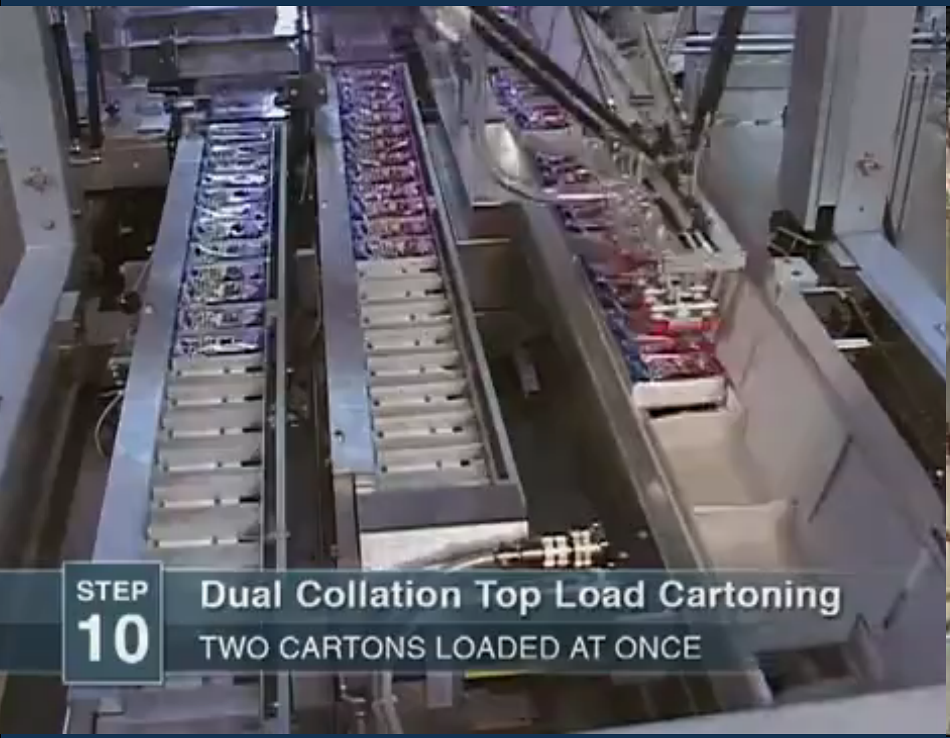
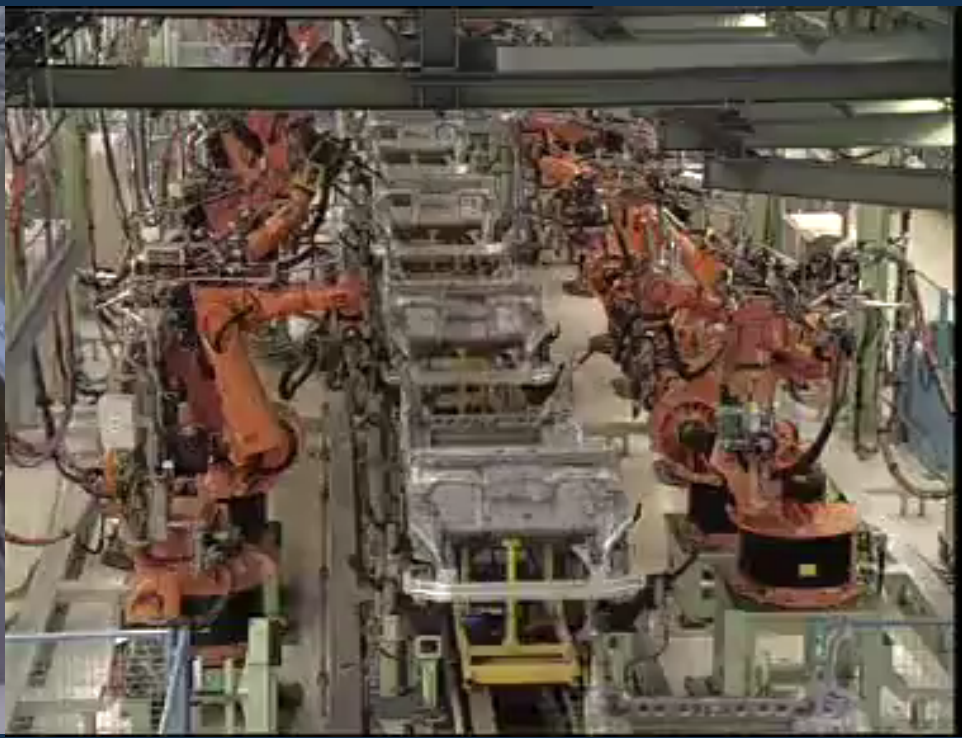
Tutorial on Learning from Demonstration

Part 1: Introduction

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Factory Robots





Industrial tasks still performed by Humans

- Manipulation tasks that require high dexterity
  → precise position and force control.
- Tasks that are versatile with limited series.

Polishing

Precise insertion

Learning from Human Demonstrations: Principle

- Transfer to the robot skills that took years for the humans to master.
- Human can quickly re-train the robot to adapt to task changes.
- The human teaches by showing how to perform the task.

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

The Multiple Promises

It offers a user-friendly means of teaching
→ No need to learn how to program the robot

It bootstraps learning by starting from a “good” example
→ Reduces search space for a good solution

Not Just Record and Replay: Generalize!

Hypotheses

Take humans as example

→ Hypothesis: Human is a good example!

Human teaches the robot by showing how to perform the task

→ Hypothesis: There is an interface to teach the robot

Which interface?

