MISC: A MIxed Strategy-Aware Model Integrating COMET for Emotional Support Conversation

Quan Tu[1*†], **Yanran Li**[2*], **Jianwei Cui**[2], **Bin Wang**[2], **Ji-Rong Wen**[1,3] and **Rui Yan**[1,3‡]
[1]Gaoling School of Artificial Intelligence, Renmin University of China
[2]Xiaomi AI Lab
[3]Beijing Academy of Artificial Intelligence
[1]{quantu,jrwen,ruiyan}@ruc.edu.cn
[2]{liyanran,cuijianwei,wangbin11}@xiaomi.com

Code： https://github.com/morecry/MISC.

（ACL-2022)

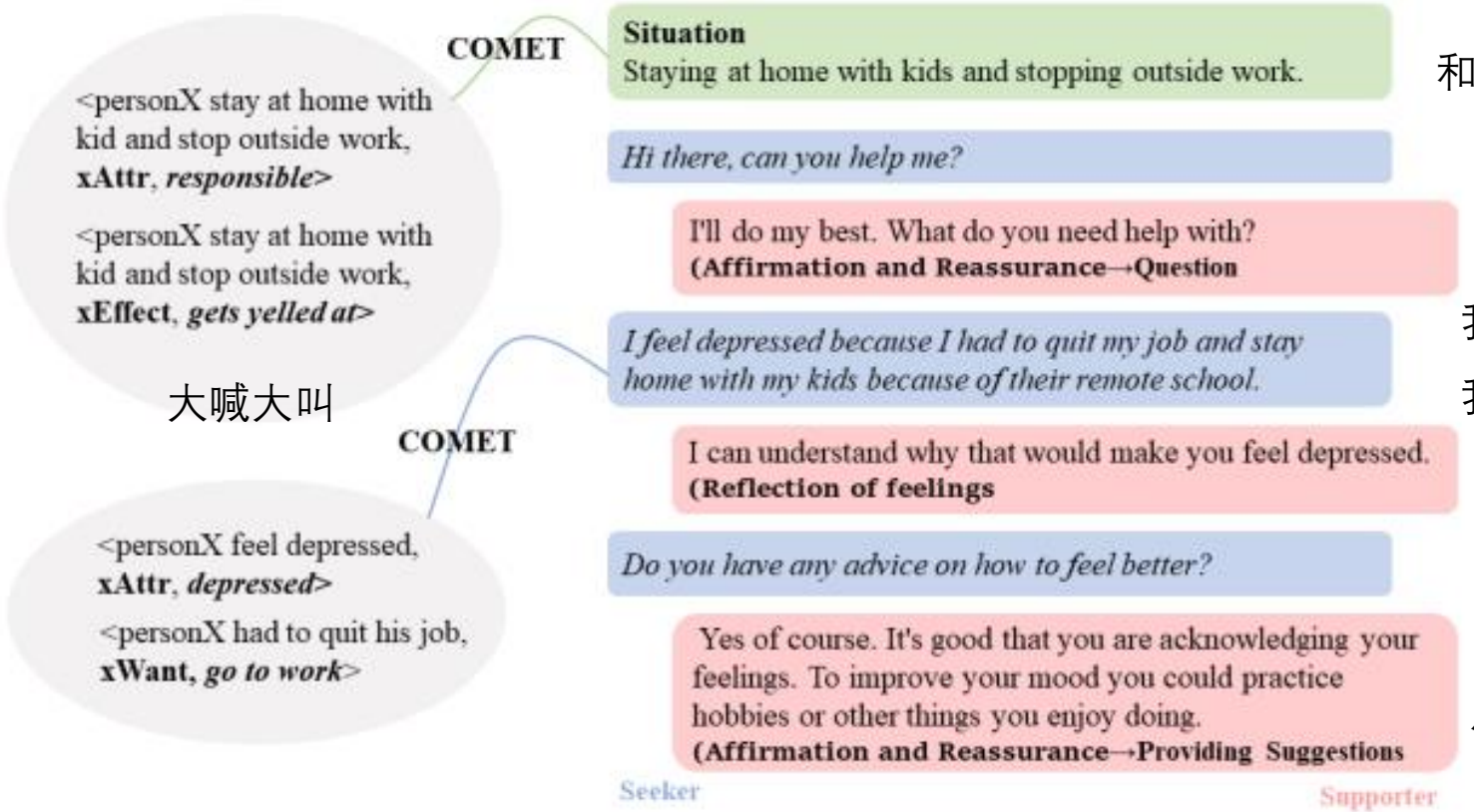
1. Introduction

2. Approach

3. Experiments

xAttr 鉴于人物 X 在事件中的角色，如何描述他们。
xEffect 事件对人 X 的影响。
xWant 事件结束后 X 可能想要做什么。

Introduction



和孩子呆在家里，停止户外工作。

我感到很沮丧，因为我不得不辞掉工作呆在家里陪我的孩子，因为他们的学校很偏远。

当然可以。你承认自己的感受是件好事。为了改善你的情绪，你可以练习爱好或其他你喜欢做的事情。

Figure 1: An Emotional Support Conversation Example.

问题 询问与问题相关的信息，以帮助寻求帮助的人阐明他们面临的问题。
自我表露 透露你有过的类似经历或与求助者分享的情绪，以表达你的同理心。
肯定和安慰 肯定求助者的优势、动机和能力，并提供安慰和鼓励。
提供建议 提供有关如何改变的建议，但要小心不要越权并告诉他们该怎么做。

Approach

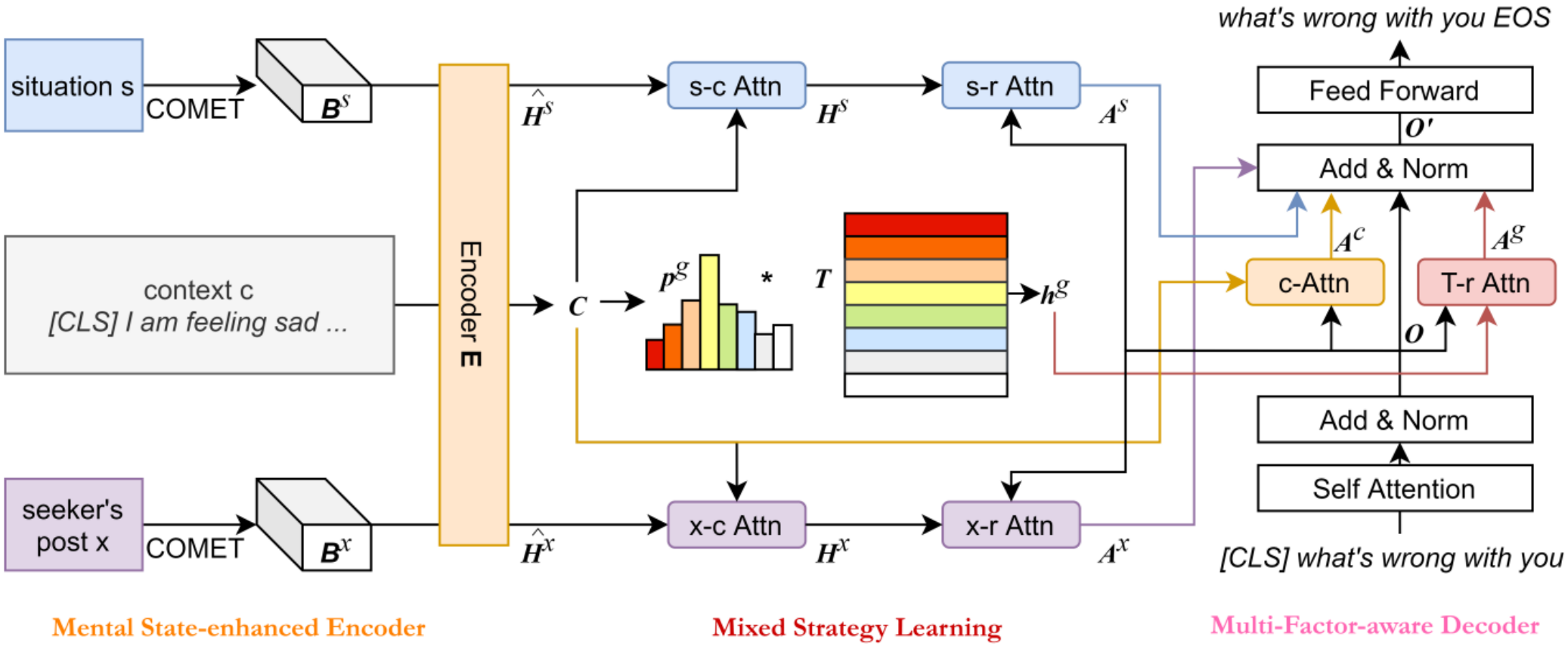
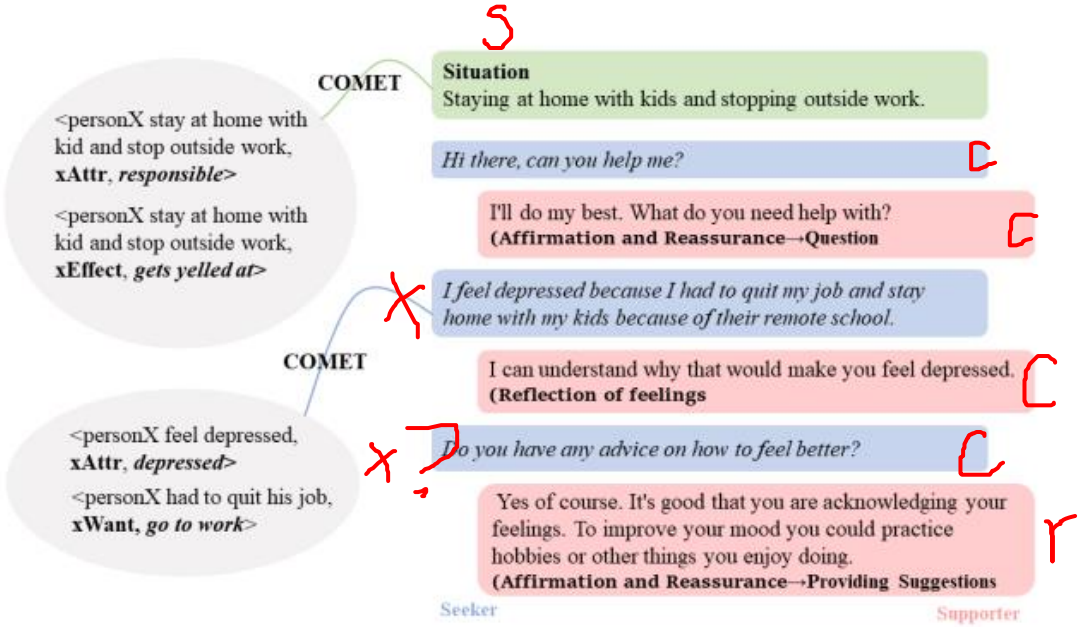
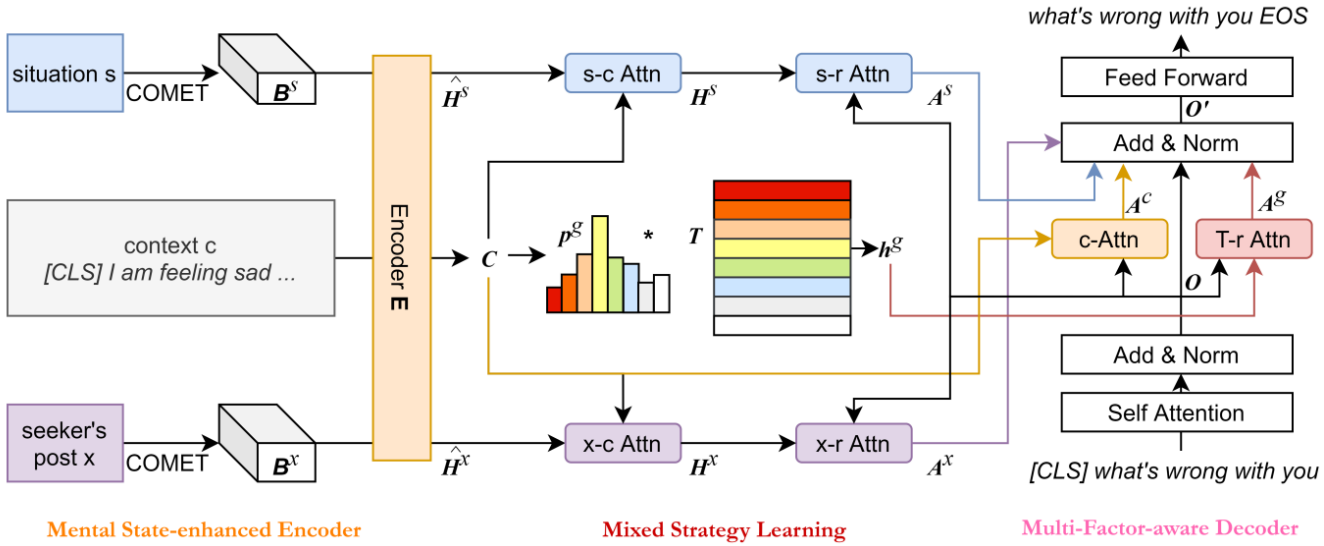


Figure 2: The overview of the proposed MISC which consists of a mental state-enhanced encoder, a mixed strategy learning module, and a multi-factor-aware decoder.



