Envisioning Future from the Past: Hierarchical Duality Learning for Multi-Turn Dialogue Generation

Ang Lv[1,*], **Jinpeng Li**[2,*], **Shufang Xie**[1], **Rui Yan**[1,3,†]
[1]Gaoling School of Artifical Intelligence, Renmin University of China
[2]Wangxuan Institute of Computer Technology, Peking University
[3]Engineering Research Center of Next-Generation Intelligent Search and Recommendation, Ministry of Education
{anglv, shufangxie, ruiyan}@ruc.edu.cn, lijinpeng@stu.pku.edu.cn

Abstract

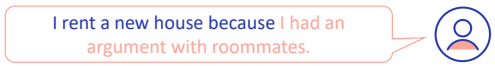
In this paper, we define a widely neglected property in dialogue text, duality, which is a hierarchical property that is reflected in human behaviours in daily conversations: Based on the logic in a conversation (or a sentence), people can infer follow-up utterances (or tokens) based on the previous text, and vice versa. We propose a hierarchical duality learning for dialogue (HDLD) to simulate this human cognitive ability, for generating high quality responses that connect both previous and follow-up dialogues. HDLD utilizes hierarchical dualities at token hierarchy and utterance hierarchy. HDLD maximizes the mutual information between past and future utterances. Thus, even if the future text is invisible during inference, HDLD is capable of estimating future information implicitly based on dialogue history and generates both coherent and informative responses. In contrast to previous approaches that solely utilize future text as auxiliary information to encode during training, HDLD leverages duality to enable interaction between dialogue history and the future. This enhances the utilization of dialogue data, leading to the improvement in both automatic and human evaluation.

1 Introduction

Dialogue generation achieves great success in recent years. In this task, a dialogue session lasts for multiple turns, and the goal is to predict the response to the previous context. Researchers have explored various aspects of dialogue generation models including context-aware (Tian et al., 2017), persona-based (Qian et al., 2018) and knowledge-grounded (Zhao et al., 2020; Tao et al., 2021), etc. Because dialogue generation is a context-aware process, there is a large body of literature on the context modeling among all research lines.

In the pilot studies, contexts are modeled as the concatenation of previous utterances (Sordoni et al., 2015), organized in a non-hierarchical way. A dialogue is later considered to have hierarchical characteristics in which both token semantics within an utterance and logic between utterances need consideration (Serban et al., 2016). The hierarchical context modeling among utterances can be further improved with a global hidden variable (Serban et al., 2017) indicating topic, emotion, or persona. At the same time, in addition to modeling text at multiple hierarchies, some researchers have also committed to mining the intensive interaction among the previous utterances sufficiently (Tao et al., 2019; Zhou et al., 2018; Wu et al., 2017). In summary, most context modeling approaches focus on the previous utterances and ignore the follow-up utterances. In recent years, researchers (Shen et al., 2018; Feng et al., 2020; Liu et al., 2022) realize that a response of high quality should not only be related to dialogue history but also connect the follow-up parts. However, they only treat

(a) Inter-duality.



(b) Intra-duality.

Figure 1: (a) Inter-duality: the dialogue is divided into two parts. According to the first two utterances, we can infer that the follow-up part is related to the scenic spot commodities. Similarly, based on the last two utterances, we can infer that the previous conversation is about tourism. (b) Intra-duality: in an utterance, given the blue part, people know that the follow-up part is the reason. Given the pink part, people know that the previous tokens may explain the result of the argument.

*Equal contribution.
†Corresponding author: Rui Yan (ruiyan@ruc.edu.cn).

Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics
Volume 1: Long Papers, pages 7382–7394
July 9-14, 2023

utterances after the target response (i.e., dialogue future) as auxiliary information to encode during training, which violates a human cognitive ability in daily conversations: given a dialogue segment, people can infer future contents; also, people can imagine what the interlocutors discussed before. In the whole process, there are interactions between visible dialogue parts and human-envisioned parts. We name the property in dialogue texts that humans take advantage of to infer unknown previous or follow-up parts as **duality**[1]. Inadequate consideration of duality ignores the interaction between dialogue history and the future, resulting in underutilized future information.

Duality in dialogues has two key characteristics. Firstly, as a property in dialogues, duality can be viewed from a hierarchical viewpoint. At token hierarchy, intra-utterance duality (abbreviated as intra-duality) is reflected in that preceding words and the subsequent words can infer each other. Similarly, at utterance hierarchy, inter-utterance duality (abbreviated as inter-duality) is reflected in that utterances before and after the target response can infer each other. Figure 1(a) and (b) illustrate dualities at two hierarchies. Secondly, due to the context-aware nature of dialogues, intra-duality does not play a positive role alone and must be tightly integrated with inter-duality.

In this paper, from the point of view of dialogue duality, we propose the **H**ierarchical **D**uality **L**earning for **D**ialogue (HDLD) in order to fully exploit dialogue text for better generation. In HDLD, a dialogue session is divided into two parts: past and future. At the utterance hierarchy, there is a forward generation model and a backward generation model working in a dual learning framework. The forward model takes the past utterances as context and predicts utterances in the future part. Predicted outputs can be used as context for the backward model to back-predict the past utterances. The backward model performs a dual process. To integrate dualities at two hierarchies, HDLD employs a cross-hierarchy distillation mechanism: guided by both future and past contexts, we distill both context-related and intra-dual knowledge from the ground truth, then we use the distilled knowledge to train the models. Through this cross-hierarchy distillation mechanism, HDLD couples hierarchies tightly, improving generation quality at multiple granularities. We design a two-stage joint optimization of the forward and the backward model that maximizes the mutual information between the past and future. Thus, during inference even if future information is invisible, HDLD is capable of envisioning future information implicitly only based on the dialogue history and generates both coherent and informative responses.

We conduct multi-turn dialogue generation experiments on two public datasets, DailyDialog (Li et al., 2017) and OpenSubtitles (Lison and Tiedemann, 2016). Compared with previous works that also consider future utterances, exploiting dialogue duality helps better utilize dialogue data, leading to the improvement in both automatic and human evaluation. To sum up, our contributions can be summarized as follows:

- We first define the dialogue duality, which was widely neglected in previous works. We propose the hierarchical duality learning framework for dialogue (HDLD) to augment multi-turn dialogue generation by exploiting duality.
- To realize HDLD, We design a cross-hierarchy distillation mechanism to couple hierarchies, and a joint optimization of two models that can be theoretically proved to augment generation.
- We show the effectiveness of HDLD in generating coherent and informative responses.

