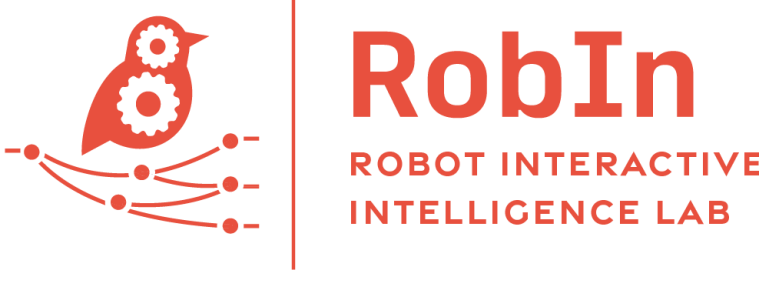




Mixed-Initiative Dialog for Human-Robot Collaborative Manipulation



Albert Yu, Chengshu Li, Luca Maccesanu, Arnav Balaji, Ruchira Ray, Ray Mooney, Roberto Martín-Martín
University of Texas at Austin, Stanford | albertyu@utexas.edu, robertomm@cs.utexas.edu



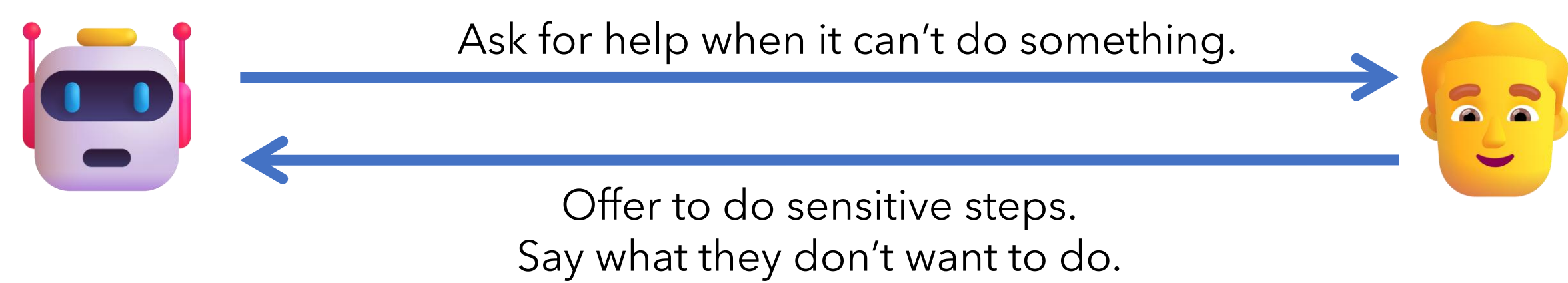
Robots can't do everything. How can human-robot teams leverage bidirectionally initiated dialog to communicate and collaborate better on **long-horizon mobile manipulation** tasks?

Problem Statement

- Robots today lack capabilities to be deployed autonomously for household tasks.
- Robots often need human help to complete complex tasks. Sometimes the human wants to perform sensitive parts of a task themselves.
- However, solely maximizing for success rate may cause the robot to ask for human help on every step. The robot should instead be helpful and alleviate human effort.
- Collaboration requires communication to distribute work suitably.

