ECAI 2023
K. Gal et al. (Eds.)

doi:10.3233/FAIA230634

Grafting Fine-Tuning and Reinforcement Learning for Empathetic Emotion Elicitation in Dialog Generation

Ying Zhu[1], Bo Wang[1;*], Dongming Zhao[2], Kun Huang[2], Zhuoxuan Jiang[3;4], Ruifang He[1] and Yuexian Hou[1]

[1] College of Intelligence and Computing, Tianjin University, Tianjin, China
[2]AI Lab, China Mobile Communication Group Tianjin Co., Ltd
[3]Faculty of Business Information, Shanghai Business School, Shanghai, China
[4]Learnable.ai Inc.

Abstract. For human-like dialogue systems, it is significant to inject the empathetic ability or elicit the opposite's positive emotions, while existing studies mostly only focus on either of the above two research lines. In this work, we propose a novel and grafted task named **Empathetic Emotion Elicitation Dialog** to make a dialog system able to possess both aspects of ability simultaneously. We do not train an empathetic dialog system and an emotion elicitation dialog system separately and then simply concatenate the responses generated by these two systems, which will cause illogical and repetitive responses. Instead, we propose a unified solution: (1) To generate empathetic responses and emotion elicitation responses within the same semantic space, we design a unified framework. (2) The unified framework has three stages which first retrieve the empathetic and emotion elicitation exemplars as external knowledge, then fine-tune the emotion/action prediction on a pre-trained language model to enhance the empathetic ability, and finally model the user feedback by reinforcement learning to enhance the emotion elicitation ability. Experiments show that our method outperforms the baselines in the response generation quality and simultaneously empathizes with the user and elicits their positive emotions.

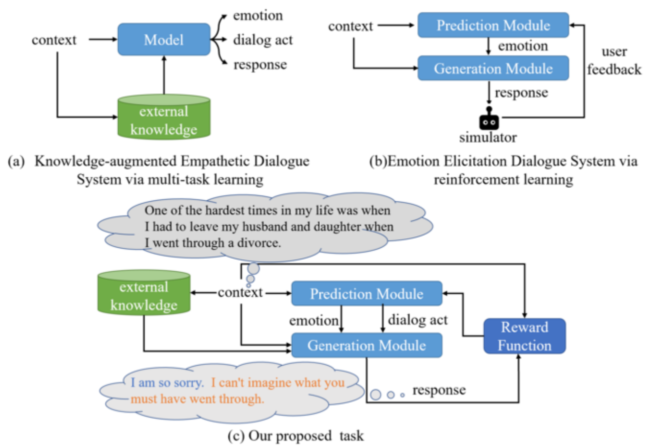

Figure 1. Comparison of our proposed Empathetic Emotion Elicitation dialog generation task with existing schemes. The response generated by our proposed task comprises a blue empathy-generic utterance and an orange empathy-special utterance.

1 Introduction

Integrating human emotions into dialogue systems is one of the essential aspects of a human-like dialog agent, which is still a challenging task. Among the various kinds of scenarios for building emotion-aware dialog systems, e.g., pre/after-sales service [25] and psychological treatment [33], injecting empathetic ability or eliciting positive emotion is currently two promising directions that attract widespread attention by the community [21, 28].

To inject the empathetic ability into a dialog system (i.e., ***empathetic*** dialog system [28]), both the *affection* and *cognition* should be involved [3]. Therefore, as shown in Figure 1(a), current systems either learn to understand the opposite user's emotion first and then to predict a close emotion and a dialog act (DA) that the system should conduct in the next dialog turn [1, 20, 34]. External knowledge bases are also introduced for reasoning to generate more informative and empathetic responses [15, 19, 30].

Besides empathizing with users, some approaches aim to elicit the user's positive emotions (i.e., ***emotion elicitation*** dialog system [21]), especially under some specific situations that need emotional intervention, e.g., depression treatment and complaint handling [6, 35]. This scenario requires the systems to take the user feedback into account dynamically at every dialog turn and then generate an encouraging response when a negative emotion is identified from the user side. Usually, a reinforcement learning method is leveraged to model the user feedback information [9, 32], as the schematic process shown in Figure 1(b).

However, our observation shows that, in real life, a human-like emotional response is usually constituted by both the empathetic part and the emotion elicitation part. For example, as shown in figure 1(c), when the user talks about her hardest time, a human-like emotional response should consist of an empathy-generic response (*"I am so sorry"*) to first empathize with the user and an empathy-special utterance response specific to context (*"I can't imagine what you must have gone through"*) to further empathize with the user and want to elicit positive emotion from the user.

Therefore, in this paper, we investigate a novel and more challenging task, in which the dialog system is expected to simultaneously possess both the empathetic and emotion elicitation abilities. The novel task is named with ***Empathetic Emotion Elicitation*** dialog and is shown in Figure 1(c). To our knowledge, this task has not been fully studied towards more anthropomorphic dialog systems,

* Corresponding Author. Email: bo_wang@tju.edu.cn

and there is no model yet that can effectively generate responses that are both empathetic and elicit positive emotions from the user.

This novel task has two new challenges. (1) We should model the emotion/DA prediction for empathy, and dynamically capture the user feedback to build the emotion elicitation ability. (2) Although a human-like emotional response comprises the empathetic part and the emotion elicitation part if we only train an empathetic dialog system and an emotion elicitation dialog system separately and simply concatenate the responses generated by these two systems, it will cause many issues, such as poor logic and a large number of repetitive parts, due the two responses are not in the same semantic space.

To address the above challenges, we adopt a grafting method that draws on the merits from both sides and proposes a unified framework: (1) To generate empathetic responses and emotion elicitation responses within the same semantic space, we design a unified framework. (2) The unified framework is designed in three stages. In the first stage, we retrieve the empathetic and emotion elicitation exemplars as external knowledge. (3) In the second stage, we fine-tune the emotion/DA prediction tasks on a pre-trained generative language model to enhance the empathetic ability. (4) In the third stage, we model the user feedback information by reinforcement learning to enhance the emotion elicitation ability. Figure 1(c) indicates the schematic process of the method.