2 Related Work

2.1 Dialogue Generation with Future Utterances

In open-domain dialogue generation, the focus of this paper, most models are context-aware, predicting the target response based on previous context representations. With the development of this field, researchers began to pay attention to the future context: NEXUS (Shen et al., 2018) introduces a posterior distribution for estimating the future information as a hidden variable and decodes responses accordingly. RegDG (Feng et al., 2020) trains a teacher model with both past and future utterances and transfers the whole-context knowledge to a student model which has only access to past utterances. ProphetChat (Liu et al., 2022) first predicts two utterances forward, generating multiple responses and future utterances. With a selector, the best response and future utterance are selected. A final response is squeezed by both history and future. Compared with HDLD, there is inadequate consideration of dialogue duality in these works.

[1]Our definition of 'duality' is different from that in papers on dual learning (He et al., 2016). We discuss details in section 2.2

They ignore the explicit interaction between the dialogue history and future, still resulting in under-utilization of future information.

2.2 Dual Learning

Dual learning is proposed to boost neural machine translation (He et al., 2016). In the dual learning framework, one agent represents the model for the primal task, which aims to build a mapping from domain X to domain Y, and the other agent represents the model for the dual task, which aims to build a mapping from domain Y back to domain X. Based on the training feedback signals generated from each other, two models are updated iteratively until convergence. By leveraging the primal-dual structure of two tasks, the training and inference of the primal and dual tasks are improved.

Many dialogue generation methods (Yang et al., 2018; Cui et al., 2019; Li et al., 2021) have mentioned 'duality' in their paper but note that there are differences in our definition of 'duality' from theirs. Duality describes the relationship between dual and primal tasks (e.g., English-to-Chinese translation and Chinese-to-English translation) in previous works. By contrast, in this paper, duality refers to a unique dialogue property that humans can take advantage of to infer dialogue future (or history) according to dialogue history (or future).

It is important to note the distinction between duality and "bi-directional" models such as Bi-RNN (Schuster and Paliwal, 1997) or MLM like BERT (Devlin et al., 2019). While these models extract features in two directions, they are not "dual." In these models, the backward features are typically merged with the forward features, rather than actively enhancing forward feature extraction as in dual learning approaches.

3 Method

We first present the preliminary in § 3.1. Then, we introduce a two-stage optimization that exploits dualities at two hierarchies in § 3.2, and some important details during training and inference in § 3.3.

3.1 Preliminary

Given a multi-turn dialogue $D = \{u_1, \ldots, u_n\}$ with $n > 2$ utterances. An utterance $u_i = \{w_i^1, w_i^2, \ldots, w_i^{|u_i|}\}$ consists of $|u_i|$ tokens. The embeddings of u_i is e_i with d dimension. Take t-th utterance as the dividing point, we divide D into two parts: the past contexts $\mathbb{P} = \{x_{1:t} | x_{1:t} = \{u_1, \ldots, u_t\}, t \in [1, n-1]\}$ and the reversed future contexts $\mathbb{F} = \{y_{n:t+1} | y_{n:t+1} = \{u_n, \ldots, u_{t+1}\}, t \in [1, n-1]\}$. In HDLD, there are a forward generation model f_{xy}: $\mathbb{P} \rightarrow \mathbb{F}$ and its dual generation model, a backward model f_{yx} : $\mathbb{F} \rightarrow \mathbb{P}$. In each model, there are one transformer encoder and two transformer decoders. At the utterance hierarchy, f_{xy} takes a $x_{1:t} \in \mathbb{P}$ as context to predict utterances in $y_{n:t+1}$, and f_{yx} takes the corresponding $y_{n:t+1} \in \mathbb{F}$ as context to predict utterances in $x_{1:t}$, which forms the horizontal cycle in Figure 2. Two decoders in each model process in dual directions to extract intra-duality features in utterances, and cross-hierarchy distillation integrates hierarchies: guided by both future and past contexts (top-down), we distill knowledge that is both context-related and intra-dual. Distilled knowledge is used to teach models (bottom-up), which forms the vertical cycle in Figure 2. Figure 3 shows the HDLD architecture. In the following description, we remove the subscript of x or y when there is no ambiguity.

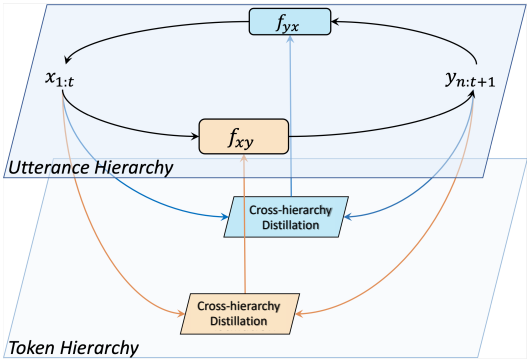


Figure 2: The concept of HDLD. Two models form a horizontal dual cycle at utterance hierarchy. Cross-hierarchy distillation forms a vertical dual cycle to couple hierarchies.

3.2 Two-Stage Optimization

As a dual learning framework, our ultimate goal is to obtain the forward model f_{xy} which generates response based on dialogue history while the f_{yx} is auxiliary. Given any aligned $x \in \mathbb{P}$ and $y \in \mathbb{F}$, to leverage inter-duality, we apply the two models consecutively, completing the horizontal dual cycle. This process yields four outputs: $x' = f_{yx}(y)$, $y' = f_{xy}(x)$, $x'' = f_{yx}(y')$, and $y'' = f_{xy}(x')$. Reconstruction errors between x'' and x, and between y'' and y reflect how well the horizontal dual cycle is completed, i.e., the utility of inter-duality. We

