*16-867*

# Human Robot Interaction

## Introduction

Instructor: Andrea Bajcsy

Carnegie Mellon University

intent ROBOTICS LAB

# Welcome!

# Professor

Andrea Bajcsy
(BYE-chee)

What to call me:

- Andrea *(if you are a grad student)*
- Prof. Bajcsy or Prof. B *(if you are undergrad)*

**Office Location**: NSH 4629
**Office Hours**: Tuesdays, 12:20-1:20pm *(after class)*
**Email**: abajcsy@cmu.edu

# Teaching Assistant

Pranay Gupta, PhD Student

Research Interests:

- Assistive driving
- Shared control

**Office Location**: NSH 4504
**Office Hours:** 4:00 - 5:00pm
**Email:** pranaygu@andrew.cmu.edu

# What is next?

Course Content

Logistics

Intro Survey

(Intro to Single-Agent Decision Making)

# Round of Introductions

- Name
- Department
- Year (Masters, PhD, …)
- Research Interests

*What makes human-robot interaction different from "typical" robotics?*

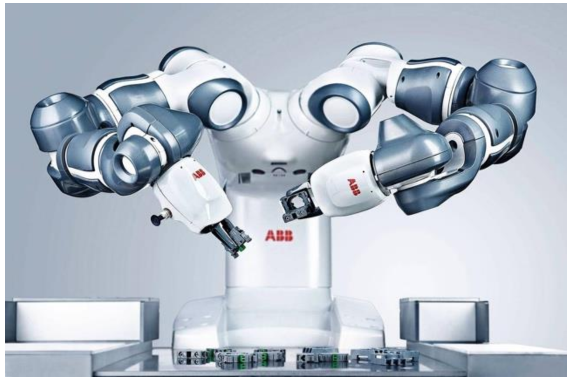

Boston Dynamics

ABB

## **Small group activity (5 min)**

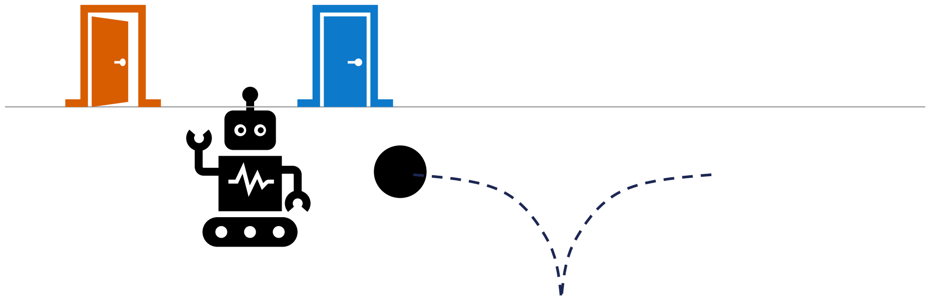
Turn to your neighbor, introduce yourself, and discuss:

*What makes human-robot interaction different from "typical" robotics?*



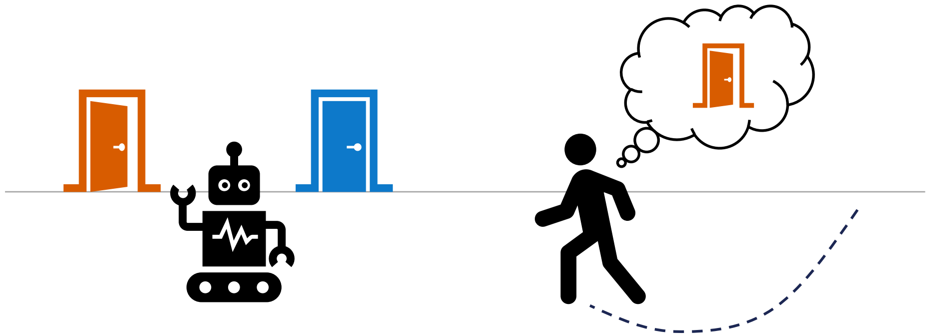

*vs.*

*Interaction*



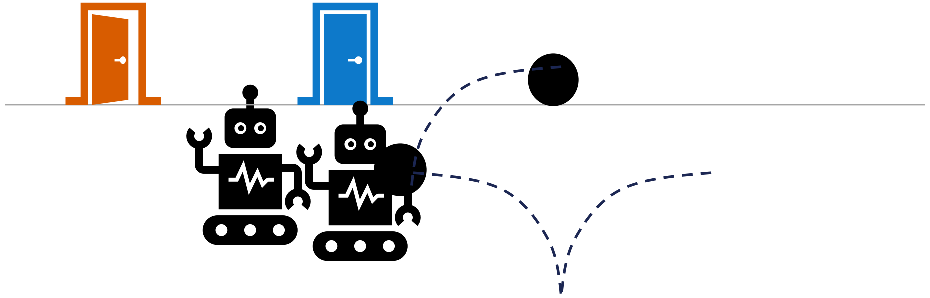
Environment driven by laws of physics

*Human Interaction*



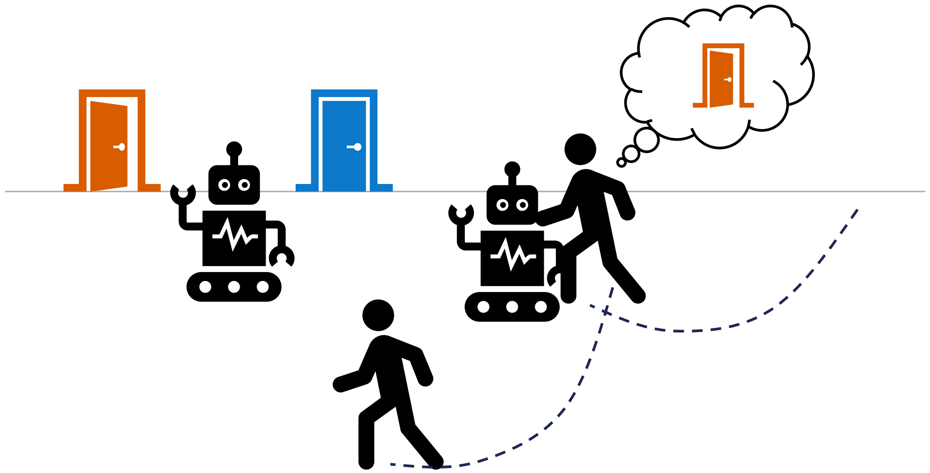
Human driven by physics and hidden **internal objectives**

*Interaction*



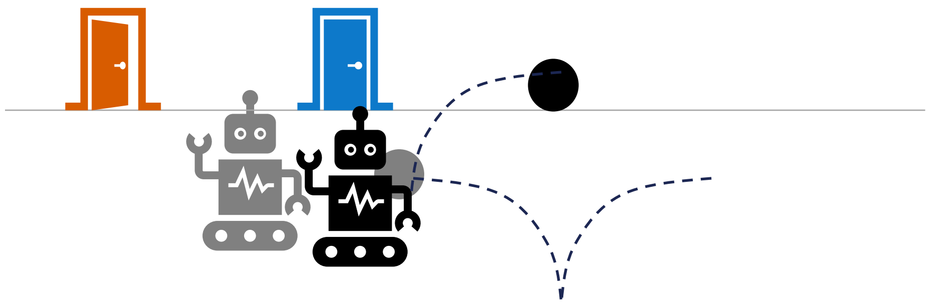
Environment can *be influenced* by robot's actions **directly**

*Human Interaction*



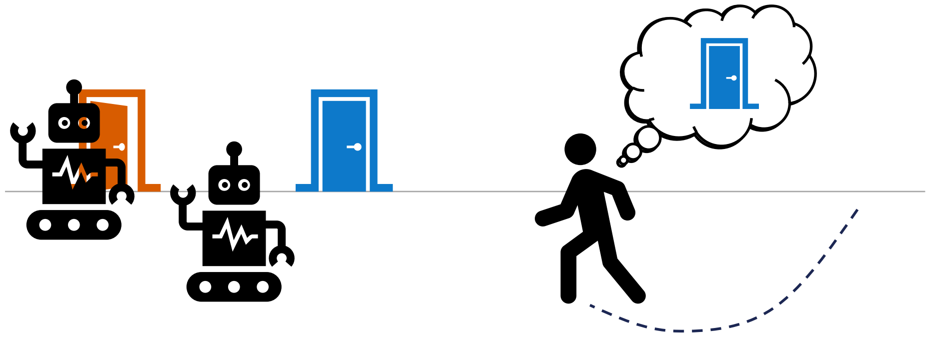
Human can *be influenced* by the robot's actions **directly**….

*Interaction*



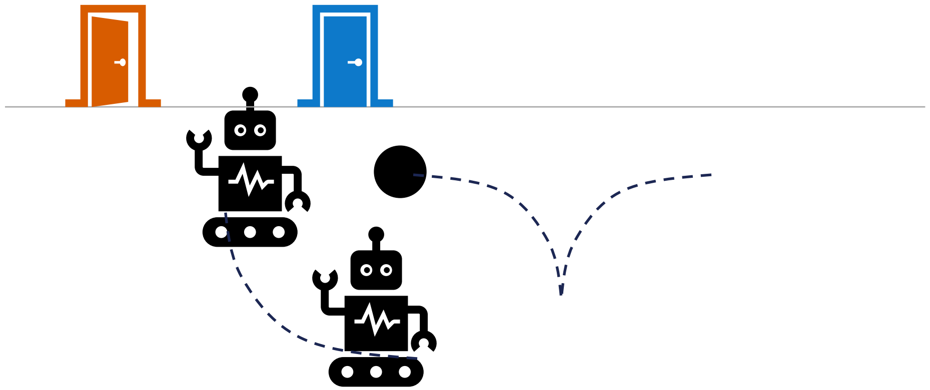
Environment can *be influenced* by robot's actions **directly**

*Human Interaction*



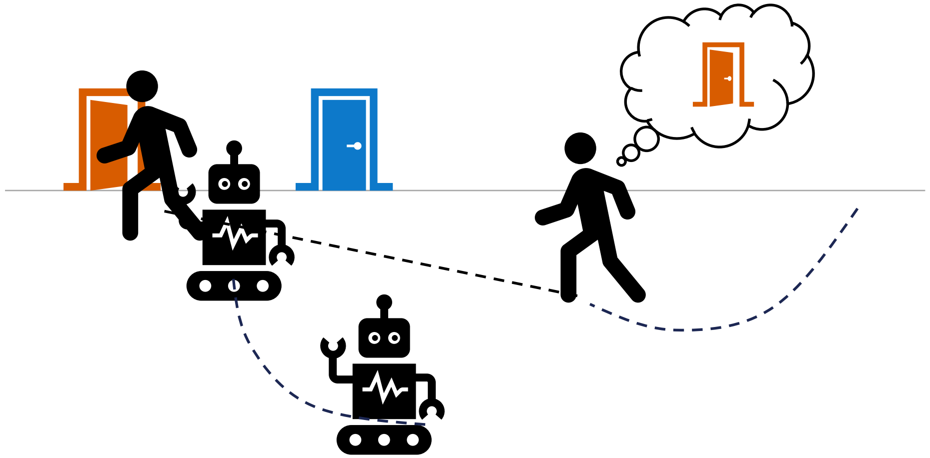
Human can *be influenced* by the robot's actions **….** and **indirectly**

*Interaction*



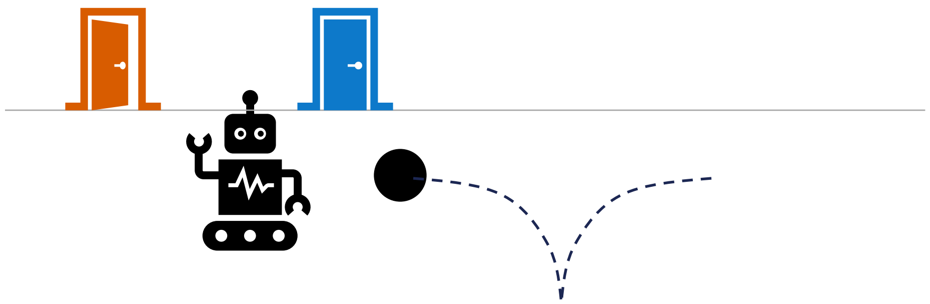
Environment can *influence* the robot's actions **indirectly**

*Human Interaction*



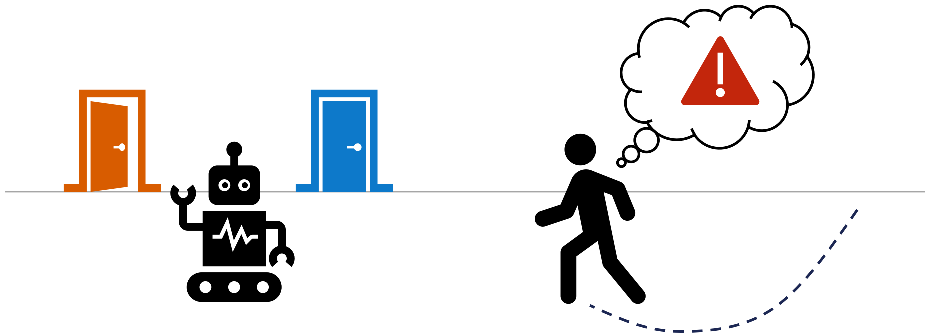
Human can *influence* the robot's behavior **directly or indirectly**

*Interaction*



"Environment" is not a stakeholder

*Human Interaction*



Human is a stakeholder! (e.g., wants to derive value from robot)

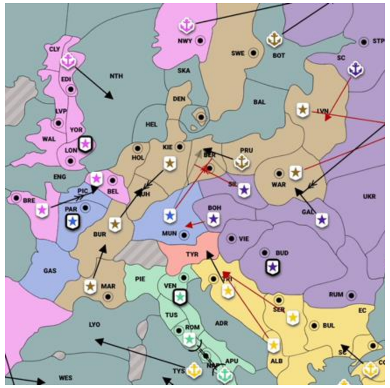
But this seems really hard to encode into our algorithms… 🤔

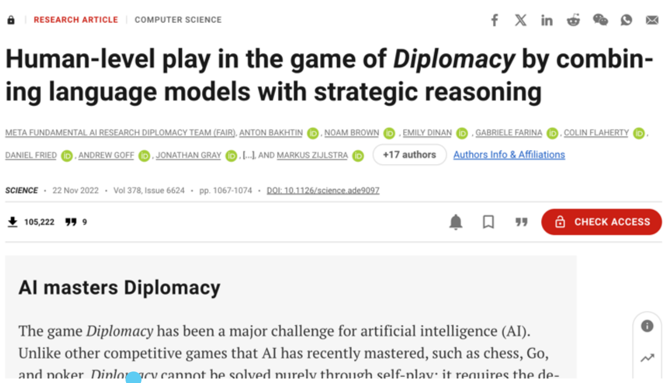
Where are people interacting with advanced autonomy the *most* right now?

# Exciting time for *interactive* Artificial Intelligence!

AlphaGo

CICERO





RESEARCH ARTICLE | COMPUTER SCIENCE

**Human-level play in the game of *Diplomacy* by combining language models with strategic reasoning**

META FUNDAMENTAL AI RESEARCH DIPLOMACY TEAM (FAIR), ANTON BAKHTIN, NOAM BROWN, EMILY DINAN, GABRIELE FARINA, COLIN FLAHERTY, DANIEL FRIED, ANDREW GOFF, JONATHAN GRAY, [...], AND MARKUS ZIJLSTRA +17 authors Authors Info & Affiliations

*SCIENCE* • 22 Nov 2022 • Vol 378, Issue 6624 • pp. 1067-1074 • DOI: 10.1126/science.ade9097

105,222 9 CHECK ACCESS

**AI masters Diplomacy**

The game *Diplomacy* has been a major challenge for artificial intelligence (AI). Unlike other competitive games that AI has recently mastered, such as chess, Go, and poker, *Diplomacy* cannot be solved purely through self-play; it requires the de-

Gemini

Supercharge your creativity and productivity

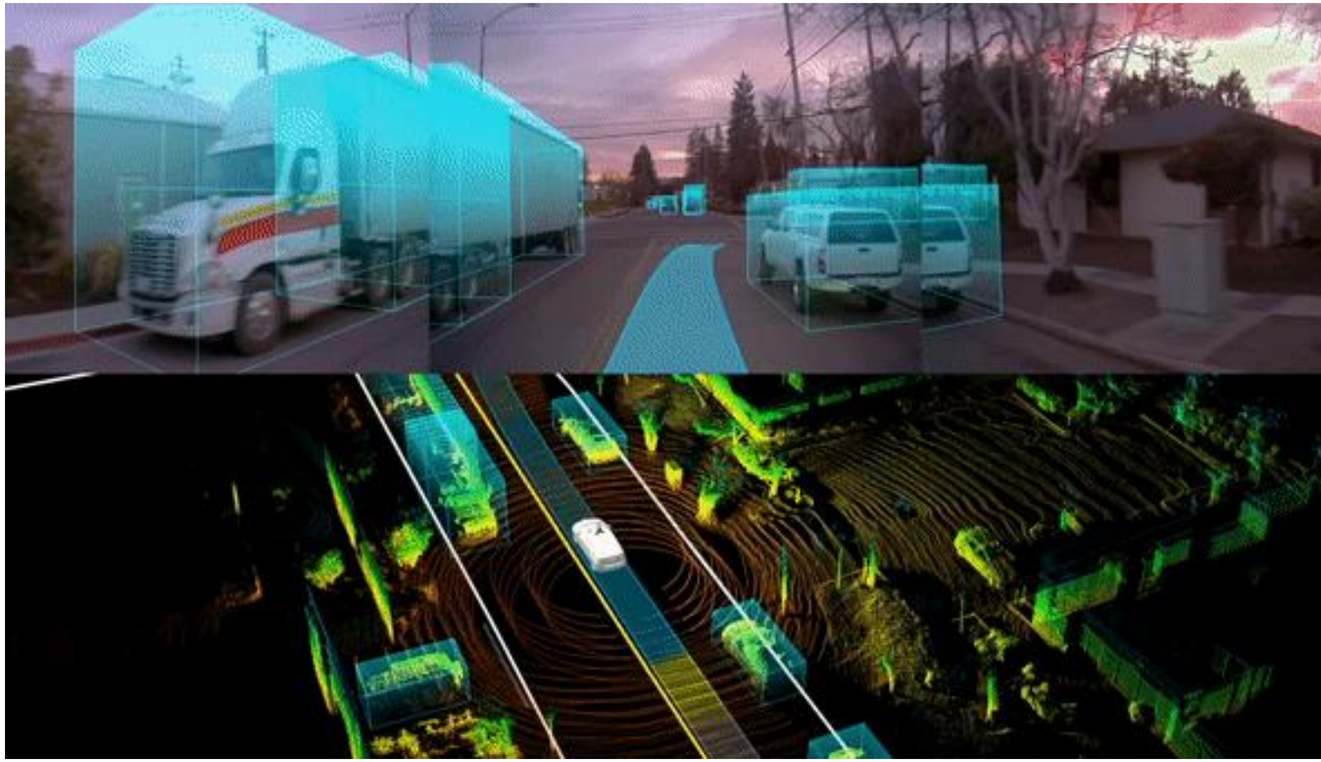
gemini.google.com


ChatGPT

2016 2022 *today*

But where are the interactive **robots**? 🤔

# How do robots interact with people today?

WAYMO

nuro

# How do robots interact with people today?


diligent

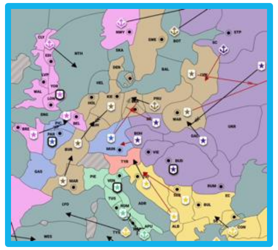
# How do robots interact with people today?


