

Unsupervised Enrichment of Persona-grounded Dialog with Background Stories

Bodhisattwa Prasad Majumder[♣] Taylor Berg-Kirkpatrick[♣]

Julian McAuley[♣] Harsh Jhamtani[◇]

[♣]Department of Computer Science and Engineering, UC San Diego

{bmajumde, tberg, jmcauley}@eng.ucsd.edu

[◇]School of Computer Science, Carnegie Mellon University

jharsh@cs.cmu.edu

Abstract

Humans often refer to personal narratives, life experiences, and events to make a conversation more engaging and rich. While persona-grounded dialog models are able to generate responses that follow a given persona, they often miss out on stating detailed experiences or events related to a persona, often leaving conversations shallow and dull. In this work, we equip dialog models with ‘background stories’ related to a persona by leveraging fictional narratives from existing story datasets (e.g. ROC-Stories). Since current dialog datasets do not contain such narratives as responses, we perform an unsupervised adaptation of a retrieved story for generating a dialog response using a gradient-based rewriting technique. Our proposed method encourages the generated response to be *fluent* (i.e., highly likely) with the dialog history, *minimally different* from the retrieved story to preserve event ordering and *consistent* with the original persona. We demonstrate that our method can generate responses that are more diverse, and are rated more engaging and human-like by human evaluators, compared to outputs from existing dialog models.

1 Introduction

Humans often rely on specific incidents and experiences while conversing in social contexts (Dunbar et al., 1997). Responses from existing chitchat dialog agents often lack such specific details. To mitigate this, some prior work has looked into assigning personas to dialog agents (Zhang et al., 2018; Majumder et al., 2020). However, persona descriptions are often shallow and limited in scope, and while they lead to improvements response specificity, they still lack the level of detail with which humans share experiences.

In this work, we propose methods to enrich dialog personas with relevant background events us-

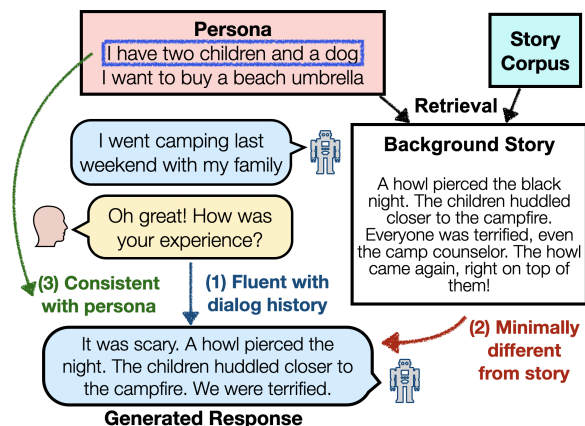


Figure 1: We enrich agent personas with ‘background stories’ from an existing corpus. We propose a gradient-based technique which encourages the generated response to be fluent with the dialog history, minimally different from the retrieved story, and consistent with the persona. The proposed approach leads to more specific and interesting responses.

ing fictional narratives from existing story datasets such as ROCStories (Mostafazadeh et al., 2016). For example, for a persona attribute ‘I have two children and a dog,’ we are able to identify a relevant narrative from a story corpus (Figure 1). However, such stories may not directly fit fluently in the dialog context. Thus, retrieved stories should be adapted to construct a response that is fluent and relevant to the context. Since existing datasets (such as PersonaChat (Zhang et al., 2018)) do not contain responses with such background stories, such adaptation has to be done in an unsupervised fashion with decoders trained to generate responses conditioned only on a dialog history and persona.

To adapt a retrieved narrative incident as a relevant background story, we use a decoding procedure which encourages the generated response to (1) be fluent with the dialog history, (2) be consistent with the original persona, and (3) be minimally different from the retrieved story. While fluency with dialog context is encouraged directly by the likelihood as per the underlying language model

the remaining two constraints are incorporated via iterative updates to the decoder output distributions at inference time. Our inference-time decoding method is different from the only recent effort by [Su et al. \(2020\)](#) that leverages non-dialog data (forum comments, book snippets) as distant labels to train dialog systems with supervision. Our contributions can be summarized as follows:

- We propose a novel approach to enrich dialog agent personas with relevant backstories, relying only on existing story datasets.
- We propose to use an unsupervised back-propagation based decoding procedure¹ to adapt the relevant stories such that the resulting response is fluent with the dialog history and consistent with the dialog agent persona. Our method works with a model trained just with dialog data i.e. without access to story corpus at training time.
- Our experiments demonstrate that the proposed approach results in much more engaging and specific dialog outputs in a persona-grounded dialog setup. This fills a gap in existing dialog models which often lack the capability to generate responses about specific events and experiences relevant to persona attributes.

