(A gentle) Introduction to Human-Robot Interaction

Ivan Marković

University of Zagreb Faculty of Electrical Engineering and Computing
Department of Control and Computer Engineering
Laboratory for Autonomous Systems and Mobile Robotics (lamor.fer.hr)

Outline

- What is Human-Robot Interaction?
- The evolution of HRI
- Design in HRI
  - Principles
  - Anthropomorphisation
- Human intention estimation
  - Bayesian Theory of Mind
- HRI applications
- Examples from FER’s robotic labs
  - LABUST, LARICS, LAMOR

Outline

- What is Human-Robot Interaction?
- The evolution of HRI
- Design in HRI
    - Principles
    - Anthropomorphisation
- Human intention estimation
    - Bayesian Theory of Mind
- HRI applications
- Examples from FER's robotic labs
    - LABUST, LARICS, LAMOR

Motivation

Expectations about robots and their roles in everyday life are quite high, and can vary a lot from person to person – cooking, laundry, iron and fold clothes, elderly assistants, good conversationalist.

The prowess of robots as assistants; however, is still rather limited.

Moravec's paradox, decades after being first expressed, still holds:

> „Anything that seems hard to people is relatively easy for machines, and anything a young child can do is almost impossible for a machine."

What is Human-robot interaction?

Human-robot interaction is a research field dedicated to understanding, designing and evaluating robot systems for use by or with human.

HRI is related to human-computer interaction (HCI), robotics, artificial intelligence, the philosophy of technology, and design.

What makes HRI unique? - the interaction of humans with social robots.

Robot definition

„A robot is an autonomous system which exists in the physical world, can sense its environment, and can act on it to achieve some goals[10].”

Note that this definition excludes teleoperated machines (no autonomy), AI bots (not in the physical worlds), simulations (no real environment sensing) and various other machines like sophisticated cookware and utensils.

Taking actions to respond to sensory inputs and to achieve what is desired is a necessary part of being a robot.

A system or machine that exists in the physical world and senses it, but acts randomly or uselessly, is not much of a robot.

[10] Matarić, Maja (2007). The robotics primer.

What is Human-robot interaction?

„Social robots are embodied agents that are part of a heterogeneous group: a society of robots or humans. They are able to recognize each other and engage in social interactions, they possess histories (perceive and interpret the world in terms of their own experience), and they explicitly communicate with and learn from each other."

K. Dautenhahn, A. Billard, Bringing up robots or—the psychology of socially intelligent robots: From theory to implementation, in: Proceedings of the Autonomous Agents, 1999.

What is Human-robot interaction?

Social robots are perceived as social actors bearing cultural meaning and having a strong impact on contemporary and future societies.

Saying that a robot is embodied does not mean that it is simply a computer on legs or wheels.

We have to understand how to design that embodiment, both in terms of software and hardware and in terms of its effects on people and the kinds of interactions they can have with such a robot.

Outline

- What is Human-Robot Interaction?
- The evolution of HRI
- Design in HRI
    - Principles
    - Anthropomorphisation
- Human intention estimation
    - Bayesian Theory of Mind
- HRI applications
- Examples from FER's robotic labs
    - LABUST, LARICS, LAMOR

The evolution of HRI

Grey Walter’s tortoise (1949)

The evolution of HRI

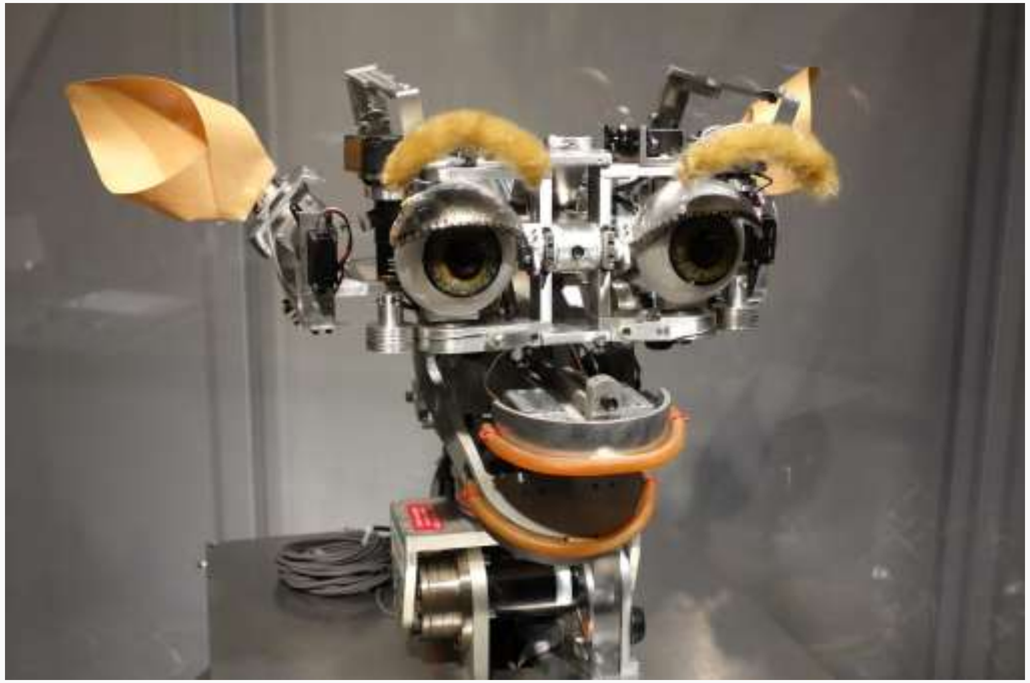
MIT Media Laboratory
Braitenbert Creatures (1980s)

The evolution of HRI

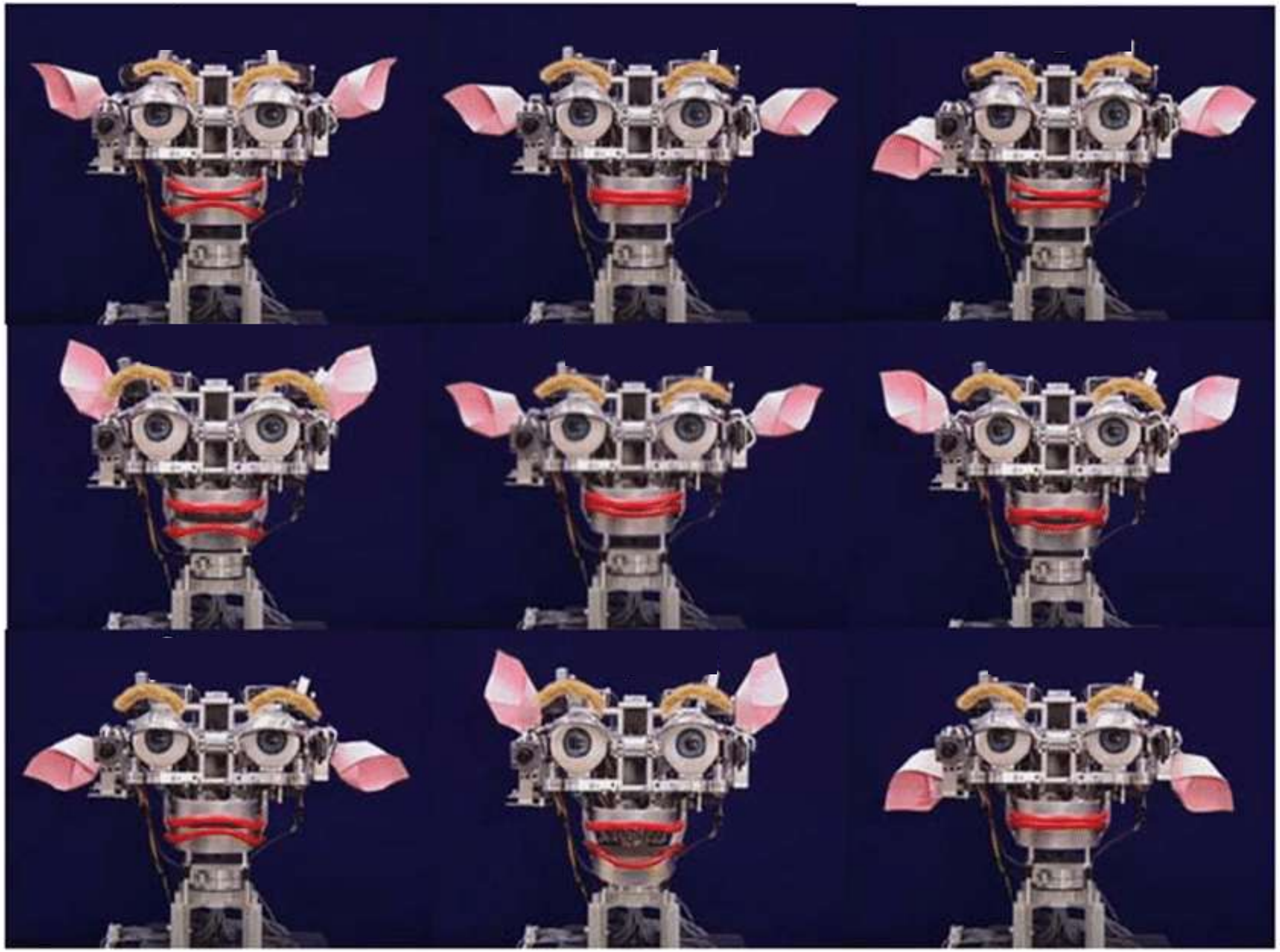
Kismet is an early example of a social robot developed at MIT in 1997.

Kismet was surprisingly effective at presenting a social presence.

