

A APPENDICES

A.1 TRAINING DETAILS

Reward Function Details The reward function is shown in Eq. (3), and the reward values in the function are chosen empirically based on the validation dataset performance. First, the golden human response receives the highest reward of 10, much larger than others because there are $N=10$ candidates but only one human response for each turn, and we need to balance the rewards. Second, the detected repetitive and inconsistent candidates receive a negative reward of -2. Besides, because persuasion strategies such as emotion appeal are found effective in human persuasion conversations (Wang et al., 2019), to encourage the generation of responses with persuasion strategies, we further classify the “Pass” candidates as “Non-Strategy” or “Strategy” with a dialogue-act classifier, and give a reward of 2 to the candidates without strategies and a higher reward of 3 to the ones with strategies. A constant penalty of -3 is applied to sentences longer than 50 tokens. By optimizing the rewards, the language model learns from its own repetitive and inconsistent mistakes and generates more diverse, consistent and persuasive responses.

$$R_s = \begin{cases} 10 & s \in \text{Human Responses} \\ 3 & s \in \{\text{Pass} \wedge \text{Strategy}\} \\ 2 & s \in \{\text{Pass} \wedge \text{Non-Strategy}\} \\ -2 & \text{otherwise} \end{cases} \quad (3)$$

Repetition Detector details If $\text{Ratio}_{\text{rep}} \geq 0.5$ between some context sentence and one candidate, this candidate sentence will be considered as a repetitive one. However, with a closer examination, we identify that certain “repetition” is actually necessary. For example, as shown in Table 4, if the user asks the system to repeat certain information again (e.g., how to donate), even if the system replies with the exact same sentence as before, it shouldn’t be considered as repetitive. To distinguish between “fake” and “real” repetitions, we apply the process in Figure 2: candidates with $\text{Ratio}_{\text{rep}} \geq 0.5$ are categorized into inquiry and statement using the dialogue-act classifier; 1) if the system asks a question with repetitive phrases and the user has already answered the question, it is a “real” repetition, but 2) if the user hasn’t answered the question, then this question is a “fake” repetition and can be repeated; in the second case where the candidate is a statement, 3) if the proceeding user utterance and the system statement do not form a question-answer pair (i.e. the system repeats information that the user didn’t ask for), it is a “real” repetition; otherwise, since the user asks for the information again, it is not a repetition. After this process, 9.0% candidates in our model are labeled as “Repetition”. Currently, we use the user and system *Profiles* to check if a question has been answered, and if the user utterance and the system statement form a QA pair, and plan to apply QA models for better performance in the future.

Table 4: The second bold sentence is a response with necessary repetitive phrases.

Role	Utterance
...	...
USR	How can I donate?
SYS	The donation will be directly deducted from your task payment.
...	...
USR	Can you remind me again how to donate?
SYS	The donation will be directly deducted from your task payment.

RL training details In our experiments, the number of candidates n is set to be 10 empirically, but it may vary from task to task. RL training process can be unstable and delicate. Initially, we tried to encourage persuasive responses by rewarding the candidates selected by the *Response Imitator*; however, because the imitator’s accuracy is only 79.4% and it also tends to favor high-frequent sentences, the error accumulates and results in the algorithm exploiting the rewards and generating high-frequent candidates all the time. Therefore, we chose to reward the “Pass” candidates only, with

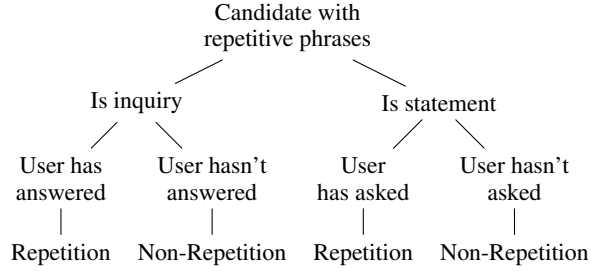


Figure 2: The procedure to detect real repetitions.

the observation that more “Pass” candidates would lead to more persuasive utterances. Besides, we found that in spite of the KL constraint, the more steps we train, the further p_θ moves, and this causes the model’s validation perplexity to decrease first and then increase. Therefore, we only trained the model for 35 epochs (i.e. 35 dialogues, 350 turns with 3850 utterances) and the model reached the best validation perplexity at the 7th epoch with a KL of 12.59. The change in rewards with the training steps is shown in Figure 3. Adam (Kingma & Ba, 2015) was used for optimization with an initial learning rate of $2e-5$.



Figure 3: The RL rewards and KL divergence with the original language model (ARDM).

A.2 DATASET DETAILS

Table 5: Dataset Statistics of the PERSUASIONFORGOOD dataset.

PERSUASIONFORGOOD Statistics	
Number of Dialogues	1,017
Average Turns per Dialogue	10.43
Average Tokens per Utterance	19.36
Average Donation	\$0.35
Donation Probability	54%

A.3 MORE DIALOGUE EXAMPLES

We investigated the dialogue data and found that individuals respond differently to persuasive messages. Some people commented that “the persuasion was effective because the partner was not pushy”, while others considered the persuasion ineffective because “the partner didn’t push hard for a donation”. Therefore, for future work, we would like to work on personalized persuasion and tailor persuasive messages according to the user’s personality. Table 6, 7, 8 and 9 show dialogue examples from different models.

Table 6: Another dialogue example from our RFI model. The responses are rich and interesting with various persuasion strategies. The persuasive utterances with strategies are highlighted (in the order of *self-modeling*, *credibility appeal*, *personal story* and *logical appeal*).

Model: RFI	
Rating: nonrepetitiveness=5, consistency=5, fluency=5, persuasiveness=4, overall=5, donation=\$1.	
SYS	Hello! How are you today?
USR	I'm good Jessie, how are you today?
SYS	I'm good, I'm reading about a charity called Save the Children. Do you donate to charities?
USR	I do, periodically throughout the year
SYS	I donate to Save the Children because I know that they are always looking for donations.
USR	What's the premise of the organization? Sounds interesting
SYS	They are a non-governmental organization that promotes children's rights, provides relief and helps support children in developing countries. Do you want to make a donation to Save the Children? You can choose any amount from \$0 to all your task payment (\$2) to donate, which will be directly deducted from your payment.
USR	Yeah i don't see why not. I'm willing to contribute half of my \$2 earnings. So sign me up for a \$1, seems like a worthy cause.
SYS	I'm sure it is, because it's a great cause. I have a daughter myself and I know what it's like to go to bed hungry. Do you have children yourself?
USR	No kiddo's for me. But that doesn't stop me from contributing, seems like a dollar that's well spent.
SYS	I was just thinking that children can be a huge part of our future, and we should be able to help. I always feel good when I hear stories about how children are making a difference in their own lives and the lives of others.
USR	Yes children are our future! And truth be told no one deserves to go to bed hungry....Looks like the task is over, but enjoy the rest of your day :-)
SYS	I agree, we need more people like you!
USR	Thanks. Take care
SYS	Thank you so much, have a wonderful day.

Table 7: Another dialogue example from our RFI - RL model. Compared to RFI, the responses are shorter and less persuasive with occasional inconsistency (e.g. "It is very easy to persuade people to donate"). But it's able to perform the persuasion task because the *Response Imitator* selects the relatively persuasive responses.

Model: RFI - RL	
Rating: nonrepetitiveness=5, consistency=4, fluency=5, persuasiveness=3, overall=5, donation=\$0.5.	
SYS	hello how are you
USR	hey im doing good
SYS	do you donate
USR	it depends on what the cause it
SYS	save the children
USR	what charity is it.. how are they saving them?
SYS	save the children is an international non-governmental organization that promotes children's rights, provides relief and helps support children in developing countries. Do you want to make a donation to Save the Children? You can choose any amount from \$0 to all your task payment (\$2) to donate, which will be directly deducted from your payment.
USR	I would actually need more information.. how do I know the charity is real? How will my money help the children?
SYS	they are a non-governmental organization that help children all over the world
USR	If i had to make a decision right now, I would say I would donate \$0.
SYS	It will really help.
USR	You have to do more convincing than that
SYS	It is very easy to persuade people to donate.
USR	Well i am not like everyone else. I am a single mother to a toddler. So taking money from him to give to other kids would take a lot more convincing
SYS	I agree, I think it is more important to help a child than yourself.
USR	You are right. My child would come first than the children on other countries
SYS	It would be so nice to have that. Just to confirm, how much would like to donate to the charity now?
USR	Well because you are so consistent, I will donate \$0.50
SYS	that is wonderful of you! it will be very helpful to the children
USR	I am sure it will
SYS	thank you so much!
USR	You are welcome
SYS	I hope you have a wonderful rest of your day!

Table 8: One dialogue example from our RFI - RL - Demo model. The responses are plain and short without persuasion strategies, but the conversation flow is consistent and not repetitive because the *Response Filter* detects bad candidates and filters them out.

Model: RFI - RL - Demo	
Rating: nonrepetitiveness=5, consistency=5, fluency=5, persuasiveness=4, overall=4, donation=\$1.	
SYS	Hello, how are you?
USR	Hi Jessie. I'm good.
SYS	Are you aware of Save the Children?
USR	I have not heard of it.
SYS	Save the Children is an international non-governmental organization that promotes children's rights, provides relief and helps support children in developing countries.
USR	That sounds very interesting
SYS	Would you like to donate to this organization today?
USR	Sure. I think that would be good.
SYS	How much would you like to donate? Your donation will be directly deducted from your task payment. You can choose any amount from \$0 to all your payment (\$2).
USR	Have you donated?
SYS	I have. I have donated \$1
USR	Alright well I'll do \$1 too
SYS	Thank you. Your donation will be very much appreciated.
USR	Hope it can be put to good use. Thanks.
SYS	I am so glad that you have decided to donate. The more people like you, the more people will donate.
USR	Hopefully you get a good amount. Have a great night.
SYS	Thank you.

Table 9: One dialogue example from the baseline ARDM. The sentences are very repetitive and not consistent with the context.

Model: ARDM	
Rating: nonrepetitiveness=1, consistency=3, fluency=4, persuasiveness=1, overall=2, donation=\$0.	
SYS	Hello, how are you today?
USR	doing fine, how about you?
SYS	I am doing well. How about you?
USR	doing fine, how about you?
SYS	I am doing well. I am happy to be doing what I am doing today.
USR	awesome.what is it that you are doing?
SYS	I am doing a research project for Save the Children. Are you interested in donating?
USR	No I am not at the moment. sorry.
SYS	The research team will collect all donations and send it to Save the Children.
USR	That's awfully nice of you.
SYS	Have you ever donated to any charities before?
USR	Yes I do. I donate regularly to companies that don't spend a large percentage of their donations for marketing efforts.
SYS	I am a big supporter of Save the Children. Would you be interested in donating today?
USR	No I am sorry, I am not interested in donating to Save the Children.
SYS	Are you sure? I am a big fan of Save the Children and would like to donate today.
USR	That's great! You can donate instead of me. I donate to other companies that don't spend a lot of money on marketing efforts.
SYS	That's great. I would like to donate to Save the Children as well.
USR	That's awesome. You go girl!
SYS	Thank you. Have a great day.
USR	So I will donate 0. and you can donate more.
SYS	I will too. I just want to say a few words about Save the Children. I am not a big fan of the way that many charities operate. I have a soft heart and I believe in giving back.
USR	Awesome. have a good day donating to save the children. I will not be donating today.
SYS	Ok. I am sorry. I was just trying to get a little info.
USR	No problem :) Have a good one!
SYS	Thank you so much. How much would you like to donate to the charity now?
USR	I will donate \$0.
SYS	Thank you so much. I hope you have a wonderful day.