

# CDL: Curriculum Dual Learning for Emotion-Controllable Response Generation

Lei Shen<sup>1,2</sup> Yang Feng<sup>1,2\*</sup>

<sup>1</sup>Key Laboratory of Intelligent Information Processing,  
Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China

<sup>2</sup>University of Chinese Academy of Sciences, Beijing, China

{shenlei17z, fengyang}@ict.ac.cn

## Abstract

Emotion-controllable response generation is an attractive and valuable task that aims to make open-domain conversations more empathetic and engaging. Existing methods mainly enhance the emotion expression by adding regularization terms to standard cross-entropy loss and thus influence the training process. However, due to the lack of further consideration of content consistency, the common problem of response generation tasks, safe response, is intensified. Besides, query emotions that can help model the relationship between query and response are simply ignored in previous models, which would further hurt the coherence. To alleviate these problems, we propose a novel framework named Curriculum Dual Learning (CDL) which extends the emotion-controllable response generation to a dual task to generate emotional responses and emotional queries alternatively. CDL utilizes two rewards focusing on emotion and content to improve the duality. Additionally, it applies curriculum learning to gradually generate high-quality responses based on the difficulties of expressing various emotions. Experimental results show that CDL significantly outperforms the baselines in terms of coherence, diversity, and relation to emotion factors.

## 1 Introduction

Infusing emotions into dialogue systems can make conversational agents more human-like and benefit the interaction between human and machine (Prendinger and Ishizuka, 2005; Prendinger et al., 2005; Partala and Surakka, 2004). In some real-life scenarios, we need to customize and control the agent’s emotion so that the agent can express a specific one. For example, in psychological counseling, the agent is supposed to express sadness to

show the sympathy and also convey happiness to cheer the patient up.

|   |                       |   |
|---|-----------------------|---|
| 1 | <i>q</i>              | It is very pleasant to have a cup of black tea with sugar on a cold day. (Happy)                    |
|   | <i>r</i> <sub>1</sub> | [Neural] It starts to cool down today.  |
|   | <i>r</i> <sub>2</sub> | [Like] I will try, thanks for your advice.  |
|   | <i>r</i> <sub>3</sub> | [Sad] I am frozen to death ...  |
|   | <i>r</i> <sub>4</sub> | [Disgust] Winner is the worst season.   |
|   | <i>r</i> <sub>5</sub> | [Angry] You know nothing!   |
| 2 | <i>q</i>              | So pets live better than humans now... (Sad)  |
|   | <i>r</i> <sub>1</sub> | [Disgust] You are so bad.   |
|   | <i>r</i> <sub>2</sub> | [Happy] Haha, you too.  |
| 3 | <i>q</i>              | We should study hard. (Neural)  |
|   | <i>r</i>              | [Disgust] You are so bad.   |
| 4 | <i>q</i>              | Happy birthday, Xinxin. May you be more beautiful, find a good person and get married soon! (Happy) |
|   | <i>r</i>              | [Happy] Haha, you too.  |

Table 1: Examples of emotion-controllable response generation (response emotions are denoted in brackets). Example 1 is one query and 6 emotional responses. Example 2 and 3 have different queries, but the responses generated with emotion “Disgust” are the same. Similar to Example 2 and 4 with emotion “Happy”. The emotions of queries are marked in parentheses.

Recently, a framework called emotional chatting machine (ECM) (Zhou et al., 2018a) was proposed to address the emotion factor in a controlled manner, which focuses on generating a response with a specific emotion (Example 1 in Table 1). In the research field of emotion-controllable response generation, ECM and its successive methods (Colombo et al., 2019; Song et al., 2019) mainly represent the given emotion category as a vector and add it to the decoding steps to influence the procedure of response generation, which would aggravate the safe response problem. For the response generation task, safe response is notorious, as the model tends to produce some generic but meaningless responses, like “Thank you”, “I don’t know”, “Yes”,

\*Yang Feng is the corresponding author.

etc. Due to the constraint of emotion factors, the scale of proper responses shrinks, and the model is more likely to map any query to a frequently-occurring response in that emotion category. That is, given “Disgust”, the response would be “You are so bad” in general, while given “Happy”, it would be “Haha, you too” (Example 2 to 4 in Table 1).

Intuitively, for a good pair of query and response, they should be in a tight relationship and have equal qualities. Then, both the query-to-response mapping and response-to-query mapping would be easier and more natural. On the contrary, it is hard for a safe response to reach the original query through back-generation, neither on the content level nor the emotion level. At the same time, the difficulties of producing various emotions are different, especially in a noisy and uneven-quality dataset. Therefore, we can evaluate the response based on the feedback from the backward process to improve the coherence (Zhang et al., 2018; Cui et al., 2019; Luo et al., 2019b) and try to learn from easy to hard data to generate appropriate and emotion-rich responses.

In this paper, we propose a new framework for emotion-controllable response generation named Curriculum Dual Learning (CDL). We take the learning of response and query generation with emotions as a dual task, and use the duality to model the mutual relation between them. The forward and backward models are trained alternatively via reinforcement learning (RL). Rewards designed here aim to encourage both emotion expression and content consistency. Specifically, emotion expression can be either explicit (embodied in some obvious emotion words) or implicit (reflected by the organization of the entire sentence). For example, “I am happy to meet her again” is explicit with the word “happy”, while “It seems like I have eaten the honey” is implicit, but the happiness can be felt when we consider the sentence as a whole. Based on these features, we use the accuracy of emotion classification of sentences and the proportion of emotion words as feedbacks for explicit and implicit emotions, respectively. For content consistency, we apply the reconstruction probability as the measurement of coherence (Section 3.1). Furthermore, in order to better utilize samples of multiple emotions from the noisy and uneven-quality dataset, we incorporate the curriculum learning (Section 3.2) into our dual learning framework (Section 3.3).

Experimental results on both automatic and human evaluations show that for a given query and an emotion category, our CDL can successfully express desired emotion as well as keep the response informative and coherent to the query.

## 2 Background

For emotion-controllable response generation, given a query  $q$  and an emotion category  $e_r$ , the goal is to generate a response  $r'$  that is not only meaningful, but also in accordance with the desired emotion.

*Emotional Chatting Machine (ECM)* (Zhou et al., 2018a) addresses the emotion factor using three new mechanisms: Emotion Category Embedding, Internal Memory, and External Memory. Specifically, 1) Emotion Category Embedding models the high-level abstraction of emotion expression by embedding emotion categories, and concatenates corresponding embedding to the input at each decoding step. 2) Internal Memory captures the change of implicit internal emotion states with read and write gates, 3) External Memory applies an external emotion vocabulary to express emotion explicitly, and finally assigns different generation probabilities to emotion and generic words. The loss function on one training sample  $(q, r)$  ( $q = q_1, q_2, \dots, q_n, r = r_1, r_2, \dots, r_m$ ) is defined as:

$$-\sum_{t=1}^m p_t \log(\mathbf{o}_t) - \sum_{t=1}^m q_t \log(\alpha_t) + \|M_{e,m}^I\|, \quad (1)$$

where  $\mathbf{o}_t$  and  $p_t$  are the predicted token distribution and gold distribution,  $\alpha_t$  is the probability of choosing an emotion word or a generic word,  $q_t \in \{0, 1\}$  is the true choice between them in  $r$ , and  $M_{e,m}^I$  is the internal emotion state at the last step  $m$ . The first term is the cross-entropy loss, the second one is used to supervise the probability of selecting an emotion or generic word, and the last one is used to ensure that the internal emotion state has been expressed completely once the generation is finished. Please refer to the original paper for more details.

## 3 CDL for Emotion-Controllable Response Generation

Since our CDL method is a combination of dual learning (DL) and curriculum learning (CL), we first present the main components of DL, including states, actions, policy and reward, then introduce

the plausibility of curriculum learning. Finally, we describe the training algorithm of CDL.

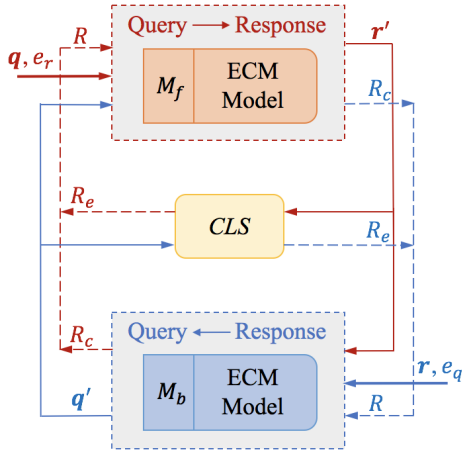


Figure 1: The architecture of dual learning.  $CLS$ ,  $M_f$  and  $M_b$  are emotion classifier, forward model and backward model, respectively. Red parts are for the forward process, while blue parts are for the backward process.

### 3.1 DL Architecture

The architecture of DL is illustrated in Figure 1. Both the forward model  $M_f$  and the backward model  $M_b$  are ECMs with independent parameters and are initialized according to the maximum likelihood estimation (MLE).  $CLS$  is a pre-trained classifier that calculates the score of implicit emotion expression.

In general,  $M_f$  generates a response  $r'$  for a given query  $q$  and emotion category  $e_r$ , and then obtains the reward  $R$  that consists of  $R_e$  from  $CLS$  and  $R_c$  from  $M_b$  (red parts in Figure 1). Similarly,  $M_b$  generates a query  $q'$  for a given response  $r$  and emotion category  $e_q$ , and obtains the reward  $R$  that consists of  $R_e$  and  $R_c$  from  $CLS$  and  $M_f$  (blue parts in Figure 1). These two models are trained alternatively via reinforcement learning (RL). Specifically, an action is the dialogue response to generate. The action space is infinite since arbitrary-length sequences can be generated. A state is denoted by the query, which is further transformed to a vector representation by the encoder. A policy takes the form of a GRU encoder-decoder and is defined by its parameters. Following the work of Li et al. (2016c); Zhang et al. (2018), we use a stochastic representation of the policy, i.e., a probability distribution over actions given states.

In order to encourage both content consistency and emotion expression, we introduce two rewards and use them to train  $M_f$  and  $M_b$ . The definition

of the two rewards for model  $M_f$  is introduced as follows<sup>1</sup>.

**Reward for emotion expression** For implicit emotion expression, a straightforward method is to employ the pre-trained classifier  $CLS$  to evaluate the emotion category of the generated response  $r'$ , and use the classification accuracy as the reward:

$$R_{e_1(q,r')} = p(e_r|r'; \varphi), \quad (2)$$

where  $\varphi$  is the parameter of  $CLS$ , and it is fixed during training. For explicit emotion expression, the reward is formulated as:

$$R_{e_2(q,r')} = n(w_{e_r})/|r'|, \quad (3)$$

where  $n(w_{e_r})$  is the number of emotion words belong to category  $e_r$ , and  $|r'|$  is the length of  $r'$ . Then, the emotion reward is defined as:

$$R_e(q,r') = R_{e_1(q,r')} + \lambda R_{e_2(q,r')}, \quad (4)$$

where  $\lambda$  controls the relative importance of implicit and explicit rewards.

**Reward for content consistency** If the response are coherent and related to the query, it will be easier to reproduce the query via back generation. Inspired by Zhang et al. (2018); Cui et al. (2019); Luo et al. (2019b), we measure the coherence by means of reconstructing  $q$  conditioned on  $r'$ . Formally, the content consistency reward is defined as:

$$R_c(q,r') = p(q|r', e_q; \eta), \quad (5)$$

where  $\eta$  is the parameter of backward model  $M_b$ , and it is fixed during the training of  $M_f$ .

**Overall reward** We use the weighted sum of the above two rewards as the final reward:

$$R(q,r') = R_c(q,r') + \gamma R_e(q,r'), \quad (6)$$

where  $\gamma$  is a hyper-parameter that controls the trade-off between  $R_c(q,r')$  and  $R_e(q,r')$ .

### 3.2 Curriculum Plausibility

Intuitively, learning from less noisy and even-quality dataset is simpler, but in this task, the data is inherently complicated as there are multiple emotions mixed in it. To better utilize the data, we integrate curriculum learning into the dual learning framework. The core of curriculum learning

<sup>1</sup>Rewards for model  $M_b$  can be computed in a similar way, where  $q'$ ,  $r$ ,  $b$  and  $f$  replace  $r'$ ,  $q$ ,  $f$  and  $b$ , respectively. Therefore, we omit them here for space limitation and brevity.

(Bengio et al., 2009) is to design an evaluation for complexity, and to provide the model with *easy* samples first, then gradually increase the difficulty. The curriculum is arranged by sorting each sample in training set according to a specific ranking standard.

Here, We reorder samples from easy, i.e., with high accuracy of emotion classification, to hard. We consider the classification accuracy after pre-training as an indicator of the learning order. Another intuitive way is to put emotionless samples (labelled as “Neural”) first and then emotional ones, however, it exhibits poor performance in our experiments. At training step  $t$ , a batch of training samples is obtained from the top  $f(t)$  portions of the entire sorted training samples. Following Platanios et al. (2019) and Cai et al. (2020), we define the function  $f(t)$  as:

$$f(t) \triangleq \min(1, \sqrt{\frac{t(1-c_0^2)}{T} + c_0^2}), \quad (7)$$

where  $c_0^2$  is set to 0.01, which means that the model starts training using the 1% easiest training samples, and  $T$  is a hyper-parameter that represents the duration of curriculum learning (curriculum length). At the early stage of the training process, the model learns from the samples in the easy part of the curriculum, where there is only one emotion category. As the advance of the curriculum, the difficulty gradually increases, as complex training samples from more different categories appear. After training  $T$  batches, training sample of each batch is drawn from the whole training set, which is the same as the conventional training procedure.

### 3.3 Training of CDL

**Optimization** We use the policy gradient method (Williams, 1992) to find parameters that lead to a larger expected reward. For the forward learning process, the expected reward of the generated response  $\mathbf{r}'$  and its approximate gradient are defined as:

$$J(\theta) = \mathbb{E}[R_{(q,r')}], \quad (8)$$

$$\nabla_{\theta} J(\theta) \simeq R'_{(q,r')} \cdot \nabla_{\theta} \log(p_{\theta}(\mathbf{r}'|q, e_r)), \quad (9)$$

where  $\theta$  is the parameter of forward model  $M_f$ ,  $R'_{(q,r')} = R_{(q,r')} - b_f$ , and  $b_f$  is the baseline value from the greedy search decoding method for  $M_f$ , which is used to reduce the variance of the estimation (Zaremba and Sutskever, 2015; Paulus et al., 2017). Analogously, for the backward learning process, the expected reward of the generated query

$q'$  and corresponding approximate gradient are defined as:

$$J(\eta) = \mathbb{E}[R_{(r,q')}], \quad (10)$$

$$\nabla_{\eta} J(\eta) \simeq R'_{(r,q')} \cdot \nabla_{\eta} \log(p_{\eta}(q'|r, e_q)), \quad (11)$$

where  $\eta$  is the parameter of backward model  $M_b$ ,  $R'_{(r,q')} = R_{(r,q')} - b_b$ , and  $b_b$  is the baseline value from the greedy search decoding method for  $M_b$ .

---

#### Algorithm 1 Curriculum dual learning algorithm for emotion-controllable response generation

---

**Input:** The training set  $\mathcal{D} = \{(q_i, e_{q_i}, r_i, e_{r_i})\}$  where each query-response pair is labelled with corresponding emotion labels  $e_{q_i}$  and  $e_{r_i}$

**Output:**  $M_f$  and  $M_b$

- 1: Pre-train  $M_f$  and  $M_b$  with  $(q_i, r_i, e_{r_i})$  and  $(r_i, q_i, e_{q_i})$ , respectively, based on Eq. 1
  - 2: Pre-train  $CLS$  with  $(q_i, e_{q_i})$  and  $(r_i, e_{r_i})$
  - 3: Sort training samples according to the ranking standard in Section 3.2 for both forward and backward learning process to get  $\mathcal{D}_f$  and  $\mathcal{D}_b$
  - 4: **for** training step  $t = 1, \dots, T$  **do**
  - 5: ▷ Train  $M_f$
  - 6:   Sample a batch  $B_{f_t}$  in  $\mathcal{D}_f$  based on Eq. 7
  - 7:   Sample  $(q, r, e_r)$  from  $B_{f_t}$
  - 8:   Generate response  $r'$  via  $M_f$
  - 9:   Compute reward  $R$  based on Eq. 6
  - 10:   Update  $\theta$  using  $R$  based on Eq. 9
  - 11:   Teacher Forcing: Update  $\theta$  with  $(q, r, e_r)$
  - 12: ▷ Train  $M_b$
  - 13:   Sample a batch  $B_{b_t}$  in  $\mathcal{D}_b$  based on Eq. 7
  - 14:   Sample  $(r, q, e_q)$  from  $B_{b_t}$
  - 15:   Generate response  $q'$  via  $M_b$
  - 16:   Compute reward  $R$  based on Eq. 6
  - 17:   Update  $\eta$  using  $R$  based on Eq. 11
  - 18:   Teacher Forcing: Update  $\eta$  with  $(r, q, e_q)$
  - 19: **end for**
- 

**Teacher Forcing** When  $M_f$  and  $M_b$  are trained with only the rewards from the dual tasks, the training process would easily collapse as it may find an unexpected way to achieve a high reward but fail to guarantee the fluency or readability of the generated text (Ranzato et al., 2015; Pasunuru and Bansal, 2018; Luo et al., 2019b). To stabilize the training process, after each update according to Eq. 9 or 11,  $M_f$  or  $M_b$  is exposed to real query-response pairs and is trained via MLE, which is also known as Teacher Forcing (Li et al., 2017; Lamb et al., 2016).

The training procedure of CDL is summarized in Algorithm 1. First, we use MLE to pre-train

$M_f$ ,  $M_b$  and  $CLS$  with query-response pairs and emotion labels in the training set. After the pre-training phase, we sort samples in the training set following the ranking standard in Section 3.2. For forward learning process, the ranking is based on responses, while for backward learning process, it is based on queries. Then, we can get two sorted training set  $\mathcal{D}_f$  and  $\mathcal{D}_b$  for each direction. Finally,  $M_f$  and  $M_b$  are optimized with rewards and the regularization of Teacher Forcing, alternatively.

## 4 Experiments

In this section, we conduct experiments to evaluate our proposed method. We first introduce some empirical settings, including dataset, hyper-parameters, baselines, and evaluation measures. Then we illustrate our results under both automatic and human evaluations. Finally, we give out some cases generated by different models and do further analyses over our method.

### 4.1 Dataset

We apply our method on the corpus of NLPCC 2017 Emotional Conversation Generation Challenge<sup>2</sup>, namely NLPCC2017 Dataset, which is an extension version of the dataset collected by Zhou et al. (2018a). The provided dataset is already segmented into Chinese words. There are over 1 million query-response pairs, in which both the query and response are labelled with one emotion tag among ‘‘Happy’’, ‘‘Angry’’, ‘‘Disgust’’, ‘‘Sad’’, ‘‘Like’’ and ‘‘Neutral’’. The dataset has been tokenized into words. We randomly split the whole dataset into training/validation/test set with the number of 1,105,487/11,720/2,000. The detailed statistics of training set are shown in Table 2.

|            | Emotion | Query     | Response |
|------------|---------|-----------|----------|
| Training   | Happy   | 120,358   | 197,528  |
|            | Angry   | 79,611    | 138,198  |
|            | Disgust | 184,427   | 197,428  |
|            | Sad     | 128,482   | 179,215  |
|            | Like    | 257,471   | 197,565  |
|            | Neutral | 335,138   | 195,553  |
|            |         | 1,105,487 |          |
| Validation |         | 11,720    |          |
| Test       |         | 2,000     |          |

Table 2: Statistics of the NLPCC2017 Dataset. In the training set, we count the number of queries and responses for each emotion category.

<sup>2</sup><http://coai.cs.tsinghua.edu.cn/hml/challenge2017/>

### 4.2 Hyper-parameter Settings

The settings of both  $M_f$  and  $M_b$  follow the default implementation details of original ECM paper (Zhou et al., 2018a), where the encoder and decoder have 2-layer GRU structures with 256 hidden cells for each layer, the embedding size of words and emotion categories are set to 100, and the vocabulary size is limited to 40,000. The minimum and maximum sentence length is set to 3 and 30, respectively. We train a TextCNN-based classifier (Kim, 2014) and the classification accuracy reaches 65.6% on the test set, which has the similar performance with those used by (Zhou et al., 2018a) and (Song et al., 2019). Before curriculum dual learning, model  $M_f$  and  $M_b$  are pre-trained 10 epochs via MLE. The optimizer is Adam (Kingma and Ba, 2015) with 0.05 initial learning rate for pre-training and  $10^{-5}$  for curriculum dual learning. The batch size is set to 64.  $\lambda$  in Eq. 4 is 0.5,  $\gamma$  in Eq. 6 is 1 and  $T$  in Eq. 7 is 100k. During curriculum dual learning, training runs until the performance on validation set does not improve.

### 4.3 Baselines

We compare our approach with four representative baselines: (1) **S2S-Attn**: The Seq2Seq model with attention mechanism as in Shang et al. (2015). (2) **EmoEmb**: A Seq2Seq variant which takes the embedding of emotion categories as additional input at each decoding position (Ficler and Goldberg, 2017; Li et al., 2016b). (3) **EmoDS**: An emotional dialogue system with lexicon-based attention and a word-based classifier (Song et al., 2019). (4) **ECM**: Emotional Chatting Machine proposed by Zhou et al. (2018a).

Additionally, we also conduct ablation study to better analyze our method as follows: (5) **CDL-emo**: CDL with emotion reward only; (6) **CDL-con**: CDL with content reward only, which is similar to the work of Zhang et al. (2018); (7) **CDL-DL**: CDL with both rewards but without curriculum learning.

### 4.4 Evaluation Measures

To better evaluate our results, we use both quantitative metrics and human judgements in our experiments.

#### 4.4.1 Automatic Metrics

For automatic evaluation, we mainly choose four kinds of metrics: 1) Embedding scores (Average,

| Method         | Embedding Metrics |              |              |              | Diversity    |              | BLEU Scores   |               | Emotion Expression |              |
|----------------|-------------------|--------------|--------------|--------------|--------------|--------------|---------------|---------------|--------------------|--------------|
|                | Avg.              | Ext.         | Gre.         | Coh.         | Dist-1       | Dist-2       | BLEU-1        | BLEU-2        | Emo-acc.           | Emo-word.    |
| S2S-Attn       | 0.497             | 0.352        | 0.328        | 0.582        | 0.035        | 0.119        | 0.0424        | 0.0073        | 0.244              | 0.285        |
| EmoEmb         | 0.532             | 0.381        | 0.356        | 0.594        | 0.040        | 0.133        | 0.0722        | 0.0164        | 0.693              | 0.436        |
| EmoDS          | 0.623             | 0.427        | 0.403        | 0.603        | 0.050        | 0.174        | 0.0976        | 0.0282        | 0.746              | 0.527        |
| ECM            | 0.625             | 0.433        | 0.405        | 0.607        | 0.052        | 0.177        | 0.1023        | 0.0332        | 0.753              | 0.562        |
| CDL-emo (ours) | 0.631             | 0.451        | 0.435        | 0.615        | 0.058        | 0.193        | 0.1162        | 0.0342        | 0.765              | 0.583        |
| CDL-con (ours) | 0.628             | 0.441        | 0.417        | 0.612        | 0.055        | 0.182        | 0.1059        | 0.0338        | 0.758              | 0.566        |
| CDL-DL (ours)  | 0.635             | 0.452        | 0.431        | 0.630        | 0.062        | 0.217        | 0.1187        | 0.0353        | 0.794              | 0.615        |
| CDL (ours)     | <b>0.642</b>      | <b>0.457</b> | <b>0.438</b> | <b>0.635</b> | <b>0.065</b> | <b>0.221</b> | <b>0.1254</b> | <b>0.0370</b> | <b>0.823</b>       | <b>0.620</b> |

Table 3: Automatic evaluation results for content and emotion measurements. The metrics Average, Extrema, Greedy, Coherence, Emotion-acc and Emotion-word are abbreviated as Avg., Ext., Gre., Coh., Emo-acc. and Emo-word., respectively.

| Method   | Like         |              | Sad          |              | Disgust      |              | Angry        |              | Happy        |              | Overall      |              |
|----------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
|          | Con.         | Emo.         | Con.         | Emo.         | Con.         | Emo.         | Con.         | Emo.         | Con.         | Emo.         | Con.         | Emo.         |
| S2S-Attn | 1.295        | 0.435        | 1.125        | 0.120        | 1.160        | 0.115        | <b>1.255</b> | 0.045        | 1.155        | 0.305        | 1.198        | 0.204        |
| EmoEmb   | 1.290        | 0.630        | 0.990        | 0.225        | 1.125        | 0.295        | 1.220        | 0.220        | 1.275        | 0.400        | 1.180        | 0.354        |
| EmoDS    | 1.375        | 0.685        | 1.210        | 0.395        | 1.200        | 0.340        | 1.225        | 0.345        | 1.260        | 0.535        | 1.254        | 0.460        |
| ECM      | 1.375        | 0.690        | 1.205        | 0.425        | 1.205        | 0.325        | 1.240        | 0.385        | 1.255        | 0.590        | 1.256        | 0.483        |
| CDL      | <b>1.395</b> | <b>0.700</b> | <b>1.245</b> | <b>0.565</b> | <b>1.235</b> | <b>0.490</b> | 1.250        | <b>0.525</b> | <b>1.305</b> | <b>0.630</b> | <b>1.286</b> | <b>0.582</b> |

Table 4: Human evaluation results. ‘‘Con.’’ and ‘‘Emo.’’ denote content and emotion, respectively.

Greedy, Extrema and Coherence)<sup>3</sup> (Liu et al., 2016; Xu et al., 2018); 2) BLEU scores (Papineni et al., 2002) in 0 to 1 scale; 3) Dist-1, Dist-2 (Li et al., 2016a) and 4) Emotion-acc, Emotion-word (Zhou et al., 2018a; Song et al., 2019).

Embedding scores and BLEU scores are used to measure the quality of generated responses in terms of content relevance. Whereas, Dist-1 and Dist-2 are used to evaluate the diversity of responses<sup>4</sup>. Emotion-acc and Emotion-word are utilized to test the emotion expression. Specifically, Emo-acc is the agreement between the ground truth labels and the predicted labels through the TextCNN classifier trained before. Emo-word is the percentage of the generated responses that contain the corresponding emotion words. Since there are no multi-emotion ground truths in the test set, we only calculate the metrics between the ground truth, labelled emotion  $e$ , and the generated response given also label  $e$  for fair comparison.

#### 4.4.2 Human Evaluation Settings

Inspired by Zhou et al. (2018a); Song et al. (2019), a human evaluation is conducted to better analyze the quality of generated responses. First, we randomly sample 200 queries from the test set. For

<sup>3</sup>We use the pre-trained word embeddings based on Sina Weibo data from <https://github.com/Embedding/Chinese-Word-Vectors>.

<sup>4</sup>We employ a popular NLG evaluation project available at <https://github.com/Maluuba/nlg-eval> for automatic evaluation.

each method except S2S-Attn, they generate six responses for six emotion categories, while S2S-Attn generates top 6 responses from beam search decoding for each query. Then, we send the triples of (query, response, emotion) to three human annotators without order, and require them to evaluate each response on both content level and emotion level independently.

Content and emotion are measured by a 3-scale rating (0, 1, 2) and a 2-scale rating (0, 1), respectively. Evaluation from the content level assesses whether a response is fluent, coherent and meaningful for the query, and evaluation from the emotion level decides if a response reveals the desired emotion.

## 4.5 Experimental Results

Now we demonstrate our experimental results on both automatic evaluation and human evaluation.

### 4.5.1 Automatic Evaluation Results

The automatic results are shown in Table 3. The top part is the results of all baseline models, and we can see that CDL outperforms the other methods on all metrics (t-test,  $p$ -value < 0.05). The improvements of CDL on *Coherence*, *Emotion-acc* and *Emotion-word* are significant, indicating that it can enhance content consistency and emotion expression simultaneously. EmoDS and ECM have similar performance, as both of them use the forward method to pay more attention on the emotion

| Query   | Method     |   | Response                                    | Translated Response               |
|---|------------|---|---|-----------------------------------|
| 创意源于生活，看你对生活的态度如何。<br>Creativity comes from life, and depends on your attitude. | S2S-Attn   |   | 生活需要改变心态。                                   | Life needs a change of mindset.   |
|   | ECM        | Like  | 希望有个好心情。                                    | Hope to have a <b>good mood</b> . |
|   |            | Sad   | 我也是很纠结的。                                    | I am also very <b>tangled</b> .   |
|   |            | Disgust   | 你太过分了。                                      | You are <b>too much</b> .         |
|   |            | Angry   | 你怎么还没睡觉?                                    | how come you are still up?        |
|   | Happy      | 哈哈，是啊。  | <b>Haha</b> , yes.                          |                                   |
| CDL   | Like       | 希望每天都有好心情。  | Hope to have a <b>good mood</b> every day.  |                                   |
|   | Sad        | 我觉得我很不淡定。   | I do not think I am calm.                   |                                   |
|   | Disgust    | 别说废话了。  | Stop <b>talking nonsense</b> .              |                                   |
|   | Angry      | 你根本不懂生活!  | You do not understand life at all!          |                                   |
| Happy   | 开心是最重要的。   | <b>Happiness</b> is the most important.             |   |                                   |
| 善良的人还是存在的。<br>Kind people still exist.  | S2S-Attn   |   | 我不做好事。                                      | I do not do good things.          |
|   | ECM        | Like  | 是的，喜欢就好。                                    | Yes, it is ok you <b>like</b> it. |
|   |            | Sad   | 是啊，我很无辜的。                                   | Yeah, I am <b>innocent</b> .      |
|   |            | Disgust   | 不是不是，不是好人。                                  | No, no, not a good person.        |
|   |            | Angry   | 你是什么意思???                                   | What do you mean???               |
|   | Happy      | 哈哈，你也是。   | <b>Haha</b> , you too.                      |                                   |
| CDL   | Like       | 你是不是在夸我?  | Are you <b>praising</b> me?                 |                                   |
|   | Sad        | 可惜我不想做。   | <b>Unfortunately</b> , I do not want to be. |                                   |
|   | Disgust    | 听起来像假话。   | It sounds like a <b>lie</b> .               |                                   |
|   | Angry      | 我一点也没觉得。  | I feel nothing at all.                      |                                   |
| Happy   | 要对生活充满希望。  | Be <b>hopeful</b> about life.                       |   |                                   |
| 我最喜欢的季节就是秋天。<br>Autumn is my favorite season.                                   | S2S-Attn   |   | 我也很喜欢。                                      | I like it, too.                   |
|   | ECM        | Like  | 我也喜欢秋天。                                     | I also <b>like</b> autumn.        |
|   |            | Sad   | 我也想念秋天。                                     | I also <b>miss</b> autumn.        |
|   |            | Disgust   | 你太过分啦。                                      | You are <b>too much</b> .         |
|   |            | Angry   | 你怎么不说我是爱?                                   | Why don't you say that I love it? |
|   | Happy      | 哈哈，我也喜欢。  | <b>Haha</b> , I like it too.                |                                   |
| CDL   | Like       | 秋天的天空非常美丽。  | The autumn sky is very <b>beautiful</b> .   |                                   |
|   | Sad        | 我很想念过去的夏天。  | I <b>miss</b> the past summer very much.    |                                   |
|   | Disgust    | 秋天的景色很难看。   | The autumn scenery is <b>ugly</b> .         |                                   |
|   | Angry      | 你有病吧?   | Are you <b>insane</b> ?                     |                                   |
| Happy   | 哈哈，要不要去秋游? | <b>Haha</b> , do you want to go to the autumn tour? |   |                                   |

Table 5: Sample responses generated by S2S-Attn, ECM and CDL (original Chinese and English translation). The colored words are the emotion words corresponding to the given emotion category.

factor. S2S-Attn can only generate fluent responses based on semantic mapping, but fail to express diverse responses.

The bottom part of Table 3 shows the results of our ablation study. Comparisons among CDL-emo, CDL-con and CDL show the effectiveness of the combined reward for both emotion expression and content consistency. In addition, we can find that with the support of curriculum learning, CDL can achieve better results than CDL-DL.

#### 4.5.2 Human Evaluation Results

The results are shown in Table 4. CDL obtains the best performance (t-test,  $p$ -value  $< 0.05$ ) on both emotion expression (0.582) and content coherence (1.286). As we can see, there is no obvious difference between EmoDS and ECM. Due to the insufficient training data of “Anger” (79,611 in queries and 138,198 in responses), S2S-Attn achieves the best content score for it, which is similar to the

results of Zhou et al. (2018a).

| Method (%) | 2-1         | 1-1         | 0-1 | 2-0  | 1-0  | 0-0  |
|------------|-------------|-------------|-----|------|------|------|
| S2S-Attn   | 10.3        | 7.2         | 2.8 | 36.4 | 26.5 | 16.8 |
| EmoEmb     | 21.8        | 12.6        | 7.5 | 24.6 | 15.3 | 18.2 |
| EmoDS      | 28.7        | 15.6        | 4.0 | 22.7 | 13.5 | 15.5 |
| ECM        | 27.1        | 12.7        | 4.5 | 23.5 | 15.4 | 16.8 |
| CDL        | <b>32.5</b> | <b>17.6</b> | 4.1 | 17.7 | 12.8 | 15.3 |

Table 6: The percentage of responses in human evaluation of *Content-Emotion* scores. 2-1 means content score is 2 and emotion score is 1.

Results of emotion and content in Table 4 are independent. To better evaluate the overall quality of the generated responses, we present results in Table 6 by considering content and emotion scores simultaneously. 32.5% of the responses generated by CDL are annotated with Emotion score 2 and Content score 1, which shows that CDL is better at producing coherent as well as emotion-rich responses.

Agreements to measure the consistency among three annotators are calculated with the Fleiss’ kappa (Fleiss and Cohen, 1973). Fleiss’ kappa for content and emotion is 0.497 and 0.825, indicating “Moderate agreement” and “Substantial agreement”, respectively.

#### 4.6 Case Study

Table 5 shows the examples generated by S2S-Attn, ECM and CDL. As can be seen from it, for a given post, there are multiple emotion categories that are appropriate for its response in the conversation. S2S-Attn generates a response with a random emotion, while ECM and CDL can utilize the specific emotion label. Compared with ECM, CDL can generate both coherent and informative responses with any desired emotion. In addition, the emotion can be expressed in either explicit or implicit manner. For example, “你/根本/不懂/生活! (You do not understand life at all!)” express anger when we read this sentence as a whole, while “美丽 (beautiful)” or “开心 (happy)” are strong emotion words to represent “Like” or “Happy”.

#### 4.7 Further Analysis of CDL

Here, we conduct a further analysis to show some characteristics of this task and the effect of CDL. Emotion lexicon size and classification accuracy after pre-training of each category ( $N(\text{correct prediction}) \div \text{category size}$ ) are listed in Table 7. We can see that the classification accuracy is not totally related to the emotion lexicon size, indicating the emotion expression is partially implicit or explicit. To better illustrate the learning efficiency of CDL, we plot the changes of *Emotion-acc* on the validation set. As shown in Figure 2, CDL accelerates the learning effectively and consistently outperforms CDL-DL.

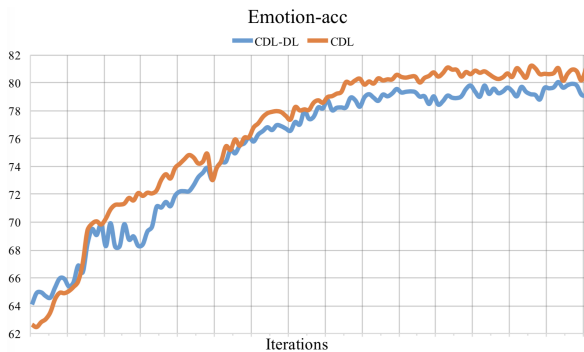


Figure 2: Comparison of CDL and CDL-DL for *Emotion-acc* on the validation set.

|                  | Like  | Sad   | Disgust | Angry | Happy |
|------------------|-------|-------|---------|-------|-------|
| Lex. Size        | 1,629 | 294   | 1,142   | 30    | 405   |
| ACC ( <i>f</i> ) | 0.653 | 0.691 | 0.609   | 0.736 | 0.818 |
| ACC ( <i>b</i> ) | 0.690 | 0.655 | 0.602   | 0.756 | 0.808 |

Table 7: Emotion lexicon size and classification accuracy after pre-training of each emotion category. “Lex.,” “ACC(*f*)” and “ACC(*b*)” represent lexicon, classification accuracy of forward process and classification accuracy of backward process, respectively.

## 5 Related Work

Responses generated by traditional open-domain dialogue systems are usually safe and generic. To produce diverse and informative responses, researchers tried to either import latent variables for model construction (Zhao et al., 2017; Serban et al., 2017; Shen et al., 2019) or utilize some extra knowledge, e.g., sentence types, personas, emotions, documents and knowledge triples/graphs (Ke et al., 2018; Li et al., 2016b; Zhou et al., 2018a; Meng et al., 2019; Zhou et al., 2018b; Niu et al., 2019). In this paper, we mainly touch on two branches of research: emotional response generation and dual learning in NLP.

### 5.1 Emotional Response Generation

Early studies have proven that dialogue systems with proper emotional expressions and reactions can directly improve user satisfaction (Prendinger and Ishizuka, 2005; Prendinger et al., 2005) and contribute to effective users’ performance (Partala and Surakka, 2004). Polzin and Waibel (2000) and Polzin and Waibel (2000) apply rule-based methods to choose emotional responses from a conversation corpus, but those rules are hard to extend to large corpora. With the advent of deep learning, some researchers utilize neural networks to solve this problem (Ghosh et al., 2017; Hu et al., 2017; Zhou and Wang, 2018; Sun et al., 2018). Besides, the Valence, Arousal, and Dominance (VAD) lexicon (Warriner et al., 2013; Mohammad, 2018) is embedded to the sequence-to-sequence model (Sutskever et al., 2014) to provide extra affective information (Asghar et al., 2018; Zhong et al., 2019).

Responses generated by above studies can simply continue the emotion of the query. To generate emotion-controllable responses, Zhou et al. (2018a) address the emotion factor in large-scale conversations, and propose ECM to generate responses based on different given emotions. After that, Colombo et al. (2019) augment ECM with



VAD embeddings and modified the loss function and decoding procedure. Song et al. (2019) use lexicon-based attention and a word-based classifier to improve the ability of emotion expression.

## 5.2 Dual Learning in NLP

He et al. (2016) propose Dual Learning (DL) for machine translation first which consider the source to target language translation and target to source language translation as a dual task. After that, Tang et al. (2017) implement a dual framework for the question answering system. Both Zhang et al. (2018) and Cui et al. (2019) use similar idea in dialogue generation task to produce coherent but not safe responses, since they find that a more diverse and specific response usually has a higher probability of being transformed back to the given query. Luo et al. (2019b) and Luo et al. (2019a) exploit DL in unsupervised text style transfer to relieve the need of parallel data.

The differences between our method and those in Section 5.1 and Section 5.2 are: (1) We consider the emotion expression and content consistency simultaneously via a DL method. (2) Instead of regarding the query as an emotionless sentence, we utilize the emotion of query, which can help model the emotion shifting and coherence to improve the quality of response. (3) To better model the changes in emotion and content between the query and response, we combine the DL method with curriculum learning, which is known to improve the effectiveness and generalization.

## 6 Conclusion

In this paper, we propose a new framework Curriculum Dual Learning (CDL) for generating emotional responses in a controlled manner. Since existing methods in this field only focus on the emotion expression of target label but fail to consider the emotion of queries, the safe response problem deteriorates and hurts the content consistency. CDL utilizes two kinds of rewards to enhance emotion and content simultaneously via dual learning. Besides, with the support of curriculum learning, it can be more efficient. Experimental results show that CDL can generate fluent, coherent, informative as well as emotional responses.

## Acknowledgements

This work was supported by National Key R&D Program of China (NO. 2017YFE0192900). We

sincerely thank the anonymous reviewers for their helpful and valuable suggestions.

## References

- Nabiha Asghar, Pascal Poupart, Jesse Hoey, Xin Jiang, and Lili Mou. 2018. Affective neural response generation. In *European Conference on Information Retrieval*, pages 154–166. Springer.
- Yoshua Bengio, Jérôme Louradour, Ronan Collobert, and Jason Weston. 2009. Curriculum learning. In *Proceedings of the 26th annual international conference on machine learning*, pages 41–48. ACM.
- Hengyi Cai, Hongshen Chen, Cheng Zhang, Yonghao Song, Xiaofang Zhao, Yangxi Li, Dongsheng Duan, and Dawei Yin. 2020. Learning from easy to complex: Adaptive multi-curricula learning for neural dialogue generation. *arXiv preprint arXiv:2003.00639*.
- Pierre Colombo, Wojciech Witon, Ashutosh Modi, James Kennedy, and Mubbasir Kapadia. 2019. Affect-driven dialog generation. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 3734–3743.
- Shaobo Cui, Rongzhong Lian, Di Jiang, Yuanfeng Song, Siqi Bao, and Yong Jiang. 2019. Dal: Dual adversarial learning for dialogue generation. In *Proceedings of the Workshop on Methods for Optimizing and Evaluating Neural Language Generation*, pages 11–20.
- Jessica Fidler and Yoav Goldberg. 2017. Controlling linguistic style aspects in neural language generation. In *Proceedings of the Workshop on Stylistic Variation*, pages 94–104.
- Joseph L Fleiss and Jacob Cohen. 1973. The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. *Educational and psychological measurement*, 33(3):613–619.
- Sayan Ghosh, Mathieu Chollet, Eugene Laksana, Louis-Philippe Morency, and Stefan Scherer. 2017. Affect-lm: A neural language model for customizable affective text generation. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 634–642.
- Di He, Yingce Xia, Tao Qin, Liwei Wang, Nenghai Yu, Tie-Yan Liu, and Wei-Ying Ma. 2016. Dual learning for machine translation. In *Advances in Neural Information Processing Systems*, pages 820–828.
- Zhiting Hu, Zichao Yang, Xiaodan Liang, Ruslan Salakhutdinov, and Eric P Xing. 2017. Toward controlled generation of text. In *Proceedings*

- of the 34th International Conference on Machine Learning-Volume 70, pages 1587–1596. JMLR. org.
- Pei Ke, Jian Guan, Minlie Huang, and Xiaoyan Zhu. 2018. Generating informative responses with controlled sentence function. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1499–1508.
- Yoon Kim. 2014. Convolutional neural networks for sentence classification. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1746–1751.
- Diederick P Kingma and Jimmy Ba. 2015. Adam: A method for stochastic optimization. In *the 3rd International Conference on Learning Representations*.
- Alex M Lamb, Anirudh Goyal Alias Parth Goyal, Ying Zhang, Saizheng Zhang, Aaron C Courville, and Yoshua Bengio. 2016. Professor forcing: A new algorithm for training recurrent networks. In *Advances In Neural Information Processing Systems*, pages 4601–4609.
- Jiwei Li, Michel Galley, Chris Brockett, Jianfeng Gao, and Bill Dolan. 2016a. A diversity-promoting objective function for neural conversation models. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 110–119.
- Jiwei Li, Michel Galley, Chris Brockett, Georgios Spithourakis, Jianfeng Gao, and Bill Dolan. 2016b. A persona-based neural conversation model. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 994–1003.
- Jiwei Li, Will Monroe, Alan Ritter, Dan Jurafsky, Michel Galley, and Jianfeng Gao. 2016c. Deep reinforcement learning for dialogue generation. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pages 1192–1202.
- Jiwei Li, Will Monroe, Tianlin Shi, Sébastien Jean, Alan Ritter, and Dan Jurafsky. 2017. Adversarial learning for neural dialogue generation. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 2157–2169.
- Chia-Wei Liu, Ryan Lowe, Iulian Serban, Mike Noseworthy, Laurent Charlin, and Joelle Pineau. 2016. How not to evaluate your dialogue system: An empirical study of unsupervised evaluation metrics for dialogue response generation. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pages 2122–2132.
- Fuli Luo, Peng Li, Pengcheng Yang, Jie Zhou, Yutong Tan, Baobao Chang, Zhifang Sui, and Xu Sun. 2019a. Towards fine-grained text sentiment transfer. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 2013–2022.
- Fuli Luo, Peng Li, Jie Zhou, Pengcheng Yang, Baobao Chang, Zhifang Sui, and Xu Sun. 2019b. A dual reinforcement learning framework for unsupervised text style transfer. *arXiv preprint arXiv:1905.10060*.
- Chuan Meng, Pengjie Ren, Zhumin Chen, Christof Monz, Jun Ma, and Maarten de Rijke. 2019. Refnet: A reference-aware network for background based conversation. *arXiv preprint arXiv:1908.06449*.
- Saif Mohammad. 2018. Obtaining reliable human ratings of valence, arousal, and dominance for 20,000 english words. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 174–184.
- Zheng-Yu Niu, Hua Wu, Haifeng Wang, et al. 2019. Knowledge aware conversation generation with explainable reasoning over augmented graphs. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 1782–1792.
- Kishore Papineni, Salim Roukos, Todd Ward, and Wei-Jing Zhu. 2002. Bleu: a method for automatic evaluation of machine translation. In *Proceedings of the 40th annual meeting on association for computational linguistics*, pages 311–318. Association for Computational Linguistics.
- Timo Partala and Veikko Surakka. 2004. The effects of affective interventions in human–computer interaction. *Interacting with computers*, 16(2):295–309.
- Ramakanth Pasunuru and Mohit Bansal. 2018. Multi-reward reinforced summarization with saliency and entailment. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2 (Short Papers)*, pages 646–653.
- Romain Paulus, Caiming Xiong, and Richard Socher. 2017. A deep reinforced model for abstractive summarization. *arXiv preprint arXiv:1705.04304*.
- Emmanouil Antonios Platanios, Otilia Stretcu, Graham Neubig, Barnabas Poczos, and Tom Mitchell. 2019. Competence-based curriculum learning for neural machine translation. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 1162–1172.
- Thomas S Polzin and Alexander Waibel. 2000. Emotion-sensitive human-computer interfaces. In *ISCA tutorial and research workshop (ITRW) on speech and emotion*.

- Helmut Prendinger and Mitsuru Ishizuka. 2005. The empathic companion: A character-based interface that addresses users' affective states. *Applied Artificial Intelligence*, 19(3-4):267–285.
- Helmut Prendinger, Junichiro Mori, and Mitsuru Ishizuka. 2005. Using human physiology to evaluate subtle expressivity of a virtual quizmaster in a mathematical game. *International journal of human-computer studies*, 62(2):231–245.
- Marc Aurelio Ranzato, Sumit Chopra, Michael Auli, and Wojciech Zaremba. 2015. Sequence level training with recurrent neural networks. *arXiv preprint arXiv:1511.06732*.
- Iulian Vlad Serban, Alessandro Sordani, Ryan Lowe, Laurent Charlin, Joelle Pineau, Aaron Courville, and Yoshua Bengio. 2017. A hierarchical latent variable encoder-decoder model for generating dialogues. In *Thirty-First AAAI Conference on Artificial Intelligence*.
- Lifeng Shang, Zhengdong Lu, and Hang Li. 2015. Neural responding machine for short-text conversation. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 1577–1586.
- Lei Shen, Yang Feng, and Haolan Zhan. 2019. Modeling semantic relationship in multi-turn conversations with hierarchical latent variables. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 5497–5502.
- Zhenqiao Song, Xiaoqing Zheng, Lu Liu, Mu Xu, and Xuan-Jing Huang. 2019. Generating responses with a specific emotion in dialog. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 3685–3695.
- Xiao Sun, Xinmiao Chen, Zhengmeng Pei, and Fujii Ren. 2018. Emotional human machine conversation generation based on seqgan. In *2018 First Asian Conference on Affective Computing and Intelligent Interaction (ACII Asia)*, pages 1–6. IEEE.
- I Sutskever, O Vinyals, and QV Le. 2014. Sequence to sequence learning with neural networks. *Advances in NIPS*.
- Duyu Tang, Nan Duan, Tao Qin, Zhao Yan, and Ming Zhou. 2017. Question answering and question generation as dual tasks. *arXiv preprint arXiv:1706.02027*.
- Amy Beth Warriner, Victor Kuperman, and Marc Brysbaert. 2013. Norms of valence, arousal, and dominance for 13,915 english lemmas. *Behavior research methods*, 45(4):1191–1207.
- Ronald J Williams. 1992. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine learning*, 8(3-4):229–256.
- Xinnuo Xu, Ondřej Dušek, Ioannis Konstas, and Verena Rieser. 2018. Better conversations by modeling, filtering, and optimizing for coherence and diversity. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 3981–3991.
- Wojciech Zaremba and Ilya Sutskever. 2015. Reinforcement learning neural turing machines-revised. *arXiv preprint arXiv:1505.00521*.
- Hainan Zhang, Yanyan Lan, Jiafeng Guo, Jun Xu, and Xueqi Cheng. 2018. Reinforcing coherence for sequence to sequence model in dialogue generation. In *IJCAI*, pages 4567–4573.
- Tiancheng Zhao, Ran Zhao, and Maxine Eskenazi. 2017. Learning discourse-level diversity for neural dialog models using conditional variational autoencoders. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 654–664.
- Peixiang Zhong, Di Wang, and Chunyan Miao. 2019. An affect-rich neural conversational model with biased attention and weighted cross-entropy loss. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pages 7492–7500.
- Hao Zhou, Minlie Huang, Tianyang Zhang, Xiaoyan Zhu, and Bing Liu. 2018a. Emotional chatting machine: Emotional conversation generation with internal and external memory. In *Thirty-Second AAAI Conference on Artificial Intelligence*.
- Hao Zhou, Tom Young, Minlie Huang, Haizhou Zhao, Jingfang Xu, and Xiaoyan Zhu. 2018b. Commonsense knowledge aware conversation generation with graph attention. In *IJCAI*, pages 4623–4629.
- Xianda Zhou and William Yang Wang. 2018. Mojtalk: Generating emotional responses at scale. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1128–1137.