2 Unsupervised Persona Enrichment with Background Stories

Given dialog history h and persona C consisting of several (typically 3-5, example shown in Figure 1) attributes, our goal is to construct a dialog response x . Our underlying model is based on the discrete persona attribute choice model from [Majumder et al. \(2020\)](#). To generate a dialog utterance x , we first sample a persona attribute $c \sim p(c|h)$ conditioned on the dialog history h . x is then generated conditioned on the dialog history and the chosen persona attribute. The underlying dialog model’s decoder is initialized with a pretrained GPT-2 model, and is fine-tuned on the PersonaChat dataset ([Zhang et al., 2018](#)). However, in our current setup, we also have to identify relevant background stories and use them to construct fluent responses at decoding time. Therefore, we propose a different decoding procedure.

To generate a response, we first sample a persona attribute $c \sim p(c|h)$. Next we retrieve stories cor-

responding to the persona attribute c (Section 2.1). However, the underlying dialog model is trained to generate responses conditioned only on the dialog history and persona. To incorporate the retrieved story in the response, we perform gradient-based inference (Section 2.2), that only assumes a left-to-right language model trained on dialog context and responses, and the story is handled at decoding time in an unsupervised fashion. We refer to the proposed method as **PABST** (Unsupervised **P**erson**A** enrichment with **B**ackground **S**Tories).

2.1 Retrieving Relevant Stories

For a persona attribute c , we aim to identify relevant stories from a story corpus. Toward this goal, we rank the stories using the F1 component of BERT-score ([Zhang et al., 2020](#)) based retrieval using the persona attribute c as the query and the highest scoring story is chosen. Note that many of the stories are written in the third person. For use as background stories, we must first transform them to first-person. Following prior work ([Brahman and Chaturvedi, 2020](#)), we identify the protagonist of such stories as the most frequently occurring character. Thereafter, we use co-reference resolution ([Lee et al., 2017](#)) to identify all words or phrases that refer to the protagonist. Finally, all words or phrases so identified are replaced with suitable first person pronouns (e.g. ‘his books’ to ‘my books’).

2.2 Gradient-based Inference

Our underlying dialog model is not trained to condition on a retrieved story, and cannot be directly used to construct a desirable response using s . To tackle this, we consider a decoding strategy which, in addition to fluency with history h , encourages response x to follow two soft constraints: (1) be minimally different from story s , and (2) be consistent with persona c .

First, we generate an initial response based only on the dialog history. Then we perform an iterative procedure which alternates between performing a forward pass on the language model to encourage fluency, and a backward pass which updates the response via back-propagation to respect the two soft constraints. However, x is discrete, and cannot be directly updated using gradients from back-propagation. Instead, we maintain and update a soft representation o of x , where o_i corresponds to the last hidden state representation for the i^{th} token position, i.e., $p(x_i) \sim \text{softmax}(W o_i / \tau)$, where τ is the temperature parameter, W is the embedding

¹Code can be found at <https://github.com/majumderb/pabst>

matrix, and $W_{o_i} \in \mathcal{R}^V$ (V is the vocabulary size). Our approach is inspired by recent works that use gradient-based decoding for text generation with soft constraints (Dathathri et al., 2020; Qin et al., 2020). Next we describe the backward and forward passes of the iterative procedure.

Backward Pass with Soft Constraints We define the following soft constraints on response x :

(1) **Divergence from story:** We want to encourage x to be *minimally different* from the story s . Following prior work (Qin et al., 2020), we compute a cross entropy loss (denoted by `cross-entr` henceforth) with story $s = \{s_1, \dots, s_T\}$ tokens as labels and W_{o_1}, \dots, W_{o_T} as the logits.

