RPTCS: A Reinforced Persona-aware Topic-guiding Conversational System

Zishan Ahmad[†§], **Kshitij Mishra**[†§], **Asif Ekbal**[†] and **Pushpak Bhattacharyya**[‡]
[†]Department of Computer Science and Engineering,
Indian Institute of Technology Patna, India
{1821cs18, kshitij_1921cs23, asif}@iitp.ac.in
[‡]Department of Computer Science and Engineering,
Indian Institute of Technology Bombay, India
pb@cse.iitb.ac.in

Abstract

Although there has been a plethora of work on open-domain conversational systems, most of these lack the mechanism of controlling the concept transitions in a dialogue. For activities like switching from casual chit-chat to task-oriented conversation, an agent with the ability to manage the flow of concepts in a conversation might be helpful. The user would find the dialogue more fascinating and engaging and be more receptive to such transitions if these concept transitions were made while taking into account the user's persona. Focusing on persona-aware concept transitions, we propose a Reinforced Persona-aware Topic-guiding Conversational System (**RPTCS**). Due to the lack of a persona-aware topic transition dataset, we propose a novel conversation dataset creation mechanism in which the conversational agent leads the discourse to drift to a set of target concepts depending on the persona of the speaker and the context of the conversation. To avoid scarcely available expensive human resources, the entire data-creation process is mostly automatic with human-in-loop only for quality checks. This created conversational dataset named **PTCD** is used to develop the **RPTCS** in two steps. First, a maximum likelihood estimation loss-based dialogue model is trained on **PTCD**. The trained model is then fine-tuned in a Reinforcement Learning (RL) framework by employing novel reward functions to assure persona, topic, and context consistency with non-repetitiveness in generated responses. Our experimental results demonstrate the strength of the proposed system with respect to strong baselines[1].

1 Introduction

Due to the abundance of conversational corpora, there has been a great interest in building open-domain conversational systems[2] (Huang et al., 2020). While capable of producing fluent responses, these systems also frequently generate generic responses devoid of any useful information (Gao et al., 2019). Human discourse is complex and difficult to replicate as it often covers a wide range of topics (Winograd, 1977). Hence, the key to a multi-turn conversation's success is striking a balance between effectively changing the subject and keeping it on-topic (See et al., 2019).

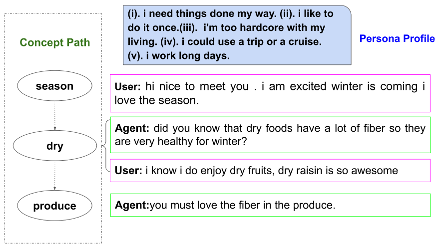

Figure 1: An illustration of the suggested task. The persona profile and the concept path to the target concept serve as the foundation for the created dialogues.

It is also seen that even if open-domain conversational systems employ topic shift, they are not able to engage user due to absence of knowledge of user's own perceptions. Therefore, to achieve a robust topic shift guiding conversational system, it has to evolve away from the prior factual information systems (Leuski et al., 2006) towards a novel blend of both task-oriented and non-task-oriented systems (Akasaki and Kaji, 2017). However, they may fail to engage user in difficult multi-turn information-seeking tasks or dialogues (Trippas et al., 2020) and may also suffer from "anomalous state of knowledge" (Belkin and Vickery, 1985) where the user has vague information requirements and is often struggling to articulate it with enough precision in a conversation. Thus, we require context-sensitive user guidance that does not assume a rigid hierarchy of the user's plans and

[§]equal contribution

[1]Codes and dataset available at https://github.com/zishan-ahmad-nlp/persona-topic-shift

[2]we use the terms conversational system/agent, dialogue system/agent, agent, and chatbot interchangeably.

work objectives. Such a guidance, can be achieved by utilising persona information of user with the ongoing dialogue. This will drive the conversation from one topic to a different one in consonance with the user's interests.

Focusing on these two aspects, we propose a reinforced persona aware topic guiding conversational system (**RPTCS**). Due to scarcity and expensiveness of human resource, first, employing minimal manual intervention, we present a novel persona aware topic guiding conversational dataset (**PTCD**) created employing a novel automatic conversation creation mechanism. Majority of the topic shift guiding conversational systems leverages only the concept space knowledge to represent the potential conversation flow (Wu et al., 2019; Zhou et al., 2020; Zhang et al., 2020; Wang et al., 2021; Qiu et al., 2022). These approaches are predicated on a restricted defined logical premise: individuals would be interested in talking about any random concept. This presumption can be too straightforward to mimic topic flows in real human conversations. In this work, we take a significant step toward developing a proactive conversational agent that is more attuned to a user's personality with a clear conversation aim. We aim at training a system that is able to decide on topic shifts and generate responses that reflect these logical topic shifts, while also considering the user persona. Figure 1 shows an example of this task. Since bringing an open-domain conversation to a task-oriented domain is one of the key use cases for this system, we chose the concepts popular in the latter domain (Eg. restaurant, travel, clothing, smartphone, etc.).

Using **PTCD**, we first train a maximum likelihood estimation (MLE) loss-based conversational employing a language model (LM) GPT-2 small (Radford et al., 2019). This trained model is then fine-tuned in a Reinforcement Learning (RL) framework using novel rewards designed to assure persona, concept, and context consistency with non-repetitiveness in the generated responses. The key contributions of this work are *four-fold* **(i).** Proposal of a new task to build a persona-aware conversation system for targeted topic transitions and creation of a novel datatset **PTCD** to tackle the task in-hand; **(ii).** A novel semi-automatic persona aware topic guiding corpus creation method using pre-trained models (can be adapted to the other similar tasks in the broad area of dialogue systems), **(iii).** Designing a novel reward function to built the proposed system **RPTCS** in an RL framework. **(iv).** Performed extensive automatic and human evaluation by designing two novel evaluation metrics to evaluate topic and context consistency to demonstrate strength of the proposed system.

2 Related Work

To have a human-like conversation and avoid off-topic response generation, several methods have been explored (Ghazvininejad et al., 2018; Harrison et al., 2020). The research community has employed ConceptNet (Speer et al., 2017) to mimick concept transitions in human conversations, (Zhou et al., 2018), such as: Zhang et al. (2020) used concept relations with 2-hop nodes, to account more thoroughly human concept shifts. However, none of the works have taken into account the persona of the user to guide the topic shift in the conversation which may act as a crucial component to have effective communication.