More technically and specifically, (1) We consider two kinds of exemplars [2], which are the template responses retrieved from external knowledge bases that can guide the generation with stylistic and thematic cues (Figure 3 shows some examples) with two separate retrieval methods. We use the Dense Passage Retrieval model (DPR) [10] to retrieve a set of semantically similar responses to the input context from the training set as the empathy-special exemplars. Also, we use semantic search to retrieve generic responses that elicit users' positive emotions from the Distress Management Conversations Knowledge Graph (HEAL) [35], which contains user feedback information to form the empathy-generic exemplars. This way, the empathetic and emotion elicitation exemplars can provide the following step with sufficient implicit knowledge. (2) With the two kinds of exemplars as partial input, we fine-tune the emotion/DA prediction tasks on a pre-trained T5 model [27] to enhance the empathetic ability. The task can train a model to generate an empathy-generic utterance and an empathy-special utterance simultaneously. (3) To enhance the emotion elicitation ability of our model, we are inspired by those dialog systems via reinforcement learning and adopt a Proximal Policy Optimization method [31]. We hope the joint training by grafting fine-tuning and reinforcement learning can better equip the model with empathy and emotion elicitation abilities.

We conduct our experiments on a large-scale public dataset [28], and the results demonstrate that our method is effective on all automatic and human evaluation metrics. Overall, the main contributions of this paper are summarized as follows:

- We propose a new task of the Empathetic Emotion Elicitation dialog system that possesses both empathetic and emotion elicitation abilities and design a unified framework.
- The unified framework is designed in three stages: exemplar retrieving, multi-task fine-tuning with retrieved exemplars for empathy enhancement, and grafted reinforcement learning for emotion elicitation enhancement.
- Through automatic and human evaluation, we confirmed that our model exceeded SOTA methods not only in the general quality of responses but also in empathizing with the user and eliciting their positive emotions.

2 Related Work

Empathetic Dialogue Generation. Integrating human emotions into a dialogue system is essential in an open-domain dialogue system. The early work in this field is to control the emotions in the generated responses, and the emotions need to be identified [36, 37]. However, in real life, the emotions will change with dialogue and will not be specified beforehand. One party in the conversation needs to understand and perceive the other party's feelings and respond appropriately. Therefore, the dialogue system needs to express the specific emotion in the generated response and be able to empathize, that is, to select an appropriate emotion by understanding the user's emotion to generate the response. With the release of an empathetic dialogue corpus named EmpatheticDialogues [28], many studies have been proposed to generate empathetic responses, such as the mixture of emotional experts [18], emotion grouping and mimicry [23]. However, empathy includes both affective and cognitive empathy [3]. So, some studies have focused on cognitive empathy recently, such as exploring the emotional cause of the context [11, 17, 26], predicting dialog act in the conversation [1, 34], introducing external knowledge bases for reasoning to help generate more informative and empathetic responses [15, 19, 30]. However, in current empathetic dialogue systems, most of the work uses the dialogue context to predict the emotion and the dialog act in the next dialog turn or introduces external knowledge bases for reasoning without dynamically taking the user feedback into account.

Emotion Elicitation Dialogue Generation. The dialogue system must consider the user's feedback and improve their emotions during the interaction. [7] is the first time to explore how users' emotions were affected by what others say. [21] builds a dialogue system that can generate more natural responses to promote users' positive emotions. [16] proposes a variational model named EmoElicitor to generate responses that can elicit users' specific emotions with the help of a pre-trained language model. Both [9] and [32] utilize the reward function of reinforcement learning to reward the action that improves users' emotional state. The difference is that [9] 's action space is the probability distribution of the generated response, while [32] 's action space is the emotion type of generated response. However, these models are limited to predicting the emotion in the next dialog turn and generating an empathetic response specific to the emotion, which needs more in empathetic dialogue systems. By contrast, our model not only predicts the emotion and the dialog act in the next dialog turn but also introduces two kinds of exemplars to help generate an empathetic response that can elicit positive emotion from the user.

3 Problem Definition

Our task is to train a dialog system with empathy and emotion elicitation abilities with three stages: exemplar retrieving, multi-task fine-tuning with retrieved exemplars, and grafted reinforcement learning.

More formally, $C = \{s_0, l_0, s_1, l_1, \ldots, s_n\}$ is the current dialogue context where s_* and l_* represent utterances from the speaker and the listener. The user plays the role of the speaker, while the dialog system plays the role of the listener. In the first stage, we retrieve a set of exemplars $T = \{T_1, T_2\}$ composed of the empathy-special exemplars T_1 and the empathy-generic exemplars T_2. In the second stage, given C, we aim to understand the user's emotion e_u and then to predict an emotion e_s and a dialog act d_s in the next turn, and use T to help guide the generation of the response r specific to e_s and d_s. In the third stage, we continue to train the model via reinforcement learning by using the user's emotional feedback f on the responses generated in the second stage.

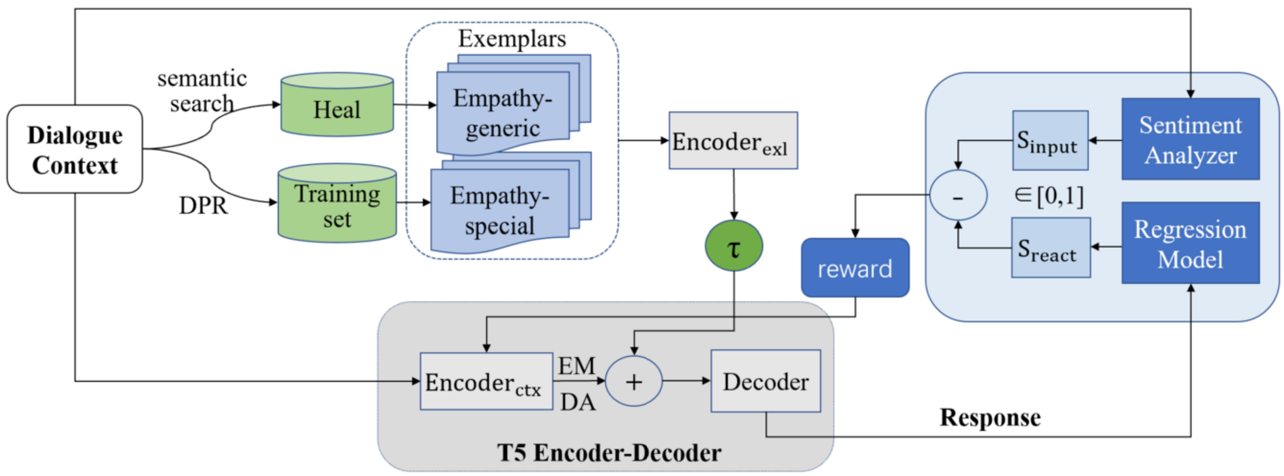

Figure 2. The unified framework has three stages: (1)exemplar retrieval;(2)multi-task fine-tuning with retrieved exemplars;(3)grafted reinforcement learning.