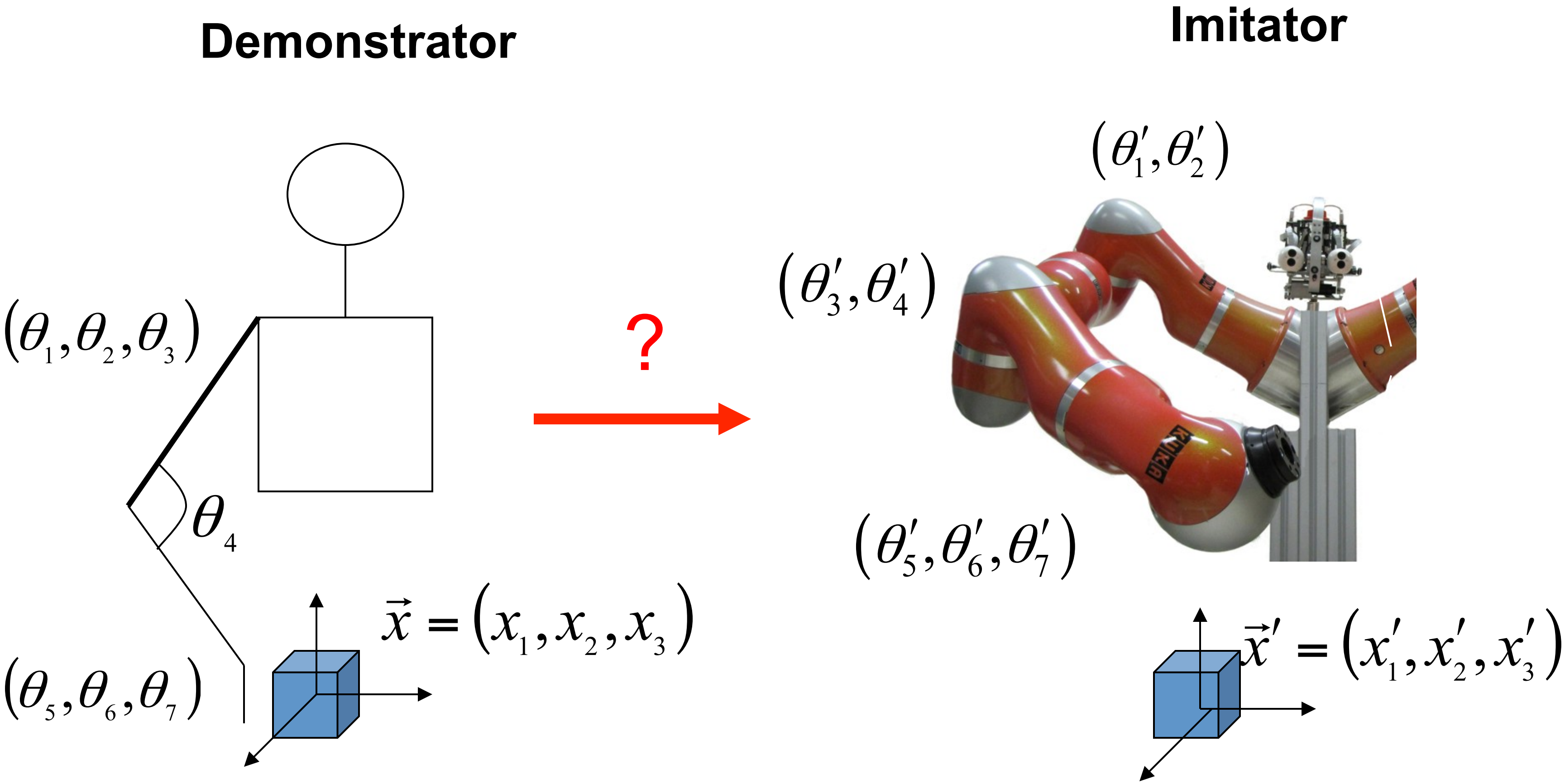
Vision:

Pros:
- Unobtrusive
- Record information on whole body.

Cons:
- Correspondence problem.
- No haptic information

Full body motion tracking using vision. Ude et al 2004

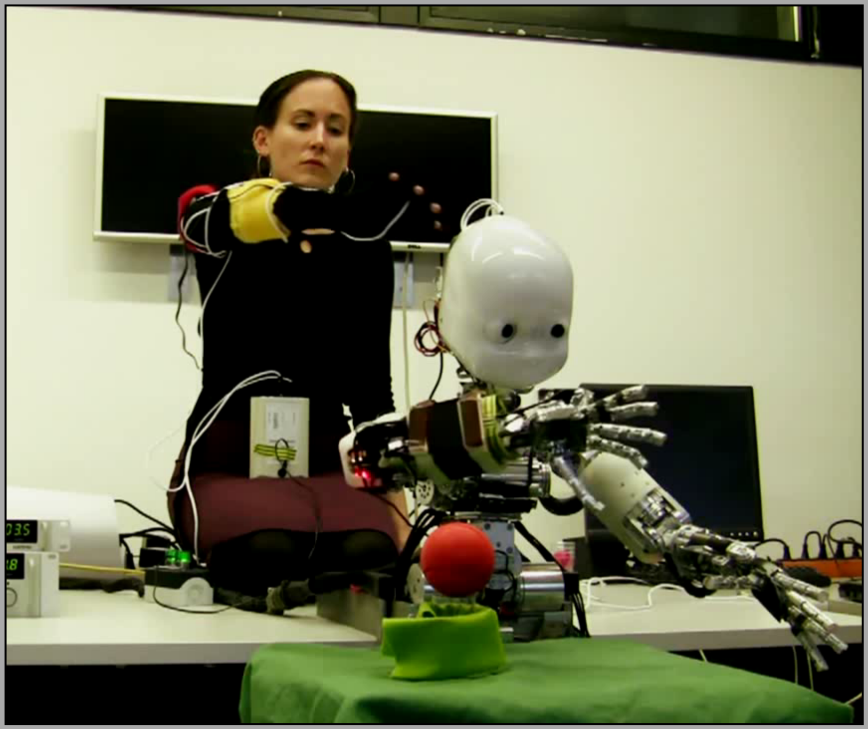
Correpondence Problem



Establish a correspondence across degrees of freedom when feasible.

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Which interface?



