Learning to Correct Erroneous Words for Document Grounded Conversations

Junyan Qiu
University of Chinese Academy of Sciences
Beijing, China
qiujunyan2018@ia.ac.cn

Haidong Zhang
Institution of Automation, Chinese Academy of Sciences
Beijing, China
zhanghaidong@ia.ac.cn

Yiping Yang
Institution of Automation, Chinese Academy of Sciences
Beijing, China
yiping.yang@ia.ac.cn

ABSTRACT

Document grounded conversation (DGC) aims to generate informative responses when talking about a document. It is normally formulated as a sequence-to-sequence (Seq2seq) learning problem, which directly maps source sequences, i.e., the context and background documents, to the target sequence, i.e., the response. These responses are normally used as the final output without further polishing, which may suffer from the global information loss owing to the auto-regression paradigm. To tackle this problem, some researches designed two-pass generation to improve the quality of responses. However, these approaches lack the capability of distinguishing inappropriate words in the first pass, which may maintain the erroneous words while rewrite the correct ones. In this paper, we design a scheduled error correction network (SECN) with multiple generation passes to explicitly locate and rewrite the erroneous words in previous passes. Specifically, a *discriminator* is employed to distinguish erroneous words which are further revised by a *refiner*. Moreover, we also apply curriculum learning with reasonable learning schedule to train our model from easy to hard conversations, where the complexity is measured by the number of decoding passes. We conduct comprehensive experiments on a public document grounded conversation dataset, Wizard-of-Wikipedia, and the results demonstrate significant promotions over several strong benchmarks.

CCS CONCEPTS

• **Computing methodologies** → **Natural language generation**.

KEYWORDS

Deep Learning, Natural Language Generation, Dialogue System, Curriculum Learning

ACM Reference Format:
Junyan Qiu, Haidong Zhang, and Yiping Yang. 2023. Learning to Correct Erroneous Words for Document Grounded Conversations. In *2023 12th International Conference on Software and Computer Applications (ICSCA 2023), February 23–25, 2023, Kuantan, Malaysia.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3587828.3587883

ICSCA 2023, February 23–25, 2023, Kuantan, Malaysia

ACM ISBN 978-1-4503-9858-9/23/02.
https://doi.org/10.1145/3587828.3587883

1 INTRODUCTION

Open domain dialogue system is designed to provide entertainment or useful information for users in natural language. It has drawn increasing attention in both research and industry communities due to the applications on chatbots such as Microsoft XiaoIce [19] and Amazong Alexa [18]. Data-driven approaches trained in an end-to-end manner are commonly applied in this field. However, these methods tend to generate generic responses with deficient information, such as *"I don't know"* and *"I like it"*. To alleviate this problem, numerous methods have been proposed to ground conversations with unstructured documents such as books and articles [10, 12].

Document grounded conversation (DGC) aims at generating informative responses that are consistent with the context by exploiting relevant documents. Sequence-to-sequence [20] is a common technique adopted in this field. However, most existing approaches just adopt one-pass decoding to generate a response without further refinement. Motivated by human cognition that people tend to polish a draft when writing sentences, deliberation network with two-pass decoding process is introduced to sequence generation tasks [10, 23]. Generally, a second-pass decoder is employed to polish the raw sequence generated by the first-pass decoder. Despite their effectiveness, these works have difficulties in locating inappropriate words, which may maintain the erroneous words while rewrite the correct ones. In this paper, we introduce a error correction network to distinguish erroneous words from correct ones, and revise the response iteratively.

Furthermore, we notice that it has divergent complexity degrees to rewrite different responses. For example, it is more difficult to respond *"How do you like Gone with the Wind?"* than *"Do you like Gone with the Wind?"*. Correspondingly, it normally takes more steps to polish responses with harder difficulties. Inspired by human education that starting from easy and gradually learning more complex concepts, we incorporate curriculum learning [2] into the response generation process for better generalization capacities.

To sum up, we develop a novel architecture named scheduled error correction network (SECN) upon the vanilla seq2seq framework to rewrite erroneous words iteratively for document grounded conversations. The SECN consists of two major components, including the *error correction network* (ECN) and *curriculum learner*. More concretely, the ECN can be divided into a discriminator and a refiner. At each iteration, the discriminator distinguishes inappropriate words from correct ones. Then the refiner incorporates responses at previous and current iterations weighted by the outputs of the discriminator. To further arrange scheduled learning process for ECN, we evaluate the complexity of each conversation according the number of iterations, and train the model from easy to hard conversations based on the evaluation results.