4 Methodology

4.1 Overview

The overview architecture of our proposed framework is illustrated in Figure 2. We divide the task into three progressive stages: Exemplar Retrieval, Multi-task Fine-tuning, and Grafted Reinforcement Learning. Specifically, in the first stage, we use context to retrieve two kinds of exemplars from two separate knowledge bases. In the second stage, we fine-tune the emotion/dialog act prediction tasks on a pre-trained generative language model for empathy. In the third stage, we adopt joint training by fine-tuning and reinforcement learning to maximize the user's emotional experience and better equip the model with empathetic and emotion-elicitation abilities.

4.2 Exemplar Retrieval

Empathy-special Exemplar. We use Dense Passage Retrieval(DPR) to obtain empathetic responses from the training set based on the specific context [10]. DPR is a dense embedding retrieval model pre-trained on Wikipedia dump, and we use the pre-trained DPR model[1] fine-tuned on the training set of EmpatheticDialogues corpus [28] as same as in [22]. Given a dialogue context, DPR can retrieve a set of possible responses from the training set as the empathy-special exemplars. These empathy-special exemplars can answer situations similar to the current context, so they are specific to the context. Given the context C as a query q and each candidate response from the training set as candidate input p, the DPR model calculates the similarity between the query and candidate input using the dot product of their embeddings. The distribution is calculated as follows:

$$sim(q,p) = E_Q(q)^\top E_P(p) \tag{1}$$

Where $E_Q(\cdot)$ and $E_P(\cdot)$ are the encoders of query and candidate input, respectively. Finally, we select the top-k responses with the highest similarity as the empathy-special exemplars T_1.

Empathy-generic Exemplar. In addition to context-specific responses, we find that some generic responses such as *"I'm sorry to hear that"* can also appease users. So our model introduces a knowledge graph HEAL[2] [35] involving user feedback for distress management conversations. It consists of five types of nodes: **(1) stressors:** the cause of distress, such as suicidal ideation; **(2) expectations:** questions asked by the speakers usually; **(3) response types:** the most common types of responses given by the listeners with different stressors, such as *"I understand how you feel."*; **(4) feedback types:** the most frequent types of feedback provided by the speakers after a response, such as *"Thanks for those words, means a lot."*; **(5) affective states:** emotional states associated with each node which can be used to determine whether the emotion state of user feedback is positive. For example, we define that *grateful* is positive, *lonely* is negative. By clustering, 47109 stressors are divided into 4363 stressor clusters, each with its topic sentence, corresponding response types, and feedback types.

For the empathy-generic exemplar, we do not use the DPR model fine-tuned on the training set of EmpatheticDialogues. To get the sentence embedding, we use sentenceBert [29] to encode context C and stressor cluster topic sentence in HEAL. Then we use the similarity between embeddings to complete the semantic search and find the top k clusters. At last, we query the responses under this cluster and filter out the responses that elicit positive feedback from users to form the empathy-generic exemplars T_2.

4.3 Emotion and Dialog Act Annotation

Emotion(EM)[3] We divide the 27 emotions and one neutral emotion proposed in [4] into six basic emotion categories (*anger, disgust, fear, happiness, sadness, surprise*) [5] and one neutral emotion. Because some of the 27 emotions are difficult to distinguish, it may lead to a decrease in the accuracy of classification.

Dialog Act(DA)[4]**.** We divide dialog act into eight categories (*questioning, acknowledging, consoling, agreeing, encouraging, sympathizing, suggesting, wishing*) and one category of others, which is proposed in [34].

Classifier and Annotation. Due to the absence of emotion and

[1] github.com/declare-lab/exemplary-empathy

[2] github.com/anuradha1992/HEAL

[3] github.com/google-research/google-research/tree/master/goemotions

[4] github.com/anuradha1992/EmpatheticIntents

dialog act labels for each utterance in the EmpatheticDialogues corpus [28], we leverage several external datasets introduced above. We fine-tuned the T5-Encoder [27] Model for EM and DA. The emotion classifier achieved an accuracy of 70% on the test set, and the DA classifier achieved an accuracy of 97%, ensuring the quality of automatic annotation. Then the EM and the DA classifier annotate an emotion label and a dialog act label for each utterance in the EmpatheticDialogues corpus.

4.4 Multi-task Fine-tuning

We use text-to-text transfer transformer(T5) [22] as the backbone of the encoder-decoder setup of our generative model. We first train the T5 model to understand the user's emotion e_u first and then to predict an emotion e_s and a dialog act d_s in the next turn through context, and use a set of exemplars T to guide the generation of the response specific to e_s and d_s.

Encoder: Following common practice, we concatenate dialogue context sentences $C = \{s_0, l_0, s_1, l_1, \ldots, s_n\}$, the input representation for each word is the sum of its token embedding (E_K), position embedding (E_P) and speaker embedding (E_S): $E(C) = E_K(C) + E_P(C) + E_S(C)$, where $E(C) \in \mathbb{R}^{k\times n_{emb}}$, k is the number of words in C, and n_{emb} is the embedding length of the words. Then they are fed into the T5 encoder Trs^{ctx}_{enc}:

$$Z = Trs^{ctx}_{enc}(E(C)) \tag{2}$$

Where $Z \in \mathbb{R}^{k\times D_{emb}}$ is the encoder output holding context representation.

We use the first last hidden state $Z_{[0]}$ of the encoder output Z as the comprehensive representation of the dialogue context. Then the representation is fed into a linear layer to obtain the user's emotion probability distribution:

$$P\left(e_u \mid C\right) = softmax\left(Z_{[0]}\boldsymbol{V}_e^{\top}\right) \tag{3}$$

Where $\boldsymbol{V}_e \in \mathbb{R}^{N_e\times D}$ is the layer weights shared with the emotion state embedding, N_e is the number of the emotion (here is 7) and D is the vector dimension.

We analyze the emotion shift pattern as in the [20] between the speaker and the listener in the EmpatheticDialogues corpus, such as when the speaker expresses happiness, the listener will also express relative happiness. So we can calculate the frequency of emotion e_i shifts to emotion e_j and normalize the shifting probability to construct an EM-EM prior matrix $M_e = [a_{i,j}] \in \mathbb{R}^{N_e\times N_e}$ based on the training and validation set of EmpatheticDialogues corpus. We use this prior matrix M_e and dialogue context to predict the emotion e_s in the next turn.