Some benchmark conversation datasets has been proposed to assess the conversation focusing on different personal attributes such as: (He et al., 2017) predicts the next topic based on social and cultural situation, ETHICS dataset constitutes the conversation based on the concept of morality (Hendrycks et al., 2020). Hsu et al. (2018) presents emotion labelled EMOTIONLINES dataset, Yu et al. (2020) proposes a social relation inference based DIALOGRE dialogue dataset. Zhang et al. (2018) proposed PERSONA-CHAT dataset to make chit-chat dialogues more engaging by conditioning them on user's profile information. Most of these datasets are collected through crowdsource workers. To obtain such a large human resource is expensive, time-consuming, and can be infeasible. We investigate an automatic data creation approach requiring minimal manual intervention as an alternative to this.

Recently, Wang et al. (2021) proposed a multi-turn topic-driven NaturalConv dataset ensuring a smooth topic shift in conversations. Our work here is different from the existing topic-shift guiding conversational systems in three aspects. First, we followed an automatic dialogue data curation process. Second, the topic shift in the conversations is guided using ConceptNet as well as the persona of the user. This ensures both goals of an effective communicator *viz.* completing the task and maintaining the face of the user. Third, we built a persona aware topic guiding conversational system in an RL framework using novel efficient and effective reward function. To the best of our knowledge, our proposed setup is novel and has not been tried before.

3 Dataset Creation

To tackle the persona-aware concept shift to a target concept, we create and propose a novel dataset. We come up with a novel semi-automatic data creation technique involving prompting a GPT-J model and human intervention for quality control. The entire dataset creation consists of the following steps:

Obtaining Seed Data: To start the few-shot dialogue generation, we required a seed utterance from a user with some assigned persona. To build In PERSONA-CHAT-grounded topic-shifting conversations"the first utterance of a conversation (focused on a persona-profile) in the PERSONA-CHAT (Zhang et al., 2018) is selected. In PERSONA-CHA", persona profiles are defined as "profiles that are natural and descriptive, and contain typical topics of human interest that the speaker can bring up in conversation". The concepts in the seed utterance is obtained by extracting words with *'nouns'*, *'adjectives'*, and *'verbs'* Part-of-Speech (PoS) tags. We call these first utterance concepts as the source concept. The objective here is to guide the conversation away from the source concept and toward a target concept. Concepts like *'travel'*, *'restaurant'*, *'shopping'*, *'electronics'*, etc. are chosen as the system's target concepts.

Concept Path Creation and Selection: To have smooth and logical transition and avoid abrupt source-to-target concept jump, we make use of Concept-Net (Speer et al., 2017). First, employing Dijkstra's algorithm, the shortest distance between all the source-concepts in the seed dialogue and all the target-concepts is calculated. This results in multiple paths between each source and target concepts. From these multiple paths only one path has to be selected for the desired concept transition in dialogues, i.e., based on the ongoing topic and persona, the conversation must be able to switch to the most suitable target and path. Although it would be ideal for the conversation to veer toward a subject (concept) relevant to the user persona, the context of the conversation may not permit it. Hence, a balance is aimed between context and persona relevance. We extend the use of RoBERTa (Liu et al., 2019) for contextual concept selection as proposed by Yasunaga et al. (2021) to concept-path selection. Let $P = \{p_1, p_2, ..., p_N\}$ be the persona profiles of a user where $p_i = \{w_1, w_2, ..., w_m\}$ is a persona sequence. Each concept-path t to the target is denoted by $C_t = \{c_1, c_2, ..., c_t\}$, where c_j is a concept in the path. For each pair of persona sequence and concept-path, persona relevance probability is computed using the RoBERTa head (LM_{head}). The probabilities are then averaged to obtain $P_{persona}$ as shown in Equation 1 and 2.

$$P_{seqp} = LM_{head}(LM_{enc}([p_i : C_k])) \tag{1}$$

$$P_{persona} = Avg(P_{seqp}[c_{1k}, c_{2k}, ..., c_{tk}]) \tag{2}$$

Here, [:] is the concatenation operation. Using the same RoBERTa LM_{head}, the path probability P_{conv} with respect to the conversational context for a conversation $D = \{U_1, U_2, ..., U_n\}$ is computed (U_i are utterances in the dialogue). To obtain P_{conv}, P_i in encoder LM_{enc} are replaced with D_p (Equation 3 and averaging the probabilities of the concepts (Equation 2).

$$P_{seqc} = LM_{head}(LM_{enc}([D_p : C_k])) \tag{3}$$

$$P_{conv} = Avg(P_{seqc}[c_{1k}, c_{2k}, ..., c_{tk}]) \tag{4}$$

From the obtained persona $P_{persona}$ and contextual P_{conv} concept-path probabilities, the path with the highest probability is selected.

Utterance Generation: After selecting seed utterances, new utterances are generated utilizing GPT-J (Wang and Komatsuzaki, 2021) (Brown et al., 2020). Taking advantage of few-shot prompting capability of the GPT-J model, few manually written sample conversations following the appropriate persona and context relevant topic transitions are used. [3]. This prompt is followed by the current seed dialogue and selected path and given as input to GPT-J, which generates the next utterance of the dialogue. To lead the conversation, all concept transitions are initiated by the agent. Top_k sampling (Fan et al., 2019; Radford et al., 2019) is used during utterance generation and 10 candidate responses for each input utterance is generated.