Motion sensors:

Pros:

- Real-time kinematic information
- Solve correspondence problem

Cons:

- Require to wear the system
- No haptic information

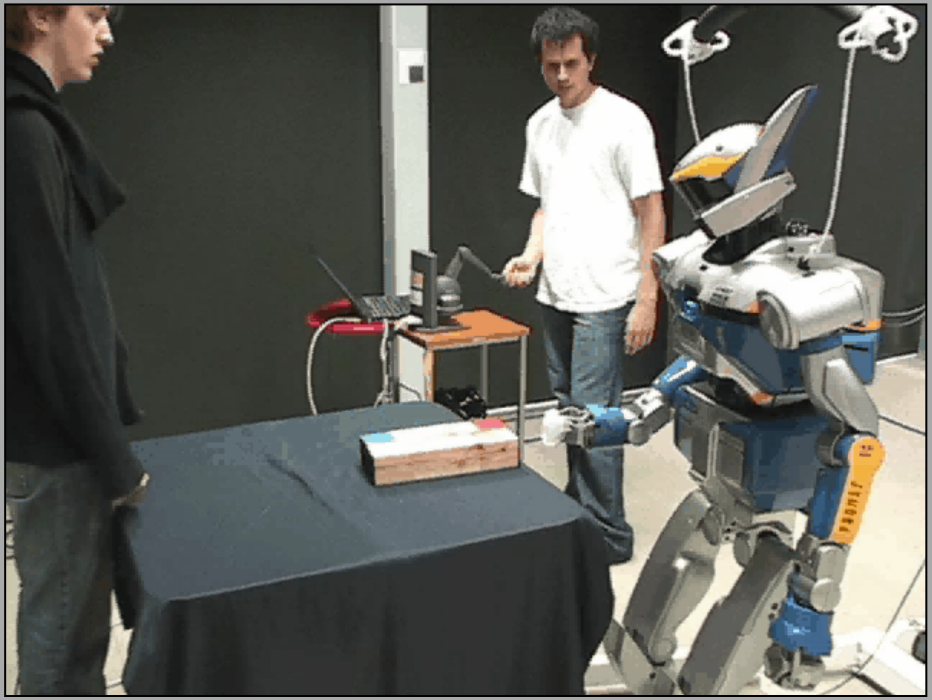
Which interface?

Kinesthetic Teaching:

Pros:
- Solve correspondence problem
- Transmit kinematic & haptic information

Cons:
- Need two hands to teach movements of a few DOFs

Which interface?



Haptic devices:

Pros:
- Solve correspondence problem
- Transmit kinematic & haptic information

Cons:
- Requires training
- User far from task location

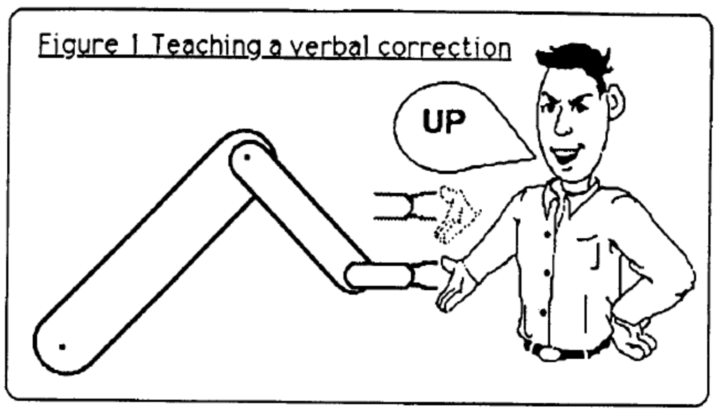
Equivalent Terms in the Literature

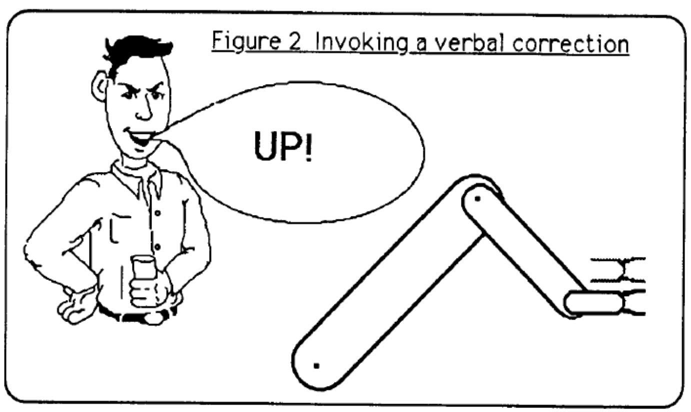
Robot Programming by Demonstration
Imitation Learning in Robots
Apprenticeship Learning
Robot Learning from Demonstration

Short History

• Programming by examples (Lauriere, 1978; Bauer, 1979)

• MacDonald, Bruce A., and Witten, Ian H. "Programming computer controlled systems by non-experts." (1987).