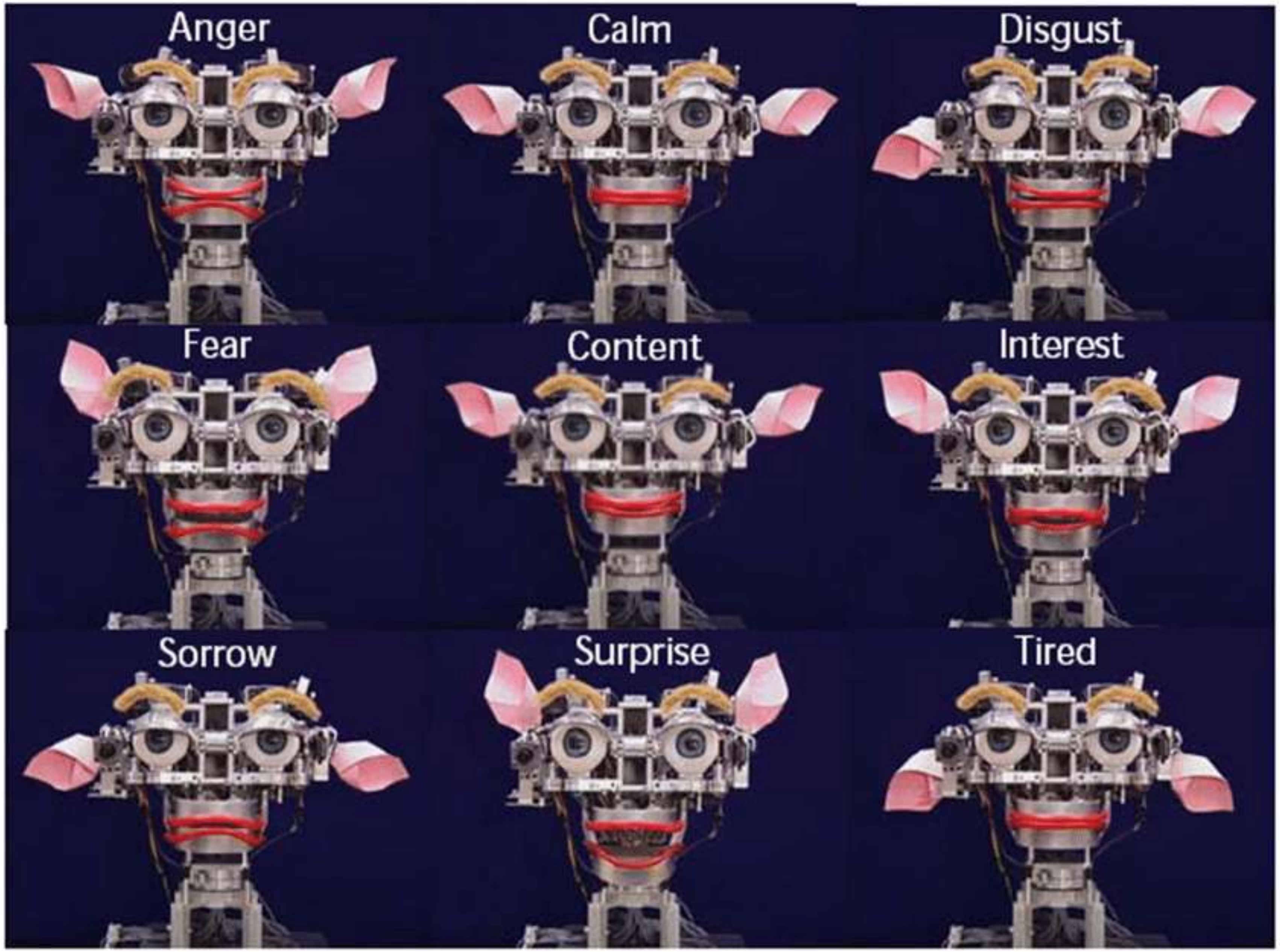
It included what is known as the „baby schema," a predisposition to treat things with big eyes and exaggerated features in social ways despite their lack of fully functional social skills.



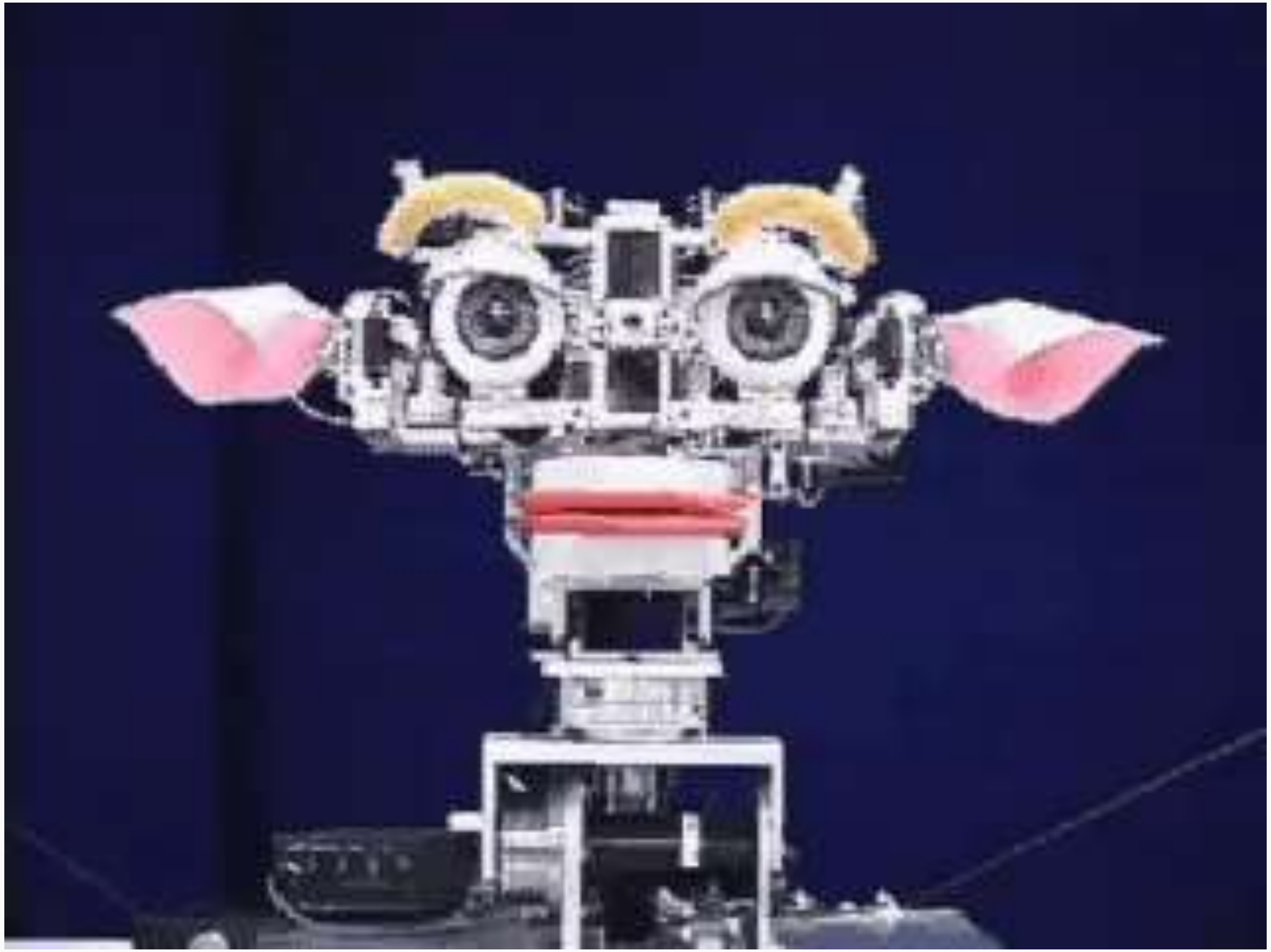
The evolution of HRI



The evolution of HRI



The evolution of HRI



The evolution of HRI

Aldebaran's (Softbank) robots are perhaps the most influential robots in the study of social robotics, especially the Nao robot.

Nao (2008)

Pepper (2014)

Romeo (2014)

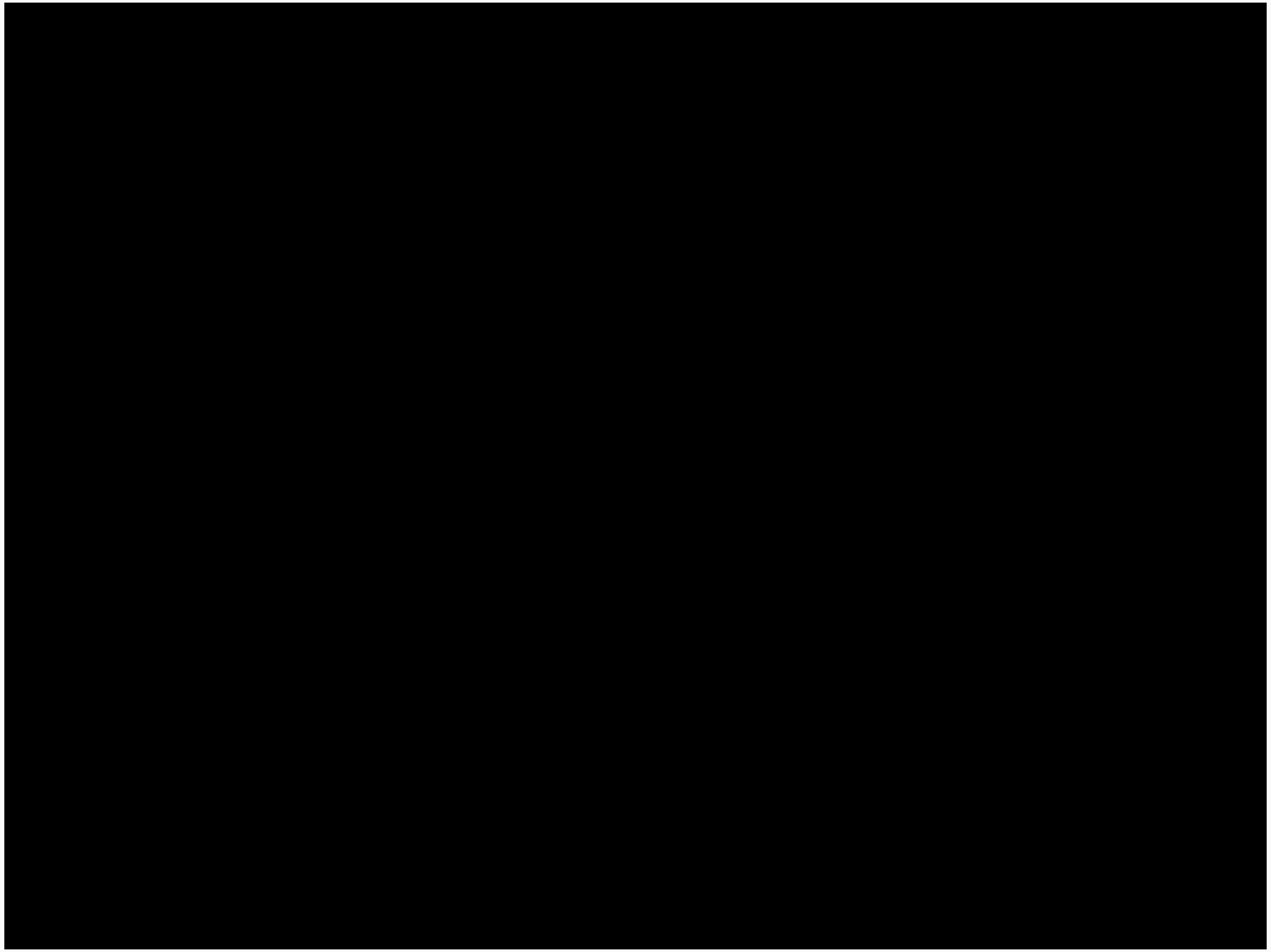
The evolution of HRI

The Keepon robot was developed by Hideki Kozima of Tohoku University, Japan in 2009.

Studies with the Keepon robot convincingly demonstrated that a social robot does not need to appear humanlike.

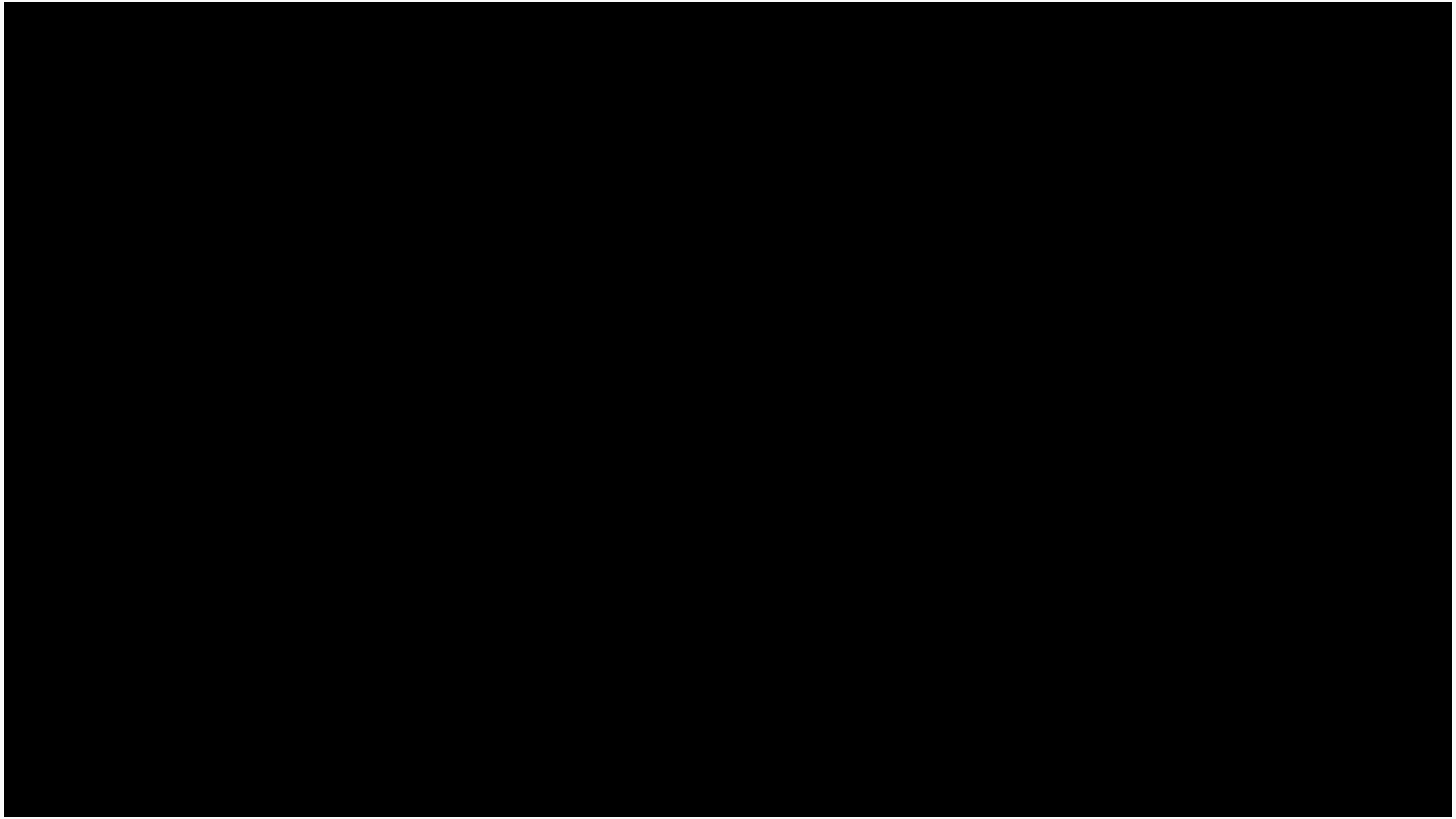
The simple form of the robot is sufficient to achieve interaction outcomes where one might assume the need for more complex and humanlike robots.

The evolution of HRI



The Paro companion and therapy robot has been particularly popular in the study of socially assistive robots in eldercare, since 2006.

The evolution of HRI

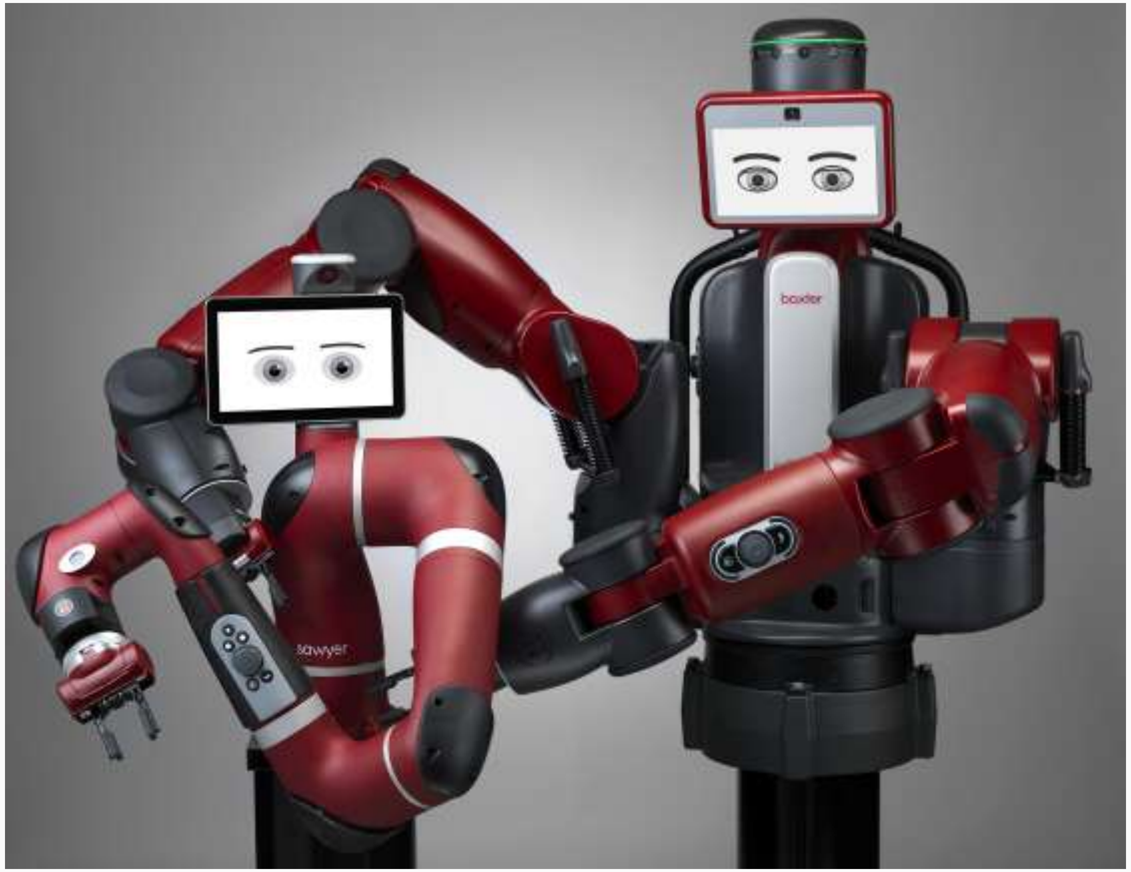


The evolution of HRI

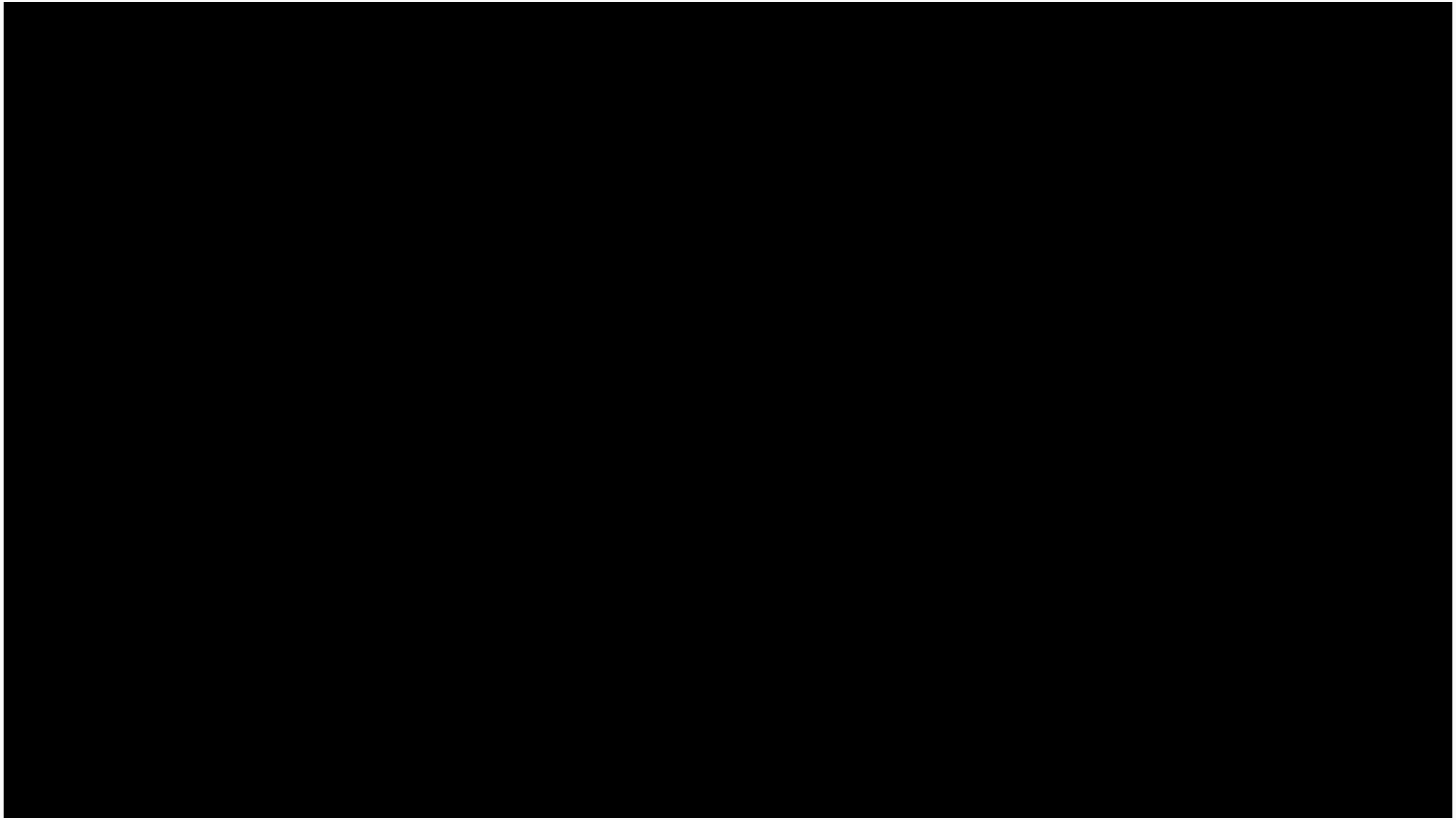
The Baxter robot (2011) of Rethink Robotics is both an industrial robot and an HRI platform.

The robot's two arms are actively compliant: in contrast to the stiff robot arms of typical industrial robots, Baxter's arms move in response to an externally applied force.

Suitable for collaborative tasks.



The evolution of HRI



Outline

- What is Human-Robot Interaction?
- The evolution of HRI
- Design in HRI
  - Principles
  - Anthropomorphisation
- Human intention estimation
  - Bayesian Theory of Mind
- HRI applications
- Examples from FER’s robotic labs
  - LABUST, LARICS, LAMOR

Design in HRI

„How does a pile of wires, motors, sensors, and microcontrollers turn into a robot that people will want to interact with?"

Design in HRI

Usually, we first think of what the robot is going to be doing.

However, in HRI, form and function are inherently interconnected and thus cannot be considered separately.

Contemporary HRI designers can choose from:

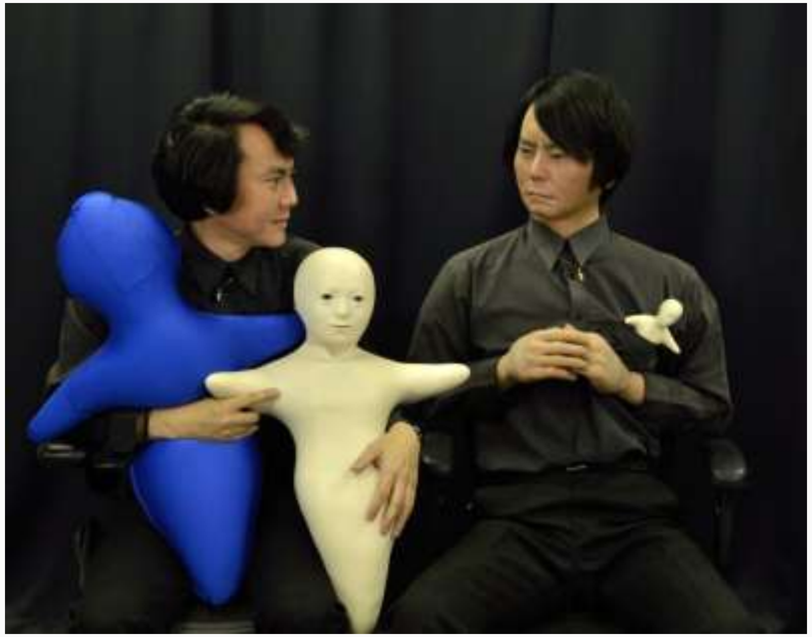
1. Androids and humanoids
2. Zoomorphic robots
3. Minimal robots for social HRI
4. Object-based interactive form

Form and function


Muu Socia, robot-patient communication facilitator

Partridge, M., & Bartneck, C. (2013). The Invisible Naked Guy: An exploration of a minimalistic robot.

Michio Okada, Toyohashi University of Technology



Hiroshi Ishiguro

Form and function

Another important concept in HRI design is the notion of affordances[1].

Affordances are perceivable relationships between an organism and its environment that enable certain actions (e.g., a chair is something to sit on, but so is a stair).

Affordances need to be explicit and one needs to incorporate user expectations and cultural perceptions.

Appearance is an important affordance because people tend to assume that the robot's capabilities will be commensurate with its appearance (eyes – can see, arms – can manipulate, say „Hello" – can converse, has facial expressions – can read emotions).

[1] Norman, Don (2008), The Design of Everyday Things

Design principles in HRI

1. Matching the form and function of the design:
   - If your robot is humanoid people will expect it to do humanlike things.
   - If this is not necessary for its purpose, such as cleaning, it might be better to stick to less anthropomorphic designs.
   - People can also be prompted to associate specific social norms and cultural stereotypes with robots through design[3]
2. Underpromise and overdeliver:
   - When people's expectations are raised by a robot's appearance and those expectations are not met by its functionality, people get disappointed and negatively evaluate the robot
   - it is better to decrease people's expectations about robots which might have been increased by how robots are portrayed in society[4]

[3] Powers, Aaron et al. (2005). Eliciting information from people with a gendered humanoid robot
[4] Paepcke, Steffi and Takayama, Leila (2010). Judging a bot by its cover: An experiment on expectation setting for personal robots

Design principles in HRI

3. Interaction expands function:
    - When confronted with a robot, people will fill in the blanks left open by the design
    - It can thus be useful for robots with limited capabilities to design them in an open-ended way
    - Worked well with the seal-like robot Paro, invokes associations with pets that people have had, but it also does not get compared to animals they know (cats and dogs, which could lead to disappointment)
    - Paro passed as a pet-like character even though its capabilities are significantly below those of a typical domestic animal or that of an actual seal baby[5]
4. Do not mix metaphors:
    - robot's capabilities, behaviors, affordances for interaction should all be coordinated
    - if the robot is an animal, it may be strange for it to talk like an adult human
    - inappropriately matched abilities, behaviors, and appearance often lead to people having a negative impression of the robot

[5] Chang, Wan-Ling and Sabanovic, Selma. Interaction expands function: Social shaping of the therapeutic robot PARO in a nursing home.

Design principles in HRI

Anthropomorphization

Anthropomorphization is the attribution of human traits, emotions, or intentions to nonhuman entities.

Derives from ánthrōpos (meaning „human”) and morphē (meaning „form”).

A robot's level of anthropomorphism is one of the main design decisions because it influences not only the robot's appearance but also the functionality it needs to deliver.

Anthropomorphization

To better understand the nature of anthropomorphism, three core factors have been suggested that determine anthropomorphic inferences about nonhuman entities[6].

1. Effectance motivation:
   - concerns our desire to explain and understand the behavior of others as social actors; activated when people are unsure about how to deal with an unfamiliar interaction partner
   - People attribute humanlike characteristics to robots to psychologically regain control over the novel situation they find themselves in – reduces stress associated with human-robot interaction
2. Sociality motivation:
   - people may turn to nonhuman entities as social interaction partners to address their feelings of situational or chronic loneliness
   - Lonely people anthropomorphize robots to a greater extent than people who are sufficiently socially connected

[6] Epley, Nicholas. On seeing human: A threefactor theory of anthropomorphism.

Anthropomorphization

3. Elicited agent knowledge:
    - the way in which people use their commonsense understanding of social interactions and actors to understand robots
    - people have been found to expect different knowledge from female and male robots

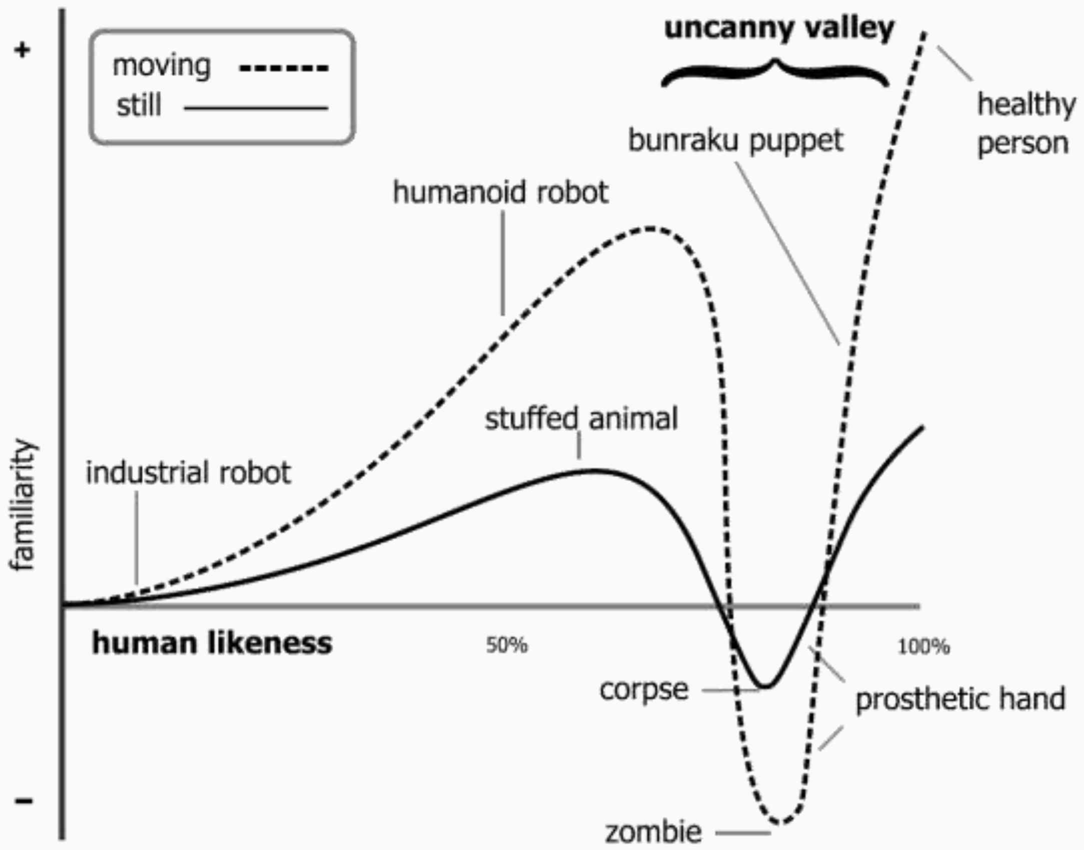
The three factors shed light on the psychological mechanisms underlying why we tend to humanize nonhuman entities.

The basic assumption is that people use self-related or anthropocentric knowledge structures to make sense of the nonhuman things.

Anthropomorphization

Uncanny valley[7] - the more humanlike robots become, the more likable they will be, until a point where they are almost indistinguishable from humans, at which point their likability decreases dramatically.

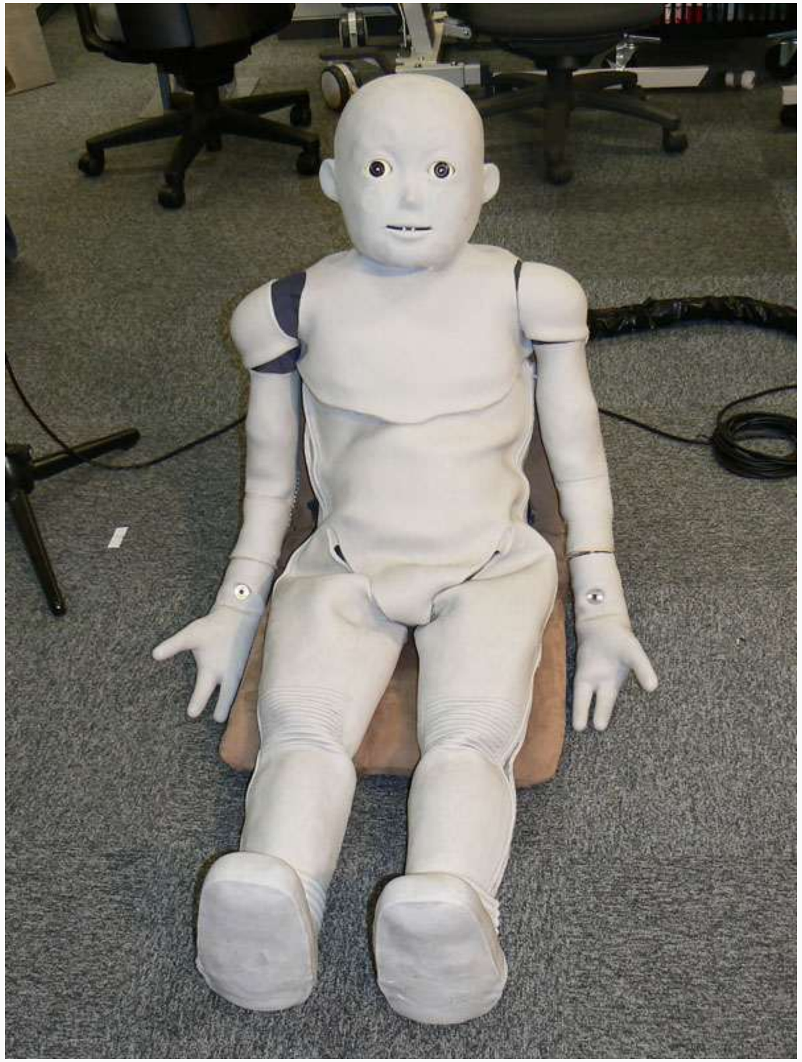
This effect is then amplified by the ability of the robot to move.

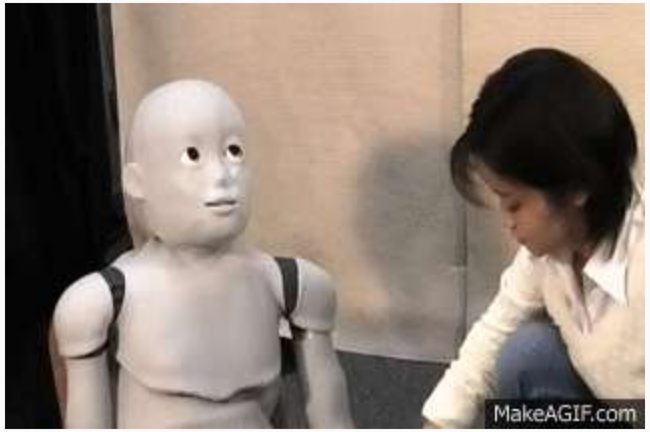


[7] Mori, Masahiro (1970). The Uncanny Valley.

Anthropomorphization

CB^2 robot (2006), Osaka University





Anthropomorphization

The following IEEE website has a ranking of robots where users can vote on how „creepy" the robot is, and the top 5 are:

The bottom 5:

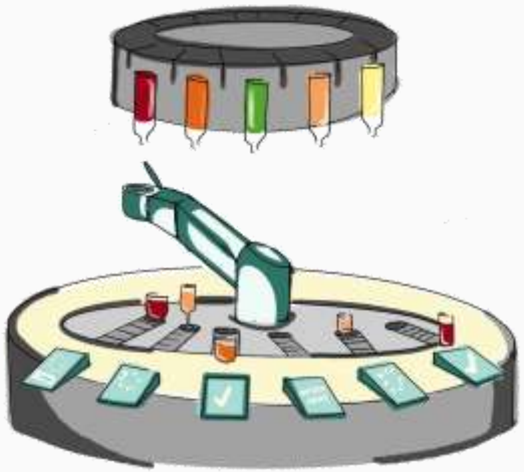
Robotic bartender

1)

2)

3)

4)



5)

6)

Outline

- Motivation
- What is Human-Robot Interaction?
- The evolution of HRI
- Design in HRI
  - Principles
  - Anthropomorphisation
- Human intention estimation
  - Bayesian Theory of Mind
- HRI applications
- Examples from FER's robotic labs
  - LABUST, LARICS, LAMOR

Human intention estimation (Theory of Mind)

Another approach to achieving natural behavior in robots is to endow them with ability to estimate human intention – instead of designing interaction responses meticulously by hand.

Theory of Mind is the ability to read desires, goals, and intentions in others. It is essential in understanding what others are thinking and what they are about to do – can be used as a tool for estimating human intentions.

E.g., the false-belief task: children must infer that another person does not possess knowledge that they possess (chocolate switch). Children usually at age 4 can give a correct answer and perform better if asked by robot – the child is acting as a teacher[15].

[15] Baratgin, Jean et al. (2020). Pragmatics in the False-Belief Task: Let the Robot Ask the Question!

Bayesian Theory of Mind

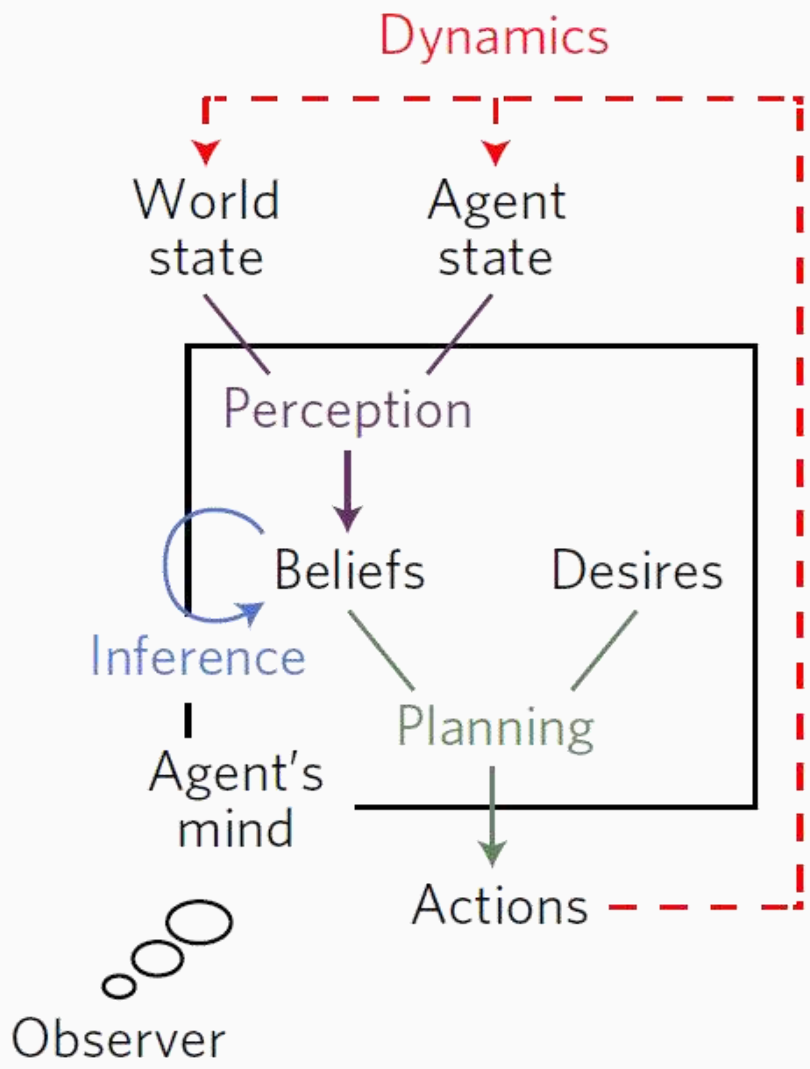
To computationally implement a ToM, researchers introduced the Bayesian ToM[15] - a model that can be seen as an approximation to ToM.

The model assumes that actions are generated via a rational planning process aiming to fulfil the agent's current desire at the lowest possible cost in the world depicted by its beliefs.

Inferring the agent's mental states is a kind of inverse planning: finding a set of causal factors (in this case, beliefs and desires) that would produce the observed action as their effect – ill-defined problem.

[15] Baker, Chris L. et al. (2017). „Rational quantitative attribution of beliefs, desires and percepts in human mentalizing."

Bayesian Theory of Mind



Bayesian Theory of Mind

Mathematically this can be captured by the Bayes' rule.

Essentially, we are looking to estimate the posterior distribution over agent's beliefs (B), desires (D), perception (P) and situation (S) given the observed agent actions (A) and the prior distribution $P(B_0, D, S)$

Likelihood factored into three components

$$P(B, D, P, S \mid A) \propto P(A \mid B_1, D) P(B_1 \mid P, B_0) P(P \mid S) P(B_0, D, S)$$

Observer's model of the agent's planning process, offers to capture how agent deviate's from the ideal; agent chooses the action in line with highest cost-effectivness w.r.t. to desires

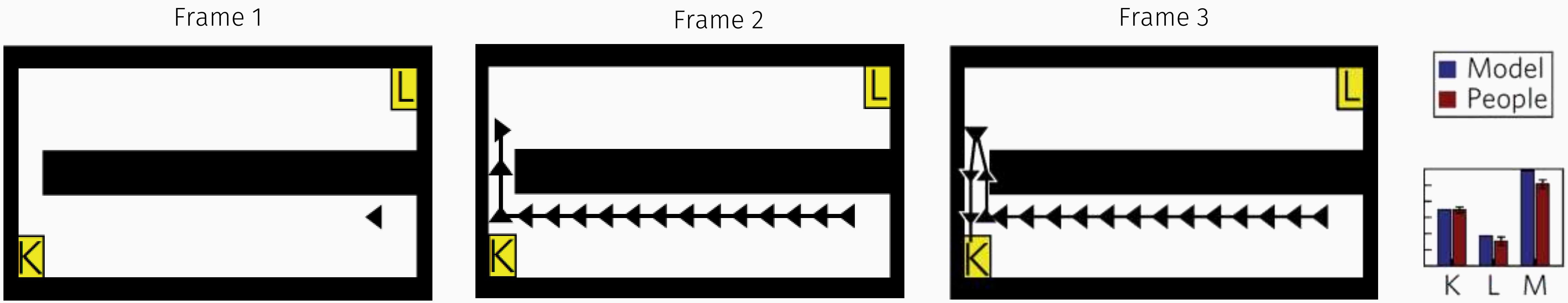
Observer's expectations about what the agent sees in a given situation, model of the agent's (visual) perception

Observer's model of the agent's belief update from initial state B_0 to B_1.

Bayesian Theory of Mind

A hungry student leaves his office looking for lunch from one of three food trucks: Korean, Lebanese or Mexican; however, there are only two spots.

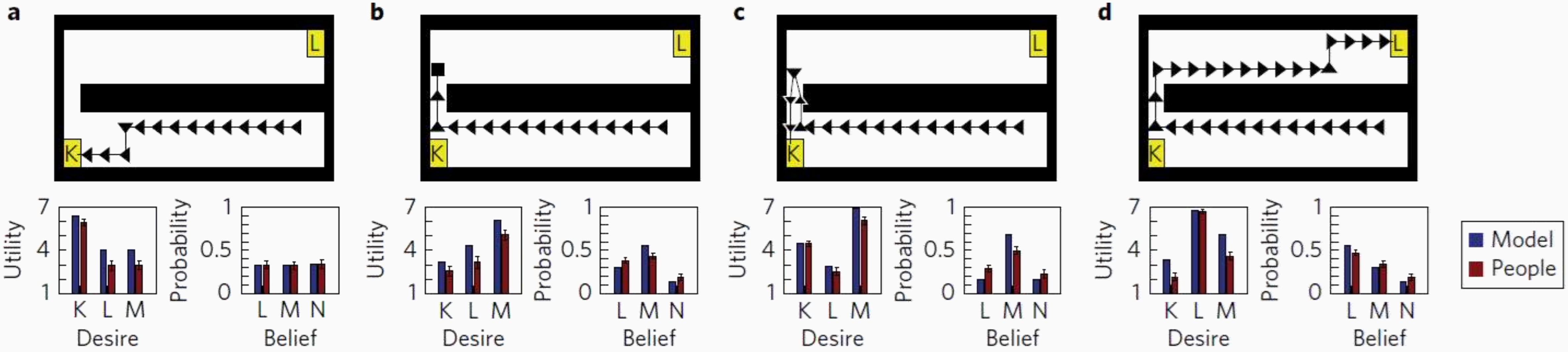
Given the three frames below, how would you order student's preference?



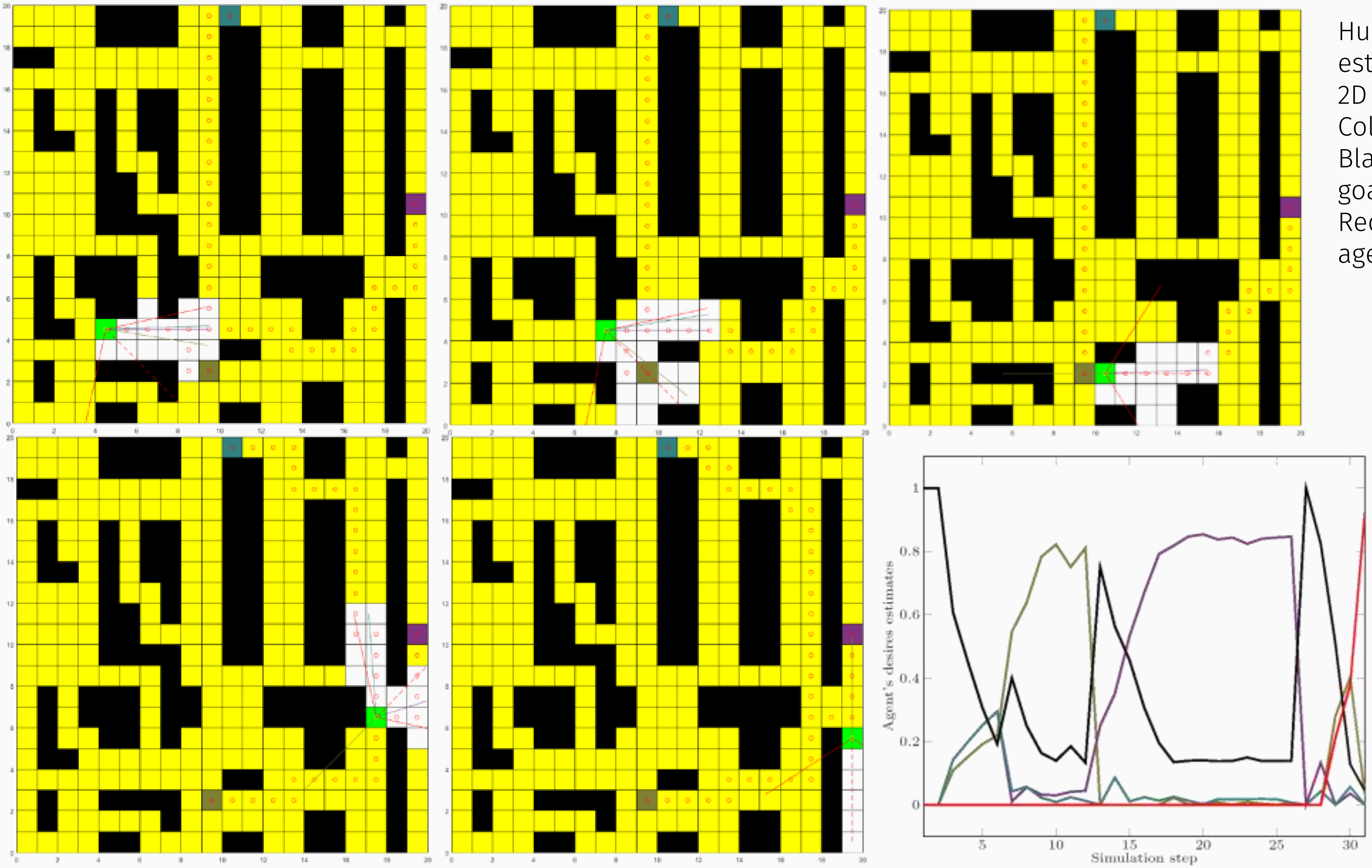
Bayesian Theory of Mind

Below we see results of a BToM algorithm applied on the food truck example with predicting inferences over:

- desires – which food does the student prefer, and
- beliefs – what the agent believes which truck is behind the wall at the start of the experiment.



Bayesian Theory of Mind



Human intention estimation in a 2D grid world. Colors = Goals, Black = unknwon goal, Red = irrational agent

Outline

- Motivation
- What is Human-Robot Interaction?
- The evolution of HRI
- Design in HRI
  - Principles
  - Anthropomorphisation
- Human intention estimation
  - Bayesian Theory of Mind
- HRI applications
- Examples from FER's robotic labs
  - LABUST, LARICS, LAMOR

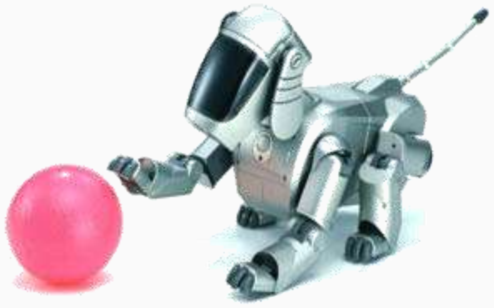
HRI applications

Human-robot interaction has numerous applications expected to make a positive difference in people's lives.

One of the pioneering products was Sony's Aibo (1999), discontinued in 2006 but relaunched in 2018.

Sony also worked on the Qrio robot, and Mayfield Robotics with Bosch on the Kuri robot – the two never made to the market.

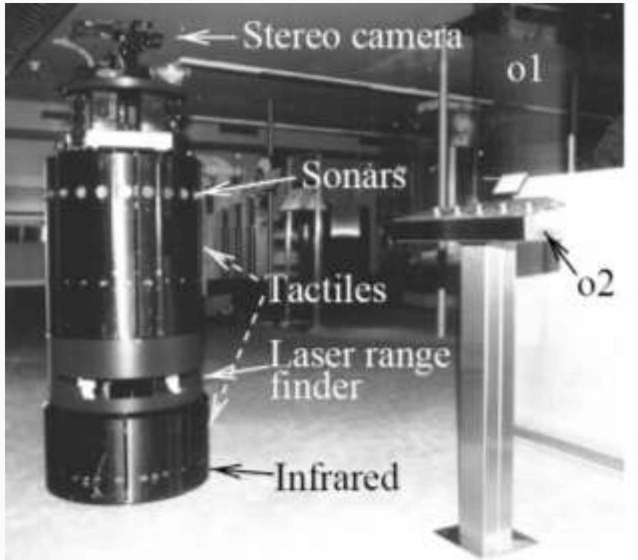
Not all robots are created to be successful applications, some are there to inspire people to think – see the Helpless robot.



HRI applications

Tour guide robots:

- move from one location to another while providing information about nearby entities;
- involves safe navigation and face-to-face interaction
- E.g., museum guides[16], retail robots that take you to the appropriate shelf, airport escort to correct gates, guides for visually impaired
- ability to display emotions can enrich the educational experience, better manage its interactions (getting frustated to make people move)



[16] Burgard, Wolfram et al. (2008). „The interactive museum tourguide robot."

HRI applications

Receptions robots:

- placed at a reception desk and interact with visitors, typically offering information through spoken-language conversation
- provide directions and would share daily stories with people
- people were sensitive to the robot's moods, and the length of their interactions with it changed based on whether the robot displayed a happy, sad, or neutral expression
- Robots have also been used as receptionists in hotels (unfortunately they were fired after 3 years in 2018)

[16] Gockley, Rachel et al. (2005). „Designing robots for long-term social interaction."

HRI applications

Sales promotion:

- robots function as proxies for store clerks, informing customers about the promotions offered by the store
- people are naturally curious about robots, these robots can easily attract the attention of potential visitors
- Typically not proactive but instead wait for visitors to initiate interaction
- Question is if proactive approach to customers is better than passive

HRI applications

Robots for entertainment:

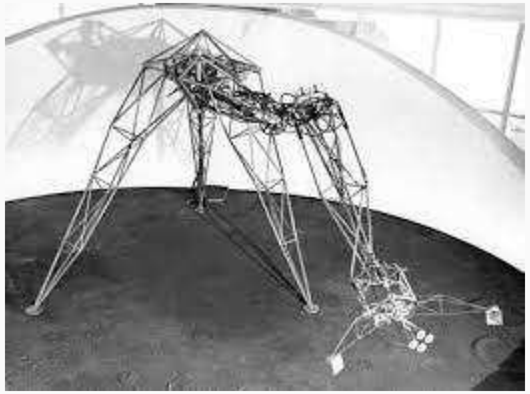
- entertainment robots have been easier to get to market because the functions they perform do not have to be as advanced
- For example: Aibo, Furby, Lego Mindstorms, Dash, Ozobot, Robosapiens, Femisapiens, Sphero, Pleo…

- Expected to stay one of the largest markets for personal robots

HRI applications

Robots for exhibitions:

- animatronic devices that are very robust; they must play the same animation script sometimes hundreds of times per day
- Most of these animatronic robots have no autonomy: they play a prerecorded script of animation timed to a soundtrack
- Disney Reserach is also developing stunt robots to explore untethered dynamic movemements



HRI applications

Service robots:

- Service robots are designed to help humans in various onerous, often called „dull, dirty, and dangerous," tasks – typically simple, repetitive and do not offer explicit interaction with people
- Cleaning robots for homes (Roomba), delivery robots in hospitals (Aetheon TUG), logistics robots (Amazon Kiva, Swisslog CarryPick, Gideon Casey), automated forklifts (Gideon Trey, Romb Technoloies)

HRI applications

Collaborative robots:

- Gaining importance in the automation industry
- Traditional industrial robots typically are stiff, strong, and have limited sensory capabilities
- In contrast, collaborative robots have safety features and a design that allow them to operate near people or even work together with people (Walt, Baxter and Sawyer, ABB Yumi, Fanuc CRX, Franka Emika, Kinova Jaco, KUKA iiwa, Universal Robots…)
- May fundamentally change the notion of collaborative teamwork






HRI applications

Self-driving cars:

- Although autonomous cars are still not widely available, most cars now have advanced driver assistance technologies (ADAS) - lane following, adaptive cruise control, automatic parking, predictive braking, pedestrian protection systems, and blind-spot warning systems
- self-driving cars require interfaces allowing them to interpret the action and intentions to passengers („why" is preferred over merely reporting that automated braking was initiated) and other traffic users
- HRI study with on perception of self-driving cars without the need for a fully self-driving car[19] (driver masked as a car seat ☺)

[19] Rothenbücher, Dirk et al. (2016). „Ghost driver: A field study investigating the interaction between pedestrians and driverless vehicles."

HRI applications

Challenges in robots applications:

- Addressing user expectations - The more humanlike the robot looks, the more human capabilities it may be expected to have. The cost of disappointing user expectations can be that the robot is perceived as incompetent, and people are therefore less willing to use it[20]
- Addiction - social robots might make people over-reliant on the social and physical interaction offered by robotic devices; robots might pamper the user, but the state is artificial but might be perceived as genuine by the human user[21]
- Attention theft – human attention is grabbed by motion and sound, especially when coming from lifelike and social device – care should be taken to identifiy when not to engage with the user[22]

[20] Paepcke, Steffi et al. (2010). „Judging a bot by its cover: An experiment on expectation setting for personal robots."
[21] Elder, Alexis (2017). „ Friendship, robots, and social media: False friends and second selves."
[22] Posner, Michael (2011). „Cognitive neuroscience of attention."

HRI applications

- Loss of interest - people pay more attention to a novel entity unfamiliar; however, such effects are usually not long lasting (from a few minutes to, at most, a few months)[23]
- Robot abuse - when left unsupervised, robots sometimes get abused by humans - shares more similarities with intimidation and bullying than with vandalism.
  Children seem especially prone to engage in robo-bullying behavior, presumably due to their strong tendency to anthropomorphize and as part of developing their social skills.
  Children became so disruptive that researchers eventually programmed the robot in a shopping mall to avoid children, especially when they were gathered in a group[24]

[23] Koang, Keng Lee et al. (2007). „Living with robots: Investigating the habituation effect in participants' preferences during a longitudinal human-robot interaction study."
[24] Brščić, Dražen (2015). „Escaping from children's abuse of social robots.."

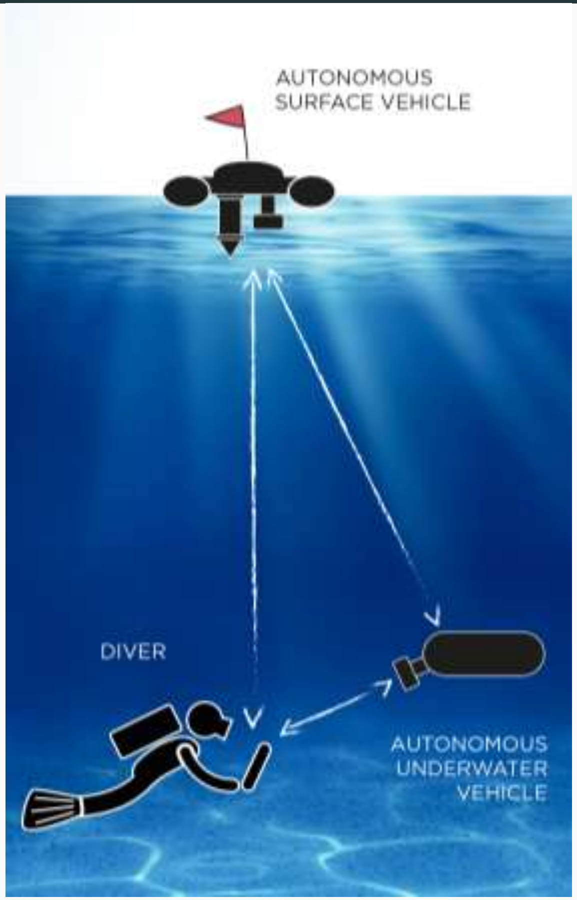
Outline

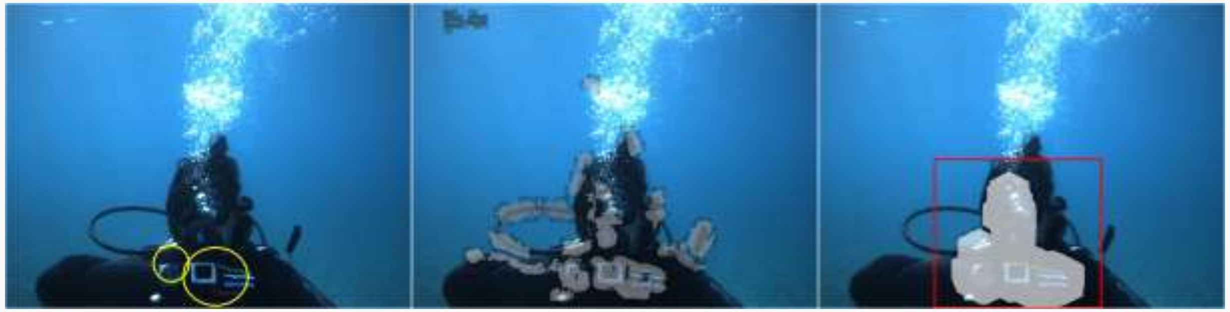
HRI applications of FER robotic labs

At the Laboratory for Underwater Systems and Technologies (LABUST) the CADDY project was conducted.

The project aimed to replace a human buddy diver with an autonomous underwater vehicle and an autonomous surface vehicle to improve monitoring, assistance, and safety of the diver's mission.

Project aimed to develop methods for determining the state of the diver, communicate with the diver, guide the diver and reconstruct the 3D diver model.





HRI applications of FER robotic labs

At the Laboratory for Robotics and Intelligent Control Systems (LARICS) the ADORE project was conducted.

The main goal of the ADORE project was the development of a robot-based diagnostic protocol and validation of a robotic diagnostic assistant in clinical settings.

The robots are used as ASD co-evaluators to help a human evaluator to assess a child's behavior objectively.

They used the Nao robots to communicate with the children in a Wizard-of-Oz setting.

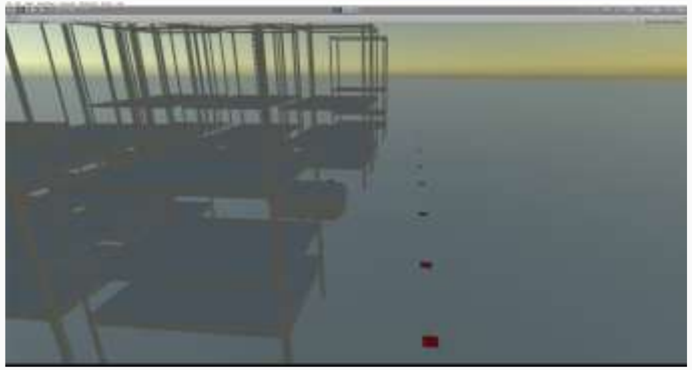
HRI applications of FER robotic labs

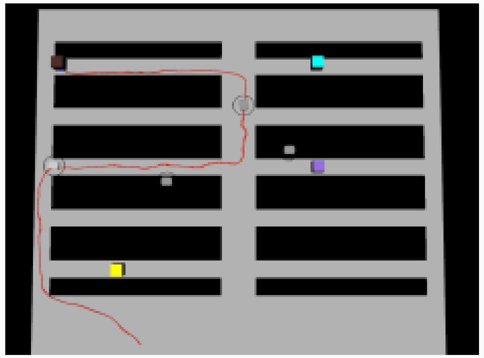
At the Laboratory for Autonomous Systems and Mobile Robotics (LAMOR) the SAFELOG project was conducted.

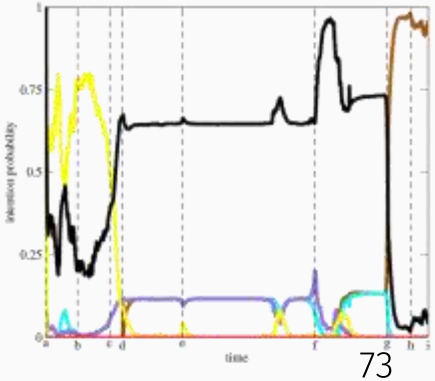
The goal of the project was to develop a holistic and certifiable safety concept, which allows the collaboration of AGVs and humans in automated flexible warehouse.

One aspect of this safety concept was human intention estimation; specifically, to estimate at which goal the human is walking to next.

Using this information robots can be rerouted to avoid the human and thus increase the overall warehouse efficiency.







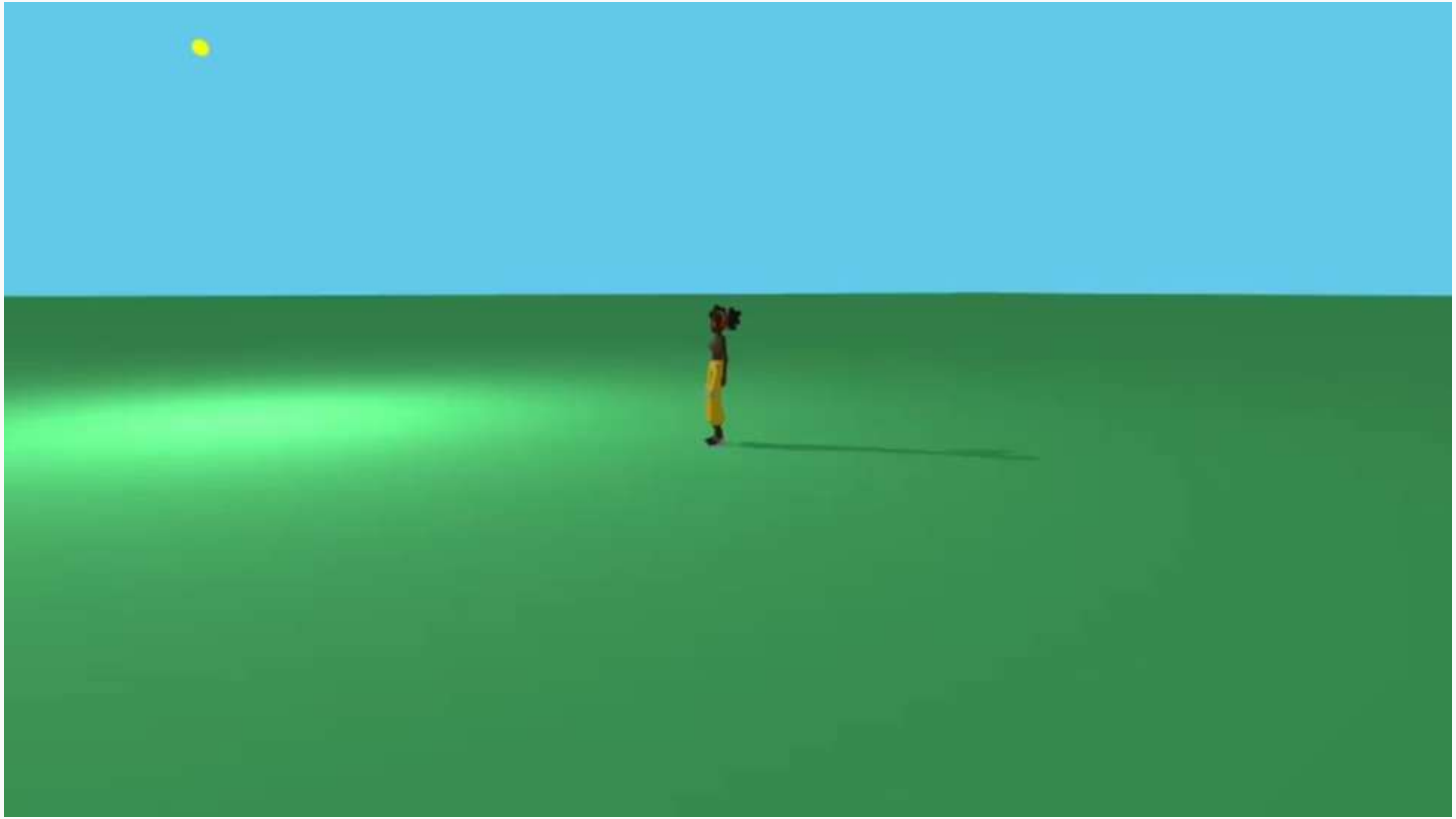
Playing chess with Baxter

Dino Bajić, Master thesis

Robot teleoperation

Filip Marić, Bachelor thesis

Human pose estimation



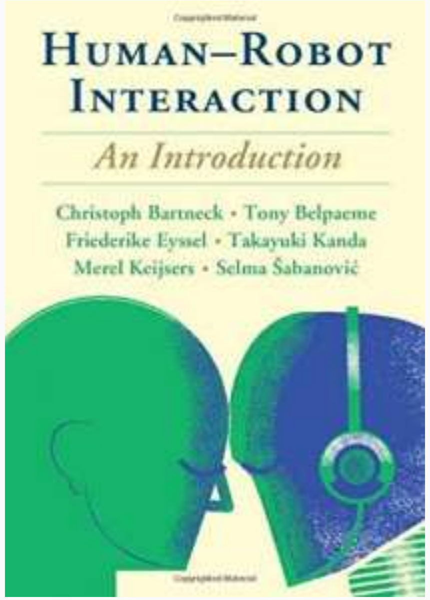
People detection and tracking

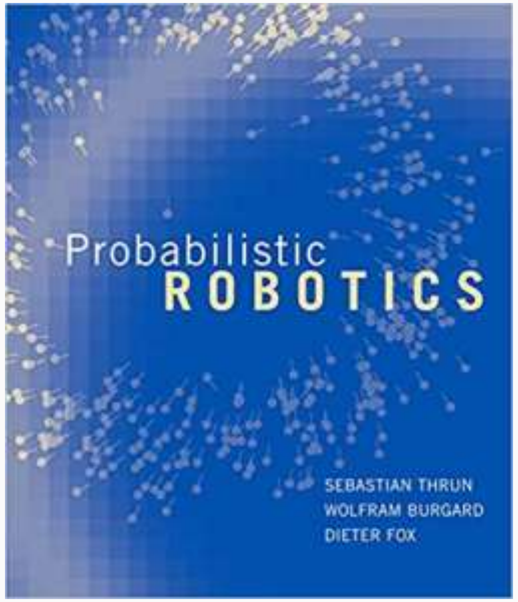
LLM robot control

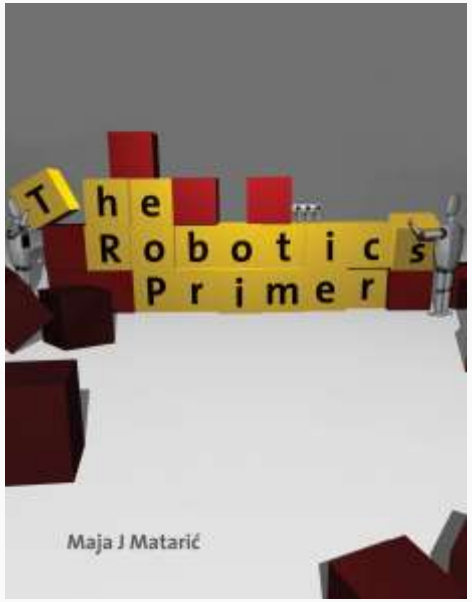
Precious Philip-Ifabiyi, Loc Pham, Atharva Patwe, Mohamed Khaled, HRI Final Project

Literature

The slides have been inspired and created by using materials from several books, research papers, and similar courses:







- Human-robot interaction design, Indiana University, Carlotta A. Berry
- Human-robot Interaction, Georgia Tech, Andrea L. Thomaz
- Principles of Human-Robot Interaction, Carnegie Mellon University, Illah R. Nourbakhsh

(A gentle) Introduction to Human-Robot Interaction

Ivan Marković

University of Zagreb Faculty of Electrical Engineering and Computing
Department of Control and Computer Engineering
Laboratory for Autonomous Systems and Mobile Robotics (lamor.fer.hr)