We obtain the predicted user's emotion label e_u using equation (3) above. Then, based on e_u, we can obtain the possible emotion shift probabilities $\boldsymbol{m}_{sft} = [a_{e_u,1}, \cdots, a_{e_u,N_e}]$. Then we can obtain the probability distribution of the emotion e_s in the next turn:

$$P\left(e_s \mid C, e_u, M_e\right) = softmax\left(W_1[Z_{[0]}; \boldsymbol{V}_e^{\top} m_{sft}] + b_1\right) \tag{4}$$

Here W_1 and b_1 are trainable parameters.

Similar to predicting an emotion in the next round, there is also the shift pattern from the emotion to the dialog act between the speaker and the listener. For example, when the speaker expresses sadness, the listener will adopt the dialog act of sympathizing. So firstly, we build an EM-DA prior matrix $M_d = [b_{i,j}] \in \mathbb{R}^{N_e\times N_d}$, here N_d is the number of the dialog act (here is 9). Then we can get the possible dialog act shift probabilities $\boldsymbol{m}_{sft} = [b_{e_u,1}, \cdots, b_{e_u,N_d}]$. After that, we can predict the probability distribution of the dialog act, which is similar to the emotion prediction in the next turn:

$$P\left(d_s \mid C, e_u, M_d\right) = softmax\left(W_2[Z_{[0]}; \boldsymbol{V}_d^{\top} m_{sft}] + b_2\right) \tag{5}$$

Here W_2 and b_2 are trainable parameters. $\boldsymbol{V}_d \in \mathbb{R}^{N_d\times D}$ is the layer weights shared with the dialog act state embedding, N_d is the number of the dialog act (here is 9), and D is the vector dimension.

For optimizing the prediction, we use a negative log-likelihood (NLL) loss, where α and β are hyper-parameters:

$$\begin{aligned} L_{pre} = &\alpha \log p\left(e_u \mid C\right) + (1-\alpha)\log p\left(e_s \mid C, e_u, M_e\right) \\ &+ \beta \log p\left(d_s \mid C, e_u, M_d\right) \end{aligned} \tag{6}$$

Decoder: Given the dialogue context, the predicted emotion and dialog act above, and a set of exemplars, the decoder should generate an empathetic response. We should inject exemplars and the predicted emotion and dialog act during decoding to control the generation process. Firstly, we use a rule-based approach to categorize a response into empathy-generic and empathy-special parts. Then we use different kinds of exemplars for different response parts. Each exemplar t_i is encoded with a T5 encoder (Trs^{exl}_{enc}) :

$$z_i = Trs^{exl}_{enc}\left(E_K\left(t_i\right)\right) \tag{7}$$

$z_i \in \mathbb{R}^{\cdot\times n_{emb}}$ is the token-level exemplar representation, which can be mean-pooled to obtain a vector representation $h_i = mean(z_i) \in \mathbb{R}^{n_{emb}}$. After that, all of the representations are aggregated to obtain the final representation:

$$\tau = \operatorname{mean}\left(\left[h_1, h_2, \ldots, h_q\right]\right) \tag{8}$$

In order to infuse the information of the dialog context and exemplars, we concatenate the exemplar representation τ to the context representation Z from Eq (2) at the token level and fed it to a fully-connected layer FC_{exl} of size n_{emb}:

$$Z_{fused} = \mathrm{FC}_{\mathrm{exl}}\left(\left[Z_i \oplus \tau\right]_{i=1}^{k}\right) \tag{9}$$

Next, we fed this fused representation $Z_{fused} \in \mathbb{R}^{k\times n_{emb}}$ to the decoder for the response generation, R is the final response of the model:

$$P_{resp} = Trs_{dec}\left(E_K\left(R_{1:t-1}\right), Z_{fused}\right) \tag{10}$$

We must also fuse the context with the predicted emotion and dialog act feature. Specifically, we first obtain the user's emotion label l_{e_u} and the emotion label l_{e_s} in the next turn where we can get from Eq (3) and Eq (4) by argmax operation. Similarly, we get l_{d_s} from Eq (5). Then we inject the emotion states and dialog act state to P_{resp} as follows:

$$P^{EM}_{resp} = W_3(P_{resp} + l_{e_u}) + tanh(W_4 P_{resp}) l_{e_s} \tag{11}$$

$$P^{DA}_{resp} = l_{d_s} \odot P_{resp} + l_{d_s} \tag{12}$$

Here W_3 and W_4 are trainable parameters. $\odot$ denotes element-wise multiplication.

Now we should project the two kinds of fused representations into vocabulary logistics space. To effectively merge two distributions, we use a gate control layer[20] to monitor the information flow. We pass the first last hidden state $Z_{[0]}$ of encoder output Z through a linear layer with sigmoid activation to obtain control value γ:

$$\gamma = \sigma\left(W_5 Z_{[0]} + b_5\right) \tag{13}$$

Here,W_5 and b_5 are trainable parameters. Then we can merge logistics to get the final probabilities:

$$P\left(y_i \mid y_{<i}; C; T; V\right)=\gamma P_{resp}^{EM}+(1-\gamma) P_{resp}^{DA} \tag{14}$$

Here, C is the context, T is a set of exemplars, and V contains the predicted emotion and dialog act states. For optimizing the generation, we adopt cross-entropy loss for it.

$$L_{\text{gen}}=-\log p\left(R_{\text{gold}} \mid C\right) \tag{15}$$

Training: We use Adam [12] to optimize the overall loss:

$$\mathcal{L}=\alpha_{gen} L_{gen}+\alpha_{pre} L_{pre} \tag{16}$$

Where α_* are hyper-parameters. We set hyper-parameter α_{gen} = 1.0, α_{pre} = 0.1, and α = 0.6, β = 0.5. The batch size is 8, and the learning rate is set to 1e-5. We trained the model for 50 epochs and adopted early stopping. The decoding strategy is top-k, and the maximum decoding step is set to 20 during inference.

4.5 Grafted Reinforcement Learning

Although the previous model can generate empathetic responses, it does not consider the user's emotional feedback and only selects an appropriate emotion and a dialog act from the dialogue context. Our model not only uses the dialogue context to predict the emotion and the dialog act but also takes the user's emotional feedback into account dynamically at every dialog turn. We design a simulation-based environment to train the model further to maximize the user's emotional experience. The whole model consists of the following four components.

