



Polite Chatbot: *A Text Style Transfer Application*

Sourabrata Mukherjee, Vojtech Hudecek, Ondrej Dusek
Institute of Formal and Applied Linguistics
Charles University, Czech Republic

Introduction

Text Style Transfer(TST)



- Change style of given input text.
 - Preserve style-independent content.
 - **Style:** *demographic attrib (personality, gender), sentiment, politeness, etc.*
-

Politeness Transfer



- A sub-task of TST
 - Converts not-so-polite text to polite text,
 - without changing content.
 - **Objective:**
 - improve communication
 - build better relationships between people.
-

EXAMPLE

Shut up! the video is starting!



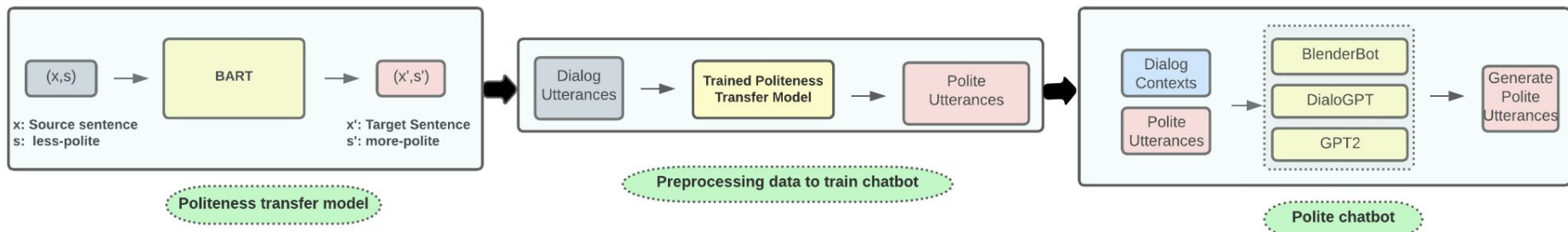
Please be quiet, the video will begin shortly.

Our Work

Overview

- *Polite chatbot*: using
 - *politeness transfer model*: generate polite synthetic dialogue pairs
 - *dialogue model*: trained on synthetic pairs
- *Evaluation*: showing
 - better politeness than baselines
 - competitive coherence to context

Proposed Architecture



- **Politeness transfer model:**
 - *politeness markers* identified and removed from polite texts
 - BART finetuned to reconstruct politeness markers
- **Synthetic polite chat data:**
 - generated by applying politeness transfer model
- **Dialogue model:**
 - BlenderBot, DialoGPT, and GPT-2 as dialog models
 - finetuned using synthetic polite data to generate polite responses

Datasets

- Politeness:
 - preprocessed & filtered Enron email sentences (Madaan et al., 2020)
 - Dialog:
 - DailyDialog (Li et al., 2017) open-domain human-human dialogues
-

Evaluation Metrics

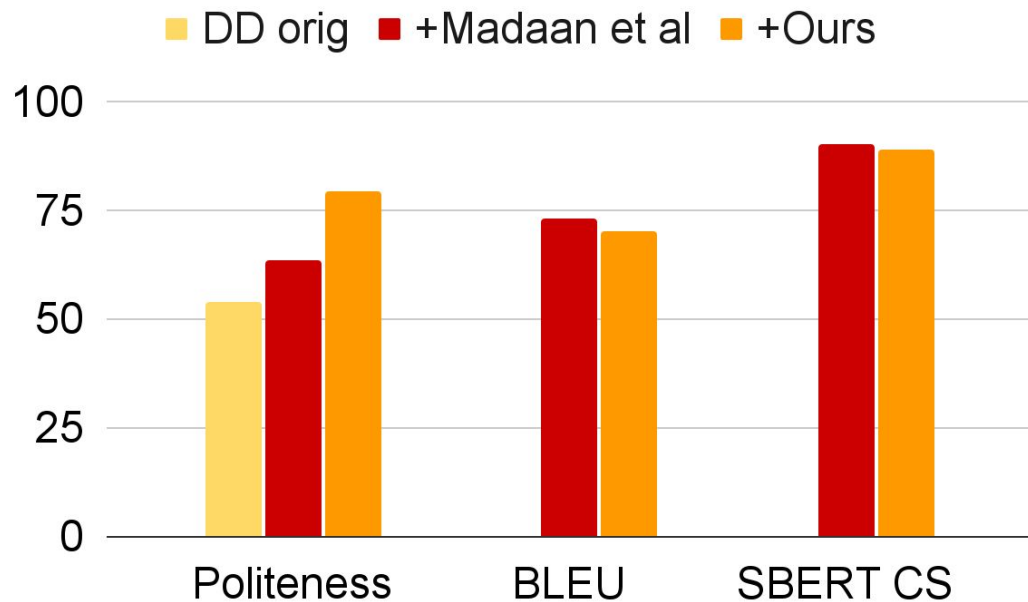
- Automatic evaluation
 - politeness: politeness multi-classifier
 - coherence: BLEU, SBERT cosine similarity
 - Human evaluation
 - politeness, coherence, fluency
 - 1-5 Likert scales
-

Results

Evaluation results of politeness transfer models

Models	PS	BLEU	CS
Madaan et al. (2020)	7.01	60.16	87.86
Ours	8.68	71.65	93.25

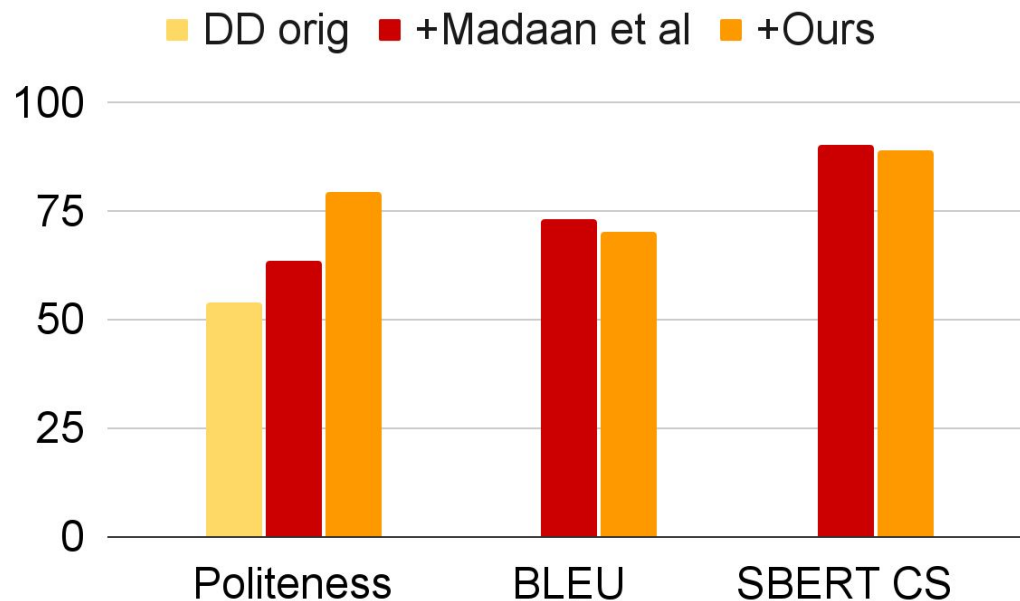
Evaluation results of politeness transfer models



Evaluation of synthetic data

Models	PS	BLEU	CS
DailyDialog (DD)	5.41	-	-
DD + Madaan et al. (2020)	6.37	73.34	90.29
DD + Ours	7.95	70.21	89.07

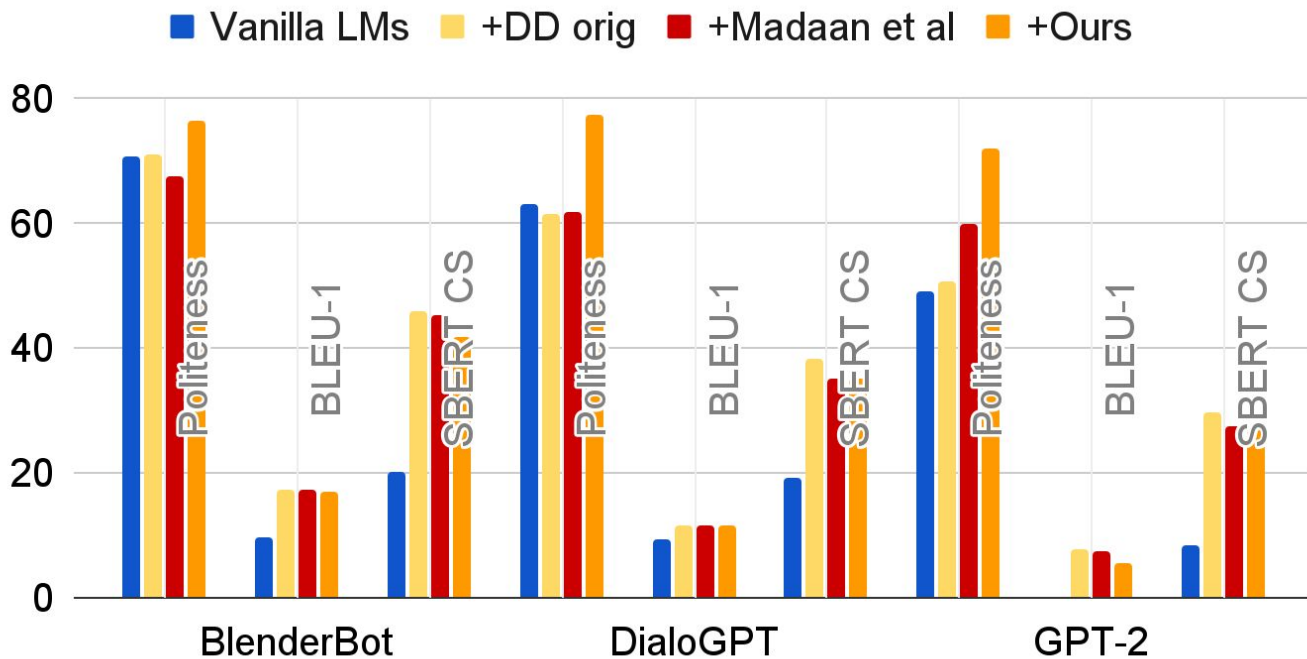
Evaluation of synthetic data



Evaluation results of polite dialog models

<i>Finetuned on</i>	BlenderBot				DialogPT				GPT-2			
	<i>PS</i>	<i>BLEU-1,2</i>		<i>CS</i>	<i>PS</i>	<i>BLEU-1,2</i>		<i>CS</i>	<i>PS</i>	<i>BLEU-1,2</i>		<i>CS</i>
Vanilla (no FT)	7.06	9.80	2.58	20.31	6.31	9.38	1.98	19.33	4.91	0.15	0.09	8.31
DailyDialog (DD)	7.11	17.2 1	7.25	45.80	6.14	11.72	2.60	38.44	5.08	7.82	2.13	29.72
DD + Madaan et al. (2020)	6.75	17.1 6	6.73	45.17	6.17	11.47	2.19	35.08	5.99	7.32	1.49	27.42
DD + Ours	7.65	17.0 3	6.85	41.80	7.75	11.44	2.57	35.03	7.20	5.65	1.03	26.80

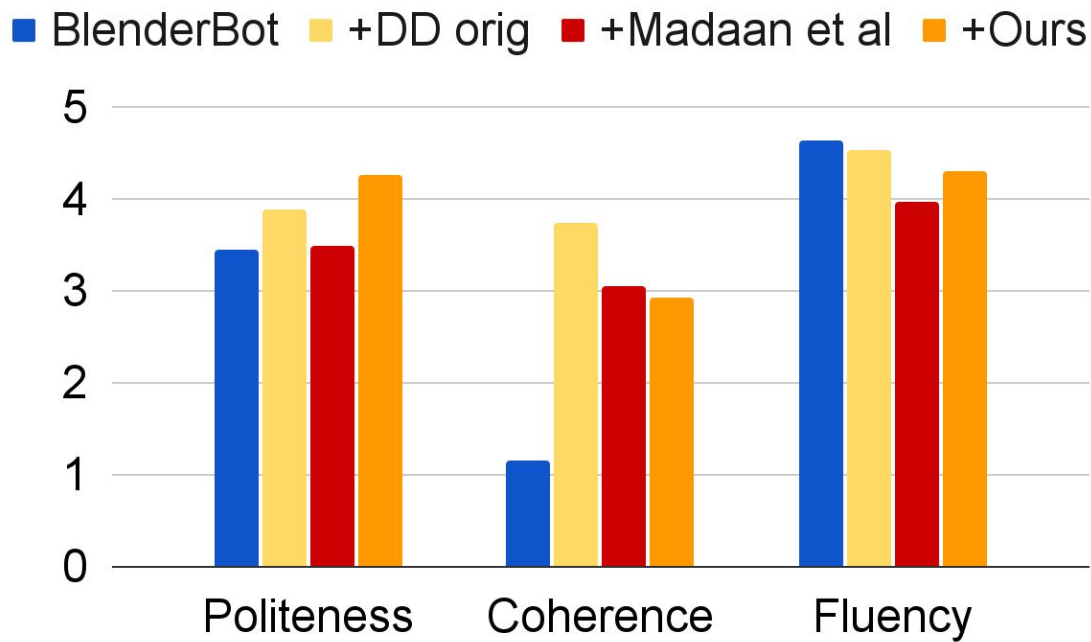
Evaluation results of polite dialog models



Human Evaluation on BlenderBot outputs

BlenderBot finetuned on	Pol	CC	Flu
Vanilla (no FT)	3.46	1.16	4.64
DailyDialog (DD)	3.90	3.74	4.54
DD + Madaan et al. (2020)	3.50	3.06	3.98
DD + Ours	4.26	2.94	4.30

Human Evaluation on BlenderBot outputs



Sample Output

Context : Nani Book Store. How can I help you? Do you have the 'The man and the Sea' by Ernest Hemingway?	
Ground Truth	Yes, we have one copy left.
Vanilla (no FT)	I would love to try it, but I don't think I'll be able to afford it.
DailyDialog (DD)	Yes, it ' s on that shelf. I ' ll get it for you right away.
DD + Madaan et al. (2020)	We do. Would you like a look at it?
DD + Ours	Yes, we do. Would you like to have a look at it? please let me know.

Conclusion

- *Text Style Transfer (TST)*: a growing research area
 - Challenges: content preservation and style transfer together, linguistic consistency, evaluation
- Our contribution:
 - A polite chatbot: a TST-based application that produces polite responses
- Future work:
 - Extend our method to other stylized response generation tasks

Thank You

mukherjee@ufal.mff.cuni.cz

 https://github.com/souro/polite_chatbot

This research was supported by Charles University projects GAUK 392221, GAUK 302120, and SWV 260575, and by the European Research Council (Grant agreement No. 101039303 NG-NLG). We would like to express our gratitude to our colleague Zdeněk Kasner for his insightful discussions and helpful feedback on this project.

