Optimal Use Of Verbal Instructions For Multi-Robot Human Navigation Guidance

Harel Yedidsion, Jacqueline Deans, Connor Sheehan, Mahathi Chillara, Justin Hart, Peter Stone, and Raymond Mooney

Learning Agents Research Group
The University of Texas at Austin

TEXAS
The University of Texas at Austin

machinelearning researchgroup

ICSR 2019
International Conference on Social Robotics
FRIENDLY ROBOTS

Indoor Human Navigation

- **The problem:**
  - Large complex buildings
  - No indoor localization
  - No reliable pedestrian odometry

- **Possible solution:**
  - Using mobile, verbally communicating robots

Robot for Human Guidance

Previous work:

1. **<u>MSR – Stationary directions robot (Bohus et al. 2014)</u>**

Actions : **Instruct**, **Gesture**

- Memorizing a long sequence of instructions is difficult
- The tendency to make mistakes increases with the length of the instruction sequence and the complexity of the environment

Robots for Human Guidance

Multiple Robots for Human Guidance

Previous work:

<u>2. UT's multi-robot human guidance system (Khandelwal et al. 2015, 2017)</u>

Actions : **Lead**, **Direct** (using arrows on screen)

- It is frustrating for the human to walk behind the robot which moves at a third of the speed of a human
- Can only direct in straight lines

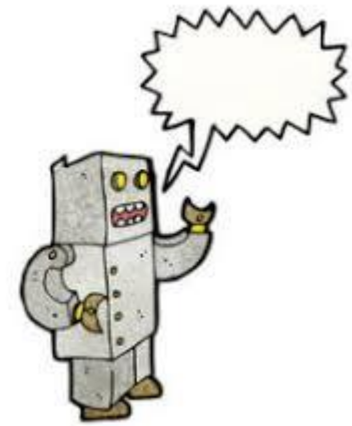
Adding Natural Language Instructions

Benefits:

➢ Instruct the human through areas which are hard for the robot to navigate

➢ Complete guidance task quickly

➢ Minimize the robots' time away from background tasks



Challenges:

➢ How to generate the natural language instructions?

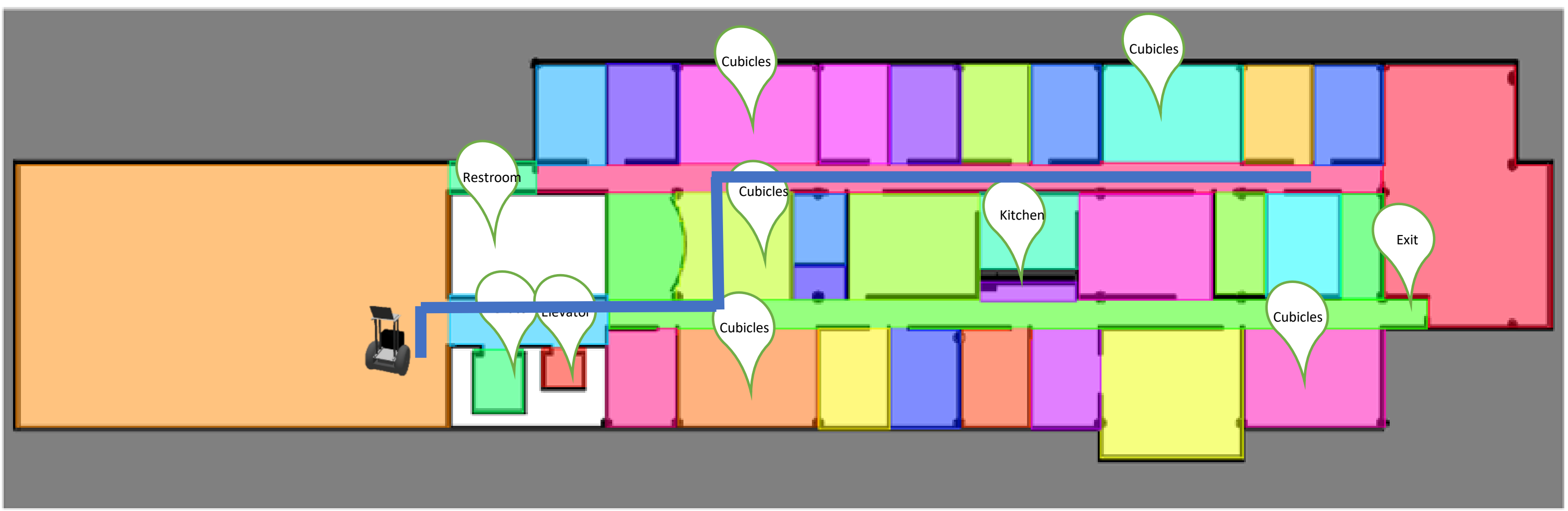
➢ How to optimize leading, instructing, and transitioning

➢ Robot implementation

➢ Coming up with a good human behavior model

Natural Language Instruction Generation

- ➢ We annotated a map with regions and landmarks
- ➢ Based on the robot's planned path we generate natural language instruction
- ➢ Template-based method using landmarks as navigational waypoints.
- ➢ Action – Preposition - Landmark



Preliminary Study

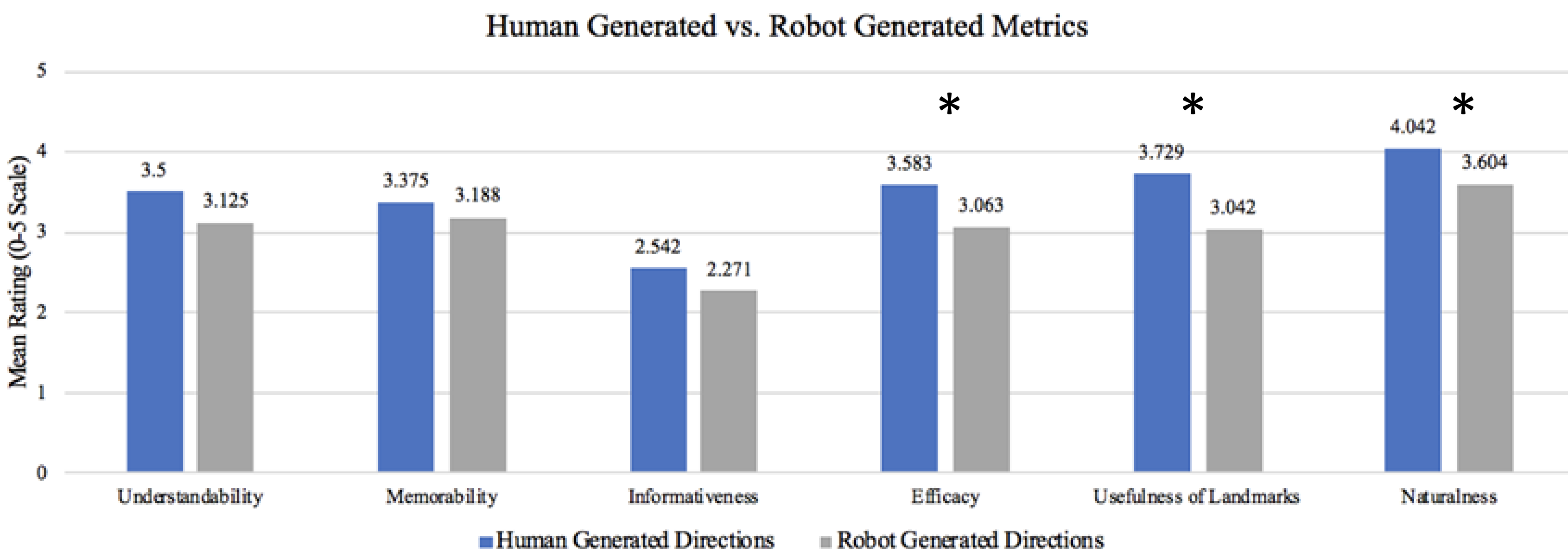
- 25 people over 4 paths either human/robot generated instructions

Location Pair	Path	Human Generated Directions	Robot Generated Directions
A	in front of 3.424 → Stacy Miller's office	Go down the hall. Next to the large whiteboard, take a left and go past the table. You'll arrive at Stacy Miller's office.	go through the hallway. turn left towards the cubicles. go through the cubicles. Stacy Miller's office is straight ahead!
B	common area → Robot soccer field	Turn right at the kitchen and go all the way down the hall. At the end, take a left past the soccer field to the door at the end of the hall.	go through the common space. turn right at the trashcan. go through the hallway. turn left at the exit sign. go through the hallway. turn right at the exit sign. robot soccer field is straight ahead!
C	back cubicles on the kitchen hallway → Justin's office	Go down the hall past the kitchen to the whiteboard and table. Take a right and pass the whiteboard and table. Take a left. Justin Hart's office is on the right.	go through the hallway. turn right at the exit sign. go through the cubicles. turn left towards the hallway. go through the hallway. turn right at the exit sign. Justin Hart's office is straight ahead!
D	area near the restrooms → Kitchen	Take a left at the first hallway towards the exit. Walk past the elevators and through the door. Continue down the hall until the kitchen is to the left.	go through the 3rd floor lab. turn left towards the hallway. go through the hallway. kitchen is straight ahead!

Directions used in the study

Preliminary Study

- The instruction generation system was almost as good as human generated instructions.



No statistically significant differences in task duration, Understandability, Memorability and Informativeness

Optimizing the Lead/Instruct combination

- For each region we measure the following properties:
  - Length of the path inside the region $length(p, l_j)$
  - Robot's traversability per region trv_j
  - Human's probability of going wrong per region cmp_j
  - Number of previously consecutive instructed regions cir_j
- Other parameters
  - Robot's speed v_r
  - Human's speed v_h
  - Robot observability factor rof
  - Duration of saying the instruction for a region t_c

Optimizing the Lead/Instruct combination

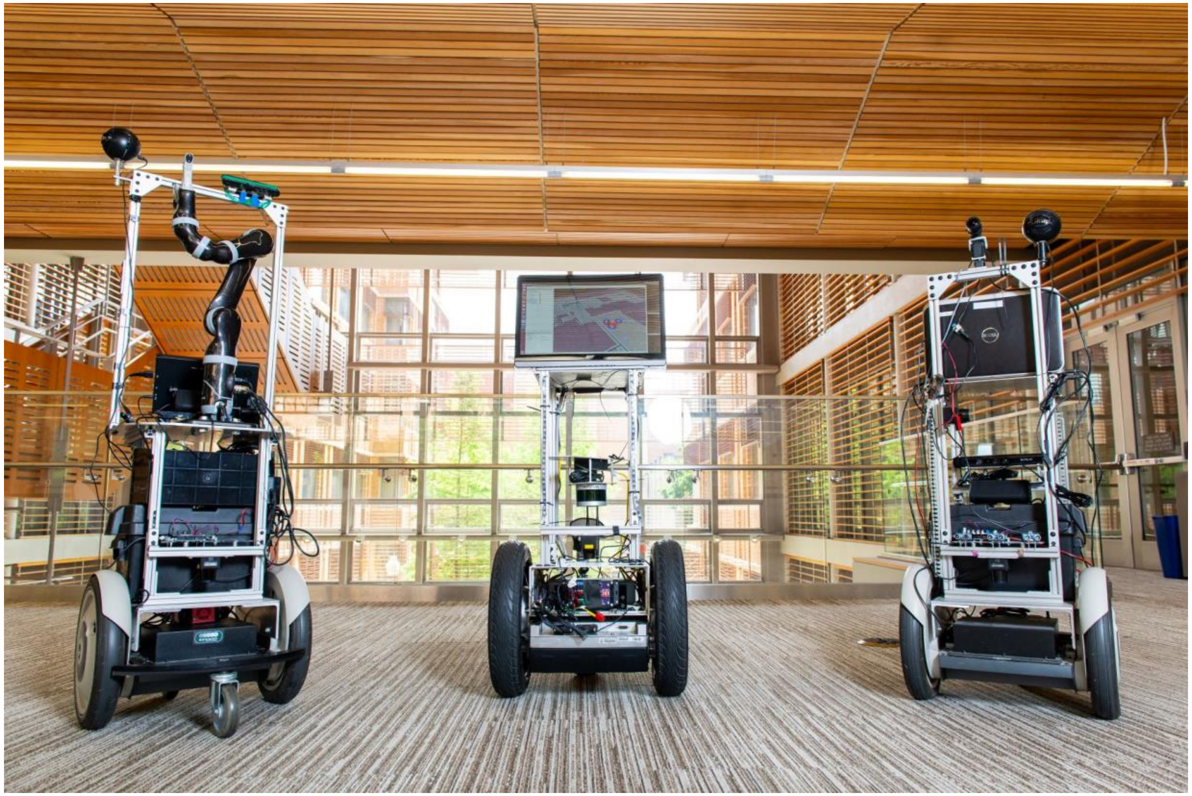
- Objective:

$$min \sum_{l_j \in L^p} t_j$$

$$t_j = \begin{cases} (length(p, l_j)/v_r)/trv_j & \text{if } action = Lead \\ (length(p, l_j)/v_h) \cdot cmp_j \cdot (cir_j + 1) + t_c & \text{if } action = Instruct \\ (length(p, l_j)/v_h) \cdot cmp_j \cdot (cir_j + 1)/rof + t_c & \text{if } Transition \quad \text{at} \quad l_j \end{cases}$$

Robot Implementation

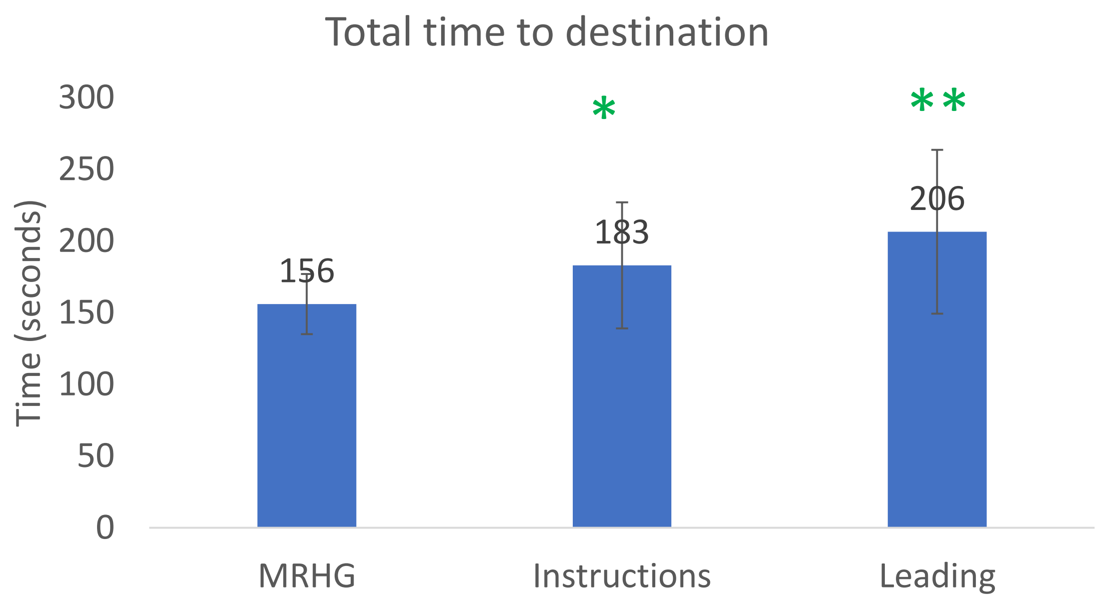
- BWIBots
- ROS
- WaveNet
- SpeechToText
- Node.js
- ROSBridge



Robot Implementation

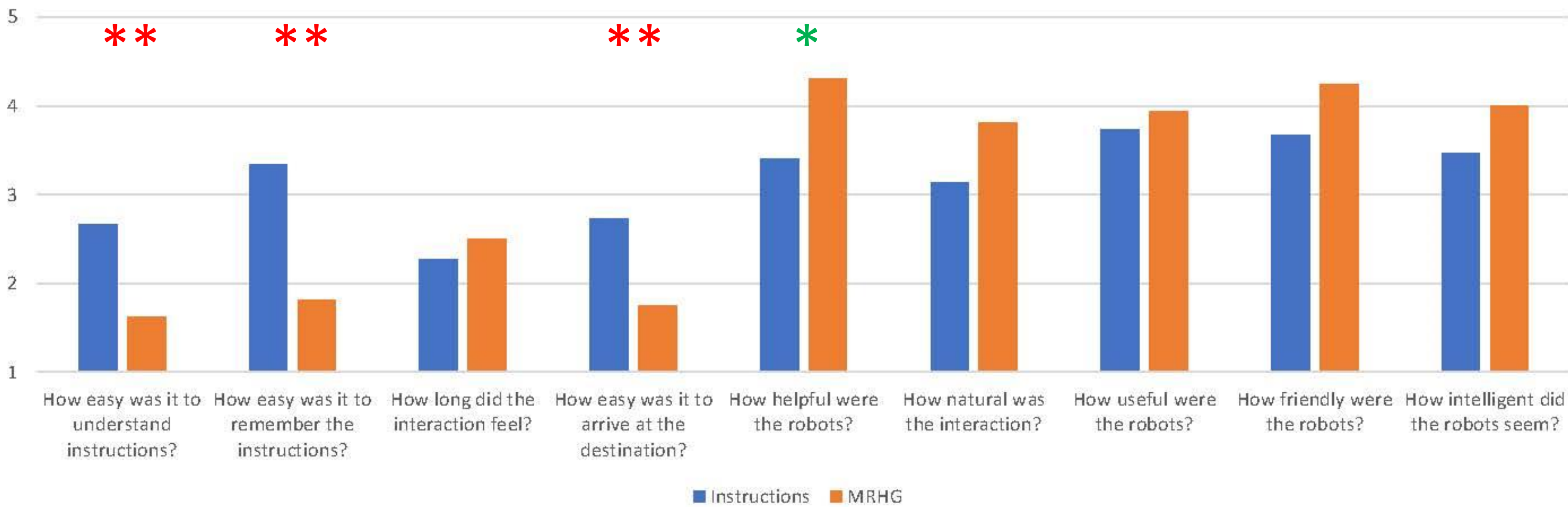
Experiments

- 30 participants without prior knowledge of GDC were recruited.
- 15 got **Instructions** only and 15 were guided by the **MRHG** system.
- For the **Leading** condition the robot ran 15 times without a human.



Survey Results

100% of the Instructions participants requested that the robot repeat the instructions and a third of them didn't make it to the destination



Better: Naturalness, helpfulness, intelligence, friendliness, and usefulness

Worse: understandability, memorability easiness, and perceived length of interaction

Conclusions

- Integrate **multi-robot coordination** with **natural-language** instruction generation.
- Use the **robots' path planner** and a landmark annotated map to generate natural language instructions.
- Tested on **human participants** and performed better than the Instructions benchmark in terms of both **success rate and time to destination.**
- Future :
  - Disfluency
  - Classification
  - Considering longer paths