To conclude, our contributions can be summarized in three-fold: (1) We present a novel error correction network to iteratively polish generated responses. To the best of our knowledge, this is the first work that distinguishes and revises erroneous words for document grounded conversations. (2) Based on the number of polishing iterations, we propose an automatic curriculum framework for the ECN to improve the performance of response generation. (3) We conduct comprehensive experiments on Wizard-of-Wikipedia and results demonstrate the validity of our proposed method.

The rest of the paper is organized as follows. In Section 2, we introduce related work in document grounded conversation and curriculum learning. In Section 3, we depict the detailed architecture of our proposed method. The performance evaluation is given in Section 4, and the conclusion is in Section 5.

2 RELATED WORK

2.1 Document Grounded Conversations

Document grounded conversation has been widely used to address the degeneration problem of responding generically in fully data-driven approaches [4, 11]. For example, Lin et al. [12] incorporated appropriate documents by using recurrent knowledge interaction among response decoding steps. Li et al. [10] encoded multi-turn utterances and relevant document knowledge by presenting an incremental transformer. Lian et al. [11] designed both prior and posterior distributions over document to facilitate knowledge selection. However, these methods generating responses with only one-pass decoding lack the deliberation process, which is a common behavior for human to polish a draft while writing sentences. Recently, deliberation networks that employ two-pass decoders has been introduced to sequence generation tasks, including machine translation [23] and dialogue generation [10, 13]. Nevertheless, very little is explored about distinguishing and correcting erroneous words in the raw responses. In this paper, we build a error correction network upon seq2seq framework revise erroneous words iteratively.

2.2 Curriculum Learning

Curriculum learning (CL) is designed to arrange meaningful training orders for the deep learning model, which is powerful to improve the generalization capacity and convergence rate of models [22]. It can be generally divided into two components. The first one is *complexity measurer* that evaluates the difficulty level of each training sample. The second one called *training scheduler* decides the sequence of data used for training based on the evaluation by the complexity measurer. It is widely applied in NLP tasks, such as natural language generation [3] and dialogue policy learning [14]. For example, [3] proposed several curricula where the dialogue complexity is evaluated by heuristic dialogue attributes, such as specificity, repetitiveness and so on. [14] proposed a curriculum learning framework with automatic complexity measurement for efficient dialogue policy learning. Since the difficulty degree to respond to different context varies and ignoring the various complexity of conversations may decrease the model's performance, in this paper, we combine curriculum learning framework with the error correction network to improve the dialogue generation.

3 PROPOSED METHOD

3.1 Architecture Overview

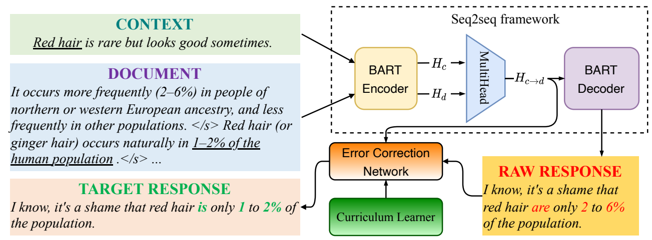

Figure 1: The overall architecture of the proposed method. Text from the context and documents that are used to generate the response is underlined

The overall architecture of our proposed method is presented in Fig. 1, which consists of three major components: (1) *seq2seq framework* generating the raw response based on the dialogue context and documents, (2) *error correction network* that polishes the raw response iteratively and (3) *curriculum learner* arranging learning schedules for the error correction network.

3.2 Seq2seq Framework

Let $C = \{x_1, x_2, \cdots, x_{\ell_c}\}$ denote the dialogue context with ℓ_c words, $\mathcal{D} = \{w_1, w_2, \cdots, w_{\ell_d}\}$ denote the documents with ℓ_d words. The target of seq2seq framework is to sequentially estimate the probability of each word in the target response $\mathcal{R} = \{y_1, y_2, \cdots, y_{\ell_r}\}$ with ℓ_r words conditioned on the context and document:

$$p(\mathcal{R}|C, \mathcal{D}) = \prod_{i=1}^{\ell_r} p(y_i|x_1 \cdots, x_{\ell_c}, w_1, \cdots, w_{\ell_d}, y_1, \cdots, y_{i-1}) \quad (1)$$

