Knowledge-Grounded Response Generation with Deep Attentional Latent-Variable Model

Hao-Tong Ye, Kai-Ling Lo, Shang-Yu Su, Yun-Nung Chen
Department of Computer Science and Information Engineering
National Taiwan University
Taipei, Taiwan

Abstract

End-to-end dialogue generation has achieved promising results without using handcrafted features and attributes specific for each task and corpus. However, one of the fatal drawbacks in such approaches is that they are unable to generate informative utterances, so it limits their usage from some real-world conversational applications. This paper attempts at generating diverse and informative responses with a variational generation model, which contains a joint attention mechanism conditioning on the information from both dialogue contexts and extra knowledge.

Introduction

Dialogue-related research can be mainly categorized into two branches: (1) task-oriented: the system trying to help users complete a certain task (2) chit-chat: the system that can handle casual conversations that do not belong to any specific domain. Recently, how to bridge these two branches has become a new research direction in conversation modeling, where the system can generate useful and fact-grounded responses via external knowledge without domain constraints (Hori and Hori 2017; Yoshino et al. 2018; Ghazvininejad et al. 2017).

Prior work showed that end-to-end neural models are capable of generating sound responses for chit-chat dialogues in a data-driven way, without using handcrafted features specific for each corpus or different task (Vinyals and Le 2015; Sordoni et al. 2015; Li et al. 2016b; Gao, Galley, and Li 2018). However, such systems still highly rely on the information stored in training corpora, which is constrained by time, space, and speakers during data collection. Also, the systems lack the direct access to external information and the knowledge-grounded mechanism; therefore they cannot effectively retrieve real-world common senses and facts in order to respond properly. This fundamental limitation makes the end-to-end systems difficult to complete tasks (Li et al. 2017; Peng et al. 2018) or generate fact-grounded chit-chat (Ghazvininejad et al. 2017).

On the other hand, for traditional dialogue systems, we can easily insert external knowledge and facts into the model with the cost of detailed hand-coding features, which requires a large amount of pre-processing and data labeling. For those tasks or corpora related to complex information or professional knowledge, pre-processing and annotations are difficult to acquire, thus making this approach impractical.

In this work, we propose an end-to-end variational model with the attention mechanism that models the interactions between dialogue contexts and external knowledge. This model is capable of balancing between scalability and generalization of neural models and provides more factual and knowledge-grounded responses compared to the traditional systems. Such extension is especially important for a conversational model deployed in systems requiring more relevant and informative interactions (e.g. the recommendation system).

To test the ability of generating knowledge-grounded responses, the seventh Dialog System Technology Challenge (DSTC7) proposed a benchmark Reddit dataset, in which the conversations are accompanied with a link to an external webpage that may contain related facts and knowledge. A dataset example is shown in Table 1, where the last two responses share the same contexts, and the fact retrieved by our model contains related knowledge given the conversation.

Proposed Approach

The task is to generate a suitable response that contains grounded knowledge or factual information given its conversation contexts.

Model Framework

The main difference between this task and others is the context-relevant facts, which are retrieved from the website links mentioned at the beginning of conversations. This external knowledge provides our model cues about how to ground the information in the response. Therefore, we first build a retrieval model to effectively obtain the fact containing the relevant knowledge, and then learn the conversation model to generate the knowledge-grounded response. Below we describe the detail of the proposed conversation model, where given a conversation con texts and its related facts, the goal is to generate the next probable sentence with informative knowledge.

Conversation:
til monty python member terry gilliam was author j . k rowling 's first choice to direct the first harry potter movie , but was rejected for chris columbus . inan interview he said " i was the perfect guy to do harry potter ... i mean , chris columbus ' versions are terrible . just dull . pedestrian " https://en.wikipedia.org/wiki/terry_gilliam
—- gilliam would have been great - but we 'd still be waiting on the second movie .
——— he should do an animated version
————— harry potter & the giant soft gradient foot
—- they hired chris columbus due to his experience directing child actors .
——— i also think he 's really good at seeing things from a kid 's imagination . those first 2 movies really seemed like someone went into my head and said " ok we 're going to film a movie here!"
——— came here to say this. iirc, he was hired specifically because he was good with kids... which gilliam had little experience with. i think they turned out very well, very true to the books.
Retrieved top-1 fact:
j . k . rowling , the author of the harry potter series , is a fan of gilliam's work . consequently , he was rowling's first choice to direct harry potter and the philosopher's stone in 2000 , but warner bros . ultimately chose chris columbus for the job . [32] in response to this decision , gilliam said that " i was the perfect guy to do harry potter . i remember leaving the meeting , getting in my car , and driving for about two hours along mulholland drive just so angry . i mean , chris columbus ' versions are terrible . just dull . pedestrian . " [33] in 2006 , gilliam said that he found alfonso cuarn ' s harry potter and the prisoner of azkaban to be " really good ... much closer to what i would've done . " [34] in retrospect , however , gilliam has stated that he wouldn't have liked to direct any potter film . in a 2005 interview with total film , he said that he would not enjoy working on such an expensive project because of interference from studio executives . [35]

Table 1: Dataset example from subreddit *todayilearned*. The horizontal lines indicate the tree-structure of the conversation; we can see that the last two responses share the same contexts. The shown fact retrieved by our model is considered the most relevant to the given conversation among all facts extracted by the official script from the wikipedia page.

Conversation Model

For each conversation, our model takes dialogue contexts $\mathbf{C}$ and context-relevant facts $\mathbf{F}$ as the input, and outputs the fact-grounded response $\mathbf{R}$. Specifically, $\mathbf{C} = \{c_i\}_{i=1}^{N_c}$, where $c_n = \{c_{n,j}\}_{j=1}^{T_n^c}$ is a sequence of word embeddings in the n-th utterance of the conversation. For the fact, $\mathbf{F} = \{f_i\}_{i=1}^{N_f}$, where $f_n = \{f_{n,j}\}_{j=1}^{T_n^f}$ is a sequence of word embeddings of n-th fact. The generated response is formulated as $\mathbf{R} = \{r_i\}_{i=1}^{T^r}$. In our model, we treat the conversation utterances and facts as two sequences, with a special token used to separate utterances in a dialogue or a fact; that is, the contexts and facts are turned into $\mathbf{C} = \{c_i\}_{i=1}^{N}$ and $\mathbf{F} = \{f_j\}_{j=1}^{M}$ respectively, where c_i and f_j are word embedding vectors.

First, we use two separate encoders, Enc_C and Enc_H, to encode the dialogue contexts and facts respectively. The encoded contexts and facts $H_C = \{h_i^c\}_{i=1}^{N}$, $H_F = \{h_j^f\}_{j=1}^{M}$ are fed into the attention module, and then the decoder generates the fact-grounded response. In our model, with the encoded contexts and facts, the decoder generates the response in an auto-regressive way, which is commonly called as a sequence-to-sequence model. For each step, the output of the decoder o_t is calculated from previous output o_{t-1} and the encoded information H_C and H_F:

$$o_t = \text{Dec}(o_{t-1}, \text{Attn}(o_{t-1}, H_C, H_F)). \tag{1}$$

The output of the decoder, o_t, is then projected to the vocabulary through a linear layer followed by a `softmax` activation. The proposed model is illustrated in Figure 1, where there are several encoders that focus on different types of information. During generation, this model proposes 1) a *fact-grounded attention* mechanism that can explicitly consider the contexts snd facts and 2) a *conditional variational generation* model that can produce diverse and informative responses. The detail of two module is described below.

Fact-Grounded Attention

In order to capture the relations between these three types of information, dialogue contexts, facts, and responses, we apply three attention variants to model their interactions (Bahdanau, Cho, and Bengio 2014): *context-only attention*, *parallel attention*, and *context-guided fact attention* detailed below.

Context-Only Attention One simple attention baseline only uses the information from contexts to generate the response. That is, with the last-step output o_{t-1} and the encoded information H_C, H_F, the attention is calculated as:

$$\begin{aligned} \text{Attn}(o_{t-1}, H_C, H_F) &= \sum_{i=1}^{N} \alpha_i^t h_i^c, \\ e_i^t &= v_c^T \tanh(W_1^c o_{t-1} + W_2^c h_i^c) + b_c, \\ \alpha^t &= \text{softmax}(e^t), \end{aligned} \tag{2}$$

where v_c, W_1^c, W_2^c, and b_c are trainable parameters.

Parallel Attention In order to well utilize the facts, one trivial solution is to consider facts as the additional contexts;

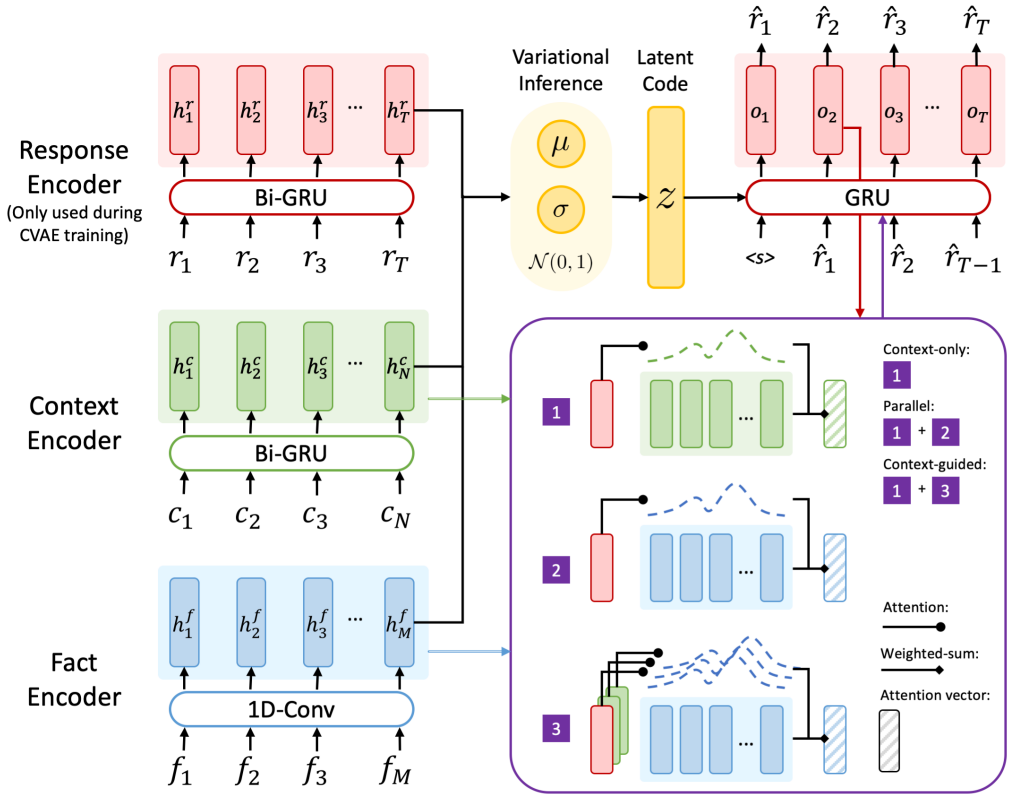

Figure 1: Illustration of the proposed model architecture.

that is:

$$\begin{aligned}\text{Attn}(o_{t-1}, H_C, H_F) &= \left[\sum_{i=1}^{N} \alpha_i^t h_i^c; \sum_{j=1}^{M} \beta_j^t h_j^f\right] \\ m_j^t &= v_f^T \tanh(W_1^f o_{t-1} + W_2^f h_j^f) + b_f, \\ \beta^t &= \text{softmax}(m^t),\end{aligned} \tag{3}$$

and α^t is the same as the context-only attention. v_f, W_1^f, W_2^f, b_f are trainable parameters.

Context-Guided Fact Attention To better model the interaction between contexts and facts, we proposed to use the information from contexts to guide the attention towards facts. Specifically, we modify the attention on facts from the parallel attention as below. We first calculate the attention distribution from contexts to facts,

$$\begin{aligned}M_{i,j} &= v_g^T \tanh(W_1^g h_i^c + W_2^g h_j^f) + b_g, \\ m_j^c &= \sum_{i=1}^{N} M_{i,j}, \\ \beta^c &= \text{softmax}(m^c).\end{aligned} \tag{4}$$

Then, for each step, we calculate the attention from the last-step output to facts, and take the mean of two distributions as the final attention distribution on facts:

$$\begin{aligned}\hat{m}_i^t &= v_o^T \tanh(W_1^o o_{t-1} + W_2^o h_i^c) + b_o, \\ \hat{\beta}^t &= \text{softmax}(\hat{m}^t), \\ \beta^t &= \frac{\hat{\beta}^t + \beta^c}{2},\end{aligned} \tag{5}$$

where v_g, v_o, W_1^g, W_2^g, W_1^o, W_2^o, b_g, and b_o are trainable parameters. Hence, the obtained attention is guided by the contextual information.

Conditional Variational Generation

The conversations in the dataset for the DSTC7 challenge are tree-like structures, where for each context, there may be more than one reference responses. This is also an important perspective for the natural conversations: for arbitrary dialogue contexts, there are usually not only one unique way to respond to it.

With the above consideration, we take the benefit from the variational autoencoder for this task, which has the better capability of capturing such relation than a simple seq2seq model (Kingma and Welling 2013; Sohn, Lee, and Yan 2015). The detail of the variational model is described below.

CVAE for Dialogue Generation For each conversation, we represent it via four random variables: the desired response R, the contexts and facts, C and F, and a latent vari-

	Time Period	Before Filter	After Filter
Train	2015-01∼2016-12	1,101,684	142,750
Dev	2017-01∼2017-06	116,858	14,875

Table 2: Statistics of the used dataset.

able z. The conditional probability $p(R, z \mid C, F)$ can be rewritten as:

$$p(R, z \mid C, F) = p(R \mid C, F, z)p(z \mid C, F). \quad (6)$$

We model the probability $p(R \mid C, F, z)$ and $p(z \mid C, F)$ using the parameters θ and ϕ respectively. Under the variational autoencoder (VAE) framework, we can interpret θ and ϕ as the decoder and the encoder; by setting up a Bayesian prior $p(z \mid C, F)$, our optimization target $p_\theta(R \mid C, F)$ becomes the variational lower bound (ELBO):

$$\begin{aligned}\log p_\theta(R \mid C, F) \geq &- \mathrm{KL}(q_\phi(z \mid R, C, F) \parallel p(z \mid C, F)) \\ &+ \mathbb{E}_{q_\phi(z|R,C,F)}[\log p_\theta(R \mid C, F, z)].\end{aligned} \quad (7)$$

In our model, the prior $p(z \mid C, F)$ is set as $\mathcal{N}(0, I)$.

Annealing Loss of KL Divergence As mentioned above, the optimization target, which is the variational lower bound of $\log p_\theta(R \mid C, F)$, is composed of two sub-goals: one is to minimize the KL divergence between the prior and the conditional encoder probability q_ϕ; another is to maximize the reconstruction probability.

It is found that the model tends to minimize the KL divergence instead of reducing the reconstruction error during early training, resulting in a KL vanishing issue. In order to alleviate the strong bias on minimization of KL divergence, we apply the annealing loss trick to scale down the effect of the KL term at the beginning of training for improving the performance (Bowman et al. 2016).

Training

The proposed model is trained to generate the responses using the CVAE objective, where the attention mechanisms enforce the responses to cover the fact-related information for *knowledge-grounded response generation*.

Experiments

To evaluate the proposed model, we conduct the experiments on the DSTC7 challenge. The used dataset and the experimental setting are described below. Then the results are analyzed in terms of objective and subjective evaluation metrics.

Dataset

The dataset used in DSTC7-Track2 is crawled from Reddit with the scripts[1], which consists of discussions from subreddits like *todayilearned*, *worldnews*, *movies*, etc. In the dataset, the posts include a link to an external webpage, from which the facts for each conversation are then extracted.

In order to encourage our conversation model to contain the factual information, we process this dataset to make sure the conversations in which the context and provided facts are relevant. The processing procedure is described as:

1. **Fact relevance**: Because the facts are extracted from the HTML source codes of webpages, some of them lack the relevant information (e.g. metadata), we use TF-IDF to rank all facts and keep the top-1 fact as the relevant knowledge for ensuring better data quality.
2. **Knowledge-grounded response**: Because the discussions in some conversations may deviate from the original topic, making all facts being irrelevant to the dialogue contexts, we thus filter out data samples where the response and the retrieved fact have no common words without considering punctuations and stopwords[2]. This procedure ensures the training data to match our goal about knowledge-grounded responses.

Due to the limitation of computation resources (one GTX 1080), we use only a subset of training data, and discard the data samples with the responses longer than 20. Table 2 shows the detailed statistics of the dataset after our processing.

Training Details

Considering that the dataset contains a large amount of Internet slangs and spoken English, we train a 100 dimension word embeddings via *GLoVe* from train and development conversations and facts (Pennington, Socher, and Manning 2014). We truncate the context to the last 100 tokens and the fact to the first 500 tokens.

The context encoder Enc_C is a 2-layer bidirectional GRU (Cho et al. 2014) with hidden size 128; the fact encoder Enc_H is a convolutional network with 1,2,3 width filters, and 128 feature maps per filter. The decoder Dec is a 2-layer unidirectional GRU with the hidden size 128. For the CVAE variants, another 2-layer bidirectional GRU with the hidden size 128 is used to encode the responses.

Our models are trained using the teacher-forcing mechanism to maximize the likelihood of generating $\mathbf{R} = \{r_i\}_{i=1}^{T^r}$. We used `adam` (Kingma and Ba 2014) with the default setting as our optimizer. During testing, we apply the beam search where the beam size is 8.

Results

In the experiments, we perform two sets of evaluation, automatic evaluation and human evaluation, to better validate our generated results.

Automatic Evaluation Our evaluation metrics include BLEU (Papineni et al. 2002), METEOR (Banerjee and Lavie 2005), NIST (Doddington 2002), diversity (Li et al. 2016a) and entropy (Zhang et al. 2018) scores. We use the implementation in the Python package nlg-eval[3] for BLEU and METEOR scores (Sharma et al. 2017), and the NLTK toolkit to calculate NIST scores. Our results are shown in Table 3.

[1] `https://github.com/DSTC-MSR-NLP/DSTC7-End-to-End-Conversation-Modeling`

[2] We used stopwords defined in spaCy.

[3] `https://github.com/Maluuba/nlg-eval`

Model	B-1	B-2	B-3	B-4	N-1	N-2	N-3	N-4	MET	Div-1	Div-2	Ent-1	Ent-2	Ent-3	Ent-4
Baseline	2.861	.566	.143	.041	.007	.007	.007	.007	2.439	.004	.012	3.937	4.958	5.504	5.996
CO	2.634	.513	.130	.038	.006	.006	.006	.006	2.417	**.013**	.028	4.137	5.392	6.148	6.734
+CVAE	2.690	.527	.145	.042	.007	.007	.007	.007	2.371	**.013**	**.030**	4.204	**5.436**	**6.239**	**7.049**
PA	3.698	.763	.200	.063	.020	.021	.021	.021	2.574	.012	.027	**4.244**	5.378	6.040	6.576
+CVAE	2.449	.538	.129	.033	.009	.009	.009	.009	2.301	.011	.027	4.120	5.338	6.125	6.775
CG	2.142	.443	.124	.040	.004	.004	.004	.004	2.258	.011	.026	4.131	5.300	6.082	6.820
+CVAE	**3.898**	**.817**	**.223**	**.074**	**.023**	**.024**	**.024**	**.024**	**2.620**	.012	.027	4.089	5.220	5.916	6.427

Table 3: The automatic evaluation results of the baselines and the proposed methods. Baseline is a context-to-response seq2seq model without attention. CO, PA, CG correspond to context-only attention, parallel attention and context-guided attention respectively. (B: BLEU; N: Nist; MET: METEOR)

	Model	Context Relevance	Interest	Fluency	Knowledge Relatedness	Average
Offline	PA	2.47±0.86	2.37±0.75	4.13±0.85	2.19±0.87	2.79
	PA+CVAE	2.40±0.81	2.38±0.77	4.00±0.92	2.10±0.86	2.72
	CG	2.25±0.83	2.18±0.76	3.86±1.07	2.02±0.83	2.58
Official	Submitted (CG+CVAE)	2.52±0.04	2.40±0.05	-	-	2.46
	Best	3.09±0.04	2.87±0.05	-	-	2.94
	Human	3.61±0.04	3.49±0.04	-	-	3.55

Table 4: Human evaluation results in our offline and the official evaluation.

It can be found that the context-guided attention model with CVAE (CG+CVAE) achieves better performance for most metrics in terms of the similarity between the generated responses and the ground truth responses. This justifies the effectiveness of our context-guided attention, because its goal is to generate responses containing more relevant knowledge, and the metrics slightly measure the relatedness. However, the context-only attention with CVAE (CO+VVAE) obtains the higher diversity, which is also important for this generation task. The results show the small improvement achieved by the proposed CVAE model in terms of the generation quality and the diversity.

Human Evaluation In order to understand the effect of our fact-grounded attention and variational generation, we conduct human evaluation on three proposed methods: the parallel attention model as our baseline (PA), compared with the parallel attention with variational generation (PA+CVAE), and the context-guided attention (CG). First, we randomly sample 100 testing samples that fulfill the following two conditions:

1. Each response has at least 3 words, because some methods tend to produce very short responses, which is hard to evaluate.
2. Due to the goal about fact-grounded generation, we make sure that the contexts and the retrieved fact have more than 3 common words for each sample, where punctuations and stop-words are not considered.

Then we conduct human evaluation for our proposed methods in a similar way to the official evaluation:

1. In addition to *relevance* and *interest*, which are asked in official evaluation, we ask the judges to evaluate two additional metrics: *fluency* and *knowledge relatedness* (to the retrieved fact) of our response.
2. Because we only pick one fact based on the contexts as our model input, we directly provide this fact to judges as the extra information for them to better evaluate *knowledge relatedness* of the response.

The results are shown in Table 4. The submitted system, the best achieved results, and human performance are also included in Table 4 for better comparison. Note that the numbers for two sets of evaluation may not be directly compared but for reference.

In the offline human evaluation, it is found that the proposed models do not achieve better performance and the difference between all models are small. From the official evaluation, our submitted results are also between *disagree* (2) and *neutral* (3) as in our evaluation, but the context-guided attention achieves slightly better numbers than other proposed models shown in the offline setting. Furthermore, the best achieved performance is about 2.94, which is also lower than *neutral* (3), implying the difficulty of this task. It is clear that there is a huge gap between the currently machine-achieved and human-achieved performance, so this task requires further investigation.

Qualitative Analysis

The above results tell that there is no significant difference between our proposed models and baselines. A sample model responses from the human evaluation set is shown in Table 5 for our qualitative analysis. In this example, adding

Retrieved top-1 fact: in the united states , centenarians traditionally receive a letter from the president , congratulating them for their longevity . nbc ' s today show has also named new centenarians on air since 1983 . centenarians born in ireland receive a 2,540 " centenarians ' bounty " and a letter from the president of ireland , even if they are resident abroad . [63] japanese centenarians receive a silver cup and a certificate from the prime minister of japan upon their 100th birthday , honouring them for their longevity and prosperity in their lives . swedish centenarians receive a telegram from the king and queen of sweden . [64] centenarians born in italy receive a letter from the president of italy . in japan , a " national respect for the aged day " has been celebrated every september since 1966 .
Conversation: – til in the united states , people who turn 100 years old receive a letter from the president , congratulating them on their longevity . – same in canada but 90 instead of 100
Ground Truth: is that the canadian exchange rate these days ? **PA Response:** they are the same thing . **PA+CVAE Response:** you can have to be a . **CG Response:** it's not the same thing in the uk .

Table 5: Model response sample.

CVAE generates a more diverse response than the parallel attention result, but may not effectively ground the knowledge in the sentence. Also, our context-guided result seems to focus more on the fact compared to other models. However, the ground truth in the data is very difficult to simulate for the current models, because it may need additional knowledge or common sense. From the current results achieved by our model, we conclude that this task still needs further investigation.

Conclusion

We describe a variational knowledge-grounded conversation system, which attempts at modeling the relations between dialogue contexts and external facts in an end-to-end fashion. It guides a potential research direction about how external information interacts with dialogues and how the machine can capture such interaction for better knowledge-grounded response generation. In the experiments on DSTC7, the results demonstrate the difficulty of this task, because almost all current models fail to generate reasonable responses. Therefore, the knowledge-grounded dialogue modeling requires further study in order to advance the machine's capacity of producing a informative and knowledgable conversation.

References

Bahdanau, D.; Cho, K.; and Bengio, Y. 2014. Neural machine translation by jointly learning to align and translate. *arXiv preprint arXiv:1409.0473*.

Banerjee, S., and Lavie, A. 2005. Meteor: An automatic metric for mt evaluation with improved correlation with human judgments. In *Proceedings of the acl workshop on intrinsic and extrinsic evaluation measures for machine translation and/or summarization*, 65–72.

Bowman, S. R.; Vilnis, L.; Vinyals, O.; Dai, A.; Jozefowicz, R.; and Bengio, S. 2016. Generating sentences from a continuous space. In *Proceedings of The 20th SIGNLL Conference on Computational Natural Language Learning*, 10–21.

Cho, K.; Van Merriënboer, B.; Gulcehre, C.; Bahdanau, D.; Bougares, F.; Schwenk, H.; and Bengio, Y. 2014. Learning phrase representations using rnn encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*.

Doddington, G. 2002. Automatic evaluation of machine translation quality using n-gram co-occurrence statistics. In *Proceedings of the second international conference on Human Language Technology Research*, 138–145. Morgan Kaufmann Publishers Inc.

Gao, J.; Galley, M.; and Li, L. 2018. Neural approaches to conversational ai. 1371–1374.

Ghazvininejad, M.; Brockett, C.; Chang, M.-W.; Dolan, B.; Gao, J.; Yih, W.-t.; and Galley, M. 2017. A knowledge-grounded neural conversation model. *arXiv preprint arXiv:1702.01932*.

Hori, C., and Hori, T. 2017. End-to-end conversation modeling track in dstc6. *arXiv preprint arXiv:1706.07440*.

Kingma, D. P., and Ba, J. 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Kingma, D. P., and Welling, M. 2013. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*.

Li, J.; Galley, M.; Brockett, C.; Gao, J.; and Dolan, B. 2016a. A diversity-promoting objective function for neural conversation models. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 110–119.

Li, J.; Monroe, W.; Ritter, A.; Galley, M.; Gao, J.; and Jurafsky, D. 2016b. Deep reinforcement learning for dialogue generation. *arXiv preprint arXiv:1606.01541*.

Li, X.; Chen, Y.-N.; Li, L.; Gao, J.; and Celikyilmaz, A. 2017. End-to-end task-completion neural dialogue systems.

In *Proceedings of the Eighth International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, volume 1, 733–743.

Papineni, K.; Roukos, S.; Ward, T.; and Zhu, W.-J. 2002. Bleu: a method for automatic evaluation of machine translation. In *Proceedings of the 40th annual meeting on association for computational linguistics*, 311–318. Association for Computational Linguistics.

Peng, B.; Li, X.; Gao, J.; Liu, J.; Chen, Y.-N.; and Wong, K.-F. 2018. Adversarial advantage actor-critic model for task-completion dialogue policy learning. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 6149–6153. IEEE.

Pennington, J.; Socher, R.; and Manning, C. 2014. Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, 1532–1543.

Sharma, S.; El Asri, L.; Schulz, H.; and Zumer, J. 2017. Relevance of unsupervised metrics in task-oriented dialogue for evaluating natural language generation. *CoRR* abs/1706.09799.

Sohn, K.; Lee, H.; and Yan, X. 2015. Learning structured output representation using deep conditional generative models. In *Advances in Neural Information Processing Systems*, 3483–3491.

Sordoni, A.; Galley, M.; Auli, M.; Brockett, C.; Ji, Y.; Mitchell, M.; Nie, J.-Y.; Gao, J.; and Dolan, B. 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*, 196–205.

Vinyals, O., and Le, Q. 2015. A neural conversational model. *ICML Deep Learning Workshop 2015*.

Yoshino, K.; Hori, C.; Perez, J.; D'Haro, L. F.; Polymenakos, L.; Gunasekara, C.; Lasecki, W. S.; Kummerfeld, J.; Galley, M.; Brockett, C.; Gao, J.; Dolan, B.; Gao, S.; Marks, T. K.; Parikh, D.; and Batra, D. 2018. The 7th dialog system technology challenge. *arXiv preprint*.

Zhang, Y.; Galley, M.; Gao, J.; Gan, Z.; Li, X.; Brockett, C.; and Dolan, B. 2018. Generating informative and diverse conversational responses via adversarial information maximization. *arXiv preprint arXiv:1809.05972*.