Skydio

# How do robots interact with people today?


hello robot™

# Here is where we are in AI…



Gemini

ChatGPT

*Fully robot-autonomous*

**Autonomy**

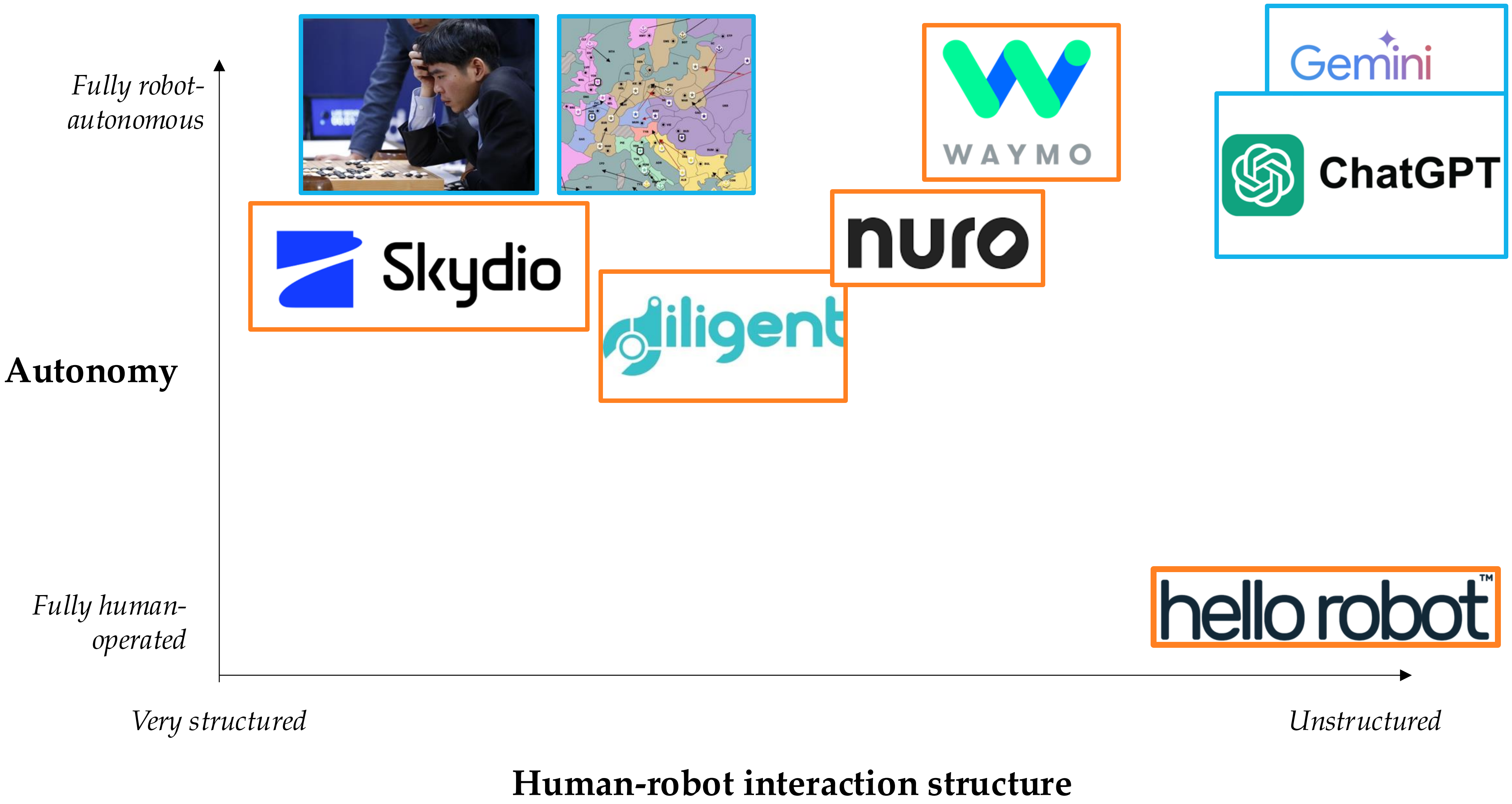
*Fully human-operated*

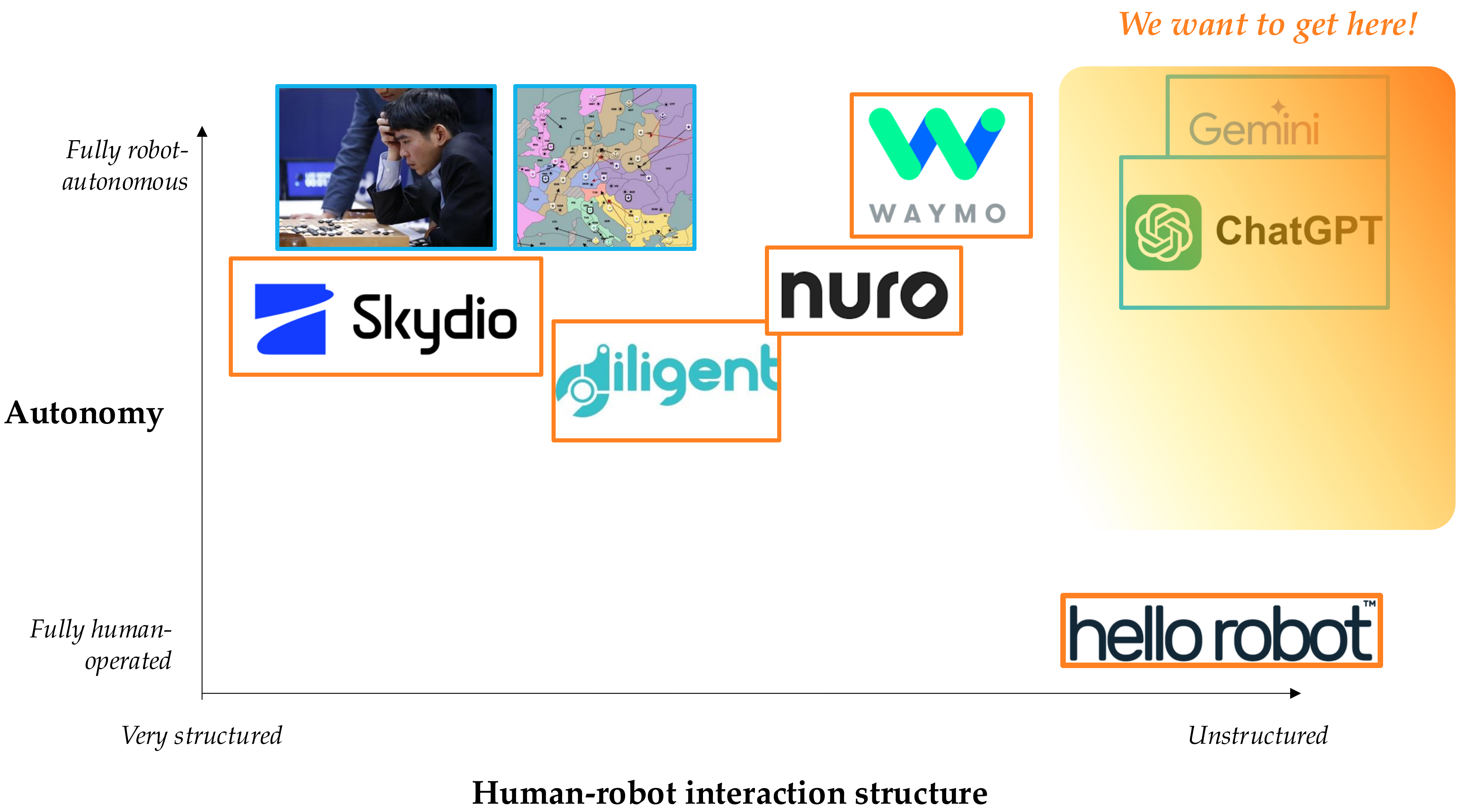
*Very structured*

*Unstructured*

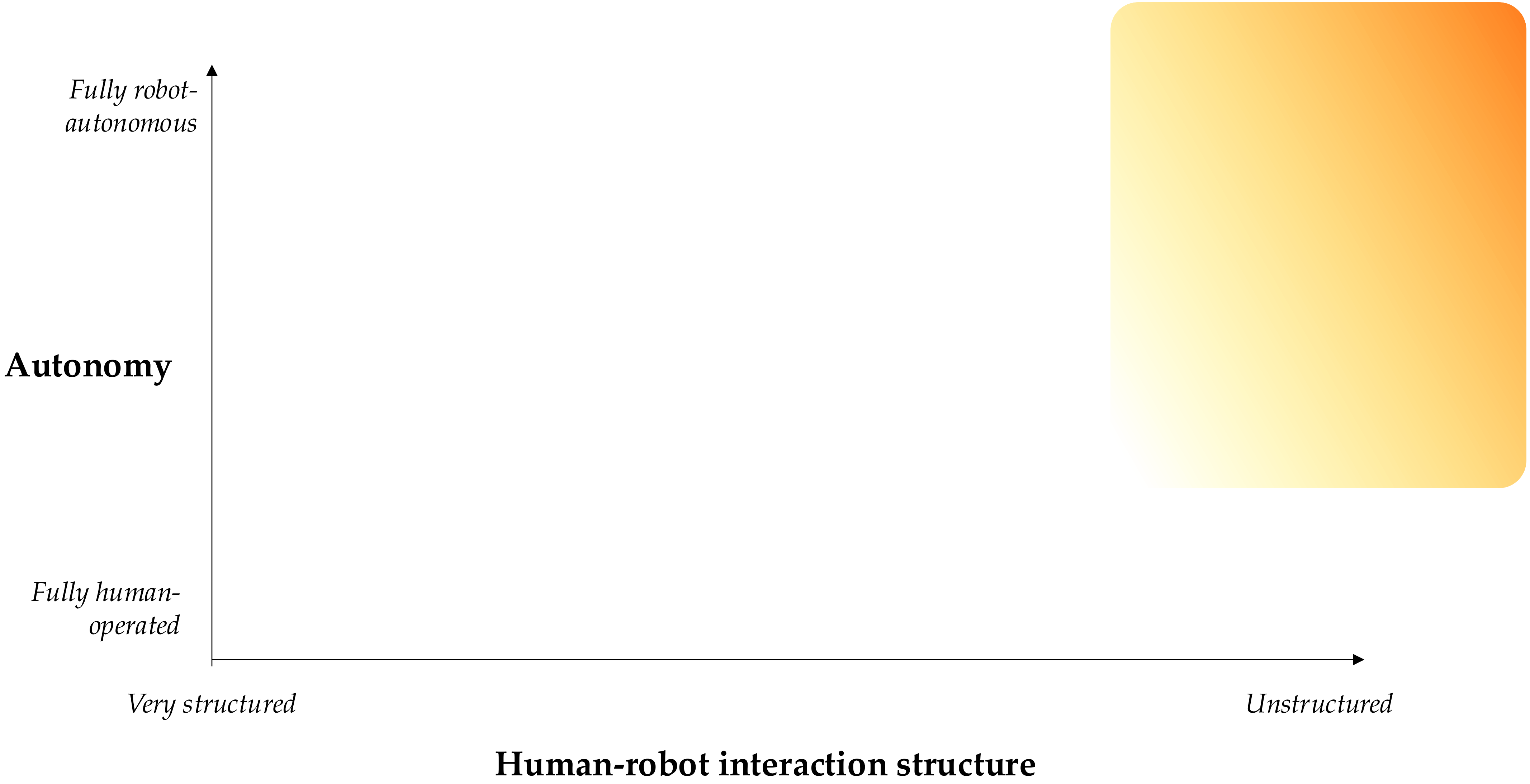
**Human-robot interaction structure**

# Here is where we are in robotics…



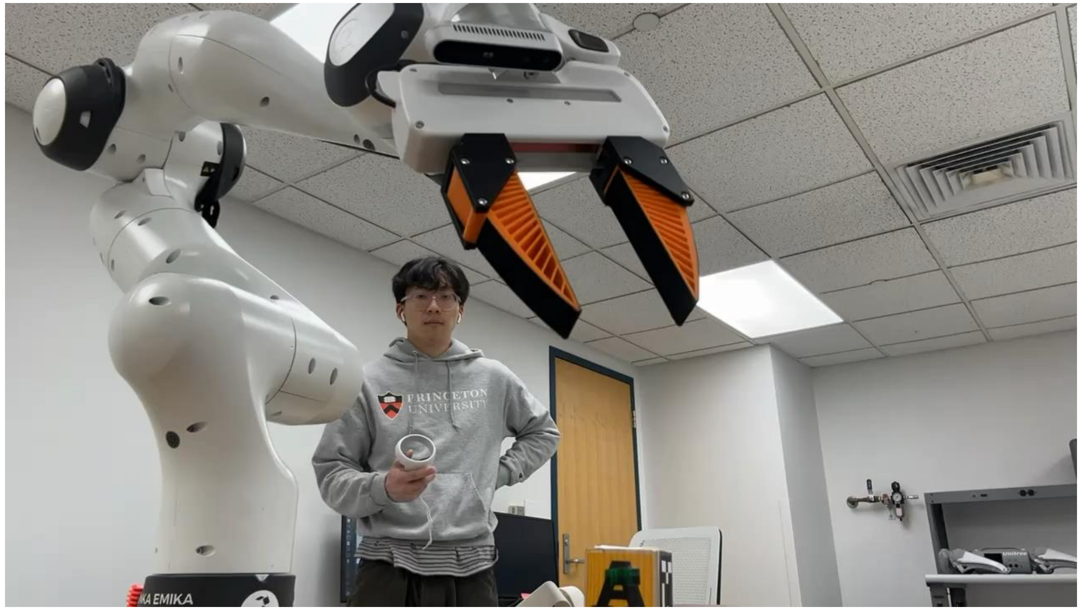

We want to get here!
Fully robot-autonomous
Autonomy
Fully human-operated
Very structured
Unstructured
Human-robot interaction structure
Skydio
diligent
nuro
WAYMO
Gemini
ChatGPT
hello robot

# **Think:** Why are robots not in millions of homes?



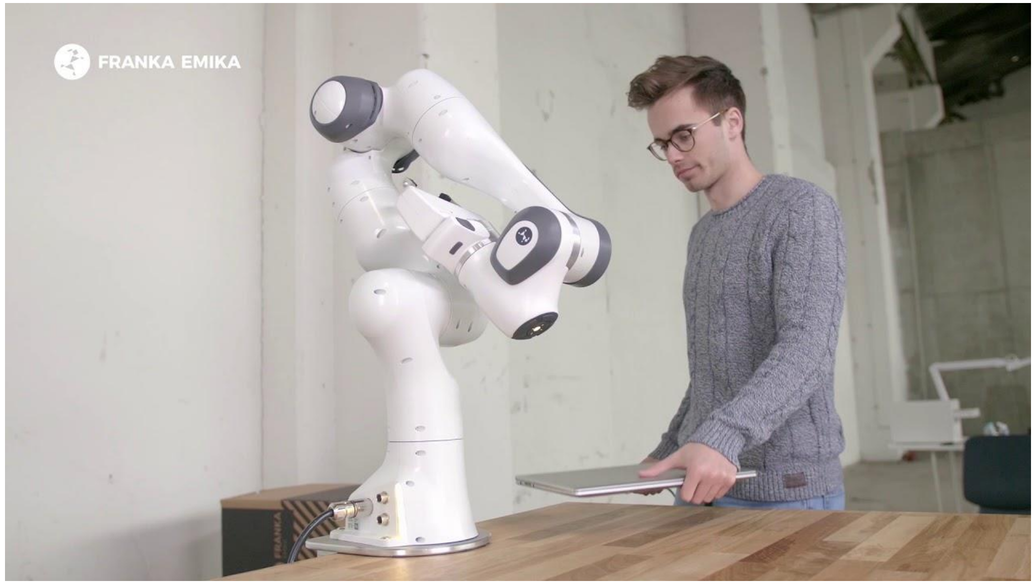
# 1. The way we program robots is rigid

*Not flexible enough to be used by everyday users for everyday tasks; requires expert knowledge*




Engineers Design Behaviors

*Ship robot* →




Users can't easily expand capabilities, or experience unexpected failures!

# 2. Hard to write down what "matters" to people

- ***Autonomy:*** *hard to design robot policies that behave according to what end-users want*
- ***Evaluation:*** *hard to write metrics that correlate with what end-users want*


"Feel the Bite: Robot-Assisted Inside-Mouth Bite Transfer using Robust Mouth Perception and Physical Interaction-Aware Control." Jenamani, et al. (2024)

# 3. Hard to model human interaction

*Human behavior is diverse: varying between individuals, environments, and over time*

# Why this course?

Take any robot application and …








1) **Model / quantify** human interaction with robots

2) **Solve** robot decision-making algorithms that are informed of/by people

3) **Identify** the frontiers of human-robot interaction

# What you will learn in this course

## Foundations

Single & multi-agent decision-making
Mathematical human models
Experimental design

## Robots *Learning* from Humans

Trajectory forecasting
Active learning
Communication

## Robots *Acting* around/with People

Shared autonomy
HRI as a Game
Safety & uncertainty quantification

# Guest Lectures

**Shared autonomy**

Dylan Losey
*Prof @ Virginia Tech*

**Robot learning from humans**

Tesca Fitzgerald
*Prof @ Yale*

# Course Logistics

**Format:** lecture + related paper reading discussions

**Typical 80-min class:**

~5 min logistics and recap

70 min lecture, invited talk, or paper discussion

# Use ***course website*** for up-to-date schedule & paper links

https://abajcsy.github.io/human-robot-interaction/

## 16-867: Human-Robot Interaction

Fall 2024

**Professor:** Andrea Bajcsy (abajcsy [at] cmu [dot] edu)
**Office Hours:** TBD
**Office Hours Location:** NSH 4629

**Lecture Time:** Tues & Thurs, 11:00 - 12:20 pm
**Lecture Location:** Wean 4623

**Teaching Assistant:** Pranay Gupta (pranaygu [at] andrew [dot] cmu [dot] edu)
**Office Hours:** XYZ
**Office Hours Location:** NSH XYZ

**Syllabus:** PDF
**Canvas:** https://canvas.cmu.edu/courses/41578

### Overview

Human-robot interaction (HRI) is a multidisciplinary field that aims to create successful interactions between people and robots. In this class, we will study algorithmic HRI topics such as mathematical human models, trajectory forecasting, shared autonomy, robot learning from human feedback, active learning, communication, and safety.

This course aims to provide an overview of the state of the art in algorithmic HRI. As such, it will cover a large number of topics, with examples drawn from foundational work and research published in the last five years. The course combines lecture, readings, in-class presentations, written reports, and a final project to engage students with the current challenges and approaches in the field. The course also emphasizes the practice of reading and discussing scientific literature to learn and communicate about the most recent progress in HRI.

### News

- [08/19/24] New room location: Wean 4623

### Schedule (tentative)

| Date | Topic | Info |
|---|---|---|
| **Week 1** Tue, Aug 27 | Lecture Introduction | • *Please check the course syllabus* |
| **Week 1** Thurs, Aug 29 | Lecture Fundamentals | Single-Agent Decision Making |
| **Week 2** Tue, Sept | Lecture Fundamentals | Probability, Entropy, Bayesian inference |
| **Week 3** Tue, Sept 10 | Lecture Mathematical Human Models | Internal state, bounded rationality, suboptimality |
| **Week 3** Thurs, Sept 12 | Paper discussion Mathematical Human Models | **Required Reading:**<br>• [P1] Where Do You Think You're Going?: Inferring Beliefs about Dynamics from Behavior. Reddy, et al. (2018)<br>• [P2] LESS is More: Rethinking Probabilistic Models of Human Behavior. Bobu, et al. (2020)<br>• [P3] The Boltzmann Policy Distribution: Accounting for Systematic Suboptimality in Human Models. Laidlaw & Dragan (2022). |
| **Week 4** Tue, Sept 17 | Lecture Trajectory Forecasting | Planning-based & learning-based; applications in manipulation, navigation |
| **Week 4** Thurs, Sept 19 | Paper discussion Trajectory Forecasting | **Required reading:**<br>• [P1] Probabilistically Safe Robot Planning with Confidence-Based Human Predictions. Bajcsy, et al. (2018)<br>• [P2] Identifying Driver Interactions via Conditional Behavior Prediction. Tolstaya, et al. (2021)<br>• [P3] ManiCast: Collaborative Manipulation with Cost-Aware Human Forecasting. Kedia, et al. (2023) |
| **Week 5** Tues, Sept 24 | Guest Lecture Shared Autonomy | **Dylan Losey** (Prof @ Virginia Tech) |
| **Week 5** Thurs, Sept 26 | Paper discussion Shared Autonomy | **Required reading:**<br>• [P1] Shared Autonomy via Hindsight Optimization. Javdani, et al. (2015)<br>• [P2] Shared Autonomy via Deep Reinforcement Learning. Reddy, et al. (2018)<br>• [P3] LILA: Language-Informed Latent Actions. Karamcheti, et al. (2021) |
| **Week 6** Tue, Oct 8 | Lecture Experimental Design | Due **Homework**<br>Designing and conducting user studies |
| **Week 6** Thurs, Oct 10 | Paper discussion Experimental Design | **Required Reading:**<br>• [P1] Review of Human Studies Methods in HRI and Recommendations. Bethel & Murphy (2010)<br>• [P2] Feel the Bite: Robot-Assisted Inside-Mouth Bite Transfer using Robust Mouth Perception and Physical Interaction-Aware Control. Jenamani, et al. (2024)<br>• [P3] Independence in the Home: A Wearable Interface for a Person with Qadriplegia to Teleoperate a Mobile Manipulator. Padmanabha, et al. (2024) |
| **Week 7** Tue, Oct 14 | *No Class (Fall Break)* | |

# Use *Canvas* for downloading / uploading assignments

16867-A

*Fall 2024*

- **Home**
- Announcements
- Syllabus
- Assignments
- Quizzes
- Grades
- Discussions
- Files
- People
- Zoom
- NameCoach
- Syllabus Registry
- Pages
- Outcomes
- Collaborations

## Recent Announcements

### Human Robot Interaction

Assign To | Edit

Welcome to **16-867: Human-Robot Interaction!**

Human-robot interaction (HRI) is a multidisciplinary field that aims to create successful interactions between people and robots. n this class, we will study algorithmic HRI topics such as mathematical human models, trajectory forecasting, shared autonomy, robot learning from human feedback, active learning, communication, and safety.

This course aims to provide an overview of the state of the art in algorithmic HRI. As such, it will cover a large number of topics, with examples drawn from research published in the last five years. The course combines lecture, readings, in-class presentations, written reports, and a final project to engage students with the current challenges and approaches

# Grading

*See **class syllabus on course website** for detailed info*

| | |
|---|---|
| Participation | (5%) |
| Homework  (x1) | (10%) |
| Paper summaries | (10%) |
| Paper presentations | (15%) |
| Class project | (60%) |

## Participation (5%)

Expected to attend class in person—this is how we will all get the most out of the class!

I understand that occasionally you may have challenges attending (e.g., illness, religious observance,..); let me know

<u>Please show up on time</u>, especially for reading days

# Homework (10%)

16-867 Human-Robot Interaction (Fall 2024) — Prof. Andrea Bajcsy

### Homework 1: Learning from Demonstration

In this homework, will walk through Maximum Entropy inverse reinforcement learning, intent inference, and intent expression in a simple grid-world environment. For programming, you should use the code provided in `hw1_code.zip` which is compatible with `Python` and uses `Jupypter Notebooks`. **The notebook itself contains details of each question and the code that you need to fill out and submit**. This document summarizes the key problems you will implement in the Jupyter notebook.

**Due Week 6**
(Tue, Oct 8)

This is a coding-based homework in **Python**.
It is *not* meant to be tedious; it is meant to **empower** you! ☺

# Paper Summaries + Presentations (25%)

**Paper discussion days:**

~10 paper reading days
3 papers per reading day

**Before class:**

write 1-2 paragraphs of paper review / takeaway / questions (must submit on Canvas)

**In class:**

Split you into small groups, discuss set of questions, I assign a representative from each group to present on the group's takeaways, and the whole class can engage on the answer

# On paper reviews

Be **compassionate** *(e.g., invert your position)*

Be **constructive** *(e.g., what would you change to improve it?)*

Be **scholarly** *(e.g., cite sources, justify disagreements with proofs or citations)*

Daniel Dennett
*Professor, Philosopher*

"You should attempt to re-express your target's position so *clearly*, *vividly*, and *fairly* that your target says,

'Thanks, I wish I'd thought of putting it that way.' "

Slide adapted from Dorsa Sadigh

# Class Project (60%)

Two options:

**Research project:**

Identify a research direction broadly relevant to this class

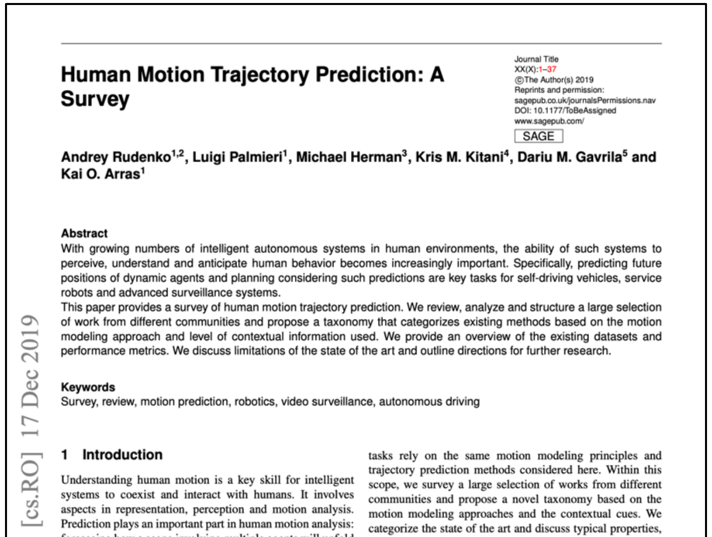
Propose and take first steps towards an original idea

**Literature survey:**

Select a topic area and rigorous way in which you will find papers

Characterize this topic area in an insightful way (e.g., open questions, common assumptions, tractable vs. theoretical gaps)

*Example of good literature survey*



**Human Motion Trajectory Prediction: A Survey**

Journal Title
XX(X):1–37

Reprints and permission:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/ToBeAssigned
www.sagepub.com/
SAGE

Andrey Rudenko[1,2], Luigi Palmieri[1], Michael Herman[3], Kris M. Kitani[4], Dariu M. Gavrila[5] and Kai O. Arras[1]

**Abstract**

With growing numbers of intelligent autonomous systems in human environments, the ability of such systems to perceive, understand and anticipate human behavior becomes increasingly important. Specifically, predicting future positions of dynamic agents and planning considering such predictions are key tasks for self-driving vehicles, service robots and advanced surveillance systems.

This paper provides a survey of human motion trajectory prediction. We review, analyze and structure a large selection of work from different communities and propose a taxonomy that categorizes existing methods based on the motion modeling approach and level of contextual information used. We provide an overview of the existing datasets and performance metrics. We discuss limitations of the state of the art and outline directions for further research.

**Keywords**

Survey, review, motion prediction, robotics, video surveillance, autonomous driving



**1 Introduction**

Understanding human motion is a key skill for intelligent systems to coexist and interact with humans. It involves aspects in representation, perception and motion analysis. Prediction plays an important part in human motion analysis:

tasks rely on the same motion modeling principles and trajectory prediction methods considered here. Within this scope, we survey a large selection of works from different communities and propose a novel taxonomy based on the motion modeling approaches and the contextual cues. We categorize the state of the art and discuss typical properties,

# Class Project (60%)

**Project proposal (0%) -- due on Tues, Sept 24**

~1 page project summary. Identify the problem, background literature, potential solution

**Mid-term report (20%) -- due on Tue, Oct 29**

~2 page writeup of progress, updated goals and timeline

**Oral project presentation (10%) -- to be scheduled for Dec. 3 & Dec. 5**

short presentations (~10 minutes but depends on number of people)

**Final project report (30%) -- due on Dec. 10**

~6 pages final report

# Survey (5 min)

https://forms.gle/nwAoLvneinkL14Cz8

*16-867*

# Human Robot Interaction

## Introduction

Instructor: Andrea Bajcsy

Carnegie Mellon University

intent ROBOTICS LAB