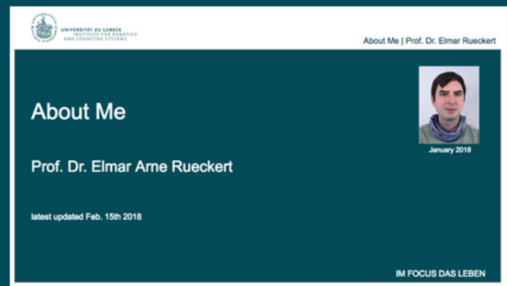
Deep Learning for Motor Control

Lübeck Summer Academy (LSA 2018)



Lübeck, July 4th, 2018
Prof. Dr. Elmar Rueckert

latest updated July 2nd 2018

Introduction & Motivation

Humanoid robots are among the most complex machines on earth.

And you will learn here how to build, teach and program them.

A brief historical review

Link to a more detailed history review

1920 **Karek Capek**: “robot” in his play “R.U.R.” (Rossum’s Universal Robots).

1941 **Isaac Asimov**: Three laws of “robotics”:

1. A robot may not injure a human being or, through inaction, allow a human being to come to harm.
2. A robot must obey orders given it by human beings except where such orders would conflict with the First Law.
3. A robot must protect its own existence as long as such protection does not conflict with the First or Second Law.

A brief historical review

1968 “**Shakey**” of the “Stanford Research Institute” defines a landmark in robotics:

- basic planning and navigation skills.
- object detection and manipulation capabilities.

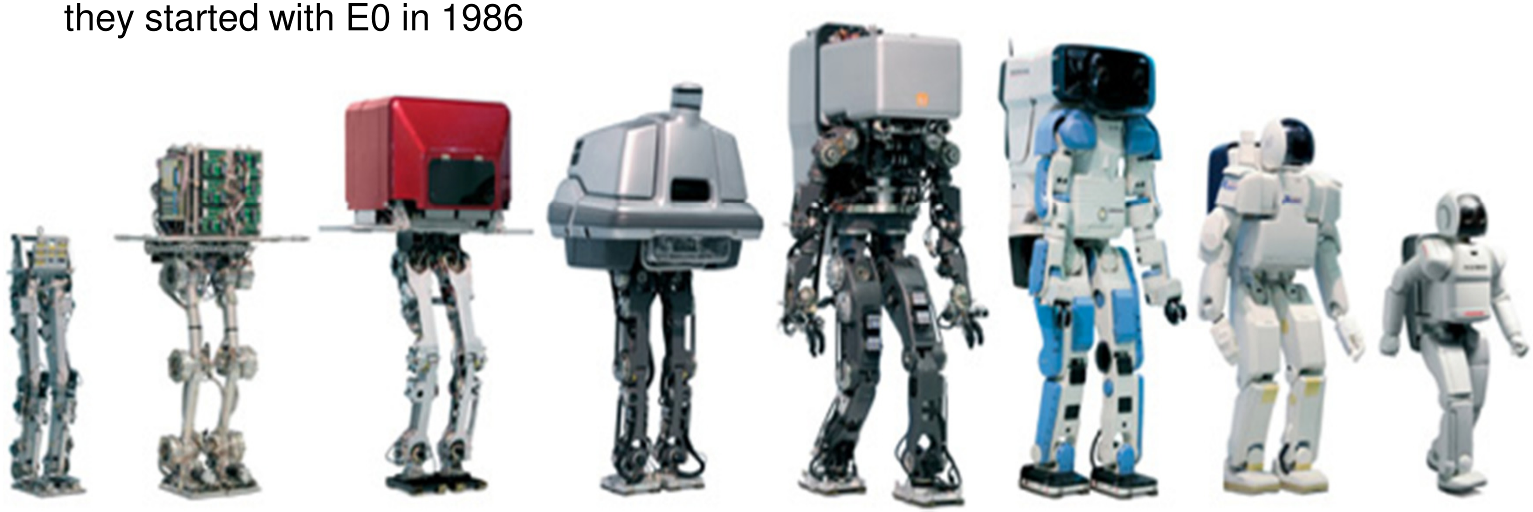
A brief historical review

1973 **Ichiro Kato** develops the first “full-scale” antrophomorphic humanoid, WABOT I.

A brief historical review

1996 **Honda** presents its P2

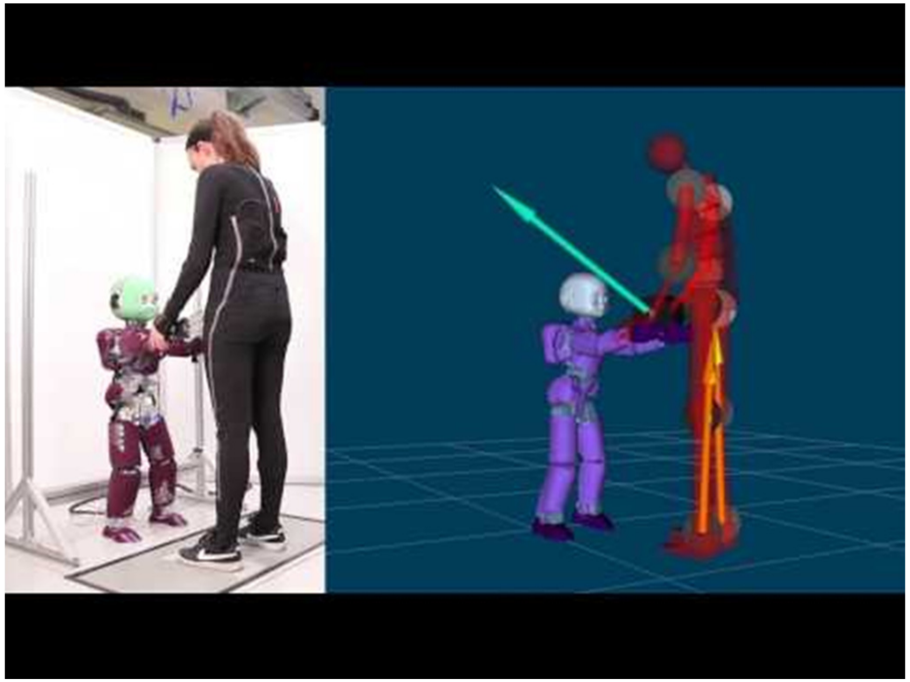
they started with E0 in 1986



the history of Hunda's humanoids

A brief historical review

2004 The Italian Institute of Technologie presents the **ICub** (intelligent man-cub).



A brief historical review

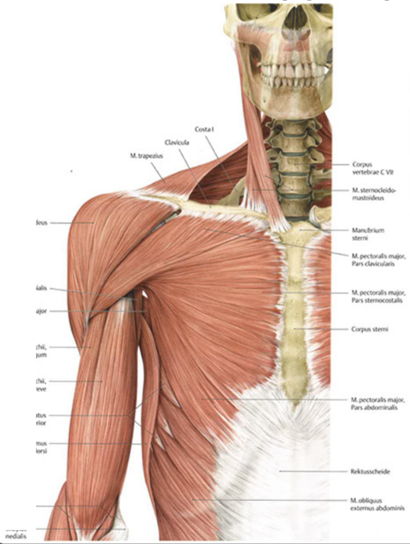
2017 Boston dynamics' **Atlas** impresses the robotics community.

