

Will I Sound Like Me? Improving Persona Consistency in Dialogues through Pragmatic Self-Consciousness

Hyunwoo Kim Byeongchang Kim Gunhee Kim

Department of Computer Science and Engineering

Seoul National University, Seoul, Korea

{hyunw.kim, byeongchang.kim}@vl.snu.ac.kr gunhee@snu.ac.kr

<https://vl.snu.ac.kr/projects/consistency>

Abstract

We explore the task of improving persona consistency of dialogue agents. Recent models tackling consistency often train with additional Natural Language Inference (NLI) labels or attach trained extra modules to the generative agent for maintaining consistency. However, such additional labels and training can be demanding. Also, we find even the best-performing persona-based agents are insensitive to contradictory words. Inspired by social cognition and pragmatics, we endow existing dialogue agents with *public self-consciousness* on the fly through an imaginary listener. Our approach, based on the Rational Speech Acts framework (Frank and Goodman, 2012), can enforce dialogue agents to refrain from uttering contradiction. We further extend the framework by learning the distractor selection, which has been usually done manually or randomly. Results on Dialogue NLI (Welleck et al., 2019) and PersonaChat (Zhang et al., 2018) dataset show that our approach reduces contradiction and improves consistency of existing dialogue models. Moreover, we show that it can be generalized to improve context-consistency beyond persona in dialogues.

1 Introduction

In the study of dialogue agents, *consistency* has been a long-standing issue. To resolve this, much research has been conducted to endow dialogue agents with *personas*. Li et al. (2016) propose to encode persona in embeddings and Zhang et al. (2018) introduce a persona-conditioned dialogue dataset. On top of these works, many efforts have been made to improve consistency.

In spite of such recent significant progress, there is much room for improving persona-based dialogue agents. We observe that even the best performing persona-based generative models (See et al., 2019; Wolf et al., 2019b; Roller et al., 2020)

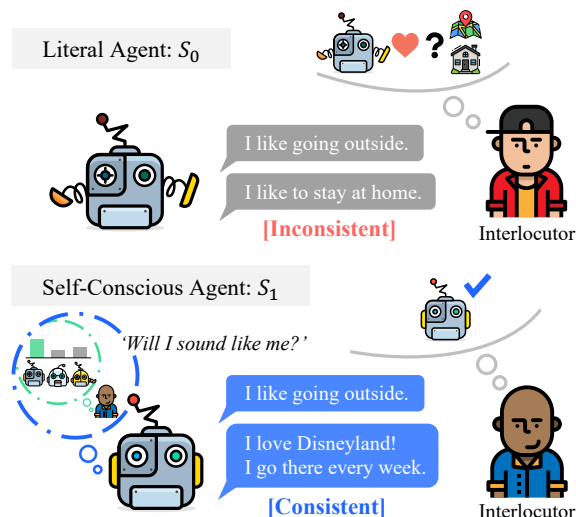


Figure 1: Illustration of the consistency issue in dialogue. While a literal dialogue agent (S_0) fails to deliver a consistent persona, our self-conscious agent (S_1) does so, by modeling an imaginary listener. Icons are designed by Nhor Phai and Vincent Le Moign.

are highly insensitive to contradictory words, and thus fail to deliver consistent persona to the interlocutor (Figure 1). Also, extra modules other than the generative model is often required for improving consistency. Recent works on consistency in persona-based dialogue actively adopt the NLI-based approach (Welleck et al., 2019; Song et al., 2019; Li et al., 2020; Song et al., 2020), which have the following prerequisites. First, they require labeled pairs of persona sentences and dialogue utterances with three categories: entailment, neutral, and contradiction. Next, methods with NLI models for rating the agent’s consistency also need to train them separately with those labels.

In this work, we step back from this NLI-based supervised approach and ponder: *how do humans maintain consistency?* We humans never learn how to be consistent. Instead, we have an innate drive for consistency to hold our beliefs and behavior in harmony (Festinger, 1962). If so, how do we

know we are consistent or not? We do not ask others. We ask ourselves by predicting how we are perceived by others. *Public self-consciousness* is this awareness of the self as a social object that can be observed and evaluated by others (Fenigstein et al., 1975). We particularly emphasize that public self-consciousness is not equivalent to the philosophical self-consciousness (or self-awareness)¹. Simply put, public self-consciousness is the concern about how oneself will be perceived by others, as opposed to the philosophical state of being conscious of self-existence.

According to Doherty and Schlenker (1991), people with high public self-consciousness tend to act more consistent with known information about themselves. They care deeply about how others will evaluate them and have a strong tendency to avoid negative evaluations (Fenigstein et al., 1975). Since inconsistency is condemned by others, one who has high public self-consciousness will try more to maintain consistency. In order to predict how we are perceived, we rely on abstract models of others (Gopnik and Wellman, 1992) and simulate others' reactions based on imagination (Hassabis et al., 2013). Inspired by this, our intuition is that self-consciousness through an imaginary listener will let dialogue agents better maintain consistency.

Modeling a listener has been one of the main topics in computational pragmatics. Our work extends this long line of work in cognitive science by making use of the Bayesian Rational Speech Acts framework (Frank and Goodman, 2012), which has been originally applied to improving informativeness of referring expressions. Since personas ought to express who we are, we adopt this framework for dialogue agents by regarding personas as targets that should be conveyed to the interlocutor. As the agent tries to generate tokens that help the imaginary listener identify the agent's persona, it can lastly generate more consistent utterances.

In summary, we take inspiration from social cognition and pragmatics to endow generative agents with self-consciousness, which makes them imagine the listener's reaction and incorporate it to the generation process for improving consistency. Our major contributions can be outlined as follows:

(1) We propose an orthogonally applicable approach for any persona-based generative agents to improve consistency without the use of additional

consistency labels and training. Moreover, it is even generalizable to improve context-consistency beyond persona in dialogue.

(2) We extend the Rational Speech Acts framework (Frank and Goodman, 2012) with two new technical features: (i) a learning method for distractor selection (*e.g.* other samples different from the given target (Andreas and Klein, 2016)), which has been usually done manually or randomly, and (ii) a different update for the listener's world prior that better preserves information of previous states.

(3) Our approach improves consistency of three recent generative agents (See et al., 2019; Wolf et al., 2019b; Roller et al., 2020) over Dialogue NLI (Welleck et al., 2019) and PersonaChat (Zhang et al., 2018). Along with large reduction in contradiction, the utterance accuracy significantly increases too.

2 Related Work

Persona & Consistency in Dialogue. Li et al. (2016) learn personas in embeddings. Zhang et al. (2018) release the *PersonaChat* dataset, a chitchat dialogue set involving two interlocutors each playing their given persona. Madotto et al. (2019) use meta-learning to adapt to new personas with few dialogue samples. Liu et al. (2020) use reinforcement learning to enhance mutual persona perception.

Recent works use extra modules or NLI labels to improve consistency. Shum et al. (2019) fill generated templates, and rank with a language model. Zhang et al. (2019) use self-supervised feature extractors for generation. Welleck et al. (2019) annotate NLI labels to the PersonaChat dataset. They train an NLI model and run pairwise comparison between candidates and persona to compute contradiction scores. The NLI approach is applied for coherence evaluation (Dziri et al., 2019), rewards to reinforcement learning agents (Song et al., 2019), finding inconsistent words (Song et al., 2020), and unlikelihood training (Li et al., 2020). They require NLI labels on the target dialogue dataset; otherwise, sharp decrease in performance is observed, due to mismatch of data distribution (Welleck et al., 2019). Such dataset-specific NLI annotations and training NLI models can be costly and time-consuming.

Compared to previous methods, the novelty of our approach is to improve consistency without NLI labels and extra modules.

Pragmatics. Our approach belongs to the general family of Bayesian Rational Speech Acts

¹<https://plato.stanford.edu/entries/self-consciousness/>

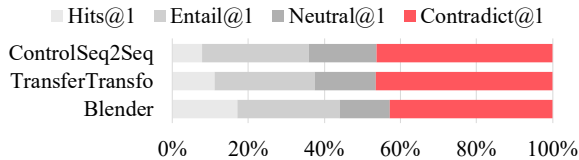


Figure 2: Proportion of Hits@1, Entail@1, Neutral@1 and Contradict@1 in the top-1 candidates returned by the models on the Dialogue NLI dataset.

	ROUGE-1	ROUGE-L	SPICE
GT Utterance	15.7	14.6	10.6
Top Entail-Utt	15.3	14.5	7.1
Contradict@1-Utt	16.3	15.9	6.6

Table 1: Comparison between ground-truth utterances, top-ranked entailing candidates and Contradict@1 utterances in ROUGE and SPICE scores.

(RSA) frameworks (Frank and Goodman, 2012) in pragmatics. It has improved informativeness in a number of NLP tasks, including reference games (Andreas and Klein, 2016), image captioning (Mao et al., 2016; Vedantam et al., 2017; Cohn-Gordon et al., 2018), instruction following (Fried et al., 2017), navigating (Fried et al., 2018), translation (Cohn-Gordon and Goodman, 2019), summarization (Shen et al., 2019) and referring expression generation (Zarrieß and Schlangen, 2019).

However, its application to the dialogue domain remains understudied. In this work, we explore how the RSA framework can be adopted in dialogue agents to alleviate the inconsistency problem. Also, we further extend the framework by making the distractor selection as a learnable process.

3 Insensitivity to Contradictory Words in Existing Persona-based Agents

Although conditional language generation has shown promising progress, maintaining consistency within the generation yet remains unsolved. From quantitative evaluation, we reveal existing generative models for dialogues are highly insensitive to contradictory words.

Dialogue NLI Evaluation. Welleck et al. (2019) introduce the Dialogue NLI dataset based on the PersonaChat dataset (Zhang et al., 2018). They collect entailing and contradictory utterances to the given persona, and release an evaluation set comprised of dialogues each with 31 utterance candidates: 10 entailing, 10 neutral, and 10 contradictory utterances with 1 ground-truth (GT) utterance. On this evaluation set, we run three recent models (See et al., 2019; Wolf et al., 2019b; Roller

Persona	I love wearing skinny jeans and shirts. I am a blonde girl with short hair.
GT Utterance	(I, 1.87) (have, 51.42) (really, 201.45) (short, 1.78) (hair, 1.30) (and, 2.81) (it, 45.25) (is, 2.19) (blonde , 461.60).
Contradict@1-Utt	(What, 60.89) (color, 103.11) (is, 1.99) (your, 1.06) (hair, 1.05) (?, 1.11) (Mine, 3.57) (is, 1.03) (brown , 17.25).

Table 2: Example of a contradictory utterance returned by the model and its GT utterance with perplexity per token. The words of entailment and contradiction to the persona are shown in blue and red, respectively.

et al., 2020) that achieve the best performance on PersonaChat. We report four ranking metrics following Welleck et al. (2019): Hits@1, Entail@1, Neutral@1 and Contradict@1. Each metric is the proportion of GT, entailing, neutral and contradictory utterances in the top-1 candidates returned by the model, respectively. The models rank the candidates by perplexity scores.

Figure 2 shows that all three models select contradictory candidates much more often than the GT utterances (see further results in Table 3). Though models are conditioned on a given persona, they are highly insensitive to contradictions.

3.1 Analysis of Contradict@1 Utterances

To investigate why insensitivity to contradiction prevails in the state-of-the-art models, we further analyze the contradictory utterances returned by the models (Contradict@1-Utt), comparing with the GT utterances and the top-ranked entailing candidates (Top Entail-Utt). Table 1 reports language metrics between the selected candidates and the given persona sentences using SPICE (Anderson et al., 2016) and ROUGE (Lin, 2004). SPICE metric measures semantic similarity and ROUGE metric measures n -gram overlaps between two sentences. Contradict@1-Utt shows *lower* SPICE scores and *higher* ROUGE scores than other utterances, implying that it may be different in semantics but similar in syntax to the given persona.

To take a closer look, we extract the contradicting words from Contradict@1-Utt and their counterparts from GT utterances to compare their average perplexity scores. In the Dialogue NLI dataset, every utterance is labeled with a triple ($entity_1, relation, entity_2$), such as “I just like to listen to rock music” with ($i, like_music, rock$).

By construction, `Contradict@1-Utt` must contain words that are contradictory to the GT utterance and the given persona. The perplexity scores of contradictory words (106.7) were considerably lower than those of the counterparts in GT utterances (280.1). Table 2 shows an example of such dialogue instance with perplexity per word. If properly conditioned with the given persona, models should show lower perplexity for the words in the persona. However, their perplexity scores are significantly higher than those of contradictory words. It reveals that models behave more as a plain language model rather than as a persona-conditioned model. Thus, guarantee of consistency for each word generation step is required for persona-based dialogue agents to resolve such issue.

4 Approach

We introduce how to endow dialogue agents with public self-consciousness, which helps them keep consistency in mind at each generation step by reflecting an imaginary listener’s distribution over personas. Since the imaginary listener arises from the plain dialogue-agent, separate training is not needed. Figure 3 illustrates its overall structure.

We present how to model public self-consciousness using the Rational Speech Acts (RSA) framework (Frank and Goodman, 2012) in Section 4.1. We then discuss learning of distractor selection as our major novelty for the RSA in Section 4.2.

4.1 Modeling the Public Self-Consciousness

We seek to build a dialogue agent who is self-conscious about its consistency without the need for training on NLI labels or rating consistency with NLI models. Given that modeling the interactions between listener and speaker is a main topic in pragmatics, we take advantage of the RSA framework (Frank and Goodman, 2012). It treats language use as a recursive process where probabilistic speaker and listener reason about each other’s intentions in a Bayesian fashion. To apply the framework to sequence generation for dialogues, we extend the incremental approach proposed for image captioning (Cohn-Gordon et al., 2018).

To generate an utterance, the agent computes the distribution of every next token u_t at timestep t in Bayesian fashion as follows.

Base Speaker S_0 . We first assume persona i is given to the base speaker, along with the dialogue

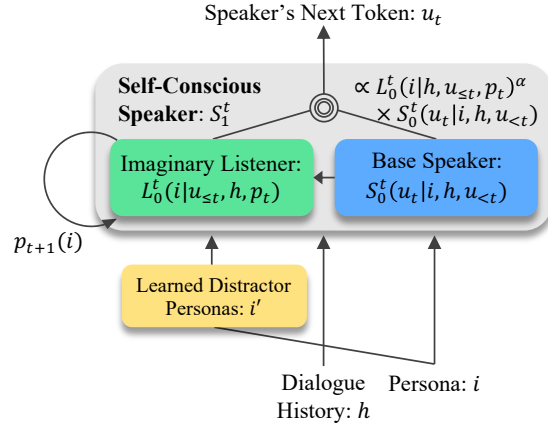


Figure 3: The proposed self-conscious agent S_1 consists of base speaker S_0 and imaginary listener L_0 . It recursively generates the next token u_t at every time t .

history h and partial utterance $u_{<t}$, as shown in Figure 3. The base speaker S_0^t returns a distribution over the next token at timestep t : $S_0^t(u_t|i, h, u_{<t})$. Any conditional dialogue agent can be used as a base speaker. See the details in Section 5.2.

Imaginary Listener L_0 . While the base speaker generates each token one at a time, the imaginary listener reasons about the speaker’s persona. The imaginary listener L_0^t is the posterior distribution of the speaker’s persona in terms of the base speaker and the world prior $p_t(i)$ over personas as follows,

$$L_0^t(i|h, u_{\le t}, p_t) \propto \frac{S_0^t(u_t|i, h, u_{<t})^\beta \times p_t(i)}{\sum_{i' \in \mathcal{I}} S_0^t(u_t|i', h, u_{<t})^\beta \times p_t(i')}. \quad (1)$$

where β on S_0^t is the listener rationality coefficient that controls the amount of information from the current timestep compared to the cumulative prior $p_t(i)$. L_0 returns a probability distribution over the personas in world \mathcal{I} , which is a finite set ($|\mathcal{I}| = 3$) comprising the given persona i and distractor personas. The distractors are different personas from other dialogue instances in the dataset. We decide world \mathcal{I} per dialogue instance through learning, which will be elaborated in Section 4.2.

Self-Conscious Speaker S_1 . With S_0^t and L_0^t , the self-conscious speaker S_1^t is defined as

$$S_1^t(u_t|i, h, u_{<t}) \propto L_0^t(i|h, u_{\le t}, p_t)^\alpha \times S_0^t(u_t|i, h, u_{<t}), \quad (2)$$

where α is the speaker rationality coefficient that determines how much the likelihood is considered. By taking the listener’s distribution into account, the speaker is now self-conscious about what persona it sounds like. Especially, the agent seeks

to be perceived as the given persona i rather than some other persona i' . The likelihood of each token being identified as the persona i acts as a bonus added to the base speaker’s token scores. Hence, tokens that are consistent to the given persona are preferred to others. The token with the highest probability is added to the partial utterance, becoming the next input $u_{<t+1}$ for the speaker.

Updating the world prior with L_0 . Starting from a uniform distribution as the initial prior $p_0(i)$, we update the world prior $p_{t+1}(i)$ according to S_1 ’s output u_t at every time step:

$$p_{t+1}(i) = L_0^t(i|h, u_{\leq t}, p_t). \quad (3)$$

Hence, $p_t(i)$ represents the cumulative state of the partial utterance up to t . Cohn-Gordon et al. (2018) report the prior update with $L_1 \propto S_0^t(u_t|i, h, u_{<t}) \times L_0^t(i|h, u_{\leq t}, p_t)$ makes little practical effect compared to a uniform prior. We find that updating the prior with Eq. (3) instead is effective. See the results in Section 5.6.

4.2 Learning to Select Distractors

Distractors (Andreas and Klein, 2016) are samples (e.g. other personas in the dataset) which are different from the given target. In previous works of RSA, the distractors to be included in world \mathcal{I} are selected manually or randomly from the dataset. However, we find that performance variance is large according to the selected distractors. We thus propose to learn distractor selection, especially based on the life-long memory network (Kaiser et al., 2017). The life-long memory network is capable of implicitly clustering similar dialogue contexts into a few slots with associated persona. Therefore, it can efficiently memorize and retrieve distractor personas for each context. In Appendix, we experiment that our approach outperforms other models including BERT-based algorithms.

To better select useful distractor personas, supervised learning is desirable. However, there is no explicit label indicating which distractors are helpful for each dialogue. We select the persona that have the best Hits@1 as the distractor label per training dialogue. The Hits@1 is the score for favoring the ground-truth next utterance (consistent and context-relevant) over other candidate utterances which are just being consistent (i.e. entailing) or contradictory to the given persona. In other words, the score represents consistency and also appropriateness at the same time. Thus, such

distractors can help the self-conscious agent to generate responses which are context-relevant and allow the imaginary listener to identify the speaker’s persona. Each training datapoint comprises a given persona, a distractor persona and dialogue context.

Memory Structure. The memory consists of three types of information: $M = (\mathbf{K}, \mathbf{v}, \mathbf{a})$. $\mathbf{K} \in \mathbb{R}^{m \times d}$ is a key matrix, where m is the number of memory slots and d is the dimension of the key vectors, which are the embedding of datapoints. The value vector $\mathbf{v} \in \mathbb{R}^m$ stores the index of a persona. $\mathbf{a} \in \mathbb{R}^m$ is an age vector, which is used for memory update. We set $m = 16,000$ and $d = 768$.

Memory Addressing. We construct the query vector \mathbf{q} for each datapoint with the *BERT-Uncased-Base* (Devlin et al., 2019) model. We use the output embedding of BERT’s [CLS] token, and normalize it to a unit length to build $\mathbf{q} \in \mathbb{R}^d$.

Using the cosine similarity between \mathbf{q} and each memory key, we can find the k nearest neighbors:

$$(n_1, n_2, \dots, n_k) = NN_k(\mathbf{q}, \mathbf{K}). \quad (4)$$

Memory Loss. Suppose that the query datapoint has a distractor label l . Among (n_1, \dots, n_k) , we denote the positive neighbor n_p as the one with $\mathbf{v}[n_p] = l$ and the negative neighbor n_b with $\mathbf{v}[n_b] \neq l$. If there are multiple positive neighbors, we pick the one with the smallest memory index. If no positive neighbor is found, we select a random key whose value is l . For the negative neighbor, we select one randomly from (n_1, \dots, n_k) . We set $k = 2048$. Then, the loss is computed as

$$\mathcal{L} = \max(\mathbf{q} \cdot \mathbf{K}[n_b] - \mathbf{q} \cdot \mathbf{K}[n_p] + \alpha, 0), \quad (5)$$

where α is a positive margin, which we set as 0.2. This loss maximizes the cosine similarity between the query \mathbf{q} and the positive key $\mathbf{K}[n_p]$, while minimizing the similarity to the negative key $\mathbf{K}[n_b]$. We finetune the query network BERT with this loss.

Memory Update. After computing the loss, memory M is updated differently for two cases. (1) If the top-1 neighbor’s value (i.e. persona) is correct ($\mathbf{v}[n_1] = l$), the key vector is updated as:

$$K[n_1] \leftarrow \frac{\mathbf{q} + K[n_1]}{\|\mathbf{q} + K[n_1]\|}. \quad (6)$$

(2) Otherwise ($\mathbf{v}[n_1] \neq l$), we make a slot for the query; we find the oldest memory slot n' according to the age vector \mathbf{a} and write

$$K[n'] \leftarrow \mathbf{q}, \quad \mathbf{v}[n'] \leftarrow l, \quad \mathbf{a}[n'] \leftarrow 0. \quad (7)$$

Training & Inference. In our *Distractor Memory* network, training corresponds to updating the memory and the parameters of the query network.

At inference, given a test example, we obtain the query by encoding the dialogue context and the persona using BERT. We find n nearest keys from the memory, and use their values (*i.e.* persona indices) as the distractor personas. We set $n = 2$.

5 Experiments

We show that our self-conscious framework can significantly improve consistency and accuracy of state-of-the-art persona-based agents on two benchmark datasets. We prove its effectiveness using both automatic and human evaluations. We also show our framework can be generalized to improve consistency of dialogue context beyond persona.

5.1 Datasets

Dialogue NLI Evaluation Set (Welleck et al., 2019). This dataset is based on PersonaChat with additional NLI annotations. Its main task is to rank next-utterance candidates given previous context. For each dialogue, they collect 31 next-utterance candidates in respect to the given persona: 10 entailing, 10 neutral and 10 contradicting candidates with 1 ground-truth utterance. In total, the evaluation set includes 542 instances.

PersonaChat dialogue (Zhang et al., 2018). This dataset involves two interlocutors who are each given a persona and asked to get to know each other while playing their roles. This task was the subject of the ConvAI2 competition (Dinan et al., 2019) at NeurIPS 2018. The competition version contains 17,878 chitchat conversations conditioned on 1,155 personas for training and 1,000 conversations conditioned on 100 personas for validation.

5.2 Experimental Setting

Base Speakers. We experiment on three pretrained models including ControlSeq2Seq (See et al., 2019), TransferTransfo (Wolf et al., 2019b), and Blender (Roller et al., 2020) as base speakers (S_0) for our self-conscious agents (S_1). The ControlSeq2Seq is a Seq2Seq model with attention trained on Twitter dataset (Miller et al., 2017) and finetuned on PersonaChat. TransferTransfo based on GPT (Radford et al., 2018) is the winner of the ConvAI2 competition in automatic evaluation. Blender, a recently released generative dialogue model, is the state-of-the-art open-domain chatbot. Our approach improves these base speakers by

Model	Hits@1 \uparrow	Entail@1 \uparrow	Contradict@1 \downarrow
ControlSeq2Seq (See et al., 2019)			
S_0	7.9	27.9	46.3
S_1	10.5	36.4	34.0
S_1 +DM	13.1	40.8	24.5
TransferTransfo (Wolf et al., 2019b)			
S_0	11.1	26.4	46.5
S_1	17.5	40.4	29.7
S_1 +DM	18.8	45.8	19.7
Blender (Roller et al., 2020)			
S_0	18.8	27.3	42.4
S_1	21.8	38.0	30.6
S_1 +DM	22.5	44.1	19.6

Table 3: Comparison of our approach (S_1) with base speakers (S_0) on the Dialogue NLI evaluation set (Welleck et al., 2019). +DM is the *Distractor Memory*. High scores in Hits@1, Entail@1 and low scores in Contradict@1 imply better consistency.

granting them the sense of self-consciousness. We defer implementation details to Appendix.

Evaluation Metrics. For Dialogue NLI, we report three ranking metrics introduced in the original paper: Hits@1, Entail@1, and Contradict@1. Each metric is the proportion of GT, entailing, and contradictory utterances in the top-1 candidates returned by the model, respectively. High scores in Entail@1 and low scores in Contradict@1 indicate better consistency with the persona.

For PersonaChat, we report Hits@1, standard F1 score, perplexity and C score, following the ConvAI2 protocol. Hits@1 is the accuracy of choosing the ground-truth next-utterance among 20 candidates as the models rank the candidates by perplexity. The C score is a metric for dialogue consistency, introduced in Madotto et al. (2019). It computes pairwise comparison between utterance u and persona sentence p_j with a pretrained NLI model. The NLI model returns 1, 0, -1 for entailment, neutrality, and contradiction, respectively. We sum the NLI scores across persona sentences per dialogue instance: $C(u) = \sum_j \text{NLI}(u, p_j)$.

5.3 Quantitative Results

Results on Dialogue NLI. Table 3 compares the performance of dialogue agents on the Dialogue NLI evaluation set. Our self-conscious agent S_1 significantly reduces Contradict@1 scores and increases the Entail@1 along with the Hits@1 accuracy of the literal agents S_0 . We remind that each entailing candidate shares the same annotated triple as the GT utterance. In other words, they have similar semantics to the GT utterance and follow the

Model	Hits@1 \uparrow	F1 \uparrow	Perplexity \downarrow	C \uparrow
ControlSeq2Seq (See et al., 2019)				
S_0	16.1	17.0	22.9	0.45
S_1	16.4	16.9	23.9	0.54
S_1 +DM	16.7	17.1	23.9	0.55
TransferTransfo (Wolf et al., 2019b)				
S_0	16.2	19.2	17.6	0.86
S_1	17.5	19.4	19.1	0.96
S_1 +DM	18.2	19.5	19.1	0.97
Blender (Roller et al., 2020)				
S_0	27.6	19.5	12.0	0.85
S_1	28.8	19.7	13.2	0.93
S_1 +DM	29.1	19.8	13.2	0.95

Table 4: Comparison of our approach (S_1) with base speakers (S_0) on PersonaChat (Zhang et al., 2018). C is the consistency score evaluated by a pretrained NLI model (Madotto et al., 2019). For TransferTransfo, we use the generative version to calculate Hits@1.

given persona. Thus, Entail@1 is a lenient version of Hits@1 (Welleck et al., 2019). The *Distractor Memory* (DM) is better than random distractor selection for S_1 across all metrics. It concludes that learned distractors are more effective than random distractors for pragmatic agents.

Results on PersonaChat. Table 4 compares the performance of different dialogue agents on the PersonaChat dataset. Our model S_1 outperforms all other generative dialogue agents in terms of consistency related metrics, *i.e.* Hits@1 and C score. Since the posterior update of our self-conscious agent revises the distribution learned by the base speaker, the increase in perplexity is natural due to the effect of regularization. Nevertheless, our approach improves the F1 score for TransferTransfo and Blender. Thus, being consistent to the given persona can also help improve the generation performance of dialogue agents.

Comparison with agents that use NLI model. We also test agents with pretrained NLI models attached (Welleck et al., 2019), denoted by +NLI in Table 5. The NLI model computes contradiction scores of each candidate utterances, and penalize its rank accordingly. Compared to base agents with no self-consciousness, our agents improve consistency in all three metrics even further when using additional NLI models. Another notable result is that our agents without NLI (S_1 +DM in Table 3) for ControlSeq2Seq and TransferTransfo even outperform the base agents with NLI (S_0 +NLI) on Hits@1. That is, our self-conscious agents achieve

Model	Hits@1 \uparrow	Entail@1 \uparrow	Contradict@1 \downarrow
ControlSeq2Seq (See et al., 2019)			
S_0 +NLI	12.7	48.2	8.1
$[S_1$ +DM]+NLI	14.4	51.7	7.0
TransferTransfo (Wolf et al., 2019b)			
S_0 +NLI	17.2	44.4	9.8
$[S_1$ +DM]+NLI	21.4	54.6	5.4
Blender (Roller et al., 2020)			
S_0 +NLI	24.9	44.7	6.0
$[S_1$ +DM]+NLI	26.6	52.0	5.7

Table 5: Comparison of our approach (S_1) with base speakers (S_0) on the Dialogue NLI evaluation set (Welleck et al., 2019) with pretrained NLI model attached.

Model	Raw		Calibrated	
	Consistent	Engaging	Consistent	Engaging
TransferTransfo (Wolf et al., 2019b)				
S_0	0.53 (0.02)	2.48 (0.03)	0.44 (0.01)	2.48 (0.01)
S_1 +DM	0.61 (0.02)	2.55 (0.03)	0.52 (0.01)	2.52 (0.01)

Table 6: Human evaluation results comparing the consistency and engagingness of the base speaker (S_0) and our self-conscious agent (S_1). Numbers in parentheses are the standard errors.

better GT accuracy even without the help of an NLI model trained on consistency labels.

5.4 Human Evaluation

We perform human evaluation via Amazon Mechanical Turk. We random sample 250 test examples, each is rated by three unique human judges in terms of (i) *Consistency* and (ii) *Engagingness*. Turkers are shown a given persona, a dialogue context, and the model’s generated utterance. For consistency, we follow Madotto et al. (2019) and ask judges to assign 1, 0, -1 to the utterance for consistency, neutrality, and contradiction, respectively. Following See et al. (2019), we evaluate the engagingness of the utterance in a 4-point scale, where higher scores are better. To alleviate annotator bias and inter-annotator variability, we apply Bayesian calibration (Kulikov et al., 2019) to the scores.

Table 6 summarizes the human evaluation results. The agent with our self-consciousness method S_1 is rated as more consistent than the base agent S_0 while maintaining a similar level of engagingness. While it can be trivial to increase consistency at the cost of engagingness (*e.g.* perfect consistency can be generated by boring utterances with very little variance), it is not the case for our agent. Since

Model	Hits@1 ↑	Entail@1 ↑	Contradict@1 ↓
Dialogue NLI (Welleck et al., 2019)			
S_0	18.8	27.3	42.4
S_1 (on context)	32.7	27.7	26.4

Model	Hits@1 ↑	F1 ↑	Perplexity ↓	C ↑
PersonaChat (Zhang et al., 2018)				
S_0	27.6	19.5	12.0	0.57
S_1 (on context)	30.5	19.9	13.5	0.58
EmpatheticDialogue (Rashkin et al., 2019)				
S_0	32.6	20.5	14.7	0.47
S_1 (on context)	34.2	20.6	15.4	0.50

Table 7: Comparison of our approach (S_1) with base speaker Blender (S_0) when conditioned on dialogue context in three datasets. We compute the consistency score C respect to the dialogue context.

our agent seeks to be heard as the given persona to the listener, self-distinctive words tend to meld into generated responses (see Figure 6). Thus, the responses from self-conscious agents have their own color, which can help improving engagingness.

Figure 4 displays selected examples of utterance generation. Each example is comprised of dialogue history, human response, and utterances generated by our method and baselines.

5.5 Consistency for Dialogue Context

We demonstrate that our self-conscious agent can be generalized to generate context-consistent utterances beyond persona. We condition the agent with its previous responses in the dialogue history; that is, i in Eq. (2) is the agent’s past responses instead of persona sentences. Hence, tokens that are inconsistent to the agent’s past response would be less favored by the model.

Table 7 reports the results of context conditioned self-conscious agents. The EmpatheticDialogue (Rashkin et al., 2019) is an open-domain dialogue dataset where a speaker describes a past emotional experience and the listener responds accordingly. Since the speaker’s descriptions should be consistent to the experience and previous utterances, it is a suitable benchmark for consistency. We model the speaker’s utterances and measure its consistency.

Our S_1 agent outperforms other literal agents on all three datasets in terms of consistency. Thus, our approach can also be applied to help agents stay more consistent to its context.

P1’s Persona	
I own a house in Florida.	
I work in it and have been at the same company for 15 years.	
I enjoy American sports	
I’ve a children and a dogs.	
Dialogue History	
⋮	
[P2] Hello, how are you today?	
[P1] Hey! Just watching a game of football with my children. You?	
[P2] That’s cool! I am an alcoholic who is recovering.	
(S_1+DM) i work in a company and i’m a workaholic.	
(S_0) i’m sorry to hear that. I’m a retired professional athlete.	
(Human) ah man congrats for trying to get back on the road !	
P1’s Persona	
My family does not support my career choices.	
My dream car is a Rolls Royce ghost.	
I often shop for shoes and watches.	
I like shopping.	
I currently work for a watch dealer.	
Dialogue History	
⋮	
[P1] I really enjoy shopping and my dream is to one day own a Rolls Royce ghost.	
[P2] Wow. I enjoy running over driving.	
[P1] Running is also quite lovely. Breathing in the lovely outside air.	
[P2] Yes it is. It clears my head when I need to as well.	
(S_1+DM) shopping is a great way to clear my head.	
(S_0) i love to shop and watch movies.	
(Human) yes , and it also helps with depression i have found.	

Figure 4: Examples of generated responses by our self-conscious agent with *Distractor Memory* (S_1 +DM) on the PersonaChat dataset (Zhang et al., 2018). We compare it with the base speaker (S_0) of TransferTransfo (Wolf et al., 2019b) and the human response (Human).

5.6 Controlling the Self-Conscious Agent

To further analyze our self-conscious agent, we conduct experiments by controlling three features

of our agent: world prior updates $p_t(i)$, listener rationality β and speaker rationality α .

World Prior Update. In the self-conscious agent, the world prior acts as a cumulative state over personas. We remind that we propose to update the world prior with L_0^t instead of L_1^t in Eq. (3). As reported in Cohn-Gordon et al. (2018), our experiments on the Dialogue NLI dataset confirm the prior update with L_1^t makes little difference in performance compared with using a uniform distribution. However, our approach with L_0^t makes significant difference, as shown in Figure 5. The reason is that the pragmatic listener $L_1^t \propto S_0^t(u_t|i, h, u_{<t}) \times L_0^t(i|h, u_{\leq t}, p_t)$ reflects the *current* S_0^t twice (*i.e.* in L_0^t and in itself) per time step. Hence, the update with L_1^t becomes more of an instantaneous prior rather than a cumulative one. On the other hand, L_0^t moderately combines the information from both S_0^t and $p_t(i)$, preserving better cumulative information.

Listener Rationality β . We add β in L_0^t to control the amount of information incorporated to the world prior $p_t(i)$. Figure 5 depicts that when β is large, the Hits@1 scores (*i.e.* the GT accuracy) drop. With a big β , the information S_0^t at current time step overrides the cumulative prior $p_t(i)$. That is, the utterance state evolves shortsightedly, ignoring the context information from the previous steps. Therefore, setting of $\beta \leq 1$ is advantageous for the self-conscious agent to incrementally decode.

Speaker Rationality α . Figure 6 shows an example of how generated responses vary according to the intensity of speaker rationality α . As α increases, the self-conscious agent reflects the listener’s distribution (*i.e.* the likelihood) more into the posterior. When α is too large, the posterior distribution is overwhelmed by the likelihood of the persona. Then, the language model degenerates to favor uttering fragments of the given persona while even ignoring the syntax. Hence, α can control the degree of copying the given condition text. An appropriate α value allows the given persona condition to blend smoothly in the utterance.

6 Conclusion

This work investigated how modeling public self-consciousness can help dialogue agents improve persona-consistency. We showed existing dialogue agents are highly insensitive to contradiction, and introduced an orthogonally applicable method using the RSA framework (Frank and Goodman, 2012) to alleviate the issue. We also designed a

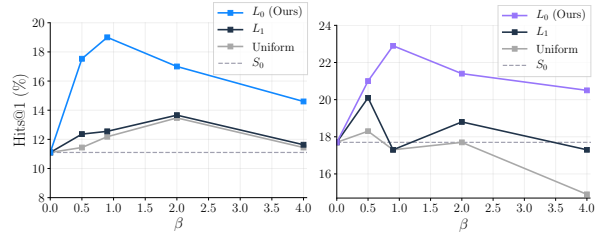


Figure 5: Performance variation of the self-conscious agent for TransferTransfo (left) and Blender (right) according to β . We compare different methods of updating the world prior $p_t(i)$ with L_0 (Ours), L_1 and a uniform prior. The dashed line is the base speaker S_0 .

Persona	Utterance
	I've 5 cats. I am a construction worker. My cats are very special to me. I enjoy building houses.
($\alpha = 0$)	i'm a construction worker. i'm going to be a vet.
($\alpha = 2$)	i work construction. i'm a construction worker.
($\alpha = 8$)	construction work is great. i build houses for my cats.
($\alpha = 10$)	construction workers earn 5 cats so building houses affords us special pets. yours? kittens! d ou

Figure 6: An example of utterance changes by controlling the speaker rationality α on the PersonaChat.

learning method for distractor selection, named *Distractor Memory* and proposed a better update for the listener’s world prior. Furthermore, we demonstrated how our approach can be generalized to improve dialogue context-consistency. Our self-conscious agents improved the base agents on the Dialogue NLI (Welleck et al., 2019) and PersonaChat (Zhang et al., 2018) dataset, without consistency labels and NLI models. An important future direction will be generating the distractors and learning the rationality coefficients.

Acknowledgements

We would like to thank Reuben Cohn-Gordon, Sean Welleck, Junhyug Noh and Jiwan Chung for their valuable comments. We also thank the anonymous reviewers for their thoughtful suggestions on this work. This research was supported by Brain Research Program by National Research Foundation of Korea (NRF) (2017M3C7A1047860), Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No. 2017-0-01772, Video Turing Test, No. 2019-0-01082, SW StarLab), and Creative Pioneering Researchers Program through Seoul National University. Gunhee Kim is the corresponding author.

References

- Peter Anderson, Basura Fernando, Mark Johnson, and Stephen Gould. 2016. SPICE: Semantic Propositional Image Caption Evaluation. In *ECCV*, pages 382–398. Springer.
- Jacob Andreas and Dan Klein. 2016. Reasoning about Pragmatics with Neural Listeners and Speakers. In *EMNLP*.
- Reuben Cohn-Gordon and Noah Goodman. 2019. Lost in Machine Translation: A Method to Reduce Meaning Loss. In *NAACL-HLT*.
- Reuben Cohn-Gordon, Noah Goodman, and Christopher Potts. 2018. Pragmatically Informative Image Captioning With Character-level Inference. In *NAACL-HLT*.
- Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *NAACL-HLT*.
- Emily Dinan, Varvara Logacheva, Valentin Malykh, Alexander Miller, Kurt Shuster, Jack Urbanek, Douwe Kiela, Arthur Szlam, Iulian Serban, Ryan Lowe, et al. 2019. The Second Conversational Intelligence Challenge (ConvAI2). *arXiv:1902.00098*.
- Kevin Doherty and Barry R Schlenker. 1991. Self-Consciousness and Strategic Self-Presentation. *Journal of Personality*, 59(1):1–18.
- Nouha Dziri, Ehsan Kamaloo, Kory W Mathewson, and Osmar Zaiane. 2019. Evaluating Coherence in Dialogue Systems Using Entailment. In *NAACL-HLT*.
- Allan Fenigstein, Michael F Scheier, and Arnold H Buss. 1975. Public and Private Self-Consciousness: Assessment and Theory. *Journal of Consulting and Clinical Psychology*, 43(4):522.
- Leon Festinger. 1962. *A Theory of Cognitive Dissonance*, volume 2. Stanford University Press.
- Michael C Frank and Noah D Goodman. 2012. Predicting Pragmatic Reasoning in Language Games. *Science*, 336(6084):998–998.
- Daniel Fried, Jacob Andreas, and Dan Klein. 2017. Unified Pragmatic Models for Generating and Following Instructions. In *NAACL-HLT*.
- Daniel Fried, Ronghang Hu, Volkan Cirik, Anna Rohrbach, Jacob Andreas, Louis-Philippe Morency, Taylor Berg-Kirkpatrick, Kate Saenko, Dan Klein, and Trevor Darrell. 2018. Speaker-follower models for vision-and-language navigation. In *NeurIPS*.
- Alison Gopnik and Henry M Wellman. 1992. Why the Child’s Theory of Mind Really is a Theory. *Mind & Language*, 7(1-2):145–171.
- Demis Hassabis, R Nathan Spreng, Andrei A Rusu, Clifford A Robbins, Raymond A Mar, and Daniel L Schacter. 2013. Imagine All the People: How the Brain Creates and Uses Personality Models to Predict Behavior. *Cerebral Cortex*, 24(8):1979–1987.
- Łukasz Kaiser, Ofir Nachum, Aurko Roy, and Samy Bengio. 2017. Learning to Remember Rare Events. In *ICLR*.
- Diederik Kingma and Jimmy Ba. 2015. Adam: A Method for Stochastic Optimization. In *ICLR*.
- Iliia Kulikov, Alexander Miller, Kyunghyun Cho, and Jason Weston. 2019. Importance of Search and Evaluation Strategies in Neural Dialogue Modeling. In *INLG*.
- Jiwei Li, Michel Galley, Chris Brockett, Georgios P Spithourakis, Jianfeng Gao, and Bill Dolan. 2016. A Persona-Based Neural Conversation Model. In *ACL*.
- Margaret Li, Stephen Roller, Iliia 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*.
- Chin-Yew Lin. 2004. Rouge: A Package for Automatic Evaluation of Summaries. In *Text Summarization Branches Out*, pages 74–81.
- Qian Liu, Yihong Chen, Bei Chen, Jian-Guang Lou, Zixuan Chen, Bin Zhou, and Dongmei Zhang. 2020. You Impress Me: Dialogue Generation via Mutual Persona Perception. In *ACL*.
- Andrea Madotto, Zhaojiang Lin, Chien-Sheng Wu, and Pascale Fung. 2019. Personalizing Dialogue Agents via Meta-Learning. In *ACL*.
- Junhua Mao, Jonathan Huang, Alexander Toshev, Oana Camburu, Alan L Yuille, and Kevin Murphy. 2016. Generation and Comprehension of Unambiguous Object Descriptions. In *CVPR*.
- A. H. Miller, W. Feng, A. Fisch, J. Lu, D. Batra, A. Bordes, D. Parikh, and J. Weston. 2017. ParlAI: A Dialogue Research Software Platform. *arXiv:1705.06476*.
- Alec Radford, Karthik Narasimhan, Tim Salimans, and Ilya Sutskever. 2018. Improving Language Understanding with Unsupervised Learning. Technical report, Technical report, OpenAI.
- Hannah Rashkin, Eric Michael Smith, Margaret Li, and Y-Lan Boureau. 2019. Towards Empathetic Open-domain Conversation Models: A New Benchmark and Dataset. In *ACL*.
- 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:2004.13637*.

- Abigail See, Stephen Roller, Douwe Kiela, and Jason Weston. 2019. What Makes a Good Conversation? How Controllable Attributes Affect Human Judgments. In *NAACL-HLT*.
- Sheng Shen, Daniel Fried, Jacob Andreas, and Dan Klein. 2019. Pragmatically Informative Text Generation. In *NAACL-HLT*.
- Michael Shum, Stephan Zheng, Wojciech Kryściński, Caiming Xiong, and Richard Socher. 2019. Sketch-Fill-AR: A Persona-Grounded Chit-Chat Generation Framework. *arXiv:1910.13008*.
- Haoyu Song, Yan Wang, Wei-Nan Zhang, Xiaojiang Liu, and Ting Liu. 2020. Generate, Delete and Rewrite: A Three-Stage Framework for Improving Persona Consistency of Dialogue Generation. In *ACL*.
- Haoyu Song, Wei-Nan Zhang, Jingwen Hu, and Tiu Liu. 2019. Generating Persona Consistent Dialogues by Exploiting Natural Language Inference. *arXiv:1911.05889*.
- Ramakrishna Vedantam, Samy Bengio, Kevin Murphy, Devi Parikh, and Gal Chechik. 2017. Context-Aware Captions from Context-Agnostic Supervision. In *CVPR*.
- Sean Welleck, Jason Weston, Arthur Szlam, and Kyunghyun Cho. 2019. Dialogue Natural Language Inference. In *ACL*.
- Thomas Wolf, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, Rémi Louf, Morgan Funtowicz, et al. 2019a. Transformers: State-of-the-art Natural Language Processing. *arXiv:1910.03771*.
- Thomas Wolf, Victor Sanh, Julien Chaumond, and Clement Delangue. 2019b. TransferTransfo: A Transfer Learning Approach for Neural Network based Conversational Agents. *arXiv:1901.08149*.
- Sina Zarrieß and David Schlangen. 2019. Know What You Don’t Know: Modeling a Pragmatic Speaker that Refers to Objects of Unknown Categories. In *ACL*.
- 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*.
- Yizhe Zhang, Xiang Gao, Sungjin Lee, Chris Brockett, Michel Galley, Jianfeng Gao, and Bill Dolan. 2019. Consistent Dialogue Generation with Self-supervised Feature Learning. *arXiv:1903.05759*.

A Results on Variants of Distractor Selection (Section 4.2)

Model	Hits@1 ↑	Entail@1 ↑	Contradict@1 ↓
ControlSeq2Seq (See et al., 2019)			
Random	8.5	32.8	37.6
Nearest	7.6	32.8	36.5
Farthest	9.4	33.6	35.4
BERT-Classifier	9.2	33.6	35.6
BERT-Ranker	9.6	33.3	35.1
DM	11.1	36.0	28.2

Table 8: Quantitative results of the proposed *Distractor Memory* (DM) and other distractor selection methods on the Dialogue NLI evaluation set (Welleck et al., 2019).

We compare our proposed *Distractor Memory* (DM) with three heuristic methods, and two variants of the pretrained BERT model (Devlin et al., 2019). As a straightforward baseline, we randomly select k personas from training set and directly use it as distractors. Second, we test the k -nearest search by speaker’s persona, denoted by Nearest; for a given persona descriptions, we find its closest training persona embedding using cosine similarity on average pooled BERT features. The third baseline denoted by Farthest is to find the k -farthest persona among the training personas.

We also compare with two variants of the BERT model. The first variant is BERT-Classifier, which takes dialogue context as input and returns the index of persona from training set as output. The second variant is bi-encoder ranking model of Miller et al. (2017), denoted by BERT-Ranker. It encodes dialogue context and candidate persona with separate BERT encoders measuring its ranking with cosine similarity. For both methods, we use top- k ranked personas as distractors and set $k = 4$ for all the methods. We use Adam optimizer (Kingma and Ba, 2015) with learning rate $2e-5$ and finetune *BERT-Uncased-Base* up to 3 epochs.

Table 8 compares the performance of different distractor selecting methods on the Dialogue NLI evaluation set (Welleck et al., 2019). We set $\alpha = 8$, $\beta = 0.5$, and $|\mathcal{I}| = 5$. The DM model outperforms all the baselines across all metrics. The Farthest shows better performance than the Nearest. It can be understood that dissimilar distractors are more effective in the Rational Speech Acts framework (Frank and Goodman, 2012). The BERT-Ranker performs the best among baselines, but not as good as ours, which validates that memorization capability is effective for selecting useful distractors.

B Implementation Details

Base Codes and Datasets. We use the ParlAI framework² (Miller et al., 2017) and Hugging-Face’s Transformers³ (Wolf et al., 2019a) to implement our models and baselines. We use Dialogue NLI (Welleck et al., 2019) and PersonaChat (Zhang et al., 2018) datasets from the ParlAI framework as is. We use the default preprocessing in ParlAI.

Training. Our self-consciousness approach improves consistency for any pretrained dialogue-agents without additional consistency labels and pretrained NLI models. Since it post-processes the output probability of pretrained dialogue-agents in a Bayesian fashion, no additional model parameters are added to the dialogue agents. Thus, it does not require any training. In the case of using the Distractor Memory (DM), first we initialize *BERT-Uncased-Base* with pretrained weights and finetune it up to 3 epochs with Adam optimizer with learning rate $2e-5$. Then we find the best distractor persona for each model and use those labels to train our DM. We train our DM on one NVIDIA TITAN Xp GPU up to 7 epochs.

Hyperparameters. For Dialogue NLI evaluation, we set the speaker rationality $\alpha = 8.0$, the listener rationality $\beta = 1.0$, and the cardinality of the world \mathcal{I} to 3. In PersonaChat evaluation, we set $\alpha = 2.0$, $\beta = 0.3$ for ControlSeq2Seq (See et al., 2019), $\alpha = 2$, $\beta = 0.9$ for TransferTransfo (Wolf et al., 2019b), and $\alpha = 2.0$, $\beta = 0.5$ for Blender 90M (Roller et al., 2020). We also set $|\mathcal{I}| = 3$. We experiment $\alpha = \{1.0, 2.0, 4.0, 8.0, 16.0\}$, $\beta = \{0.3, 0.5, 0.9, 1.0, 2.0, 4.0\}$, and $|\mathcal{I}| = \{2, 3, 5\}$. We choose the hyper-parameter configuration showing the best performance in Hits@1 for Dialogue NLI and F1 score for PersonaChat. The posterior distribution of our self-conscious agents are computed deterministically. For our Distractor Memory, we set the memory key matrix as $\mathbf{K} \in \mathbb{R}^{m \times d}$, where $m = 16000$ and $d = 768$. We set the number of nearest neighbor $k = 2048$.

Inference. We use greedy decoding for all methods. The average runtime for our self-conscious approach is dependent on the base dialogue agents and the cardinality of world \mathcal{I} which can be run in parallel like beam search.

Evaluation. We follow the evaluation of the ParlAI framework. Following Madotto et al. (2019),

we use the finetuned BERT-based NLI model⁴ to compute the C score.

C Dialogue Examples

Figure 7 shows selected examples of generated responses. In each set, we show given persona, dialogue context, human responses, and generated responses by our self-conscious agent and the base speaker. We use TransferTransfo (Wolf et al., 2019b) as a base speaker.

²<https://parl.ai/>

³<https://huggingface.co/transformers/>

⁴<https://github.com/HLTCHKUST/PAML>.

P1's Persona	P1's Persona
I've 5 cats. I am a construction worker. My cats are very special to me. I enjoy building houses.	I own a house in Florida. I work in it and have been at the same company for 15 years. I enjoy American sports I've a children and a dogs.
Dialogue History	Dialogue History
⋮	⋮
[P2] It is going very great. I just have homework to do [P1] Oh what are you in school for? [P2] High school. On my sophomore year. [P1] Are you planning to go to college ? [P2] I am ! I want to be a professional soccer player.	[P2] Hello, how are you today? [P1] Hey! Just watching a game of football with my children. You? [P2] That's cool! I am an alcoholic who is recovering.
(S₁+DM) i work construction. i'm a construction worker. (S₀) i'm a construction worker. i'm going to be a vet. (Human) exciting ! i played soccer in high school too , long ago	(S₁+DM) i work in a company and i'm a workaholic. (S₀) i'm sorry to hear that. I'm a retired professional athlete. (Human) ah man congrats for trying to get back on the road !
P1's Persona	P1's Persona
I'm a huge star wars fan. I try various coffees as a hobby. I enjoy poetry. I played football for a division a college.	My family does not support my career choices. My dream car is a Rolls Royce ghost. I often shop for shoes and watches. I like shopping. I currently work for a watch dealer.
Dialogue History	Dialogue History
⋮	⋮
[P1] I'm good, taking a break from my assignments before heading to Europe. [P2] I went to Spain then, learned I love cooking paella. What team are you for? [P1] I'm with Ohio state. Born and raised in Ohio. [P2] Awesome. What do you do for a living?	[P1] I really enjoy shopping and my dream is to one day own a Rolls Royce ghost. [P2] Wow. I enjoy running over driving. [P1] Running is also quite lovely. Breathing in the lovely outside air. [P2] Yes it is. It clears my head when I need to as well.
(S₁+DM) i play football for a local college. (S₀) i'm a student. i'm a student (Human) i'm a student , going to school for veterinary medicine .	(S₁+DM) shopping is a great way to clear my head. (S₀) i love to shop and watch movies. (Human) yes , and it also helps with depression i have found.

Figure 7: Examples of generated responses by our self-conscious agent with *Distractor Memory* (S_1+DM) on the PersonaChat dataset (Zhang et al., 2018). We compare it with the base speaker (S_0) of TransferTransfo (Wolf et al., 2019b) and the human response (Human).