TREA: Tree-structure Reasoning Schema for Conversational Recommendation

Wendi Li[1,2], Wei Wei[1,2,✉], Xiaoye Qu[1],
Xianling Mao[3], Ye Yuan[4], Wenfeng Xie[4], Dangyang Chen[4]
[1]Cognitive Computing and Intelligent Information Processing (CCIIP) Laboratory, Huazhong University of Science and Technology
[2]Joint Laboratory of HUST and Pingan Property & Casualty Research (HPL)
[3]Department of Computer Science and Technology, Beijing Institute of Technology
[4]Ping An Property & Casualty Insurance company of China
[1]{wendili,weiw,xiaoye}@hust.edu.cn [3]maoxl@bit.edu.cn
[4]{yuanye503,xiewenfeng801,chendangyang273}@pingan.com.cn

Code:t https://github.com/WindyLee0822/TREA.

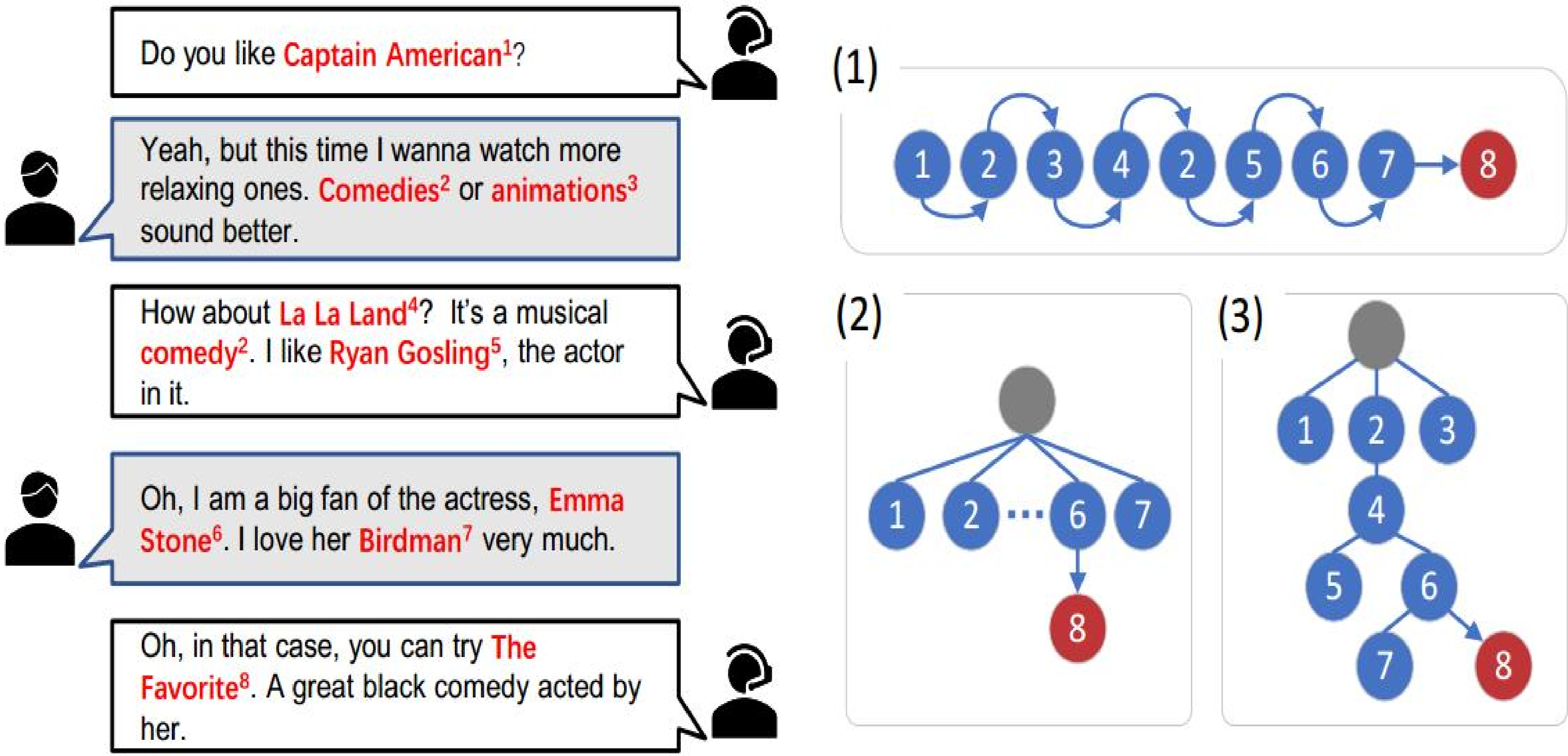
—— ACL 2023

2023.11.12 • ChongQing

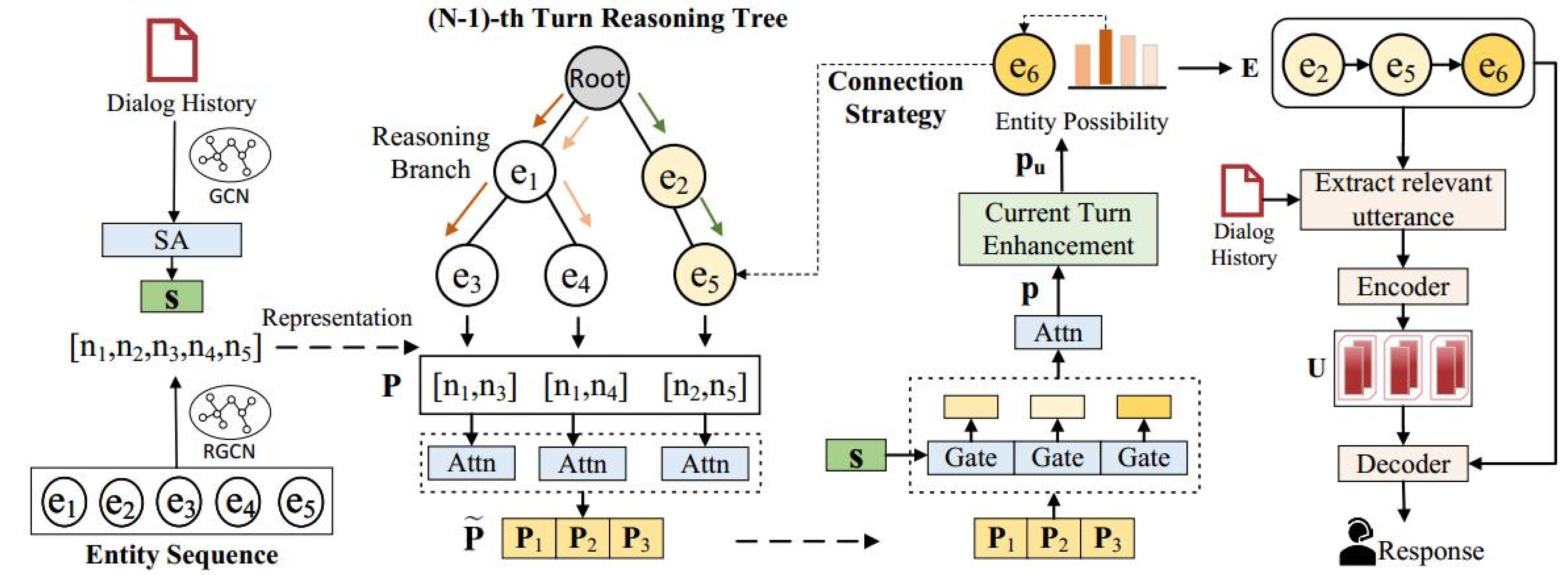
Reported by ZhenDong Huanzg

Motivation

(1)these methods fail to model the complex causal relations among mentioned entities, owing to the diversity of user interest expression and the frequent shift of conversation topic