Many Challenges in motor skill learning

More than robotics ...

The challenges in understanding humans and in building intelligent humanoids are converging!

~ 700 muscles
~ 100 joints
~ 100 x 10^6 photo receptors
~ 10^2 FA-I receptors per fingertip

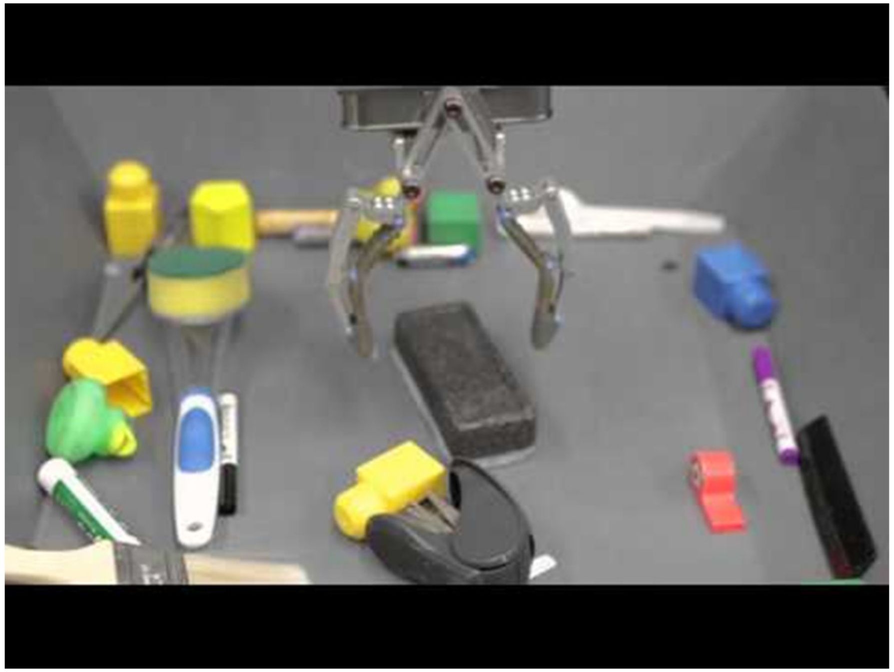


- 53 degrees of freedom
- 4 force/torque sensors
- 1.8 x 10^6 photo receptors
~ 2000 tactile sensors

In **humans** we suffer from noise, accuracy, delays.

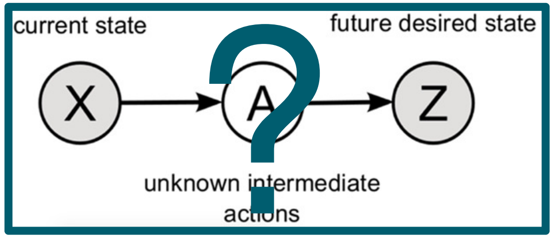
Despite **robot** vision is richer and more precise, robot motion is faster and more accurate their motor skills are inferior, **why?**

Is Deep Learning the answer?



How do we humans learn and plan?

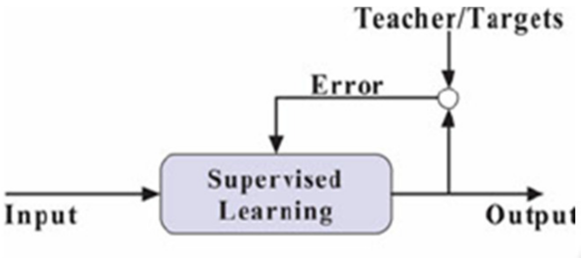
We exploit:

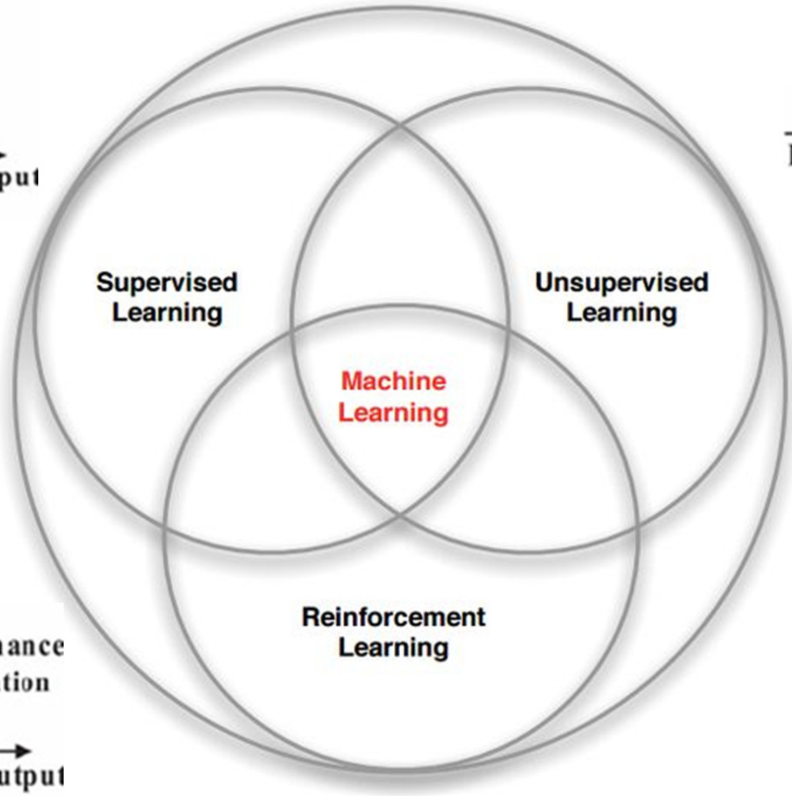


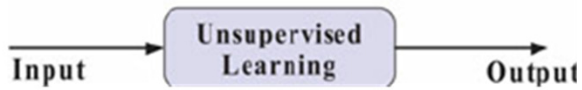
- Redundancy
- Variability
- Flexibility
- Structure

to generalize knowledge

Types of Learning







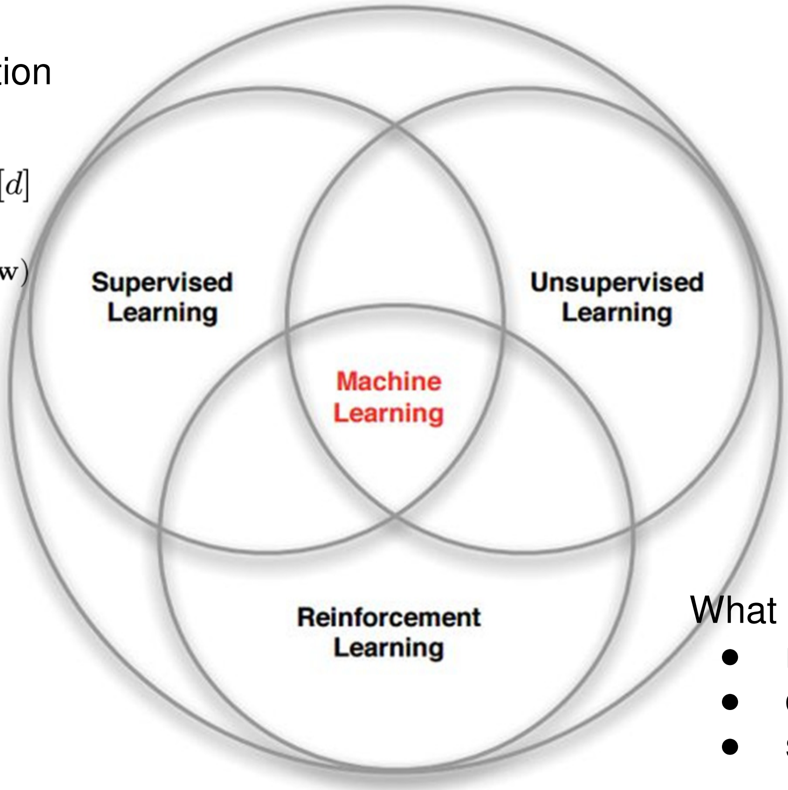
Shouyi Wang, Wanpracha Chaovalitwongse, Robert Babuska, Machine Learning Algorithms in Bipedal Robot Control, DOI: 10.1109/TSMCC.2012.2186565.

Types of Learning

- Classification
- Linear Regression / Function Approximation, i.e.,

$$\mathbf{w}^{[d]} = (\boldsymbol{\Psi}^T\boldsymbol{\Psi} + \lambda\mathbf{I})^{-1}\boldsymbol{\Psi}^T\tilde{\mathbf{f}}^{[d]}$$

$$J = \frac{1}{2}(\tilde{\mathbf{f}} - \mathbf{f})^T(\tilde{\mathbf{f}} - \mathbf{f}) = \frac{1}{2}(\tilde{\mathbf{f}} - \boldsymbol{\Psi}\mathbf{w})^T(\tilde{\mathbf{f}} - \boldsymbol{\Psi}\mathbf{w})$$

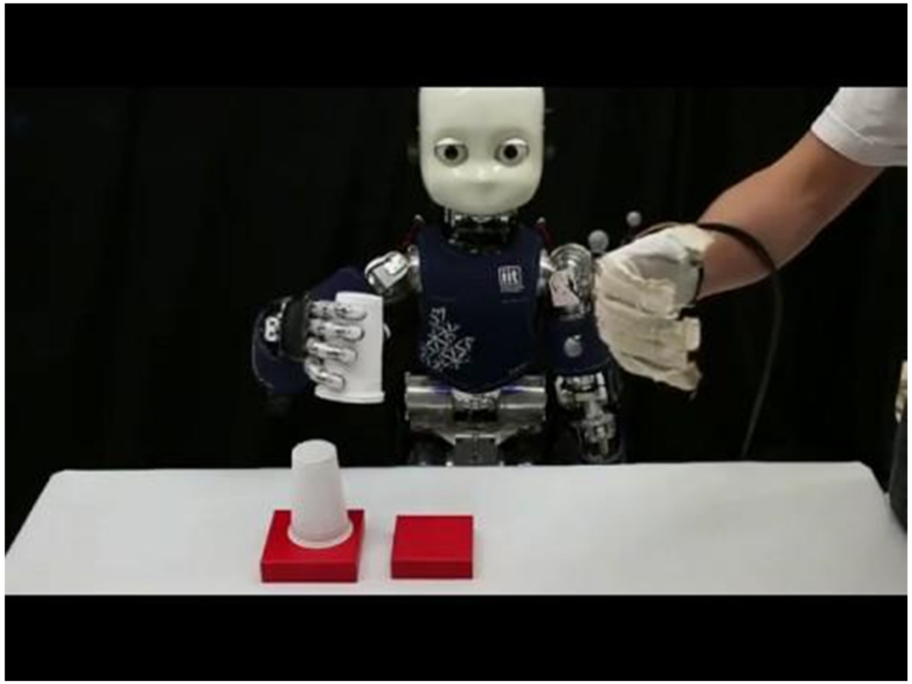


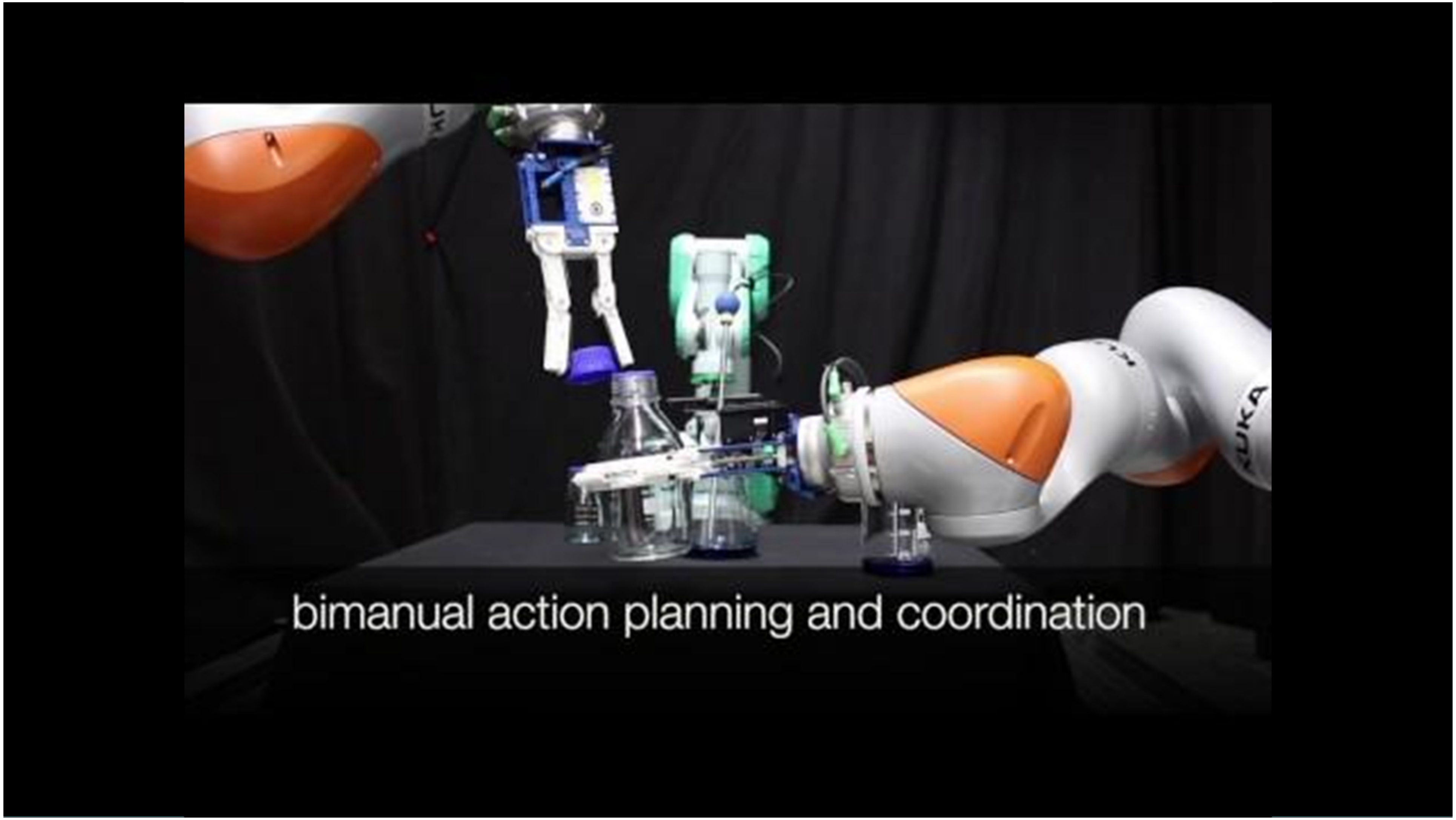
- Clustering, e.g., k-means
- Feature learning, e.g., autoencoder
- Dimensionality reduction, e.g., PCA

What is different in RL:

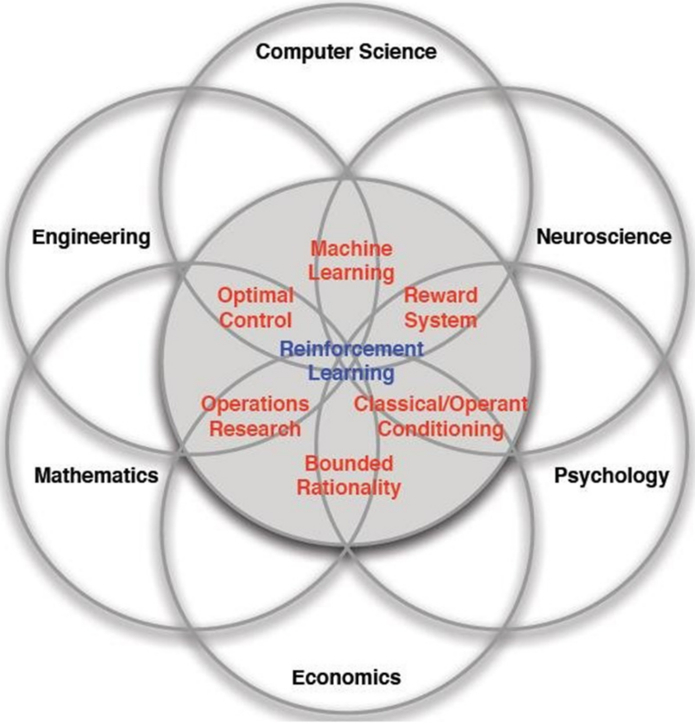
- no supervisor, just trial and error
- delayed feedback
- sequential non i.i.d data