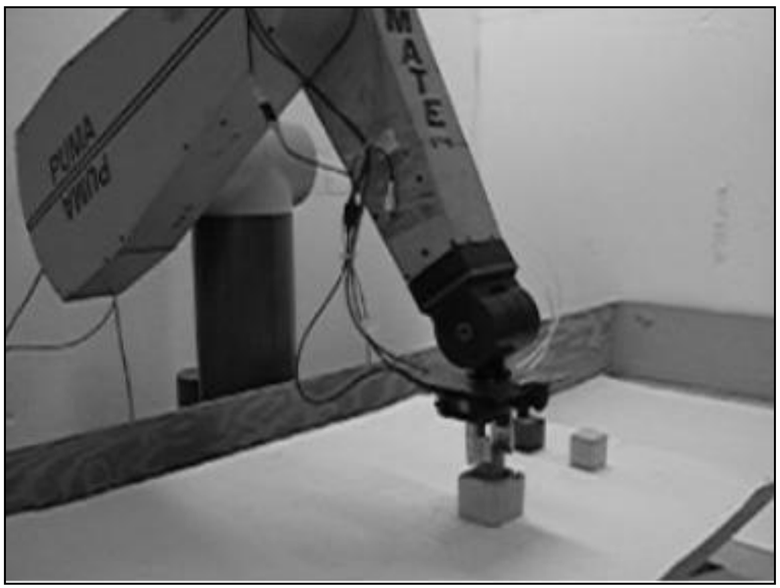


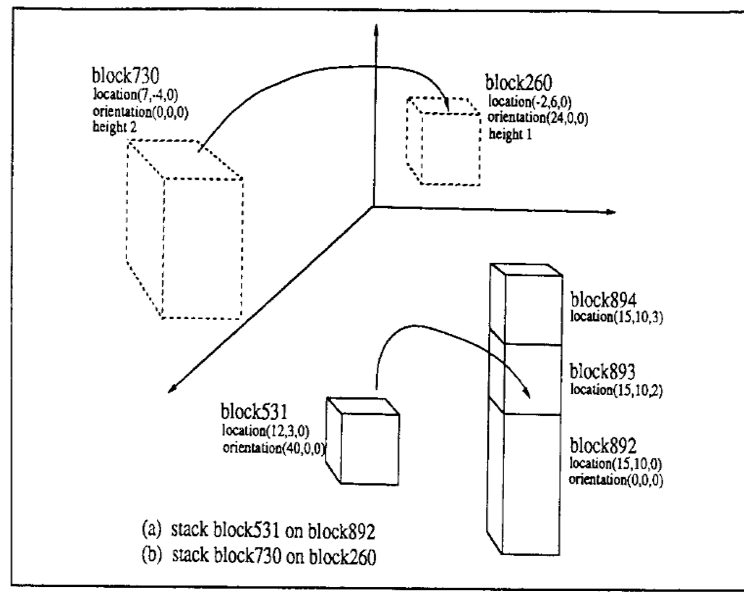


ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Short History

- Discrete State and Action System :
  - Block world + Predefined set of actions
- Learn sequencing across state-actions





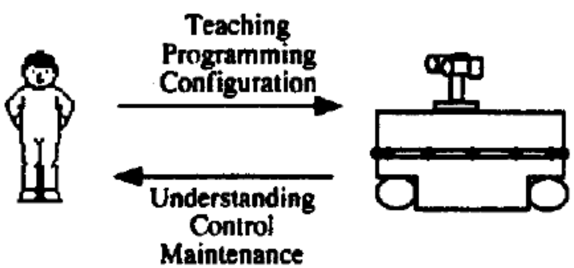
Block world

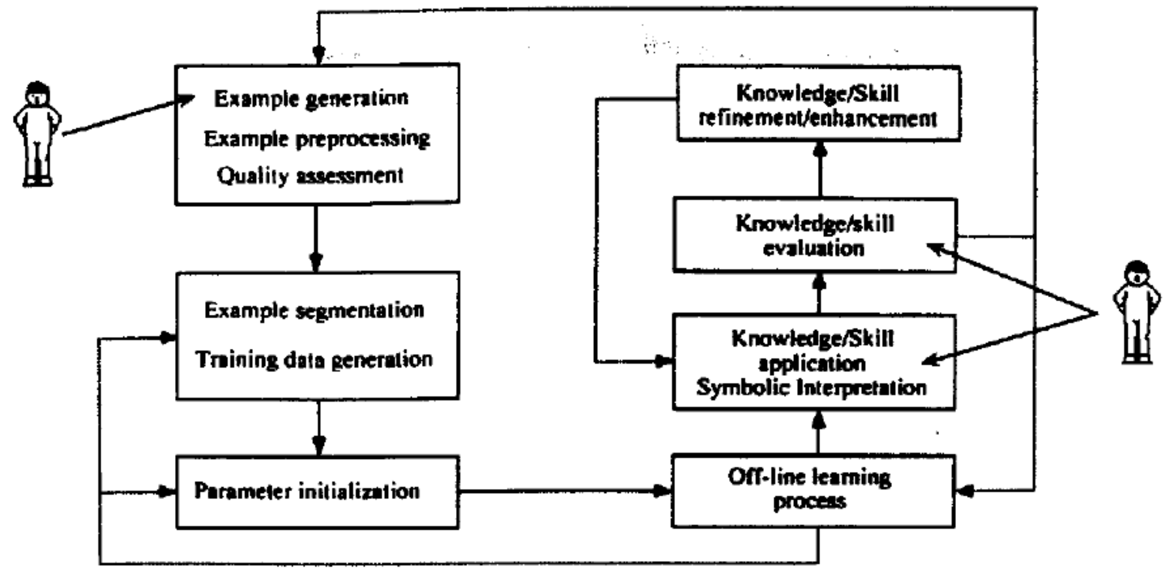
- Heise, 1989, Tech. Report, Univ. of Alberta
- Kuniyoshi et al. IEEE Trans. on Robotics and Automation,1994.

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Short History

First work to question how to correct bad teaching:
→ detects wrong / unnecessary / inefficient manipulations
→ do not burden the human to redo a demonstration when imperfect