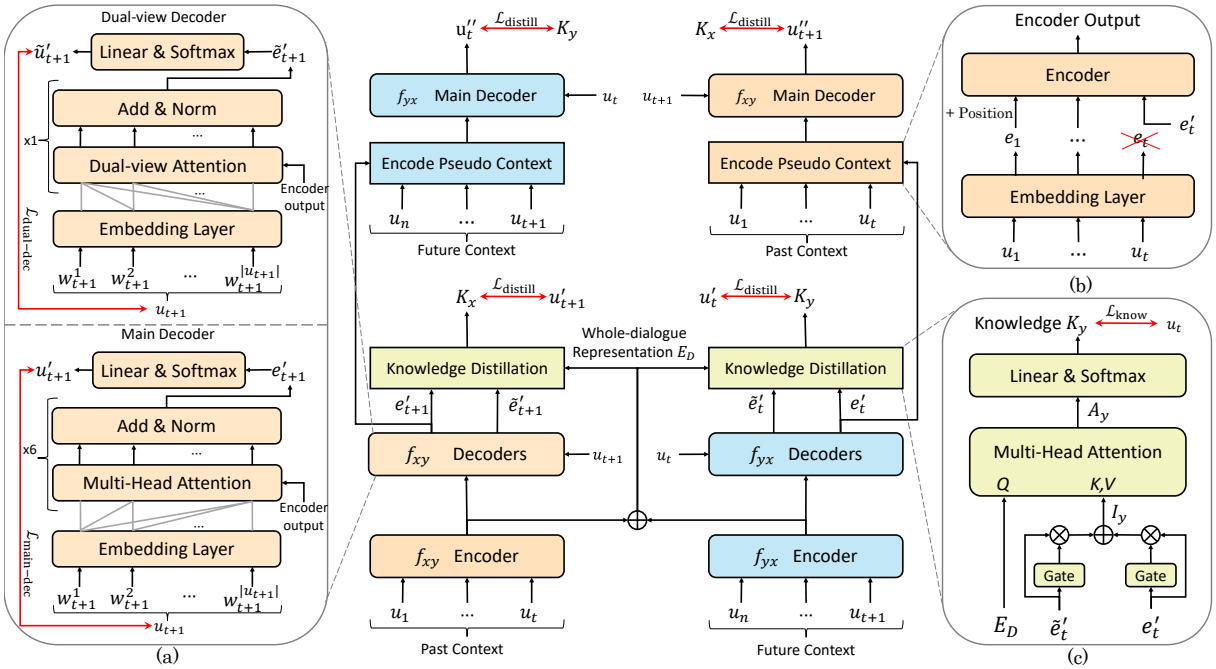


Figure 3: Model structure. In the first stage (bottom half of the figure), each model take its own context to generate responses. Based on the whole dialogue and two decoders' outputs (figure a), knowledge is distilled (figure c). Two models share the distillation module. In the second stage (top half of the figure), a model's outputs in the previous stage are used to make the pseudo context for the other model (figure b). Red arrows indicate the loss calculation between two terms. $\oplus$ represents matrix addition. $\otimes$ represents Hadamard product.

minimize reconstruction errors by minimizing the objective:

$$\begin{aligned} &-\sum_{x\in\mathbb{P}} \log p(x''=x|x;f_{xy},f_{yx}) \\ &-\sum_{y\in\mathbb{F}} \log p(y''=y|y;f_{yx},f_{xy}). \end{aligned} \tag{1}$$

However, it is challenging to directly optimize the final output x'' and y'', which is hindered by the error accumulation in practice. Thus, we find the upper bound of Eq. 1 which is more practical. Take the logarithmic probability $-\log p(x''=x|x;f_{xy},f_{yx})$ as an example, we have:

$$\begin{aligned} &-\log p(x''=x|x;f_{xy},f_{yx}) \\ &=-\log \sum_{y'} p(x''=x,y'|x,f_{xy},f_{yx}) \\ &=-\log \sum_{y'} p(x''=x|x,y',f_{xy},f_{yx})p(y'|x;f_{xy},f_{yx}) \\ &\le -\sum_{y'} p(y'|x,f_{xy}) \log p(x''=x|y',f_{yx}). \end{aligned} \tag{2}$$

The inequality is derived from two facts: (1) x'' is only directly related to y' and f_{yx} and y' is only directly related to x and f_{xy} and (2) the concavity of logarithm. Therefore, we can derive the upper bound of Eq. 1:

$$\begin{aligned} &-\sum_{x\in\mathbb{P}} \log p(x''=x|x;f_{xy},f_{yx}) \\ &-\sum_{y\in\mathbb{F}} \log p(y''=y|y;f_{yx},f_{xy}) \\ &\le -\sum_{x\in\mathbb{P}}\sum_{y'\in\mathbb{F}} p(y'|x,f_{xy}) \log p(x''=x|y',f_{yx}) \\ &-\sum_{y\in\mathbb{F}}\sum_{x'\in\mathbb{P}} p(x'|y,f_{yx}) \log p(y''=y|x',f_{xy}). \end{aligned} \tag{3}$$

We can optimize the upper bound by simultaneously minimizing each probability term within it. This means that in order to exploit inter-duality, it is not necessary to complete the horizontal dual cycle by consecutively applying two models and directly optimizing the final outputs. Instead, we can optimize the intermediate outputs, which divides the training into two stages. In the first stage, f_{xy} generates y' and f_{yx} generates x'. We fit y' and x' to their ground truth y and x, respectively. In the second stage, f_{xy} predicts y'' and f_{yx} predicts x'', which fit to y and x again, respectively. Formally,

we minimize the negative log-likelihoods:

$$\begin{aligned} L_{inter} = & \mathbb{E}\left[-\log p\left(y' = y|x, f_{xy}\right)\right] \\ & + \mathbb{E}[-\log p\left(y'' = y|x', f_{xy}\right)] \\ & + \mathbb{E}[-\log p\left(x' = x|y, f_{yx}\right)] \\ & + \mathbb{E}[-\log p\left(x'' = x|y', f_{yx}\right)]. \end{aligned} \tag{4}$$

Another advantage of the two-stage optimization is that it allows for easier consideration of intra-duality by assigning the top-down and bottom-up parts of the cross-hierarchy distillation to each stage. We delve into the detailed implementation of Eq. 4 in each stage.

Stage One. In the first stage, we apply both models once, and utilize top-down distillation to distill integrated hierarchical duality knowledge. However, reconstructing the entire dialogue history x and future response y can be challenging due to their often lengthy nature, consisting of multiple sentences. To address this challenge, we relax the constraint and allow each model to reconstruct only one subsequent utterance in the other part of the dialogue. For example, in the case of f_{xy}, $x_{1:t}$ serves as the context and u_{t+1} is the target, while for f_{yx}, $y_{n:t+1}$ acts as the context and u_t is the target. Both models encode their respective contexts and pass the encoder output to their respective main decoders. The generated responses from the main decoders in the first stage are denoted as $u^{'}_{t+1}$ and $u^{'}_{t}$. To optimize f_{xy} and f_{yx}, we fit their outputs to the ground truth u_{t+1} and u_t by minimizing the negative log-likelihoods:

$$\begin{aligned} \mathcal{L}_{\text{main-dec}} = & \mathbb{E}\left[-\log p(u_{t+1}|x_{1:t})\right] \\ & + \mathbb{E}\left[-\log p(u_t|y_{n:t+1})\right]. \end{aligned} \tag{5}$$

The dual-view decoder in each model operates on the same encoder outputs as the main decoder but processes them in a reverse direction, which enables us to leverage intra-duality, as the cross attention mechanism in decoders limits tokens from attending to their subsequent tokens. To achieve this, we flip the attention mask used in the main decoder along the main diagonal to create the mask for the dual-view decoder. The generated responses from the dual-view decoders in the first stage are represented as $\tilde{u}^{'}_{t+1}$ and $\tilde{u}^{'}_{t}$. Figure 3(a) provides an illustration of the behaviors of both the main decoder and the dual-view decoder.

