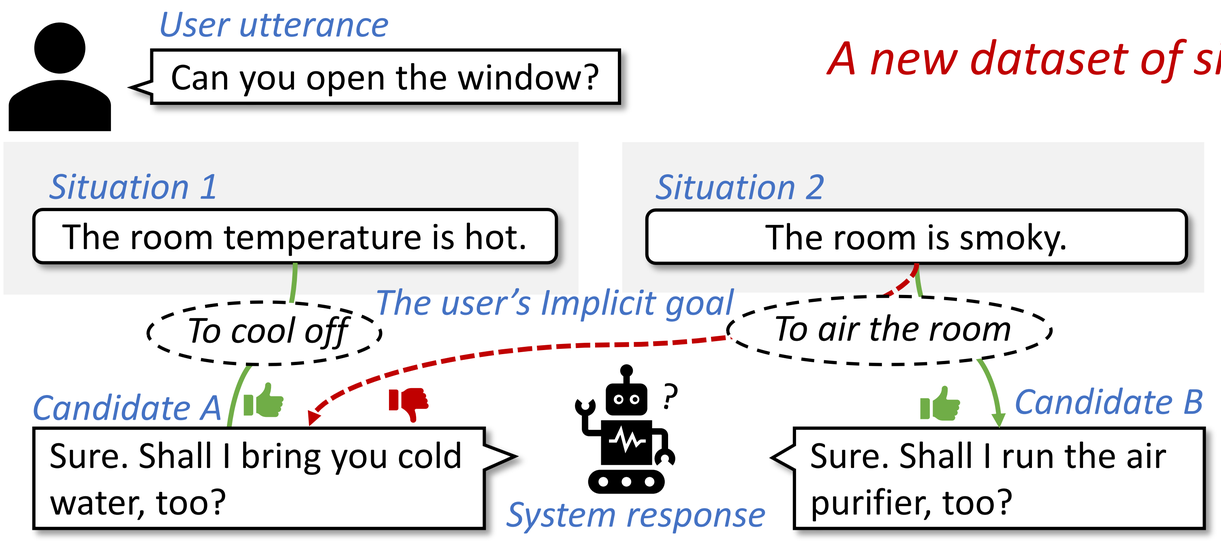
A Textual Dataset for Situated Proactive Response Selection

Naoki Otani[1], Jun Araki[2], HyeongSik Kim[2], and Eduard Hovy[3]
[1]Megagon Labs (*Work done at Carnegie Mellon University), [2]Robert Bosch LLC, [3]University of Melbourne

#5850



Proactive conversational assistance: Addressing the implicit user goal

A new dataset of single-turn conversations + situational information

Highlights

We represent situational information by **free-form English texts.**
- Six semantic categories: location, possession, time/date, …

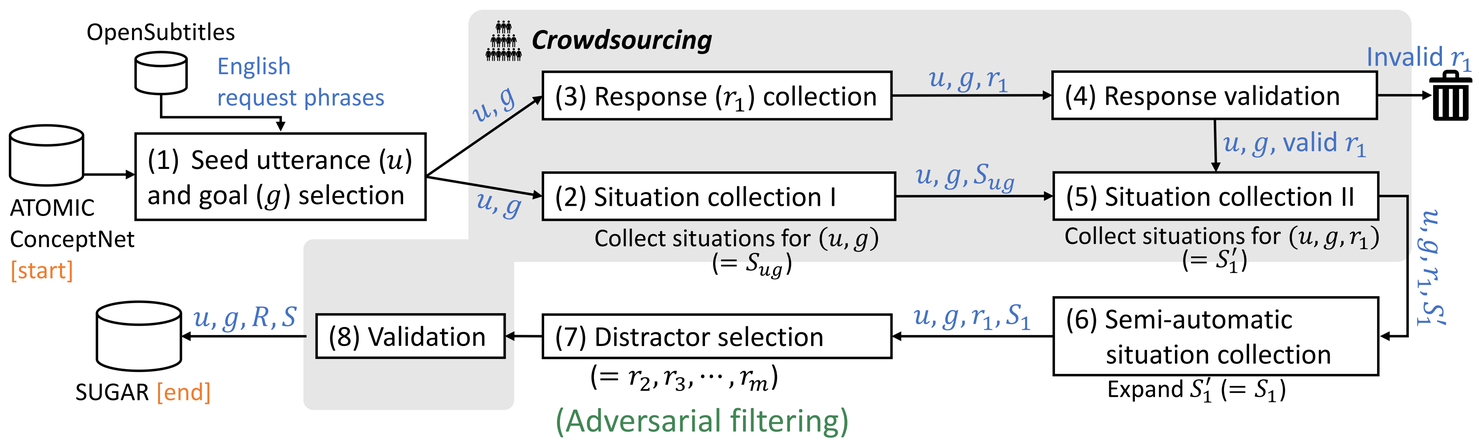
We introduce a dataset of **1,760** single-turn conversations.
- Language: English
- Target scenario: help-seeking (user request → proactive response)
- Crowdsourcing + Language generation model + Manual verification

Our experiments show that response systems **can be misled by distractors.**
- Selection accuracy decreases only with ~5 irrelevant statements in input.

Representation of Situation: Short English statements that describe *observable* facts of the current world state

Category	Definition	Example
Location	Information about the user's current location	The user is home. / The user is in the kitchen. / The user is in the office.
Possession	Information about what the user possesses	The user owns a car. / There are apples in the kitchen.
Time	Information about time	It's midnight now. / It's morning now.
Date	Information about date and season	Today is the user's birthday. / It's summer now.
Behavior	Information about the user's behavior	The user just woke up. / The user has a flight to catch this afternoon.
Environment	Information about non-user entities and eventualities	The room temperature is hot. / The user's car has a flat tire.

Resource collection pipeline: Crowdsourcing + Language Generation model + Manual verification



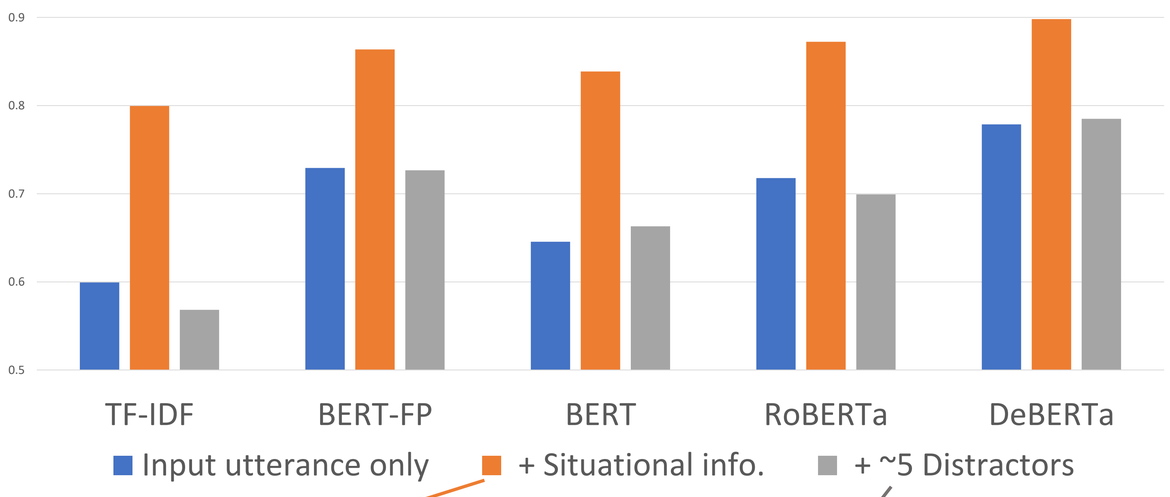
Utterance	Please turn on the TV. (u)
Situations	It is evening now. [user] is home. [user] is in the living room. [user] is sitting on the couch. [user] has a TV in the house. [user] has an outfit on the bed. [user] has drinks and snacks in the kitchen. [user] has game cards on the shelf. The TV is off. [someone]'s birthday is today. There are several sports games available to watch. (r_1) There is a basketball game scheduled. (S)
Responses	Sure. Would you like me to check today's sports listings? *(Best)* (r_1) Sure. Shall I pour a drink and bring some snacks for the game? *(Acceptable)* (r_2) Sure, shall I select an outfit for you? *(Bad)* (r_3)

Response Selection Experiments: RQ: Can systems use situational statements to find optimal responses?

User request (u)
+ Situational statements (S) → Ranker → Score of the response candidate
+ Response candidate (r)

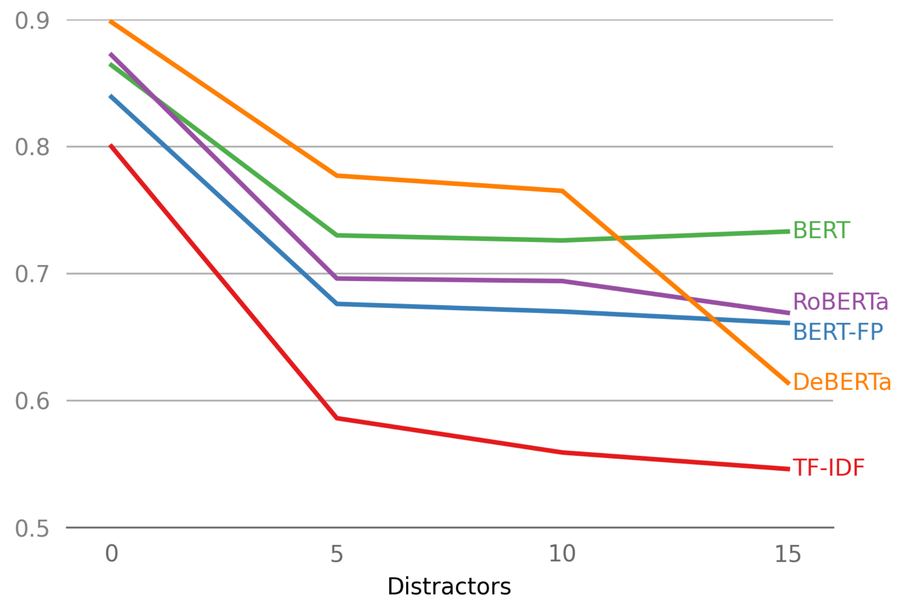
Ranker
- TF-IDF encoder
- Transformer encoder (BERT, BERT-FP[(Han+21)], RoBERTa, DeBERTa)

Training: optimizing rankers for a max-margin loss
Evaluation: Precision@1, nDCG@3 (5-fold cross-validation)
Results (P@1):



Relevant situational statements gave a performance boost

When a few distractors were added, scores dropped by large margins



RoBERTa and DeBERTa (higher-performing systems) were more easily misled by distractors

Conclusion

- We introduced a new dataset of single-turn help-seeking conversations **augmented with situational statements**
- Our dataset was created through crowdsourcing and a neural language generation model followed by multiple manual verification steps.
- Response selection experiments show that systems can benefit from situational information but at the same time can be misled by distractors.

Future work:
(1) Exploring representations of situational information (multi-modal representations? adequacy?)
(2) Extension to response generation (Refer to our paper @ NLP4ConvAI workshop)
Naoki Otani, Jun Araki, HyeongSik Kim, and Eduard Hovy. 2023. On the Underspecification of Situations in Open-domain Conversational Datasets. NLP4ConvAI workshop