Supervised Learning from demonstrations




KUKA
bimanual action planning and coordination

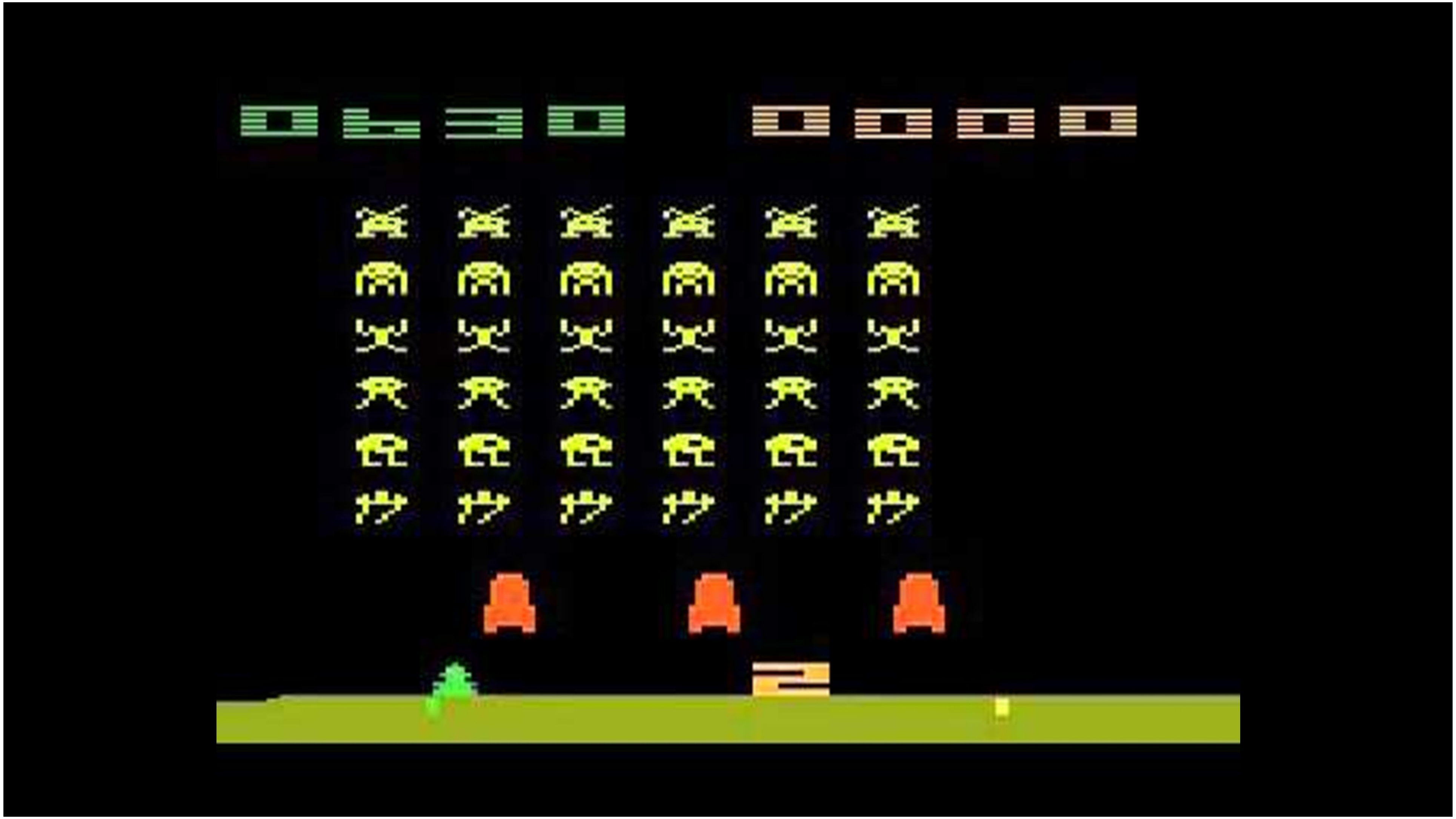
Reinforcement Learning



by David Silver, have a look at his great video lecture.

Some Examples of Reinforcement Learning


0470
4 215
2X
ACTIVISION


0630
0000
2


DEEPMIND AI
LEARNED HOW TO WALK

Related Reading on Reinforcement Learning

Books:

- Sutton & Barto 1998. **An Introduction to Reinforcement Learning**, *MIT Press.* free online version
- Szepesvari 2010. **Algorithms for Reinforcement Learning**, Morgan and Claypool. free online version

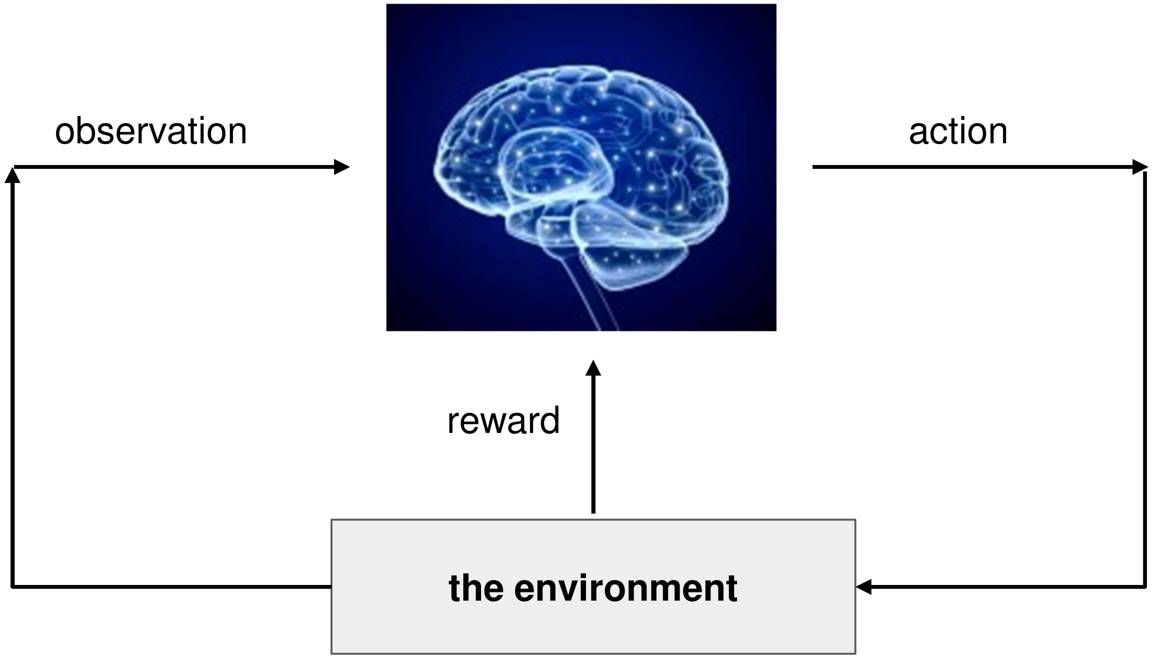
Video Lectures:

- videolectures.net *on Reinforcement Learning*
- coursea.org *on Robotics*

Lecture notes:

- Humanoid Robotics *by Prof. Dr. Elmar Rueckert, University of Luebeck.*
- Lecture notes on learning methods *by Prof. Dr. Marc Toussaint, University Stuttgart.*
- Lecture notes on dynamics *by Prof. Dr. Russ Tedrake, Massachusetts Institute of Technology.*