Utterance Selection: Best response out of the 10 candidates is selected by taking into account persona-entailment, current concept relevance, and conversation context relevance. To assess the level of persona-entailment in the produced outputs, a Natural Language Inference (NLI) model using the Dialog NLI dataset is trained (Welleck et al.,

[3]Example of the designed prompt is given in Section A.2 of the appendix

	Total	Train	Test	Valid
#Dialogues	2,586	1,945	406	236
#Utternaces	13,746	10,310	2,436	1,000
#Unique Concepts	1,843	1,505	617	314
#Unique Paths	1,738	1,372	316	184
Avg Path Length	3.55	3.54	3.57	3.57

Table 1: Statistics of the dataset created

2019). by fine-tuning a pre-trained BERT (Devlin et al., 2018) model[4]. The probability score of entailment $P_{ent}^{p_i}$ of a given utterance U_k with respect to a persona p_i are obtained by final *softmax* layer. Averaging the $P_{ent}^{p_i}$ gives the final entailment score $P_{ent}^{U_k}$ for a candidate utterance U_k. This entailment score is calculated for both the user and agent utterances to counteract the contradiction to a user's preferences.

To make the generated response pertinent to the conversation's context, contextual probability score of the candidate utterance U_k is computed. To do so, last utterance in the conversation U_{k-1} is taken in a sequence $S_{ctx} = U_{k-1}$ *[SEP]* U_k, which in turn is given as input to BERT to obtain the U_k's probability score $P_{ctx}^{U_k}$ with the next sentence.

To ensure the appropriateness of the created utterance with the current concept, a concept-relevance score is calculated.Again the BERT model is used to encode the utterance U_k and current concept c_j to obtain the embedding representation U_k^{emb}, and c_j^{emb} respectively. Then, the concept-relevance score $P_{cpt}^{U_k}$ is computed by taking the cosine similarity between U_K^{emb} and c_j^{emb}. Equation 5 is used to calculate each candidate's final score for each speech.

$$P_{final}^{U_k} = \alpha P_{ctx}^{U_k} + \beta P_{ent}^{U_k} + \gamma P_{cpt}^{U_k} \quad (5)$$

We set the values of the constants α, β, and γ as 0.38, 0.30, and 0.32 through several quality evaluations (Detailed in Section A.1.2). The candidate with the highest P_{final} score is chosen as the next utterance.

Manual Filtering of Dialogues: Once the complete conversational dataset is obtained, by iteratively employing the above steps. These dialogues are quality-checked by human evaluators. Each utterance is scored using two parameters, *viz.* Humanness and Concept Consistency. Each utterance is evaluated with a score of 0, 1, or 2, w.r.t these two parameters where 0 denotes the lowest and 2 denotes the highest. Three human experts with post-graduate qualifications were asked to rate all the generated utterances. These professionals are regular employees in our research group and have 2 years of expertise in related fields. Following the ratings, dialogues with utterances having a score of 0 by any expert for 'humanness' or 'concept consistency' are eliminated.

[4]Test set accuracy of NLI model is obtained as 88.43%.

4 Persona-aware Topic-guiding Conversational System

We build our proposed system **RPTCS** in two phases. In first phase, an MLE-loss-based conversational system (MLCS) is trained to learn the user and agent's utterances distribution and obtain natural language interaction. In the second phase, this MLCS is fine-tuned in an RL framework employing proximal policy optimization (PPO) (Schulman et al., 2017)ss method to generate persona-aware topic guiding utterances.

4.1 Phase 1

A multi-turn conversation can be represented as $C = \{a_0, u_0, ..., a_{T-1}, u_{T-1}\}$, where, a and u denotes the agent's and user's utterances. Each conversation is further attributed with a persona $p = \{p_1, p_2, ..., p_m\}$ having a set of m persona statements of user, a topic-path $tp = \{tp_1 \rightarrow tp_2 \rightarrow .. \rightarrow tp_k\}$, with k topics, had to be followed in conversation, and with each of the agent's utterance, a $< topic >$ (any one tp_i from tp) to which conversation has to be guided. Following (Wu et al., 2021), the probability distributions over the conversation C's utterances concatenated with user's persona p, topic-path tp, and agent's $< topic >$ to be guided are decomposed into two LMs *viz.* one for user and other for agent denoted as ρ_u and ρ_a, respectively. Given the conversation context, the LMs ρ_u and ρ_a predict the next token in an agent's generated response $r = \{r_1, r_2, ..., r_t\}$ with t tokens. The joint probability for an utterance u_i or a_i can be formulated as:

$$\rho_u(u_i|u_{<i}, a_{<i}) = \prod_{j=1}^{t_{u_i}} \rho(r_j|r_{<j}, p, tp, u_{<i}, a_{<i}) \quad (6)$$

$$\rho_a(a_i|u_{<=i}, a_{<i}) = \prod_{j=1}^{t_{a_i}} P(r_j|r_{<j}, tp_{<i}, u_{<=i}, a_{<i}) \quad (7)$$

User Persona	(i). i enjoy killing sea creatures. (ii). i'm a fan of animals. (iii).i'm practically a chef ! (iv). my father was in the car industry. (v). i enjoy life.	
Seed User Utterance	hey ! i'm a happy camper this evening . just finished making dinner .	
Selected Concept Path	['camper', 'backpack', 'food', 'restaurant']	
Speaker	***Generated Utterance***	***Concept***
Agent	i am a fan of backpacking. i know a lot of people who are happy in their camp.	backpack
User	i would love to go backpacking but i dont have enough money to afford one.	backpack
Agent	i have been known to carry around my backpack full of food, i just have a big appetite.	food
User	i just love food.	food
Agent	you must know your restaurants then.	restaurant

Table 2: Sample dialogues from the corpus generated using our method.

Finally, the conversational system $\rho_\theta(C)$ defined on conversation C is trained by maximizing the likelihood estimation.

$$\rho_\theta(C) = \prod_{T=0}^{T-1} \rho_u(u_i|u_{<i}, a_{<i})\rho_a(a_i|u_{<=i}, a_{<i}) \tag{8}$$

4.2 Phase 2

To generate persona-aware and topic-consistent responses $\rho_\theta(C)$ (MLCS) is fine-tuned in an RL framework. For a given context, $\rho_\theta(C)$ generates *n* possible candidates. These candidate responses are then quality-checked in terms of persona awareness, topic consistency, context consistency, and repetition using respective rewards.

4.2.1 Rewards

To achieve utterance-persona and utterance-topic consistency in generated responses, we design two novel task-specific rewards *viz.* Utterance-Persona consistency (R_1) and Utterance-Topic consistency (R_2). Similarly, to ensure conversation properties like contextual correctness and non-repetitiveness, two generic rewards are designed *viz.* Context consistency (R_3) and Non-Repetition (R_4). Lastly, a compound reward function R, considering both task-specific and generic rewards is computed, which outputs the end reward value for the generated candidate response.

Utterance-Persona consistency Reward: The essence of engaging response generation also relies on persona of the user, hence a conversational system should be able to maintain the utterance-persona consistency in the generated responses. This problem of utterance consistency with persona statements of user can be characterized as a natural language inference (NLI) problem, having three labels *viz.* entailment, neutral and contradiction. Entailed responses are consistent with persona whereas contradictory responses are inconsistent, hence, should be penalized. To build our NLI model, BERT (Devlin et al., 2018) is employed as a classifier, which takes input persona p and generated candidate response r with a $< SEP >$ tag, and outputs one of three classes through a hot vector $[c_e, c_n, c_c]$ (entailment c_e, neutral c_n or contradiction c_c). To achieve the respective class probabilities, a softmax is applied on this hot-vector, i.e. $[prob_{c_e}, prob_{c_n}, prob_{c_c}] = softmax([c_e, c_n, c_c])$. The predicted entailed probability is used to design the reward R_1 which can be written as:

$$R_1 = prob_{c_e}(r_T, p) \tag{9}$$

where r_T represents the generated response at turn T. It can be inferred that R_1 will reward more, if entailment probability is high.

Utterance Topic Consistency Reward: A topic guiding conversational system should not deviate from the topic in-hand. It should be forced to generate topic consistent utterances by rewarding the ones which employ the required topic. To formulate the utterance topic consistency, we considered *cosine* similarity between topic and utterance which is calculated using Sentence-BERT (Reimers and Gurevych, 2019). R_2 can be written as:

$$R_2 = \cos(r_T, tp_T) \tag{10}$$

where tp_T represents the topic at turn T. Higher cosine similarity values will lead to higher rewards for utterance topic consistency.

Context Consistency Reward: A conversational system is required to generate context consistent responses. Therefore, to assess context consistency, we devise a reward by calculating *cosine* similarity (using Sentence-BERT (Reimers and Gurevych, 2019)) between generated response and agent's and user's utterance at turn T. R_3 can be formulated as:

$$R_3 = \frac{1}{2} \times (\cos(r_T, a_T) + \cos(r_T, u_T)) \tag{11}$$

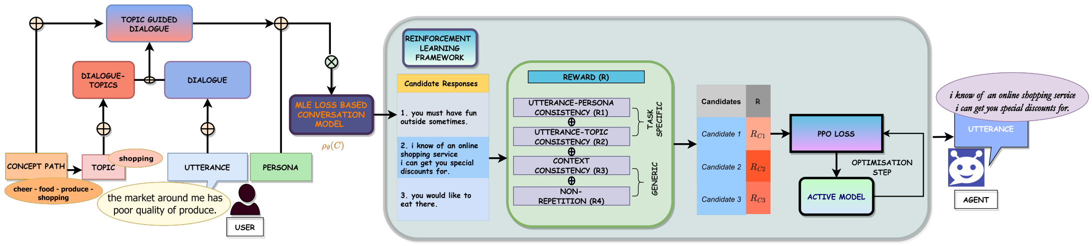

Figure 2: Architectural overview of the proposed **RPTCS**. It is first initialized with trained MLE-loss based model $\rho_\theta(C)$, then is fine-tuned employing PPO-loss utilizing novel reward R to build a persona aware topic guiding conversational system (**RPTCS**)

.

Non-Repetition Reward: Due to repetitive same response generation in a conversation, user's may not show engagement in the conversation. Hence, the generated responses should be diverse with each turn of the conversation. To ensure this, we design a Non-repetition reward R_4 by computing Jaccard distance (Jaccard, 1912; Tanimoto, 1958) between generated responses r_T and r_{T-1} at turns T and $T-1$, respectively. R_4 can be given as:

$$R_4 = 1 - \left(\frac{gr_{T-1} \cap gr_T}{gr_{T-1} \cup gr_T} \right) \tag{12}$$

Reward Function: Lastly, to train the whole system, reward function R is formulated using weighted sum of all these four rewards. To approximate a better function, the sum of all weights has been taken equal to 1.

$$R = \delta_1 R_1 + \delta_2 R_2 + \delta_3 R_3 + \delta_4 R_4 \tag{13}$$

where $\delta_1 + \delta_2 + \delta_3 + \delta_4 = 1$. The obtained value of R assesses the quality of generated response, which is further used to optimize the PPO loss, such that model can learn to generate persona, topic, and context-consistent responses.

4.2.2 Policy

A probability mapping function $\mathcal{P}_\theta$ representing the probability of generating an utterance r consisting of L tokens gives the policy.

$$\mathcal{P}_\theta(r_{1:L}|x) = \prod_{l=0}^{L} \mathcal{P}_\theta(r_l|y_{<l}, x) \tag{14}$$

Proximal Policy Optimisation: Policy updates at each each step are done using PPO method to ensure low variance. It seeks improvement on certain parameters to update the existing policy such that it is not too different from the old policy. First, expected reward is maximized using gradient ascent on loss function $J(\theta)$,

$$\nabla_\theta J(\theta) = Ep_{r \sim \mathcal{P}_\theta}[\nabla_\theta \log \mathcal{P}_\theta(r) \hat{A}_r] \tag{15}$$

Second, large deviations from old policy are restricted by replacing the log term with an importance sampling and clipping is performed to prevent catastrophic forgetting. It relies on specialized clipping without any KL-divergence term (Kullback and Leibler, 1951) or any constraint in the objective function.

$$L^{\text{CLIP}}(\theta) = \hat{E}[\min(pr_r(\theta)\hat{A}_r, \text{clip}(pr_y(\theta), 1-\varepsilon, 1+\varepsilon)\hat{A}_r)]$$

Here, $pr_r(\theta) = \mathcal{P}_\theta^{new} / \mathcal{P}_\theta^{old}$ denotes the ratio of probabilities of the generated response between the new and old policies. ε is the clipping range and $\hat{A}_y$ is the estimated advantage which is the normalized rewards in our case. Lastly, parameters are updated using the following steps:

$$\theta_{k+1} = \underset{\theta}{\text{argmax}} \underset{s,a \sim \mathcal{P}_{\theta_k}}{E} [L^{\text{CLIP}}] \tag{16}$$

5 Experiments

5.1 Implementation Details

MLE loss based conversational system (MLCS) is trained by employing two pre-trained GPT-2 small (Radford et al., 2019) models, one for user and other for agent. To fine-tune trained MLCS in RL-setting, $n = 3$ is selected as per better loss, after experimenting with different values of candidate responses i.e. $n = 2, 3, 4, 5, 10$. The candidate responses are decoded using nucleus sampling (Holtzman et al., 2019) with temperature $T = 0.8$ and probability $p = 0.9$. To train the RPTCS $seed\_value = 10$, $human\_reward = 10$, $max\_candidate\_length = 50$ is adopted with $optimizer = AdamW$ (Loshchilov and Hutter,

2017) and learning rate $\alpha = 2e-05$, $\varepsilon = 0.2$ and $epochs = 17$. The reward weight combination of $0.3, 0.3, 0.2, 0.2$ are chosen as the final weights for δ_1, δ_2, δ_3, and δ_4, respectively (detailed weight optimization is given in Table 5 of appendix). Due to space restrictions, detailed implementation details of the NLI model to entail the utterance persona consistency, and MLCM are given in Section A.1.3 of the appendix.

5.2 Evaluation Metrics

Both automatic and human evaluations are performed to assess the performance of the proposed system **RPTCS**. To evaluate all three tasks of utterance-persona, utterance-topic and context consistency, three metrics *viz.* **U-PCon**, **U-TCon**, and **CxCon** respectively, are adopted. **U-PCon** computation can be formulated as below:

$$U-PCon = \frac{\sum_{j=1}^{n} class_i(NLI)}{n} \tag{17}$$

where, $class(NLI)$ gives one of three classes 1 (entailment), 0 (neutral), -1 (contradiction) and n is the number of generated responses. **U-TCon** is calculated as *cosine-similarity* between the generated response r_T and topic tp_T, i.e. $\cos(r_T, tp_T)$. **CxCon** calculates the average METEOR (Banerjee and Lavie, 2005) score for n number of generated responses with respect to the ground truth response a_T at turn T. It can be given as:

$$CxCon = \frac{METEOR(r_T, a_T)}{n} \tag{18}$$

Further, to evaluate the quality of the generated responses, Perplexity (**PPL**) and response length (**R-LEN**) are calculated.

Human evaluation are performed by three evaluators (regular employees in our research group) with postgraduate experience and having proficiency in similar tasks [5]. Each evaluator is asked to interact with the proposed system 20 times and evaluate the conversations in terms of task-specific and generic metrics. Former includes **PerAw**, **TopGu** - to assess the persona awareness and topic guidance respectively in generated conversations. Latter includes **Fluen**, **Const**, and **N-Rep** - to check fluency, consistency, and non-repetitiveness of generated responses in interacted conversations. All metrics are evaluated on an integer scale of 1-5 [6].

[5] Human evaluators were paid as per our university norms.

[6] The scale 1-5 denotes low to high intensity such as PerAw = 1 denotes highly persona awareness and PerAw = 5 denotes no persona awareness.

6 Results and Analysis

RPTCS is compared with two baselines: **(1.)** ARDM (Wu et al., 2021): A MLE loss based self-play conversation model, which trains two GPT-2 medium models (one for user and one for agent) alternatively. Here, as we are assessing only agent's performance, hence we train only one GPT-2 small model. **(2.)** RPTCS-R: **RPTCS** with $R = 0$. Further, to check the effects of each of the three aspects *viz.* concept-path, topic and persona, three respective variants of ARDM *viz.* ARDM+CP (ARDM considering concept-path), ARDM+CP+T (ARDM considering concept-path and topic), and ARDM+CP+T+P (ARDM considering concept-path, topic and persons) are also trained to compare. Lastly, to assess the importance of both task-specific (Ts) and generic (Ge) rewards, **RPTCS** is also compared with RPTCS-Ts ($R = \delta_3 R_3 + \delta_4 R_4$) and RPTCS-Ge ($R = \delta_1 R_1 + \delta_2 R_2$).

Automatic Evaluation: It can be noticed in Table 3 that the proposed **RPTCS** performs better than all four baselines *viz.* ARDM, RPTCS-R, RPTCS-Ts and RPTCS-Ge in terms of all the four metrics *viz.* **U-PCon**, **U-TCon**, **PPL** and **R-LEN**. For task-specific metrics **U-PCon**, and **U-TCon**, **RPTCS** achieves better scores of *95.8%*, and 0.414, respectively, with a significant difference of <18.9, 0.124>, <7.7, 0.095>, <8.4, 0.1>, <6.7, 0.042>, and <3.5, 0.047> than the baselines ARDM, ARDM+CP+T+P, RPTCS-R, RPTCS-Ts and RPTCS-Ge, respectively. It can also be inferred that the difference of **U-TPer**, **U-TCon**, **CxCon** scores decreased in order ARDM>RPTCS-R>RPTCS-Ge>RPTCS-Ts. This shows the importance of task-specific rewards in our proposed system **RPTCS** and it can be argued that utterance-persona and utterance-topic consistency rewards do force the system to adapt towards generating persona and topic-consistent responses.

It can also be observed in the Table 3 that **RPTCS** obtains better **PPL** = 6.14 score than that of ARDM, ARDM+CP+T+P, RPTCS-R, RPTCS-Ts and RPTCS-Ge with a difference of 2.53, 1.51, 1.10, 0.89, and 0.84, respectively. This may be due to the Context consistency reward which drives the model to generate responses consistent with the conversation context, which, in turn, leads to the generation of much more natural and fluent responses. Further, obtained a score of **R-LEN** = 15.23 is also better than that of ARDM, ARDM+CP+T+P, RPTCS-R, RPTCS-

Model	U-PCon	U-TCon	CxCon	PPL	R-LEN
ARDM (Wu et al., 2021)	76.9%	0.290	0.134	8.67	13.11
ARDM+CP	77.2%	0.303	0.145	8.34	13.21
ARDM+CP+T	77.4%	0.314	0.147	8.12	13.27
ARDM+CP+T+P	88.1%	0.319	0.160	7.65	13.86
RPTCS-R	87.4%	0.314	0.152	7.24	13.42
RPTCS-Ts	89.1%	0.372	0.161	7.03	14.27
RPTCS-Ge	92.3%	0.367	0.166	6.98	14.69
RPTCS	**95.8%**	**0.414**	**0.178**	**6.14**	**15.23**

Table 3: Results of automatic evaluation

Model	PerAw	TopGu	Fluen	Const	N-Rep
ARDM	2.69	2.13	3.11	3.83	2.94
ARDM+CP+T+P	3.41	2.64	3.57	4.02	3.21
RPTCS-R	3.31	2.66	3.65	3.98	3.26
RPTCS-Ts	3.40	2.70	3.71	4.12	3.42
RPTCS-Ge	3.62	2.81	3.84	4.04	3.37
RPTCS	**3.82**	**2.96**	**4.01**	**4.14**	**3.55**

Table 4: Results of human evaluation

Ts, and RPTCS-Ge with a difference of 2.12, 1.37, 1.81, 0.99, and 0.54, respectively. This indicates that the **RPTCS** is able to generate longer responses, hence, showcasing more engagingness with the user. It can be due to the incorporation of all four rewards where R_1, R_2, and R_3 play the crucial role of persona, topic, and context consistency, and R_4 maintain the non-repetitiveness, hence, driving the agent to build the rapport with a user as well as be on the goal topic by generating diverse and interactive responses. Engagingness rewards - forcing the model to generate more interactive and engaging responses. Lastly, it can also be seen in Table 3 that **PPL ter** and **R-LEN** scores of **RPTCS** decreased in order ARDM>RPTCS-R>RPTCS-Ge>RPTCS-Ts>**RPTCS**, hence, strengthening our hypothesis for the requirement of both task-specific and generic rewards to generate persona aware topic consistent responses.

Human Evaluation: Table 4 shows the human evaluation results for all five models *viz.* ARDM, ARDM+CP+T+P, RPTCS-R, RPTCS-Ts, RPTCS-Ge and ARDM, RPTCS-R, RPTCS-Ts, and **RPTCS**. It can be noted that **RPTCS** yields better scores of **PerAw**, **TopGu**, **Fluen**, **Const** and **N-Rep** as compared to the the baselines, ARDM, RPTCS-R, RPTCS-Ts and RPTCS-Ge. Scores of **Fluen**: 4.01, **Cons**: 4.14, and **N-Rep**:3.55 implies that all the four rewards R_1, R_2, R_3, and R_4 play a critical role in obtaining most fluent, consistent, and non-repetitive responses as compared to other four models. Further, in terms of **PerAw**, and **TopGu**, **RPTCS** attains well scores of 3.82, and 2.96, respectively, showcasing the importance of rewards R_1, and R_2. Therefore, it can be inferred that employing utterance-persona and utterance-topic consistency rewards helps the proposed model to generate persona-aware responses and is able to guide the conversation keeping the topic intact.

7 Conclusion

For an open domain conversation to be successful, proper concept transition befitting the context is essential. Further, persona awareness of the user may lead to user adaptive response generation, hence, resulting in more engaging and interactive conversations. Therefore, a conversational agent should be able to transition the concept efficiently as well as build a rapport with the user by understanding his/her persona. To encompass both of these aspects, we proposed here a Reinforced Persona-aware Topic-guiding Conversational System (**RPTCS**). First, we create a persona-aware topic transition dataset (PTCD) by leveraging the few-shot prompt feature of the language model. Second, employing GPT-2 small, we train an MLE loss-based conversational model (MLCM) on PTCD. Lastly, using a novel designed reward function to ensure aspects of persona, topic, and context consistency with non-repetitiveness, we fine-tune MLCM adopting PPO loss optimizer in an RL framework. Automatic and human evaluation results strengthen the design and use of rewards and concludes that our proposed model **RPTCS** achieves state-of-the-art performance compared to the strong baselines. Our results also concludes that **RPTCS** is able to retain and facilitate persona awareness, naturalness, and consistency at par in an ongoing dialogue.

In the future, we would like to look into personality traits such as age, gender, etc. to model a persona-aware conversational system.

8 Limitations

RPTCS has also some limitations. First, to create the data, GPT-J is used which requires a large GPU memory size (here, 40 GB). Further, empirical analysis for each of the possible combinations of different rewards weights may lead to model training and validation time to months. Hence, some heuristics should be used to choose the set of combinations of rewards and reward weights (such as here, we restricted the reward weight sum as 1).

When interacting with the system, if users continuously state short and direct responses such as *'Yes'*, *'I don't know'*, *'No'*, *'I can'*, *'Okay'*, then the system first tries to respond by inquiring about topic like *'restaurant'*, *'job'*, *'shopping'* or *'travel'* but after two or three turns starts deviating and generating out of the context or hallucinated responses. It is also seen that sometimes model starts attending persona statements of the user frequently and generate most of the time only persona awared responses which tend to be out of context. Hence, the model should be forced to generate only relevant responses and persona attention should be controlled. This opens up the door for future studies to build a controlled persona aware topic guiding conversational system.

9 Ethical considerations

We use a freely available dataset under a Creative Commons license to create our new dataset. The dataset has been used only for academic purposes, and in complete compliance with the license. The dataset created in this work will be made available only after filling and signing an agreement declaring that the data will be used only for research purposes. The annotation for manual evaluations was done by human experts, who are regular employees of our research group and are paid in accordance with the institute's policy. There are no other issues to declare.

10 Acknowledgement

The research reported in this paper is an outcome of the project "Autonomous Goal-Oriented and Knowledge-Driven Neural Conversational Agents", sponsored by Accenture LLP. Kshtij Mishra acknowledges the support of the "Prime Ministers Research Fellowship (PMRF)" Program of the Government of India.

References

Satoshi Akasaki and Nobuhiro Kaji. 2017. Chat detection in an intelligent assistant: Combining task-oriented and non-task-oriented spoken dialogue systems. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1308–1319.

Satanjeev Banerjee and Alon Lavie. 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*, pages 65–72.

Nicholas J Belkin and Alina Vickery. 1985. *Interaction in information systems: A review of research from document retrieval to knowledge-based systems.* 025.04 BEL. CIMMYT.

Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. 2020. Language models are few-shot learners. *Advances in neural information processing systems*, 33:1877–1901.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

Angela Fan, Mike Lewis, and Yann Dauphin. 2019. Strategies for structuring story generation. *arXiv preprint arXiv:1902.01109*.

Jianfeng Gao, Michel Galley, Lihong Li, et al. 2019. Neural approaches to conversational ai. *Foundations and trends® in information retrieval*, 13(2-3):127–298.

Marjan Ghazvininejad, Chris Brockett, Ming-Wei Chang, Bill Dolan, Jianfeng Gao, Wen-tau Yih, and Michel Galley. 2018. A knowledge-grounded neural conversation model. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32.

Vrindavan Harrison, Juraj Juraska, Wen Cui, Lena Reed, Kevin K Bowden, Jiaqi Wu, Brian Schwarzmann, Abteen Ebrahimi, Rishi Rajasekaran, Nikhil Varghese, et al. 2020. Athena: Constructing dialogues dynamically with discourse constraints. *arXiv preprint arXiv:2011.10683*.

He He, Anusha Balakrishnan, Mihail Eric, and Percy Liang. 2017. Learning symmetric collaborative dialogue agents with dynamic knowledge graph embeddings. *arXiv preprint arXiv:1704.07130*.

Dan Hendrycks, Collin Burns, Steven Basart, Andrew Critch, Jerry Li, Dawn Song, and Jacob Steinhardt. 2020. Aligning ai with shared human values. *arXiv preprint arXiv:2008.02275*.

Ari Holtzman, Jan Buys, Li Du, Maxwell Forbes, and Yejin Choi. 2019. The curious case of neural text degeneration. *arXiv preprint arXiv:1904.09751*.

Chao-Chun Hsu, Sheng-Yeh Chen, Chuan-Chun Kuo, Ting-Hao Huang, and Lun-Wei Ku. 2018. Emotionlines: An emotion corpus of multi-party conversations. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*.

Minlie Huang, Xiaoyan Zhu, and Jianfeng Gao. 2020. Challenges in building intelligent open-domain dialog systems. *ACM Transactions on Information Systems (TOIS)*, 38(3):1–32.

Paul Jaccard. 1912. The distribution of the flora in the alpine zone. 1. *New phytologist*, 11(2):37–50.

Diederik P. Kingma and Jimmy Ba. 2015. Adam: A method for stochastic optimization. In *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*.

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

Anton Leuski, Ronakkumar Patel, David Traum, and Brandon Kennedy. 2006. Building effective question answering characters. In *Proceedings of the 7th SIGdial Workshop on Discourse and Dialogue*, pages 18–27.

Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. 2019. Roberta: A robustly optimized bert pretraining approach. *arXiv preprint arXiv:1907.11692*.

Ilya Loshchilov and Frank Hutter. 2017. Decoupled weight decay regularization. *arXiv preprint arXiv:1711.05101*.

Adam Paszke, Sam Gross, Soumith Chintala, Gregory Chanan, Edward Yang, Zachary DeVito, Zeming Lin, Alban Desmaison, Luca Antiga, and Adam Lerer. 2017. Automatic differentiation in pytorch.

Liang Qiu, Yizhou Zhao, Jinchao Li, Pan Lu, Baolin Peng, Jianfeng Gao, and Song-Chun Zhu. 2022. Valuenet: A new dataset for human value driven dialogue system. *IEEE Trans. on Pattern Analysis and Machine Intelligence (TPAMI)*, 44(5):2468–2484.

Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, Ilya Sutskever, et al. 2019. Language models are unsupervised multitask learners. *OpenAI blog*, 1(8):9.

Nils Reimers and Iryna Gurevych. 2019. Sentence-bert: Sentence embeddings using siamese bert-networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 3982–3992.

John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. 2017. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*.

Abigail See, Stephen Roller, Douwe Kiela, and Jason Weston. 2019. What makes a good conversation? how controllable attributes affect human judgments. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 1702–1723.

Yusuxke Shibata, Takuya Kida, Shuichi Fukamachi, Masayuki Takeda, Ayumi Shinohara, Takeshi Shinohara, and Setsuo Arikawa. 1999. Byte pair encoding: A text compression scheme that accelerates pattern matching.

Robyn Speer, Joshua Chin, and Catherine Havasi. 2017. Conceptnet 5.5: An open multilingual graph of general knowledge. In *Thirty-first AAAI conference on artificial intelligence*.

Taffee T Tanimoto. 1958. Elementary mathematical theory of classification and prediction.

Johanne R Trippas, Damiano Spina, Paul Thomas, Mark Sanderson, Hideo Joho, and Lawrence Cavedon. 2020. Towards a model for spoken conversational search. *Information Processing & Management*, 57(2):102162.

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.

Ben Wang and Aran Komatsuzaki. 2021. GPT-J-6B: A 6 Billion Parameter Autoregressive Language Model. `https://github.com/kingoflolz/mesh-transformer-jax`.

Xiaoyang Wang, Chen Li, Jianqiao Zhao, and Dong Yu. 2021. Naturalconv: A chinese dialogue dataset towards multi-turn topic-driven conversation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pages 14006–14014.

Sean Welleck, Jason Weston, Arthur Szlam, and Kyunghyun Cho. 2019. Dialogue natural language inference. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 3731–3741.

Terry Winograd. 1977. A framework for understanding discourse. *Cognitive processes in comprehension*, 63:88.

Thomas Wolf, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, Rémi Louf, Morgan Funtowicz, et al. 2019. Huggingface's transformers: State-of-the-art natural language processing. *arXiv preprint arXiv:1910.03771*.

Qingyang Wu, Yichi Zhang, Yu Li, and Zhou Yu. 2021. Alternating recurrent dialog model with large-scale pre-trained language models. In *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume*, pages 1292–1301.

Wenquan Wu, Zhen Guo, Xiangyang Zhou, Hua Wu, Xiyuan Zhang, Rongzhong Lian, and Haifeng Wang. 2019. Proactive human-machine conversation with explicit conversation goals. *arXiv preprint arXiv:1906.05572*.

Michihiro Yasunaga, Hongyu Ren, Antoine Bosselut, Percy Liang, and Jure Leskovec. 2021. Qa-gnn: Reasoning with language models and knowledge graphs for question answering. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 535–546.

Dian Yu, Kai Sun, Claire Cardie, and Dong Yu. 2020. Dialogue-based relation extraction. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 4927–4940.

Houyu Zhang, Zhenghao Liu, Chenyan Xiong, and Zhiyuan Liu. 2020. Grounded conversation generation as guided traverses in commonsense knowledge graphs. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 2031–2043.

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 *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 2204–2213.

Hao Zhou, Minlie Huang, Tianyang Zhang, Xiaoyan Zhu, and Bing Liu. 2018. Emotional chatting machine: Emotional conversation generation with internal and external memory. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32.

Hao Zhou, Chujie Zheng, Kaili Huang, Minlie Huang, and Xiaoyan Zhu. 2020. Kdconv: A chinese multi-domain dialogue dataset towards multi-turn knowledge-driven conversation. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 7098–7108.

A Appendix

A.1 Implementation Details

A.1.1 Data Creation Models Details

All the models were implemented using PyTorch (Paszke et al., 2017). The BERT model trained on the Persona-NLI dataset was implemented using the transformers library (Wolf et al., 2019). The BERT model was initialized with the weights of 'bert-base-uncased'. The model was trained with an initial learning rate of 1e-4 with a linear schedule and a warmup (Vaswani et al., 2017), using the Adam Optimizer (Kingma and Ba, 2015). Mini-batches of size 8 were used during training.

A.1.2 Values for utterance selection constants:

The values of α, β, and γ were determined empirically. Initially, all the values were set to 0.33. Then to ensure a balance between concept-consistency, persona-entailment, and context-relevance, these values are varied only in the range of 0.30 to 0.40. It was found that a higher value of α (i.e. of context-relevance) resulted in more human-like utterances by avoiding abrupt transitions. The value of γ lower than 0.32 resulted in generating a higher number of concept-agnostic utterances. In this manner, the final values of α=0.38, β=0.30, and γ=0.32 are obtained.

A.1.3 Maximum Likelihood Estimation loss based Conversational Model

To train MLE loss based dialogue model, two pre-trained GPT-2-medium models (Radford et al., 2019) with 345M parameters are used to model the persuadee's and persuader's utterances. Here, pre-trained GPT-2-medium model consists of 24-layers, 1024 hidden units and 16 heads. We use Byte-Pair Encoding (Shibata et al., 1999) to tokenize the words. The dialogue model is trained with a learning rate = 3e-5, using AdamW optimizer (Kingma and Ba, 2015) with 100 warm-up steps and dropout rate of 0.1.

A.1.4 Specifications of Computational Resource

To train the transformer based NLI model, MLE-loss based conversational model and proposed **RPTCS**, following configurations are used:

- **GPU:** A100-PCIE-40GB.
- **CUDA Support:** CUDA 11.x (or later.
- **Memory clock:** 1215 MHz.
- **Total board power:** 250 W.
- **GPU clocks:** Base: 765 MHz, Boost: 1410 MHz.
- **Memory Size:** 40 GB.
- **Memory Type:** HBM2.
- **Bus Width:** 5120 bits.

A.1.5 Model Run time Specifications

RPTCS takes approximately 3 mins/epoch to train the model, hence for 30 epochs, it took 90 minutes to train the model. Further, if we try to perform validation along with training, considering three candidate responses per utterance per dialogue, **RPTCS** takes approximately 30 mins/epoch, hence, total time it took for 900 minutes (15 hours) to train and validate the model. Finally, to evaluate the model, the testing of proposed system takes approximately 5 minutes for 200 utterances.

REWARD WEIGHT OPTIMIZATION				
δ_1	δ_2	δ_3	δ_4	PPL
0.4	0.4	0.1	0.1	6.53
0.1	0.1	0.4	0.4	6.82
0.2	0.2	0.3	0.3	6.64
0.2	0.3	0.3	0.2	6.44
0.3	0.2	0.3	0.2	6.30
0.3	**0.3**	**0.2**	**0.2**	**6.14**

Table 5: Weight Optimisation using different values of δ.

A.2 Data Creation

In total 22 concepts were used as targets for obtaining the path. The complete list of the concepts is as follows: (i). travel, (ii). journey, (iii). voyage, (iv). outing, (v). restaurant, (vi). hotel, (vii). inn, (viii). diner, (ix). cafe, (x). canteen, (xi). bar, (xiii). shopping, (xiv). mall, (xv). market, (xvi). grocery, (xvii). electronics, (xviii). mobile, (xix). laptop, (xx). computer, (xxi). smartphone, and (xxii). camera.

We use six-shot prompts to generate the synthetic data. We provide a sample of the one-shot version of the prompt below. At the end of the sequence, we also append the persona, path, and utterances for which the next utterance needs to be generated. The six-shot prompt follows the same pattern with six examples in the input sequence. We restrict our prompt to six shots since the maximum sequence length allowed in GPT-J is 2,048:

<startdial>
<persona> i enjoy sprinting and long races. | i am at the end of my career. | i recently started a position helping others with daily challenges | i used to be very unhealthy. </persona>
<transitions> retirement -> retreat -> travel </transitions>
<topic> topic remain | retirement </topic> user: i am thinking about my upcoming retirement .
<topic> topic transition | retirement -> retreat </topic> agent: you should go on a retreat, it will help clear your mind and keep you healthy .
<topic> topic remain | retreat </topic> user: i've been thinking about it, but i am a little short on money .
<topic> topic transition | retreat -> travel </topic> agent: maybe you could go on a trip to a country with low cost of living.
</enddial>

<startdial>
<persona> like to tinker with machines. | i had a run-in with the law | i enjoy seeing nature. | i am not an honest person. </persona>
<transitions> scene -> photo -> camera </transitions>
<topic> topic remain | scene </topic> user: the sunset makes for a great scenery .
<topic> topic transition | scene -> photo </topic> agent: