Contextual Data Augmentation for Task-Oriented Dialog Systems

Dustin Axman, Avik Ray, Shubham Garg, and Jing Huang

Amazon Alexa
{dax, avikray, gargshu}@amazon.com, jinghuang.zhu@gmail.com

Abstract. Collection of annotated dialogs for training task-oriented dialog systems have been one of the key bottlenecks in improving current models. While dialog response generation has been widely studied on the agent side, it is not evident if similar generative models can be used to generate a large variety of, and often unexpected, user inputs that real dialog systems encounter in practice. Existing data augmentation techniques such as paraphrase generation do not take the dialog context into consideration. In this paper, we develop a novel dialog augmentation model that generates a user turn, conditioning on full dialog context. Additionally, with a new prompt design for language model, and output re-ranking, the dialogs generated from our model can be directly used to train downstream dialog systems. On common benchmark datasets MultiWoZ and SGD, we show that our dialog augmentation model generates high quality dialogs and improves dialog success rate by as much as 8% over baseline.

Index Terms: task-oriented dialog, data augmentation, dialog-state tracking

1 Introduction

Users of commercial voice assistants and chatbots (e.g. Alexa, Siri, Google assistant) are able to accomplish various tasks by interacting with them via natural language conversation. Task-oriented dialog models form the core technology behind these applications, which understands users' natural language utterances [1,2], keeps track of the conversation [3,4], performs requested tasks (e.g. API calls) [5,6], and generates appropriate meaningful response to the user [7,8].

Training neural task-oriented dialog models [9,10,11], requires a large amount of annotated data, which is difficult to obtain for model developers. While crowdsourcing and dialog simulation based on agent interplay [12,13] addresses this issue to a certain extent, these are slow and don't provide sufficient coverage of different natural language (NL) user turn surface form variations. Recently, large pre-trained language models (e.g. GPT-2 [14], T5 [15]) have been successfully used to generate fluent agent dialog responses, both with dialog context [16,8,17] or without it [18,19]. However, it is unclear if similar models can capture the large variation of user turn distribution in such task-oriented dialogs. Previous work on data augmentation for spoken language understanding has largely focused on generating paraphrases of user utterance, with a specific goal and set of entities [20,21,22]. However, such utterances again fail to provide sufficient coverage of the large semantic space possible between dialog turns, and may not improve performance of downstream task-oriented dialog systems. As an example, in Table 4, dialog 1, the user says in the first turn $U_1 =$ *"please put me in touch with the local police, i was just robbed"*. A valid variation of this user turn which fits in the dialog context as generated by our model is $U_2 =$ *"I was robbed and I am looking for help"*. Note that, U_1 and U_2 are not semantically equivalent paraphrases (U_2 doesn't explicitly request police).

Therefore, in this paper, we propose a novel dialog augmentation model, using BART [23], which can generate variations of a user turn, when conditioned on past and future dialog turns. Unlike dialog response generation, our model does not have the strict requirement of conveying a desired fixed action response, and can also leverage the future turns in the dialog. We show that using future context is indeed beneficial for the dialog augmentation task. In addition, we propose a new NL context prompt design to delineate between user and system turns, which better aligns with the language model pre-training task, and significantly improves quality of generated utterances and their positive impact on downstream tasks. On benchmark MutiWoZ [24] and SGD [25] datasets, using dialogs generated from our augmentation model can significantly improve dialog success rate and goal accuracy compared to state-of-the-art baseline models.

2 Our Dialog Augmentation Model

In this section, we describe our dialog augmentation model. Let $D = \{U_i, B_i, S_i\}_{i=1}^{n}$ denote a task oriented dialog with n turns, where U_i denote the user request, B_i represent the belief state, and S_i denote the system response, for the i-th turn. We want to train a dialog augmentaion model $\mathcal{M}_d$ which can generate a user turn U'_t for $t \in [n]$ given D, such that $D' = \{U_1, B_1, \ldots, U'_t, B_t, S_t, \ldots, B_n, S_n\}$ is a valid dialog. We refer $\{U_i, B_i, S_i\}_{i=1}^{t-1}$ as the **past context** and $\{U_i, B_i, S_i\}_{i=t+1}^{n}$ as the **future context**.

Base models: Large pre-trained language models have been successfully used as backbone for common NLP tasks [14,23,15], modeling task-oriented dialog systems [9,10], and dialog response generation [18,19]. In this work we also use BART as the base model for our dialog augmentation. BART is pre-trained with *text infilling* task where portions of the input text are masked using special mask tokens, and the model is tasked to re-generate these missing portions at the output. We leverage this task to augment dialog turns. Suppose we want to generate a variation of user turn U_t: we mask the t-th user turn, and construct the input sequence as $X = (U_1, S_1, \ldots, S_{t-1}, [\text{MASK}], S_{t+1}, \ldots, U_n, S_n)$. The output is the user turn $Y = U_t$. The base models are fine-tuned with (X, Y) pairs by masking one turn for every dialog in the training set, and trained with cross-entropy loss. Note that, we do not use the belief states in this base model.

Table 1: Results of the our BART dialog augmentation model and its ablations on MultiWoZ 2.1 dataset. First two columns show extrinsic evaluation metrics, while the remaining columns present intrinsic evaluation metrics.

Models	Inform ↑	Success ↑	BLEU	BERT Score	BLEURT
Soloist [10] (no augmentation)	0.873	0.733	—	—	—
T5 paraphrase augmentation	0.903	0.722	0.592	0.956	0.307
Augmentation with past/future contexts	0.930	0.794	0.188	0.890	-0.547
+ “user:/ system:”	0.910	0.772	0.196	0.890	-0.518
+ re-rank	**0.936**	0.751	0.188	0.890	-0.547
+ “user:/ system:” + re-rank	0.928	**0.816**	0.259	0.908	-0.092
+ “user:/ system:” + re-rank (no future context)	0.917	0.740	0.171	0.887	-0.431
+ “user:/ system:” + re-rank + BS slots	0.932	0.785	0.280	0.909	-0.054
+ “user:/ system:” + re-rank + BS slots (no future context)	0.921	0.765	0.229	0.900	-0.204
+ BS slots	0.916	0.728	0.218	0.896	-0.399

Dis-entangling user/system turns: Such a dialog augmentation model would struggle to learn the differences between user and system turn distribution, and often generate generic and uninformative turns such as *“thank you”*, *“you’re welcome. enjoy!”*. To encourage the model to better learn the nuances of user and system turn distribution, we add special user and system prompts (e.g. *user:/system:*) before every user U_i and system S_i turn respectively. We also add the user token before the mask tokens at the input, when we want to augment one user turn. Note that our model input design involves natural language prompts, as opposed to using special schema tokens (e.g. ⟨user⟩/⟨system⟩) used in previous work [17,22,11]. This results in a better alignment of the input to BART’s pre-training task and generates high quality of dialogs.

Output re-ranking: Our post-generation re-ranking is done by generating 20 top augmentations (found through greedy search on a lattice with 25 beams). We compute the Bleurt score [26] between each generation and the true turn U_t that we are currently augmenting. The highest-score augmentation is returned.

Note that, a key difference of our approach from previous dialog response generation work [16,8,17] is that our model has access to both past/future dialog context; unlike response generation model which only has access to past turns. Additionally, response generation is typically a much more constrained problem due to its usage being generation of system responses in an online context, which conveys a specific intent/API response along with returned entities. In our dialog data augmentation problem, there is no need to restrict the output to be a strict paraphrase of the original user turn. Doing so harms the performance of downstream task-oriented dialog systems, as shown in Section 3.3. Instead, we want the model to generate user turns that are rare and unseen in the training data, and that fit semantically within the provided past/future dialog context.

3 Experiments

In this section, we present intrinsic and extrinsic evaluation results of our proposed dialog augmentation model.

Table 2: Extrinsic evaluation of our BART dialog augmentation model on MultiWoZ 2.1 dataset, in a low resource settings. We observe that augmenting data from our model helps downstream Soloist model achieve higher Inform and Success rates.

Models	20% **data**		50% **data**	
	Inform ↑	**Success ↑**	**Inform ↑**	**Success ↑**
Soloist [10] (no augmentation)	0.549	0.386	0.622	0.494
Augmentation w/ past/ future contexts + “user:/ system:” + re-rank	0.559	0.413	0.789	0.620

3.1 Datasets

We experiment on common benchmark datasets MultiWoZ 2.1, and SGD for multi-domain task-oriented dialogs.

The **MultiWoZ 2.1 dataset** [24], is a consolidated and cleaned version of its earlier version [27]. It is widely used as a benchmark for evaluation of task-oriented dialog models. This dataset contains task-oriented dialogs from multiple domains (e.g. Restaurant, Hotel, Attraction, Taxi, Train, Hospital, Bus, and Police). The dataset contains 8,438 training, 1,000 dev, and 1,000 test dialogs.

The **Schema Guided Dialog (SGD) dataset** [25], is a large task-oriented dialog benchmark dataset containing 22,825 dialogs covering 16 different domains (e.g. Flights, Hotels, Events, Services, Alarm etc.), and split into 16,142 train, 2,482 dev, and 4,201 test dialogs. The dialogs are represented with a flexible and unified schema, which facilitates easier integration of new domain services via zero/few shot dialog state tracking.

3.2 Metrics and setup

Our evaluation is split into two groups, intrinsic and extrinsic. For **intrinsic evaluation**, in MultiWoZ, we generate an augmented turn for every user turn of the test dialogs. We compute intrinsic evaluation metrics between augmented turn and ground truth user turn to gauge the augmentation quality. For the **intrinsic metrics** we compute **BLEU** [28], **BertScore** [29], **Bleurt** [26].

To evaluate use of the generated dialogs for helping downstream task-oriented dialog systems, we perform **extrinsic evaluation** as follows. We augment 1 randomly selected user turn in each of a fixed percentage p of the training dialogs

Table 3: Results of our BART dialog augmentation model and its ablations on SGD dataset using the SG–DST model [25].

Models	Active Int Acc ↑	Req Slot F1 ↑	Avg GA ↑	Joint GA ↑
SG–DST [25] (no augmentation)	0.870	**0.968**	0.559	0.241
Augmentation with past/future contexts	**0.902**	0.965	0.569	0.249
+ "user:/ system:"	0.901	0.965	0.560	0.250
+ re-rank	0.898	0.965	0.572	**0.257**
+ "user:/ system:" + re-rank	0.899	0.966	**0.573**	0.250
+ "user:/ system:" + re-rank + BS slots	0.901	0.966	0.570	0.244

(we use $p = 5\%$ in MultiWoZ, and $p = 25\%$ for SGD[1]). These augmented dialogs (with the augmented turn replacing the original) are added into this training set. For MultiWoZ, we choose Soloist [10] as the baseline, train it on this augmented training set and evaluate on the MultiWoZ test set using the most commonly used **Inform rate** and the **Success rate** metrics [10]. In SGD, we select the dialog-state tracking (DST) baseline model Schema guided DST (SG–DST) introduced in [25]. We evaluate the performance of SG–DST on the test split using the metrics **Active intent accuracy**, **Requested slots F1**, **Average goal accuracy**, and **Joint goal accuracy** as defined in [25]. For each dataset we use the same extrinsic metrics as used in the original papers [10,25] for easier comparison with corresponding baselines. On MultiWoZ, while models such as LAVA [30] can achieve a better performance than Soloist, we consider it less suited for practical applications due to its complex multi-step RL based training.

As a baseline **paraphrase augmentation** model, we fine-tune a T5 model [15] using the paraphrase generation task on a set of paraphrases from MultiWoZ which were collected using the same method as in [31]. Due to unavailability of such paraphrases for SGD dataset, we do not study this baseline for SGD.

Training details: We fine tune base BART model for 4 epochs on the task defined above, using batch size of 8, and learning rate of 2×10^{-5}. The encoder and embeddings are frozen during training. We use 4 eval beams. For extrinsic evaluation, we train the Soloist model with default hyper-parameters using the original code [32]. The SGD–DST model was trained for 70 epochs using the default hyper-parameters of the original implementation [33]. We train both baseline models on the original training sets using the same hyper-parameters. All models were trained using a machine with single V100 GPU. Training BART dialog augmentation model requires about 3 hours, training Soloist model takes approximately 7 hours, and training SG–DST model takes about 26 hours.

[1] Since SGD is a larger dataset, we observed that it takes more augmentations to cause a significant improvement.

3.3 Results and discussion

In Table 1 we present the results on MultiWoZ dataset. We observe that T5 paraphrase based augmentation achieves high intrinsic metrics since the paraphrases of the user turn are very similar to the original ground-truth user utterance. However, augmenting with paraphrases does not necessarily improve extrinsic metrics of downstream Soloist model. Our models achieve lower intrinsic metric compared to paraphrasing as expected, since they produce more variations of user turn, different from the ground-truth. However, the greater semantic coverage by our model leads to higher extrinsic metrics for downstream Soloist model. Our best augmentation model is the model with "system:"/"user:" prompting as well as a post augmentation re-ranking which significantly outperforms baseline Soloist model on all extrinsic metrics achieving an 8% improvement in Success rate and 5% improvement in Inform rate (row 6 in Table 1). This validates that augmentations generated from our model indeed help in improving downstream task-oriented dialog models. We also conclude that, traditional intrinsic metrics do not correlate well with extrinsic metrics for the dialog augmentation task. Since extrinsic metrics are more important for success of task-oriented dialog models, for the remaining section we mainly focus on these.

Table 3 presents the results on SGD with SG–DST (no augmentation) as the baseline. Although SGD is a much larger dataset compared to MultiWoZ, our BART augmentation model can still improve performance over baseline across several metrics (3.2% intent accuracy, 1.4% average goal accuracy, and 1.6% joint goal accuracy). We do not observe improvement in slot F1 score since we didn't perform any re-annotation of slots in the generated turn, which can result in some missing annotations. We also observe that "system:"/ "user:" prompting is less beneficial than re-ranking in SGD. We hypothesize that in SGD it is easier for the model to differentiate user/ system distribution and thus making prompting less effective compared to MultiWoZ dataset.

Results in low-resource settings: Our dialog augmentation models can benefit the low resource dialog applications by boosting the performance of the dialog systems. We test this on the smaller MultiWoZ dataset, which can better emulate a low resource setting. We further sample 20%, and 50% training dialogs from MultiWoZ training set, and augment additional 5% of the sampled dialogs using our best performing model, BART + "user: /system:" prompts + re-rank. We then train the Soloist model on this augmented training data and evaluate the extrinsic metrics on the full test set. We compare the performance of this augmented model with baseline model trained only on the sampled training data (**no augmentation**). From the results in Table 2 we observe that adding the augmented dialogs greatly improves both the Inform and Success rate of downstream Soloist model in low resource settings: by 16.7% Inform rate and 12.6% Success rate for the 50% training data experiment.

3.4 Ablation studies

We conduct several ablations designed to explore the impact of individual model components on our extrinsic and intrinsic metrics: the post augmentation re-

Table 4: Augmented dialog examples in MultiWoZ 2.1. The original turn with '*' around them (referred as GT) was augmented, and the augmented turn from our model (referred as AUG) is italicized.

	Dialog 1
user (GT)	*please put me in touch with the local police, i was just robbed.*
user (AUG)	*I was robbed and I am looking for help*
system	i have information for the parkside police station, is this close to your location?
user	i don't know, could you just give me the postcode and phone?
system	the phone number is 01223358966 and the post code is cb11jg.
user	thank you for your help
system	you're welcome. the police will help you immediately. goodbye.
	Dialog 2
user	please see if you can find me a 3-star hotel on the east
system	i'm sorry we do not have any east hotels that meet your criteria. would you like to try a different part of town?
user	yes, please look for any 3-star hotel -s in town that include free parking . east is still preferred.
system	none fits your criteria. can we change the area?
user	try something in the center of town.
system	the gonville hotel is an expensive, 3-star hotel in the centre of town. would you like me to book you a room?
user	yes please . it's for 2 people and 2 nights starting on thursday.
system	booking was successful. reference number is : 2v7ewg9e. is there anything else i can help you with?
user (GT)	*how about a place to dine in the expensive price range serving indian food.*
user (AUG)	*I'm also looking for a place to dine that serves indian food*
system	i have many. what area would you prefer dining in?
user	i would like somewhere in the same area as the hotel.
system	i have 5 all located in the centre
user	then please pick 1 and give me their phone number band postcode.
system	saffron brasserie is located at hills road city centre, post code cb21la. their phone number is 01223354679. can i help with anything else today?
user	that is everything . thank you for your help.
system	you are welcome.

ranking ("re-rank"), user/system prompts with ("user:/ system:"), removal of future context from the model input ("No future context"), as well as comparison with a variant that inserts belief state slots into the prompt ("BS slots"). The results are shown in Tables 1 and 3. We note that while increases in extrinsic metrics are objectively positive, intrinsic metrics are more open to interpretation. For example, having a high lexical/semantic similarity between an augmentation and original user turn does not always indicate the most useful augmentations from a downstream impact perspective.

Importance of future context: To study if conditioning augmentations on future context indeed helps, we conduct an ablation on MultiWoZ where we remove the future context, and augment based on only past and current turns (no future context), which is the usual settings in response generation [16,18]. We observe that removing future context indeed harms downstream performance e.g. degrading Success rate by 7.6% in our best model configuration.

Effect of re-ranking: We can see that re-ranking gives some of our best results in both SGD and MultiWoZ. In SGD, models with re-ranking achieve best goal accuracy, while in MultiWoZ extrinsic results are best when re-ranking is combined with user/system prompting. In all models except the base augmentation model, re-ranking improves scores on similarity metrics such as BERTScore, BLEU, and BLEURT. This is expected, because re-ranking based on Bleurt scores encourages selection of augmentations with greater semantic similarity to original user turn.

Using belief state slots: Using information from belief state B_t, has been shown to be effective in generating relevant system dialog responses [17,18]. Our final ablation experiment (**+ BS slots**), studies the impact of adding entities/slots from belief state B_t to the base model input X, just before the mask token. We convert the slots to their natural language template[2] similar to [21,17]. This encourages the model generate an augmented turn U'_t containing the same set of entities. This is also validated by a consistent increase in intrinsic semantic similarity metrics when BS slots are added to the input. Although proven to be effective in response generation, for both MultiWoZ and SGD, it does not offer improvement over our best model in extrinsic metrics. We hypothesize that this is because these slots over-constrain the augmentation to generate close paraphrases of the original user turn.

3.5 Example augmented dialogs

In the Table 4 we present some example dialogs generated by our augmentation model in MultiWoZ dataset. The turn with '*' around them was augmented with our model (also referred as ground-truth GT). The augmented dialog generated by our model is italicized (also referred as AUG). In **Dialog 1** we can see that our augmentation of the first turn is very similar to the original even though it was generated without seeing the original turn as input. In **Dialog 2**, augmentation preserves the user intent and entity "indian", but drops the term "expensive". In future work, we want to research more into effective ways of leveraging this current turn entity information. In both the examples, our model generates a new user turn variation which fits in the dialog context, although not being a strict paraphrase of the original user turn.

4 Conclusion

We develop a novel dialog augmentation model, which can generate new user turn variations given both past and future dialog context, and current dialog state. We carefully design natural language prompts for pre-trained language models. Together with a re-ranking model, our data augmentation approach generates high quality dialogs that can augment existing dialog datasets. We further show that the augmentation data from our model greatly improve dialog completion and success rates of SOTA task-oriented dialog systems. Using ablation study, we also highlight an important tradeoff: generating accurate paraphrases of user turns does not necessarily improve downstream task-oriented dialog systems. Instead, generating more variations of user inputs that fits the given context, would result in better performance.

[2] For example, to augment user turn $U_t =$ *"i need train reservations from norwich to cambridge"* containing entities {"norwich", "cambridge"}, we include natural language template phrase *"train departing norwich, train destination cambridge"* before the mask token in the input X.

References

1. D. Hakkani-Tür, G. Tür, A. Celikyilmaz *et al.*, "Multi-domain joint semantic frame parsing using bi-directional RNN-LSTM," in *Interspeech*, 2016.
2. C.-W. Goo, G. Gao, Y.-K. Hsu *et al.*, "Slot-gated modeling for joint slot filling and intent prediction," in *NAACL-HLT*, 2018.
3. N. Mrksic, D. Ó. Séaghdha, T. Wen *et al.*, "Neural belief tracker: Data-driven dialogue state tracking," in *ACL 2017*, 2017.
4. L. Chen, B. Lv, C. Wang *et al.*, "Schema-guided multi-domain dialogue state tracking with graph attention neural networks," in *AAAI*, 2020.
5. T. Wen, Y. Miao, P. Blunsom *et al.*, "Latent intention dialogue models," in *ICML*, 2017.
6. B. Peng, X. Li, J. Gao, J. Liu, and K. Wong, "Deep dyna-q: Integrating planning for task-completion dialogue policy learning," in *ACL*, 2018.
7. T. Wen, M. Gasic, N. Mrksic, P. Su, D. Vandyke, and S. J. Young, "Semantically conditioned lstm-based natural language generation for spoken dialogue systems," in *EMNLP*, 2015.
8. Y. Zhang, S. Sun, M. Galley *et al.*, "DIALOGPT : Large-scale generative pre-training for conversational response generation," in *ACL*, 2020.
9. E. Hosseini-Asl, B. McCann, C. Wu *et al.*, "A simple language model for task-oriented dialogue," in *NeurIPS*, 2020.
10. B. Peng, C. Li, J. Li *et al.*, "SOLOIST: building task bots at scale with transfer learning and machine teaching," *TACL*, vol. 9, 2021.
11. Y. Yang, Y. Li, and X. Quan, "UBAR: towards fully end-to-end task-oriented dialog system with GPT-2," in *AAAI*, 2021.
12. P. Shah, D. Hakkani-Tür, G. Tür *et al.*, "Building a conversational agent overnight with dialogue self-play," *arXiv arXiv:1801.04871*, 2018.
13. C.-W. Lin, V. Auvray, D. Elkind *et al.*, "Dialog simulation with realistic variations for training goal-oriented conversational systems," *arXiv preprint arXiv:2011.08243*, 2020.
14. A. Radford, J. Wu, R. Child *et al.*, "Language models are unsupervised multitask learners," *OpenAI blog*, 2019.
15. C. Raffel, N. Shazeer, A. Roberts *et al.*, "Exploring the limits of transfer learning with a unified text-to-text transformer," *JMLR*, vol. 21, 2020.
16. J. Gu, Q. Wu, C. Wu, W. Shi, and Z. Yu, "PRAL: A tailored pre-training model for task-oriented dialog generation," in *ACL/IJCNLP*, 2021.
17. M. Kale and A. Rastogi, "Template guided text generation for task-oriented dialogue," in *EMNLP*, 2020.
18. ——, "Text-to-text pre-training for data-to-text tasks," in *INLG*, 2020.
19. X. Xu, G. Wang, Y. Kim, and S. Lee, "Augnlg: Few-shot natural language generation using self-trained data augmentation," in *ACL/IJCNLP*, 2021.
20. Y. Hou, Y. Liu, W. Che, and T. Liu, "Sequence-to-sequence data augmentation for dialogue language understanding," in *Proc. of COLING*, 2018.
21. Z. Zhao, S. Zhu, and K. Yu, "Data augmentation with atomic templates for spoken language understanding," in *Proc. of EMNLP-IJCNLP*, 2019.
22. H. Lin, L. Xiang, Y. Zhou, J. Zhang, and C. Zong, "Augmenting slot values and contexts for spoken language understanding with pretrained models," in *Interspeech*, 2021.
23. M. Lewis, Y. Liu, N. Goyal *et al.*, "BART: denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension," in *ACL*, 2020.

24. M. Eric, R. Goel *et al.*, “Multiwoz 2.1: A consolidated multi-domain dialogue dataset with state corrections and state tracking baselines,” in *LREC*, 2020.
25. A. Rastogi, X. Zang, S. Sunkara *et al.*, “Towards scalable multi-domain conversational agents: The schema-guided dialogue dataset,” in *AAAI*, 2020.
26. T. Sellam, D. Das, and A. P. Parikh, “Bleurt: Learning robust metrics for text generation,” *arXiv preprint arXiv:2004.04696*, 2020.
27. P. Budzianowski, T. Wen *et al.*, “Multiwoz - A large-scale multi-domain wizard-of-oz dataset for task-oriented dialogue modelling,” in *EMNLP*, 2018.
28. K. Papineni, S. Roukos, T. Ward, and W.-J. Zhu, “Bleu: a method for automatic evaluation of machine translation,” in *ACL*, 2022.
29. T. Zhang, V. Kishore, F. Wu, K. Q. Weinberger, and Y. Artzi, “Bertscore: Evaluating text generation with BERT,” in *ICLR*, 2020.
30. N. Lubis, C. Geishauser, M. Heck, H. Lin, M. Moresi, C. van Niekerk, and M. Gasic, “LAVA: latent action spaces via variational auto-encoding for dialogue policy optimization,” in *COLING*, 2020, pp. 465–479.
31. S. Gao, Y. Zhang, Z. Ou, and Z. Yu, “Paraphrase augmented task-oriented dialog generation,” in *ACL*, 2020.
32. Baolin Peng and Chunyuan Li and Jinchao Li and Shahin Shayandeh and Lars Liden and Jianfeng Gao, “Soloist,” https://github.com/pengbaolin/soloisthttps://github.com/pengbaolin/soloist, 2021.
33. Rastogi et al., “Schema guided dst,” https://github.com/google-research/google-research/tree/master/schema_guided_dst, 2020.