(2) **Consistency to persona:** We want x to be *consistent with persona attribute c* . Consider a classifier $q_\phi(o, c)$ which predicts the probability of x (or rather the soft representation o of x) entailing c . The classifier $q_\phi(o, c)$ is a bag-of-words classification head on decoder hidden states o , fine-tuned on the Dialogue-NLI dataset (Welleck et al., 2019) to predict whether pairs of persona attributes and responses are entailed or not. The objective to maximize can be written as:

$$\mathcal{L}(c, s; o) = \lambda_c \log q_\phi(o, c) - \lambda_d \text{cross-entr}(s, W_o)$$

where λ_c and λ_d are hyper-parameters. We update o through back-propagation by computing the gradient $\nabla_o \mathcal{L}(c, s; o)$, while keeping the model parameters constant. Let the resulting o after the gradient-based updates be denoted by o^b .

Forward Pass to Encourage Fluency Next we perform a forward pass of the underlying dialog model, with the goal of regularizing the hidden states towards the unmodified language model values. On computing the forward pass at the j^{th} token, we mix the final hidden states o_j^f from the forward pass with o_j^b computed in the backward pass, via weighted addition to get the resulting $o_j = \gamma \times o_j^f + (1 - \gamma) \times o_j^b$, where $\gamma \in (0, 1)$ is a hyperparameter. The resulting o_j is used for computing the logits at the next time step $j + 1$.

We initialize the output response by performing greedy decoding from the underlying dialog model, conditioned on the dialog history and persona attribute. Then we iteratively update o by alternate backward and forward passes. We sample the final response $x \sim \text{softmax}(W_o/\tau)$. In practice, we found that 5 iterations are sufficient to generate good quality outputs.

Method	Training	Decoding	D-1	D-2	ENTR
W/o Story Data					
TRANSFERO	PERSONA	Nucleus	0.05	0.11	1.21
DISCCHOICE	PERSONA	Nucleus	0.15	0.25	1.25
DISCCHOICE	CS-KB	Nucleus	0.87	1.07	2.04
With Story Data					
DISCCHOICE	PSEUDO	Nucleus	0.91	2.45	2.89
DISCCHOICE	MULTITASK	Nucleus	0.99	2.54	2.71
DISCCHOICE	PERSONA	RETRIEVAL	2.56	9.67	3.86
PABST (Ours)	PERSONA	Grad. Inf.	1.56	3.57	3.21

Table 1: Diversity metrics on the PersonaChat test set. D-1/2 is the % of distinct uni- and bi-grams. ENTR is the geometric mean of n-gram entropy. Grad. Inf. is the unsupervised gradient-based decoding as opposed to Nucleus sampling (Holtzman et al., 2020).

3 Experiments

We evaluate methods in terms of their capability to generate diverse, fluent and engaging responses. Hyperparameters are noted in Appendix §A.

Datasets We experiment with the PersonaChat dialog dataset (Zhang et al., 2018) consisting of 131,438 utterances for training, 15,602 for validation, and 15,024 for testing. For stories, we use the training split of the ROCStories dataset (Mostafazadeh et al., 2016), that consists of 78,529 stories, each typically of 4 to 5 sentences.

Baselines We consider two broad groups of models as baselines: (1) **Without access to story corpus:** We use finetuned GPT2 (TRANSFERO) on PersonaChat, and the discrete persona attribute choice model (DISCCHOICE) from Majumder et al. (2020). We also consider a version of DISCCHOICE which enriches personas with inferences from a commonsense knowledge base (CS-KB). (2) **Baselines using story corpus:** To allow DISCCHOICE models to generate story-like responses, we adapt an alternative training regime (PSEUDO) from (Su et al., 2020), where we randomly replace some of the target dialog responses with retrieved stories—treating them as pseudo labels. Finally, we also consider a MULTITASK training setup from (Su et al., 2020), wherein the decoder is trained on PersonaChat as well as with a language modeling objective on ROCStories. We additionally consider a RETRIEVAL baseline that uses the retrieved story verbatim as the dialog response.

3.1 Automatic Evaluation

We hypothesize that that the proposed approach to leverage external non-dialog data can increase the diversity of the generated responses. Following

PABST vs. Aspect	TRANSFERO		DISCCHOICE		RETRIEVAL		PSEUDO		MULTITASK		w/o DNLI		Gold	
	win	loss	win	loss	win	loss	win	loss	win	loss	win	loss	win	loss
Sensible	83.8	5.6	71.4	10.1	72.7	15.7	94.3	3.1	92.8	1.5	81.7	6.8	29.1	33.6
Engaging	96.7	1.8	82.6	6.7	87.2	5.3	94.7	2.1	94.1	2.6	70.9	14.8	70.2	15.6

Table 2: Pairwise comparison (% win/loss cases, ties not reported) between responses from PABST and from other baselines as well as gold. All differences are significant with $p < 0.05$. Cohen’s Kappa (Cohen, 1960) for sensibility and engagement were 0.79 and 0.82 respectively. Complete details of the human evaluations are in Appendix §B.

