ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) | 978-1-7281-6327-7/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICASSP49357.2023.10097068

CONTRASTIVE LEARNING WITH DIALOGUE ATTRIBUTES FOR NEURAL DIALOGUE GENERATION

Jie Tan[1], *Hengyi Cai*[2], *Hongshen Chen*[2], *Hong Cheng*[1], *Helen Meng*[1], *Zhuoye Ding*[2]

[1]The Chinese University of Hong Kong, China, [2]Data Science Lab, JD.com, China

{jtan, hcheng, hmmeng}@se.cuhk.edu.hk,hengyi1995@gmail.com,ac@chenhongshen.com,dingzhuoye@jd.com

ABSTRACT

Designing an effective learning method remains a challenge in neural dialogue generation systems as it requires the training objective to well approximate the intrinsic human-preferred dialogue properties. Conventional training approaches such as maximum likelihood estimation focus on modeling general syntactic patterns and may fail to capture intricate conversational characteristics. Contrastive dialogue learning offers an effective training schema by explicitly training a neural dialogue model on multiple positive and negative conversational pairs. However, constructing contrastive learning pairs is non-trivial, and multiple dialogue attributes have been found to be crucial for governing the human judgments of conversations. This paper proposes to guide the response generation with attribute-aware contrastive learning to improve the overall quality of the generated responses, where contrastive learning samples are generated according to various important dialogue attributes each specializing in a different principle of conversation. Extensive experiments show that our proposed techniques are crucial to achieving superior model performance.

***Index Terms*—** Dialogue Generation, Contrastive Learning, Conversational Attributes and Adversarial Perturbations

1. INTRODUCTION

Recent years have witnessed a surge of interests in neural dialogue generation [1, 2, 3, 4, 5] Typically, dialogue generation models utilize the given conversational histories as context to construct the response via maximum likelihood estimation (MLE). Though effective, some issues, including limited informativeness in generated responses, are observed on dialogue models learned by MLE [6, 7]. Endeavors combating such defects attempt to feed extra information, for example, topics [8], personas [2], and external knowledge [9], to the dialogue model, or enhance the model itself using latent variables [10].

Another line of research focuses on improving the response quality by redesigning the learning scheme of dialogue models [6, 7, 11, 12]. Among these efforts, one promising training strategy for improving the neural dialogue generation is contrastive learning [13, 14, 15]. The key idea is to incentivize the dialogue model to explicitly perceive the difference between positive and negative utterances, and thus enhance the distinctiveness of the generated responses.

However, naively contrasting the given response may result in a sub-optimal solution, and the negative utterances sometimes are inappropriately sampled without considering the intrinsic dialogue properties. Directly applying the contrastive learning to the dialogue generation task faces two issues. **First**, while it suffices to measure general conversational matching relations by minimizing the distance between positive context-response pairs and maximizing differences of negative pairs, the vanilla contrastive learning approach fails to capture multi-modality conversational aspects including response specificity and text repetitions, which affects the overall conversation quality and requires deliberate modeling to improve the dialogue model performance [16]. **Second**, the conventional contrastive learning method barely generates positive and negative samples conditioned on the limited training set. Since the key ingredient of contrastive dialogue learning is to leverage the discrimination of the positive samples against the negative distribution to guide the training, being able to create informative contrastive samples which largely exhibit the conversational attributes acts as the key driver to facilitate dialogue learning. However, the discrete nature of text makes it challenging to establish universal rules for generating effective contrastive samples, let alone concerning various conversational aspects during the contrastive sampling procedure.

In this paper, we propose an attribute-aware contrastive dialogue learning framework that is capable of contrasting training examples regarding a committee of conversational aspects, and enhancing the generated response towards the desired polarity along various attribute axes. As shown in Fig. 1, our model consists of an encoder-decoder dialogue generation model and a series of attribute indicators. We first train the encoder-decoder and indicators and get the latent representation of the input example. Different from the previous approaches that blindly draw negatives without considering intrinsic human-preferred conversational properties, the contrastive samples in this work can be obtained by iteratively performing perturbations on the latent representation, until it can be identified as target polarity by the attribute indicator. We then leverage these generated latent examples to conduct contrastive learning, helping the dialogue model learn with more informative pairs. Empirical results confirm that responses generated by our approach are preferred over those by competitive baselines on two public datasets.

2. METHODOLOGY

2.1. Encoder-Decoder Backbone

Given an input context $\mathbf{x} = \{x_1, x_2, \cdots, x_{T_\mathbf{x}}\}$, the dialogue generation model aims to generate an appropriate response $\mathbf{y} = \{y_1, y_2, \cdots, y_{T_\mathbf{y}}\}$. Given a training set $\mathbb{D} = \{(\mathbf{x}, \mathbf{y})_i\}_{i=1}^N$, the neural dialogue model parameterizes the conditional distribution $p_\theta(\mathbf{y}|\mathbf{x})$ as follows:

$$\begin{aligned} p_\theta(\mathbf{y} \mid \mathbf{x}) &= \prod_{t=1}^{T_\mathbf{y}} p_\theta\left(y_t \mid \mathbf{y}_{<t}, \mathbf{x}\right) \\ p_\theta\left(y_t \mid \mathbf{y}_{<t}, \mathbf{x}\right) &= \mathbf{softmax}\left(\boldsymbol{W}_o \mathbf{h}_t + \boldsymbol{b}_o\right) \\ \mathbf{h}_t &= g\left(y_{t-1}, \mathbf{H}_\mathbf{x}; \boldsymbol{\theta}\right), \mathbf{H}_\mathbf{x} = f\left(\mathbf{x}; \boldsymbol{\theta}\right), \end{aligned} \tag{1}$$

where $f(\cdot)$ and $g(\cdot)$ represent the encoder and decoder respectively, $\boldsymbol{W}_o$ and $\boldsymbol{b}_o$ are trainable parameters, and $\boldsymbol{\theta}$ is the parameter set of the dialogue generation model. $\mathbf{H}_\mathbf{x}$ and $\mathbf{H}_\mathbf{y} = \left[\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_{T_\mathbf{y}}\right]$ denote the concatenation of encoder and decoder hidden states respectively. By adopting MLE as the learning approach, the training objective of the neural dialogue model can be defined as follows:

$$\mathcal{L}_{\text{MLE}}(\boldsymbol{\theta}) = -\sum_{(\mathbf{x},\mathbf{y})\in\mathbb{D}} \sum_{t=1}^{T_\mathbf{y}} \log p_\theta\left(y_t \mid \mathbf{y}_{<t}, \mathbf{x}\right). \tag{2}$$

2.2. Attribute Indicator

In order to draw effective contrastive samples that exhibit specific conversational attribute $a_i \in \{a_1, a_2, \cdots, a_\mathcal{K}\}$, we first design a committee of conversational attribute indicators, $\{\mathcal{A}_i\}_{i=1}^{\mathcal{K}}$, to predict the polarity of the input example regarding a specific dialogue aspect, and then generate positive and negative samples for contrastive learning by iteratively performing perturbations on the latent dialogue representation, until the representation can be identified as target polarity by the attribute indicator. The attribute indicator $\mathcal{A}_i$ predicts the corresponding binary attribute polarity $\mathbf{z}_{a_i}$ as follows:

$$p\left(\mathbf{z}_{a_i} \mid \mathbf{x}, \mathbf{y}\right) = \mathbf{sigmoid}\left(\mathcal{A}_i\left(\left[\psi\left(\mathbf{H}_\mathbf{x}\right); \psi\left(\mathbf{H}_\mathbf{y}\right)\right]\right)\right) \tag{3}$$

where $\mathcal{A}_i$ can be implemented as a multilayer feed-forward network, $[;]$ denotes the concatenation operation, ψ represents the linear transformation followed by an average pooling.

Assuming $\boldsymbol{\phi}_{a_i}$ is the parameter set of a specific conversational attribute indicator, we define the training objective as

$$\mathcal{L}_{att} = \sum_{(\mathbf{x},\mathbf{y})\in\mathbb{D}} \sum_{i=1}^{\mathcal{K}} -\log p\left(\mathbf{z}_{a_i} \mid \mathbf{H}_\mathbf{x}, \mathbf{H}_\mathbf{y}; \boldsymbol{\phi}_{a_i}\right). \tag{4}$$

We exploit four fine-grained dialogue attributes in this work: 1) **Specificity (Spe.)** measures "level of detail" of a sentence [17]. We calculate its value is based on the normalized inverse response frequency [18, 16]; 2) **Repetitiveness (Rep.)**: inappropriate repetition of words significantly degrades the quality of the dialogue response. We estimate the repetitiveness of a given input utterance by identifying repeating content words following [16]; 3) **Relatedness (Rel.)** quantifies how well a dialogue response semantically relates to its context and is assessed using the cosine similarities between the response and its corresponding context representations; 4) **Informativeness (Inf.)**: good responses are characterized by communicating comprehensive information in the responding utterance. We measure such property by dividing the sentence entropy value by its length. We discretize the aforementioned attribute values into two bins ("positive" and "negative").

2.3. Contrastive Dialogue Learning with Attribute-aware Adversarial Perturbations

Generation of Negative Examples The goal of generating an attribute-aware negative example is to find an optimal response representation that is very close to $\mathbf{H}_\mathbf{y}$ but performs negative on the corresponding attribute. As shown in Fig. 1, taking the attribute "Relatedness" as an example, we add a small perturbation $\mathbf{r} = \left[\boldsymbol{r}_1, \boldsymbol{r}_2, \cdots, \boldsymbol{r}_{T_\mathbf{y}}\right]$ to $\mathbf{H}_\mathbf{y}$ to generate a negative sample, such that its probability of performing positive on the attribute "Relatedness", i.e., $p\left(\mathbf{z}_{\text{Rel.}} = 1 \mid \mathbf{H}_\mathbf{x}, \mathbf{H}_\mathbf{y} + \mathbf{r}; \boldsymbol{\phi}_{\text{Rel.}}\right)$, is minimized. Formally, the optimal perturbation is given by

$$\mathbf{r}^* = \underset{\mathbf{r}, \|\mathbf{r}\| \leq \epsilon}{\arg\min} \log p(\mathbf{z} = 1 \mid \mathbf{H}_\mathbf{x}, \mathbf{H}_\mathbf{y} + \mathbf{r}; \hat{\boldsymbol{\phi}}), \tag{5}$$

where ϵ is the perturbation radius, and $\hat{\boldsymbol{\phi}}$ denotes a copy of the indicator parameters $\boldsymbol{\phi}$ that is deemed to be constant during the process of negative sample construction. We omit the attribute subscript for brevity.

However, the exact optimal solution $\mathbf{r}^*$ to the minimization objective in Eq.(5) is computationally intractable. Following [19], we use the linear approximation of this value by linearizing $\log p\left(\mathbf{z} = 1 \mid \mathbf{H}_\mathbf{x}, \mathbf{H}_\mathbf{y}\right)$ around $\mathbf{H}_\mathbf{y}$ and the negative example can be then defined as:

$$\widetilde{\mathbf{H}}_\mathbf{y} = \mathbf{H}_\mathbf{y} - \frac{\epsilon \mathbf{g}}{\|\mathbf{g}\|_2}, \ \ \mathbf{g} = \nabla_{\mathbf{H}_\mathbf{y}} \log p(\mathbf{z} = 1 \mid \mathbf{H}_\mathbf{x}, \mathbf{H}_\mathbf{y}; \hat{\boldsymbol{\phi}}). \tag{6}$$

Generation of Positive Examples We employ a two-stage approach to generate an attribute-aware positive response representation performing positive on the corresponding conversational attribute while preserving the original semantic information. As shown in Figure 1, we first add a perturbation with radius γ on $\mathbf{H}_\mathbf{y}$ such that it maximizes the probability of performing positive on the given conversational attribute:

$$\overline{\mathbf{H}}_\mathbf{y} = \mathbf{H}_\mathbf{y} + \frac{\gamma \mathbf{q}}{\|\mathbf{q}\|_2}, \ \ \mathbf{q} = \nabla_{\mathbf{H}_\mathbf{y}} \log p(\mathbf{z} = 1 \mid \mathbf{H}_\mathbf{x}, \mathbf{H}_\mathbf{y}; \hat{\boldsymbol{\phi}}). \tag{7}$$

Then we add another perturbation on $\overline{\mathbf{H}}_\mathbf{y}$ to ensure the perturbed sample maintain the same semantic information, by approximately minimizing the following $\mathcal{KL}$ divergence between the original and perturbed conditional distribution:

$$\mathcal{L}_{kl}(\boldsymbol{\theta}) = \sum_{(\mathbf{x},\mathbf{y})\in\mathbb{D}} \sum_{t=1}^{T_\mathbf{y}} \mathcal{KL}\left(p_{\hat{\theta}}\left(y_t \mid \mathbf{y}_{<t}, \mathbf{x}\right) \| p_\theta\left(\overline{y}_t \mid \overline{\mathbf{y}}_{<t}, \mathbf{x}\right)\right) \tag{8}$$

where $\hat{\theta}$ is the constant copy of the dialogue model parameters. Consequently, the positive sample is finally defined to be $\ddot{\mathbf{H}}_\mathbf{y} = \overline{\mathbf{H}}_\mathbf{y} - \gamma \mathbf{p}/\|\mathbf{p}\|_2$, $\mathbf{p} = \nabla_{\overline{\mathbf{H}}_\mathbf{y}} \mathcal{L}_{kl}(\boldsymbol{\theta})$.

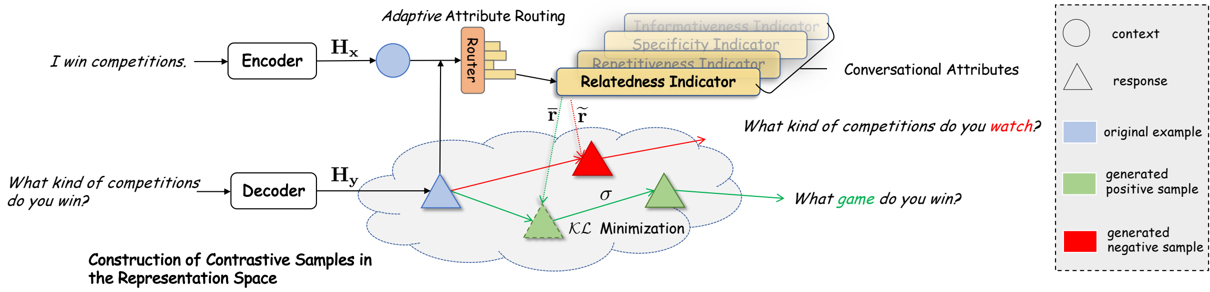

Fig. 1: Our proposed framework first projects the context and response to its latent representation, and then generates an attribute-aware contrastive pair under the guidance of the attribute indicator. The dialogue generation model is finally trained to jointly maximize the log likelihood and the contrastive learning objective with the generated attribute-aware contrastive pairs.

Contrastive Dialogue Learning Objective The naive contrastive learning framework trains the dialogue model to learn the representations of the ground-truth response by contrasting the positive pairs with the negatives as follows:

$$\mathcal{L}_{\text{naive}}(\boldsymbol{\theta}) = -\sum_{(\mathbf{x},\mathbf{y})\in\mathbb{D}} \log \frac{\mathcal{F}(\mathbf{H_x}, \mathbf{H_y})}{\sum_{\mathbf{H}'_\mathbf{y}\in\mathcal{S}\cup\{\mathbf{H_y}\}} \mathcal{F}(\mathbf{H_x}, \mathbf{H}'_\mathbf{y})}, \tag{9}$$

where $\mathcal{S}$ is a collection of negative examples that are in the same batch not paired with the corresponding context. With the temperature parameter τ and cosine similarity function $\mathbf{sim}(.,.)$, $\mathcal{F}$ defines the similarity between the sentence representation pair:

$$\mathcal{F}(\mathbf{H_x}, \mathbf{H_y}) = \mathbf{exp}\left(\mathbf{sim}\left(\psi(\mathbf{H_x}), \psi(\mathbf{H_y})\right)/\tau\right). \tag{10}$$

To encourage the dialogue generation model to identify the conversation-centric features during contrastive learning, we further consider the generated attribute-aware negative sample $\widetilde{\mathbf{H}}_\mathbf{y}$ as an additional negative to perform contrastive learning:

$$\mathcal{L}_{neg}(\boldsymbol{\theta}) = -\sum_{(\mathbf{x},\mathbf{y})\in\mathbb{D}} \log \frac{\mathcal{F}(\mathbf{H_x}, \mathbf{H_y})}{\sum_{\mathbf{H}'_\mathbf{y}\in\mathcal{S}\cup\{\mathbf{H_y}, \widetilde{\mathbf{H}}_\mathbf{y}\}} \mathcal{F}(\mathbf{H_x}, \mathbf{H}'_\mathbf{y})}. \tag{11}$$

With the attribute-aware positive examples, it is straightforward to contrast them from the negatives in Eq.(11) to further induce the model to distinguish the high-quality response with desired dialogue attributes from the less plausible ones. We derive the following training objective by incorporating $\ddot{\mathbf{H}}_\mathbf{y}$:

$$\mathcal{L}_{pos}(\boldsymbol{\theta}) = -\sum_{(\mathbf{x},\mathbf{y})\in\mathbb{D}} \log \frac{\mathcal{F}\left(\mathbf{H_x}, \ddot{\mathbf{H}}_\mathbf{y}\right)}{\sum_{\mathbf{H}'_\mathbf{y}\in\mathcal{S}\cup\{\ddot{\mathbf{H}}_\mathbf{y}, \widetilde{\mathbf{H}}_\mathbf{y}\}} \mathcal{F}(\mathbf{H_x}, \mathbf{H}'_\mathbf{y})}. \tag{12}$$

Finally, the training objective of our proposed learning framework can be formulated as:

$$\mathcal{L} = \mathcal{L}_{pos} + \mathcal{L}_{neg} + \mathcal{L}_{\text{MLE}} + \mathcal{L}_{kl} + \mathcal{L}_{att}. \tag{13}$$

Adaptive Attribute Routing To make attribute-aware contrastive learning more instance-adaptive and effective, we further employ a routing mechanism to adaptively select the conversational attribute for constructing contrastive examples based on the incoming context-response pair, inspired by [20]. The router takes the aggregation representation of a context-response pair as input and then routes it to $\mathcal{K}$ attributes. In this work, we implement the router as a multi-layer feed-forward network, by utilizing a differentiable gumbel-softmax [21] approximation when selecting the result dialogue attribute.

3. EXPERIMENT SETTINGS

Datasets: DailyDialog [22] and PersonaChat [2].
Baseline Models: (i) **HRED**: a generalized sequence-to-sequence model with a hierarchical RNN encoder [23]; (ii) **TRANSFORMER**: an encoder-decoder architecture relying on attention mechanisms [24]; (iii) **BLENDER**: a large-scale pretrained open-domain dialogue model [25].
Comparison Learning Approaches: (i) **ADVERSARIAL**: an adversarial training framework [1]; (ii) **MMI**: a training objective which maximums the mutual information between the context and response [6, 26]; (iii) **DEEPRL**: a reinforcement learning framework with heuristic reward functions [27]; (iv) **CVAE**: a conditional variational auto-encoder learning framework [10] ; (v) **DIALOGWAE**: a conditional Wasserstein auto-encoder framework [11]; (vi) **GCL**: a contrastive learning framework which conducts contrastive learning between a target dialogue model and a pretrained reference model [12].
Automatic Evaluation Metrics: (i) BLEU [28]; (ii) Dist-1 and Dist-2 [6] that evaluate the diversity of the generated responses; (iii) Entropy-based metrics (Ent-1, Ent-2) [29] that show the empirical n-gram distribution of the generated responses.
Implementation and Reproducibility: For HRED, the utterance encoder, context encoder and decoder are bidirectional GRUs [30] with 300 hidden units in each direction. Regarding the TRANSFORMER model, following [31], we set its hidden size, attention heads and the number of hidden layers as 300, 2 and 2, respectively, and use pre-trained Glove vectors [32] to initialize the word embeddings. Regarding the BLENDER model, we use its small version with 90M parameters provided by HuggingFace [33]. For optimization, we use Adam [34] with a learning rate of 0.001 with gradient clipping.

4. RESULTS AND DISCUSSION

Performance on Baseline Models We evaluate the performance of our proposed framework using several baselines. From the results in Table 1, we observe that: 1) Our proposed framework brings improvements across various model architectures regarding most evaluation metrics; 2) The diversity of the responses is promisingly high when applying our proposed approach to BLENDER, indicating that conducting the attribute-aware contrastive learning on the pre-trained dialogue

Models	DailyDialog					PersonaChat				
	BLEU	Dist-1	Dist-2	Ent-1	Ent-2	BLEU	Dist-1	Dist-2	Ent-1	Ent-2
HRED	1.63	0.79	2.43	6.69	10.24	2.10	0.25	0.86	6.74	10.44
HRED (♠)	**1.73**	**0.84**	**2.73**	**6.72**	**10.25**	**2.24**	**0.30**	**1.06**	**6.80**	**10.60**
TRANSFORMER	2.06	0.79	2.99	6.73	10.25	2.05	0.31	1.34	7.00	10.96
TRANSFORMER (♠)	**2.41**	**1.24**	**6.00**	**6.96**	**10.90**	**2.09**	**0.52**	**2.26**	**7.01**	**11.07**
BLENDER	0.98	5.74	21.98	5.62	6.50	**1.95**	1.78	6.12	5.92	7.54
BLENDER (♠)	**2.46**	**7.72**	**32.32**	**6.48**	**8.34**	1.39	**4.19**	**21.97**	**6.56**	**8.77**

Table 1: Automatic evaluation results (%). "♠" denotes that the model is trained using our proposed framework. Significant results in each group are highlighted with **bold**.