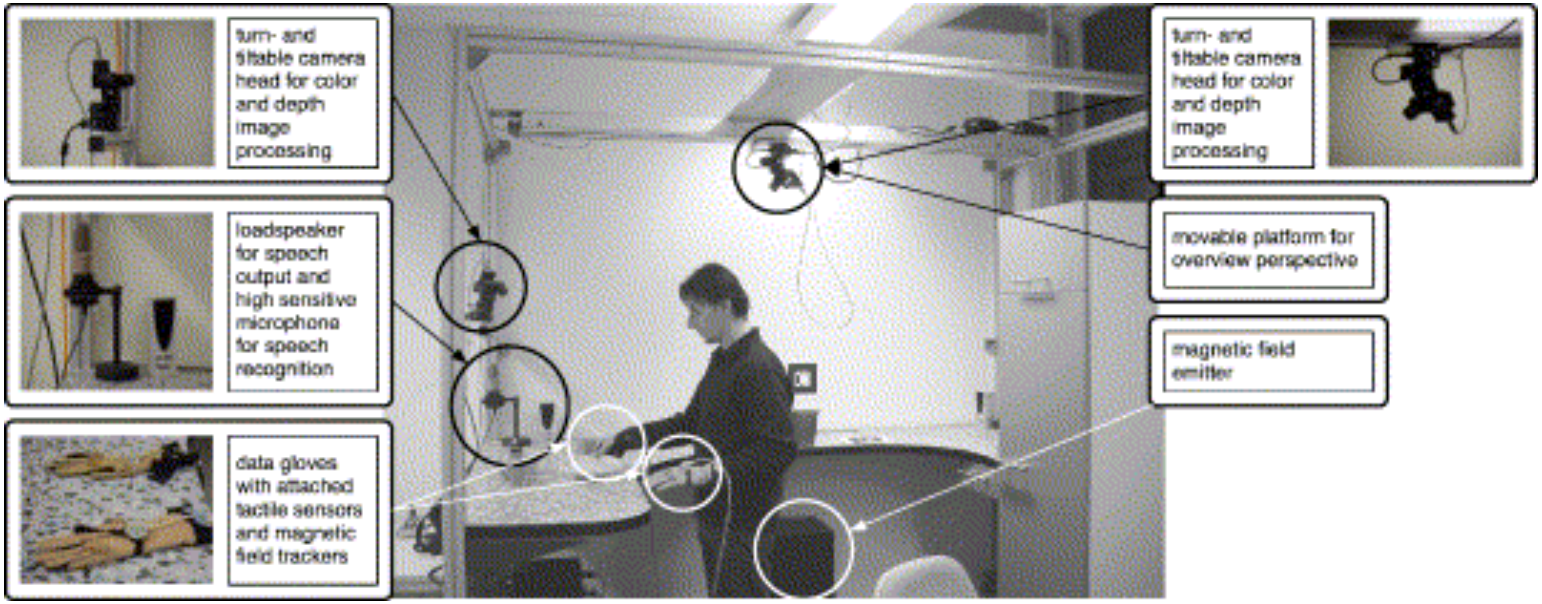
• Kaiser, Holger and Dillmann. "Obtaining good performance from a bad teacher." *Programming by Demonstration vs. Learning from Examples Workshop at ML*. Vol. 95. 1995.
• Holger, Kaiser, and Dillmann. "What can robots learn from humans?." *Annual Reviews in Control* 20 (1996): 167-172.

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Short History

One-Short Learning

- Segmentation of demonstration into primitives
- Classification of gestures into predefined states (e.g. grasp, collision)
- Built-in controller for producing sequences of states

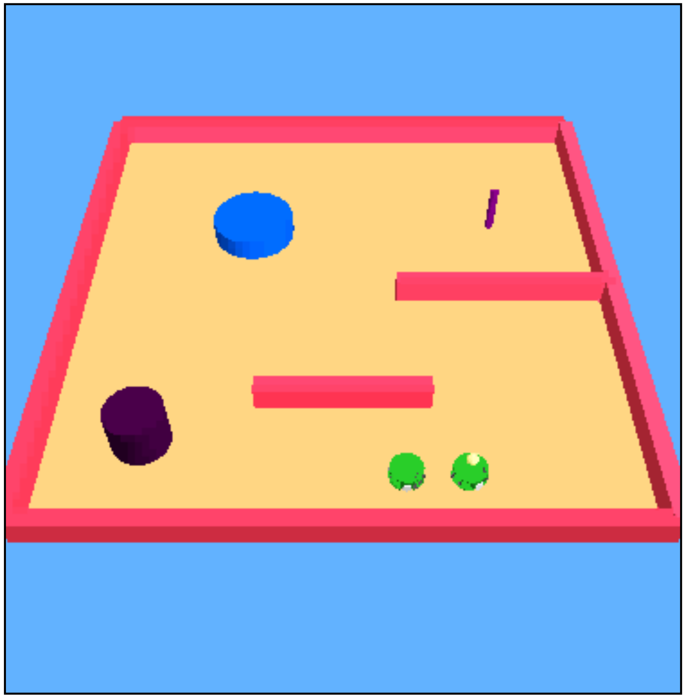


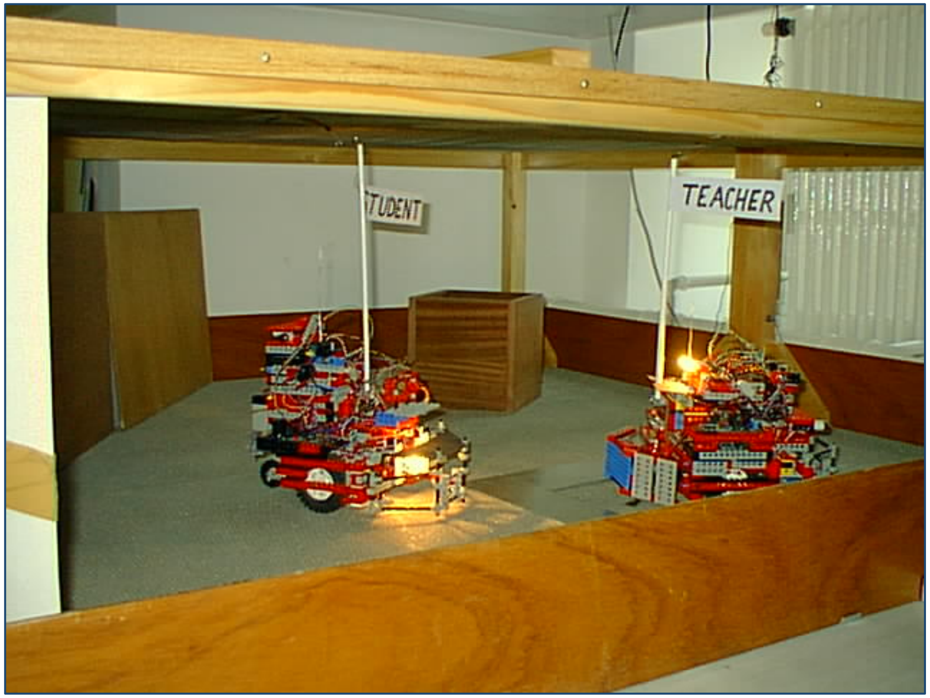
- Dillmann et al, Robotics & Autonomous Systems, 2001.
- Steil et al., Robotics and Autonomous Systems, 2004
- Aleotti et al, Robotics & Autonomous Systems, 2004.

