j'}^{af}, K_{j'}^{bf}))} \tag{4}$$

$$\mathcal{F}(h_i, h_j, K_j^{af}, K_j^{bf}) = \text{LeakyReLU}(a^\top [W_h h_i \,\|\, W_h h_j + W_e K_j^{af} + W_e K_j^{bf}]) \tag{5}$$

Method

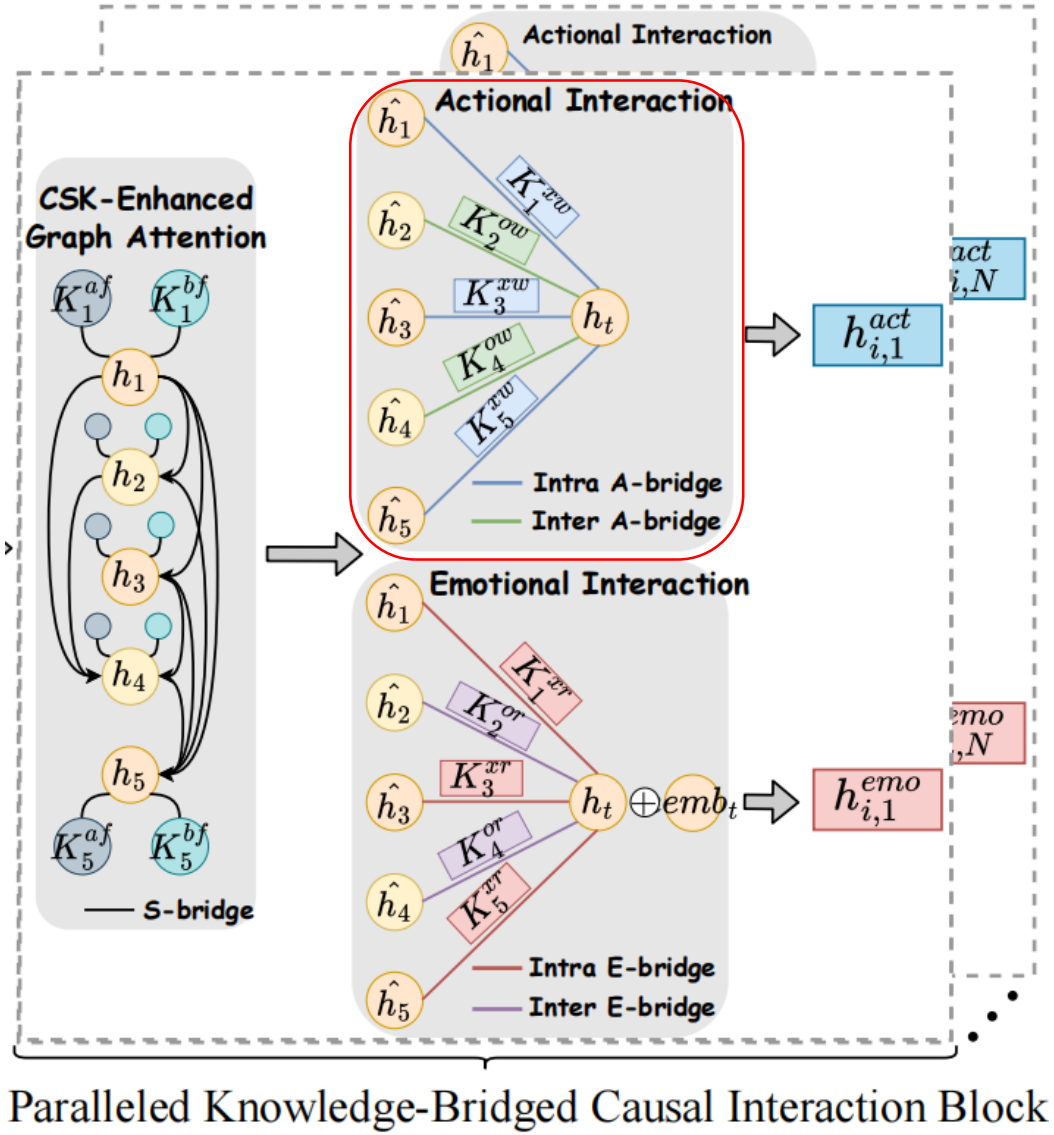


Paralleled Knowledge-Bridged Causal Interaction Block

$$Q^{emo} = f_q(h_t + eemb_t),\ K^{emo} = f_k(\hat{h}_i) + f_e(K_i^r)$$
$$s_i^{emo} = \text{softmax}(\frac{Q^{emo}(K^{emo})^\top}{\sqrt{d_h}}) \quad (6)$$

$$h_i^{emo} = s_i^{emo} V^{emo} + s_i^{emo} Q^{emo}$$
$$V^{emo} = f_v(\hat{h}_i) + f_e(K_i^r) \quad (7)$$