Dialog History

Agent: How was your last trip?

User: It was great. I went to Utah and saw the Canyonlands.

Retrieved Story:

I went on a trip. I went to Florida. When we got there, there was no air conditioning. I was irritated. It was too hot.

DISCCHOICE: That’s great. You must be adventurous.

RETRIEVAL: I went on a trip. I went to Florida. When we got there, there was no air conditioning. I was irritated. It was too hot.

PSEUDO: I have a house in Florida.

MULTITASK: I like Florida. I went to the beaches.

PABST ($\lambda_d = 5$): I went on a trip. I went to Florida. I have a house there. We do not have air condition. It was hot.

PABST ($\lambda_d = 1$): That’s great. I frequently go to Florida. I have a house there but it gets hot in summer.

Dialog History

Agent: I work as stunt double. What about you?

User: I’m an engineer. I just got back from work. How’s yours?

Retrieved Story:

I was a working actor. For my new role I was a stunt double. I started working on new stunts. When I jumped off a building yesterday I got hurt. I had to be rushed to the hospital.

DISCCHOICE: That’s great. I am looking for a job.

RETRIEVAL: I was a working actor. For my new role I was a stunt double. I started working on new stunts. When I jumped off a building yesterday I got hurt. I had to be rushed to the hospital.

PSEUDO: I got a new job. I am an actor.

MULTITASK: I started a new role. I got hurt.

PABST ($\lambda_d = 5$): I was working as a stunt double. I jumped off a building. I got hurt. I had to be rushed to the hospital.

PABST ($\lambda_d = 1$): Yesterday while I was working, I jumped off a building and I got hurt. I had to be taken to the hospital.

Table 3: Generations from different models. More examples are in Appendix §C.

prior work (Li et al., 2016), we report the percentage of distinct uni-grams and bi-grams (**D-1** and **D-2** respectively). Note that these values do not capture the actual frequency distribution of different word types. Therefore, we also report the geometric mean of entropy values of empirical frequency distributions of n-grams of words ($n \in \{1, 2, 3\}$) (Jhamtani et al., 2018), denoted by **ENTR**.

We observe that methods that use story data show much higher diversity compared to methods that do not (Table 1). Among methods using story data, gradient-based decoding (PABST) performs better than DISCCHOICE trained with PSEUDO or MULTITASK. Note that just using RETRIEVAL outputs as-is leads to even more diverse outputs than PABST. However, they are much less sensible with the context, as shown in human evaluations.

3.2 Human Evaluation

Since we do not have ground truth story-like responses in the dialog dataset, we perform human evaluation with 150 test examples to investigate if PABST generates responses that are 1) **sensible** with the dialog history and 2) **engaging**. We hired two Anglophone (Lifetime HIT acceptance % > 85) annotators for every test sample. The order of the systems present in the interface is randomized.

A snapshot of the human evaluation interface is provided in Appendix §C. All differences in values from human evaluations are significant with $p < 0.05$ from bootstrap tests on 1000 subsets of size 50. Cohen’s Kappa (Cohen, 1960) to measure inter-annotator agreement for sensibility and engagement were 0.79 and 0.82 respectively.

From the results (shown in Table 3), we note that in comparison to responses from baselines, responses from PABST are more engaging and more sensible with respect to the dialog history. We further make following observations. Firstly, using the gradient-based decoding approach with retrieved stories (PABST) works significantly better than using distant supervision with stories data (PSEUDO and MULTITASK). Secondly, background stories provide sufficient detail for an engaging conversation compared to DISCCHOICE which expands persona attributes using commonsense knowledge (Majumder et al., 2020). Finally, we also observe that PABST performs worse when we do not use the consistency constraint (w/o DNLI).

Choice of λ_d We also experiment with different values of the weight for the divergence term (λ_d) in \mathcal{L} : High ($\lambda_d = 5$), Moderate ($\lambda_d = 1$),

and Low ($\lambda_d = 0.05$). We consider 100 samples for this experiment. We attribute a high λ_d to responses strictly copying the story. We find that PABST (moderate λ_d) wins 81.2% and 69.1% cases against PABST (high λ_d) on ‘sensible’ and ‘engaging’ response criteria respectively. Similarly, PABST (moderate λ_d) wins 93.2% and 84.7% cases against PABST (low λ_d) in terms of sensibility and engagement respectively.