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Short History

Following: simple imitation mechanism / while following the teacher, the learner robot learns to associate a word with a meaning in terms of sensory inputs.





Demiris & Hayes, 1994, 1996;
Dautenhahn, *Robotics & Autonomous Systems* 1995
Billard & Dautenhahn, *Robotics & Autonomous Systems*, 1998
Moga, Gaussier, Applied Artificial Intelligence, 2000
Kaiser et al, *Robotics & Autonomous Systems, 2002*

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Overview of current research areas

Low-Level Skills
- Trajectories
- Force profiles

High-Level Skills
- Combination of actions
- Speech-directed teaching

Batch learning versus incremental learning

Combined with other techniques
- Bootstrap reinforcement learning
- Inverse optimal control

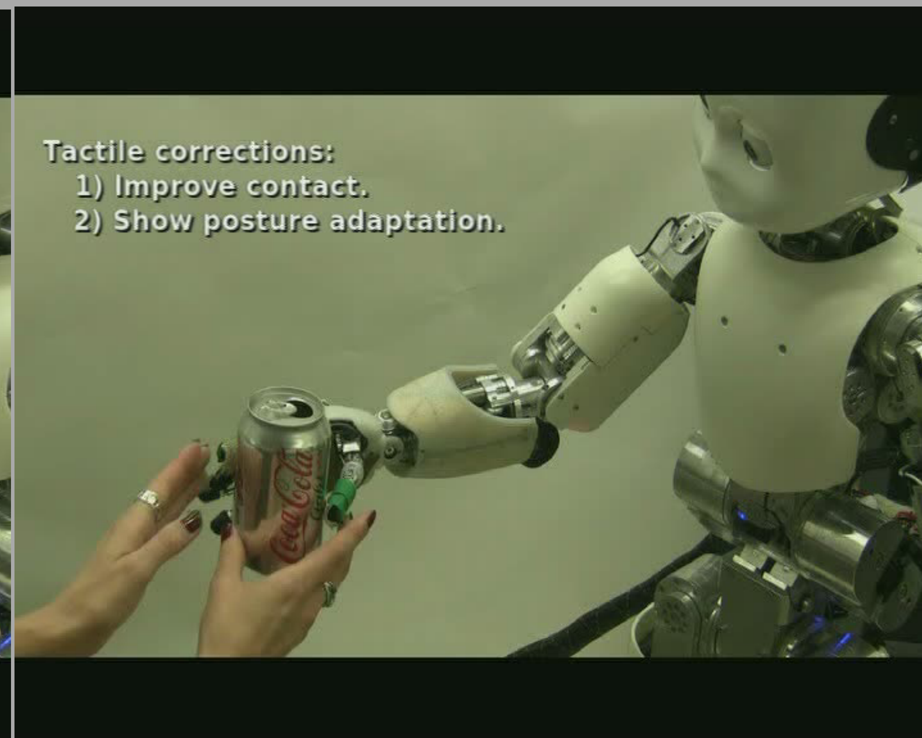
User-studies to assess:
- Interfaces
- Effectiveness of algorithm

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

http://www.scholarpedia.org/article/Robot_learning_by_demonstration

Kinesthetic Teaching using Tactile Sensing

No Adaptation



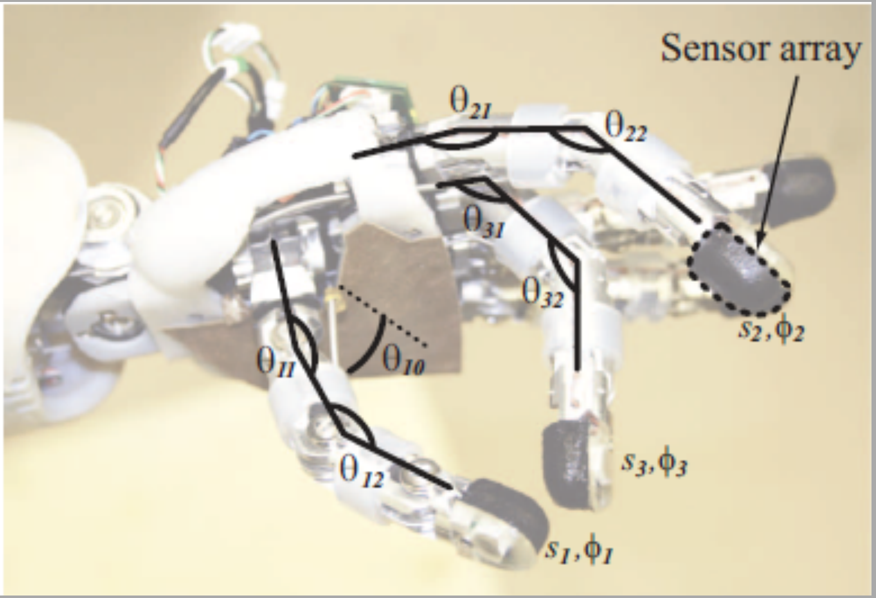
Teaching adaptive behavior

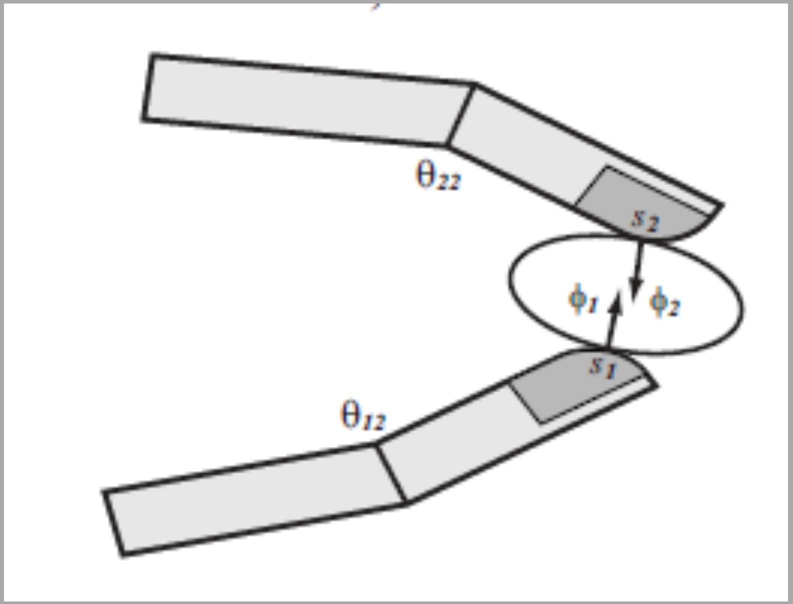
ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Sauser, Argall, Metta and Billard, Autonomous Robots, 2011

Kinesthetic Teaching using Tactile Sensing

Learn a probabilistic mapping $p(\phi, s, \theta)$ between contact signature of the object (normal force ϕ and tactile response s) and fingers' posture θ.

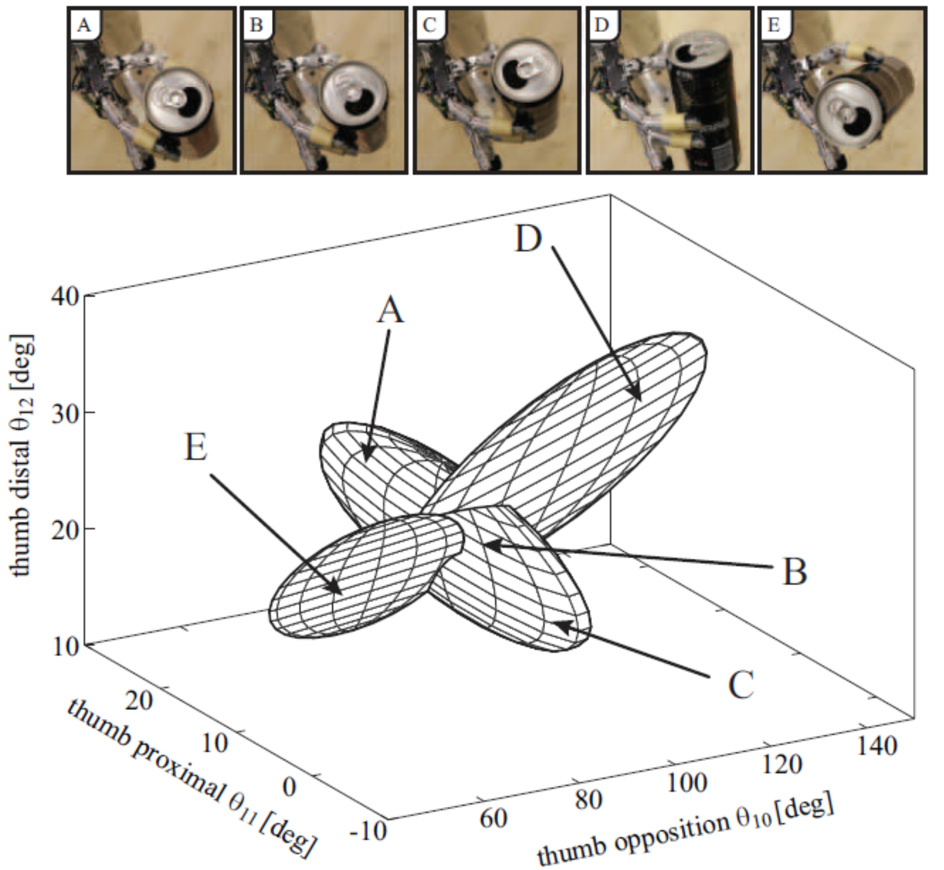




Sauser, Argall, Metta and Billard, Autonomous Robots, 2011

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Kinesthetic Teaching using Tactile Sensing



$$\hat{\theta} = E\left\{p\left(\theta \mid s,\phi\right)\right\}, \quad \hat{s} = E\left\{p\left(s \mid \theta,\phi\right)\right\}$$