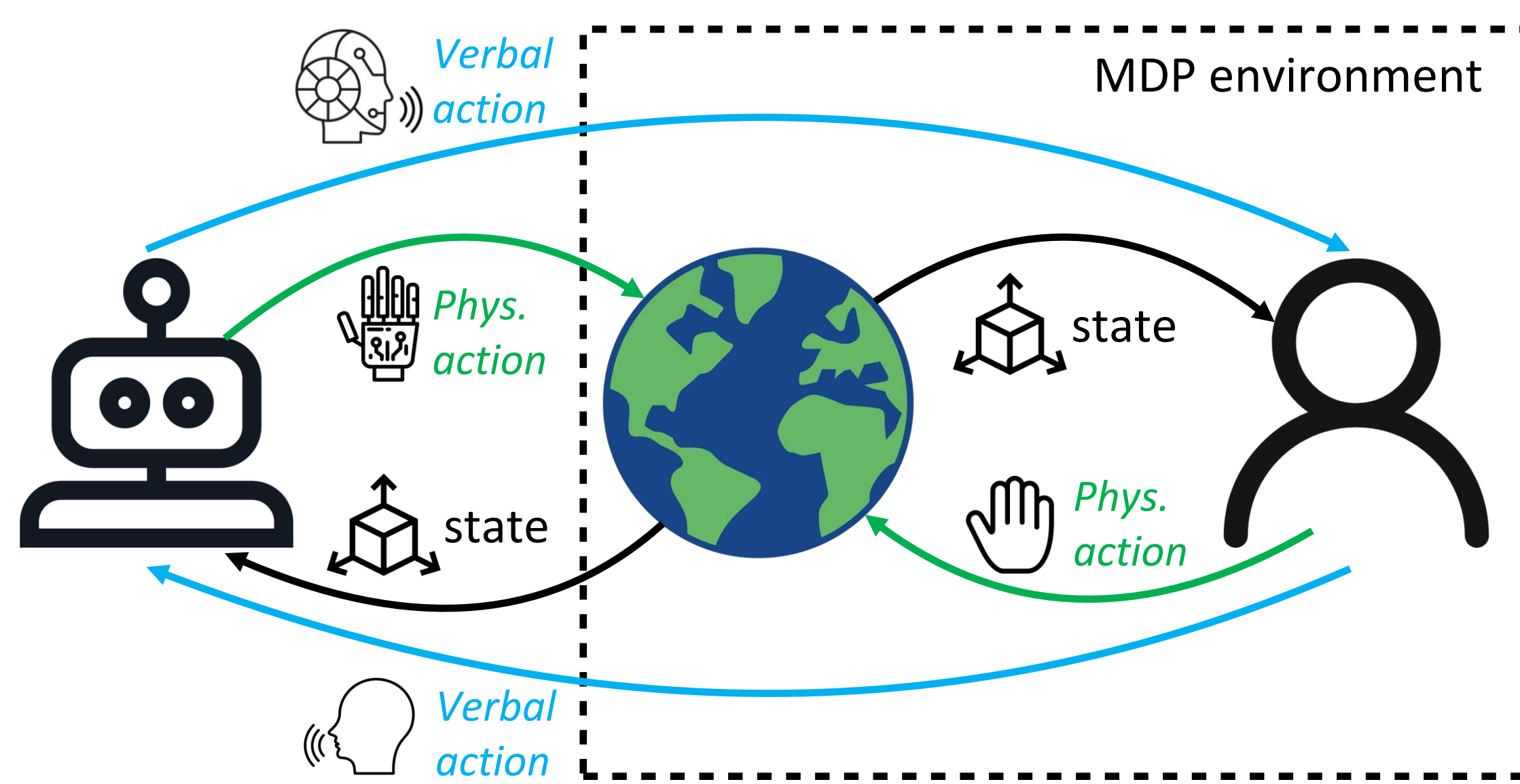
Insight: Mixed-initiative Dialog Improves Collaboration

- For efficient collaboration, both agents should be able to initiate + act from dialog.



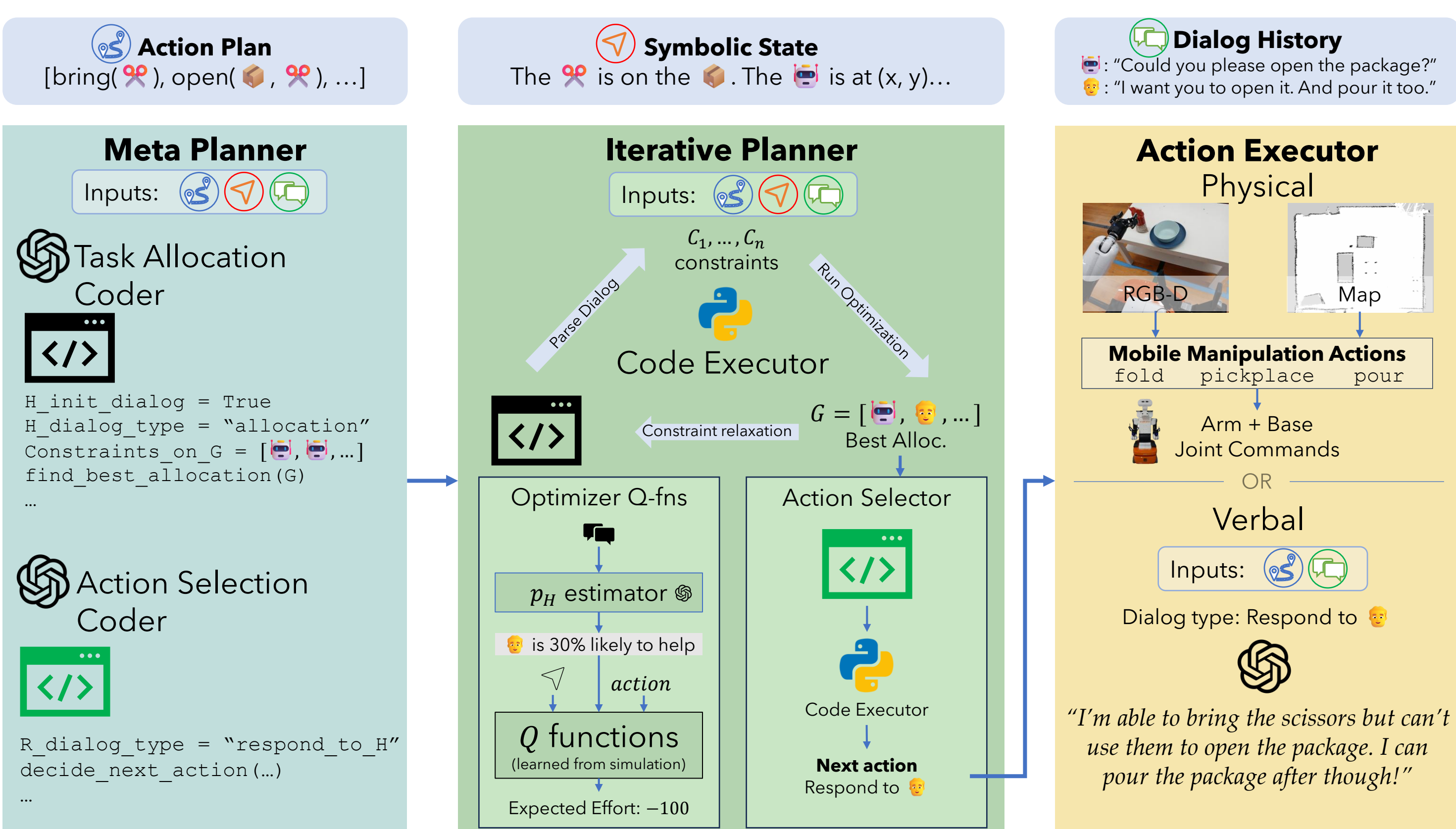
- This communication and collaboration must adapt to human behavior. For instance, the robot may attempt steps it is less proficient at if the human is unhelpful or busy.

Problem Formulation



Optimization Problem: Allocate steps to 🤖 or 👤.
Maximize task success. Minimize 👤 effort. Adhere to 👤-specified constraints.
Core Problem: Communicate this allocation to 🤖. Understand 👤 lang. constraints.

MicoBot Architecture



Our system can both initiate dialog (first row) and understand and carry out human-initiated dialog (second row).
Mixed-Initiative Dialog: a paradigm with free-flowing questions and answers from both the human and robot (Carbonell, 1970).

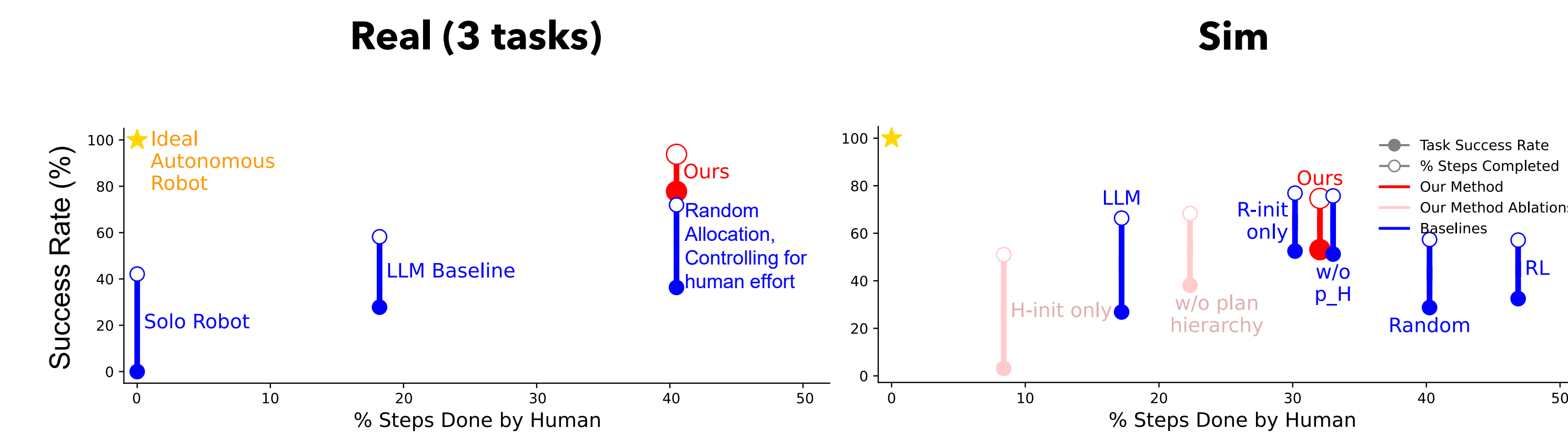


3 Real World Tasks

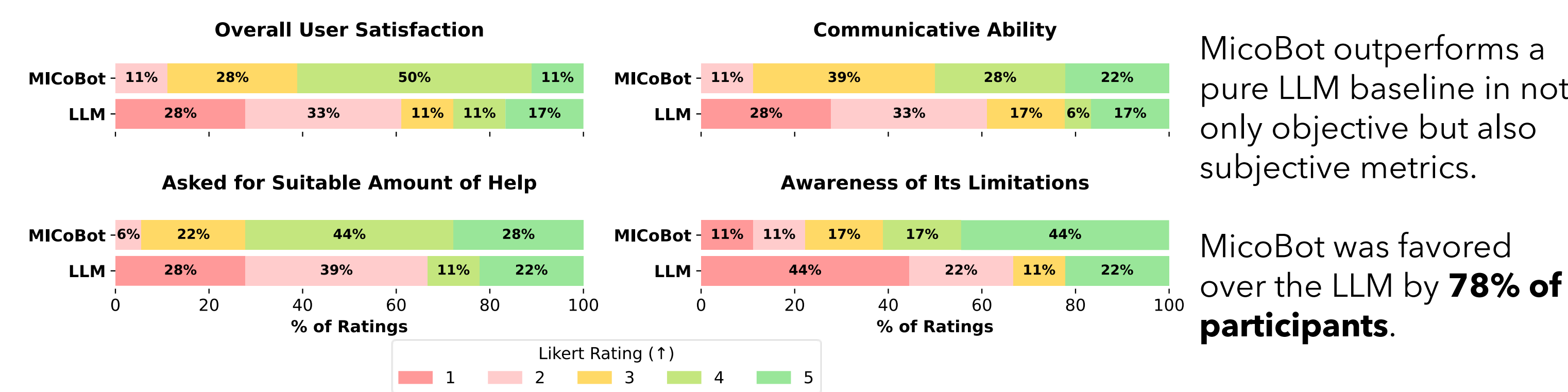


Results: Success x Human Effort

In real human user studies ($n = 18$), and thousands of rollouts in simulation, MicoBot outperforms a pure LLM baseline and all ablations, achieving the best success rate for a fixed amount of human effort.

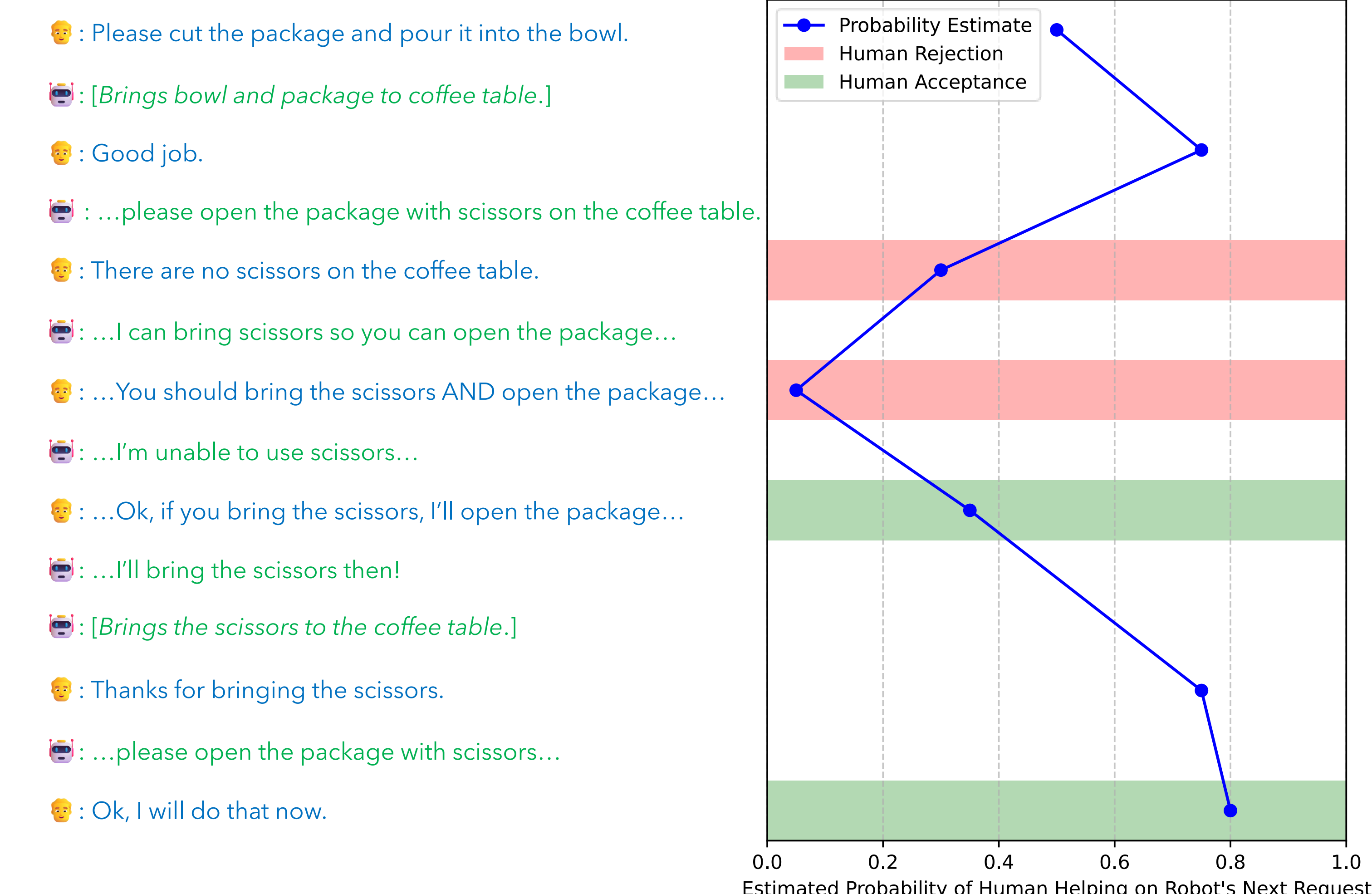


Human User Study Ratings



Estimating Human Helpfulness in User Study

A dialog interaction where MicoBot convinces an initially reluctant human to help perform a step it is incapable of. MicoBot's estimate of human helpfulness (p_{help}) reacts to the latest human language.



Main Takeaways

- MicoBot is the first work to enable mixed-initiative dialog for real-world HRI.
- Mixed-initiative dialog improves real-world mobile manipulation task success compared to single-initiative dialog and adaptability to diverse human collaborators.
- MicoBot outperforms a pure LLM baseline in the real world by leveraging affordance knowledge learned from simulation trajectories.

MicoBot outperforms a pure LLM baseline in not only objective but also subjective metrics.
MicoBot was favored over the LLM by **78% of participants**.