Ablation Study		BLEU	Dist-1	Dist-2	Ent-1	Ent-2
Dialogue Attributes	None	2.06	0.79	2.99	6.73	10.25
	(a) Specificity	2.19	0.99	4.08	6.84	10.60
	(b) Repetitiveness	2.07	0.87	3.11	6.76	10.45
	(c) Relatedness	2.31	0.89	3.47	6.84	10.62
	(d) Informativeness	2.09	1.08	4.59	6.87	10.66
Framework Variants	*w/o* $\mathcal{L}_{neg}$	2.24	0.98	4.15	6.87	10.63
	w/o $\mathcal{L}_{pos}$	2.28	0.94	4.26	6.86	10.70
	Naive Contrastive Learning	2.36	0.89	3.60	6.87	10.58
	w/o Router	2.35	0.84	3.41	6.91	10.63
FULL version		**2.41**	**1.24**	**6.00**	**6.96**	**10.90**

Table 2: Ablation test (%) on DailyDialog.

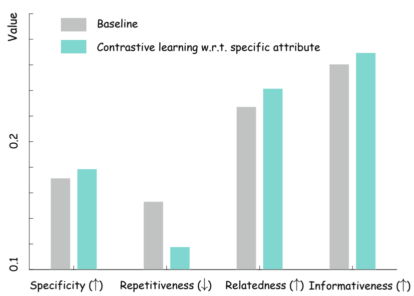

Fig. 2: The effect of individual dialogue attribute.

representations can steer more informative expressions; 3) We also notice that the performance gain regarding the BLEU metric is relatively limited. We anticipate that since our proposed approach is able to construct semantically valid positive training examples during dialogue learning, it implicitly exposes the model to various meaningful outputs beyond the ground-truth responses and induces the dialogue model to generate more diverse utterances.

Ablation Study on Various Dialogue Attributes The results with TRANSFORMER are shown in Table 2. For each ablation, we use only one conversational attribute for constructing the contrastive samples. We observe that drawing contrastive examples regarding single attribute is also beneficial and the nearly best performance is achieved with multiple attributes. Fig. 2 presents the performance gain on specific attribute when constructing the contrastive examples with corresponding attribute indicator, which indicates that the proposed learning scheme does help the model to improve the response quality along various conversational aspects contrapuntally.

Ablation Study on Framework Variants From Table 2, we observe that: 1) Removing either the attribute-aware negative contrastive loss or the attribute-aware positive contrastive loss leads to a performance drop regarding most evaluation metrics, indicating that using the attribute-aware negative or positive sample alone is insufficient, as only one direction of guidance is provided; 2) Naively conducting the contrastive learning using in-batch negatives leads to significant performance degradation, proving the effectiveness of constructing contrastive instances with attribute-aware adversarial perturbations; 3) Disabling the adaptive attribute routing strategy results in a large decrease in model performance, indicating that choosing the instance-adaptive conversational attribute for contrastive sample generation benefits model training.

Learning Approach	Learning Effectiveness(%)					Human Evaluation			
	BLEU	Dist-1	Dist-2	Ent-1	Ent-2	**Win**	**Loss**	**Tie**	**Kappa**
ADVERSARIAL	1.27	0.57	1.46	6.64	10.00	53%	8%	39%	0.5393
MMI	1.46	0.89	4.07	6.71	10.38	48%	11%	41%	0.5837
DEEPRL	1.42	0.68	2.22	6.57	10.18	52%	9%	39%	0.5632
CVAE	1.23	0.92	4.12	6.76	10.53	46%	8%	46%	0.5196
DIALOGWAE	1.50	1.20	5.35	7.15	11.29	40%	12%	48%	0.5372
GCL	1.49	1.05	5.91	6.89	10.70	47%	10%	43%	0.5381
OURS	2.41	1.24	6.00	6.96	10.90	-	-	-	-

Table 3: Learning effectiveness (%) of dialogue training approaches and human evaluation results on DailyDialog.

Comparison with Various Dialogue Learning Approaches As shown in Table 3, we observe that the proposed learning schema outperforms all the previous dialogue training approaches on the majority of evaluation metrics. Note that naively contrasting the given response with sampled negative utterances in the limited training set underperforms our approach, verifying the effectiveness of constructing contrastive examples with attribute-aware adversarial perturbations.

Human Evaluation We randomly sample 100 test messages from DailyDialog to conduct human evaluations, by presenting the input context and the corresponding responses generated by our model and the comparison model to the annotators, and asking them to compare the quality of the two responses in terms of coherence, language consistency, fluency and informativeness. From Table 3, we notice that the dialogue model trained using our approach is more adept to produce human-preferred responses, and the kappa scores indicate that the annotators came to a fair agreement on judgment.

5. CONCLUSION

In this paper, we propose an attribute-aware contrastive training framework, which learns to improve response generation by explicitly capturing human-concerned conversation aspects contrastively and thus is able to induce the dialogue model to generate responses towards a particular attribute polarity correspondingly. Experiments show that the proposed framework can boost the performance of existing dialogue systems.

6. ACKNOWLEDGEMENTS

This research was fully supported by the Centre for Perceptual and Interactive Intelligence (CPII) Ltd under the Innovation and Technology Commissions' InnoHK program.

References

[1] Jiwei Li, Will Monroe, Tianlin Shi, Sébastien Jean, Alan Ritter, and Dan Jurafsky, "Adversarial learning for neural dialogue generation," in *EMNLP*, 2017, pp. 2157–2169.

[2] Saizheng Zhang, Emily Dinan, Jack Urbanek, Arthur Szlam, Douwe Kiela, and Jason Weston, "Personalizing dialogue agents: I have a dog, do you have pets too?," in *ACL*, 2018, pp. 2204–2213.

[3] Minlie Huang, Xiaoyan Zhu, and Jianfeng Gao, "Challenges in building intelligent open-domain dialog systems," *ACM Trans. Inf. Syst.*, vol. 38, pp. 21:1–21:32, 2020.

[4] Zhou Yu, Alan W. Black, and Alexander I. Rudnicky, "Learning conversational systems that interleave task and non-task content," in *IJCAI*, 2017, pp. 4214–4220.

[5] Lei Shen, Haolan Zhan, Xin Shen, and Yang Feng, "Learning to select context in a hierarchical and global perspective for open-domain dialogue generation," in *ICASSP*, 2021, pp. 7438–7442.

[6] Jiwei Li, Michel Galley, Chris Brockett, Jianfeng Gao, and Bill Dolan, "A diversity-promoting objective function for neural conversation models," in *NAACL*, 2016, pp. 110–119.

[7] Hainan Zhang, Yanyan Lan, Jiafeng Guo, Jun Xu, and Xueqi Cheng, "Reinforcing coherence for sequence to sequence model in dialogue generation.," in *IJCAI*, 2018, pp. 4567–4573.

[8] Chen Xing, Wei Wu, Yu Wu, Jie Liu, Yalou Huang, Ming Zhou, and Wei-Ying Ma, "Topic aware neural response generation," in *AAAI*, 2017, pp. 3351–3357.

[9] Emily Dinan, Stephen Roller, Kurt Shuster, Angela Fan, Michael Auli, and Jason Weston, "Wizard of wikipedia: Knowledge-powered conversational agents," in *ICLR*, 2019.

[10] Tiancheng Zhao, Ran Zhao, and Maxine Eskenazi, "Learning discourse-level diversity for neural dialog models using conditional variational autoencoders," in *ACL*, 2017, pp. 654–664.

[11] Xiaodong Gu, Kyunghyun Cho, Jung-Woo Ha, and Sunghun Kim, "Dialogwae: Multimodal response generation with conditional wasserstein auto-encoder," in *ICLR*, 2019.

[12] Hengyi Cai, Hongshen Chen, Yonghao Song, Zhuoye Ding, Yongjun Bao, Weipeng Yan, and Xiaofang Zhao, "Group-wise contrastive learning for neural dialogue generation," in *EMNLP*, 2020, pp. 793–802.

[13] Raia Hadsell, Sumit Chopra, and Yann LeCun, "Dimensionality reduction by learning an invariant mapping," in *CVPR*, 2006, pp. 1735–1742.

[14] Aäron van den Oord, Yazhe Li, and Oriol Vinyals, "Representation learning with contrastive predictive coding," *CoRR*, vol. abs/1807.03748, 2018.

[15] Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton, "A simple framework for contrastive learning of visual representations," in *ICML*, 2020, pp. 1597–1607.

[16] Abigail See, Stephen Roller, Douwe Kiela, and Jason Weston, "What makes a good conversation? how controllable attributes affect human judgments," in *NAACL*, 2019, pp. 1702–1723.

[17] Annie Louis and Ani Nenkova, "Automatic identification of general and specific sentences by leveraging discourse annotations," in *IJCNLP*, 2011, pp. 605–613.

[18] Ruqing Zhang, Jiafeng Guo, Yixing Fan, Yanyan Lan, Jun Xu, and Xueqi Cheng, "Learning to control the specificity in neural response generation," in *ACL*, 2018, pp. 1108–1117.

[19] Ian J Goodfellow, Jonathon Shlens, and Christian Szegedy, "Explaining and harnessing adversarial examples," in *ICLR*, 2014.

[20] William Fedus, Barret Zoph, and Noam Shazeer, "Switch transformers: Scaling to trillion parameter models with simple and efficient sparsity," *CoRR*, vol. abs/2101.03961, 2021.

[21] Eric Jang, Shixiang Gu, and Ben Poole, "Categorical reparameterization with gumbel-softmax," in *ICLR*, 2017.

[22] Yanran Li, Hui Su, Xiaoyu Shen, Wenjie Li, Ziqiang Cao, and Shuzi Niu, "Dailydialog: A manually labelled multi-turn dialogue dataset," in *IJCNLP*, 2017, pp. 1735–1742.

[23] Iulian Vlad Serban, Alessandro Sordoni, Yoshua Bengio, Aaron C. Courville, and Joelle Pineau, "Building end-to-end dialogue systems using generative hierarchical neural network models," in *AAAI*, 2016, pp. 3776–3784.

[24] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin, "Attention is all you need," in *NIPS*, 2017, pp. 5998–6008.

[25] Stephen Roller, Emily Dinan, Naman Goyal, Da Ju, Mary Williamson, Yinhan Liu, Jing Xu, Myle Ott, Eric Michael Smith, Y-Lan Boureau, and Jason Weston, "Recipes for building an open-domain chatbot," in *EACL*, 2021, pp. 300–325.

[26] Yizhe Zhang, Michel Galley, Jianfeng Gao, Zhe Gan, Xiujun Li, Chris Brockett, and Bill Dolan, "Generating informative and diverse conversational responses via adversarial information maximization," in *NIPS*, 2018, pp. 1815–1825.

[27] Jiwei Li, Will Monroe, Alan Ritter, Michel Galley, Jianfeng Gao, and Dan Jurafsky, "Deep reinforcement learning for dialogue generation," in *EMNLP*, 2016, pp. 1192–1202.

[28] Kishore Papineni, Salim Roukos, Todd Ward, and Wei-Jing Zhu, "Bleu: a method for automatic evaluation of machine translation," in *ACL*, 2002, pp. 311–318.

[29] Iulian Vlad Serban, Alessandro Sordoni, Ryan Lowe, Laurent Charlin, Joelle Pineau, Aaron C. Courville, and Yoshua Bengio, "A hierarchical latent variable encoder-decoder model for generating dialogues," in *AAAI*, 2017, pp. 3295–3301.

[30] Kyunghyun Cho, Bart van Merrienboer, Çaglar Gülçehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," in *EMNLP*, 2014, pp. 1724–1734.

[31] Zhaojiang Lin, Andrea Madotto, Jamin Shin, Peng Xu, and Pascale Fung, "Moel: Mixture of empathetic listeners," in *EMNLP-IJCNLP*, 2019, pp. 121–132.

[32] Jeffrey Pennington, Richard Socher, and Christopher D Manning, "Glove: Global vectors for word representation," in *EMNLP*, 2014, pp. 1532–1543.

[33] Thomas Wolf, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, Remi Louf, Morgan Funtowicz, Joe Davison, Sam Shleifer, Patrick von Platen, Clara Ma, Yacine Jernite, Julien Plu, Canwen Xu, Teven Le Scao, Sylvain Gugger, Mariama Drame, Quentin Lhoest, and Alexander Rush, "Transformers: State-of-the-art natural language processing," in *EMNLP*, 2020, pp. 38–45.

[34] Diederik P. Kingma and Jimmy Ba, "Adam: A method for stochastic optimization," in *ICLR*, 2015.