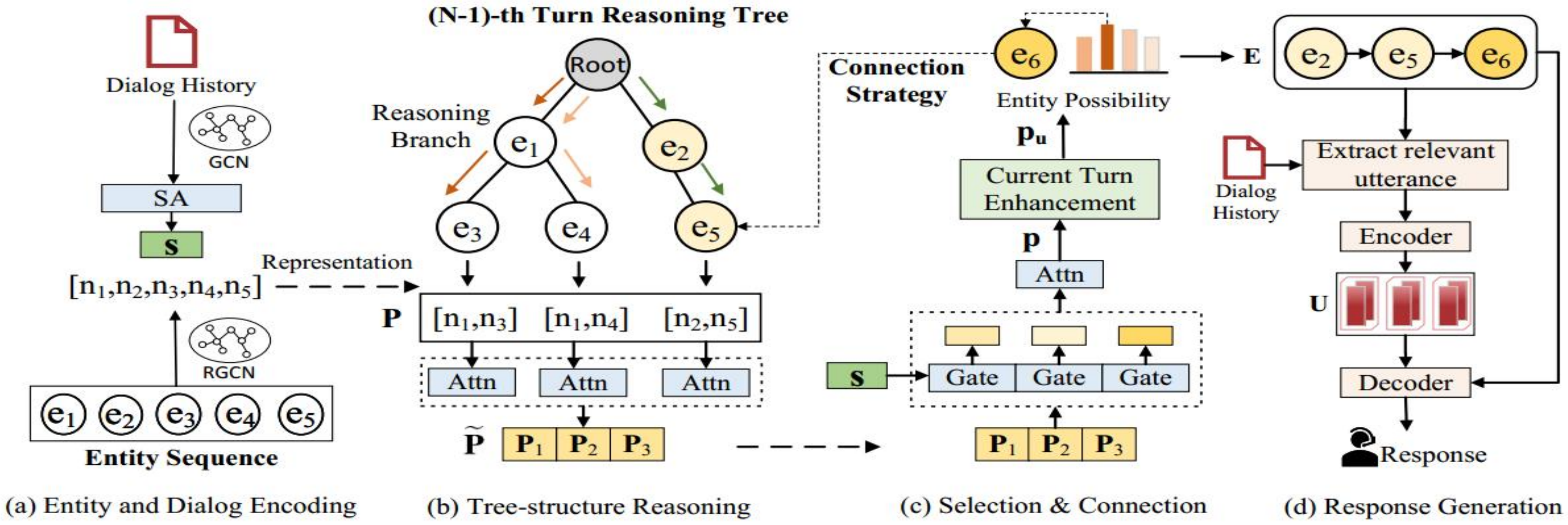


Overview



(a) Entity and Dialog Encoding (b) Tree-structure Reasoning (c) Selection & Connection (d) Response Generation

Method



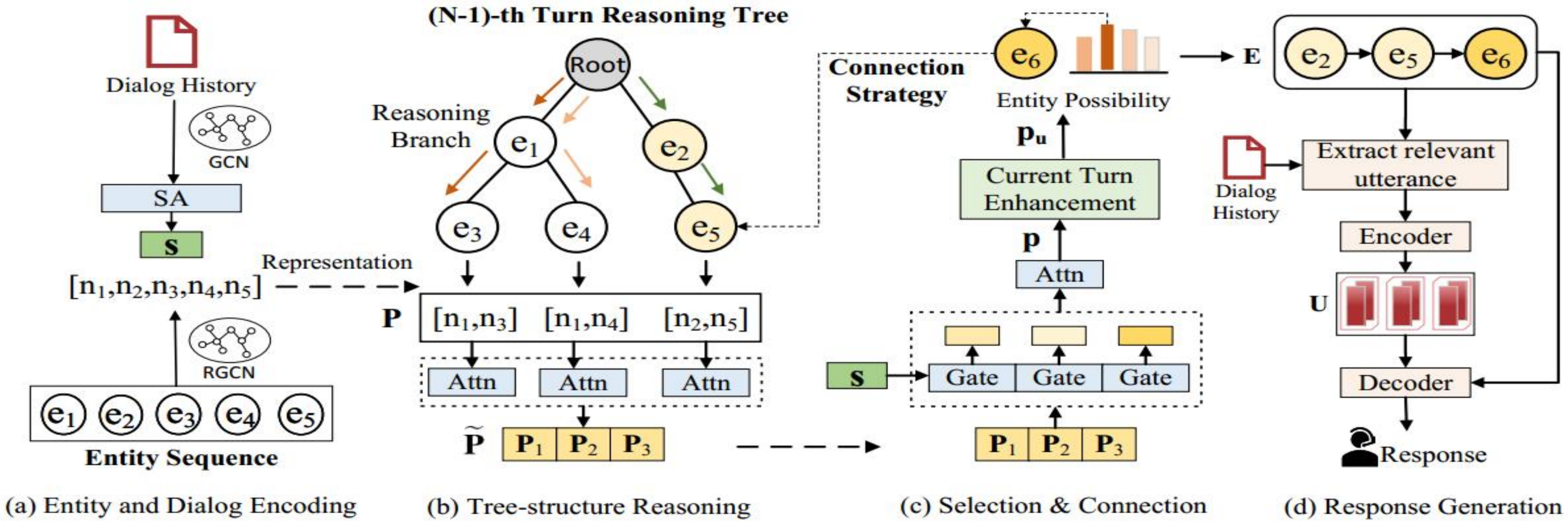
$$\mathbf{n}_e^{l+1} = \sigma\Big(\sum_{r\in\mathcal{R}} \sum_{e'\in\mathcal{N}_e^r} \frac{1}{Z_{e,r}} \mathbf{W}_r^l \mathbf{n}_{e'}^l + \mathbf{W}^l \mathbf{n}_e^l\Big) \quad (1)$$

$$\widetilde{\mathbf{P}} = \mathrm{Attn}(\mathbf{P}) = \mathbf{P}\,\alpha_r$$
$$\alpha_r = \mathrm{Softmax}(\mathbf{b}_r \tanh(\mathbf{W}_r \mathbf{P})) \quad (2)$$

$$\mathbf{p} = \mathrm{Attn}(\gamma \widetilde{\mathbf{P}} + (1-\gamma)\mathbf{s})$$
$$\gamma = \sigma(\mathbf{W}_s \mathrm{Concat}(\widetilde{\mathbf{P}}, \mathbf{s})) \quad (3)$$

$$\mathbf{p}_u = g(\mathbf{p}, g'(\mathrm{Attn}(\mathbf{e}_c), \mathrm{Attn}(\mathbf{s}_c))) \quad (4)$$

Method



$$\mathcal{P}_r^u = \mathrm{Softmax}([\mathbf{p}_u\mathbf{e}_0^{\mathrm{T}}, \cdots, \mathbf{p}_u\mathbf{e}_n^{\mathrm{T}}]) \quad (5)$$

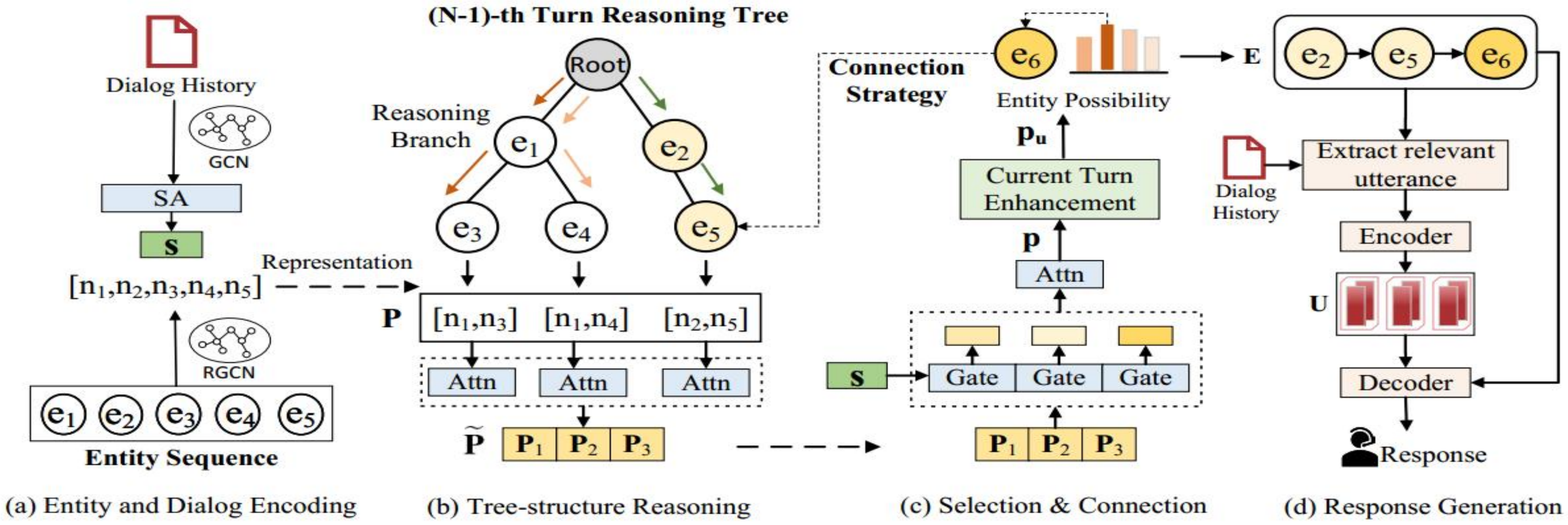
$$\mathbf{R}^l = \mathrm{Decoder}(\mathbf{R}^{l-1}, \mathbf{E}, \mathbf{U}) \quad (6)$$

$$\mathbf{R}^b = \mathrm{FFN}(\mathrm{Concat}(\mathrm{Attn}(\mathbf{E}), \mathbf{R}^l)) \quad (7)$$

$$\mathcal{P}_g = \mathrm{Softmax}(\mathbf{R}^l\mathbf{V}^{\mathrm{T}} + \mathbf{R}^b\mathbf{W}^v) \quad (8)$$

$$\mathcal{L}_I = \sum_{i \neq j} \mathrm{sim}(\widetilde{\mathbf{p}}_i, \widetilde{\mathbf{p}}_j) = \sum_{i \neq j} \frac{\widetilde{\mathbf{p}}_i \widetilde{\mathbf{p}}_j}{|\widetilde{\mathbf{p}}_i| \cdot |\widetilde{\mathbf{p}}_j|} \quad (9)$$

Method



$$\mathcal{L}_a = \lambda_c \mathrm{sim}(\mathbf{p}_c, \mathbf{s}_c) + (1 - \lambda_c)\mathrm{sim}(\mathbf{p}, \mathbf{s}) \quad (10)$$

$$\mathcal{L}_g = -\frac{1}{N}\sum_{t=1}^{N} \log \mathcal{P}_g^t(s_t | s_1, s_2, \ldots, s_{t-1}) \quad (12)$$

$$\mathcal{L}_r = -\sum_{u}\sum_{e_i} \log \mathcal{P}_r^u[e_i] + \lambda_I \mathcal{L}_I + \lambda_a \mathcal{L}_a \quad (11)$$

Experiments

Dataset	ReDial						TG-ReDial					
Method	R@10	R@50	Dist-3	Dist-4	Bleu-2	Bleu-3	R@10	R@50	Dist-3	Dist-4	Bleu-2	Bleu-3
ReDial	0.140	0.320	0.269	0.464	0.022	0.008	0.002	0.013	0.529	0.801	0.041	0.010
KBRD	0.150	0.336	0.288	0.489	0.024	0.009	0.032	0.077	0.691	0.997	0.042	0.012
KGSF	0.183	0.377	0.302	0.518	0.025	0.009	0.030	0.074	1.045	1.579	0.046	0.014
RevCore	0.204	0.392	0.307	0.528	0.025	0.010	0.029	0.075	1.093	1.663	0.047	0.014
CR-Walker	0.187	0.373	0.338	0.557	0.024	0.009	-	-	-	-	-	-
CRFR	0.202	0.399	0.516	0.639	-	-	-	-	-	-	-	-
C^2-CRS	0.208	0.409	0.412	0.622	0.027	0.012	0.032	0.078	1.210	1.691	0.048	0.015
UCCR	0.202	0.408	0.329	0.564	0.026	0.011	0.032	0.075	1.197	1.668	0.049	0.014
TREA	**0.213***	**0.416***	**0.692***	**0.839***	**0.028***	**0.013***	**0.037***	**0.110***	**1.233***	**1.712***	**0.050***	**0.017***

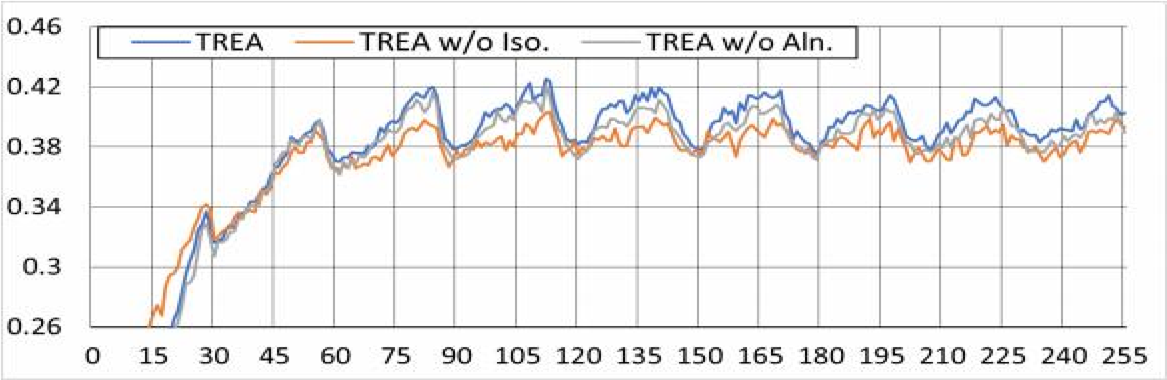
Table 1: Automatic evaluation results on two datasets. Boldface indicates the best results. Significant improvements over best baseline marked with *.(t-test with $p < 0.05$)

Experiments

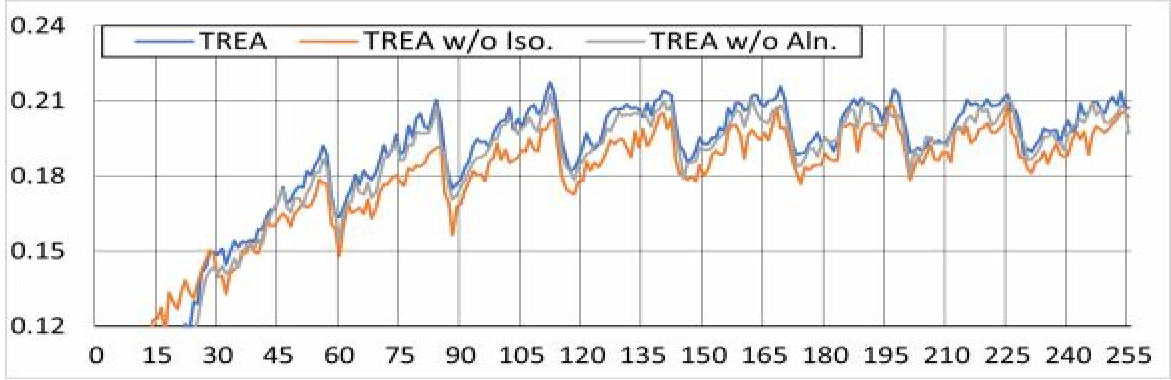
Method	Rel.	Inf.	Flu.	Kappa
RevCore	1.98	2.22	1.53	0.78
CR-Walker	1.79	2.15	1.68	0.77
C^2-CRS	2.02	2.25	1.69	0.66
UCCR	2.01	2.19	1.72	0.72
TREA	**2.43**	**2.26**	**1.75**	0.75

Table 2: Human evaluation results on the conversation task. Rel., Inf. and Flu. stand for Relevance, Informativeness and Fluency respectively. Boldface indicates the best results (t-test with $p < 0.05$).

