CerealBar: building agents that participate in situated, collaborative natural language interactions

Executing Instructions in Situated Collaborative Interactions

Alane Suhr, Claudia Yan, Charlotte Schluger*, Stanley Yu*, Hadi Khader**, Marwa Mouallem**, Iris Zhang, Yoav Artzi. **: Equal contribution

Okay, pick up yellow hearts and run past me toward the bush sticking out, on the opposite side is 3 green stars

Goal: select sets of three cards with distinct colors, shapes, and counts

Leader view

Follower view

Leader
- Plans which cards to collect
- Moves and collects cards
- Delegates tasks using natural language
- Has full observability of the game board

Follower
- Follows the leader's instructions
- Moves and collects cards
- Has more steps per turn than the leader
- Only observes a first-person view

The CerealBar Scenario

Spatial reasoning
Collaborative interaction
Sequential instructions
User interaction

Model

An interpretable neural network agent

Instruction →

Stage 1: Plan Prediction

Plan distributions

- Goal cards
- Path to cards
- Cards to avoid
- Obstacles

Stage 2: Action Generation

→ LEFT
→ FORWARD

Data

1,202 human-human games
Vocabulary: 3,641 word types

	Median	Max
Score	9.0	19
# Instructions per game	24.0	40
Instruction length	13.0	55
Follower actions per instruction	8.0	50

Training

Learning to recover from errors

- **Problem**: human-human interactions don't provide signal about agent error recovery
- **Solution**: create and add such examples to training set during learning

Sampled Act.
Static Oracle

Instruction 1:
Go to the left of the tree

Instruction 2:
Get the card straight ahead

Trained agent makes a mistake

New training example showing error recovery

Agent Partner

- Auxiliary loss on Stage 1 activations to predict if the example requires error-recovery
- Encourages the model to learn how to reason differently about error-recovery examples

Code, demos, and videos at:
lil.nlp.cornell.edu/cerealbar

Evaluation

A new protocol for sequential instruction evaluation

Instruction-level metrics don't measure effects of error propagation

Interaction-level metrics poorly utilize data to measure error propagation due to failures early in the interaction

New evaluation examples

Solution: evaluate the model's performance when starting in each instruction in each interaction until the end of the interaction

Full Model

8.7	15.4	96.8
6.5	17.9	95.6

Prop. Instr. Followed   Prop. Points Scored

Learned agent effectively collaborates with human leaders

Interaction allows humans to adapt language and behavior to best use the agent

Agent Partner   Human Partner

	Mean Points	Max Points
Agent	6.2	14
Human	12.7	20

With automated followers:
- Shorter instructions (12.3 vs. 8.5 tokens)
- Smaller vocab (1037 word types vs. 578)