Qualitative Analysis Table 3 shows responses generated by different baselines. We observe that PABST is able to follow the retrieved story (same as output from RETRIEVAL) while modifying the response to be conversation-like and sensible with dialog history. Responses from other baselines remain verbose or incoherent. Mirroring the human evaluation, we observe that choosing a higher λ_d makes the model to almost repeat the retrieved story but a lower value smooths the output to make it more sensible with the ongoing dialog.

4 Related Work

A desired impact of the proposed approach is increase in diversity of the generated responses. To tackle the issue of diversity in dialog model outputs, prior work has focused on decoding strategies such as diversity-promoting sampling (Holtzman et al., 2020); training strategies such as discouraging undesirable responses via unlikelihood training (Li et al., 2020); model changes such as using stochastic variables (Serban et al., 2017); and using external data such as forum data (Su et al., 2020) or external knowledge bases (Majumder et al., 2020). In contrast to these, our proposed method generates responses with background stories using a gradient-based decoding approach.

One of the steps in our proposed approach is to retrieve relevant stories from an external corpus. Prior work has explored using retrieval of similar dialog instances as an initial step in improving response diversity and other human-like desiderata in dialog (Roller et al., 2020; Weston et al., 2018). Distant supervision by using retrieved text snippets as pseudo responses has been explored in prior work (Su et al., 2020; Roller et al., 2020). We use an external data source to improve dialog responses, a theme shared with some efforts in other tasks such as machine translation (Khandelwal et al.). The use of narrative text in dialog has been explored in prior work, mostly as a ‘script’ or template for conversation (Xu et al., 2020; Zhu et al., 2020).

We adapted a BERT-based retrieval method (Zhang et al., 2020) in our case to retrieve relevant story given dialog context and use retrieved story in the decoding phase.

Gradient-based for text generation with soft constraints has been explored in prior work (Dathathri et al., 2020; Qin et al., 2020). Song et al. (2020) focused on generating response which are consistent to given persona. Differently, we use a gradient-based decoding to generate a dialog response while honoring constraints such as consistency to persona and similarity to retrieved story.

5 Conclusion

We propose a method to enrich persona-grounded dialog with background stories at the inference time only using an existing corpus of non-conversational narratives—opening up new ways to generate enriched and engaging responses. One of the limitations of PABST is the assumption of the background story at every turn. As future work, we can include a decision step to decide if we need to incorporate a background story or not, given the dialog history. We can further explore ways to use retrieved stories over multiple turns instead of a single turn.

Acknowledgements

We thank anonymous reviewers for providing valuable feedback. BPM is partly supported by a Qualcomm Innovation Fellowship and NSF Award #1750063. Findings and observations are of the authors only and do not necessarily reflect the views of the funding agencies.

Impact Statement

In this work, we discuss ways to make a dialog system to generate more engaging responses. Since we use a finetuned version of a pretrained generative model, we inherit the general risk of generating biased or toxic language, which should be carefully filtered. Furthermore, the generations may incorporate biases that are already present in the dialog dataset and story dataset due to crowd-sourced data collection. Hence, we cautiously advise any developer who wishes to use a different story dataset for the background stories to be aware of the biases present in the dataset. Finally, we also note that experiments in this paper are limited only to English language.

