Utterance Selection based on Sentence Similarities and Dialogue Breakdown Detection on NTCIR-12 STC Task

Hiroaki SUGIYAMA
NTT Communication Science Labs., NTT corp., Japan

Objective
Utterance selection for chat-oriented systems from a number of candidates

User utterance
I like classical music, especially Beethoven.

Sentence similarity based retrieval
•Retrieve sentences that resemble the user utterance using word2vec.
  \[ Sim(s_1, s_2) = v_{s_1} \cdot v_{s_2} \]
  \[ v_s = \frac{\sum_{w \in W_s} v_w}{|W_s|} \]
  \( v_w \): word2vec based vector

Dialogue breakdown detection based reranking
•Detects inappropriate utterances that cause dialogue breakdown.
•We leverage the estimated appropriateness to reranking of the candidates.

System utterance
I think Beethoven is thick sounds with presence.

Proposed approach

Utterance candidates

Microblog Corpus

Utterance selection for chat-oriented systems from a number of candidates

Experiments

Approach | Correct=2 1-best | Correct=2 5-best | Correct=1&2 1-best | Correct=1&2 5-best
--- | --- | --- | --- | ---
W/o DBD | 0.0921 | 0.0698 | 0.2639 | 0.2318
W/ DBD | 0.0876 | 0.0677 | **0.2946** | **0.2333**

Results and Discussion
© DBD reranking is effective to filter inappropriate utterances.
© DBD does not have enough sensitivity to distinguish labels 1 and 2, because the DBD was trained with dialogue system's utterances that are less appropriate than the tweets.

Future work
•Examine fine-tuned DBD system with the tweets.

Implement
- Word2vec
  •Develop with 150M tweets (2013)
- Dialogue breakdown detection (DBD)
  •Six layer perceptron [Sugiyama 2015]
  •Trained with DBD corpus [Higashinaka, 2015]