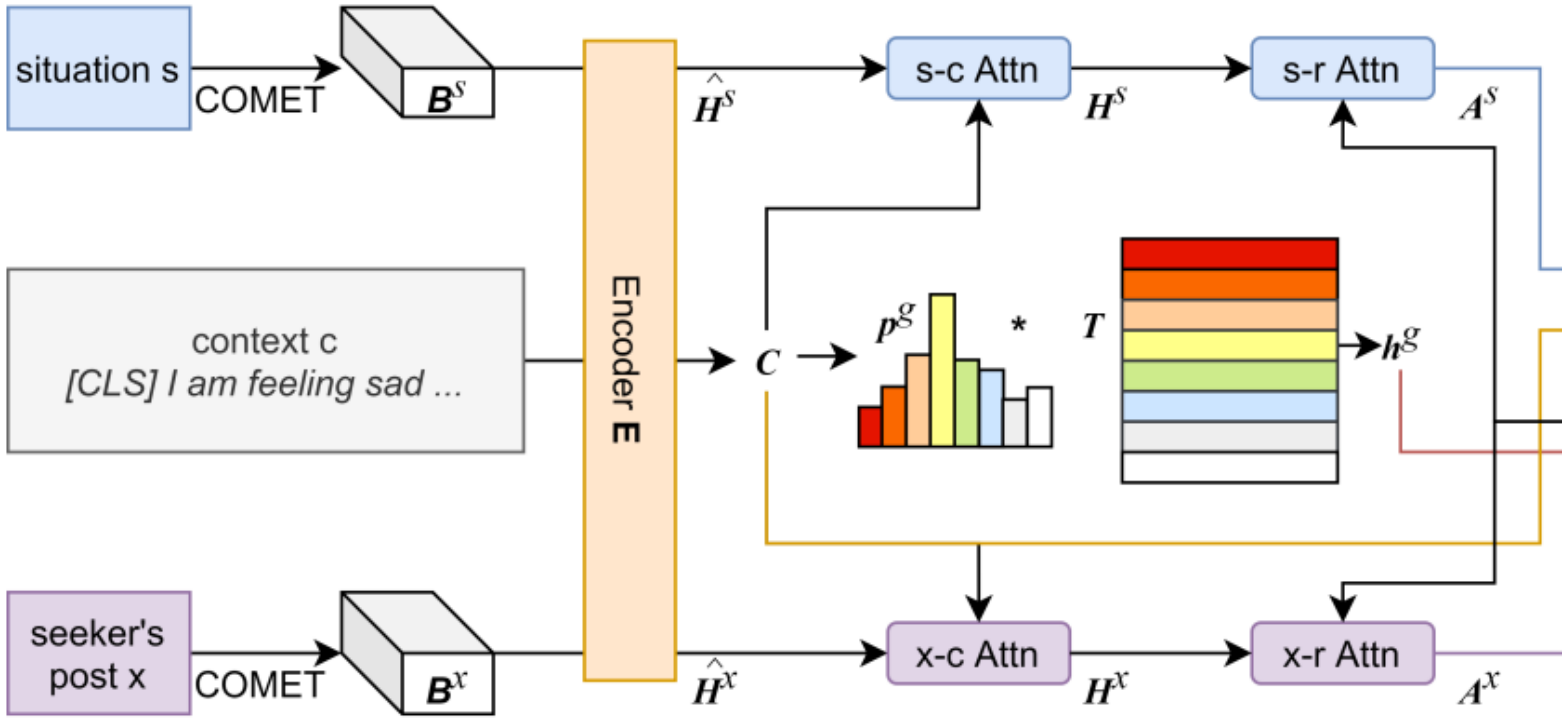


Problem Formulation

the target is to estimate the probability distribution $p(\boldsymbol{r}|\boldsymbol{c})$ of the dataset $\mathcal{D} = \{\boldsymbol{c}^{(i)}, \boldsymbol{r}^{(i)}\}_{i=1}^{N}$ where $\boldsymbol{c}^{(i)} = (\boldsymbol{u}_1^{(i)}, \boldsymbol{u}_2^{(i)}, ..., \boldsymbol{u}_{n_i}^{(i)})$

the seeker's last post (ut- terance) as $\boldsymbol{x}$. the probability distribution $p(\boldsymbol{r}|\boldsymbol{c}, \boldsymbol{s}, \boldsymbol{x})$.

Approach



Mental State-enhanced Encoder

Mixed Strategy Learning

$$\hat{\boldsymbol{H}}^s = [\boldsymbol{h}^s_{1,1}, \boldsymbol{h}^s_{2,1}, ..., \boldsymbol{h}^s_{N_{st},1}]$$
$$\boldsymbol{h}^s_j = \mathrm{E}(\boldsymbol{b}^s_j) \quad (3)$$

$$\boldsymbol{Z} = \mathtt{softmax}(\hat{\boldsymbol{H}}^s \cdot \boldsymbol{C}^{\mathrm{T}}) \cdot \boldsymbol{C}$$
$$\boldsymbol{H}^s = \mathrm{LN}(\hat{\boldsymbol{H}}^s + \boldsymbol{Z}) \quad (4)$$

Mental State-Enhanced Encoder

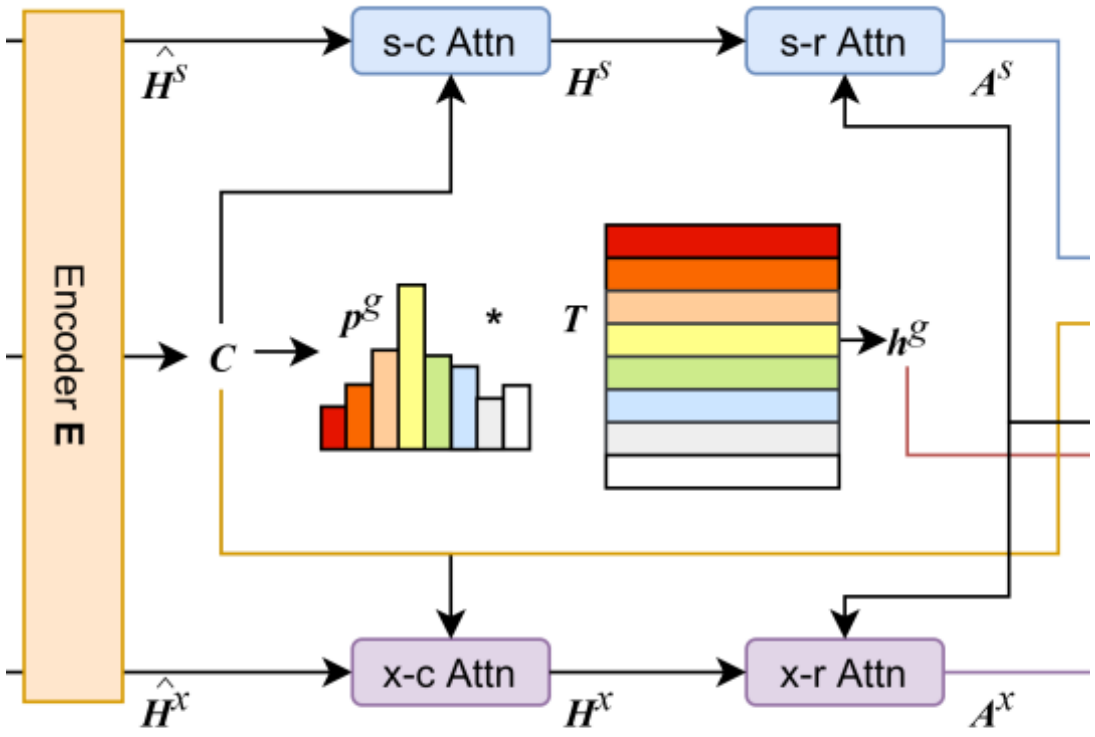
$$\boldsymbol{C} = \mathrm{E}(\mathtt{CLS}, \boldsymbol{u}_1, \mathtt{EOS}, \boldsymbol{u}_2, ..., \boldsymbol{u}_{n_i}) \quad (1)$$

$$\boldsymbol{B}^s = \bigcup_{j=1}^{N_r} \mathtt{COMET}(rel_j, \boldsymbol{s}) \quad (2)$$

where N_r is the number of pre-defined relations in COMET, and rel_j stands for the j-th specific relation, such as `xAttr` and `xReact`.[1] Note that given a certain event-relation pair, COMET is able to generate multiple "tails" of free-form mental state information, $\boldsymbol{B}^s$ is a set of N_s mental state blocks, i.e., $\boldsymbol{B}^s = \{\boldsymbol{b}^s_j\}^{N_s}_{j=1}$. Similarly, we can obtain the set of mental state blocks $\boldsymbol{B}^x$ using the seeker's last post $\boldsymbol{x}$.

xAttr 鉴于人物 X 在事件中的角色，如何描述他们。

xReact 人物X 对事件的反应。



Mixed Strategy Learning

Mixed Strategy Learning Module

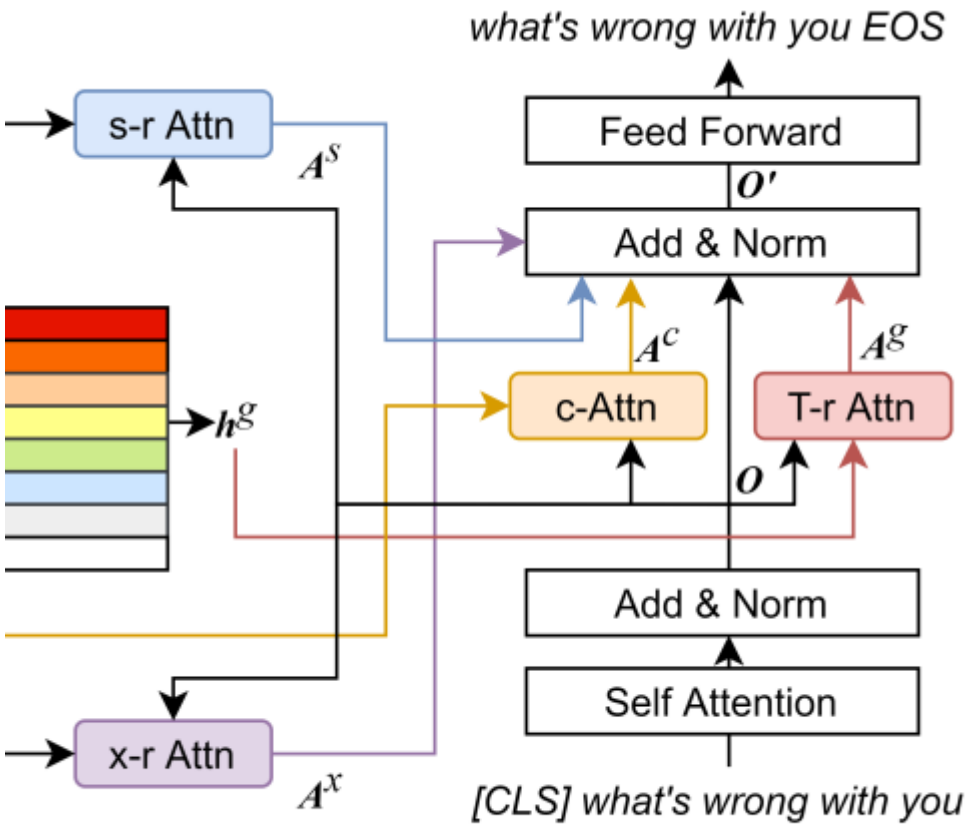
$$\boldsymbol{p}^g = \texttt{MLP}(\boldsymbol{C}_1) \tag{5}$$

where MLP is a multi-layer perceptron, and $\boldsymbol{p}^g$ records the probabilities of each strategy to be used.

$$\boldsymbol{h}^g = \boldsymbol{p}^g \cdot \boldsymbol{T} \tag{6}$$

response generation. Here, we masterly learn from the idea of VQ-VAE's codebook to represent strategy(Oord et al., 2017). The strategy codebook $\boldsymbol{T} \in \mathbb{R}^{m \times d}$ represent m strategy latent vectors (here $m = 8$) with the dimension size d. By weighting $\boldsymbol{T}$ using $\boldsymbol{p}^g$, we are able to obtain a comprehensive strategy representation $\boldsymbol{h}^g$

Approach



Learning

Multi-Factor-aware Decoder

Multi-Factor-Aware Decoder

$$
\begin{aligned}
\boldsymbol{A}^c &= \texttt{CROSS-ATT}(\boldsymbol{O}, \boldsymbol{H}) \\
\boldsymbol{A}^s &= \texttt{CROSS-ATT}(\boldsymbol{O}, \boldsymbol{H}^s) \\
\boldsymbol{A}^x &= \texttt{CROSS-ATT}(\boldsymbol{O}, \boldsymbol{H}^x) \\
\boldsymbol{A}^g &= \texttt{CROSS-ATT}(\boldsymbol{O}, \boldsymbol{h}^g) \\
\boldsymbol{O}' &= \texttt{LN}(\boldsymbol{A}^c + \boldsymbol{A}^s + \boldsymbol{A}^x + \boldsymbol{A}^g + \boldsymbol{O})
\end{aligned} \tag{7}
$$

$$
\begin{aligned}
\mathcal{L}_r &= -\sum_{t=1}^{n_r} \texttt{log}(p(r_t|\boldsymbol{r}_{j<t}, \boldsymbol{c}, \boldsymbol{s}, \boldsymbol{x})) \\
\mathcal{L}_g &= -\texttt{log}(p(g|\boldsymbol{c}, \boldsymbol{s}, \boldsymbol{x})) \\
\mathcal{L} &= \mathcal{L}_r + \mathcal{L}_g
\end{aligned} \tag{8}
$$

where n_r is the length of response, g is the true strategy label, $\mathcal{L}_g$ is the loss of predicting strategy, $\mathcal{L}_r$ is the loss of predicting response, and $\mathcal{L}$ is combined objective to minimize.

Experiments

Category	Train	Dev	Test
# dialogues	14117	1764	1764
Avg. # words per utterance	17.25	17.09	17.11
Avg. # turns per dialogue	7.61	7.58	7.49
Avg. # words per dialogue	148.46	146.66	145.17

Table 1: The statistics of processed ESConv dataset.

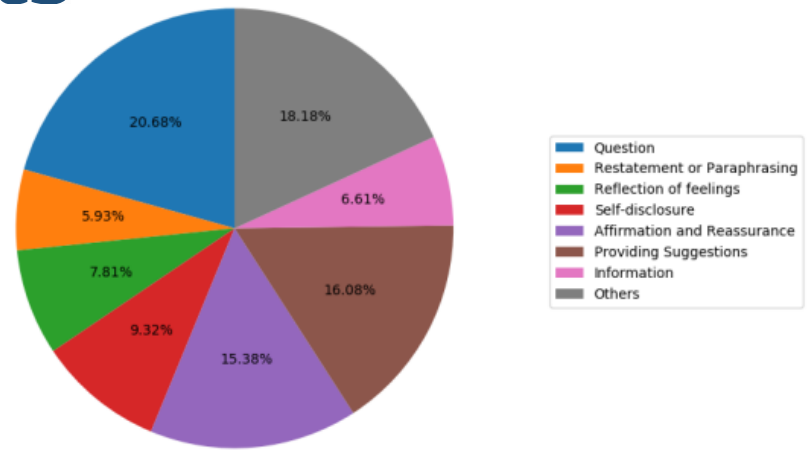


Figure 6: The strategy distribution in the original ESConv dataset.

Model	ACC(%)↑	PPL↓	D-1↑	D-2↑	B-2↑	B-4↑	R-L↑	M(%)↑
Transformer	-	89.61	1.29	6.91	6.53	1.37	15.17	10.33
MT Transformer	-	89.52	1.28	7.12	6.58	1.47	14.75	10.27
MoEL	-	133.13	2.33	15.26	5.93	1.22	14.65	9.75
MIME	-	47.51	2.11	10.94	5.23	1.17	14.74	9.49
BlenderBot-Joint	28.57	18.49	4.12	17.72	5.78	1.74	16.39	9.93
MISC	**31.63**	**16.16**	**4.41**	**19.71**	**7.31**	**2.20**	**17.91**	**11.05**

Table 2: Automatic Evaluation Results on ESConv.

Experiments

Model	Flu.	Know.	Emp.
Transformer	0.62	0.31	0.29
MT Transformer	0.78	0.34	0.82
MoEL	0.36	0.80	0.33
MIME	1.13	0.27	0.35
BlenderBot-Joint	**1.87**	0.74	1.21
MISC	1.84	**1.06**	**1.44**

Table 3: Manual Evaluation Results. The Fleiss Kappa score (Fleiss and Cohen, 1973) reaches 0.445, indicating a moderate level of agreements.

Model	D-1 ↑	B-2 ↑	R-L ↑	M(%) ↑
MISC	**4.41**	**7.31**	**17.91**	**11.05**
w/o g	3.85	7.09	16.75	9.85
w/o s	4.39	6.35	17.05	10.06
w/o x	4.27	6.49	17.03	10.09

Table 4: Evaluation Results of Ablation Study.

Experiments

Situation	
Seeker	My boyfriend and I recently broke up due to long-distance relationship and the impact COVID-19...
Context	
Seeker	Hi, I'm hoping you can help me get through this difficult time.
Supporter	Hi, can you please explain the difficult time you are having?
Seeker	My boyfriend broke up with me and I've been upset. We live in different parts of the country and...
	.....
Seeker	Have you ever gone through a break up?
Response	
Transformer	I am so sorry that you are not able to get your friends and family. (*Inconsistency*)
MT Transformer	I have had a similar situation and I understand. It was the only way I was in similar situation, I have found that way to get out of my life. I have ... and i was in a similar situation. (*Repetition*)
MIME	I know it is hard, but it is always good to do things out for yourself. (*Poor Empathy*)
MoEL	I am so sorry that you are feeling better. I had a relationship with your boyfriend! (*Aggressiveness*)
BlenderBot-Joint	I have not but i have been through a lot of breakups... (*Contradiction*)
MISC	I have been through a breakup and I can tell you that it was very painful. I have had to move on from a relationship because of it. (*Strong Empathy*)
Ground-Truth	It sounds like you are really stressed over the breakup and you want to know how to move forward.

Table 5: Responses generated from MISC and other compared models. Some words are omitted due to space limit.

Experiments

Component	D-1↑	B-2↑	R-L↑	M(%)↑
MISC	**4.41**	**7.31**	**17.91**	**11.05**
MISE	3.94	7.09	16.93	10.53

Table 6: Results of MISC with Different Emotions.

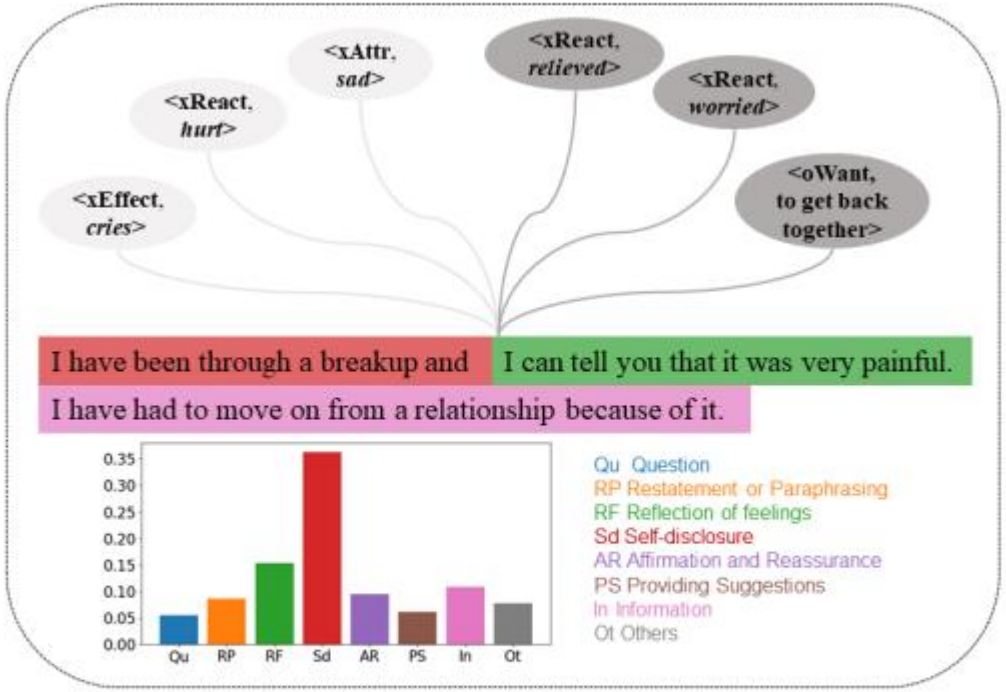


Figure 4: The visualization of how the MISC organizes the response under the effect of multiple factors.

Experiments

Strategy	D-1↑	B-2↑	R-L↑	M(%)↑
Mixture	4.41	**7.31**	**17.91**	**11.05**
Single	**4.79**	6.30	17.01	10.22

Table 7: Comparison of different strategy modeling.

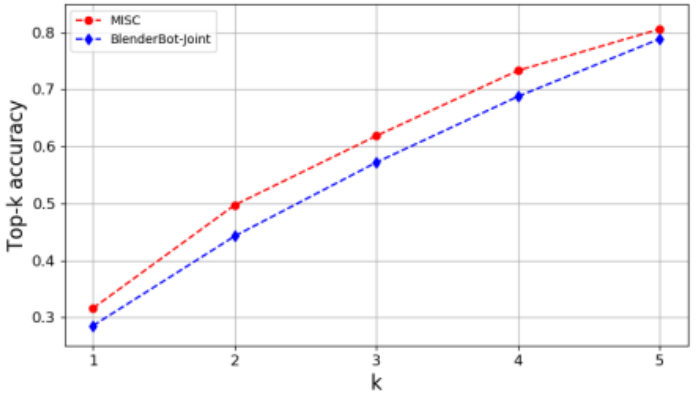
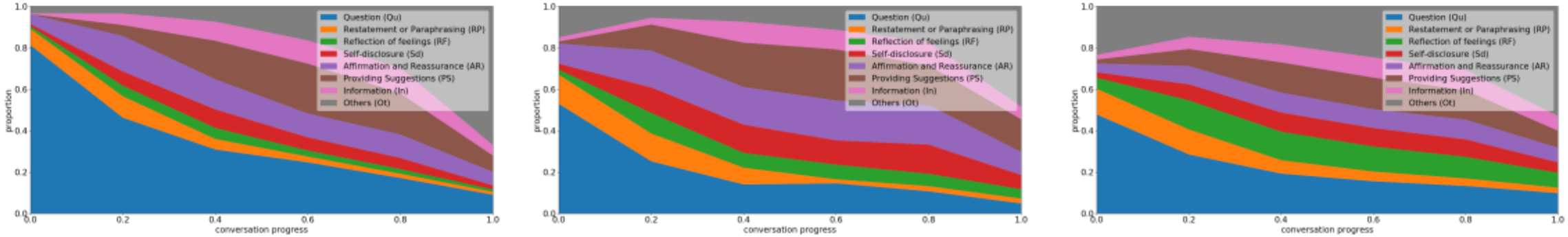


Figure 5: The Top-k Strategy Prediction Accuracy.



(a) predicted by the MISC. (b) from the test set. (c) predicted by the BlenderBot-Joint.

Figure 3: The strategy distribution in the different stage of conversation.

Thank you !