A Appendices

A.1 Sequence Sampling in Reinforcement Learning

The generator G transfers a source sentence X into a sentence in target style. In this work, we use beam search of width k to find a reference target sentence $Y^{\text{ref}}_{1:T'}$. In RL, we need to estimate the reward of each action y_t in the reference sentence $Y^{\text{ref}}_{1:T'}$. Fig. 3 shows the sampling and scoring process.

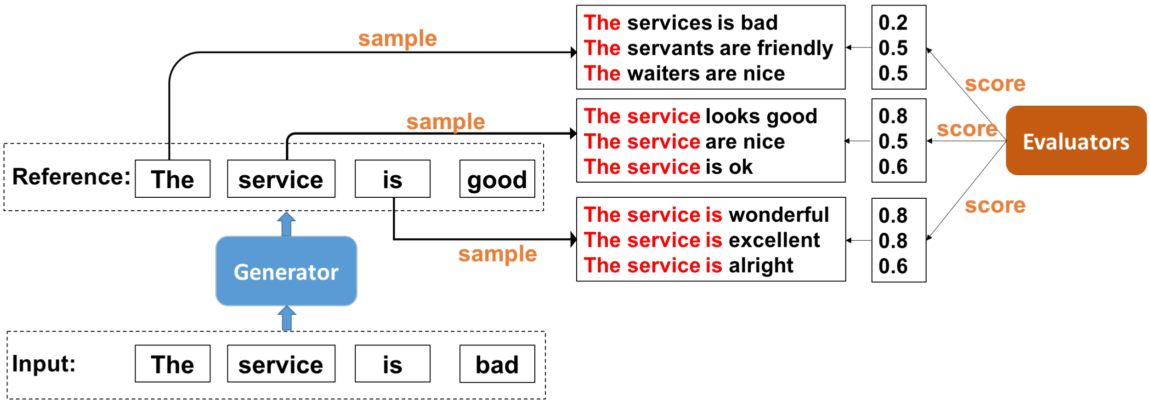

Figure 3: Sequence sampling: red words are sub-sequences in reference target sentence based on which the remaining sub-sequences are sampled. The sampled complete sentences are sent to the evaluator for scoring.

Suppose that the reference target sentence $Y^{\text{ref}}_{1:T'}$ is "The service is good". At the first time step, i.e., when $t = 1$, we start from the sub-sequence $Y^{\text{ref}}_{1:1}$, i.e., the sub-sentence "The". We use multinomial sampling to roll out "The" to complete sentences, which are "The service is bad", "The servants are friendly" and "The waiters are nice" in Fig. 3. These sampled sentences are then sent to the evaluator for scoring in terms of style, content and fluency. Their scores are 0.2, 0.5 and 0.5 respectively, and we average them as the action score $f(y_1, s_1)$ of the first word y_1 at its state s_1. The score $f(y_1, s_1) = 0.4$ is sent back to the generator, which will be used to obtain the reward $r(y_1, s_1)$ as described in Eq. 10. Similarly when $t = 2$, we sample three complete sentences based on the sub-sentence "The service": "The service looks good", "The service are nice" and "The service is ok".

A.2 Experiments

Table 5: Semantic and style scores given by our evaluator on all systems.

	Sentiment				Formality			
	Negative-to-Positive		Positive-to-Negative		Informal-to-Formal		Formal-to-Informal	
Metric	Semantic	Style	Semantic	Style	Semantic	Style	Semantic	Style
CA	**-1.293**	0.806	**-1.346**	0.818	-1.212	0.646	-1.281	0.851
MDS	-1.412	**0.855**	-1.662	0.822	-1.508	0.568	-1.445	**0.878**
RLS	-1.315	0.846	-1.458	**0.847**	**-0.935**	**0.782**	**-0.903**	0.872

Automatic evaluation metrics. We reported the automatic evaluation results of all text style transfer systems in Table 3, where we used the evaluation metrics adopted by previous works (Fu et al., 2018; Santos et al., 2018). Here we report the style and semantic scores given by the evaluator in our system in Table 5. Recall that semantic score given by our evaluator was the negative of word movers' distance between the generated sentence and the source sentence divided by the sentence length. The larger the semantic score was, the better the content was preserved in the generated sentence. As for the style evaluation, we used a bidirectional recurrent neural network as style classifier. It predicted the likelihood that an input sentence was in target style, which was taken as the *style score* of the generated sentences. Again, the larger the style score was, the better the generated sentence fitted in target style.

As shown in Table 5, the results given by the semantic and style modules of our evaluator are very similar to those given by Fu et al.. In sentiment transfer task, CA model does best in content preservation and MDS does best in transfer strength. As for FT, our model outperforms the two baselines in terms of semantic and style scores.

Table 6: Example transferred sentences of all systems.

Type	Transferred sentence
Negative-to-positive	Source: I 've noticed the food service sliding down hill quickly this year . CA: I have enjoyed the food here and this place is perfect . MDS: Food is the best staff . RLS: I 've noticed the food service was perfect this time .
	Source: The chicken tenders did n't taste like chicken , wtf ? CA: The food tastes good , just like spicy ! MDS: And the food is the food in the food in well . RLS: . The chicken were like chicken, you can find what you want .
Positive-to-negative	Source: I recommend ordering the " special chicken " really good ! CA: I would give the pizza ... how they really really good ? MDS: They are the worst customer service . RLS: I would say chicken were very bad .
	Source: My experience was brief , but very good . CA: My experience was ok , but , very good . MDS: Worst , i would never go to going back . RLS: My experience was bad .
Informal-to-formal	Source: Well that is just the way it is i guess . CA: It is the best thing i think that is not . MDS: That is for the way . RLS: It is the way I think .
	Source: Like i said he already knows that you like him , so just take a deep breathe and ask him . CA: I think that she likes you , but perhaps you will get a relationship and and ask her . MDS: If you find him and i think that you have been in a relationship . RLS: I believe he knows that you like him, so go to ask him .
Formal-to-informal	Source: Well, if you are really attracted to this guy, then smile and speak nicely to him . CA: If you to tell her the way that is you and get married . MDS: The way of guys are not if you are not . RLS: Well , if you really like this guy , then smile to him .
	Source: Men are unintelligent! What person understands the meaning behind their behavior? CA: Men are not of his meaning . MDS: Men are understands all men are not ? RLS: Men are stupid ! Why girl loves the mind ?

Examples and Analysis. We list some example transferred sentences given by our model and two baseline systems in Table 6. In the first example of negative-to-positive transfer, our model adheres to the topic of food service while baselines change to topic of food. Similarly in the first example of positive-to-negative transfer, our model preserves the topic of chicken while CA model talks about pizza and MDS model talks about customer service. Semantic similarity as explicit semantic constraints in our model is shown to be better at preserving the topic of source sentences.

There is still space to improve content preservation in all models. In the second example of informal-to-formal transfer, all transferred sentences miss the segment of "take a deep breathe" in the source sentence. In the second example of formal-to-informal transfer, the three transferred sentences miss part of source information. The source sentence is a rhetorical question, which truly means "people hardly understand the meaning behind their behavior". This is a hard example, and all models do not capture its semantic meaning accurately.

FT task is more challenging compared with ST given that the sentence structure is more complicated with a larger vocabulary in the formality dataset. Its difficulty is also reflected by the degraded transfer performance of all systems as reported in Table. 3. From the examples in Table 6, our transferred sentences are more fluent than the outputs of two baselines in FT. The language model in our system plays an important role in making the model's outputs more fluent.