Sauser, Argall, Metta and Billard, Autonomous Robots, 2011

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

LEARNING ALGORITHMS AND SYSTEMS LABORATORY

Teaching Compliant Control

Too stiff – Spills the liquid

Too compliant – spill the liquid outside the glass

Kronander and Billard, IEEE Trans. on Haptics, 2013

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Teaching Compliant Control

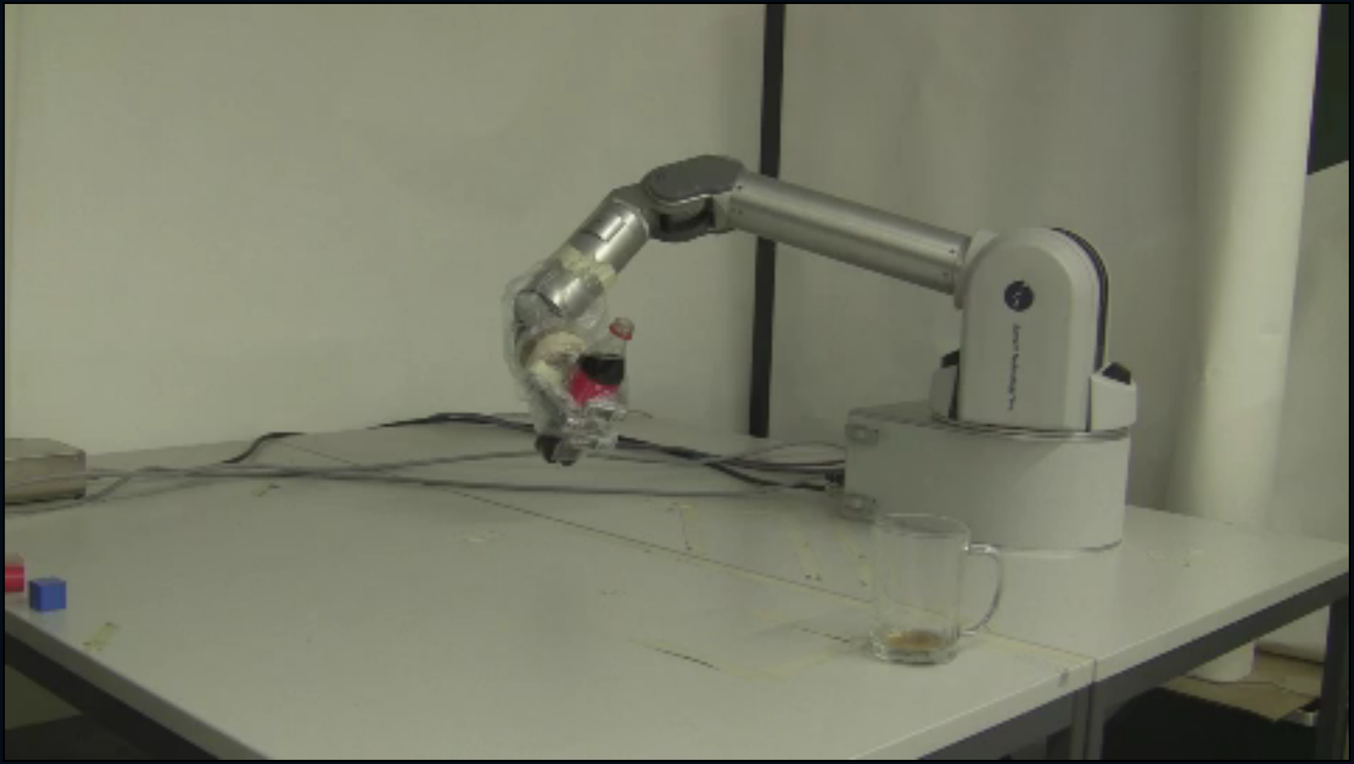
Decrease stiffness

Increase stiffness

Kronander and Billard, IEEE Trans. on Haptics, 2013

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Teaching Compliant Control



Kronander and Billard, IEEE Trans. on Haptics, 2013

Swiss National Centre of Competence in Research

Combining LfD with Optimal Control

- One cannot use directly demonstration as the dynamics of robot differ from human dynamics
- Use human demonstrations to learn the reward of a LQR controller.
- Use optimal control (LQR) to generate trajectories close to demonstrations.

Atkeson & Schaal, Intern. Conf. on Machine Learning, 1997;

Combining LfD with Reinforcement Learning

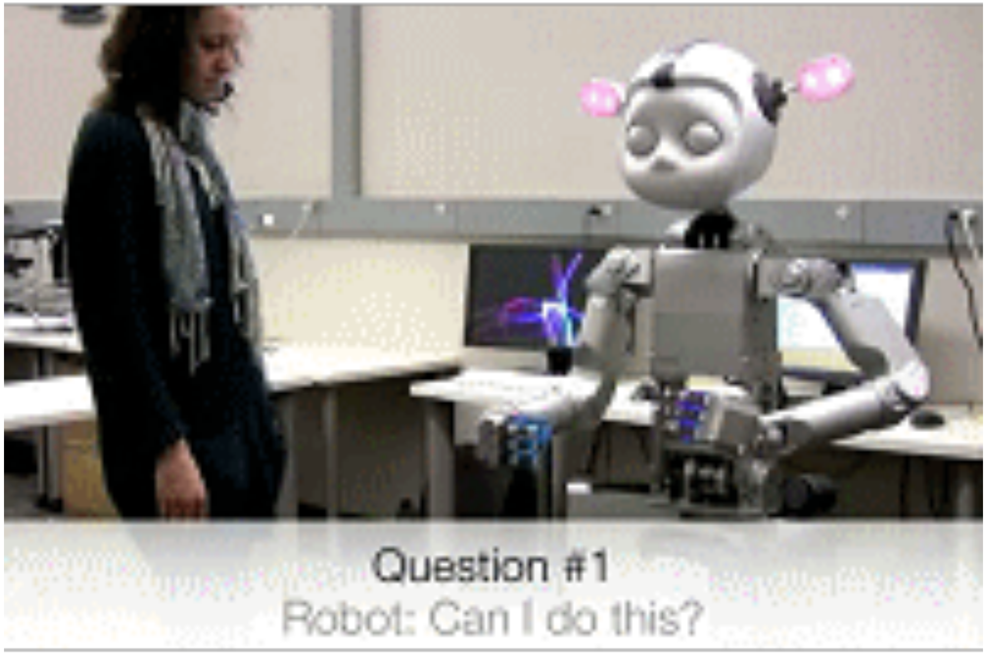
- Use human demonstrations to initialize the parameters of the controller.
- One cannot use directly demonstration as the dynamics of robot differ from human dynamics
- User reinforcement learning to search for solutions nearby the demonstrations.

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Kormushev et al, Int. Conf. on Robotics and Intelligent Systems,2010

Refinement through verbal interaction

- Robot has initial set of reaching skills
- Robot provided with a dialogue system to query the teacher
- Teacher modifies the controller through verbal guidance



Cakmak & Thomaz, Intern. Conf. on Human-Robot Interaction, 2012

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Companies selling LfD/PbD products

Yumi robot by ABB

Baxter Robot by Rethink Robotics

LWR by KUKA

UR3, Universal Robotics

Welding Robots, Robotik

Carbon Robotics

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE