State of the Art and Research Needs

- Fred Proctor
National Institute of Standards and Technology, Retired

OUTLINE

$$\begin{bmatrix} X_x & Y_x & Z_x & P_x \\ X_y & Y_y & Z_y & P_y \\ X_z & Y_z & Z_z & P_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

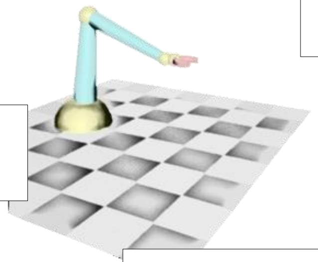
Key Concepts

Dexterous Manipulation

Mobility

Sensing and Perception

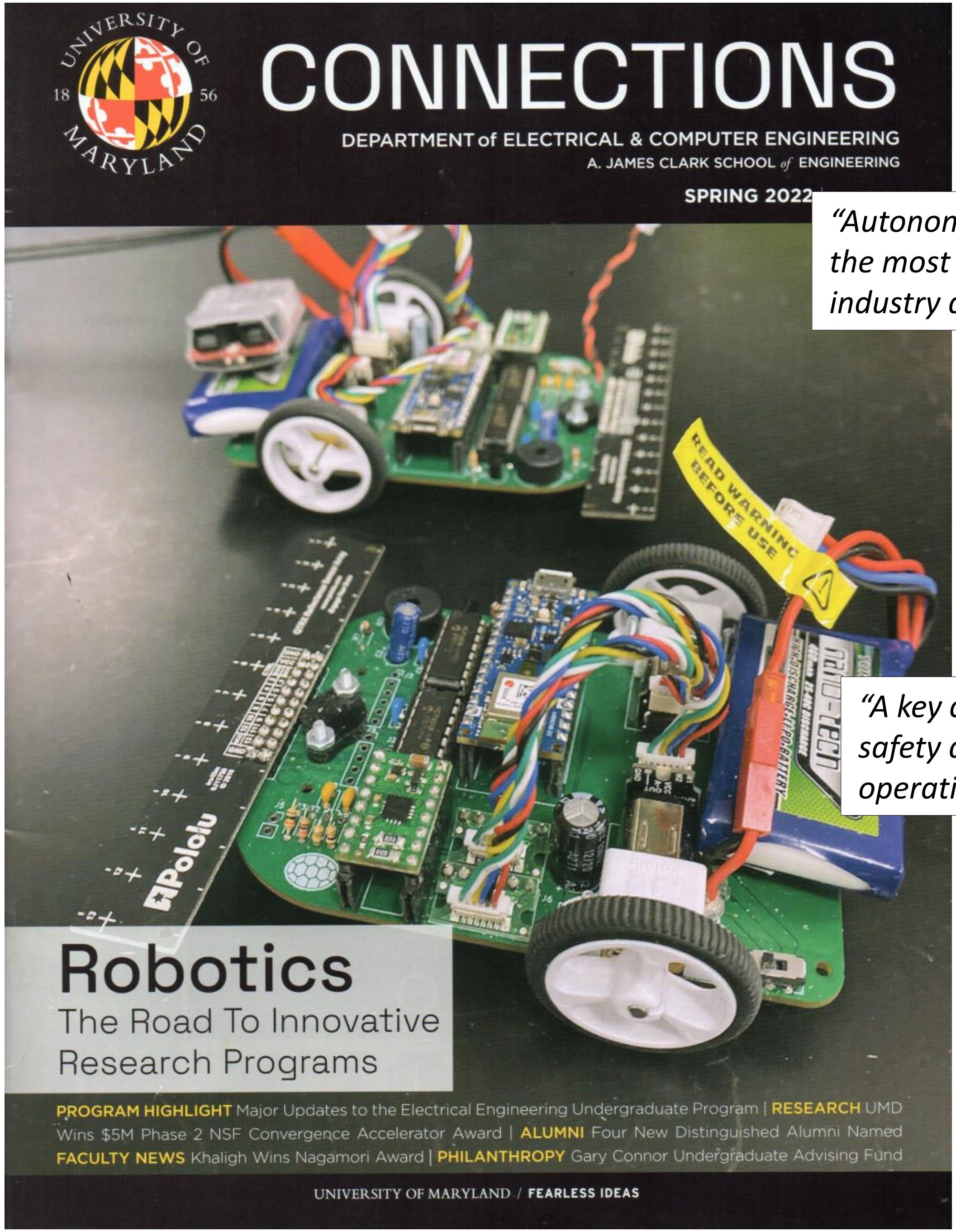
Collaborative Robots

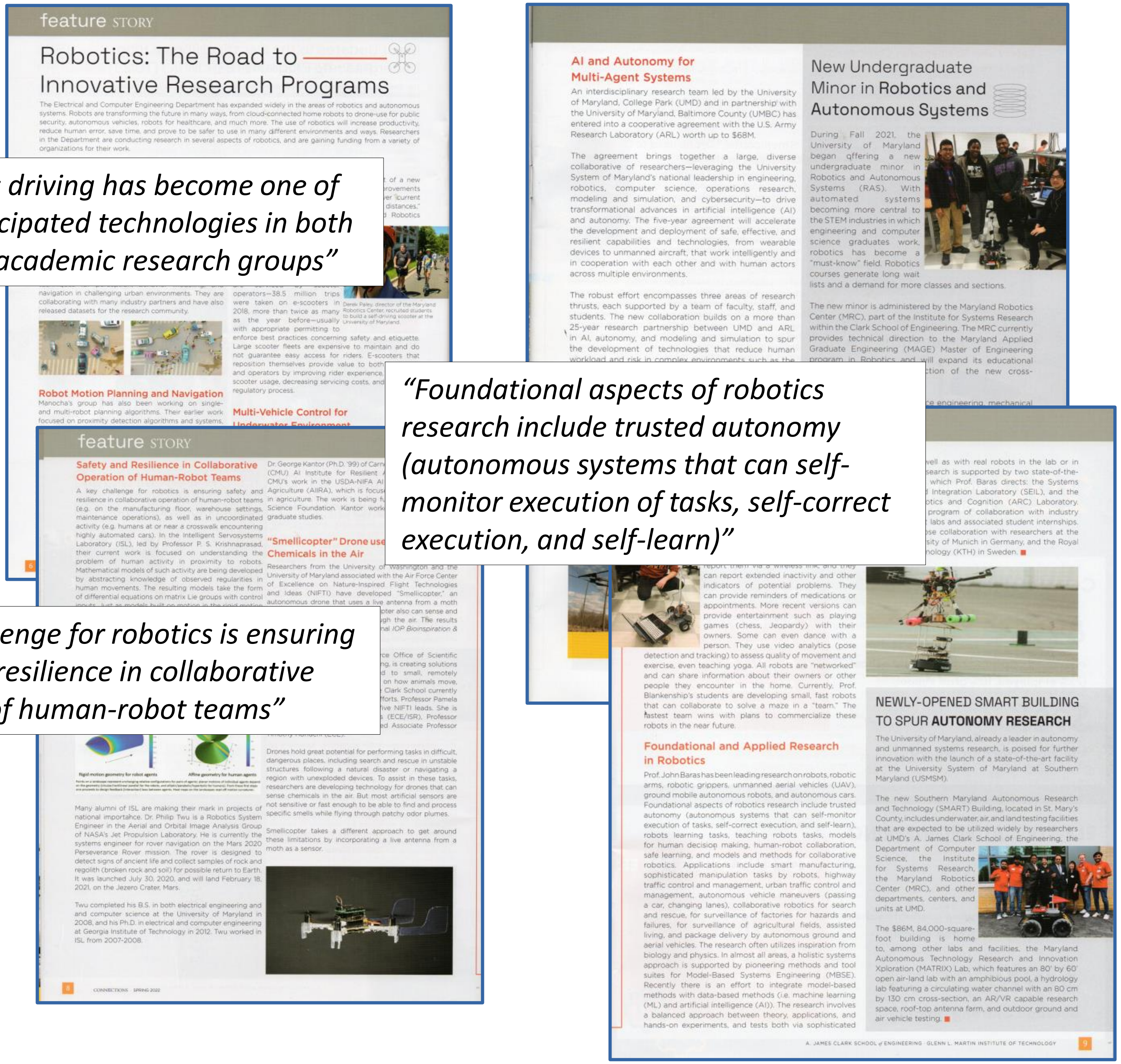


Simulation

Planning

Artificial Intelligence





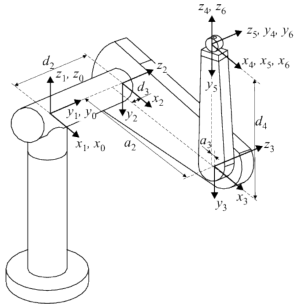
"Autonomous driving has become one of the most anticipated technologies in both industry and academic research groups"

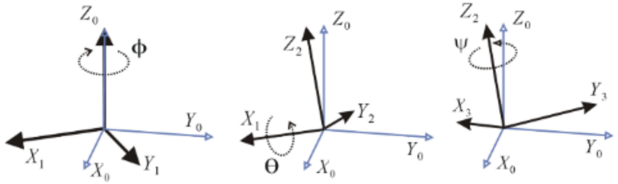
"Foundational aspects of robotics research include trusted autonomy (autonomous systems that can self-monitor execution of tasks, self-correct execution, and self-learn)"

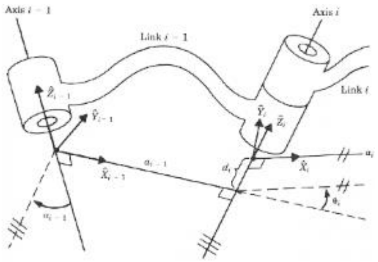
"A key challenge for robotics is ensuring safety and resilience in collaborative operation of human-robot teams"

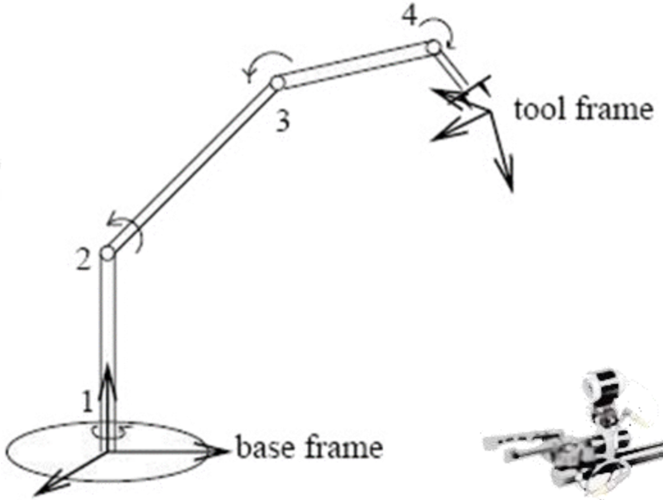
KEY CONCEPTS

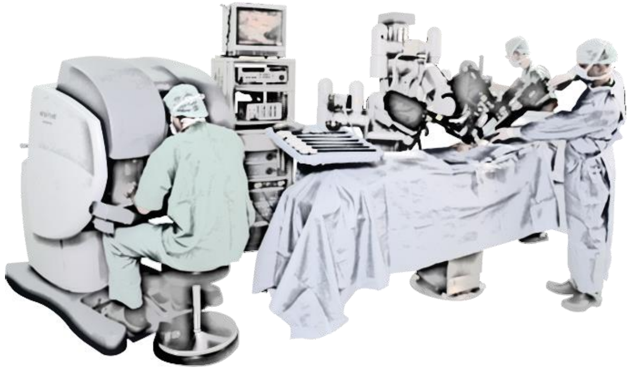
$$\begin{bmatrix} X_x & Y_x & Z_x & P_x \\ X_y & Y_y & Z_y & P_y \\ X_z & Y_z & Z_z & P_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

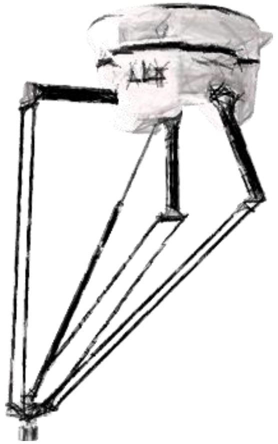




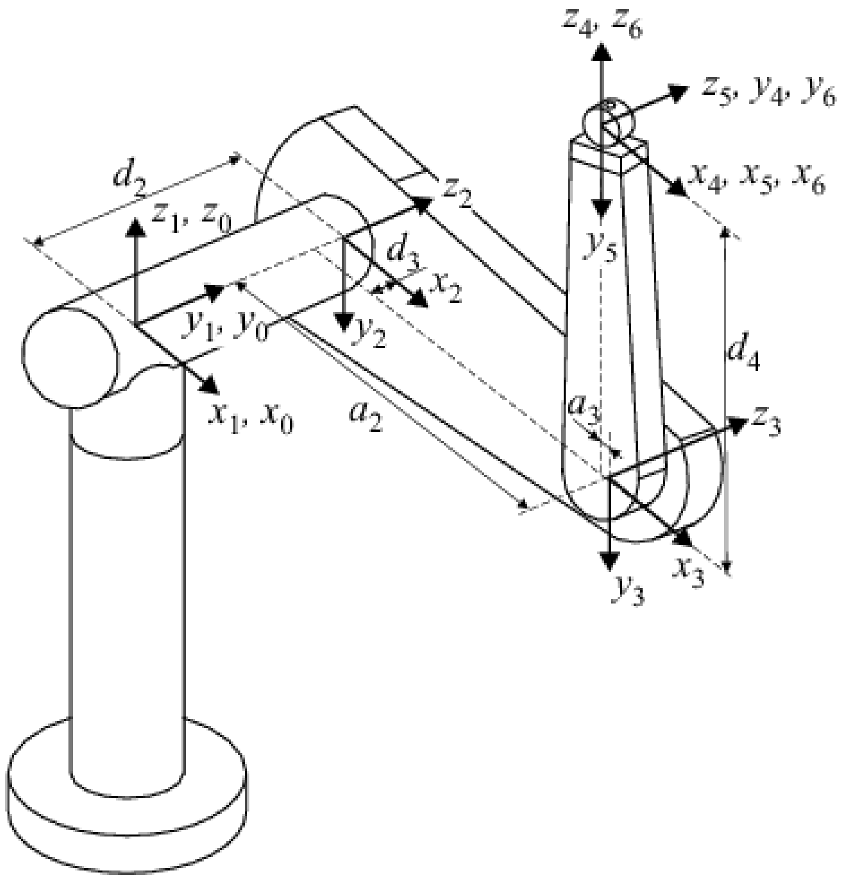


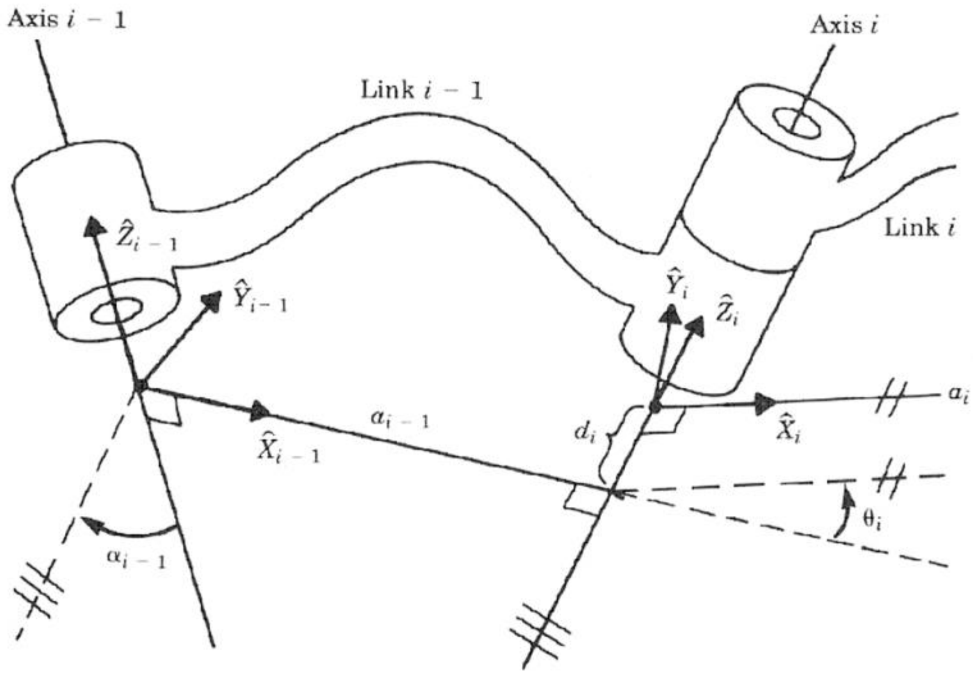






Kinematics define the geometric relationship between robot *links* and the *joints* that connect them

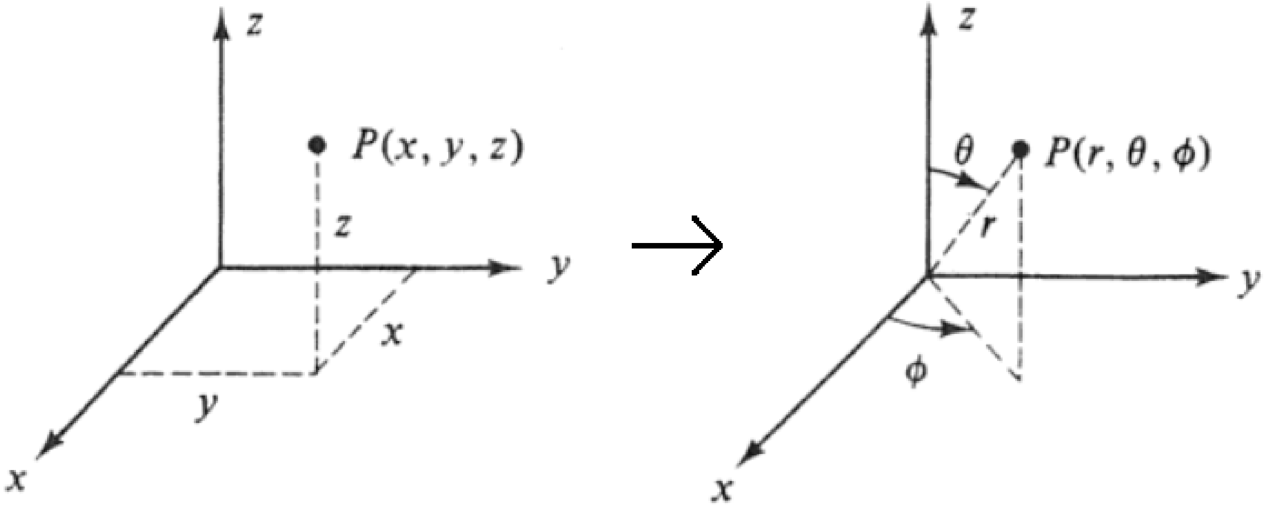




The Denavit-Hartenberg standard form is widely used to represent kinematics, with 4 parameters for each joint-link pair

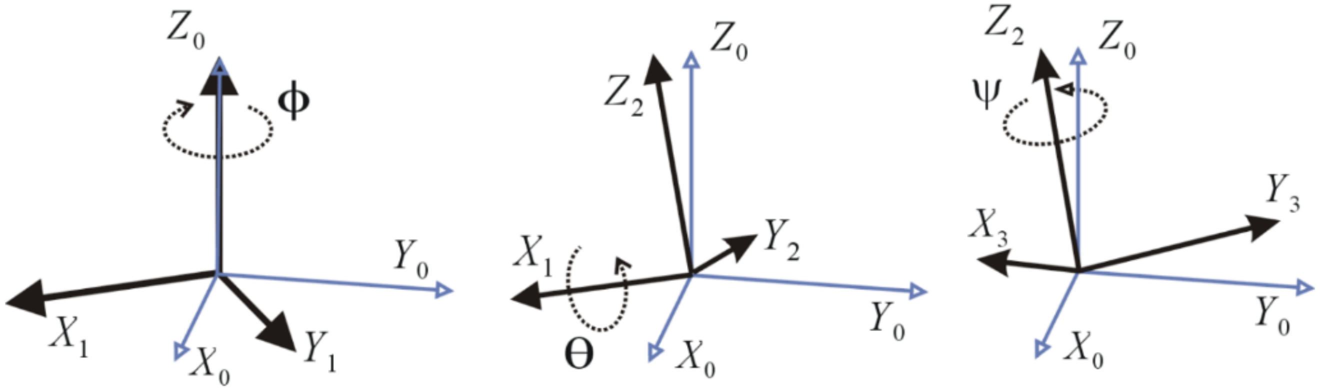
Credits: John Craig, Introduction to Robotics

The position and orientation of an object in our 3D space takes 3 numbers for position and 3 for orientation



There are several equivalent ways to represent position, e.g., X-Y-Z or spherical coordinates

Likewise for orientation, e.g., roll-pitch-yaw or Euler angles



Rotation matrices let you chain together rotations using matrix multiplication

$$R_x(\theta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}$$

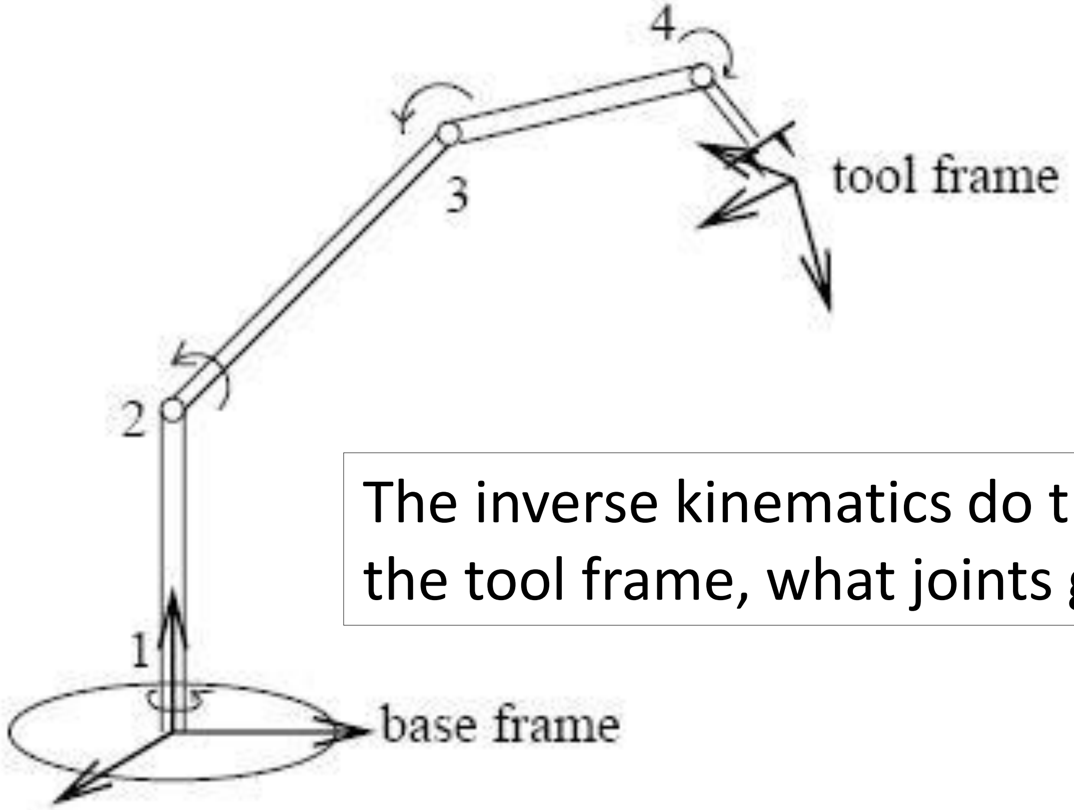
$$R_y(\theta) = \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix}$$

$$R_z(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

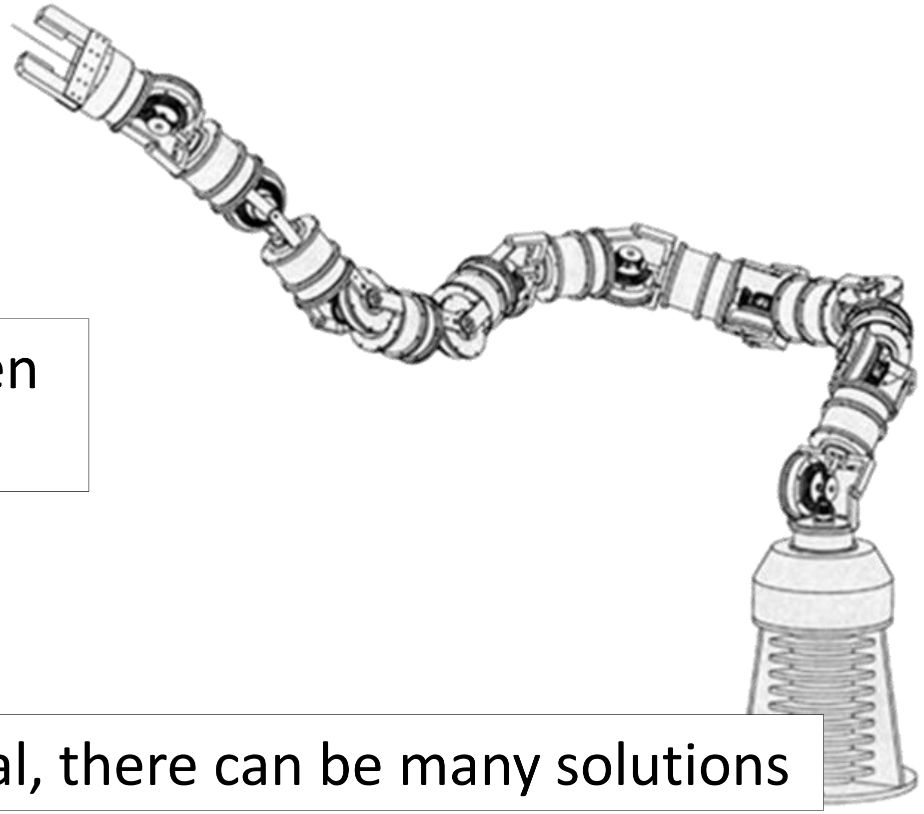
$$\begin{bmatrix} X_x & Y_x & Z_x & P_x \\ X_y & Y_y & Z_y & P_y \\ X_z & Y_z & Z_z & P_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Homogeneous transforms are the full position and orientation version, padded with some extra fluff to make square matrices amenable to multiplication and inversion

The *forward kinematics* relate the pose at the end of the robot (its *tool frame*) with respect to the *base frame*, for a given set of joint values



The inverse kinematics do the opposite – given the tool frame, what joints get us there?



In general, there can be many solutions

Dynamics is the aspect of robot structures that handle speeds, forces and torques

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \\ \omega_x \\ \omega_y \\ \omega_z \end{bmatrix} = \left[\frac{dh(q)}{dq} \right]_{6\times n} \begin{bmatrix} \dot{q}_1 \\ \dot{q}_2 \\ \vdots \\ \dot{q}_n \end{bmatrix}_{n\times 1}$$

The *Jacobian matrix* relates changes in joint values (e.g., joint speeds) to changes in the Cartesian position (e.g., linear- and angular speed of the wrist)

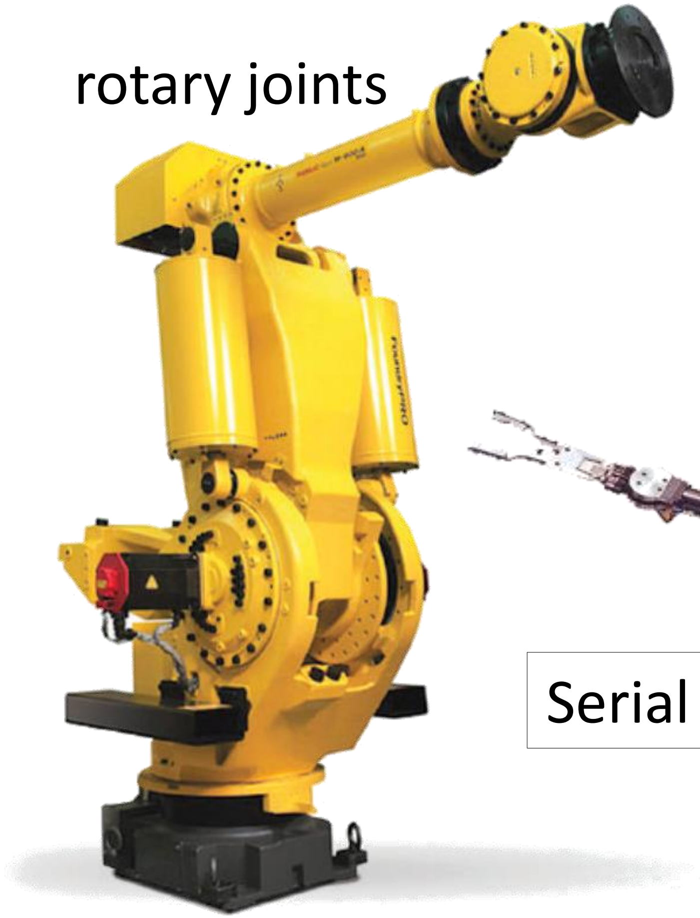
$$J = \left(\frac{dh(q)}{dq} \right)_{6\times n} = \begin{bmatrix} \frac{\partial h_1}{\partial q_1} & \frac{\partial h_1}{\partial q_2} & \cdots & \frac{\partial h_1}{\partial q_n} \\ \frac{\partial h_2}{\partial q_1} & \frac{\partial h_2}{\partial q_2} & \cdots & \frac{\partial h_2}{\partial q_n} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial h_6}{\partial q_1} & \frac{\partial h_6}{\partial q_2} & \cdots & \frac{\partial h_6}{\partial q_n} \end{bmatrix}_{6\times n}$$

In general, it is not a square matrix; its *pseudo-inverse* can be used to pick from many configurations (over-actuated) or find a best fit (under-actuated)

It is also used to related joint torques to wrist forces and vice-versa

Types of Robot Kinematics

rotary joints



prismatic joint

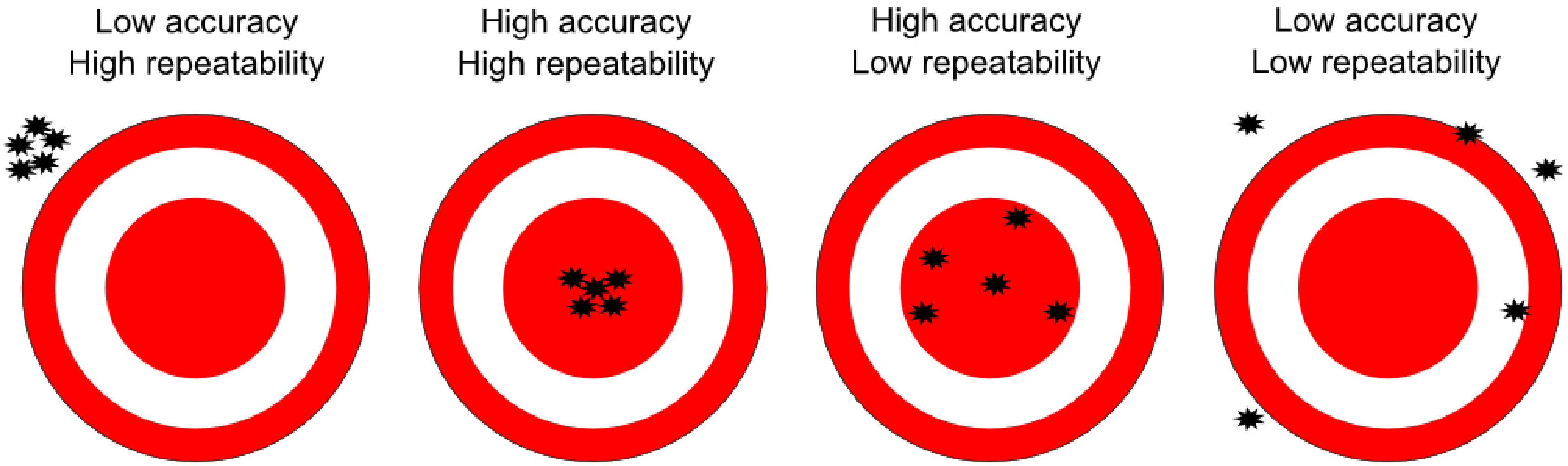
Serial links

Parallel links

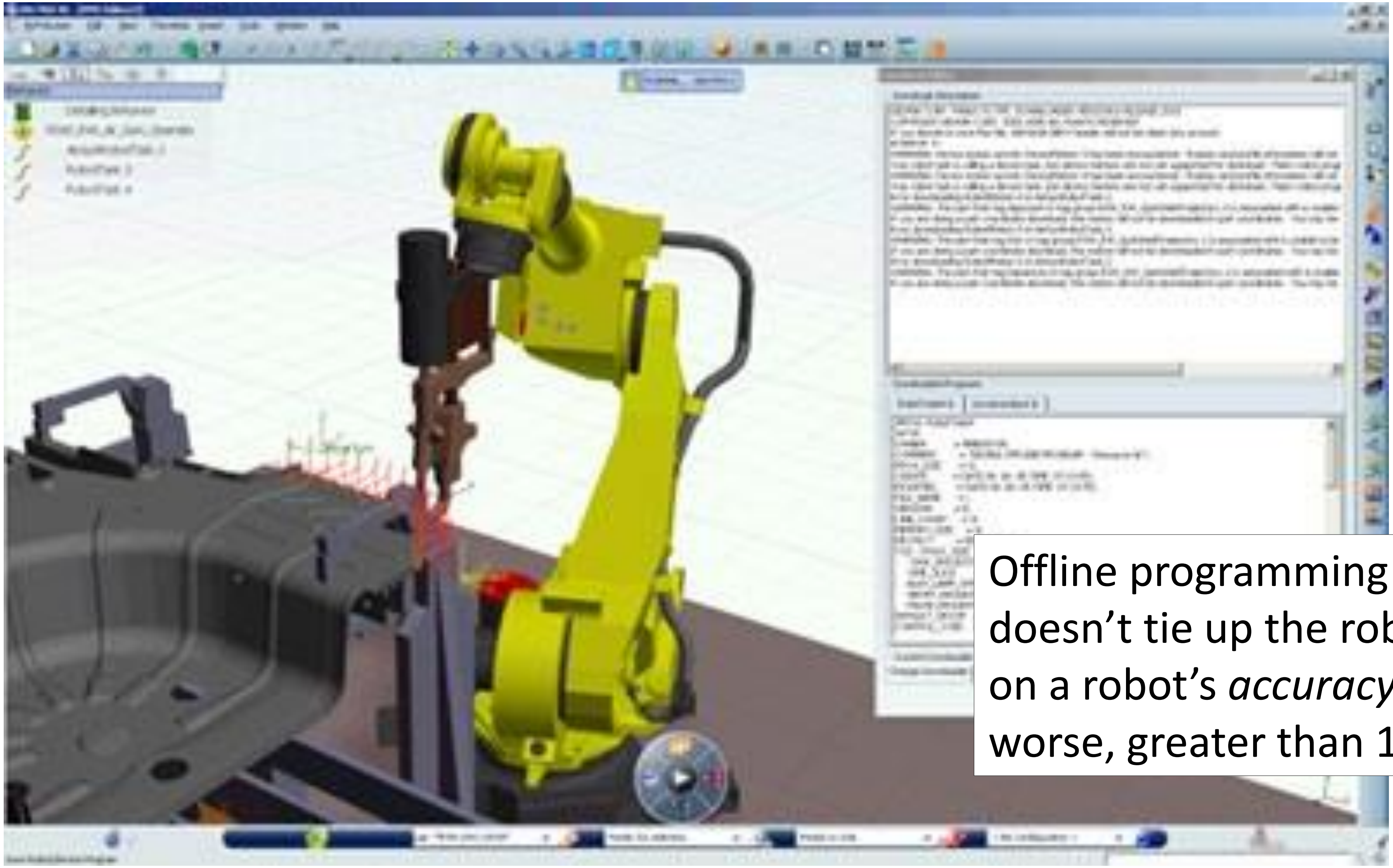
Credits: ABB; Fanuc; Unimation

Teach programming is the dominant programming method – manually guiding the robot, recording poses, playing them back

It’s tedious and time consuming, but gives good results since it relies on the robot’s *repeatability*, typically below 1 mm



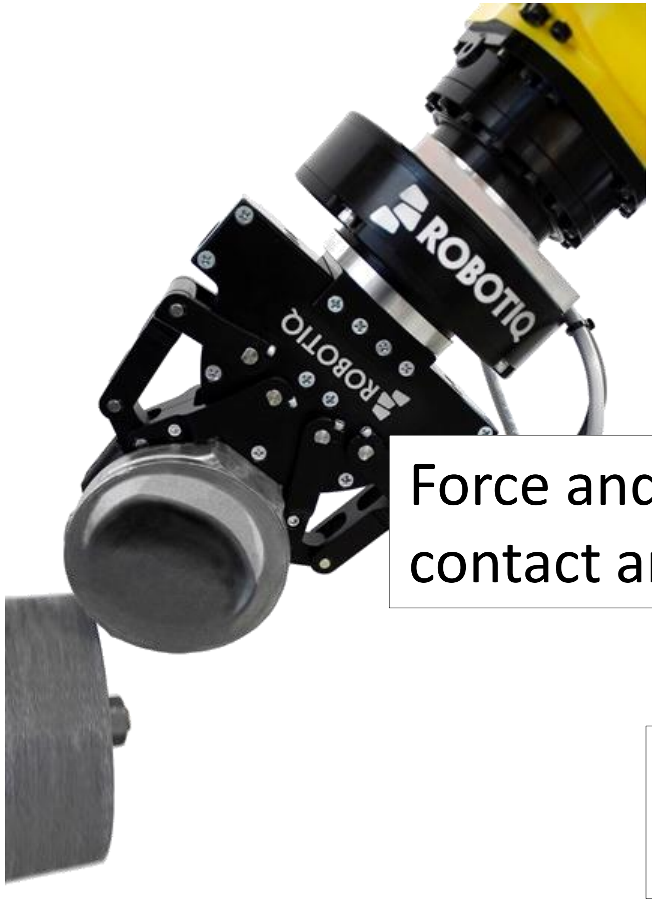
Credits: TheFabricator.com



Offline programming is faster, and doesn’t tie up the robot, but it relies on a robot’s *accuracy*, typically much worse, greater than 1 mm

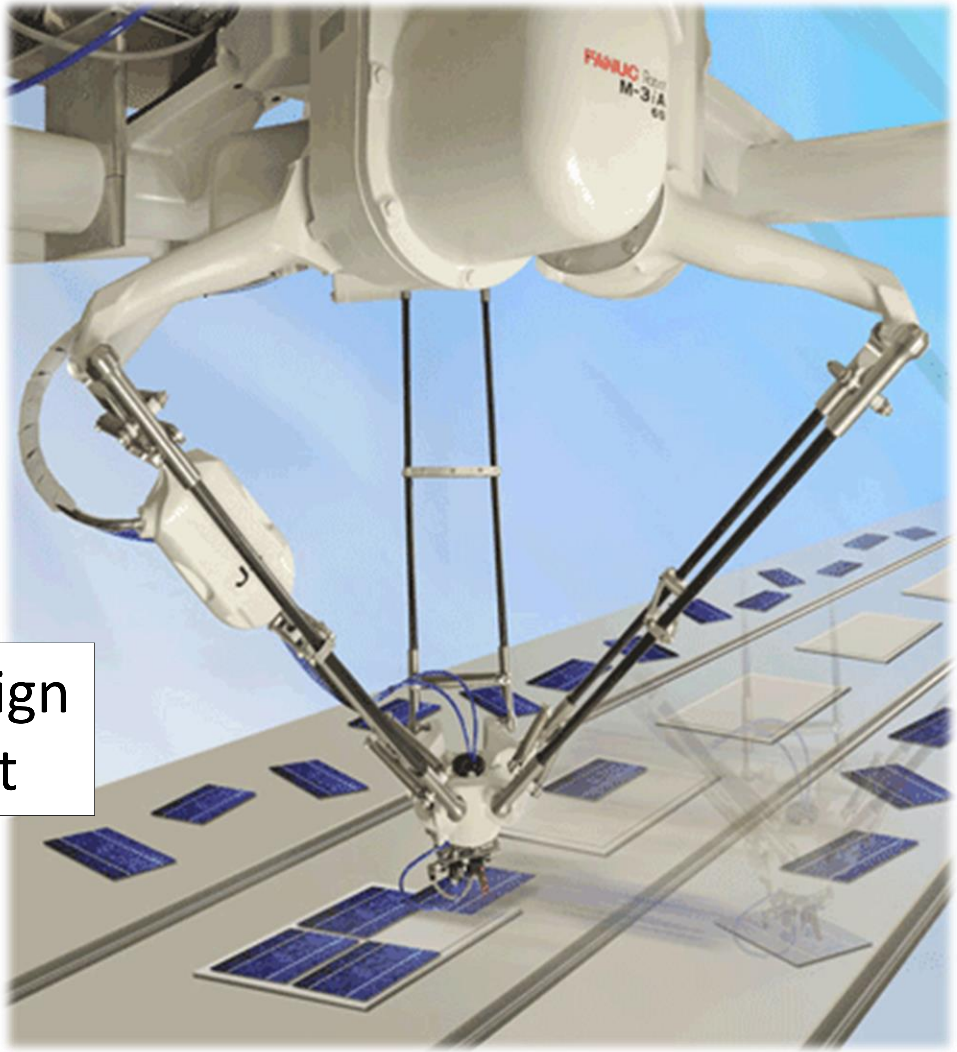
Credits: Delmia

Sensors can improve the accuracy of robots in demanding applications



Force and torque sensors can detect contact and guide insertion tasks

Vision sensing can precisely align semiconductor chip placement



Credits: Fanuc; Robotiq

Haptic devices “push back” on operators, conveying forces felt by the robot

3D vision microscopes and haptic feedback are used during microsurgery

Credits: Kuka; Intuitive Surgical

Mobile robots are usually *teleoperated* by a person

Some robots are *autonomous*, and can plan actions and take corrective actions themselves by sensing the environment

Credits: iRobot; John Deere; Qinetiq

Useful robots are connected to other things

Robot *controllers* provide many options for input/output and networking

The nice thing about standards is there are so many to choose from

Internet connectivity and standards are becoming more common for industrial devices

Credits: Adept

INTERNET OF THINGS RELATED STANDARDS

Below is a partial listing of IEEE standards related to the Internet of Things.

IEEE 754™-2008 - IEEE Standard for Floating-Point Arithmetic

IEEE 802.1AS™-2011 - IEEE Standard for Local and Metropolitan Area Networks - Timing and Synchronization for Time-Sensitive Applications in Bridged Local Area Networks

IEEE 802.1Q™-2011 - IEEE Standard for Local and metropolitan area networks--Media Access Control (MAC) Bridges and Virtual Bridged Local Area Networks

IEEE 802.3™-2012 - IEEE Standard for Ethernet

IEEE 802.3.1™-2011 - IEEE Standard for Management Information Base (MIB) Definitions for Ethernet

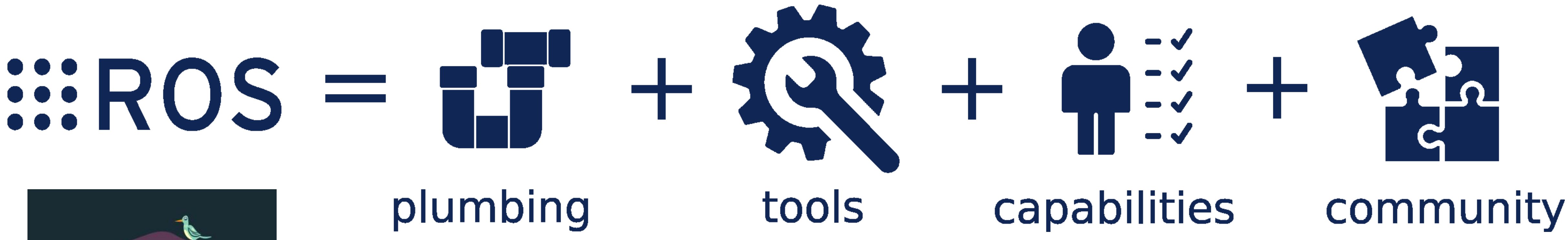
IEEE 802.11™-2012 - IEEE Standard for Information Technology--Telecommunications and information exchange between systems--Local and metropolitan area networks--Specific requirements Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications Amendment 10: Mesh Networking

IEEE 802.11ad™-2012 - IEEE Standard for Local and Metropolitan Area Networks - Specific Requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications - Amendment 3: Enhancements for Very High Throughput in the 60 GHz Band

IEEE 802.15.1™-2005 - IEEE Standard for Information Technology - Telecommunications and Information Exchange Between Systems - Local and Metropolitan Area Networks - Specific Requirements. - Part 15.1: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Wireless Personal Area Networks (WPANs)

IEEE 802.15.2™-2003 - IEEE Recommended Practice for Information Technology - Telecommunications and Information Exchange Between Systems - Local and Metropolitan Area Networks - Specific Requirements Part 15.2: Coexistence of Wireless Personal Area Networks With Other Wireless Devices Operating in Unlicensed Frequency Bands

The Robot Operating System (ROS) is a popular open-source framework for developing robot applications



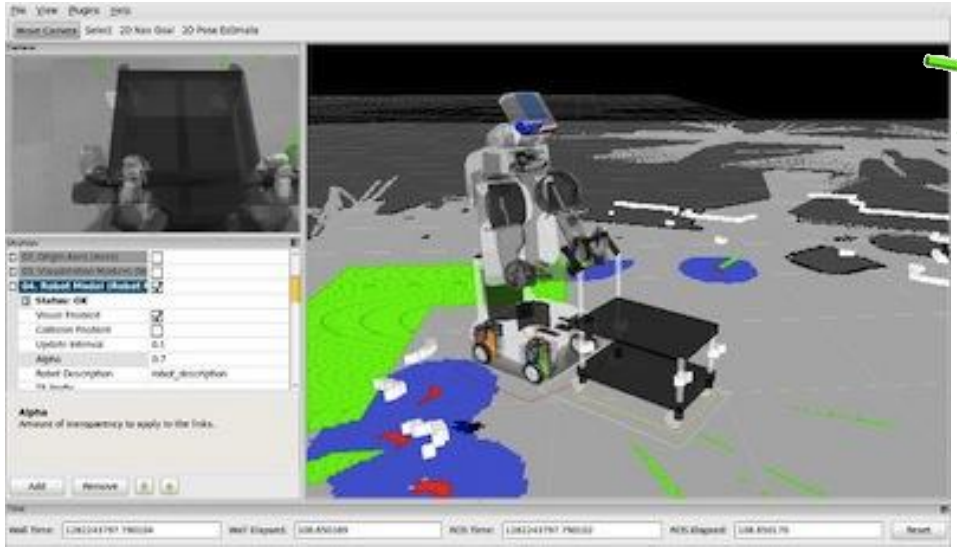
ROS 1 is the original Linux-based system

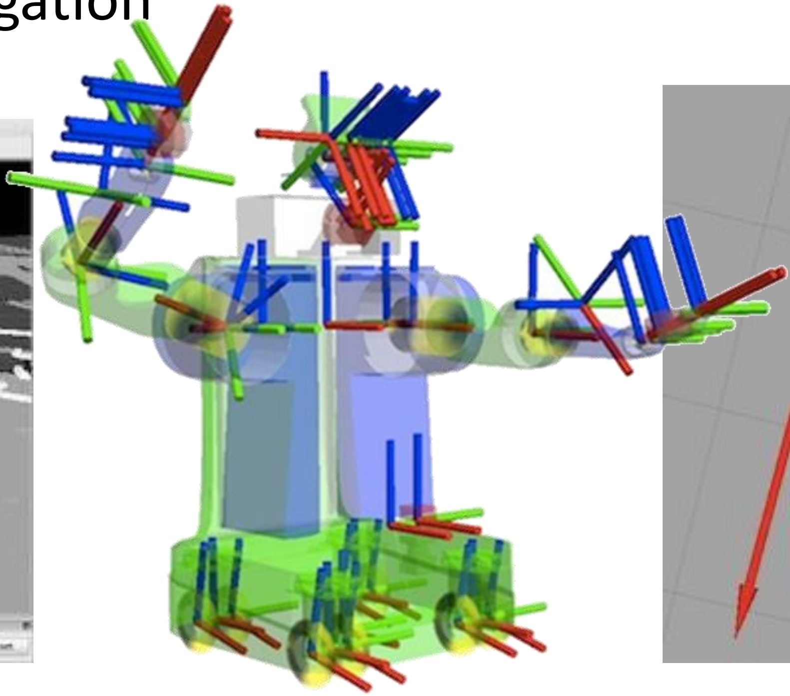
ROS 2 works on Linux, Windows, MacOS

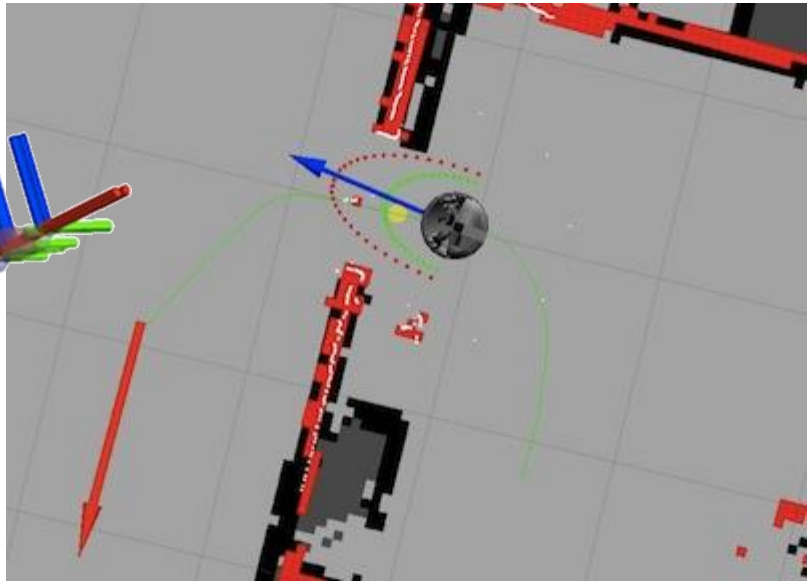
Micro-ROS works on embedded systems like Arduino and Raspberry Pi

ROS

- Standard Message Definitions for Robots
- Robot Geometry Library
- Robot Description Language
- Visualization and Diagnostics
- Pose Estimation
- Localization, Mapping, and Navigation

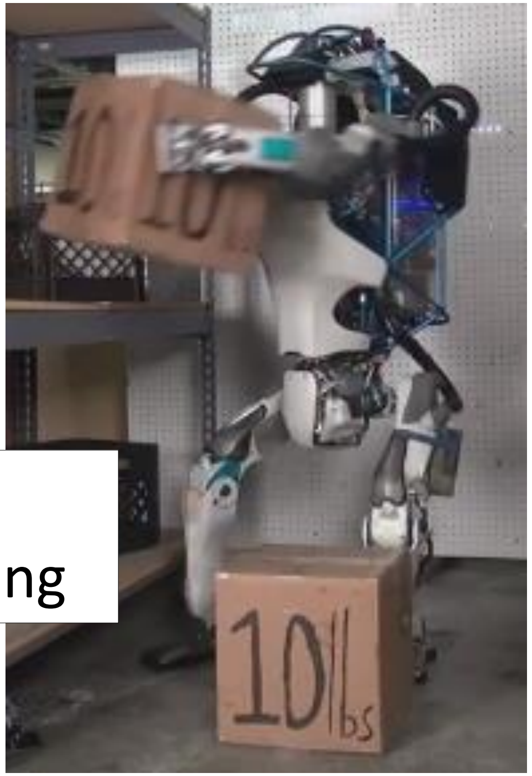






Robots Everywhere!

Bin picking



Box loading

Package delivery

Surgery

Room delivery

Mining

Farming

Credits: Amazon; Autonomous Solutions; Boston Dynamics; Fanuc; UC Berkeley; Savioke

Software Development

- *Source Code Management* (SCM): version control, issue tracking, concurrent development
  - Version control supports file versioning, branching parallel versions for later merging, tagged releases
  - Examples: Concurrent Versioning System (CVS), Git, Mercurial
  - Host sites enable web-based collaboration
  - Examples: SourceForge, GitHub (ROS), BitBucket
- *Continuous Integration* (CI): source code merged and built continually, possibly several times daily
  - You define build rules and tests to run; code is built and tested automatically; results are posted
  - Often offered as a feature in hosted sites
  - Examples: Jenkins, Travis (ROS *Build Farm*)

Software Development

- *Virtualization*: virtual machines give you a complete computer system, emulated on a (much faster) computer
  - *Virtual Machines* are the full computer: hardware, operating system, and system software
  - *Containers* share the same hardware and operating system, but give you your own copy of resources such as system libraries and the root file system
  - Examples: Unix `chroot`, VMWare, VirtualBox, Docker
- *Cloud Computing*: your software (VM, Container, …) runs on a server somewhere externally (e.g., a data center on the Internet)
  - *Edge Computing* runs some of the computation close to the source, such as on servers in your facility
  - *Fog Computing* is similar, for Internet of things (IoT)
  - Examples: Amazon Web Services, Microsoft Azure

Pair Programming

- *Pair programming*: assistants within integrated development enviroments (IDEs) that help programmers by autocompleting source code
- Artificial intelligence (AI) programs trained on huge bodies of source code form the foundation; IDEs determine the intent of a sequence of source code and suggest the rest
- GitHub's Copilot is a prominent example, using OpenAI's Codex model
- Microsoft's Intellisense and Intellicode are similar
- Somewhat controversial: trained on open source that has a variety of licenses; does an auto-complete require the license of what was used to (mostly) train it [*] ?

* spectrum.ieee.org/ai-code-generation-ownership

sentiments.ts | write_sql.go | parse_expenses.py | addresses.rb

```typescript
#!/usr/bin/env ts-node

import { fetch } from "fetch-h2";

// Determine whether the sentiment of text is positive
// Use a web service
async function isPositive(text: string): Promise<boolean> {
  const response = await fetch(`http://text-processing.com/api/sentiment/`, {
    method: "POST",
    body: `text=${text}`,
    headers: {
      "Content-Type": "application/x-www-form-urlencoded",
    },
  });
  const json = await response.json();
  return json.label === "pos";
}
```

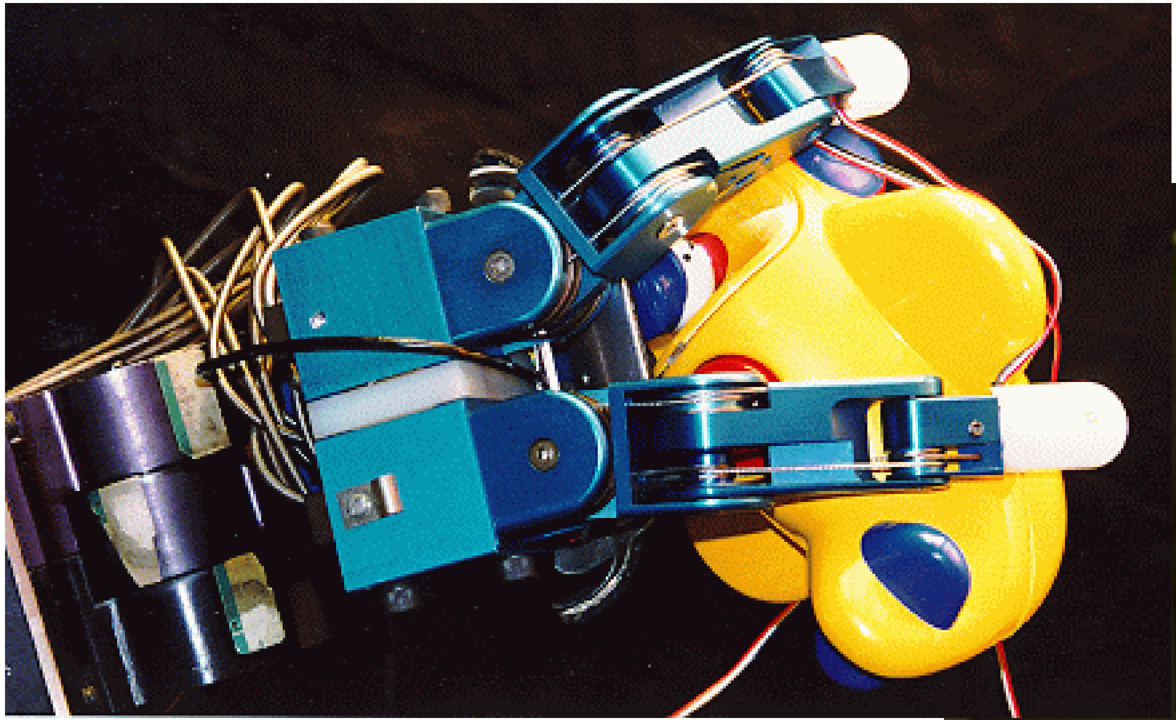
Copilot

You enter your intent using a comment; Copilot fills out the body

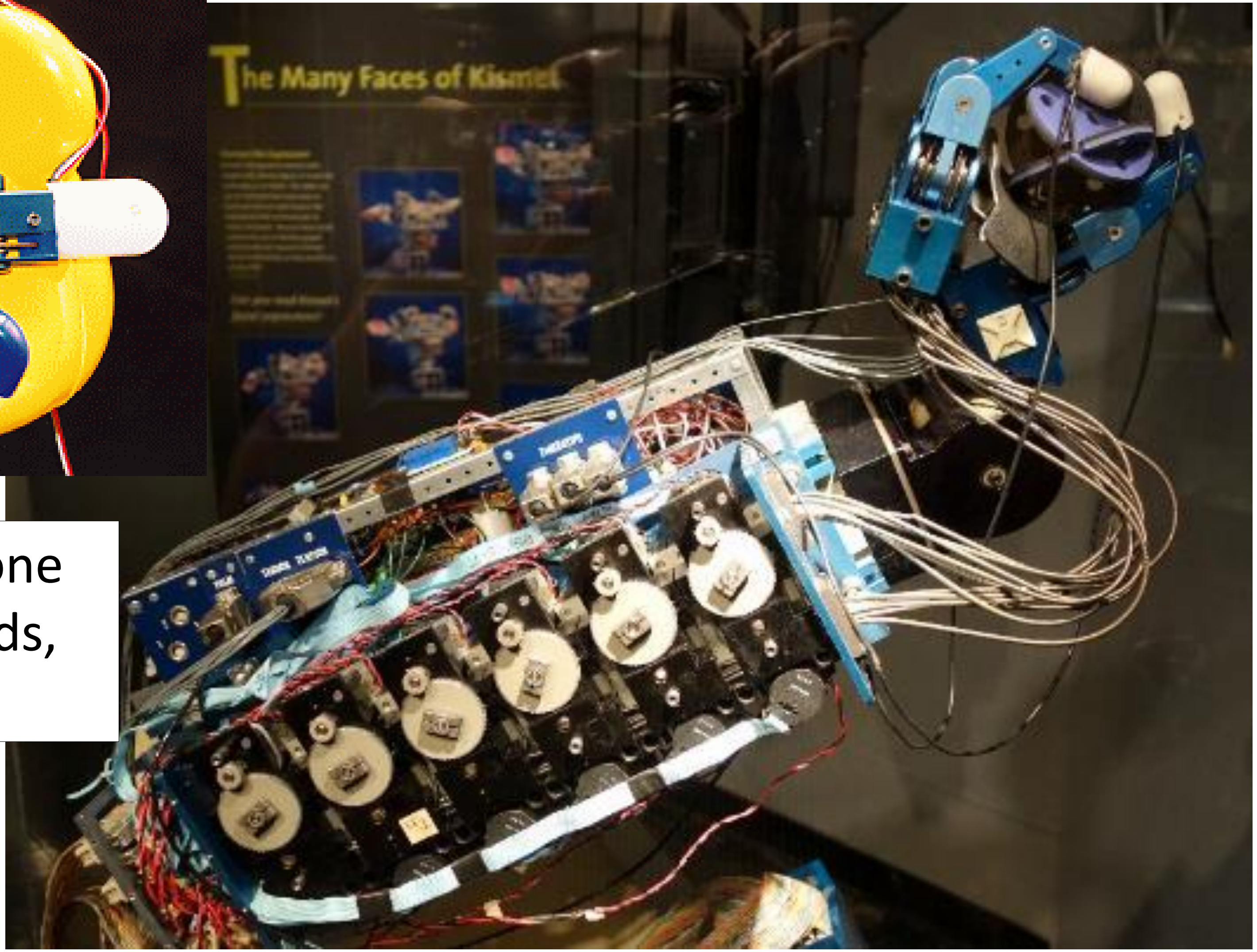
DEXTEROUS MANIPULATION

Traditional jaw- and vacuum-type grippers are common today, with continued innovation

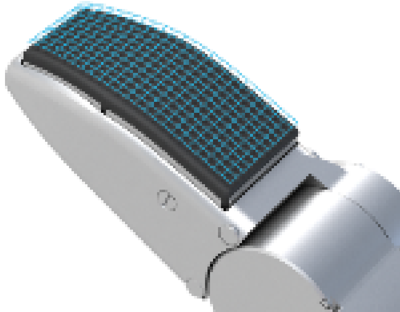
Credits: ABT Robotic; Empire; HTE Technologies; Materilise; Schunk



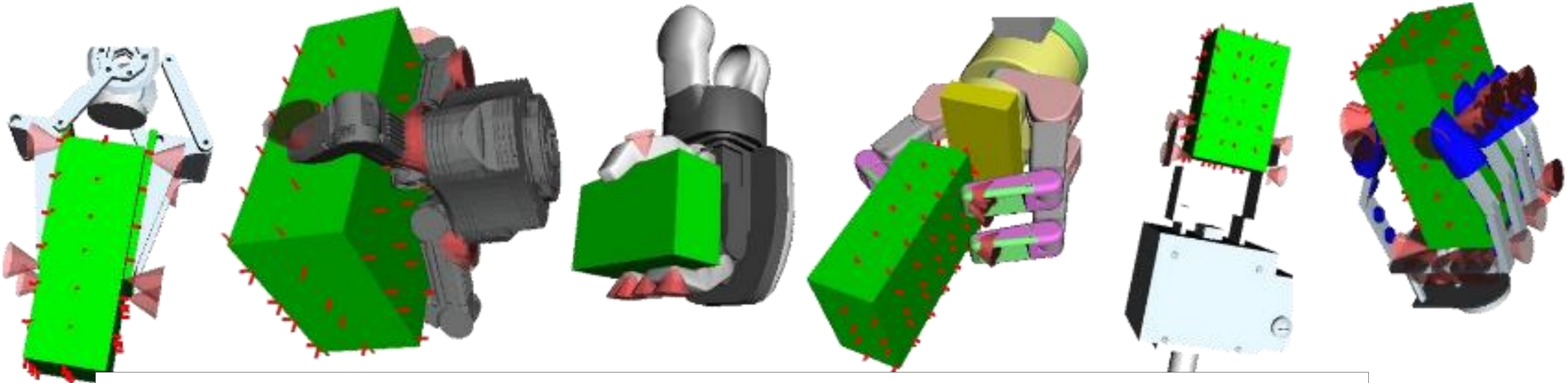
The Stanford/JPL Hand, c. 1985, one of the first anthropomorphic hands, actuated by cables



Dexterous hands today: active research, some commercially available, including tactile sensing



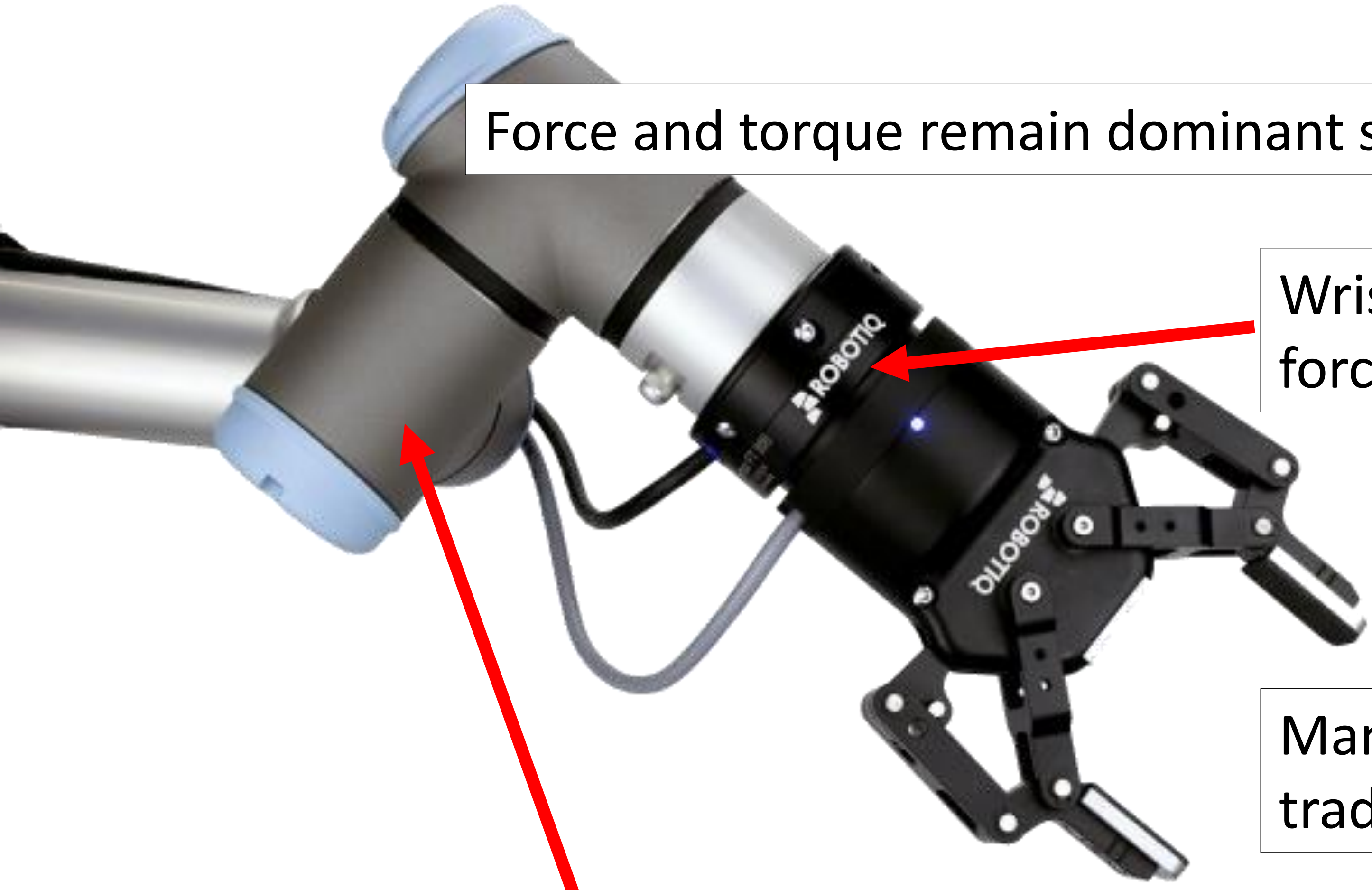
Credits: Barrett; Empire; Robotiq; Schunk; Shadow; SynTouch; U of Washington



Where to put the fingers? Grasp planning is an active research area

Credits: OpenRAVE; Rosen Diankov

Force and torque remain dominant sensing modes for robotic manipulation



Wrist-mounted force sensing measures forces at the application’s area of interest

Many interesting techniques can trade off accuracy vs cost

Torque sensing at the robot’s joints can compensate for gravity loads, and allow people to directly manipulate the robot

Credits: Robotiq; Universal

Kitting Demo

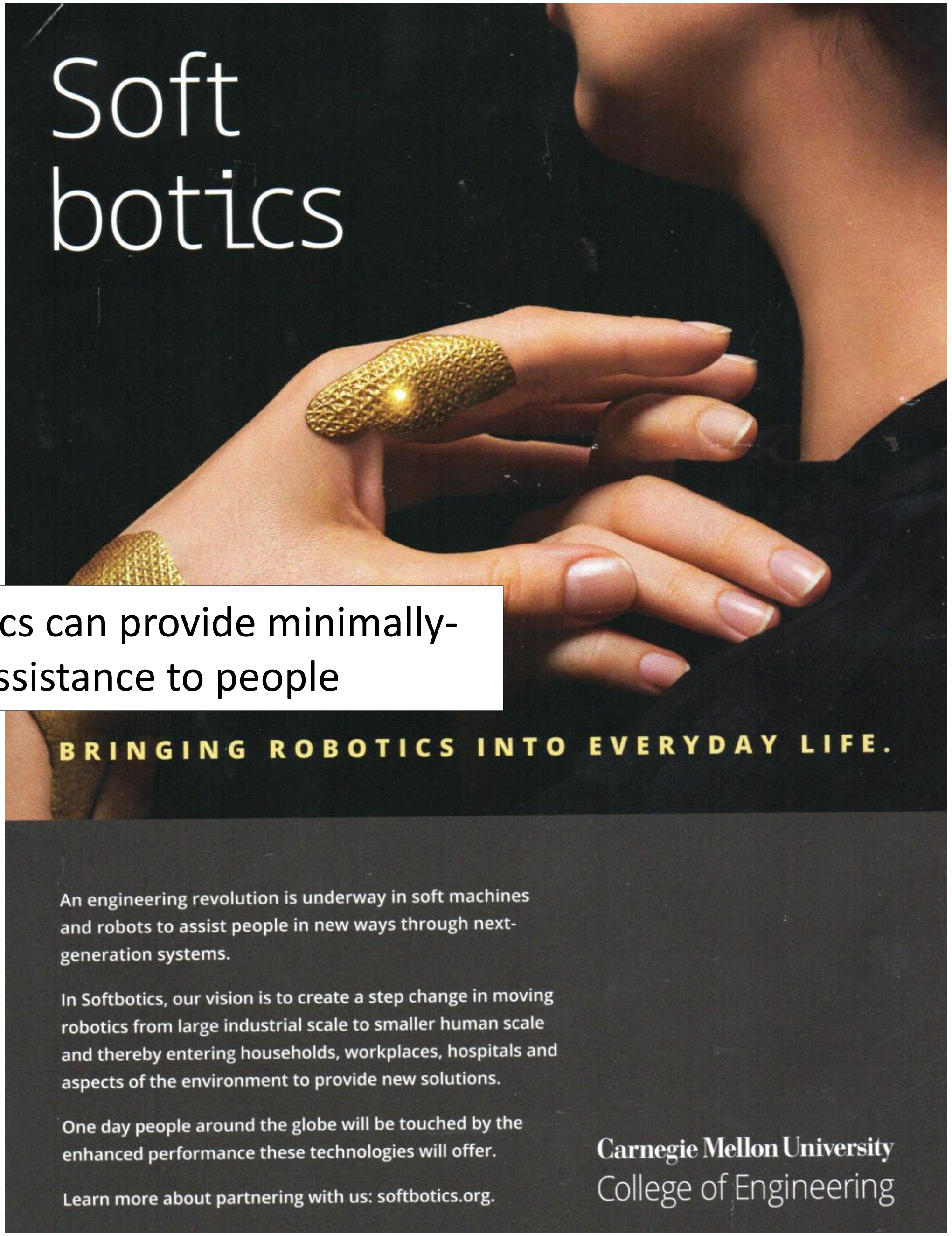
Yaskawa Motoman SDA10 & Robotiq Adaptive Gripper

YASKAWA
MOTOMAN ROBOTICS

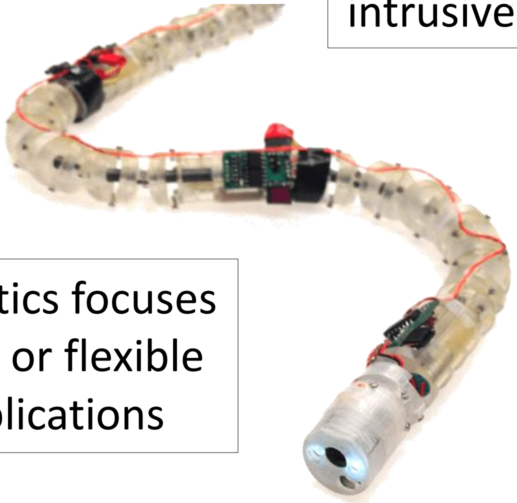
ROBOTIQ

Soft robotics is an emerging technology

"Soft robot tentacle is scarily effective"



Soft robotics can provide minimally-intrusive assistance to people

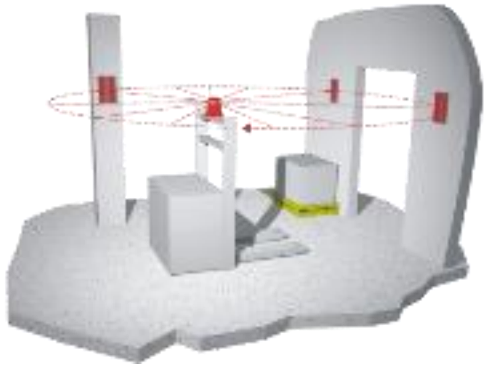


Continuum- and soft robotics focuses on the study of redundant or flexible mechanisms and their applications

Credits: Electronics Weekly, Carnegie-Mellon, Georgia Tech, MIT, Shenyang University of Technology

MOBILITY

Traditional automated guided vehicles (AGVs) perform well in structured environments, with many varieties

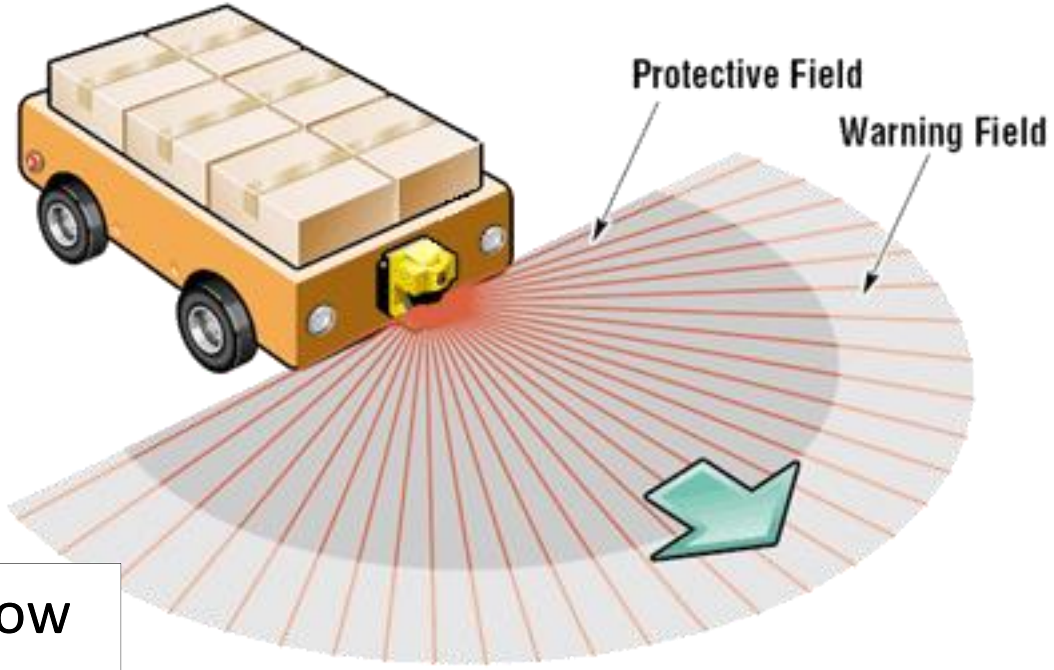


Infrastructure modifications may be needed: embedded guide wires, magnetic paint, reflective beacons, QR tags, elevator controller integration

Credits: Aethon; Creform; Egemin; Götting KG; Sky-Trax

Safety is always a concern

Contact bumpers are common, and were required by safety standards



Standards have been updated to allow non-contact scanning sensors

Credits: AGV International; CoreCon; Tapeswitch autonomous vehicles (self-driving cars)

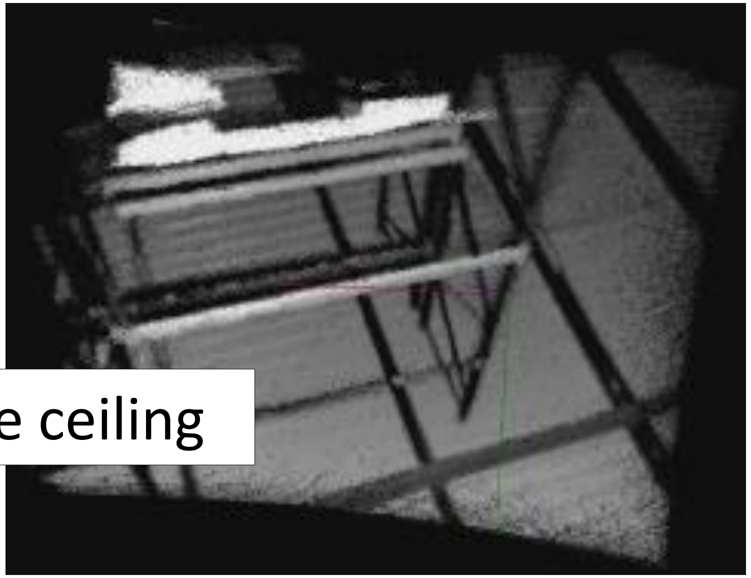
2007

cameras + lidar + GPS + maps

2018

2022

Infrastructure-free installations are the goal: wireless communication, on-board localization and mapping

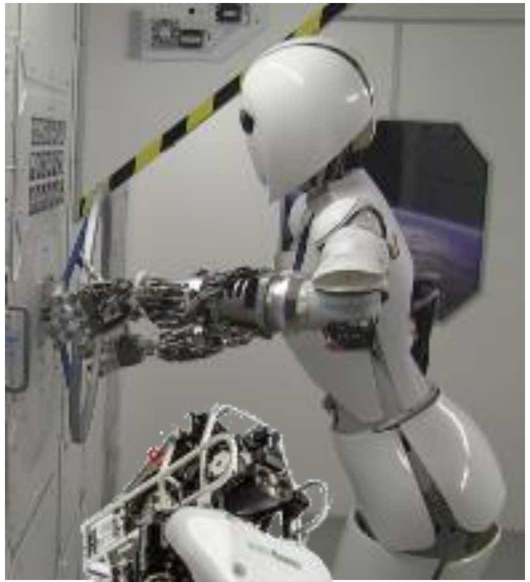


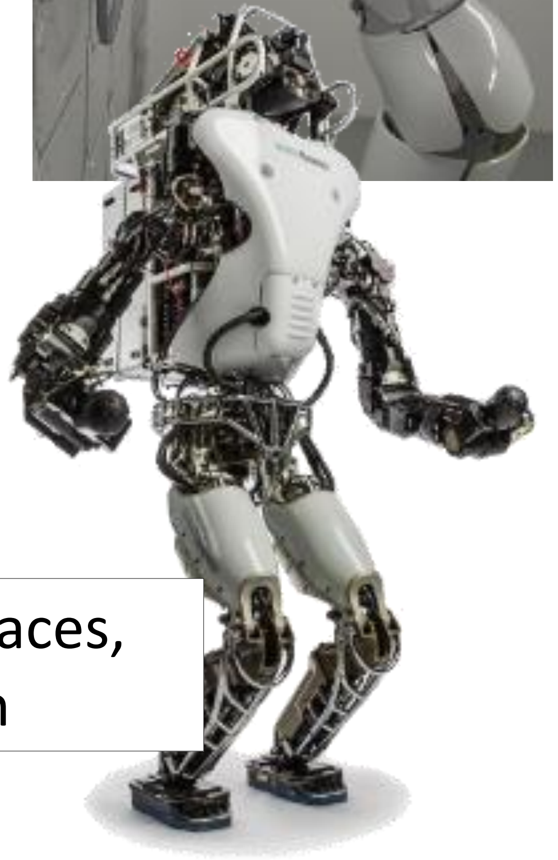
SLAM using images of the warehouse ceiling

Credits: Stanford U/MDV-Mohr; Waymo; Tesla

Mobile manipulation is an active R&D topic

The 2015 Amazon Picking Challenge posed a practical problem





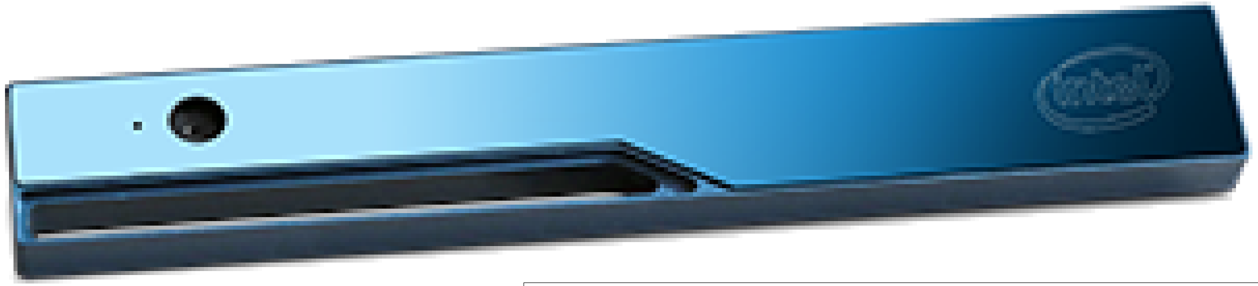
Full human-like abilities for walking on uneven surfaces, climbing steps, and opening doors is on the horizon

Credits: Boston Dynamics; Kuka; TU Berlin/RBO; U of Bremen/DFKI


Boston Dynamics

https://www.youtube.com/watch?v=tF4DML7FIWk

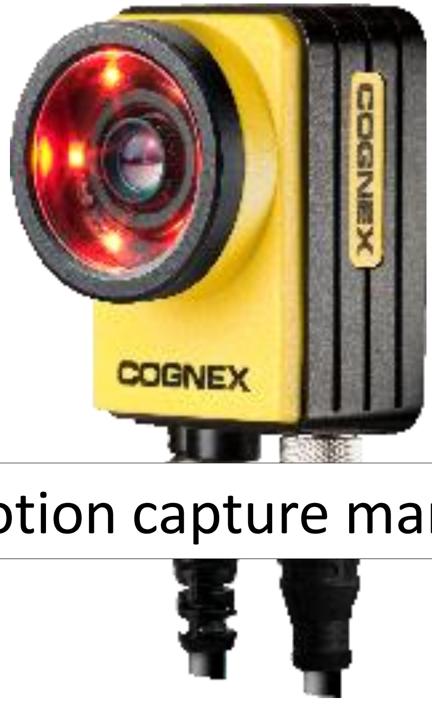
SENSING AND PERCEPTION



Consumer-grade sensors are being used in serious research, and in industrial applications

Credits: ASUS; Intel; Microsoft; Occipital

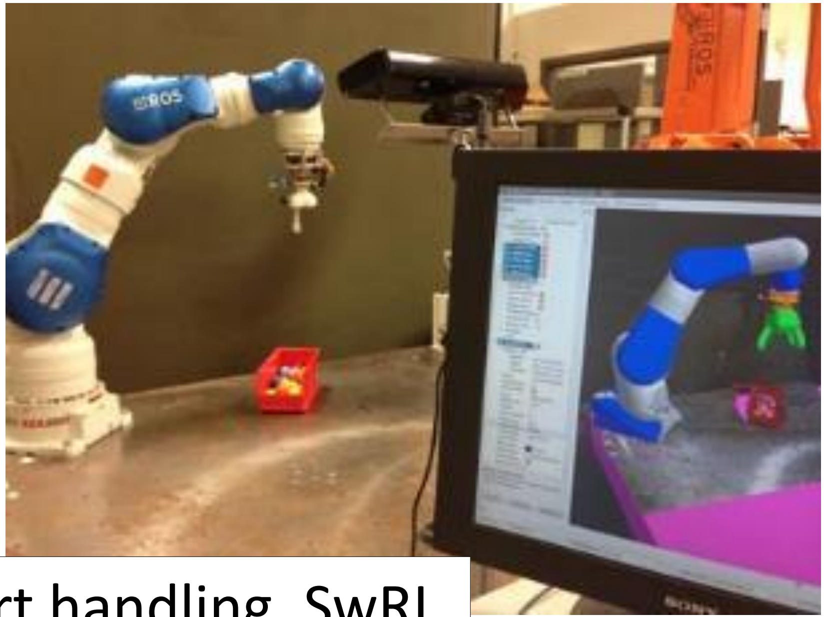
Lidar provides hi-resolution 2D and 3D range images directly



3D pose measurement is common, driven by the motion capture market

Credits: Cognex; Hokuyo; Optitrack; SICK; Velodyne; Vicon

Robot teaching, U of Tarapacá



Robot part handling, SwRI

Box depalletizing, Motoman

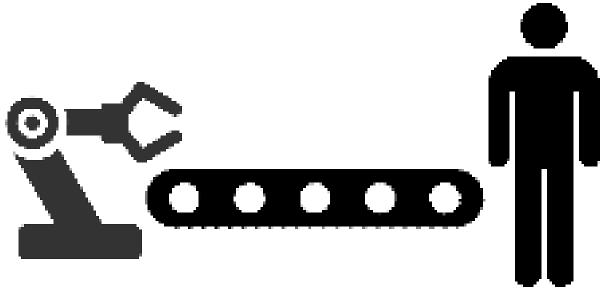

Chicken fileting, SINTEF

COLLABORATIVE ROBOTICS

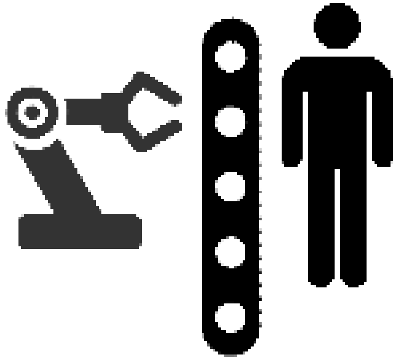
- Interactive activity involving the robot and a person
- Independent motion of the robot and person
- The robot and person share a workspace at the same time
- Physical contact between the robot and a person, directly or indirectly

Separate

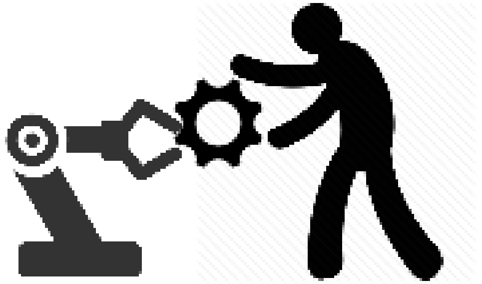
Sequential



Simultaneous



Supportive

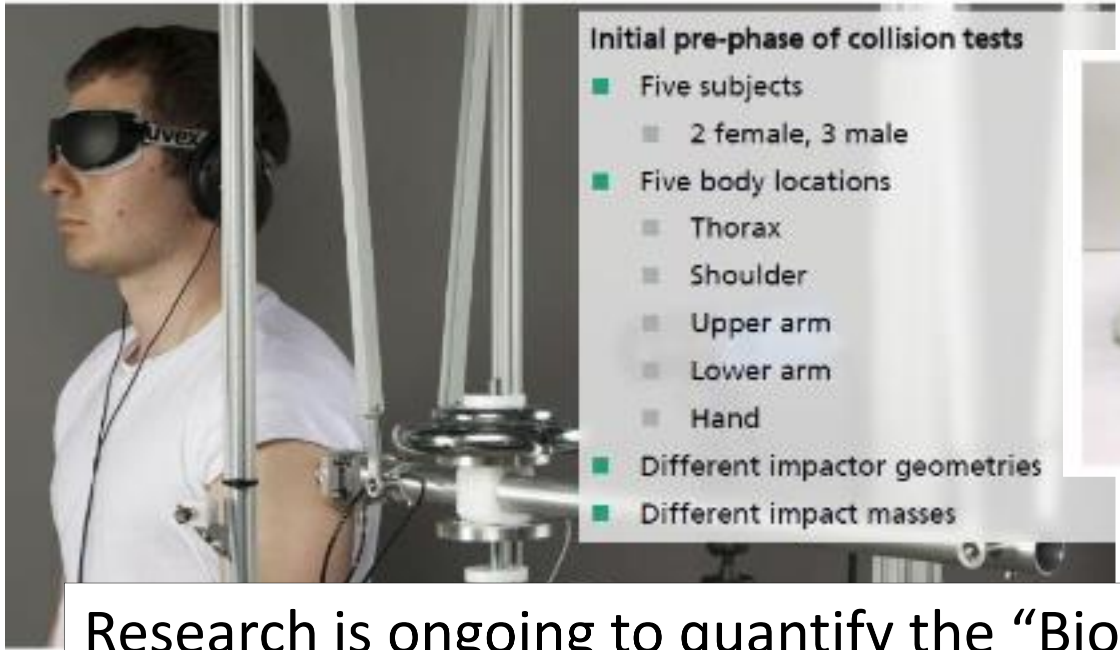


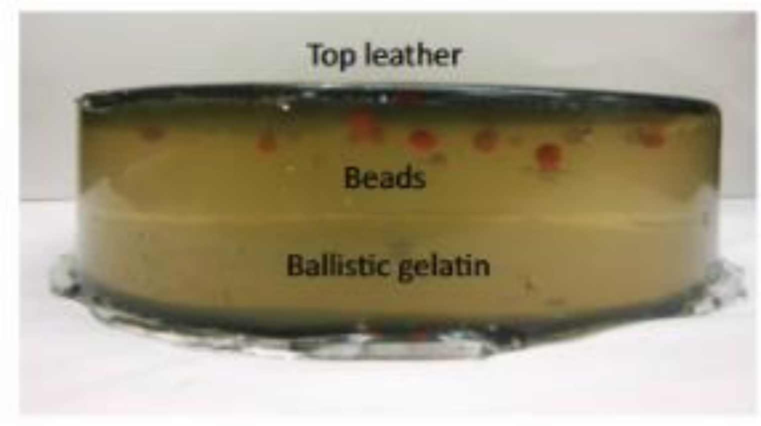
Speed and separation monitoring systems reduce the robot speed when people are close

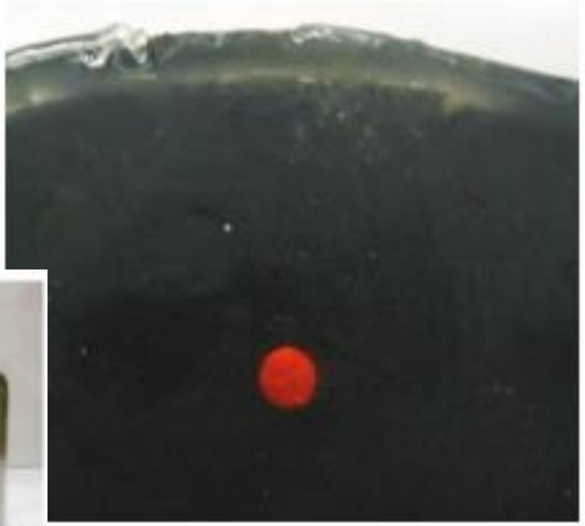
Power and force limiting on robots makes them safer

Credits: ABB; F&P; Fanuc; Kuka; Rethink; Universal

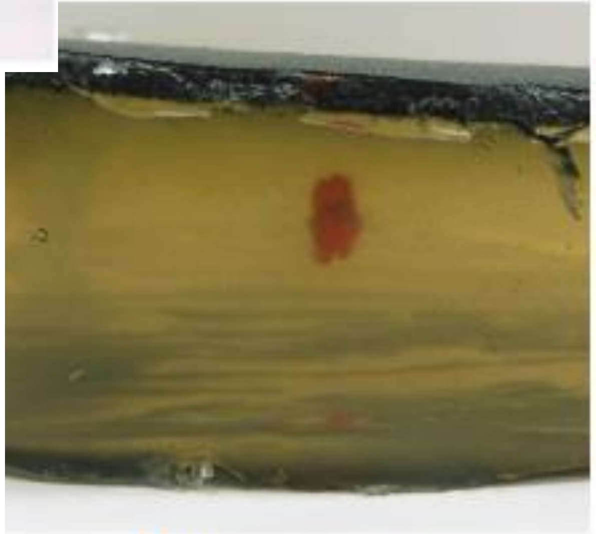
How badly can robots hurt people?







(a) Before impact testing



(b) After impact testing

Research is ongoing to quantify the “Biomechanical Limit Criteria” using human subjects

Credits: Fraunhofer IFF/Magdeburg; NIST

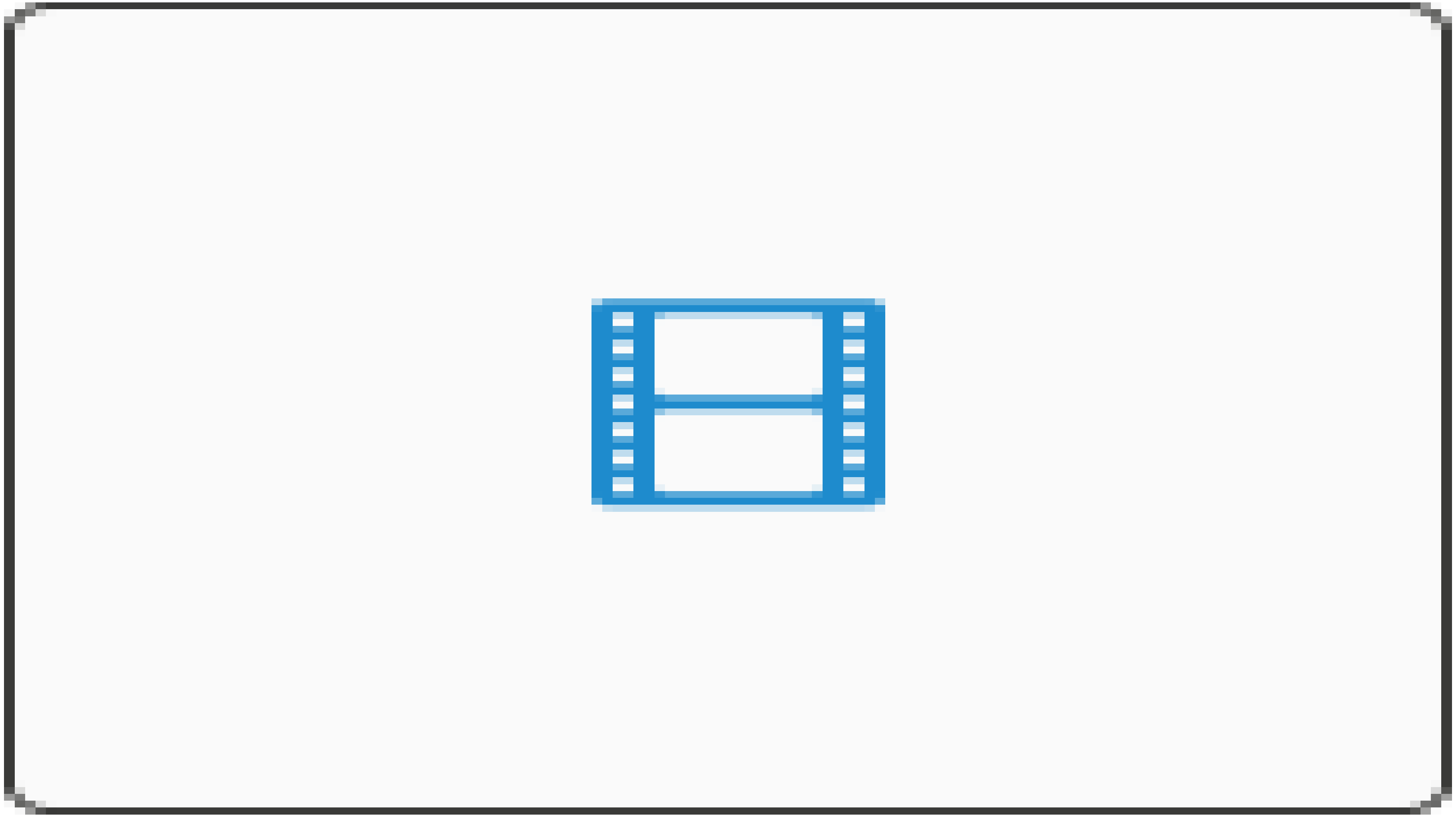
Sharing a kit unloading task

An interactive CNC grinding application

Working together to prepare samples

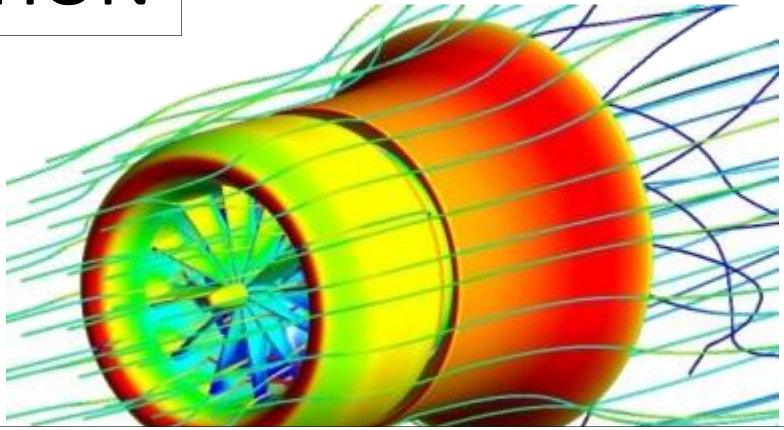
Hand guiding lets people move the robot directly

Credits: ABB; Kuka; Rethink; Universal



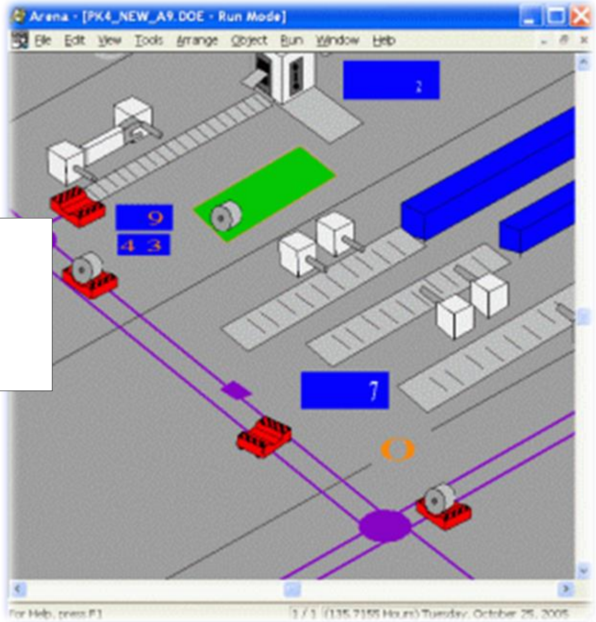
Robots with joint torque sensing can be hand guided

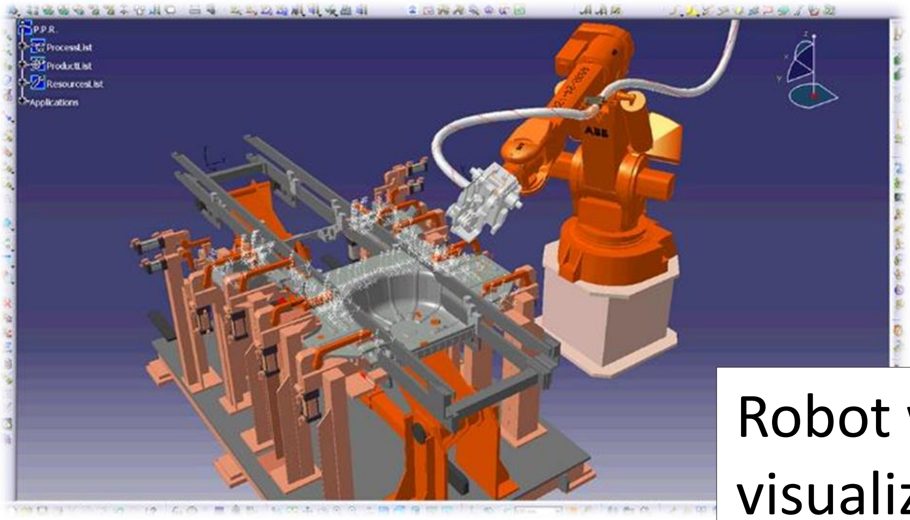
SIMULATION



Computational fluid dynamics + heat transfer

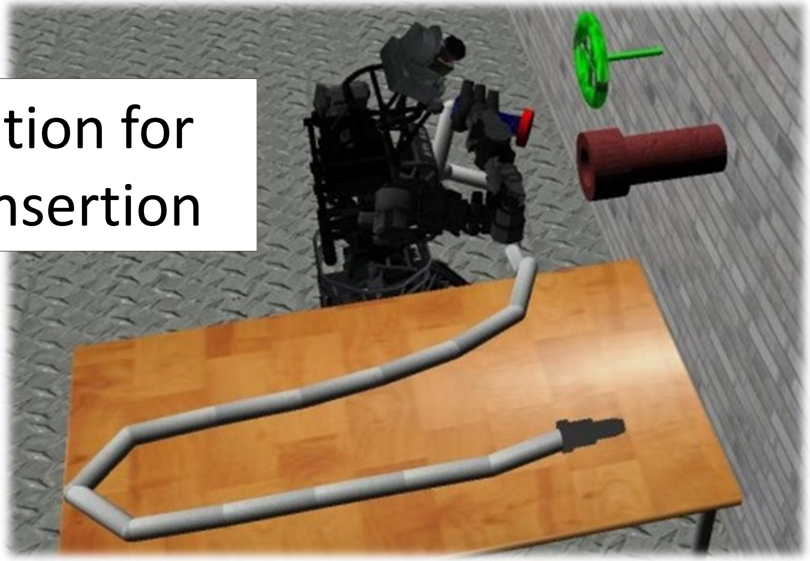
Discrete event simulation for estimating cycle time





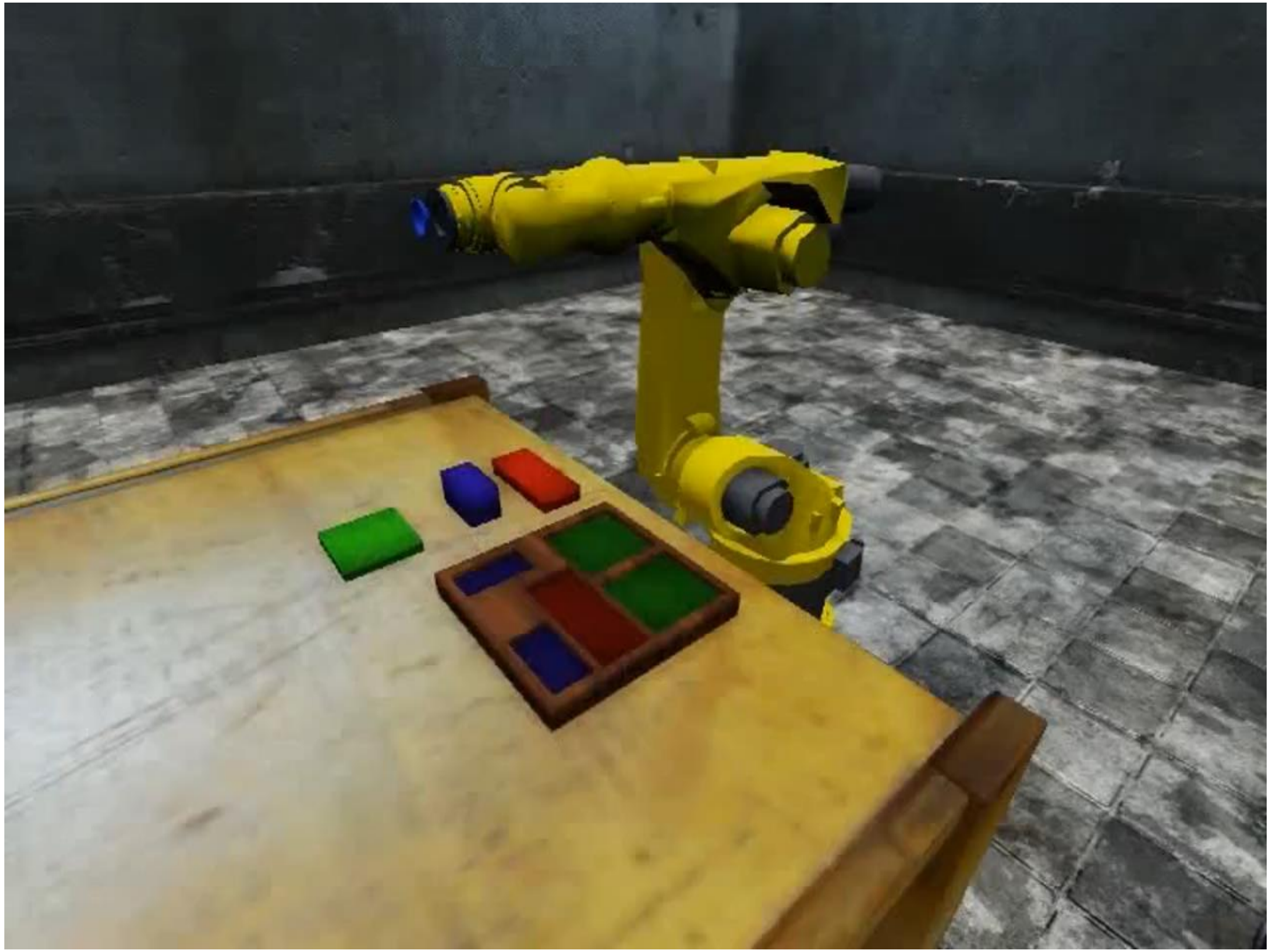
Robot workcell visualization

Physics simulation for robotic hose insertion



Credits: ANSYS; Arena; DARPA; Delmia

You can “cheat” in simulation to work around unmodeled effects



Objects can be made to snap to the robot, or magically appear

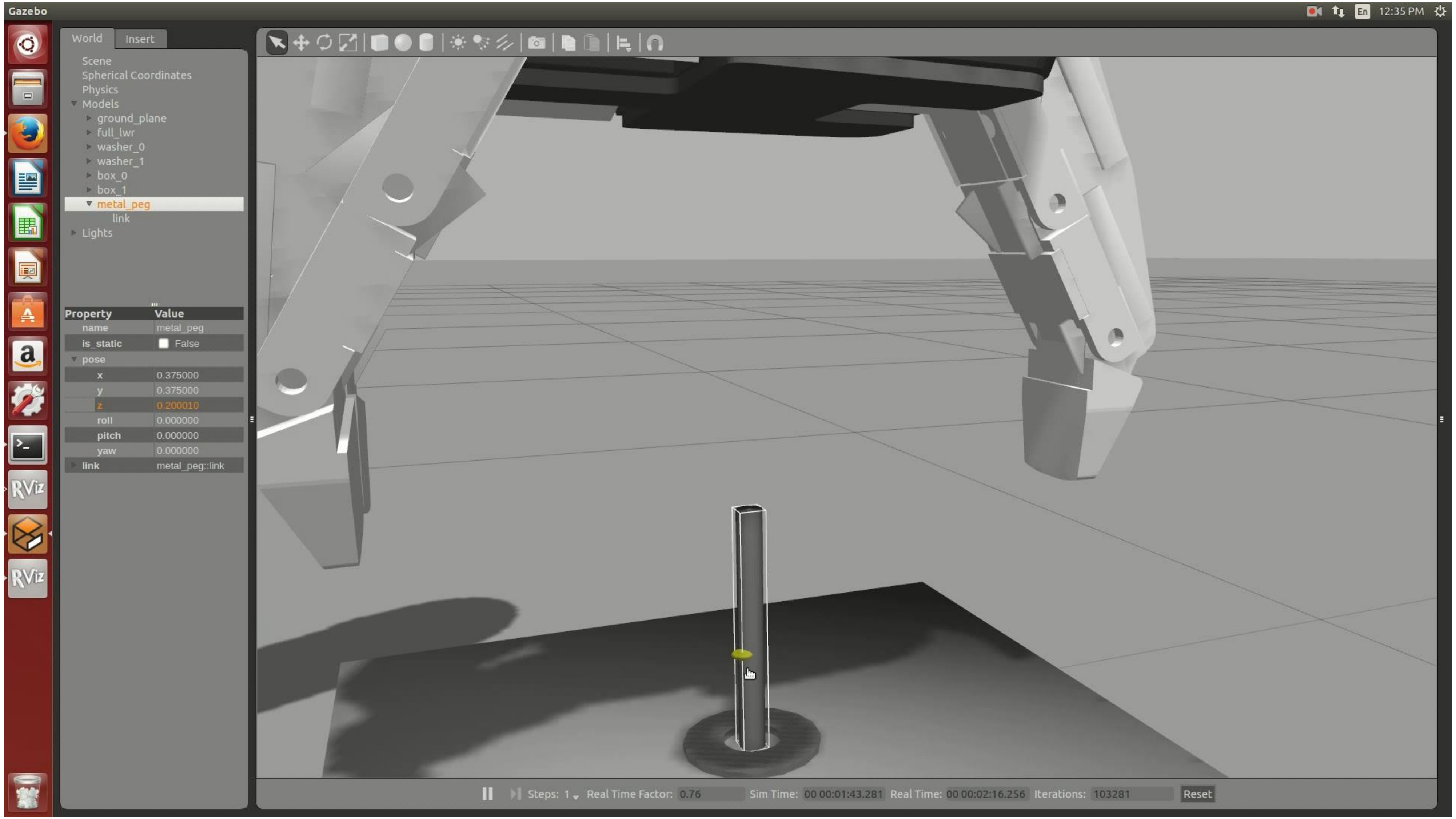
Physics simulations use today's high-performance graphical processing units (GPUs)

GPU arrays can provide scalable simulation power

Software tools like CUDA simplify tedious programming

and fill rooms

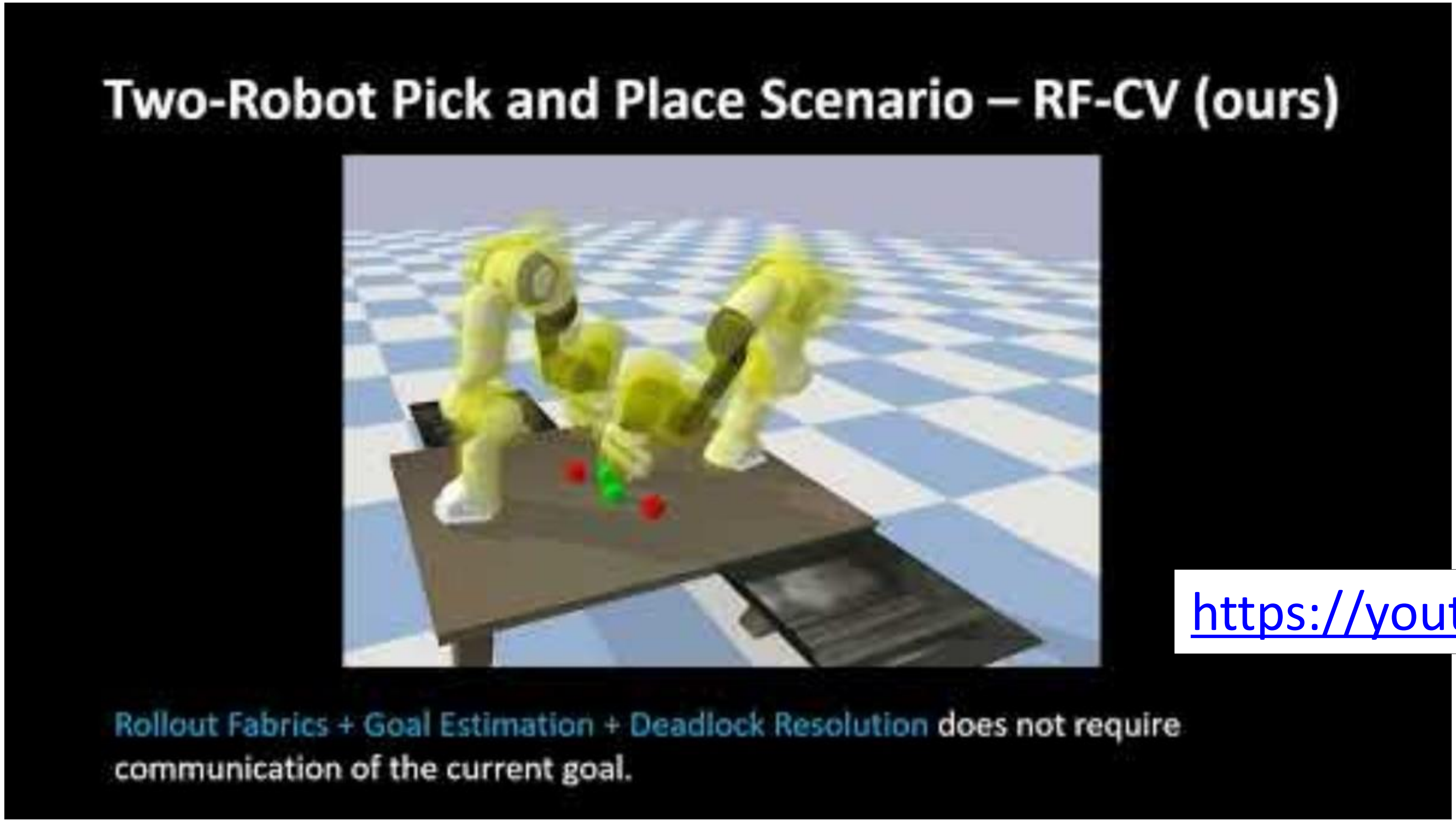
Credits: Elsa; HPC Paris; nVIDIA



Physics simulation in ROS and Gazebo showing robot and object interactions

PLANNING

Planning robot motion in changing environments is becoming more prevalent

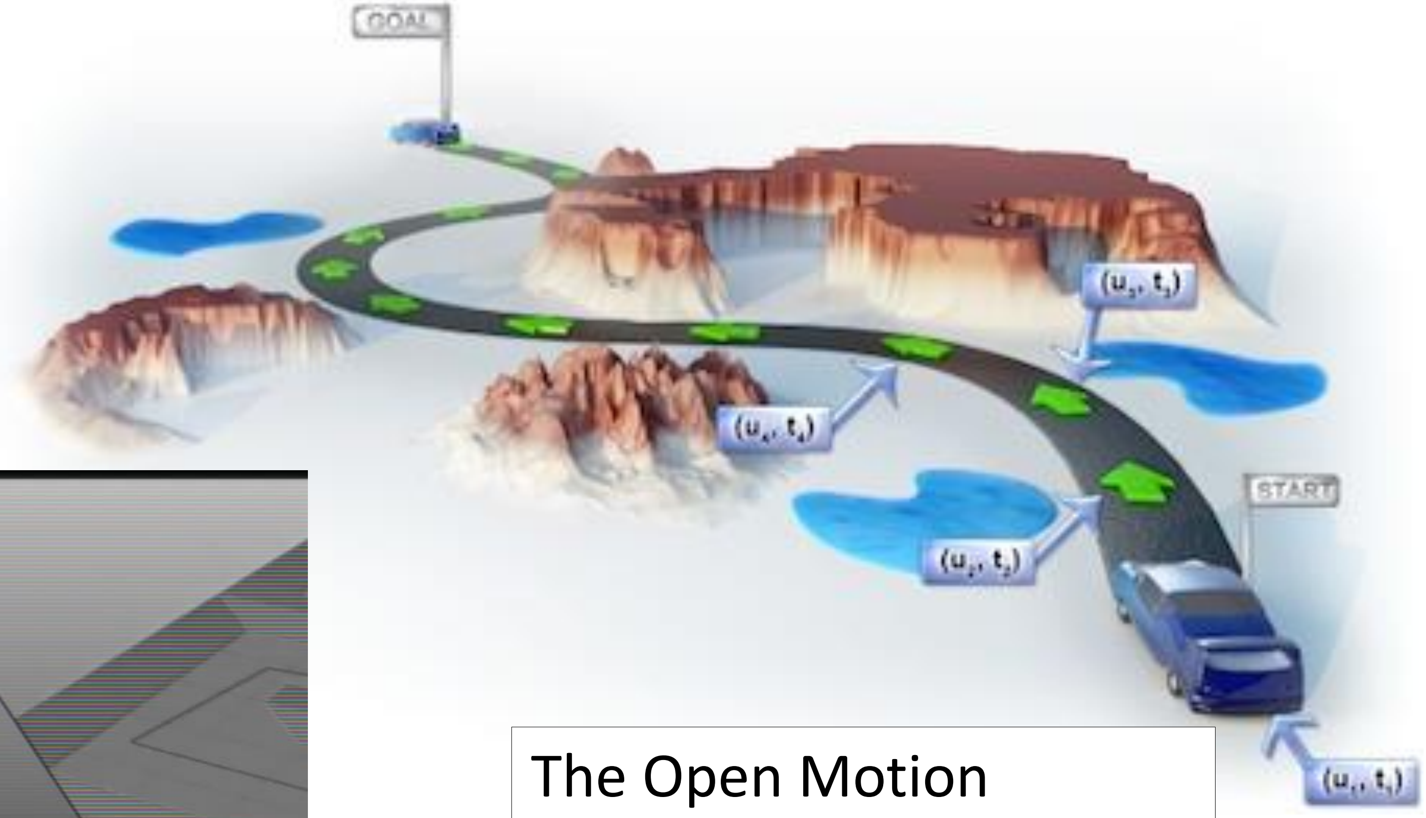


Extra degrees of freedom allow the selection of configurations that avoid collisions

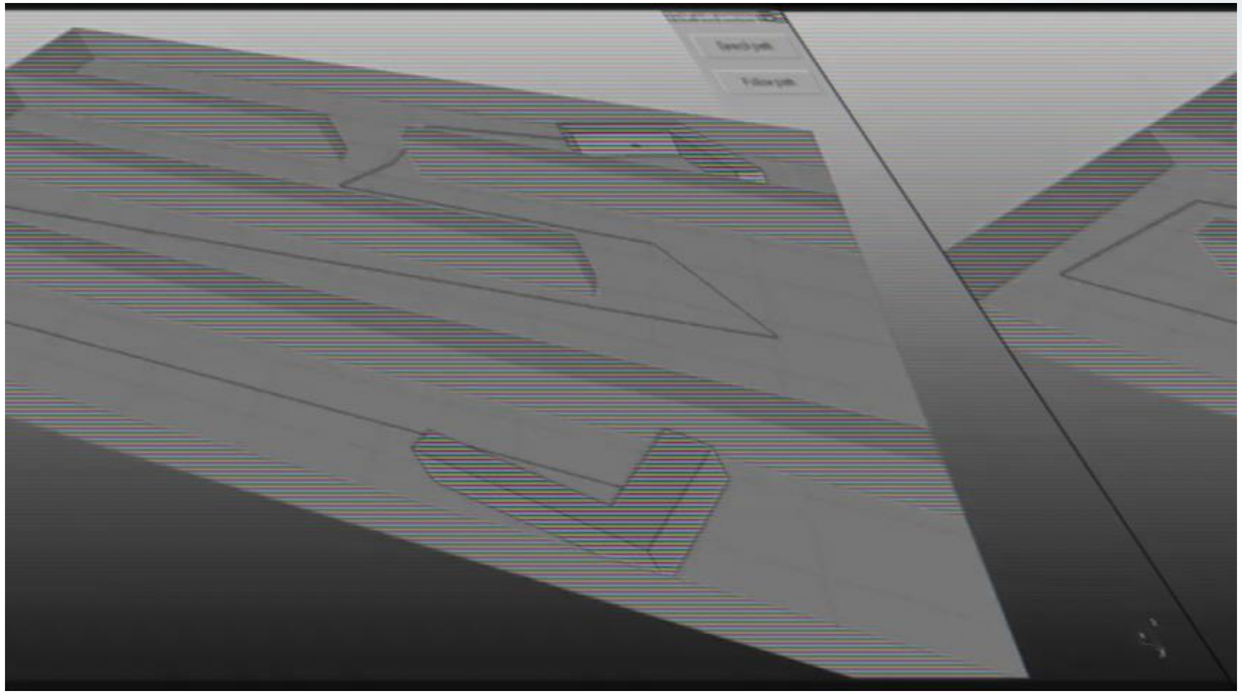
https://youtu.be/jaJBrSecDcM

Credits: Autonomous Multi-Robots Laboratory at the Delft University of Technology

ROS planning “plug-ins” can be used for various applications



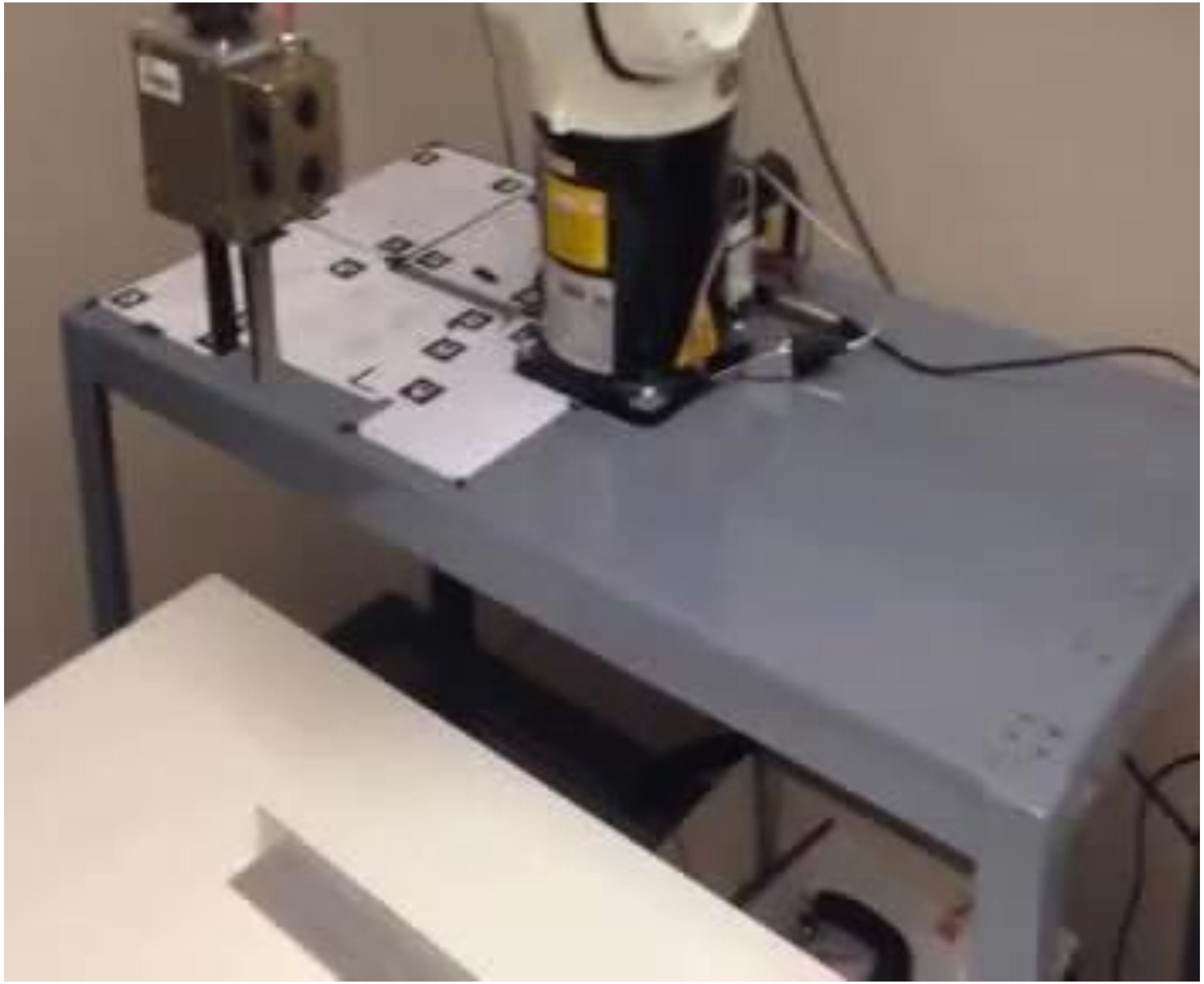
The Open Motion Planning Library (OMPL)



Path planning visualized in Gazebo

Credits: Timotei Robotics; Rice University

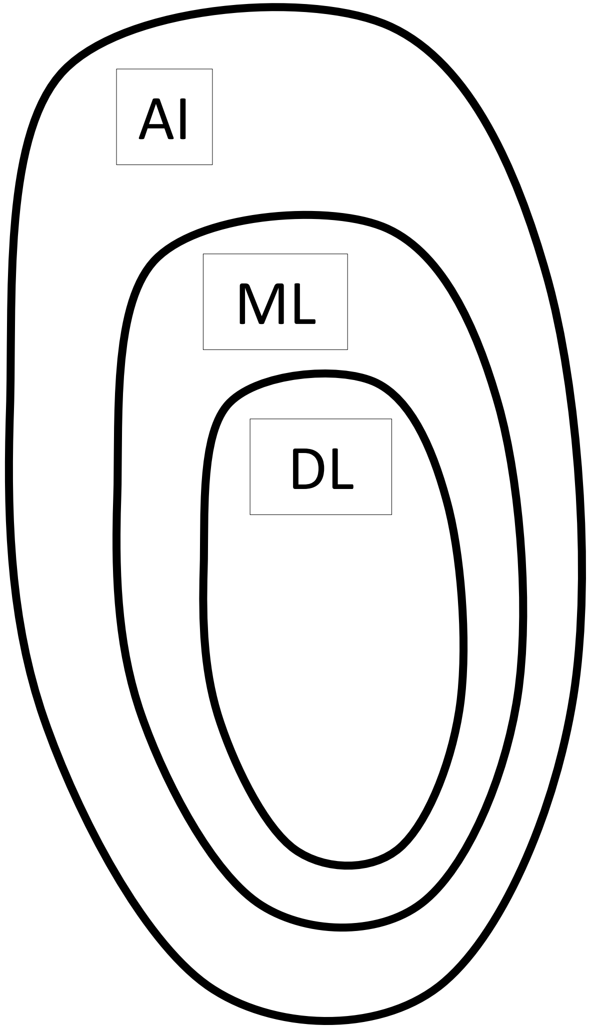
Knowledge-based planning uses information about actions and consequences to find unexpected solutions to problems



The search spaces can be extremely large; heuristics can be used to prune out probable dead ends, focus on promising leads

Credits: Georgia Tech

ARTIFICIAL INTELLIGENCE



Artificial Intelligence: the science and engineering of making intelligent machines
– Rules engines, expert systems, symbolic computing

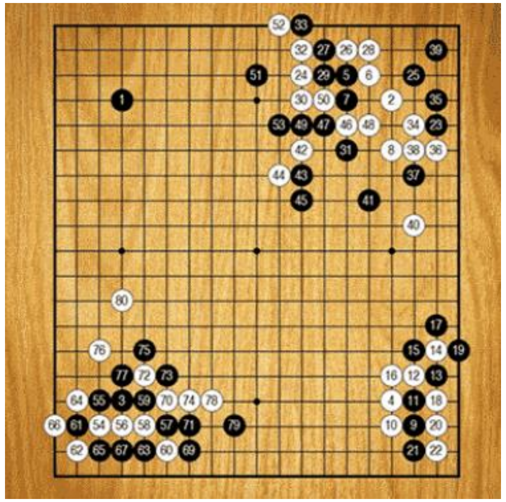
Machine Learning: the field of study that gives computers the ability to learn without being explicitly programmed
– Neural networks, objective function optimization

Deep Learning: neural networks with many layers

- Unsupervised learning: That thing is like this other thing.
- Supervised learning: That thing is a “double bacon cheese burger”. Labeling can be laborious, often manual.
- Reinforcement learning: Eat that thing because it tastes good and will keep you alive longer.

Credits: Skymind.ai

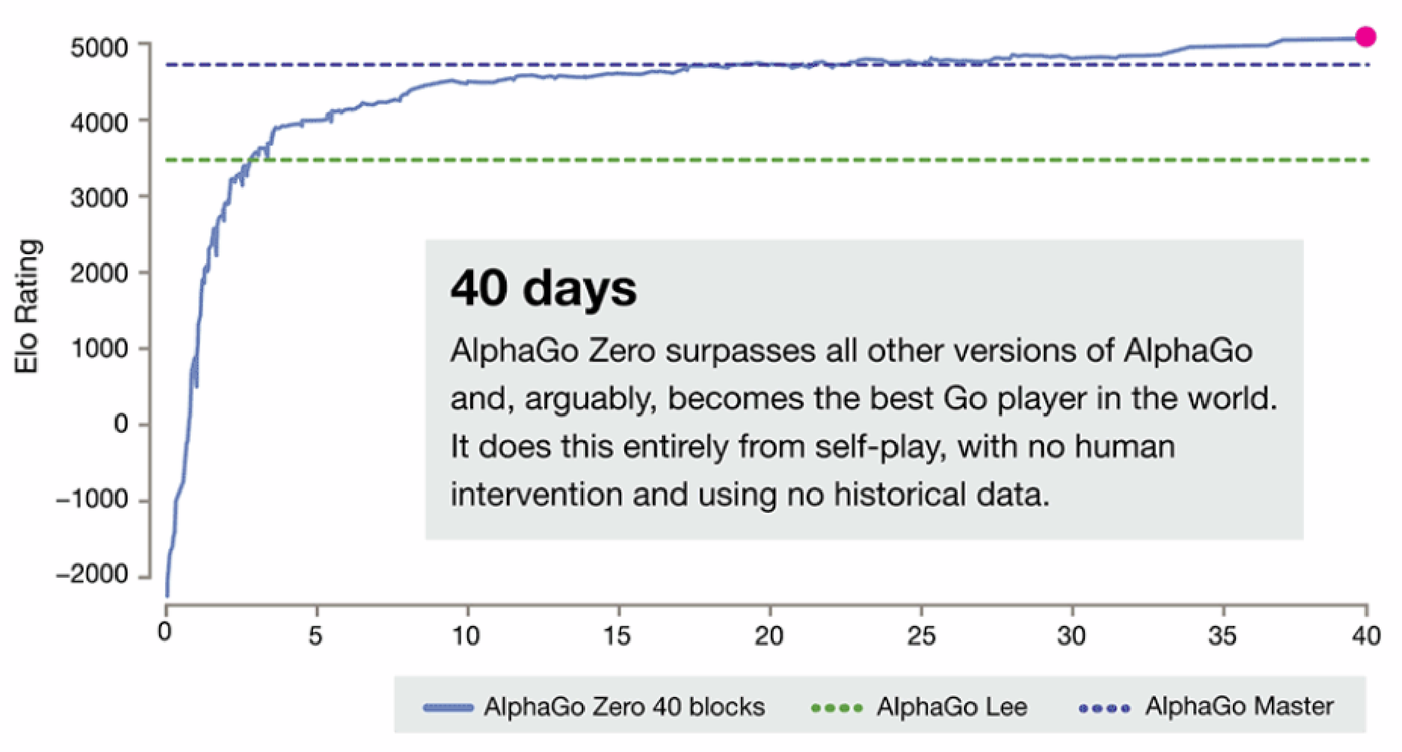
AlphaGo Zero: Deep reinforcement learning from scratch



Game of Go: place stones on a 19x19 grid to surround more of your opponent's stones

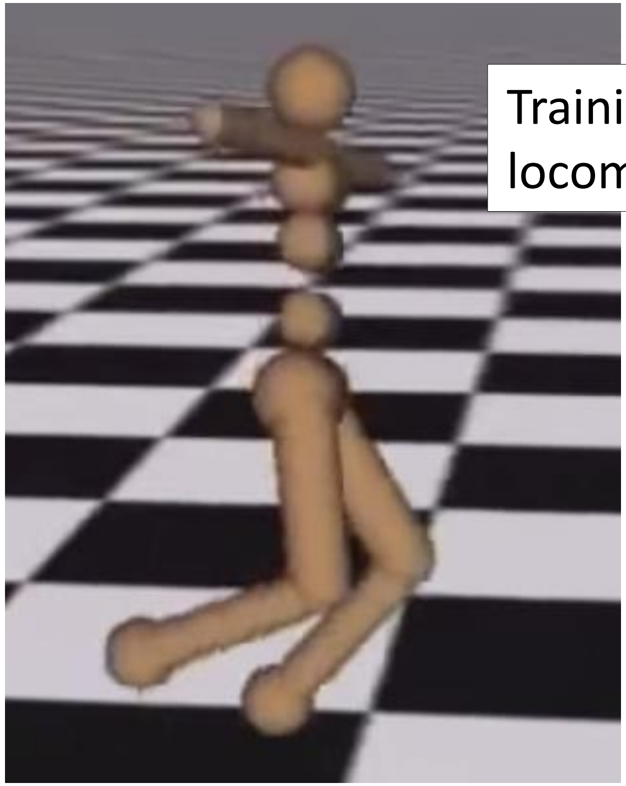
Much more complex than chess; at least 10^{170} board positions

- AlphaGo was the first computer program to beat the world champion, in March 2016
- AlphaGo trained on thousands of professional games
- AlphaGo **Zero** learns to play by playing games against itself, starting from completely random play



Credits: Deepmind.com

AI and DRL for Robotics



Training in simulation: locomotion; robo-sumo

Control of difficult-to-model systems: NASA’s Super Ball Bot

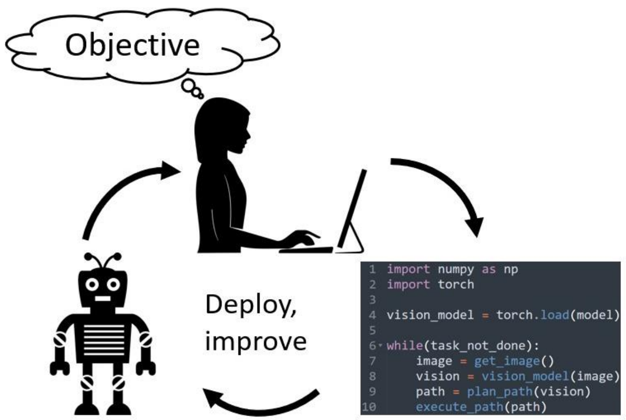
Understanding tactile sensing for planning

Credits: Prof. Pieter Abbeel et al.

ChatGPT – an AI Chatbot

- A *chatbot* is a computer program that uses natural language to converse with people
- ChatGPT uses the GPT-3 large-language model as the foundation, and has been applied to many non-chat uses:
  - writing and debugging computer programs
  - composing artwork, music, poetry, and student essays
  - answering test questions

Robotics today: engineer *in the loop*



Microsoft 2023-2-20

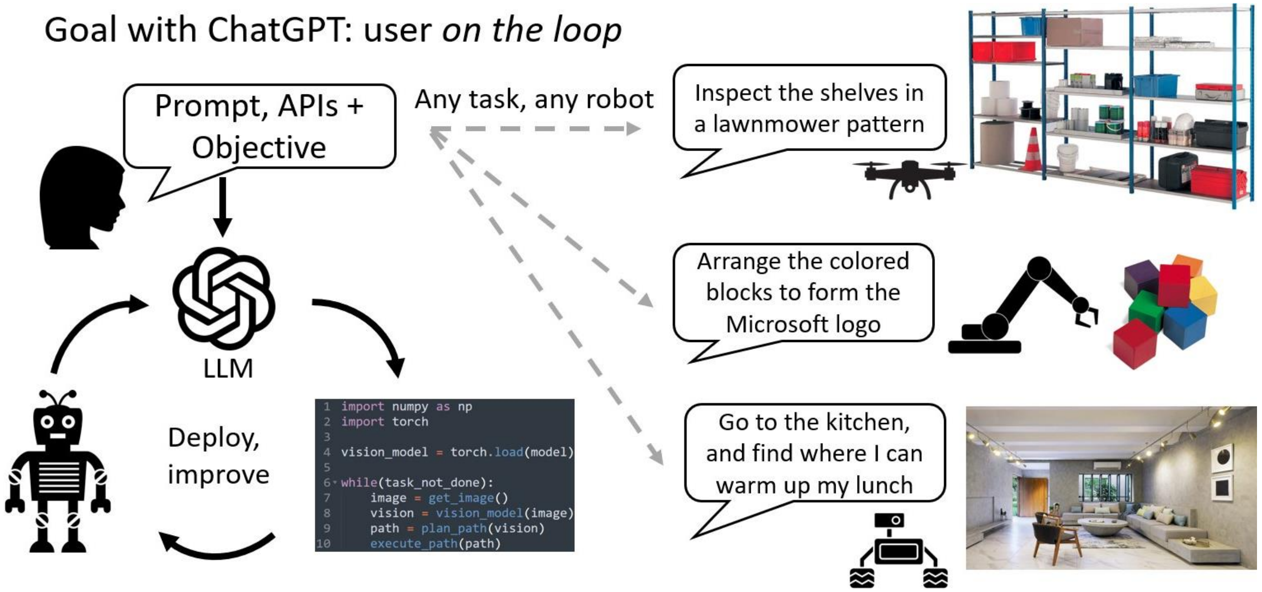
ChatGPT for Robotics:
Design Principles and Model Abilities

Sai Vemprala*, Rogerio Bonatti*, Arthur Bucker , and Ashish Kapoor
Microsoft Autonomous Systems and Robotics Research

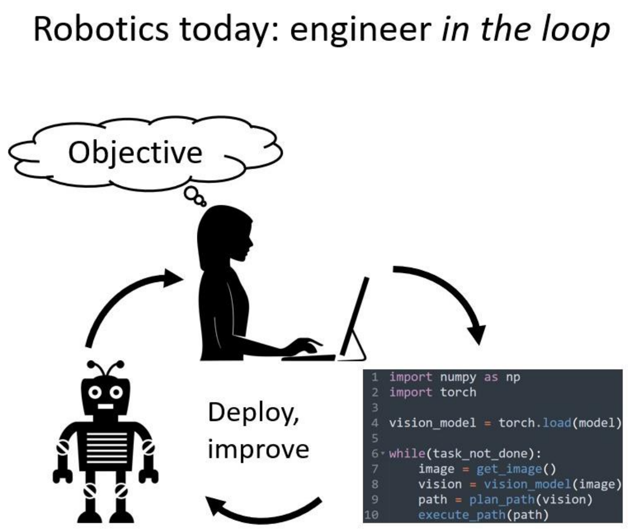
This paper presents an experimental study regarding the use of OpenAI's ChatGPT [1] for robotics applications. We outline a strategy that combines design principles for prompt engineering and the creation of a high-level function library which allows ChatGPT to adapt to different robotics tasks, simulators, and form factors. We focus our evaluations on the effectiveness of different prompt engineering techniques and dialog strategies towards the execution of various types of robotics tasks. We explore ChatGPT's ability to use free-form dialog, parse XML tags, and to synthesize code, in addition to the use of task-specific prompting functions and closed-loop reasoning through dialogues. Our study encompasses a range of tasks within the robotics domain, from basic logical, geometrical, and mathematical reasoning all the way to complex domains such as aerial navigation, manipulation, and embodied agents. We show that ChatGPT can be effective at solving several of such tasks, while allowing users to interact with it primarily via natural language instructions. In addition to these studies, we introduce an open-sourced research tool called *PromptCraft*, which contains a platform where researchers can collaboratively upload and vote on examples of good prompting schemes for robotics applications, as well as a sample robotics simulator with ChatGPT integration, making it easier for users to get started with using ChatGPT for robotics.

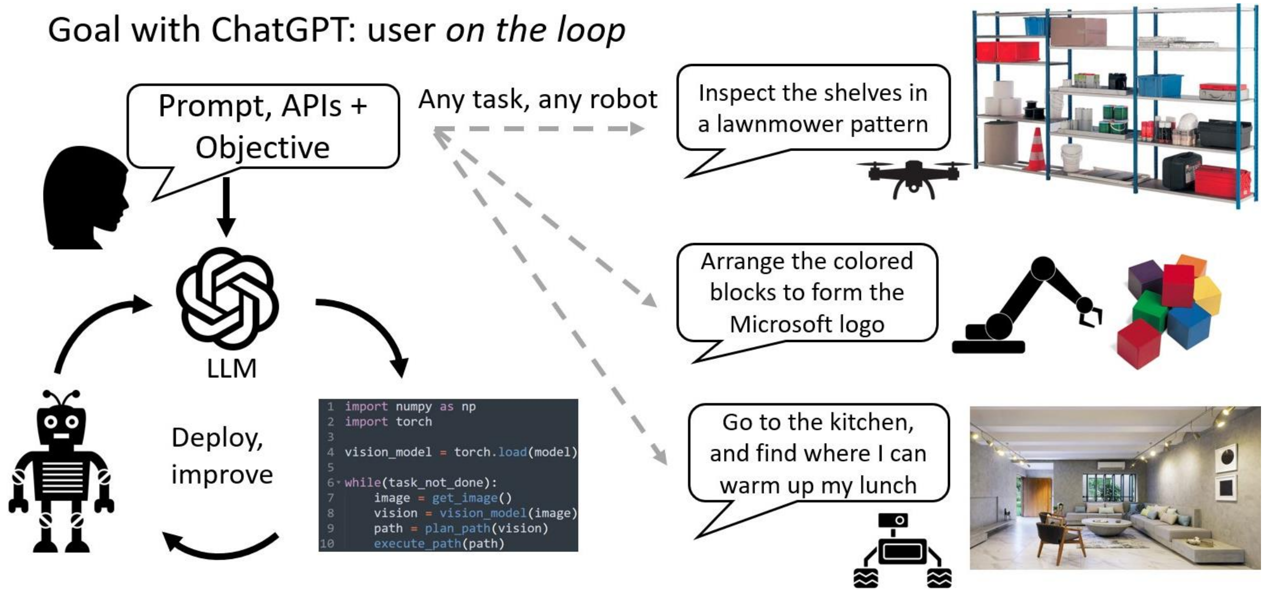
Videos and blog: aka.ms/ChatGPT-Robotics
PromptCraft, AirSim-ChatGPT code: https://github.com/microsoft/PromptCraft-Robotics

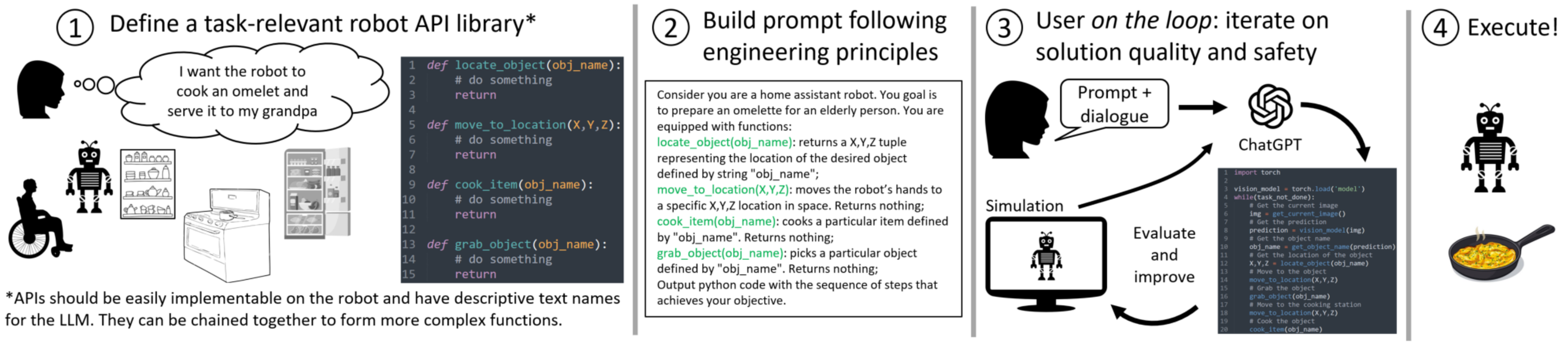
Goal with ChatGPT: user *on the loop*



ChatGPT for Robotics







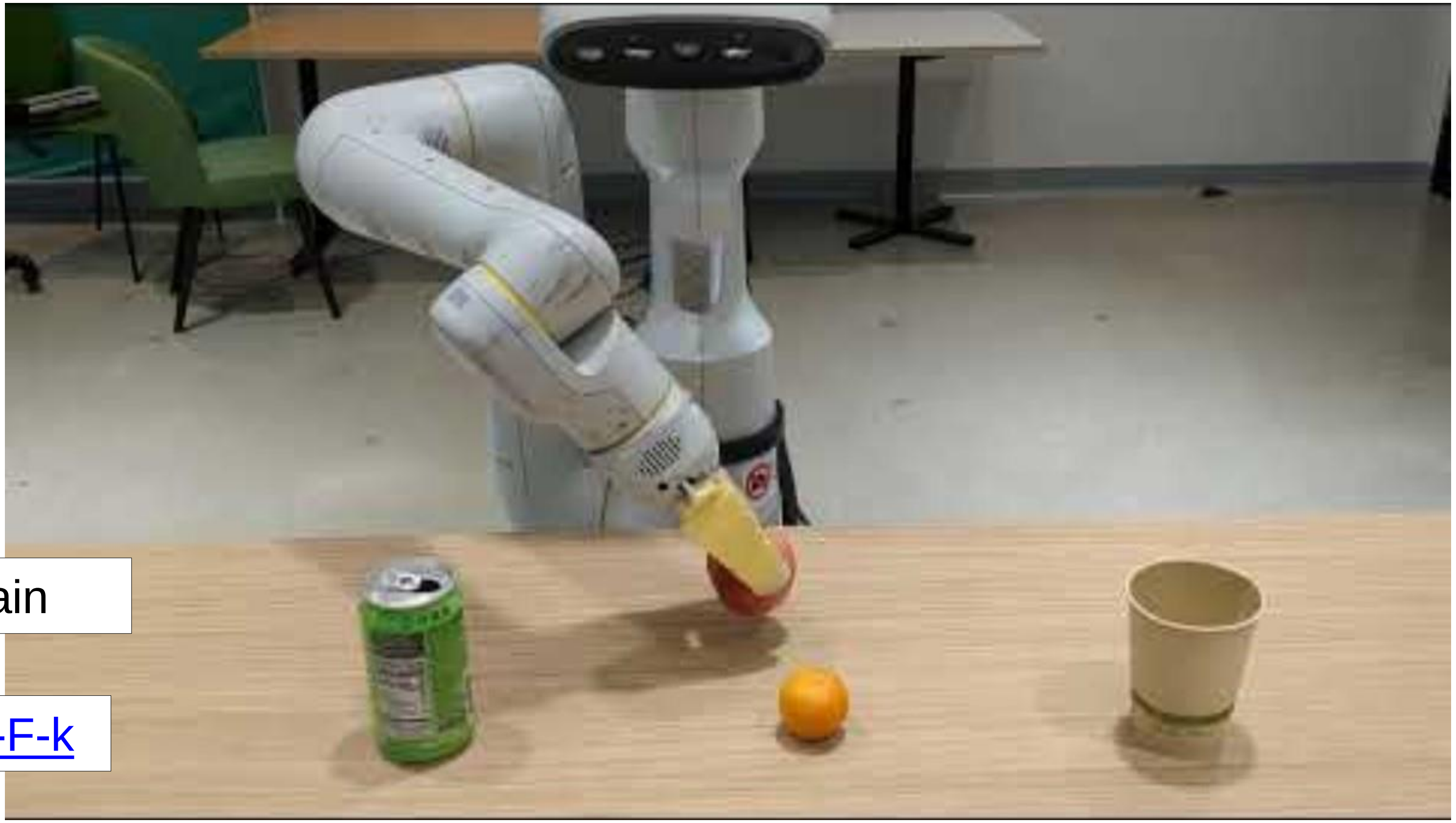
www.microsoft.com/en-us/research/group/autonomous-systems-group-robotics/articles/chatgpt-for-robotics/

Addressing the Big Data training set problem for robotics

Unfortunately, the highly successful generative AI formula doesn’t easily carry over into robotics because the Internet is not full of robotic-interaction data in the same way that it’s full of text and images

In 2023, 32 robotics laboratories in North America, Europe, and Asia started the RT-X Project to assemble data, resources, and code to make general-purpose robots a reality

The RT-X dataset contains nearly a million robotic trials for 22 types of robots and behaviors including picking and placing objects, assembly, and specialized tasks like cable routing, with about 500 skills and interactions with thousands of different objects – the largest open-source dataset of real robotic actions in existence



https://spectrum.ieee.org/global-robotic-brain

https://youtu.be/qSARoad-F-k

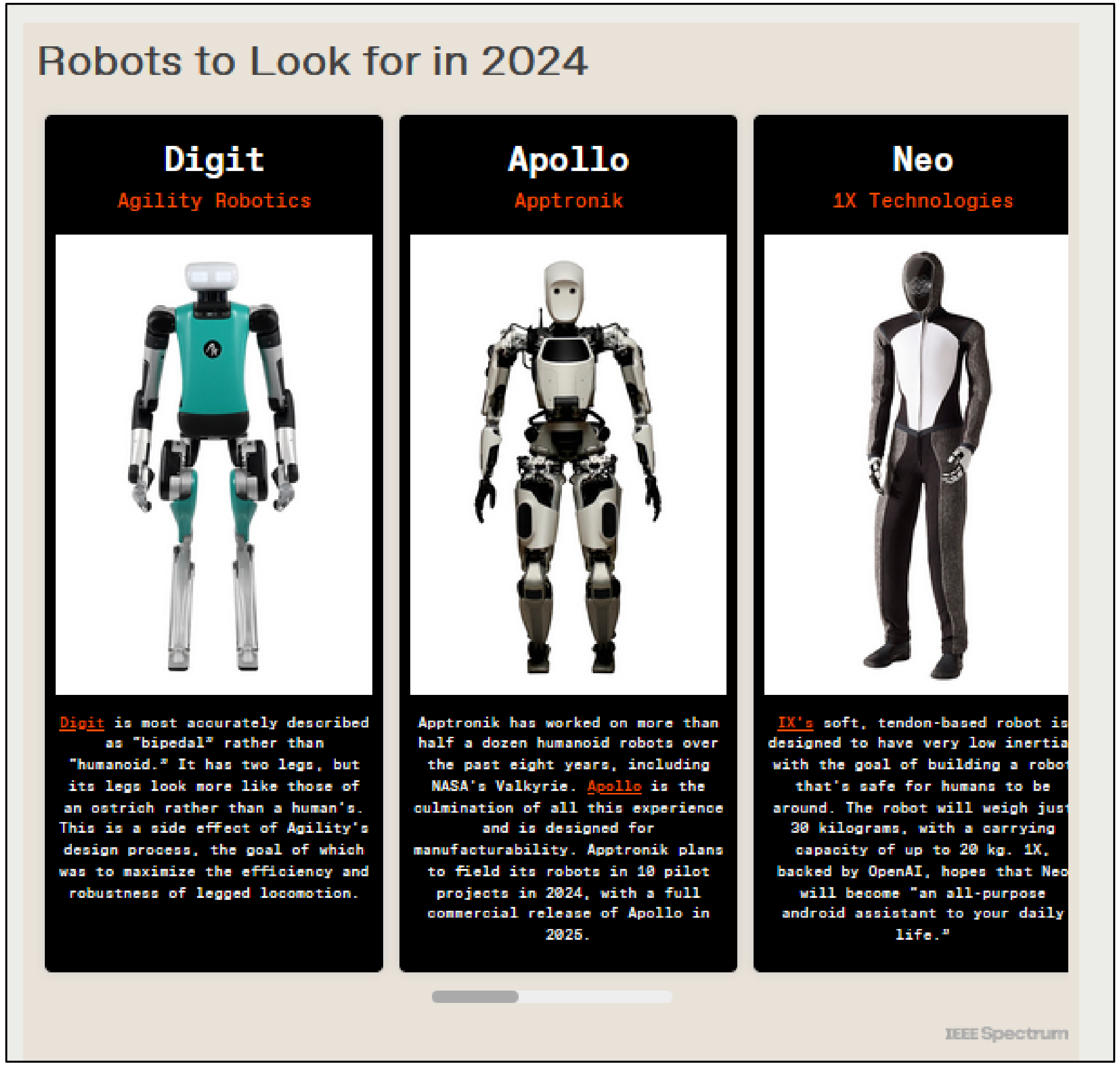
After the DARPA Robotics Challenge ended in 2015 with several of the robots successfully completing the entire final scenario, the obvious question was: When would humanoid robots make the transition from research project to a commercial product?

FEATURE | ROBOTICS

Humanoid Robots Are Getting to Work › Humanoids from Agility Robotics and seven other companies vie for jobs

30 DEC 2023 | 7 MIN READ

Amazon is testing Agility Robotics' Digit for use in its warehouse operations, one of the companies it invested in as part of the Amazon Industrial Innovation Fund



Robots to Look for in 2024

Digit
Agility Robotics

Digit is most accurately described as "bipedal" rather than "humanoid." It has two legs, but its legs look more like those of an ostrich rather than a human's. This is a side effect of Agility's design process, the goal of which was to maximize the efficiency and robustness of legged locomotion.

Apollo
Apptronik

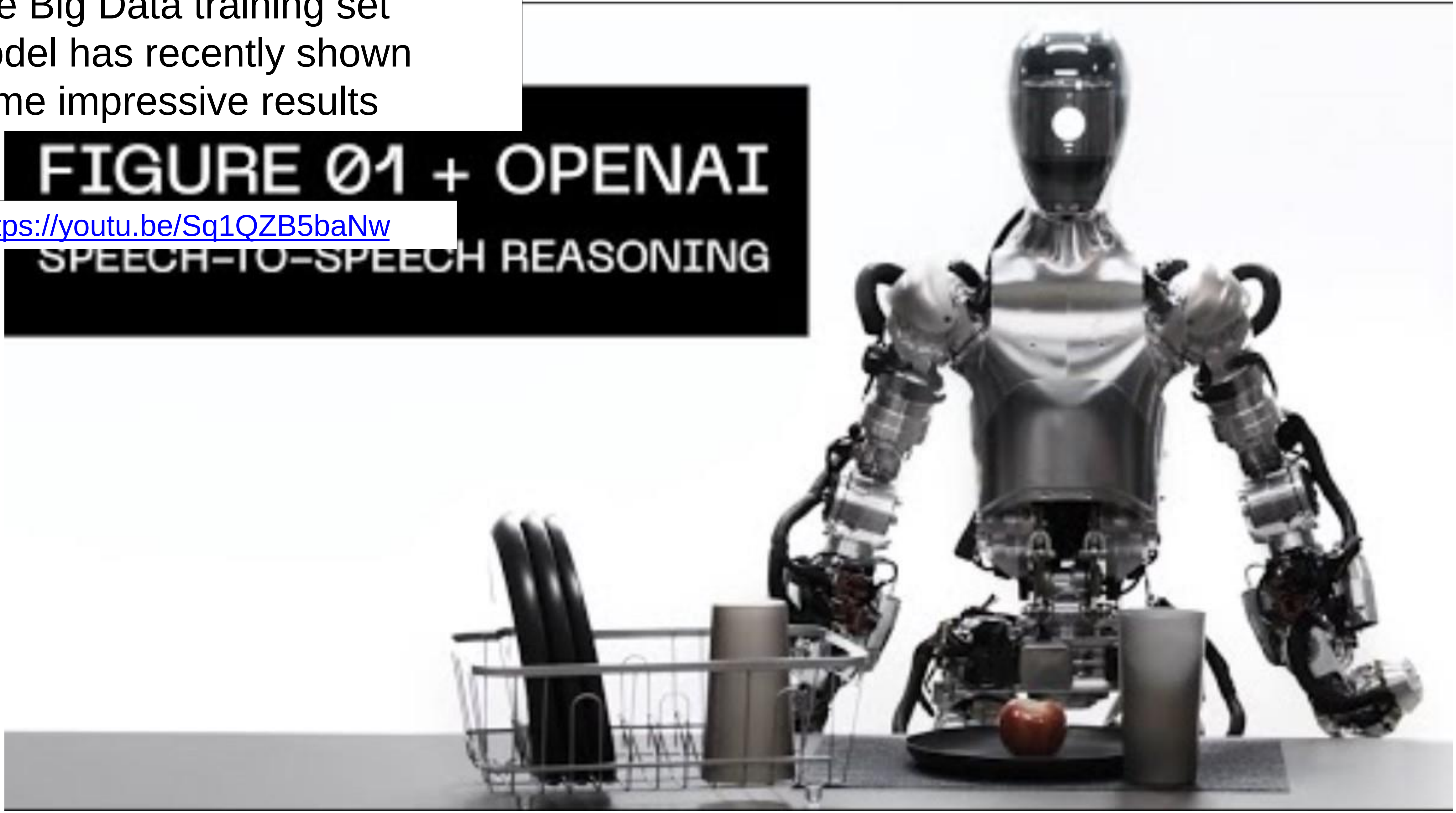
Apptronik has worked on more than half a dozen humanoid robots over the past eight years, including NASA's Valkyrie. Apollo is the culmination of all this experience and is designed for manufacturability. Apptronik plans to field its robots in 10 pilot projects in 2024, with a full commercial release of Apollo in 2025.

Neo
1X Technologies

IX's soft, tendon-based robot is designed to have very low inertia with the goal of building a robot that's safe for humans to be around. The robot will weigh just 30 kilograms, with a carrying capacity of up to 20 kg. 1X, backed by OpenAI, hopes that Neo will become "an all-purpose android assistant to your daily life."

IEEE Spectrum

The Big Data training set model has recently shown some impressive results



https://youtu.be/Sq1QZB5baNw


Questions?