13

Development of transformer-based conversational systems

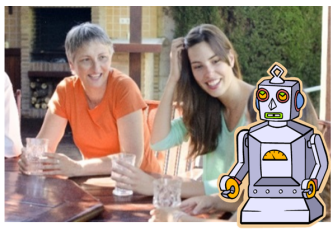
Recipes for enjoy-talking conversational systems

Abstract

We are studying a social dialogue system that satisfies people's desire for dialogue through natural conversation.
We have developed a deep-learning-based Japanese social dialogue system, which is pre-trained with the largest-scale Japanese dialogue data obtained from Twitter and fine-tuned with high-quality dialogue data that NTT has cultivated over many years of research. We also conducted a quantitative analysis of the utterances of the constructed system and identified remaining issues such as contradictions and discontinuous topics.
We believe that the desire to communicate with others is one of our fundamental desires. We aim to realize a social dialogue system as a partner that continuously satisfies this need for dialogue.

Dialogue systems chatting with people

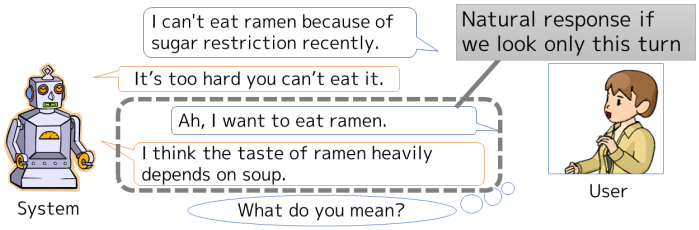
Dialogue systems are recently expected as daily conversational partners.



- **Anytime, anywhere**
  (With no limitation of time and places)
- **Relaxing, ease**
  (No need to be shy nor to worry about the other's evaluation)
- **Deep topics**
  (Easy to talk about private topics and fine-grained favorites)

Issues of conventional systems

- Even a natural utterance as a one-turn response may **become unnatural when considering the context**.
- Too many combinations of utterances complicate considering contexts without appropriately abstracting the meaning of words.

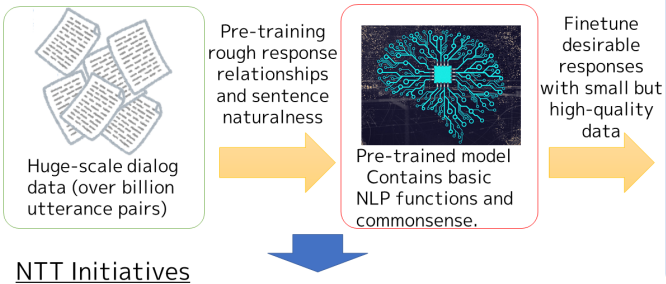


Deep-learning-based dialogue system

Deep-learning-based dialogue systems with huge-scale pre-training* are rapidly developed in English.

*Pre-training: A method to learn sentence naturalness or rough response patterns using huge data in advance. It shows a good performance in machine translation and question answering.

→It dramatically improved dialogue systems' performance by realizing highly natural sentence generation and highly accurate abstraction of the meaning of words.



Finetune with dialogue **about profile**

System	User
Hello. Thank you and best regards.	
	Nice to meet you!
I'm an esthetician. What do you do for a living?	
	I'm an office worker.
You look busy with work. What do you do on your days off?	
	I sleep on my days off.
I see. I love my pet rabbit. Do you have a pet?	
	I don't have one. The rabbit is cute!
Yes, it's soothing. Do you have any hobbies?	

Finetune with dialogue **about favotites**

System	User
Hello. Thank you and best regards, Do you have any hobbies?	
	Hello. My hobby is watching "How about Wednesday".
How about Wednesday? I love it. Which series do you like?	
	I like the confrontational archipelago.
That's a masterpiece! I like Europe revenge. Do you have a favorite episode?	
	I burst out laughing no matter how many times I watch the episode that Mr. Yasuda reverses.
Yasuda is interesting! I also like the combination with Yo Oizumi.	
	Yeah. Their talk is great.
I laughed at the laughter of Yo Oizumi.	

NTT Initiatives

- Development of the largest deep-learning-based Japanese social dialogue system
  - Model Size: 1.6 billion Parameters (3 times larger than the existing Japanese model)
  - Pre-training data: 2.1 billion pairs, 512 GB Twitter reply pairs
  - Finetune data: 200,000 pairs of high-quality dialog data (accumulated by NTT in past studies)
- Error analysis of unnatural utterances
  - Most of the utterances are natural, but about 7% of the utterances are strongly unnatural.
  - Subjective evaluation significantly decreased when contradictions, topic skipping, or utterances that differ from the facts are included.

Awarded the highest prize in the 3rd dialogue system live competition, where social dialogue systems competed.

This study was supported by the Grant-in-Aid for Scientific Research on Innovative Areas "Communicative intelligent systems towards a human-machine symbiotic society" (Issue No. 19 H05693).

References

[1] H. Sugiyama, H. Narimatsu, M. Mizukami, T. Arimoto, Y. Chiba, T. Meguro, H. Nakajima, , "Development of conversational system talking about hobby using Transformer-based encoder-decoder model," in *Proc.* Special Interest Group on Spoken Language Understanding and Dialogue Processing (*SIG-SLUD)*, Vol. B5, No. 02, pp. 104-109, 2020 (in Japanese).
[2] H. Sugiyama, T. Meguro, Y. Yoshikawa, J. Yamato, "Improving Dialogue Continuity using Inter-Robot Interaction," in *Proc.* IEEE International Symposium on Robot and Human Interactive Communication (*RO-MAN)*, pp. 105-112, 2018.

Contact

Hiroaki Sugiyama / Interaction Research Group, Innovative Communication Laboratory
Email: cs-openhouse-ml@hco.ntt.co.jp