Next, we introduce the top-down part of the cross-hierarchy distillation. Denote the mixture of outputs from the main decoder and the dual-view decoder as intra-dual features I:

$$\begin{aligned} I_x &= e^{'}_{t+1} \otimes \sigma(W_I e^{'}_{t+1}) + \tilde{e}^{'}_{t+1} \otimes \sigma(\tilde{W}_I \tilde{e}^{'}_{t+1}), \\ I_y &= e^{'}_{t} \otimes \sigma(W_I e^{'}_{t}) + \tilde{e}^{'}_{t} \otimes \sigma(\tilde{W}_I \tilde{e}^{'}_{t}), \end{aligned} \tag{6}$$

where $W_I, \tilde{W}_I \in \mathbb{R}^{d\times d}$, the $\otimes$ denotes Hadamard product, and σ denotes the `sigmoid` function. $e^{'}_{\cdot}$ denotes hidden state embeddings of the corresponding $u^{'}_{\cdot}$. The representations of the entire dialogue, denoted as E_D, play a crucial role in distilling both context-related and intra-dual knowledge from the intra-dual features I using the multi-head attention mechanism. In this mechanism, I serves as both the *key* and *value*, while E_D acts as the *query*. The output of the multi-head attention, denoted as A, is then mapped to match the vocabulary size $|V|$, resulting in the final knowledge representation K. The detailed process is illustrated in Figure 3(c), where the linear layer weight is represented by W_A of size $\mathbb{R}^{|V|\times d}$:

$$\begin{aligned} A_x &= \text{softmax}(\frac{E_D I_x^T}{\sqrt{d}}) I_x, \\ A_y &= \text{softmax}(\frac{E_D I_y^T}{\sqrt{d}}) I_y, \\ K_x &= \text{softmax}(W_A A_x), \\ K_y &= \text{softmax}(W_A A_y). \end{aligned} \tag{7}$$

Stage Two. In the second stage, we use predicted responses in stage one to construct pseudo contexts and based on which, two models are applied again to complete the horizontal dual cycle. Meanwhile, the bottom-up distillation is carried out to complete the vertical dual cycle.

As shown in Figure 3(b), we substitute $e^{'}_{t}$ and $e^{'}_{t+1}$ obtained in the previous stage for e_t and e_{t+1} in embedded $x_{1:t}$ and $y_{n:t+1}$, respectively. By utilizing $e^{'}_{t}$ and $e^{'}_{t+1}$ to construct pseudo contexts, we preserve the gradient flow, allowing the joint optimization of the two models. These modified context representations can be referred to as embedded pseudo contexts, denoted as $x^{'}_{1:t}$ and $y^{'}_{n:t+1}$. Next, f_{xy} and f_{yx} predict each one's target again based on respective pseudo contexts. We denote the generated responses from main decoders in stage two as $u^{''}_{t+1}$ and $u^{''}_{t}$, respectively.

Instead of using discrete ground truth sentences, we employ more informative intra-dual knowledge K_x and K_y to guide the decoding process of the two models. To enhance the suitability of the knowledge as labels, we further augment the target word probability in the soft distributions K_x

and K_y, following the method proposed by (Wang et al., 2021). The guidance of the decoding process is achieved through bottom-up distillation, where we minimize the Kullback-Leibler divergence (KL) (Kullback and Leibler, 1951) between the outputs of the main decoders in both stages and the knowledge representation:

$$\begin{aligned}\mathcal{L}_{\text{distill}} = \frac{1}{|\mathbb{P}|}\sum_{K_x}(&\text{KL}[K_x||p(u^{'}_{t+1}|x_{1:t})] \\ &+ \text{KL}[K_x||p(u^{''}_{t+1}|x^{'}_{1:t})]) \\ + \frac{1}{|\mathbb{F}|}\sum_{K_y}(&\text{KL}[K_y||p(u^{'}_{t}|y_{n:t+1})] \\ &+ \text{KL}[K_y||p(u^{''}_{t}|y^{'}_{n:t+1})]).\end{aligned} \tag{8}$$

Moreover, we optimize both the dual-view decoder outputs and knowledge to be similar to ground truth in semantics:

$$\begin{aligned}\mathcal{L}_{\text{dual-dec}} &= \mathbb{E}\left[-\log p(\tilde{u}^{'}_{t+1} = u_{t+1}|x_{1:t})\right] \\ &\quad + \mathbb{E}\left[-\log p(\tilde{u}^{'}_{t} = u_{t}|y_{n:t+1})\right], \\ \mathcal{L}_{\text{know}} &= \mathbb{E}\left[-\log p(K_x = u_{t+1}|E_D, e^{'}_{t+1}, \tilde{e}^{'}_{t+1})\right] \\ &\quad + \mathbb{E}\left[-\log p(K_y = u_{t}|E_D, e^{'}_{t}, \tilde{e}^{'}_{t})\right].\end{aligned} \tag{9}$$

For clear demonstration, we use red arrow to indicate all loss calculations in Figure. 3.

3.3 Training and Inference

During training, the training loss function $\mathcal{L}$ is defined as follows:

$$\mathcal{L} = \mathcal{L}_{\text{main-dec}} + \mathcal{L}_{\text{dual-dec}} + \mathcal{L}_{\text{know}} + \mathcal{L}_{\text{distill}}. \tag{10}$$

During inference, **only the encoder and the main decoder of f_{xy} remain active**. Therefore, HDLD does not incur any additional inference costs than a vanilla transformer. It is feasible because to consider the mutual information $I(\mathbb{P}, \mathbb{F})$ between $\mathbb{P}$ and $\mathbb{F}$:

$$\begin{aligned}I(\mathbb{P}, \mathbb{F}) = &-\sum_{x\in\mathbb{P}}\sum_{y\in\mathbb{F}} p(x,y)\log p(x,y) \\ &+ \sum_{x\in\mathbb{P}}\sum_{y\in\mathbb{F}} p(x,y)\log p(x|y) \\ &+ \sum_{x\in\mathbb{P}}\sum_{y\in\mathbb{F}} p(x,y)\log p(y|x)\end{aligned} \tag{11}$$

In Eq. (11), $p(x, y)$ is constant and is decided by data, and HDLD maximizes $p(x|y)$ and $p(y|x)$ according to Eq. (4) through the joint optimization. Thus, Eq. 11, the mutual information between dialogue history and future is maximized. Given the past context, HDLD is able to implicitly envision future information with the maximized mutual information.

Datasets	Train	Valid	Test
DailyDialog	41558	3966	3659
OpenSubtitles	223893	22495	22413

Table 1: Dataset statistics.

4 Experimental Setups

Baselines We compare HDLD with baselines in two groups. The first group introduces future utterances to enhance the model: **NEXUS** (Shen et al., 2018) introduces a code space learned from the whole dialogue and samples from the code space during testing to estimate future context. **RegDG** (Feng et al., 2020) trains a teacher model that has access to the whole dialogue to teach a student model which only takes past utterances as input. **ProphetChat** (Liu et al., 2022) first predicts two steps forward, generating multiple responses and future utterances, which are used to squeeze the final response along with the dialogue history.

The second group considers the relationship among tokens in an utterance comprehensively: **D2GPO** (Li et al., 2020) augments the training with a data-dependent Gaussian prior distribution, which is generated in pre-procession based on the training set. **AdaLabel** (Wang et al., 2021) introduces an auxiliary decoder that uses a bidirectional attention to dynamically estimate a token distribution at each time step. Besides, we also compare to **HRED** (Serban et al., 2016), a pilot work for hierarchical context modeling, and **Transformer** (Vaswani et al., 2017), the backbone of AdaLabel, D2GPO and HDLD.

Datasets We conduct multi-turn dialogue generation experiments on two public datasets: DailyDialog (Li et al., 2017) and OpenSubtitles (Lison and Tiedemann, 2016). DailyDialog contains high quality multi-turn dialogues collected from daily conversations while OpenSubtitles contains dialogues from movies. As a preprocessing step, we exclude dialogue sessions shorter than 4 utterances from the dataset, and we create samples where the contexts consist of no more than 100 tokens. We split a pre-processed dialogue into two parts: dialogue history and future. Baselines that utilize

future utterances have access to two parts and those do not utilize future utterances only take the historical part as contexts. In this way, we ensure all experiments have access to the same amount of training samples, which is a fair setting. Due to our unique task settings, we cannot directly cite any numbers from other papers and we reproduced every baseline. Detailed datasets statistics are in Table 1.

Hyper-parameters We use grid-search to select the best hyper-parameters. The search ranges for learning rate and batch size are $\{0.00006, 0.00008, 0.00010, 0.00012\}$ and $\{64, 128, 256, 520, 720, 1000\}$, respectively. The adopted learning rate for HDLD is 0.00008 and the mini-batch size is 128 on DailyDialog and 720 on OpenSubtitles. As for other important settings: the warmup steps are 4000. We use Adam optimizer with $\beta = (0.9, 0.98)$. Both attention dropout and activation dropout are 0.1. We select the best parameters based on perplexity on the validation set. All experiments adopt greedy decoding. Early stop patience is 10 epochs on DailyDialog and 5 on OpenSubtitles. Some baseline-specific hyper-parameters not mentioned follow the original papers. The training is performed on two A40 GPUs.

Evaluation Metrics We consider three automatic evaluation metrics: BLEU (Papineni et al., 2002), Distinct (Dist) (Li et al., 2016), and BLEURT (Sellam et al., 2020). BLEU score measures the word overlap between generated responses and the ground truth. Dist score measures the ratio of unique n-grams in the generated responses. Because these two metrics are only sensitive to lexical variation, we evaluate BLEURT, an advanced learned semantic-sensitive evaluation metric.

For human evaluation, we recruit five evaluators to manually judge 500 samples from all models in blind testing. All responses are re-capitalized and de-tokenized fairly. Evaluators rate samples from three aspects: *readability* (**Read.**): how readable the response is; *coherency* (**Coh.**): whether a response is coherent with the context; and *informative* (**Info.**): whether the response is informative enough. In each aspect, the evaluator can score at '0' for bad, '1' for borderline, or '2' for good.

The salary for each evaluator is 1 dollar per 10 samples. To give a fair salary, we first evaluate 50 samples by ourselves, calculate the time and effort, and set this amount (samples evaluated by ourselves are just for evaluating the salary, which is not given to evaluators and not reported in the final results).

5 Results

Comparison with Prior Works Table 2 shows the automatic evaluation results. HDLD outperforms all baselines on all metrics. HDLD doubles scores on most metrics of baselines using no future information, i.e., Transformer and HRED, and we gain improvement by a large margin than other strong baselines. Thanks to exploited inter-duality, HDLD connects the past and future contexts so it matches references best (highest BLEU scores). Related to both the past and future contexts, there are few general and meaningless responses so HDLD achieves the highest diversity scores. Except above lexical metrics, we also achieve the highest semantic-sensitive BLEURT scores. From automatic evaluation, our motivation of making responses meaningful and specific by utilizing hierarchical duality is proved effective.

Human evaluation results are shown in Table 4. The highest readability shows that intra-duality in an utterance is leveraged and thus responses are more natural. The highest coherency and informative scores demonstrate that HDLD makes full use of dialogue texts and has the best context modeling capability, resulting in responses connecting and related to both past and future contexts. Overall, our model generates the highest-quality responses.