References

- Faeze Brahma and Snigdha Chaturvedi. 2020. [Modeling protagonist emotions for emotion-aware storytelling](#). In *EMNLP*, pages 5277–5294.
- Jacob Cohen. 1960. A coefficient of agreement for nominal scales. *Educational and psychological measurement*, 20(1):37–46.
- Sumanth Dathathri, Andrea Madotto, Janice Lan, Jane Hung, Eric Frank, Piero Molino, Jason Yosinski, and Rosanne Liu. 2020. [Plug and play language models: A simple approach to controlled text generation](#). In *ICLR*.
- Robin IM Dunbar, Anna Marriott, and Neil DC Duncan. 1997. Human conversational behavior. *Human nature*, 8(3):231–246.
- Ari Holtzman, Jan Buys, Li Du, Maxwell Forbes, and Yejin Choi. 2020. [The curious case of neural text degeneration](#). In *ICLR*.
- Harsh Jhamtani, Varun Gangal, Eduard Hovy, Graham Neubig, and Taylor Berg-Kirkpatrick. 2018. [Learning to generate move-by-move commentary for chess games from large-scale social forum data](#). In *ACL 2018*.
- Urvashi Khandelwal, Angela Fan, Dan Jurafsky, Luke Zettlemoyer, and Mike Lewis. [Nearest neighbor machine translation](#). *CoRR*.
- Kenton Lee, Luheng He, Mike Lewis, and Luke Zettlemoyer. 2017. [End-to-end neural coreference resolution](#). In *EMNLP*.
- Jiwei Li, Michel Galley, Chris Brockett, Jianfeng Gao, and Bill Dolan. 2016. [A diversity-promoting objective function for neural conversation models](#). In *NAACL HLT*.
- Margaret Li, Stephen Roller, Ilya Kulikov, Sean Welleck, Y-Lan Boureau, Kyunghyun Cho, and Jason Weston. 2020. [Don't say that! making inconsistent dialogue unlikely with unlikelihood training](#). In *ACL*.
- Bodhisattwa Prasad Majumder, Harsh Jhamtani, Taylor Berg-Kirkpatrick, and Julian J. McAuley. 2020. [Like hiking? you probably enjoy nature: Persona-grounded dialog with commonsense expansions](#). In *EMNLP*.
- Nasrin Mostafazadeh, Nathanael Chambers, Xiaodong He, Devi Parikh, Dhruv Batra, Lucy Vanderwende, Pushmeet Kohli, and James F. Allen. 2016. [A corpus and evaluation framework for deeper understanding of commonsense stories](#). *CoRR*, abs/1604.01696.
- Lianhui Qin, Vered Shwartz, Peter West, Chandra Bhagavatula, Jena D. Hwang, Ronan Le Bras, Antoine Bosselut, and Yejin Choi. 2020. [Back to the future: Unsupervised backprop-based decoding for counterfactual and abductive commonsense reasoning](#). In *EMNLP*.
- Stephen Roller, Emily Dinan, Naman Goyal, Da Ju, Mary Williamson, Yinhan Liu, Jing Xu, Myle Ott, Kurt Shuster, Eric M Smith, et al. 2020. [Recipes for building an open-domain chatbot](#). *arXiv preprint arXiv:2004.13637*.
- Iulian Vlad Serban, Alessandro Sordani, Ryan Lowe, Laurent Charlin, Joelle Pineau, Aaron C. Courville, and Yoshua Bengio. 2017. [A hierarchical latent variable encoder-decoder model for generating dialogues](#). In *AAAI*.
- Haoyu Song, Wei-Nan Zhang, Jingwen Hu, and Ting Liu. 2020. [Generating persona consistent dialogues by exploiting natural language inference](#). In *AAAI*.
- Hui Su, Xiaoyu Shen, Sanqiang Zhao, Xiao Zhou, Pengwei Hu, Randy Zhong, Cheng Niu, and Jie Zhou. 2020. [Diversifying dialogue generation with non-conversational text](#). In *ACL*.
- Sean Welleck, Jason Weston, Arthur Szlam, and Kyunghyun Cho. 2019. [Dialogue natural language inference](#). In *ACL*.
- Jason Weston, Emily Dinan, and Alexander H. Miller. 2018. [Retrieve and refine: Improved sequence generation models for dialogue](#). In *SCAI@EMNLP*.
- Thomas Wolf, Victor Sanh, Julien Chaumond, and Clement Delangue. 2019. [Transfertransfo: A transfer learning approach for neural network based conversational agents](#). *CoRR*, abs/1901.08149.
- Jun Xu, Zeyang Lei, Haifeng Wang, Zheng-Yu Niu, Hua Wu, and Wanxiang Che. 2020. [Enhancing dialog coherence with event graph grounded content planning](#). In *IJCAI*.
- Saizheng Zhang, Emily Dinan, Jack Urbanek, Arthur Szlam, Douwe Kiela, and Jason Weston. 2018. [Personalizing dialogue agents: I have a dog, do you have pets too?](#) In *ACL*.
- Tianyi Zhang, Varsha Kishore, Felix Wu, Kilian Q. Weinberger, and Yoav Artzi. 2020. [Bertscore: Evaluating text generation with BERT](#). In *ICLR*.
- Yutao Zhu, Ruihua Song, Zhicheng Dou, Jian-Yun Nie, and Jin Zhou. 2020. [Scriptwriter: Narrative-guided script generation](#). In *ACL*.

A Implementation Details

We obtain the PersonaChat dataset from ParlAI repository². ROCStories dataset is obtained from the repository of original release³. We adapted codes from original PPLM (Dathathri et al., 2020) repository⁴ and modified them for our own objective function.

Network architecture For the generator network, we use GPT2 (Transformer with 12 layers, 768 hidden size, 12 heads—`gpt2-small`⁵) following the state-of-the-art model (Wolf et al., 2019) from Conv-AI2 competition. The decoder has total of 124 Million parameters. We used the pre-trained decoder model obtained from (Majumder et al., 2020).

Hyperparameters PABST does not require any training since we perform gradient-based decoding at the inference time. For our best method, in objective function \mathcal{L} , we use λ_d as 1 and λ_c as 1. We keep generation length to be 100 to encourage longer generations. We train the consistency classifier using code from PPLM repository⁶. The weight γ for mixing forward and backward passes was set to 0.45. For PSEUDO, we substitute a random 30% of the original target responses with retrieved stories.

B Human Evaluations Setup

Figure 2 shows a sample question for the pairwise comparison between response generated by PABST and a baseline for sensibility and engagingness. A similar setup was used to measure performance between PABST variants with different λ_d values (0.5, 1, 5).

C Qualitative Examples

Tables 4 and 5 show generations from PABST and baselines for sampled text examples.

²<http://parl.ai/downloads/personachat/personachat.tgz>

³<https://www.cs.rochester.edu/nlp/rocstories/>

⁴<https://github.com/uber-research/PPLM>

⁵<https://github.com/huggingface/transfer-learning-conv-ai>

⁶https://github.com/uber-research/PPLM/blob/master/run_pplm_discrim_train.py

Dialog History

Agent: I volunteer in shelters.

User: It is a noble act.

DISCCHOICE: Thank you. I like helping people.

RETRIEVAL: I was at my house bored. I couldn't figure out what to do. I looked on TV for answers. I saw a commercial about homeless people. So I decided to volunteer at a homeless shelter.

PSEUDO: Last week, I went to the local animal shelter on the pet adaptation day.

MULTITASK: Thank you! What about you?

PABST ($\lambda_d = 5.0$): I was bored. I wanted to help people. I looked on TV about homeless people. So I decided to volunteer at a homeless shelter.

PABST ($\lambda_d = 1.0$): Thank you. I was inspired by a TV commercial and I volunteered at the homeless shelter.

Table 4: Generations from different models

Dialog History

Agent: I enjoy American sports. What sports do you like?

User: I love playing volleyball. I was in the university team.

DISCCHOICE: That's great. I like adventure sports.

RETRIEVAL: I always thought golf was for old people. One day, my dad invited me to play golf with him. I reluctantly accepted his invitation. On the golf range, I saw lots of people of my age. I realized golf was a sport for everyone to enjoy regardless of age.

PSEUDO: Last week, I went to play golf. My dad invited me. I was reluctant.

MULTITASK: That's great! I play volleyball during weekend.

PABST ($\lambda_d = 5.0$): I like playing golf. I always thought it is for old people. Recently I went to a golf range. But I saw lots of people of my age. I realized golf is a sports for everyone.

PABST ($\lambda_d = 1.0$): That's great. I like playing golf. I always thought it is for old people. Recently I went to a golf range. But I saw lots of people of my age. I realized golf is a sports for everyone.

Table 5: Generations from different models

Instructions (Click to collapse)

This task requires basic English language understanding.

For each instance, you will have to read the dialog history between two people **A** and **B**. We expect you to respond on the following for the candidates shown for **A**'s response:

- 1) Sensible: Which response do you think is more sensible with the dialog history?
- 2) Engaging: Which response do you think is more engaging/interesting?

1. Dialog History:

A's turn: How was your last trip?
B's turn: It was great. I went to Utah and saw the Canyonlands.

Candidates for A's next turn:

Response R1: That's great. I frequently go to Florida. I have a house there but it gets hot in summer.

Response R2: I have a house in Florida.

1.1 Which response do you think is more sensible with the dialog history?

R1 is better Both have similar fluency R1 is worse

1.2 Which response do you think is more engaging/interesting?

R1 is more engaging Both have similar engagement level R1 is less engaging

Figure 2: Human evaluation setup for pairwise comparison between PABST and another baseline