State/Action: At each time step t, the state S_t is C ,where C is the current context. The action space is as same as the emotion space, i.e., $A=\{e_1, e_2, \ldots, e_{N_e}\}$. Thus, at each stage, the action is to assign the user's emotion label from the current context. Then we can use the EM-EM shift matrix M_e and EM-DA shift matrix M_d to influence the selection of the emotion and dialog act in the next turn.

Policy: A stochastic policy π_Θ is adapted to sample the user's emotion label at the current state S_t. The policy function is defined by a softmax function, which is the probability distribution over all emotion labels. As the probability of each emotion differs from one to another, the policy tends to choose the emotion with a more significant likelihood, even though other choices may derive the same reward. Then using the selected emotion label of the user, we can get the emotion and dialog act in the next turn using Eq (4) and Eq (5).

Reward: We train a regression model to estimate the user's sentiment score after receiving a generated response to reward the responses more likely to improve the user's emotional state. We use VADER [8] to generate a sentiment score in [-1,1] for each user utterance as ground truth, -1, 1, and 0 being extremely negative, extremely positive, and neutral, respectively. Then we use the T5 encoder to encode the response R to get the final hidden state of encoder output Z. We consider the starting token vector $Z_{[0]}$ and fed to a classifier:

$$f_{score}(R)=tanh\left(\mathbf{w}^{\top} Z_{[0]}+b\right) \tag{17}$$

Where w and b are trainable parameters. We use MSE loss to optimize the model.

We also get the user's sentiment score before the conversation by VADER, a sentiment analyzer. The difference between the user's sentiment score before the dialogue and the score obtained from Eq (17) is used as a reward signal for reinforcement learning to help select the emotion and dialog act that can elicit the user's positive emotion.

Training: We apply Proximal Policy Optimization [31], a stable policy-based reinforcement learning algorithm using a constant clipping mechanism as the soft constraint, for dialog policy optimization:

$$J_\pi(\theta)=E_{s, a \sim \pi}\left[\min \left\{\beta_t \hat{A}_t, \operatorname{clip}\left(\beta_t, 1-\epsilon, 1+\epsilon\right) \hat{A}_t\right\}\right] \tag{18}$$

$\hat{A}_t=R_t-\hat{V}_\phi\left(s_t\right)$ is the estimated advantage, where $R_t=\sum_{\tau=t}^T$ is accumulated reward, $\hat{V}_\phi$ is the estimated value function of state S_t with parameters ϕ, $\beta_t=\frac{\pi_\theta(a_t|s_t)}{\pi_{\theta_{old}}(a_t|s_t)}$ is the ratio of the probability under the new and old policies, δ is TD residual, ε are hypeparameters, which is set to 0.2 here.

5 Experiment

5.1 Dataset

We conduct our experiments on an open-sourced EmpatheticDialogues corpus[5] [28], a large-scale English empathetic dialogue benchmark dataset. It contains 24850 dyadic dialogues between speakers and listeners. This dataset consists of 32 emotional situations and conversations related to the situation. Speakers share their experiences based on the situation, and then the listener responds empathetically based on the speaker's situation. We split them into training, validation, and testing sets by 80%,10%, and 10%.

5.2 Comparison Models

We compare our model with the following competitive baselines[6]:

- **MIME[23]**: A Transformer-based model uses polarity-based emotion clusters and emotion mimicry.
- **EmpDG[14]**: A Multi-resolution Interactive adversarial model utilizes coarse-grained dialogue-level emotions, fine-grained token-level emotions, and interactive user feedback.
- **LEMPEx[22]**: A T5-based model first use exemplars to cue the generative model but only the empathy-special exemplars.
- **HappyBot[32]**: A seq2seq with attention model trained via reinforcement learning with three kinds of reward function: forward, improvement, and simulation. We choose the improvement reward function for the best performance.

5.3 Automatic Evaluation

In order to evaluate the generative performance of the model, we adopt the widely used BLEU score [24] and Distinct-1/2 [13]. The BLEU score can compare generated text against references in language generation tasks. Distinct-1/2 evaluates the generated diversity by measuring the ratio of unique unigrams/bigrams in the response. Table 1 shows the automatic evaluation results. First, our model outperforms all compared models for empathetic dialogue generation and significantly improves all metrics. Our model achieves about 9.7% progress on BLEU compared to the best baseline MIME, which indicates that our model is more likely to generate ground truth responses than the models for empathetic dialogue generation. And our model achieves significant improvement on Distinct-1 and Distinct-2 compared to LEMPEx, which indicates our model's superiority

[5] github.com/facebookresearch/EmpatheticDialogues

[6] The first three are empathetic dialog systems, while the last one is the work for emotion elicitation dialogue generation.

in generating informative responses at the unigrams and bigrams level. This superiority is attributed to the used exemplars to guide the generation of informative responses on the premise of ensuring the BLEU scores of the generated response. Second, compared to the models for emotion elicitation dialog generation, our model achieves about 6.92% and 30.82% improvement on Distinct-1 and Distinct-2, respectively, compared to the best baseline HappyBot. The automatic evaluation results confirm the effectiveness of joint training by grafting fine-tuning and reinforcement learning in the quality of response generation, and our method can achieve the best generation quality on both the BLEU and Distinct simultaneously.

Table 1. Automatic evaluation between our model and SOTA models

Models	BLEU	Distinct-1	Distinct-2
MIME	8.76	0.63	4.29
EmpDG	8.61	1.81	6.96
LEMPEx	7.88	1.41	14.66
HappyBot	/	2.60	17.00
Ours	**9.61**	**2.78**	**22.24**

5.4 Human Ratings

In human evaluation, we randomly sample 100 dialogues from the testing set. Given the dialogue context and all models' generated responses, we ask three crowd-sourcing workers to assign a score from 1 to 5 (1:not at all, 3:OK, 5:very good) to the responses based on the aspects of Empathy, Relevance, Fluency, and Informative. Positive Emotion Elicitation Ability score is -1, 0, or 1, representing no, remain, or yes. The five specific aspects are (1) **Empathy(Emp)**: whether the response shows an understanding of the user's feelings and experiences and expresses appropriately; (2) **Relevance(Rel)**: whether the response is relevant to the dialogue context; (3) **Fluency(Flu)**: whether the response is readable and grammatically correct; (4) **Informative(Inf)**: whether the response contains valuable information; (5) **Positive Emotion Elicitation Ability(Eli)**: whether the response can improve the user's emotional state. To avoid the influence of model order in the evaluation process, we randomly shuffled the responses of compared models in each session.

Table 2 summarizes human ratings of all models[7] on the EmpatheticDialogues dataset. Our model achieves the best performance in most aspects, which verifies that our model can generate more empathetic, relevant, and informative responses with the guidance of empathy-generic and empathy-special exemplars. As we can see, the fluency evaluation of most models is similar, and our model is only slightly worse than the best model. And our model achieves the best scores in Positive Emotion Elicitation Ability, which indicates that our model can improve the user's emotional state with the help of reinforcement learning compared to the models for empathetic dialogue generation. The human evaluation results confirm that joint training by grafting fine-tuning and reinforcement learning can better equip the model with empathetic and emotion elicitation abilities.

5.5 Ablation Study

We perform ablation studies for our model to better analyze the main components' relative contributions shown in Table 2,3,4. **Does the Reinforcement Learning work?** To investigate the effect of Reinforcement Learning (RL) component, we trained the model via

[7] Due to the outdated and unofficial maintenance of HappyBot's code, we do not conduct a human evaluation on the model.

Table 2. Results of human evaluation among our model and baselines. The agreement among the annotators is measured by Fleiss's kappa.

Models	Emp	Flu	Rel	Inf	Eli
MIME	3.51	3.89	3.23	2.71	0.51
EmpDG	3.32	4.28	3.10	2.30	0.32
LEMPEx	3.52	**4.32**	2.90	2.81	0.60
Ours	**3.82**	4.30	**3.55**	**2.90**	**0.75**
w/o RL	3.10	4.22	3.30	2.51	0.35
w/o RL + w/o EM&DA	2.91	3.82	3.29	2.52	0.31
w/o exemplars	3.38	4.15	2.90	2.50	0.44
kappa	0.45	0.55	0.51	0.58	0.47

retrieval-based multi-task fine-tuning (i.e., without the RL component) and to verify that whether reinforcement learning can elicit the user's positive emotion by taking the user's emotional feedback into account. We also removed the prediction of the emotion and dialog act and just trained the model via the retrieval-based fine-tuning mode (i.e., without RL or EM&DA) to verify the effectiveness of reinforcement learning. We can observe that the Empathy and Positive Emotion Elicitation Ability decreases significantly without RL, which indicates that the reward function of RL makes a remarkable contribution to empathy and emotion elicitation. Besides, we observe that further removing the EM&DA prediction task also leads to a performance drop, indicating that the multi-task learning can also equips the model with empathetic and emotion elicitation abilities. And both the BLEU score and Distinct-1/2 scores show a certain degree of decrease, which also proves the effectiveness of reinforcement learning in the quality of response generation.

Table 3. Impact of reinforcement learning

Models	BLEU	Distinct-1	Distinct-2
Ours	**9.61**	**2.78**	**22.24**
w/o RL	9.58	2.32	18.51
w/o RL + w/o EM&DA	9.53	2.57	19.51

Do the two kinds of exemplars work? First, we remove the retrieval of the two kinds of exemplars (i.e., without exemplars) shown in Table 2 and 4. We can see that after removing the guidance of exemplars, BLEU score declines, which indicates the effectiveness of the guidance of exemplars. And relevance also shows a significant decrease, demonstrating that the exemplars can help generate a response more relevant to the context. But distinct slightly increased, possibly due to the guidance of exemplars playing the role of constraints on the diversity of generated responses. To further verify the effectiveness of these two types of exemplars, we remove either of them separately and Table 4 shows the results. We can find that the two kinds of exemplars both work. And the empathy-special exemplars retrieved from the training set are more effective.

Table 4. Impact of two kinds of exemplars

Models	BLEU	Distinct-1	Distinct-2
Ours	**9.61**	2.78	22.24
w/o exemplars	8.54	**2.89**	22.60
w/o empathy-generic	9.44	2.88	**22.67**
w/o empathy-special	8.98	2.86	22.53

5.6 Case Study

In Table 4, we provide some cases to compare the response generated by our model with those generated by the baselines on the EmpatheticDialogues corpus. Intuitively, compared to other models, our model tends to generate responses consisting of an empathy-generic

Table 5. Generated responses from baselines and our model. The response comprises a blue empathy-generic utterance for empathy and an orange empathy-special utterance for further empathy and emotion elicitation.

Context	Speaker: I went for a walk the other day and one of the neighbor's dogs was loose, a bit white husky, he was aggressive
MIME	oh no, that is so sad. I am sorry you had to deal with that. I hope it was a good thing to deal !
EmpDG	that is terrible. did you try to him?
LEMPEx	Oh man, I am so sorry. What happened?
Ours	**Oh my goodness. Did you get home back?**
Ground-Truth	Was the owner not around? Did the dog cause any harm?
Context	Speaker: My sister recently paid off her house. I felt so envious of her as I'd been trying to pay mine down.
MIME	I am so sorry to hear that. I am so sorry to hear that. I am so sorry to hear that.
EmpDG	that is a good thing to be grateful for you.
LEMPEx	Thats a relief. I'm sure that it's not a great relationship
Ours	**That's good! That's a hard thing to do. You have to do that and then have a good time for her.**
Ground-Truth	That will be such a wonderful feeling when you do. I know it's hard when you see someone reach the "finish line" first – are you able to turn that envy into inspiration?
Context	Speaker: I tripped and fell in front of some people.
MIME	oh no ! I am so sorry to hear that. I hope you were not a good person.
EmpDG	oh no! did you find out who did you?
LEMPEx	Oh my gosh, I hope you were able to find it.
Ours	**Oh no, what happened? Are you going to do anything about this or did you feel bad about it?**
Ground-Truth	Ouch! Was it a serious fall?
Context	Speaker: I'm 38 weeks pregnant, and could have my baby any day now! Listeners: Congrats! That's so exciting do You feel ok? Speakers: Yes, a bit nervous because my first birth was hard, but I know I'm ready and have everything we need for the baby!
MIME	i am sure you will be fine.
EmpDG	i am sure you will be fine.
LEMPEx	You must feel really happy! I hope so :)
Ours	**That's awesome. I hope you will be fine and your family will be happy!**
Ground-Truth	YAY! good job! He/she is going to be beautiful

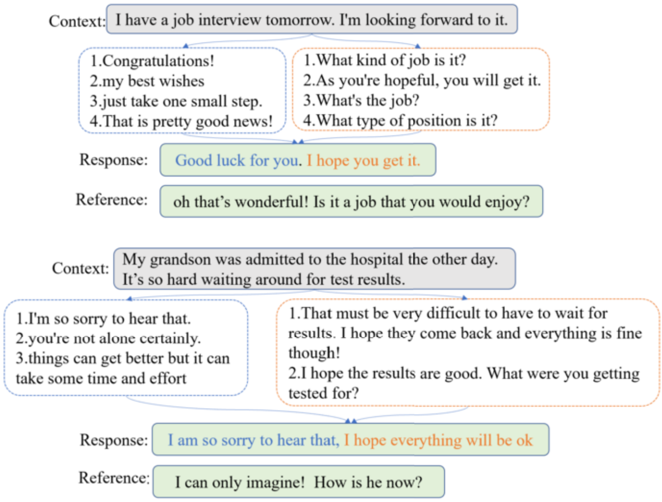

Figure 3. Impact of retrieved exemplars. The exemplars are composed of empathy-generic exemplars in the blue dashed box and empathy-special exemplars in the orange dashed box.

utterance for empathy and an empathy-special utterance for further empathy and emotion elicitation.

In the first case, when the speaker shows afraid towards his neighbor's dog, other baselines also show the negative emotion and generate the responses with negative mood particles such as *"Oh no, that is so sad"*, *"that is terrible"*. Although these can also empathize with users, the response generated by our model expressed empathetic surprise such as *"oh my goodness"*. And the second half of the sentence (*"Did you get home back?"*) shows concern for the user, which can further empathize with the user and elicit positive emotions from the user. **In the second case**, it can better demonstrate that other baselines are more likely to present negative emotions to empathize with users when facing negative emotions from them. When the user feels jealous of his sister's achievement, other baselines empathize with the user by feeling regretful. But our model first affirms what the user's sister has done and then encourages the user, which is more consistent with the emotion and the dialog act shown in the ground truth. **In the third case**, compared to other baselines, our model alleviates the user's embarrassment by asking the user more about his feeling. **In the fourth case of multiple-turn dialogue**, compared to other models, our model shows stronger excitement and explores more valuable information.

Impact of retrieved exemplars. By our observation, we find that an empathy-generic utterance and an empathy-special utterance usually constitute a human-like emotional response. So we further validate the effectiveness of exemplars that we retrieve for these two parts of the response and how these exemplars participated in the response generation. The results are shown in Figure 3.

In the first case, the former is an empathy-generic utterance such as *"Congratulations!"*, *"my best wishes"* which can express wishes generally, and the other is an empathy-special utterance that asks the user about specific situations. These two parts together form the response. In the second case, the former can also show worried like the user, and the other comfort the user.

6 Conclusion and Future work

In this paper, we propose a novel and more challenging task named the Empathetic Emotion Elicitation dialog system, which aims to simultaneously empathize with users and elicit users' positive emotions. We design a unified framework with three stages. We first retrieve exemplars and then fine-tune the emotion/DA prediction tasks on a pre-trained generative language model to enhance the empathetic ability. And finally, we model the user information by reinforcement learning to enhance the emotion elicitation ability. Extensive experiments verify the superiority of our model on automatic and human evaluation. In the future, our work will consider generating multiple rounds of empathetic dialogue to maximize the benefits of reinforcement learning.

References

[1] Mao Yan Chen, Siheng Li, and Yujiu Yang, 'Emphi: Generating empathetic responses with human-like intents', in *NAACL 2022, Seattle, WA, United States, July 10-15, 2022*, pp. 1063–1074, (2022).

[2] Mingda Chen, Qingming Tang, Sam Wiseman, and Kevin Gimpel, 'Controllable paraphrase generation with a syntactic exemplar', in *ACL 2019, Florence, Italy, July 28- August 2, 2019, Volume 1: Long Papers*, pp. 5972–5984, (2019).

[3] M. H. Davis, 'Measuring individual differences in empathy: Evidence for a multidimensional approach.', *Journal of Personality & Social Psychology*, **44**(1), 113–126, (1983).

[4] Dorottya Demszky, Dana Movshovitz-Attias, Jeongwoo Ko, Alan S. Cowen, Gaurav Nemade, and Sujith Ravi, 'Goemotions: A dataset of fine-grained emotions', in *ACL 2020, Online, July 5-10, 2020*, pp. 4040–4054, (2020).

[5] P. Ekman, 'Facial expressions of emotion: New findings, new questions', *Psychological Science*, **3**(1), 34–38, (1992).

[6] R. Elliott, A. C. Bohart, J. C. Watson, D. Murphy, and E. &. Outcome, 'Therapist empathy and client outcome: An updated meta-analysis', *Psychotherapy Theory Research & Practice*, **55**(4), (2018).

[7] T. Hasegawa, N. Kaji, N. Yoshinaga, and M. Toyoda, 'Predicting and eliciting addressee's emotion in online dialogue', in *Meeting of the Association for Computational Linguistics*, (2013).

[8] Clayton J. Hutto and E. Gilbert, 'Vader: A parsimonious rule-based model for sentiment analysis of social media text', in *International Conference on Weblogs and Social Media*, (2014).

[9] Jiun-Hao Jhan, Chao-Peng Liu, Shyh-Kang Jeng, and Hung-Yi Lee, 'Cheerbots: Chatbots toward empathy and emotionusing reinforcement learning', *CoRR*, **abs/2110.03949**, (2021).

[10] Vladimir Karpukhin, Barlas Oguz, Sewon Min, Patrick S. H. Lewis, Ledell Wu, Sergey Edunov, Danqi Chen, and Wen-tau Yih, 'Dense passage retrieval for open-domain question answering', in *EMNLP 2020, Online, November 16-20, 2020*, pp. 6769–6781, (2020).

[11] Hyunwoo Kim, Byeongchang Kim, and Gunhee Kim, 'Perspective-taking and pragmatics for generating empathetic responses focused on emotion causes', in *EMNLP 2021, Virtual Event / Punta Cana, Dominican Republic, 7-11 November, 2021*, pp. 2227–2240, (2021).

[12] D. Kingma and J. Ba, 'Adam: A method for stochastic optimization', *Computer Science*, (2014).

[13] J. Li, M. Galley, C. Brockett, J. Gao, and B. Dolan, 'A diversity-promoting objective function for neural conversation models', *Computer Science*, (2015).

[14] Qintong Li, Hongshen Chen, Zhaochun Ren, Pengjie Ren, Zhaopeng Tu, and Zhumin Chen, 'Empdg: Multi-resolution interactive empathetic dialogue generation', in *COLING 2020, Barcelona, Spain (Online), December 8-13,2020*, pp. 4454–4466, (2020).

[15] Qintong Li, Piji Li, Zhaochun Ren, Pengjie Ren, and Zhumin Chen, 'Knowledge bridging for empathetic dialogue generation', in *AAAI 2022, Thirty-Fourth Conference on Innovative Applications of Artificial Intelligence,EAAI2022 Virtual Event, February 22- March 1, 2022*, pp. 10993–11001, (2022).

[16] S. Li, S. Feng, D. Wang, K. Song, and W. Wang, 'Emoelicitor: An open domain response generation model with user emotional reaction awareness', in *Twenty-Ninth International Joint Conference on Artificial Intelligence and Seventeenth Pacific Rim International Conference on Artificial Intelligence IJCAI-PRICAI-20*, (2020).

[17] Yanran Li, Ke Li, Hongke Ning, Xiaoqiang Xia, Yalong Guo, Chen Wei, Jianwei Cui, and Bin Wang, 'Towards an online empathetic chatbot with emotion causes', in *SIGIR '21: The 44th International ACM SIGIR Conference on Research and Development in Information Retrieval, Virtual Event, Canada, July 11-15, 2021*, pp. 2041–2045, (2021).

[18] Zhaojiang Lin, Andrea Madotto, Jamin Shin, Peng Xu, and Pascale Fung, 'Moel: Mixture of empathetic listeners', in *EMNLP-IJCNLP 2019, Hong Kong, China, November 3-7, 2019*, pp. 121–132, (2019).

[19] Ye Liu, Wolfgang Maier, Wolfgang Minker, and Stefan Ultes, 'Empathetic dialogue generation with pre-trained roberta-gpt2 and external knowledge', in *IWSDS 2021, Singapore*, volume 943, pp. 67–81, (2021).

[20] Yuhan Liu, Jun Gao, Jiachen Du, Lanjun Zhou, and Ruifeng Xu, 'Empathetic response generation with state management', *arXiv e-prints*, (2022).

[21] N. Lubis, S. Sakti, K. Yoshino, and S. Nakamura, 'Positive emotion elicitation in chat-based dialogue systems', *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 1–1, (2019).

[22] Navonil Majumder, Deepanway Ghosal, Devamanyu Hazarika, Alexander F. Gelbukh, Rada Mihalcea, and Soujanya Poria, 'Exemplars-guided empathetic response generation controlled by the elements of human communication', *IEEE Access*, **10**, 77176–77190, (2022).

[23] Navonil Majumder, Pengfei Hong, Shanshan Peng, Jiankun Lu, Deepanway Ghosal, Alexander F. Gelbukh, Rada Mihalcea, and Soujanya Poria, 'MIME: mimicking emotions for empathetic response generation', in *EMNLP 2020, Online, November 16-20, 2020*, pp. 8968–8979, (2020).

[24] Kishore Papineni, Salim Roukos, Todd Ward, and Wei-Jing Zhu, 'Bleu: a method for automatic evaluation of machine translation', in *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics, July 6-12, 2002, Philadelphia, PA, USA*, pp. 311–318, (2002).

[25] Verónica Pérez-Rosas, Rada Mihalcea, Kenneth Resnicow, Satinder Singh, and Lawrence C. An, 'Understanding and predicting empathic behavior in counseling therapy', in *ACL 2017, Vancouver, Canada, July 30 - August 4, Volume 1: Long Papers*, pp. 1426–1435, (2017).

[26] Yushan Qian, Bo Wang, Ting-En Lin, Yinhe Zheng, Ying Zhu, Dongming Zhao, Yuexian Hou, Yuchuan Wu, and Yongbin Li, 'Empathetic response generation via emotion cause transition graph', *CoRR*, **abs/2302.11787**, (2023).

[27] Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J. Liu, 'Exploring the limits of transfer learning with a unified text-to-text transformer', *J. Mach. Learn. Res.*, **21**, 140:1–140:67, (2020).

[28] Hannah Rashkin, Eric Michael Smith, Margaret Li, and Y-Lan Boureau, 'Towards empathetic open-domain conversation models: A new benchmark and dataset', in *ACL 2019, Florence, Italy, July 28- August 2, 2019, Volume 1: Long Papers*, pp. 5370–5381, (2019).

[29] Nils Reimers and Iryna Gurevych, 'Sentence-bert: Sentence embeddings using siamese bert-networks', in *EMNLP-IJCNLP 2019, Hong Kong, China, November 3-7, 2019*, pp. 3980–3990, (2019).

[30] Sahand Sabour, Chujie Zheng, and Minlie Huang, 'CEM: commonsense-aware empathetic response generation', in *Thirty-Sixth AAAI Conference on Artificial Intelligence*, pp. 11229–11237, (2022).

[31] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov, 'Proximal policy optimization algorithms', *CoRR*, **abs/1707.06347**, (2017).

[32] Jamin Shin, Peng Xu, Andrea Madotto, and Pascale Fung, 'Generating empathetic responses by looking ahead the user's sentiment', in *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 7989–7993. IEEE, (2020).

[33] Richard Thwaites and James Bennett-Levy, 'Conceptualizing empathy in cognitive behaviour therapy: Making the implicit explicit', *Behavioural and Cognitive Psychotherapy*, **35**(5), 591–612, (2007).

[34] Anuradha Welivita and Pearl Pu, 'A taxonomy of empathetic response intents in human social conversations', in *COLING 2020, Barcelona, Spain (Online), December 8-13,2020*, pp. 4886–4899, (2020).

[35] Anuradha Welivita and Pearl Pu, 'Heal: A knowledge graph for distress management conversations', *Proceedings of the AAAI Conference on Artificial Intelligence*, **36**(10), 11459–11467, (Jun. 2022).

[36] Hao Zhou, Minlie Huang, Tianyang Zhang, Xiaoyan Zhu, and Bing Liu, 'Emotional chatting machine: Emotional conversation generation with internal and external memory', in *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence*, pp. 730–739, (2018).

[37] Xianda Zhou and William Yang Wang, 'Mojitalk: Generating emotional responses at scale', in *ACL 2018, Melbourne, Australia, July 15-20, 2018, Volume 1: Long Papers*, pp. 1128–1137, (2018).