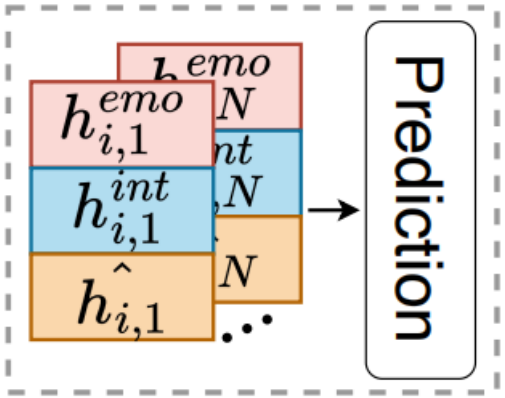
Method



$$Q^{act} = f'_q(h_t),\ K^{act} = f'_k(\hat{h}_i) + f'_e(K^r_i)$$
$$s^{act}_i = \mathrm{softmax}(\frac{Q^{act}(K^{act})^\top}{\sqrt{d_h}}) \quad (8)$$

$$h^{act}_i = s^{act}_i V^{act} + s^{act}_i Q^{act}$$
$$V^{act} = f'_v(\hat{h}_i) + f'_e(K^r_i) \quad (9)$$

Method



$$\tilde{h}_i = \hat{h}_i + h_i^{emo} + h_i^{act} \tag{10}$$

$$\hat{y}_i = \text{sigmoid}(\text{MLP}(\|_{n=1}^{N} \tilde{h}_{i,n})) \tag{11}$$

Experiments

	Train	Valid	Test
Positive Causal Pairs	7,027	328	1,767
Negative Causal Pairs	20,646	838	5,330
Num. of Dialogue	834	47	225
Num. of Utterance	8,206	493	2,405

Table 1: Dataset statistics

Experiments

Model	Neg. F1	Pos. F1	macro F1
ECE Methods			
KAG	86.35	58.18	72.26
Adapted	88.18	64.53	76.36
ECPE Methods			
ECPE-2D	94.96	55.50	75.23
ECPE-MLL	94.68	48.48	71.59
RankCP	97.30	33.00	65.15
CEE Methods			
RoBERTa-Base	88.74	64.28	76.51
RoBERTa-Large	87.89	66.23	77.06
KEC	88.85	66.55	77.70
KBCIN (Ours)	89.65	**68.59**	**79.12**

Table 2: Comparison of our model against state-of-the-art baselines of CEE, ECE and ECPE.

Experiments

S-bridge	E-bridge	A-bridge	Pos. F1	macro F1
✔	✔	✔	**68.59**	**79.12**
✘	✔	✔	67.47	78.63
✔	✘	✔	66.33	77.78
✔	✔	✘	66.92	78.37
✘	✘	✘	57.59	71.81

Table 3: Results of ablation study.

Experiments

Emotion	Pos. F1	macro F1
Gold Emotion	**68.59**	**79.12**
Predicted Emotion	67.51	78.43
No Emotion	64.05	76.73

Table 4: Results of our proposed KBCIN with different ways of using emotion information.

Experiments

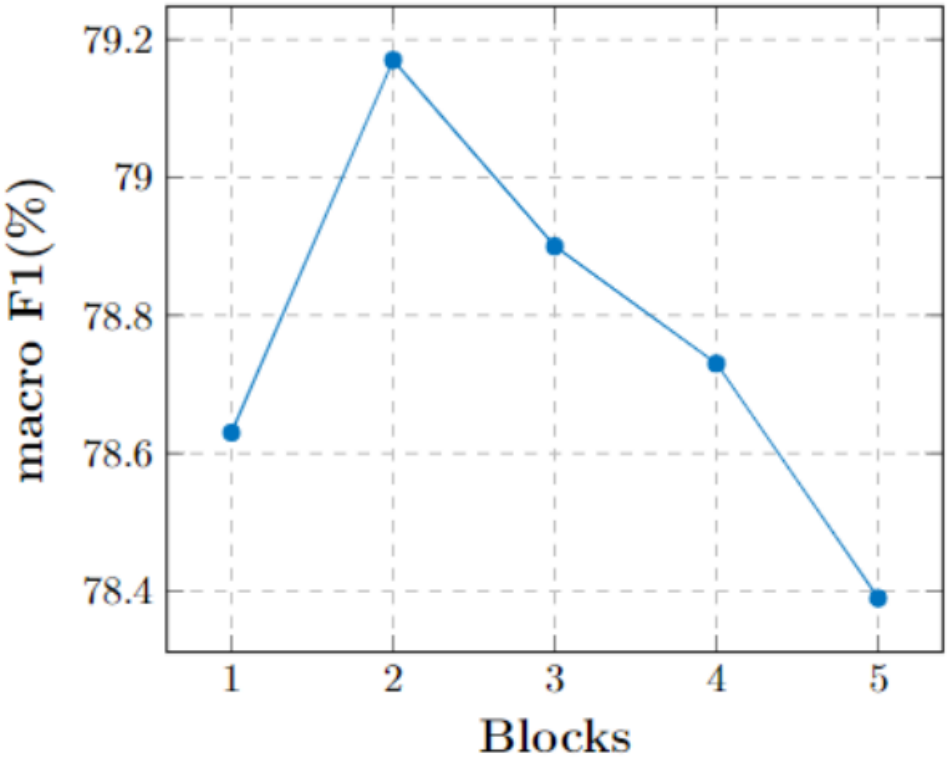


Figure 4: Results of our proposed model with different numbers of Knowledge-Bridged Causal Interaction block.

Experiments

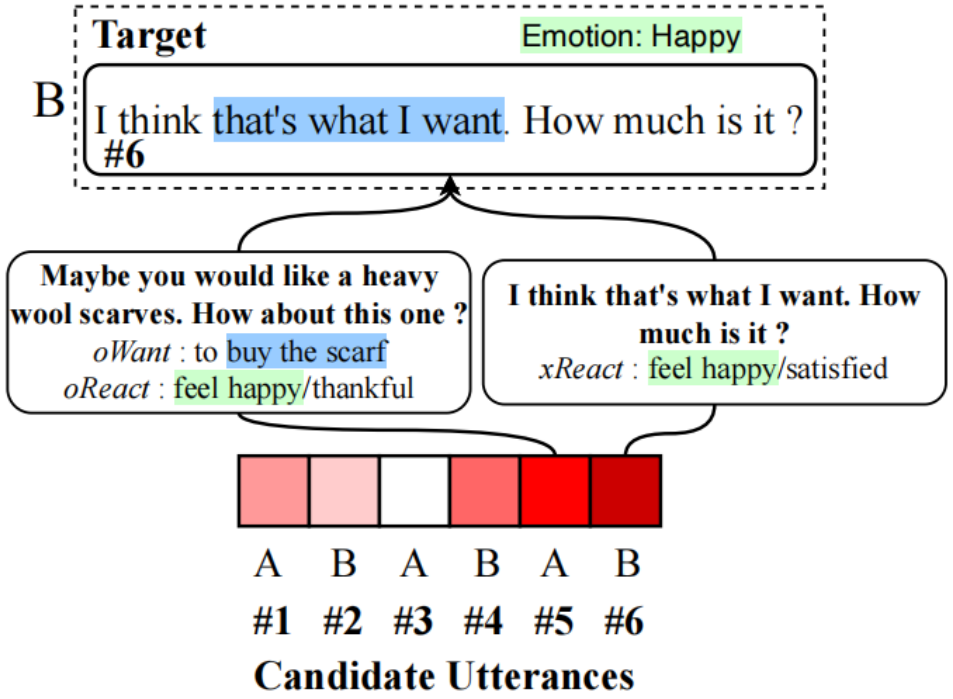


Figure 5: A case that our model makes the right prediction, along with the visualizations for attention weights of the candidate utterances (obtained by summing up the measuring results after the process of Emotional Interaction and Actional Interaction from each block). The darker color mean larger attention weights.

Thanks!