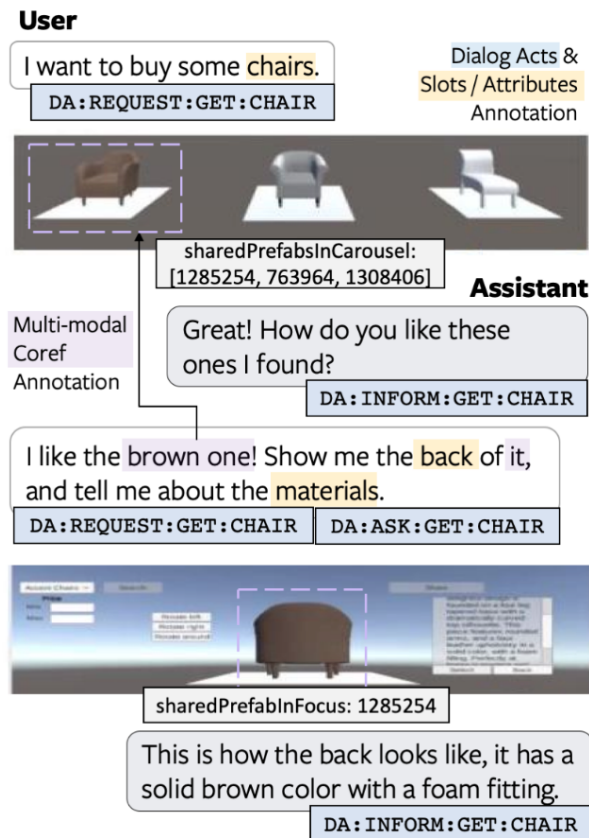


Situated Interactive MultiModal Conversations (SIMMC) Challenge 2020

Facebook is organizing the Situated Interactive Multimodal Conversations (SIMMC) Track for [DSTC9](#), and we look forward to your participation!

The SIMMC challenge aims to lay the foundations for the real-world assistant agents that can handle multimodal inputs, and perform multimodal actions. We thus focus on task-oriented dialogs that encompass a situated multimodal user context in the form of a co-observed image or virtual reality (VR) environment. The context is dynamically updated on each turn based on the user input and the assistant action. Our challenge focuses on our SIMMC datasets, both of which are shopping domains: (a) furniture (grounded in a shared virtual environment) and, (b) fashion (grounded in an evolving set of images).

Organizers: Ahmad Beirami, Eunjoon Cho, Paul A. Crook, Ankita De, Alborz Geramifard, Satwik Kottur, Seungwhan Moon, Shivani Poddar, Rajen Subba



Example from SIMMC-Furniture Dataset

Latest News

- [June 22, 2020] Challenge announcement. Training / development datasets are released.

Important Links

- [Task Description Paper](#)
- [Challenge Registration](#)
- [Data Formats](#)
- Baseline Details: [MM Action Prediction](#), [MM Response Generation](#), [MM-DST](#)
- [Challenge Instructions](#)

Timeline

Date	Milestone
June 22, 2020	Training & development data released
Sept 28, 2020	Test-Std data released, End of Challenge Phase 1
Oct 5, 2020	Entry submission deadline, End of Challenge Phase 2
Oct 12, 2020	Final results announced

Track Description

Tasks and Metrics

We present three sub-tasks primarily aimed at replicating human-assistant actions in order to enable rich and interactive shopping scenarios.

Sub-Task #1	Multimodal Action Prediction
Goal	To predict the correct Assistant API action(s) (classification)
Input	Current user utterance, Dialog context, Multimodal context

Output	Structural API (action & arguments)
Metrics	Perplexity, Mean Average Precision

Sub-Task #2	Multimodal Dialog Response Generation & Retrieval
Goal	To generate Assistant responses or retrieve from a candidate pool
Input	Current user utterance, Dialog context, Multimodal context, (Ground-truth API Calls)
Output	Assistant response utterance
Metrics	Generation: BLEU-4, Retrieval: Accuracy@k, Entropy

Sub-Task #3	Multimodal Dialog State Tracking (MM-DST)
Goal	To track user belief states across multiple turns
Input	Current user utterance, Dialogue context, Multimodal context
Output	Belief state for current user utterance
Metrics	Intent F1, Slot F1

Evaluation

For the DSTC9 SIMMC Track, we will do a two phase evaluation as follows.

Challenge Period 1: Participants will evaluate the model performance on the provided `devtest` set. At the end of Challenge Period 1 (Sept 28), we ask participants to submit their model prediction results and a link to their code repository.

Challenge Period 2: A `test-std` set will be released on Sept 28 for the participants who submitted the results for the Challenge Period 1. We ask participants to submit their model predictions on the `test-std` set by Oct 5. We will announce the final results and the winners on Oct 12.

Challenge Instructions

(1) Registration

- Fill out [this form](#) to register at DSTC9. Check “Track 4: Visually Grounded Dialog Track” along with other tracks you are participating in.
- After the registration, we highly encourage the participants to notify the SIMMC Track Organizers (simmc@fb.com) of their registration, to help us communicate any future updates on the codebase, the datasets, and the challenge track.

(2) Download Datasets and Code

- Git clone our repository to download the datasets and the code. You may use the provided baselines as a starting point to develop your models.

```
$ git lfs install
$ git clone https://github.com/facebookresearch/simmc.git
```

(3) Reporting Results for Challenge Phase 1

- Submit your model prediction results on the `devtest` set at simmc@fb.com (more instructions to follow later).
- We will then send you the `test-std` set (with ground-truth labels hidden) on Sept 28.

(4) Reporting Results for Challenge Phase 2

- Submit your model prediction results on the `test-std` set at simmc@fb.com (more instructions to follow later).
- We will evaluate the participants’ model predictions using the same evaluation script for Phase 1, and announce the results.

Contact

Questions related to SIMMC Track, Data, and Baselines

Please contact simmc@fb.com, or leave comments in the Github repository.

DSTC Mailing List

If you want to get the latest updates about DSTC9, join the [DSTC mailing list](#).

Citations

If you want to publish experimental results with our datasets or use the baseline models, please cite the following articles:

```
@article{moon2020situated,  
  title={Situating and Interactive Multimodal Conversations},  
  author={Moon, Seungwhan and Kottur, Satwik and Crook, Paul A and De, Ankita and Poddar, Shivani and Levin, Theodore and Whitney, David and Difrancia, Daniel and Beirami, Ahmad and Cho, Eunjoon and Subba, Rajen and Geramifard, Alborz},  
  journal={arXiv preprint arXiv:2006.01460},  
  year={2020}  
}
```

```
@article{crook2019simmc,  
  title={SIMMC: Situating Interactive Multi-Modal Conversational Data Collection And Evaluation Platform},  
  author={Crook, Paul A and Poddar, Shivani and De, Ankita and Shafi, Semir and Whitney, David and Geramifard, Alborz and Subba, Rajen},  
  journal={arXiv preprint arXiv:1911.02690},  
  year={2019}  
}
```

NOTE: The [paper](#) above describes in detail the datasets, the NLU/NLG/Coref annotations, and some of the baselines we provide in this challenge. The paper reports the results from an earlier version of the dataset and with different train-dev-test splits, hence the baseline performances on the challenge resources will be slightly different.

License

SIMMC is released under [CC-BY-NC-SA-4.0](#), see [LICENSE](#) for details.