Ablation Study We study three variants of HDLD by adding following approaches to a vanilla transformer: (1) **+Inter.**: utilizing inter-duality, i.e., let two transformer models form a horizontal dual cycle; (2) **+Intra.**: utilizing intra-duality, i.e., adding a dual-view decoder in a vanilla transformer to obtain knowledge for teaching the model itself; (3) **+Inter. + Intra.**: combining previous two approaches but do not combine two hierarchies with cross-hierarchy distillation; In variant (1), there is no way to distill knowledge so we minimize the negative log-likelihood in two stages. Table 3 shows results. On two datasets, the inter-duality approach improves performance and outperforms the first group baselines except ProphetChat in Table 2. By only using intra-dual features, or by combining two duality approaches without cross-hierarchy distillation, most metrics are negatively affected. It reveals that without inter-duality, intra-duality does

Models	DailyDialog							OpenSubtitles						
	BLEURT	BLEU-1,2,3,4				Dist-1,2		BLEURT	BLEU-1,2,3,4				Dist-1,2	
HRED	2.49	17.19	7.76	4.35	2.53	1.67	6.41	2.07	14.10	2.10	0.49	0.17	0.17	0.66
NEXUS	2.52	17.33	7.40	4.10	2.34	2.46	8.13	1.85	14.13	2.08	0.52	0.15	0.20	0.83
Transformer	2.92	19.69	10.15	6.43	4.37	2.18	13.94	2.59	13.95	2.17	0.60	0.21	0.21	1.20
RegDG	2.64	21.14	11.04	7.39	5.48	2.12	9.74	2.62	13.48	1.93	0.51	0.18	0.20	1.35
D2GPO	3.00	22.17	11.98	7.86	5.55	2.08	14.87	2.61	13.64	2.20	0.61	0.21	0.21	1.25
ProphetChat	3.18	23.21	14.17	10.62	8.25	4.05	22.67	2.63	13.77	2.21	0.62	0.21	0.35	2.11
AdaLabel	3.23	24.16	14.80	10.85	8.63	3.95	22.20	2.60	13.25	2.23	0.60	0.21	0.33	2.10
HDLD	**3.41**	**27.52**	**18.26**	**14.30**	**12.02**	**4.37**	**23.23**	**2.64**	**14.24**	**2.26**	**0.62**	**0.24**	**0.42**	**2.21**

Table 2: Automatic evaluation results. The improvement of HDLD passes the t-test with $p < 0.03$ in all metrics.

Models	DailyDialog							OpenSubtitles						
	BLEURT	BLEU-1,2,3,4				Dist-1,2		BLEURT	BLEU-1,2,3,4				Dist-1,2	
Transformer	2.92	19.69	10.15	6.43	4.37	2.18	13.94	2.59	13.95	2.17	0.60	0.21	0.21	1.20
+Inter.	2.96	22.33	11.87	7.84	5.56	2.90	18.77	2.63	14.29	2.23	0.61	0.21	0.38	2.13
+Intra.	2.80	18.38	8.61	5.05	3.23	2.25	13.31	2.62	13.89	2.10	0.54	0.17	0.17	1.08
+Inter. + Intra.	2.86	21.30	10.59	6.44	4.25	2.63	15.87	**2.69**	**14.52**	2.22	**0.64**	0.23	0.25	1.33
HDLD (+Inter. + Intra. + Cross.)	**3.41**	**27.52**	**18.26**	**14.30**	**12.02**	**4.37**	**23.23**	2.64	14.24	**2.26**	0.62	**0.24**	**0.42**	**2.21**

Table 3: Ablation study on the duality hierarchy on DailyDialog and OpenSubtitles.

Models	DailyDialog			OpenSubtitles		
	Read.	Coh.	Info.	Read.	Coh.	Info.
Ground-Truth	1.75	1.81	1.68	1.64	1.72	1.54
HRED	1.08	0.89	0.94	0.91	0.46	0.41
NEXUS	1.06	0.98	1.01	0.95	0.49	0.42
Transformer	1.14	1.01	1.04	0.99	0.55	0.53
RegDG	1.14	1.10	1.04	1.07	0.62	0.58
D2GPO	1.28	1.05	1.18	1.07	0.57	0.52
ProphetChat	1.22	1.03	1.10	1.09	0.62	0.55
AdaLabel	1.29	1.09	1.22	1.05	0.66	0.59
HDLD	**1.36**	**1.18**	**1.28**	**1.10**	**0.68**	**0.61**

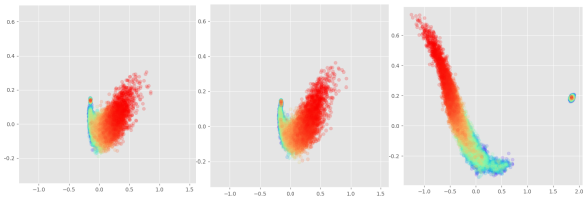
Table 4: Human evaluation results on readability, coherency, and informative. The kappa value of the evaluation is 0.67.

Dialogue Text	A: That dress is very pretty. Why don't you like it? B: It's too loud. A: _____ B: Well, you know, those styles or colors don't suit me.
Ground Truth	We've been looking around for many hours. What on earth are you looking for?
HRED	Would you like to see the menu with you?
NEXUS	I don't think you've got a lot of things. You can't do it.
Transformer	what's the color color like?
D2GPO	yes, it's my favorite.
RegDG	I'm afraid it's too late. I don't think it's too cold.
ProphetChat	I don't think so.
AdaLabel	Well , it's just a bit over there.
HDLD	Oh, I see. How about this one?

Table 5: A case showing effectiveness of HDLD.

not work well in dialogues, which confirms our motivation of coupling hierarchies. Finally, HDLD significantly improves the performance on almost all metrics. The study reveals that (1) coupled hierarchical duality is effective, and (2) cross-hierarchy distillation is necessary and non-trivial.

Further, we provide some cases to demonstrate the impact of dualities. For inter-duality, we demonstrate a case to illustrate the effectiveness of HDLD in Table 5. Although only the past context is visible to HDLD, it still generates the response which connects the past and the future utterances best. On the contrary, responses from baselines are either unrelated to the dialogue topic or can not lead to future conversation. For intra-duality, we visualize trained embedding layers of Transformer, Transformer with inter-duality approach, and HDLD in Figure 4. Comparing three sub-figures, we find that without considering intra-duality, token embeddings tend to be optimized to be indistinguishable. The result is: representations are similar whatever the inputs are, so the model prefers outputting certain specific tokens, which makes responses less diverse. Only exploiting inter-duality has little effect on token representations which are fine-grained and of low-hierarchy. By contrast, in HDLD, tokens are much more distinguishable. Because semantic information can be represented better, the context modeling ability is improved and generated responses are more diverse and meaningful.



(a) Transformer (b) Trans. + Inter. (c) HDLD

Figure 4: Visualization of embeddings from three models. Red represents the most frequent words while blue represents the least frequent word.

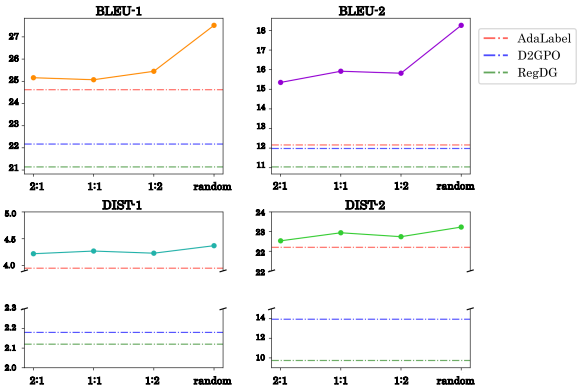


Figure 5: Past:future length ratios impact on results.

Analysis of Length Intuitively, the ratio between the length of the past context and that of the future context has an impact on the performance. We analyze this ratio. In the above experiments, we randomly divide the dialogue session into two parts to obtain as many training samples as possible. In this study, we keep the number of training samples constant by reusing randomly divided training data while the ratio constraint is satisfied by masking redundant utterances. Because dialogues in OpenSubtitles are too short to support this ablation, we only report results on DailyDialog in Figure 5. We report the variation of four metrics and for ease of comparison, we also report the performance of AdaLabel, RegDG, and D2GPO as dashed horizontal lines. It is clear that random data division performs best. Although constraining the ratio to constant hurts performance, it still outperforms all baselines. According to this study, in different dialogue sessions, responses depend to a different degree on the contexts of the past and the future. Therefore, it is better to fully utilize data without such a ratio constraint.

6 Conclusion

This paper first augments the dialogue generation task with hierarchical duality in dialogue text which was widely ignored. We propose a hierarchical dual learning framework HDLD to utilize both inter-duality and intra-duality, and couple them tightly through the cross-hierarchy distillation mechanism. Both automatic experiments and human judgment demonstrate the effectiveness of hierarchical duality in generating dialogues, which helps HDLD outperform strong baselines. In the future, more under-utilized properties except for duality in dialogues are to be explored. We will further explore the hierarchical characteristics in natural language.

Limitations

Since there are two transformer-based models in HDLD, the major limitation of HDLD is the training cost. It costs twice training time than most baselines. Here is the training time comparison on OpenSubtitles:

Model	Minutes Per Epoch
AdaLabel	7.13
D2GPO	7.16
RegDG	7.83
Ours	14.32

To address this limitation, a potential improvement direction is to fuse the inter-duality approach into one transformer, e.g., sharing parameters between the f_{xy} and f_{yx}. However, we find the significant performance drop when doing so. In the future, the trade-off between performance and cost needs more attention. Besides, like most generative models, if there is malicious information in data, there is no guarantee to avoid bad responses to users, which is a potential risk that needs to pay attention.

Acknowledgement

This work was supported by National Natural Science Foundation of China (NSFC Grant No. 62122089), Beijing Outstanding Young Scientist Program NO. BJJWZYJH012019100020098, and Intelligent Social Governance Platform, Major Innovation & Planning Inter-disciplinary Platform for the "Double-First Class" Initiative, Renmin University of China.

References

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, Minneapolis, Minnesota. Association for Computational Linguistics.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. Bert: Pre-training of deep bidirectional transformers for language understanding.

Shaoxiong Feng, Xuancheng Ren, Hongshen Chen, Bin Sun, Kan Li, and Xu Sun. 2020. Regularizing dialogue generation by imitating implicit scenarios. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 6592–6604, Online. Association for Computational Linguistics.

Di He, Yingce Xia, Tao Qin, Liwei Wang, Nenghai Yu, Tie-Yan Liu, and Wei-Ying Ma. 2016. Dual learning for machine translation. *Advances in neural information processing systems*, 29:820–828.

Solomon Kullback and Richard A Leibler. 1951. On information and sufficiency. *The annals of mathematical statistics*, 22(1):79–86.

Jinpeng Li, Yingce Xia, Rui Yan, Hongda Sun, Dongyan Zhao, and Tie-Yan Liu. 2021. Stylized dialogue generation with multi-pass dual learning. In *Advances in Neural Information Processing Systems*.

Jiwei Li, Michel Galley, Chris Brockett, Jianfeng Gao, and Bill Dolan. 2016. 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, San Diego, California. Association for Computational Linguistics.

Yanran Li, Hui Su, Xiaoyu Shen, Wenjie Li, Ziqiang Cao, and Shuzi Niu. 2017. DailyDialog: A manually labelled multi-turn dialogue dataset. In *Proceedings of the Eighth International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 986–995, Taipei, Taiwan. Asian Federation of Natural Language Processing.

Zuchao Li, Rui Wang, Kehai Chen, Masso Utiyama, Eiichiro Sumita, Zhuosheng Zhang, and Hai Zhao. 2020. Data-dependent gaussian prior objective for language generation. In *International Conference on Learning Representations*.

Pierre Lison and Jörg Tiedemann. 2016. OpenSubtitles2016: Extracting large parallel corpora from movie and TV subtitles. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, pages 923–929, Portorož, Slovenia. European Language Resources Association (ELRA).

Chang Liu, Xu Tan, Chongyang Tao, Zhenxin Fu, Dongyan Zhao, Tie-Yan Liu, and Rui Yan. 2022. ProphetChat: Enhancing dialogue generation with simulation of future conversation. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 962–973, Dublin, Ireland. Association for Computational Linguistics.

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 of the Association for Computational Linguistics*, pages 311–318, Philadelphia, Pennsylvania, USA. Association for Computational Linguistics.

Qiao Qian, Minlie Huang, Haizhou Zhao, Jingfang Xu, and Xiaoyan Zhu. 2018. Assigning personality/profile to a chatting machine for coherent conversation generation. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence*, pages 4279–4285.

M. Schuster and K.K. Paliwal. 1997. Bidirectional recurrent neural networks. *IEEE Transactions on Signal Processing*, 45(11):2673–2681.

Thibault Sellam, Dipanjan Das, and Ankur Parikh. 2020. BLEURT: Learning robust metrics for text generation. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 7881–7892, Online. Association for Computational Linguistics.

Iulian Serban, Alessandro Sordoni, Yoshua Bengio, Aaron Courville, and Joelle Pineau. 2016. Building end-to-end dialogue systems using generative hierarchical neural network models. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 30.

Iulian Serban, Alessandro Sordoni, Ryan Lowe, Laurent Charlin, Joelle Pineau, Aaron Courville, and Yoshua Bengio. 2017. A hierarchical latent variable encoder-decoder model for generating dialogues. *Proceedings of the AAAI Conference on Artificial Intelligence*, 31(1).

Xiaoyu Shen, Hui Su, Wenjie Li, and Dietrich Klakow. 2018. NEXUS network: Connecting the preceding and the following in dialogue generation. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 4316–4327, Brussels, Belgium. Association for Computational Linguistics.

Alessandro Sordoni, Michel Galley, Michael Auli, Chris Brockett, Yangfeng Ji, Margaret Mitchell, Jian-Yun Nie, Jianfeng Gao, and Bill Dolan. 2015. A neural network approach to context-sensitive generation of conversational responses. In *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 196–205, Denver, Colorado. Association for Computational Linguistics.

Chongyang Tao, Changyu Chen, Jiazhan Feng, Ji-Rong Wen, and Rui Yan. 2021. A pre-training strategy for zero-resource response selection in knowledge-grounded conversations. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 4446–4457.

Chongyang Tao, Wei Wu, Can Xu, Wenpeng Hu, Dongyan Zhao, and Rui Yan. 2019. One time of interaction may not be enough: Go deep with an interaction-over-interaction network for response selection in dialogues. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 1–11, Florence, Italy. Association for Computational Linguistics.

Zhiliang Tian, Rui Yan, Lili Mou, Yiping Song, Yansong Feng, and Dongyan Zhao. 2017. How to make context more useful? an empirical study on context-aware neural conversational models. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 231–236.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In *Advances in neural information processing systems*, pages 5998–6008.

Yida Wang, Yinhe Zheng, Yong Jiang, and Minlie Huang. 2021. Diversifying dialog generation via adaptive label smoothing. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 3507–3520, Online. Association for Computational Linguistics.

Yu Wu, Wei Wu, Chen Xing, Ming Zhou, and Zhoujun Li. 2017. Sequential matching network: A new architecture for multi-turn response selection in retrieval-based chatbots. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 496–505, Vancouver, Canada. Association for Computational Linguistics.

Min Yang, Wenting Tu, Qiang Qu, Zhou Zhao, Xiaojun Chen, and Jia Zhu. 2018. Personalized response generation by dual-learning based domain adaptation. *Neural Networks*, 103:72–82.

Xueliang Zhao, Wei Wu, Can Xu, Chongyang Tao, Dongyan Zhao, and Rui Yan. 2020. Knowledge-grounded dialogue generation with pre-trained language models. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 3377–3390, Online. Association for Computational Linguistics.

Xiangyang Zhou, Lu Li, Daxiang Dong, Yi Liu, Ying Chen, Wayne Xin Zhao, Dianhai Yu, and Hua Wu. 2018. Multi-turn response selection for chatbots with deep attention matching network. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1118–1127.

ACL 2023 Responsible NLP Checklist

A For every submission:

☑ A1. Did you describe the limitations of your work?
Following instructions, we add Limitations after Conclusion.

☑ A2. Did you discuss any potential risks of your work?
In Limitations.

☑ A3. Do the abstract and introduction summarize the paper's main claims?
The main claims in the paper are stated in the abstract and in the introduction.

☒ A4. Have you used AI writing assistants when working on this paper?
Left blank.

B ☑ Did you use or create scientific artifacts?

We use open source code and public dataset.

☑ B1. Did you cite the creators of artifacts you used?
We have cited all datasets we use. We have cited open pre-trained models.

☑ B2. Did you discuss the license or terms for use and / or distribution of any artifacts?
We discuss in appendix A.

☐ B3. Did you discuss if your use of existing artifact(s) was consistent with their intended use, provided that it was specified? For the artifacts you create, do you specify intended use and whether that is compatible with the original access conditions (in particular, derivatives of data accessed for research purposes should not be used outside of research contexts)?
Not applicable. Left blank.

☐ B4. Did you discuss the steps taken to check whether the data that was collected / used contains any information that names or uniquely identifies individual people or offensive content, and the steps taken to protect / anonymize it?
Not applicable. Left blank.

☐ B5. Did you provide documentation of the artifacts, e.g., coverage of domains, languages, and linguistic phenomena, demographic groups represented, etc.?
Not applicable. Left blank.

☑ B6. Did you report relevant statistics like the number of examples, details of train / test / dev splits, etc. for the data that you used / created? Even for commonly-used benchmark datasets, include the number of examples in train / validation / test splits, as these provide necessary context for a reader to understand experimental results. For example, small differences in accuracy on large test sets may be significant, while on small test sets they may not be.
In appendix B, we report the dataset statistics.

C ☑ Did you run computational experiments?

In Experiments.

☑ C1. Did you report the number of parameters in the models used, the total computational budget (e.g., GPU hours), and computing infrastructure used?
In terms of parameters, we report the bone architecture of most baselines: transformer with 6 encoder layers and 6 decoder layers, which has certain amount of parameters. In appendix A, we report GPU version. In Limitations, we report computational budget.

The Responsible NLP Checklist used at ACL 2023 is adopted from NAACL 2022, with the addition of a question on AI writing assistance.

☑ C2. Did you discuss the experimental setup, including hyperparameter search and best-found hyperparameter values?
In 4.1.2, appendix A and appendix C, we discuss experimental setup in detail.

☑ C3. Did you report descriptive statistics about your results (e.g., error bars around results, summary statistics from sets of experiments), and is it transparent whether you are reporting the max, mean, etc. or just a single run?
We report p-value in t-test and kappa value of human evaluation agreement.

☑ C4. If you used existing packages (e.g., for preprocessing, for normalization, or for evaluation), did you report the implementation, model, and parameter settings used (e.g., NLTK, Spacy, ROUGE, etc.)?
In 4.1.4, we report evaluation metrics. In appendix A, we report our script is based on NLTK.

D ☑ Did you use human annotators (e.g., crowdworkers) or research with human participants?
In 4.1.4.

☑ D1. Did you report the full text of instructions given to participants, including e.g., screenshots, disclaimers of any risks to participants or annotators, etc.?
In 4.1.4, we summarized three aspects of evaluation instructions. Before human evaluation, we have de-tokenized and re-capitalized the outputs for a fair and solid evaluation, and thus the instructions are relatively concise. In appendix D, we discuss more detials on human evaluation.

☑ D2. Did you report information about how you recruited (e.g., crowdsourcing platform, students) and paid participants, and discuss if such payment is adequate given the participants' demographic (e.g., country of residence)?
In appendix D, we discuss these.

☐ D3. Did you discuss whether and how consent was obtained from people whose data you're using/curating? For example, if you collected data via crowdsourcing, did your instructions to crowdworkers explain how the data would be used?
Not applicable. Left blank.

☐ D4. Was the data collection protocol approved (or determined exempt) by an ethics review board?
Not applicable. Left blank.

☐ D5. Did you report the basic demographic and geographic characteristics of the annotator population that is the source of the data?
Not applicable. Left blank.