To achieve that, we choose BART [8], a large scale pre-trained language model that generalizes bidirectional encoder and unidirectional decoder, as the backbone of seq2seq framework. Firstly the dialogue context and documents are fed into the BART encoder to obtain hidden contextual word representations:

$$\mathbf{H}_c = \text{BART}_{encoder}(C) \quad (2)$$

$$\mathbf{H}_d = \text{BART}_{encoder}(\mathcal{D}) \quad (3)$$

To incorporate document information into the context, we leverage multi-head attention [21] and obtain the document-driven context representations:

$$\mathbf{H}_{c \to d} = \text{MultiHead}(\mathbf{H}_c, \mathbf{H}_d, \mathbf{H}_d) \quad (4)$$

Given the number of attention heads n_{head} hidden nodes d_m, MultiHead(Q, K, V) is defined as:

$$\mathbf{Q}^n = \mathbf{Q}\mathbf{W}_Q^n, \ \mathbf{K}^n = \mathbf{K}\mathbf{W}_K^n, \ \mathbf{V}^n = \mathbf{V}\mathbf{W}_V^n \quad (5)$$

$$\mathbf{head} = \frac{\mathbf{Q}^n(\mathbf{K}^{nT})}{\sqrt{d_m}}\mathbf{V}^n \quad (6)$$

$$\text{MultiHead}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = [\mathbf{head}^1, \mathbf{head}^2, \cdots, \mathbf{head}^{n_{head}}] \quad (7)$$

where $[\cdot,\cdot,\cdots,\cdot]$ denotes the concatenation operation. Notably, it is referred to as the multi-head self-attention if $\mathbf{Q}$, $\mathbf{K}$ and $\mathbf{V}$ are the same sequence embedding. Otherwise it is called multi-head cross-attention. Then, the document-driven context representations are fed into the BART decoder to calculate the response distributions:

$$\mathbf{h}_i^r = \mathrm{BART}_{decoder}(\mathbf{h}_{<i}^r, \mathbf{H}_{c\to d}) \tag{8}$$

where each $\mathbf{h}_i^r$ is used to compute the distribution for the next token: $p(y_{i+1}|\mathbf{h}_{\leq i}^r, \mathbf{H}_{c\to d}) = \mathrm{softmax}(\mathbf{W}^{vocab}\mathbf{h}_i^r)$ and $\mathbf{W}^{vocab}$ is a trainable matrix that converts $\mathbf{h}_i^r$ to logits over the vocabulary.

3.3 Error Correction Network

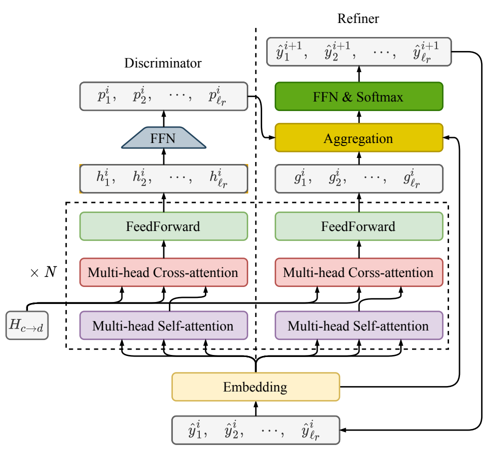

Figure 2: Architecture of error correction network. Layer normalization and residual connection are omitted for simplicity.

Given the raw responses generated by the seq2seq framework, we devise an error correction network (ECN) to polish it within a maximum number of iterations K. ECN consists of two components, i.e., the discriminator and refiner, as show in Fig. 2. At each iteration i, it is fed with the response representation at the previous iteration. The discriminator is responsible for distinguishing erroneous words at the previous iteration and produces probabilities of words deemed inappropriate. The refiner revises the sequence by aggregating response embeddings at previous turn based on the probability produced by the discriminator.

Both the discriminator and refiner contain a stack of N identical transformer layers (Transformer), where each layer is composed of three sub-layers, including multi-head self-attention, multi-head cross-attention and feedforward network. The detailed architecture is illustrated in Fig. 2. A residual connection [6] is employed around each sub-layer followed by a layer normalization [1]. We will introduce the discriminator and refiner in the following.

Discriminator. The discriminator takes as input responses $\hat{\mathcal{Y}}_i = \{\hat{y}_1^i, \hat{y}_2^i, \cdots, \hat{y}_{\ell_r}^i\}$ generated by the refiner at last iteration[1]. We employ transformer blocks to obtain their corresponding hidden representations:

$$[\mathbf{h}_1^i, \mathbf{h}_2^i, \cdots, \mathbf{h}_{\ell_r}^i] = \mathrm{Transformer}(\phi(\hat{\mathcal{Y}}_i), \mathbf{H}_{c\to d}) \tag{9}$$

where $\phi(\cdot)$ is the function that converts tokens to embeddings. Then the representations are fed into a feedforward network (FFN):

$$p_j^i = \mathrm{sigmoid}(\mathbf{W}\mathbf{h}_j^i + \mathbf{b}) \tag{10}$$

where $\mathbf{W}$ and $\mathbf{b}$ are trainable parameters. $\mathbf{p}^i = [p_i^i, p_2^i, \cdots, p_{\ell_r}^i] \in [0,1]$ means the probability of the corresponding word needed to be devised. If $p_j^i < p, \forall j \in [1, 2, \cdots, \ell_r]$[2], we assume that all words at previous iteration are correct and the iteration terminates. p is a hyperparameter which is set to 0.5 in this paper. To train the discriminator, we devise a supervised target that indicates whether the word should be revised as follows:

$$\varphi_j = \begin{cases} 0 & \text{if } y_j = \hat{y}_j^i \\ 1 & \text{otherwise} \end{cases} \tag{11}$$

The objective is defined as minimizing the cross entropy loss as follows:

$$\mathcal{L}_{DIS} = \sum_{j=1}^{\ell_r} -\varphi_j \log p_j^i \tag{12}$$

Refiner. Similar to the discriminator, we employ transformer blocks to map responses generated at last iteration $\hat{\mathcal{Y}}^i$ to hidden representations $\mathbf{G}^i = [\mathbf{g}_1^i, \mathbf{g}_2^i, \cdots, \mathbf{g}_{\ell_r}^i]$. Then, we aggregate $\hat{\mathcal{Y}}^i$ and $\mathbf{G}^i$ weighted by $\mathbf{p}^i$. The aggregated representation is fed into a classifier to update responses at the last iteration. Mathematically:

$$\mathbf{f}_j^i = p_i^j \cdot \mathbf{g}_j^i + (1 - p_j^i) \cdot \phi(\hat{y}_j^i) \tag{13}$$

$$\mathbf{q}_j^{i+1} = \mathrm{softmax}(\mathbf{W}^{vocab}\mathbf{f}_j^i + \mathbf{b}^{vocab}) \in \mathbb{R}^{|\mathcal{V}|} \tag{14}$$

where $\mathbf{W}^{vocab}$ and $\mathbf{b}^{vocab}$ are trainable parameters that map $\mathbf{f}_i^j$ to probabilities over the vocabulary $\mathcal{V}$. $\hat{y}_j^{i+1}$ is decided by the largest probability of the corresponding word in the vocabulary. The response is treated as the final response $\hat{\mathcal{Y}}$ if the polishing process terminates, i.e., it reaches the maximum number of iteration K or the termination condition is satisfied. We apply a cross entropy loss $\mathcal{H}(\cdot)$ to quantify the difference between the true and generated responses:

$$\mathcal{L}_{ECN} = \sum_j v_j \mathcal{H}(\hat{\mathcal{Y}}_j, \mathcal{Y}_j) \tag{15}$$

where v_j is a parameter that decides whether the j_{th} sample is applied for training at current stage. We will introduce it in the following section.

[1]Notably, the discriminator regards the raw response generated by the seq2seq framework as the initial input at the first iteration.
[2]We call this *termination condition* for simplicity.

3.4 Curriculum Learner

In this section, we aim to decide the order of model training. Since it has different difficulty degrees for the model to respond, the conversations are sorted from easy to hard. Then, we train the model from simple ones and gradually select more complicated samples to join the training process of the model.

Complexity Evaluation. The complexity of responding to different context diverges significantly which may result in great difficulties for the model to learn efficiently. Thus, it is important to evaluate dialogue complexity and arrange training schedules for the model. To this end, we propose a novel complexity evaluation based on the number of iterations k needed to perfect the response. The more iterations are needed, the harder it is to respond.

Training Scheduler. We leverage a parameter λ to control the complexity of conversation selected to train the ECN model. Mathematically, for the j_{th} sample, v_j in Equation 15 is defined as follows:

$$v_j = \begin{cases} 1 & \text{if } k < \lambda \\ 0 & \text{otherwise} \end{cases} \tag{16}$$

If k is smaller than a threshold λ, then this sample is deemed as a easy one and should be used for model training. Otherwise, it is too hard for the model to train on at the current training step. To be specific, λ is defined as linear pacing function regarding the current training step t given the total number of training steps T:

$$\lambda = K \cdot (\lambda_0 + \frac{1 - \lambda_0}{T} \cdot t) \tag{17}$$

where λ_0 is a hyperparameter controlling the learning pace. In this paper, λ_0 is set to 0.3.

4 EXPERIMENTS

4.1 Datasets and Experiment Details

We demonstrate the effectiveness of SECN on a public document grounded conversation dataset, Wizard-of-Wikipedia (WoW) [5]. WoW is a chit-chat dataset between two interlocutors. One of them with access to the relevant document plays the role of a knowledge expert (or referred to as the wizard), while the other one as a curious learner (the apprentice) [5]. The background documents are retrieved from Wikipedia. Each wizard turn is associated with about 60 sentences and each sentence contain about 30 words [15].

In this paper, we use base version of BART with 139M parameters to construct the seq2seq framework. The number of transformer layers N is set to 6. Adam [7] is applied to optimize the model parameters with learning rate of 1×10^{-5}. The maximum length of the context, document and response is set to 256, 512 and 64 respectively, and we will truncate the text that exceeds the length limits. The maximum number of polishing iterations K is set to 5. We utilize perplexity (**PPL**), **BLEU** [16] and Distinct (**DIST**) [9] to evaluate the quality of generated responses.

4.2 Baselines

We carry out contrast experiments to compare our model with following strong baselines:

- **ITDD:** A transformer-based model that represented multi-turn dialogue context and documents incrementally, and designed a deliberation decoder to generate responses in two passes [10].
- **DRD:** A low-resource model that devised a disentangled response decoder with copy mechanism and designed a two-stage framework to learn it in a low-resource setting [24].
- **TSLF:** A three-stage learning framework to learn knowledge-grounded dialogue generation with few resource [15].
- **DGG-FA:** Prabhumoye et al. [17] introduced context driven representation and document headed attention to enable encoder-decoder models to focus on documents.

Table 1: Comparison between SECN and baselines on Wizard-of-Wikipedia. † is produced based on the code provided by [10]. The results of DRD and TSLF are reported in the full data setting. The best results are marked in bold.

Models	PPL(%)↓	BLEU-2/4↑		DIST-1/2↑	
ITDD†	19.8	10.5	7.1	9.3	32.1
DRD	23.0	11.5	5.5	-	-
TSLF	**14.4**	13.9	6.7	9.5	38.3
DGG-FA	-	15.1	8.2	-	-
SECN(Ours)	14.9	**15.7**	**9.6**	**12.3**	**42.1**
-discriminator	21.3	11.4	6.7	8.9	37.1
-error correction network	25.6	10.3	6.3	8.7	33.1
-curriculum learner	15.1	13.1	8.8	10.9	40.7

4.3 Experiment Results

Table 1 shows the main results of our model and other baselines. It is clear that our model achieves new state-of-the-art performance in terms of BLEU and DIST, and achieves comparable PPL with TSLF even thought it is pre-trained on two auxiliary datasets to improve its representation ability. The results indicate the superiority of our model in generating coherent and informative responses. Particularly, when compared with ITDD, which also employs multi-pass response refinement, our method outperforms it with a considerable margin in all metrics. This is because the discriminator, which distinguishes erroneous words, can provide explicit guidance to the refiner to correct them.

To further evaluate the effectiveness of each component, we conduct several ablation experiments. The results are presented in the last block of Table 1. It is clear that removing part of the modules affects the performance differently. Specifically, the probability of identifying erroneous words is discarded to verify the validity of the discriminator. Formally, Equation 13 is changed to $\mathbf{f}_j^i = \mathbf{g}_i^j + \phi(y_i^j)$. As can be seen from the results, the performance decreases significantly in all metrics, which is similar to the performance of ITDD. The result provides quantitative evidence to our claim that the capability of the discriminator to distinguish erroneous words facilitates generating more fluent and coherent responses. When the entire error correction network is discarded, it exhibits even more severe performance degradation. This is because the refiner can revise erroneous words in previous iterations based on evaluation of the discriminator. It is insufficient to generate responses with high quality with only one-pass decoding. Furthermore, casting off the curriculum learner also drops the performance mildly. It demonstrates that the model benefits from meaningful learning schedule by gradually increasing the difficulty of conversations.

4.4 Discussion

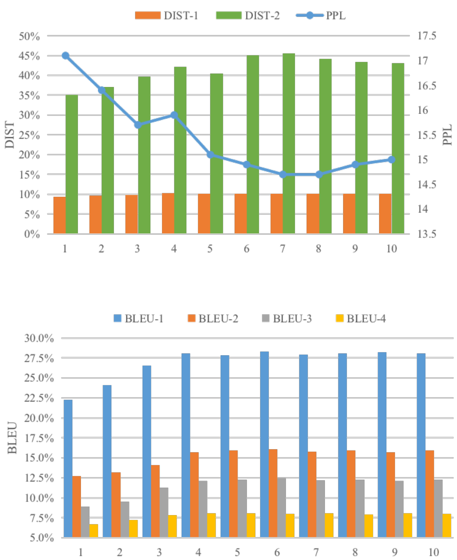

Figure 3: Performance of SECN with different K

Additionally, we are curious about what impact will different choices of K (i.e., the maximum number of refinement rounds) have on the final results. As shown in Fig. 3, the performance generally gets better as K increases from 1 to 5, indicating that the quality of responses can be improved by conducting multi-pass refinements. However, as K continues to grow, the performance hardly improves, which suggests that the optimal value of K should be carefully explored. Since it normally takes more time and computation resources with larger K, we set $K = 5$ to balance the trade-off between performance and efficiency.

4.5 Case Study

Table 2 presents an example of some responses generated by different baselines. It is obvious that our model can generate more coherent and informative responses than others. More specifically, response generated by the seq2seq model is generally consistent with the context and document. However, there still exists redundant pieces with syntax errors, i.e., *"Yes, yes"* at the beginning of the response. As for SECN with discriminator discarded, the appropriate words generated by the seq2seq framework are revised while the erroneous words remains. When the discriminator is equipped, SECN can accurately locate the erroneous words and revise them with the refiner in multiple iterations.

Table 2: Responses generated by different baselines, golden means the ground truth. wo. DIS means SECN without discriminator.

Document	D1: In 1873, German immigrants Adolph Coors and Jacob Schueler from Prussia emigrated to the United States and established a brewery in... D2: Coors operates a brewery in Golden, Colorado, that is the largest single brewery facility in the world. ...
Context	Who founded Coors Brewing Company?
Golden	R: In 1873, Adolph Coors and Jacab Schueler founded the Brewery.
Seq2seq	R: Yes yes, Adolph and and Jacob Schueler.
wo. DIS	R: Yes is, Adolph and and and Schueler in 1873.
SECN	R: In 1873, Adolph Coors and Jacob Schueler established a Brewery.

To further demonstrate the validity of the refiner, we list responses generated by the seq2seq framework and the SECN. Examples are presented in Table 3, the SECN improves the response quality in two dimensions: (1) Informativeness. It can not only respond properly, but also enrich the response by providing auxiliary information given the background document (case 1). (2) Fluency. Additionally, the SECN is able to modify some ungrammatical pieces in response (case 2 and 3).

Table 3: Responses generated by Seq2seq and SECN. Only relevant documents are presented in the table to save space.

Document: The appearance of blue and green results from the Tyndall scattering of light in the stroma... Context: Blue eyes are quite gorgeous.	Seq2seq: Yes, I agree.
	SECN: Yes, I agree. It results from the Tyndall scattering of light in the stroma.
Context: Skunks are repulsive animals.	Seq2seq: Yes, it is.
	SECN: Yes, they are.
Context: I love dogs, particularly Huskies, I have two. What about you?	Seq2seq: I like like dogs too.
	SECN: I like dogs too.

5 CONCLUSION

In this paper, we propose a novel architecture, scheduled error correction network (SECN) upon the traditional seq2seq framework. It is comprised of two components: 1) error correction network (ECN) that locates and correct erroneous words, and 2) curriculum learner arranging meaningful learning schedules for the ECN. Extensive experiments demonstrate that our model generates more coherent and informative responses, and brings significant improvements over several strong baselines.

REFERENCES

[1] Jimmy Lei Ba, Jamie Ryan Kiros, and Geoffrey E Hinton. 2016. Layer normalization. *arXiv preprint arXiv:1607.06450* (2016).
[2] Yoshua Bengio, Jérôme Louradour, Ronan Collobert, and Jason Weston. 2009. Curriculum learning. In *Proceedings of the 26th annual international conference on machine learning*. 41–48.
[3] 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. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 34. 7472–7479.

[4] Xiuyi Chen, Feilong Chen, Fandong Meng, Peng Li, and Jie Zhou. 2021. Unsupervised knowledge selection for dialogue generation. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*. 1230–1244.

[5] Emily Dinan, Stephen Roller, Kurt Shuster, Angela Fan, Michael Auli, and Jason Weston. 2018. Wizard of Wikipedia: Knowledge-Powered Conversational Agents. In *International Conference on Learning Representations*.

[6] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 770–778.

[7] Diederik P Kingma and Jimmy Ba. 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* (2014).

[8] Mike Lewis, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Veselin Stoyanov, and Luke Zettlemoyer. 2020. BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. 7871–7880.

[9] Jiwei Li, Michel Galley, Chris Brockett, Jianfeng Gao, and William B 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*. 110–119.

[10] Zekang Li, Cheng Niu, Fandong Meng, Yang Feng, Qian Li, and Jie Zhou. 2019. Incremental Transformer with Deliberation Decoder for Document Grounded Conversations. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. 12–21.

[11] Rongzhong Lian, Min Xie, Fan Wang, Jinhua Peng, and Hua Wu. 2019. Learning to Select Knowledge for Response Generation in Dialog Systems. In *IJCAI International Joint Conference on Artificial Intelligence*. 5081.

[12] Xiexiong Lin, Weiyu Jian, Jianshan He, Taifeng Wang, and Wei Chu. 2020. Generating informative conversational response using recurrent knowledge-interaction and knowledge-copy. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. 41–52.

[13] Cao Liu, Shizhu He, Kang Liu, and Jun Zhao. 2019. Vocabulary pyramid network: Multi-pass encoding and decoding with multi-level vocabularies for response generation. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. 3774–3783.

[14] Sihong Liu, Jinchao Zhang, Keqing He, Weiran Xu, and Jie Zhou. 2021. Scheduled Dialog Policy Learning: An Automatic Curriculum Learning Framework for Task-oriented Dialog System. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*. 1091–1102.

[15] Shilei Liu, Xiaofeng Zhao, Bochao Li, Feiliang Ren, Longhui Zhang, and Shujuan Yin. 2021. A Three-Stage Learning Framework for Low-Resource Knowledge-Grounded Dialogue Generation. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*. 2262–2272.

[16] 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*. 311–318.

[17] Shrimai Prabhumoye, Kazuma Hashimoto, Yingbo Zhou, Alan W Black, and Ruslan Salakhutdinov. 2021. Focused Attention Improves Document-Grounded Generation. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. 4274–4287.

[18] Ashwin Ram, Rohit Prasad, Chandra Khatri, Anu Venkatesh, Raefer Gabriel, Qing Liu, Jeff Nunn, Behnam Hedayatnia, Ming Cheng, Ashish Nagar, et al. 2018. Conversational ai: The science behind the alexa prize. *arXiv preprint arXiv:1801.03604* (2018).

[19] Heung-Yeung Shum, Xiao-dong He, and Di Li. 2018. From Eliza to XiaoIce: challenges and opportunities with social chatbots. *Frontiers of Information Technology & Electronic Engineering* 19, 1 (2018), 10–26.

[20] Ilya Sutskever, Oriol Vinyals, and Quoc V Le. 2014. Sequence to sequence learning with neural networks. *Advances in neural information processing systems* 27 (2014).

[21] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is All you Need. *Advances in Neural Information Processing Systems* 30 (2017), 5998–6008.

[22] Xin Wang, Yudong Chen, and Wenwu Zhu. 2021. A survey on curriculum learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2021).

[23] Yingce Xia, Fei Tian, Lijun Wu, Jianxin Lin, Tao Qin, Nenghai Yu, and Tie-Yan Liu. 2017. Deliberation networks: Sequence generation beyond one-pass decoding. *Advances in neural information processing systems* 30 (2017).

[24] Xueliang Zhao, Wei Wu, Chongyang Tao, Can Xu, Dongyan Zhao, and Rui Yan. 2019. Low-Resource Knowledge-Grounded Dialogue Generation. In *International Conference on Learning Representations*.