The Sequential Decision Making Framework

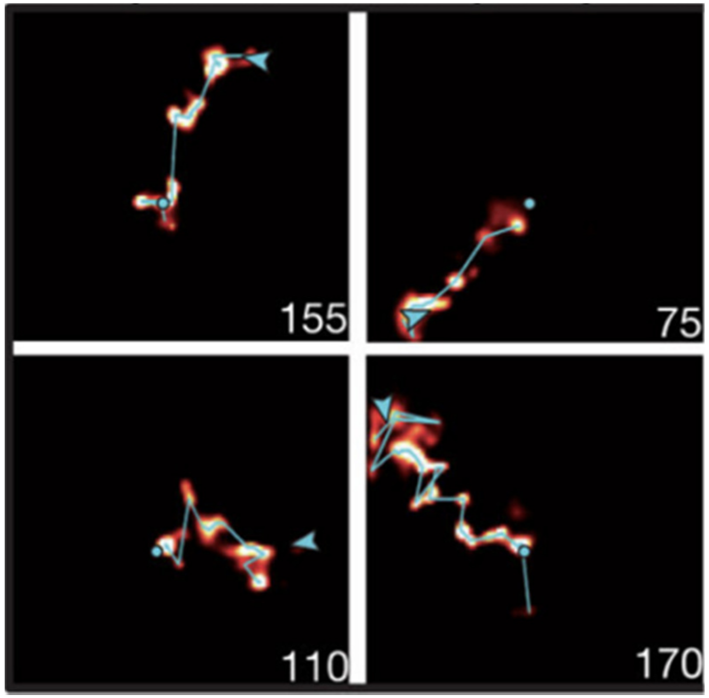


The execution of action $\mathbf{a}_t$ leads to

- a new observation $\mathbf{o}_{t+1}$
- a potential reward $\mathbf{r}_{t+1}$

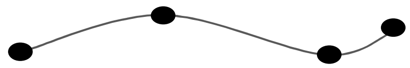
Evidence

Decision making & planning in few ms



The stochastic process for planning

$p(\underline{\boldsymbol{x}} \mid \quad) =$

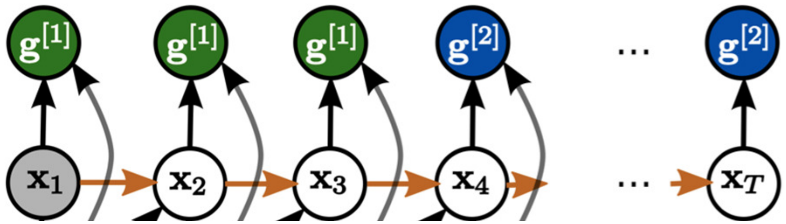


The stochastic process for planning

$$p(\underline{\boldsymbol{x}}) = p(\boldsymbol{x}_0) \prod_{t=1}^{T} \mathscr{T}(\boldsymbol{x}_t | \boldsymbol{x}_{t-1})$$

The stochastic process for planning

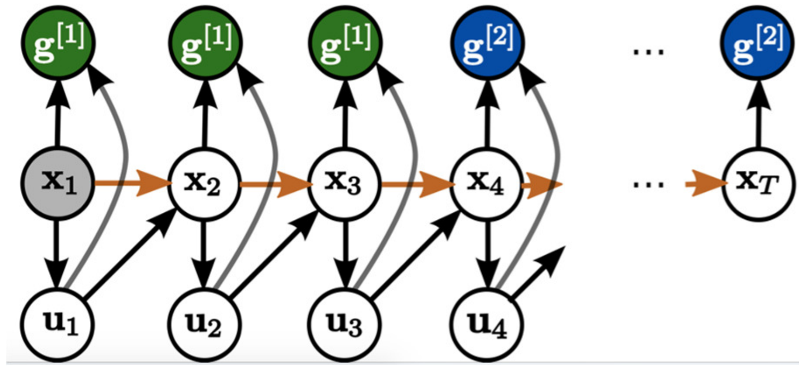
$$p\left(\underline{\boldsymbol{x}}|r=1\right)=\frac{1}{\mathcal{Z}}p\left(r|\underline{\boldsymbol{x}}\right)p\left(\boldsymbol{x}_0\right)\prod_{t=1}^{T}\mathcal{T}\left(\boldsymbol{x}_t|\boldsymbol{x}_{t-1}\right)$$



The stochastic process for planning

$$p(\underline{\boldsymbol{x}}|r=1)=\frac{1}{\mathscr{Z}}p(r|\underline{\boldsymbol{x}})\,p(\boldsymbol{x}_0)\prod_{t=1}^{T}\mathscr{T}(\boldsymbol{x}_t|\boldsymbol{x}_{t-1})$$

Cannot be implemented in RNNs!



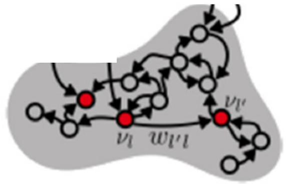
However a RNN can implement forward sampling from a learned distribution

$$p(\underline{\boldsymbol{x}}|r=1) = \frac{1}{\mathcal{Z}} p(r|\underline{\boldsymbol{x}})\, p(\boldsymbol{x}_0) \prod_{t=1}^{T} \mathscr{T}(\boldsymbol{x}_t|\boldsymbol{x}_{t-1})$$

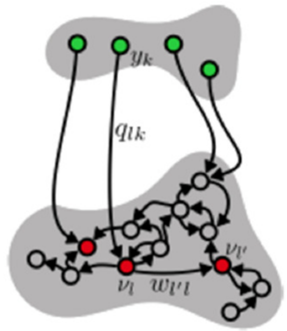
- Reward modulated Hebbian Learning
- Supervised Model Learning (CD)

From Random walks to planning

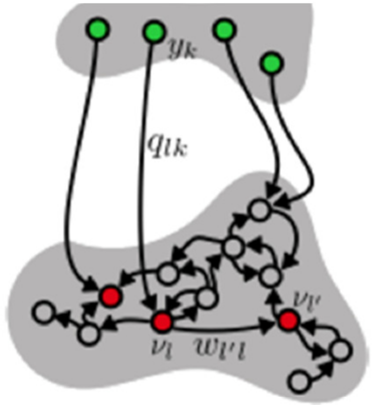
$$p(\underline{\boldsymbol{x}}) = p(\boldsymbol{x}_0) \prod_{t=1}^{T} \mathscr{T}(\boldsymbol{x}_t|\boldsymbol{x}_{t-1})$$



$$p(\underline{\boldsymbol{x}}|r = 1) = \frac{1}{\mathscr{Z}} p(r|\underline{\boldsymbol{x}}) p(\boldsymbol{x}_0) \prod_{t=1}^{T} \mathscr{T}(\boldsymbol{x}_t|\boldsymbol{x}_{t-1})$$



Models for inference and planning



Neural model distribution

$$q(\underline{\nu};\boldsymbol{\theta}) = p(\boldsymbol{\nu}_0)\prod_{t=1}^{T}\prod_{k=1}^{K}\rho_{t,k}{}^{\nu_{t,k}}\left(1-\rho_{t,k}\right)^{1-\nu_{t,k}}$$

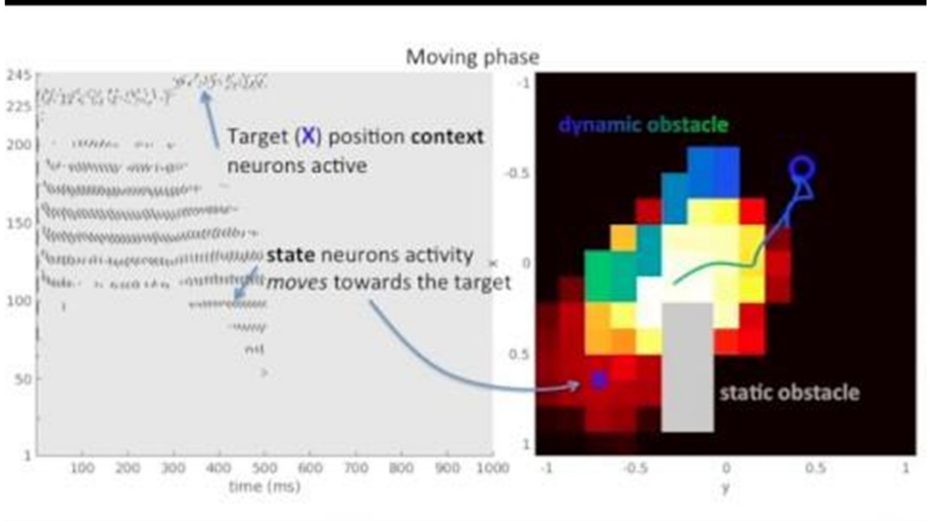
$$= p(\boldsymbol{\nu}_0)\prod_{t=1}^{T}\mathscr{T}(\boldsymbol{\nu}_t \mid \boldsymbol{\nu}_{t-1})\,\phi_t(\boldsymbol{\nu}_t;\boldsymbol{\theta})$$

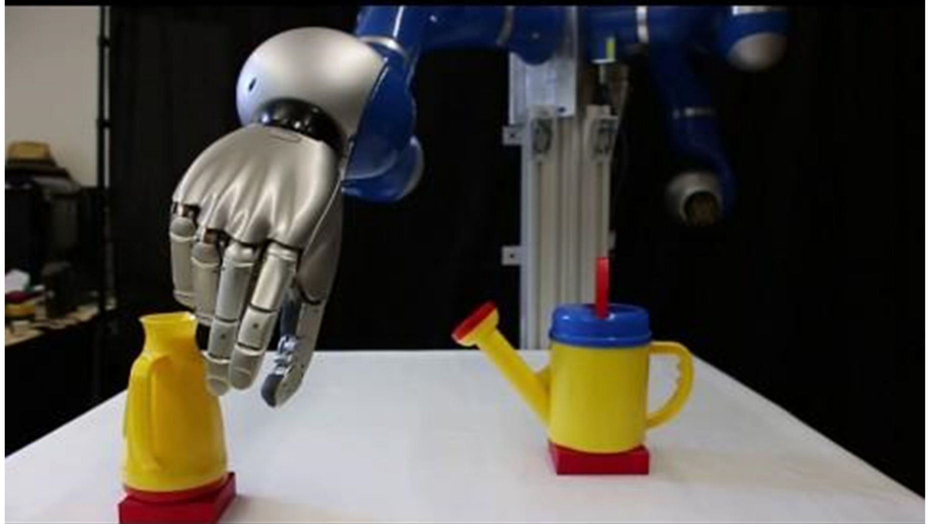
$$\mathscr{T}(\boldsymbol{\nu}_t \mid \boldsymbol{\nu}_{t-1}) = \exp\left(\sum_{i=1}^{K} w_{ki}\,\nu_{t-1,i}\,\nu_{t,k}\right)$$

$$\phi_t(\boldsymbol{\nu}_t;\boldsymbol{\theta}) = \frac{\exp\left(\sum_{j=1}^{N}\theta_{kj}\,y_{t-1,j}\,\nu_{t,k}\right)}{\sum_{l=1}^{K}\exp\left(u_{t,l}\right)}$$

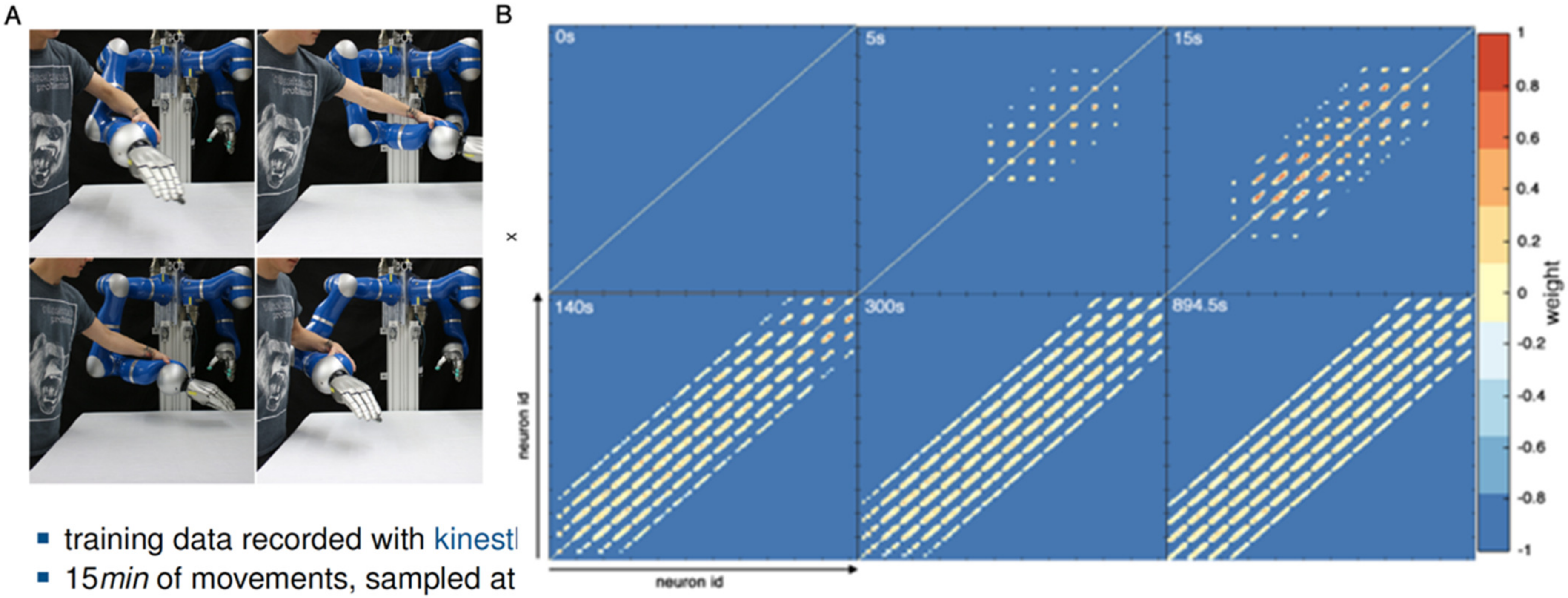
- Supervised Model Learning (CD)
- Reward modulated Hebbian Learning