Experiments



(a) Recall@50



(b) Recall@10

Figure 3: Performance comparison of TREA and its two variants. One step (X-axis) denotes parameter updates for 20 batches of training data.

Experiments

Dataset	ReDial		TG-ReDial	
Method	R@10	R@50	R@10	R@50
TREA	0.214	0.418	0.037	0.110
TREA w/o Iso.	0.202	0.405	0.028	0.079
TREA w/o Aln.	0.209	0.412	0.035	0.103
TREA w/o IA.	0.201	0.403	0.026	0.076

Table 3: Ablation results on the recommendation task. (t-test with $p < 0.05$)

Experiments

Figure 4: 2D projection of KG embeddings trained by TREA (the above) and TREA w/o Iso. (the below) to illustrate the impact of the isolation loss $\mathcal{L}_I$. Embeddings are projected through t-SNE with Perplexity set to 10 and the Iterations set to 13.)

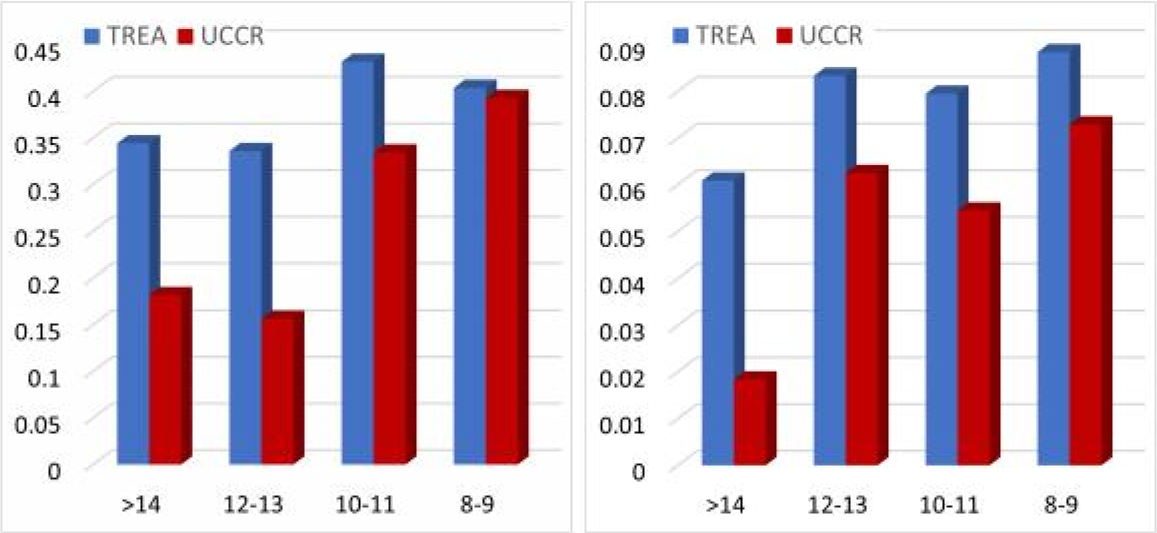
Experiments

Model	Dist-4	Bleu-3	PPL(↓)	Rel.
TREA	0.839	0.013	4.49	2.43
TREA w/o Ent.	0.799	0.012	4.56	2.28
TREA w/o Utt.	0.764	0.011	4.61	2.13
TREA w/o EU.	0.789	0.011	4.78	2.10

Table 4: Evaluation results on the ablation study of the generation task. Fleiss's kappa values of Rel. all exceed 0.65.

Experiments

Table 4: Evaluation results on the ablation study of the generation task. Fleiss's kappa values of Rel. all exceed 0.65.



(a) ReDial (b) TG-ReDial

Figure 5: Evaluation results (R@50) of TREA and UCCR on data of different converstaion rounds.

Thanks!