Predictive models of navigation skills



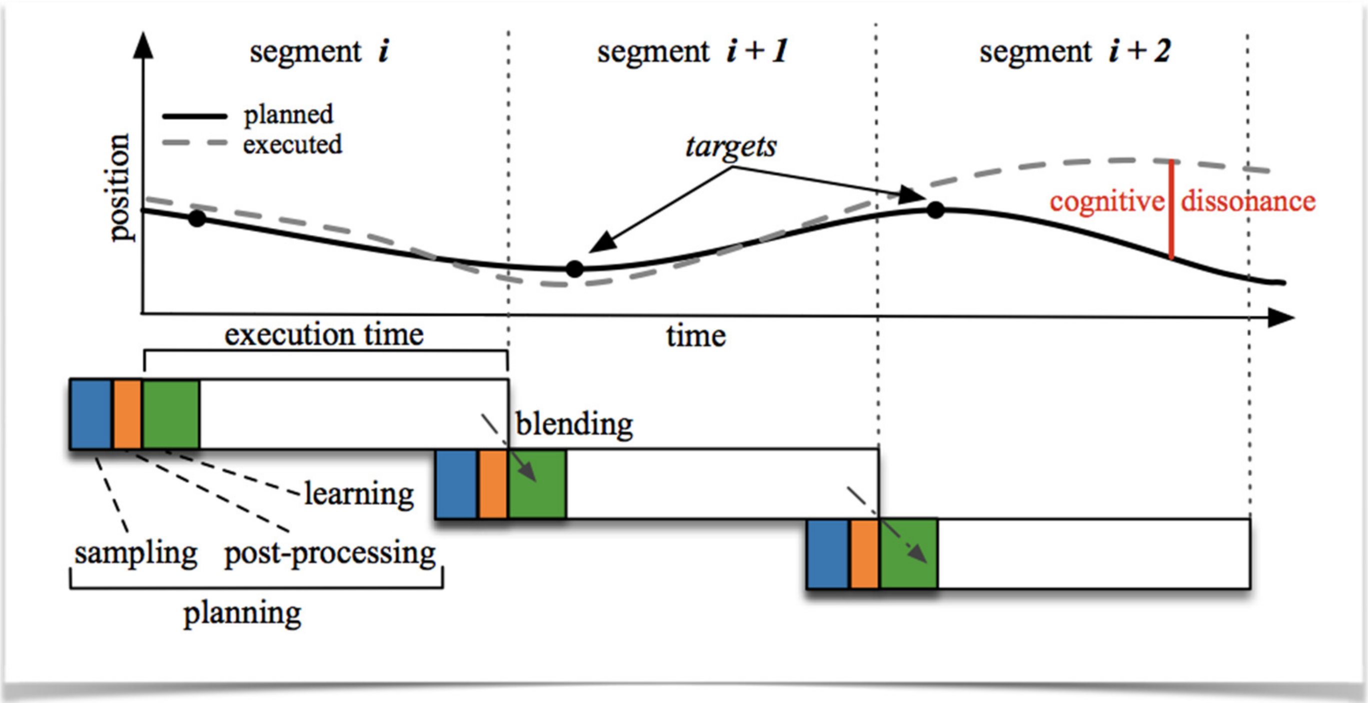


Model Learning in 15 Minutes



- training data recorded with kinestl
- 15*min* of movements, sampled at

Real Time Adaptation and Control



[redacted]

more at: https://rob.ai-lab.science/publications/

Paraschos, Alexandros; Rueckert, Elmar; Peters, Jan; Neumann, Gerhard
Probabilistic Movement Primitives under Unknown System Dynamics Journal Article
Advanced Robotics (ARJ), 2018.

Tanneberg, Daniel; Peters, Jan; Rueckert, Elmar
Online Learning with Stochastic Recurrent Neural Networks using Intrinsic Motivation Signals Inproceedings
Proceedings of the Conference on Robot Learning (CoRL), 2017.

Rueckert, Elmar; Kappel, David; Tanneberg, Daniel; Pecevski, Dejan; Peters, Jan
Recurrent Spiking Networks Solve Planning Tasks Journal Article
Nature Publishing Group: Scientific Reports, 6 (21142), 2016.

Rueckert, Elmar; Neumann, Gerhard; Toussaint, Marc; Maass, Wolfgang
Learned graphical models for probabilistic planning provide a new class of movement primitives Journal Article
Frontiers in Computational Neuroscience, 6 (97), 2013.

Summary: Deep Learning for Motor Control

- In robotics, we need to **compute torques** from **noisy** and **high dimensional** data
- Can capture and exploit **correlations** for **predictions**
- **Low** dimensional **feature** representation for **learning**
- Generate **Stroke-based** and **Rhythmic** Movements and Feedback
- **Transfer learning** through mixture models (extension)
- Can be combined with **deep convolutional neural networks**

Thank you for your attention!

Contact:

Universität zu Lübeck
Institute for Robotics and Cognitive Systems
Ratzeburger Allee 160
Building 64, Room 94
23538 Lübeck, Deutschland

Telefon: +49 (0) 451 3101 5209
E-Mail: rueckert@rob.uni-luebeck.de

https://rob.ai-lab.science

Disclaimer: