Probabilistic approaches in LfD

286 IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART B: CYBERNETICS, VOL. 37, NO. 2, APRIL 2007

On Learning, Representing, and Generalizing a Task in a Humanoid Robot

Sylvain Calinon, Florent Guenter, and Aude Billard, *Member, IEEE*

Sylvain Calinon

FOLLOW

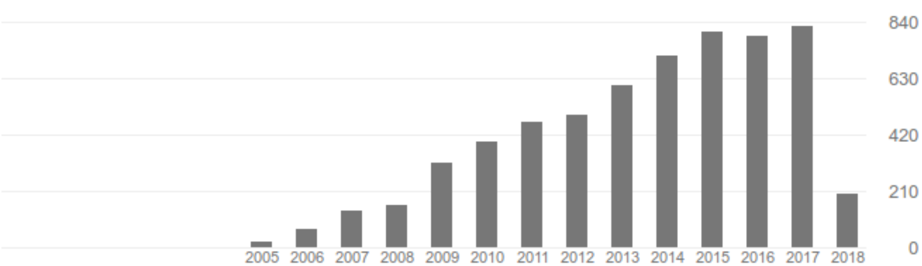
Idiap Research Institute

Verified email at idiap.ch - Homepage

robotics machine learning

ARTICLES CITED BY CO-AUTHORS

	All	Since 2013
Citations	6061	3951
h-index	35	33
i10-index	61	57



Aude Billard

FOLLOW

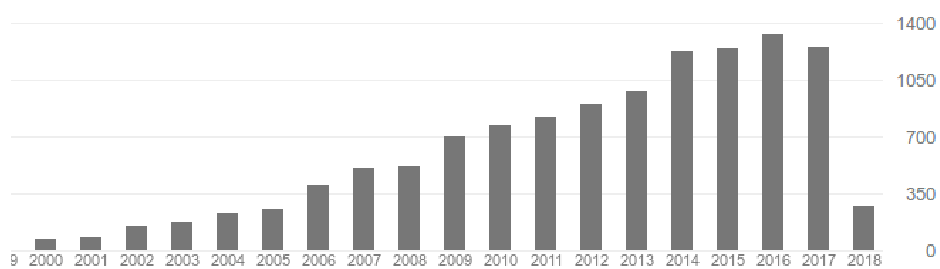
Full Professor, EPFL

Verified email at epfl.ch - Homepage

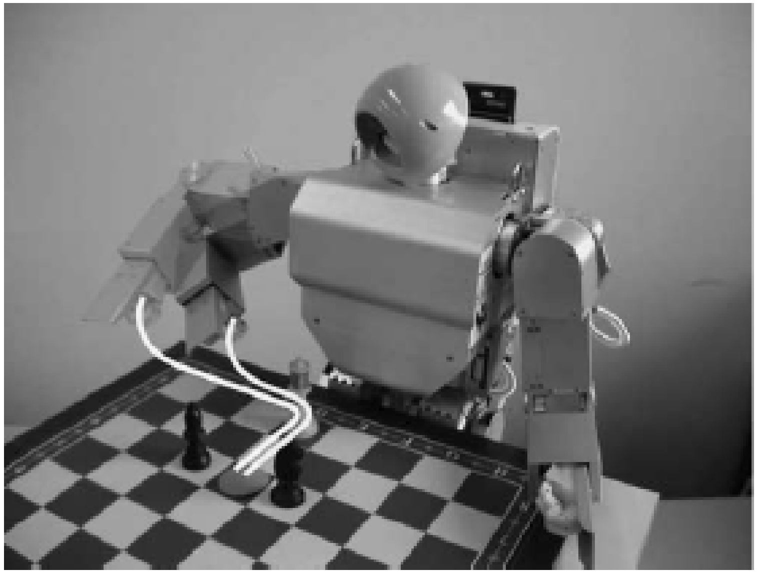
Robotics Machine Learning Robot Learning Human-Robot Interaction Programming by Demonstr…

ARTICLES CITED BY

	All	Since 2013
Citations	12172	6334
h-index	58	39
i10-index	148	115

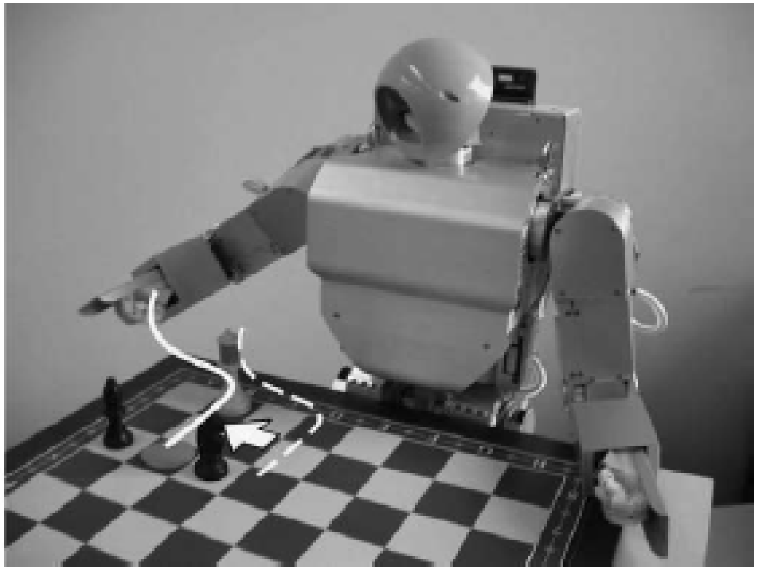


What to imitate?



- The spatio-temporal variations and correlations across the variables
- Weak correlations at the beginning of the motion
- strong spatio-temporal correlation for grabbing the piece and pushing it toward the desired location without hitting the other pieces on the chessboard

How to imitate?



- the demonstrated joint angles and hand path can be mutually exclusive in the imitator space, it is not possible to fulfill both constraints at the same time
- the trajectory which gives the optimal tradeoff between satisfying the constraints of the task (spatio-temporal correlations across the variables) and its own body constraints.

4N-1K

- Different demonstrations of the same task
- Probabilistical estimation of relevance to extract the important aspects of the task

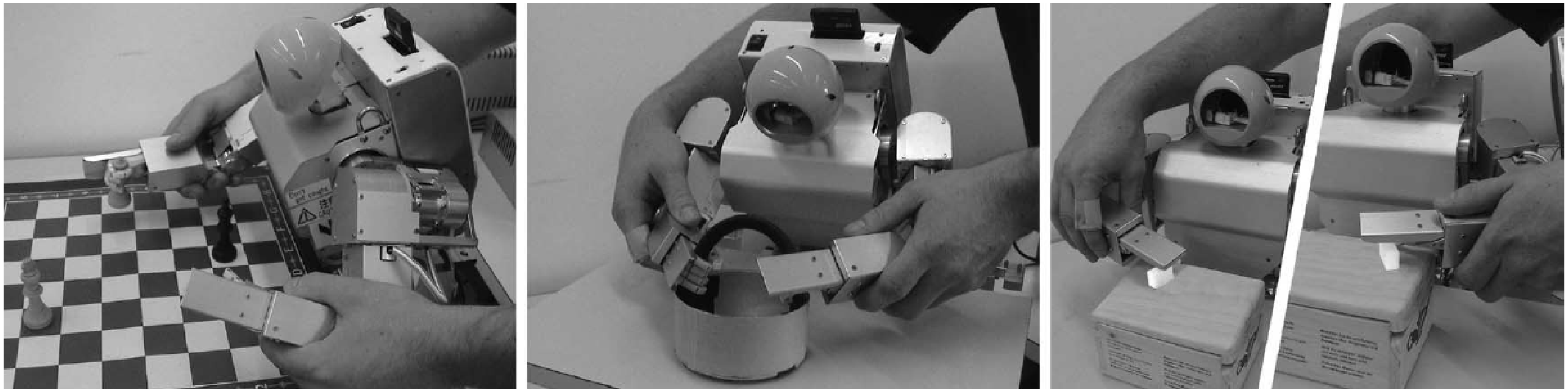
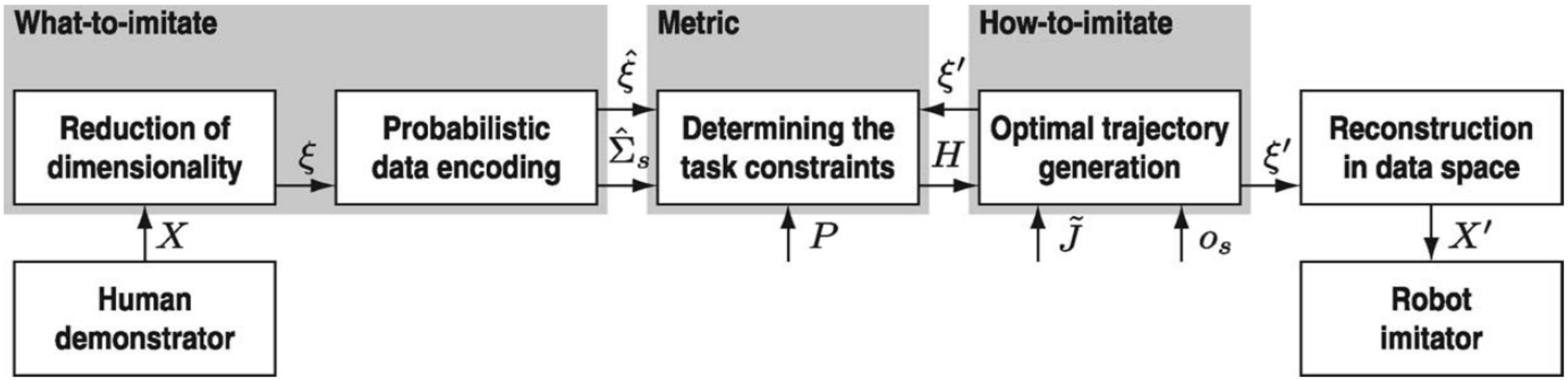


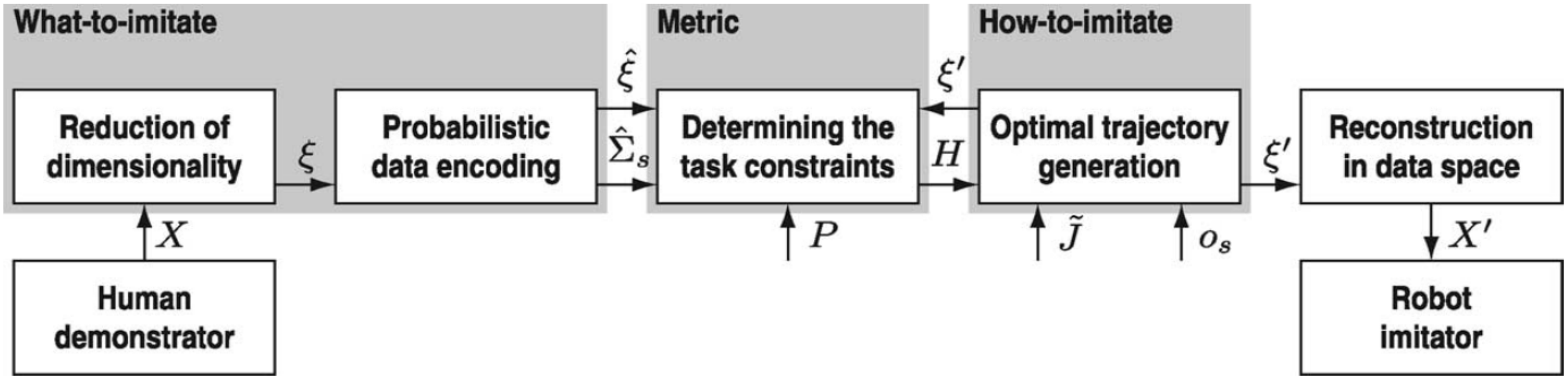
Fig. 3. Teaching through kinesthetics for the three experiments conducted. Chess task: Grabbing and moving a chess piece two squares forward. Bucket task: Grabbing and bringing a bucket to a specific position. Sugar task: Grabbing a piece of sugar and bringing it to the mouth, using either the right or left hand.

Architecture

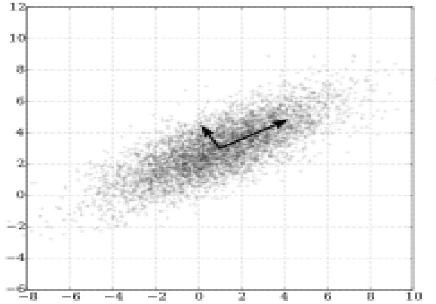


- The signals are encoded in three stage process
  - Determine the latent space of the motion
  - Temporarily align the signals
  - Determine probabilistic representation of data in latent space

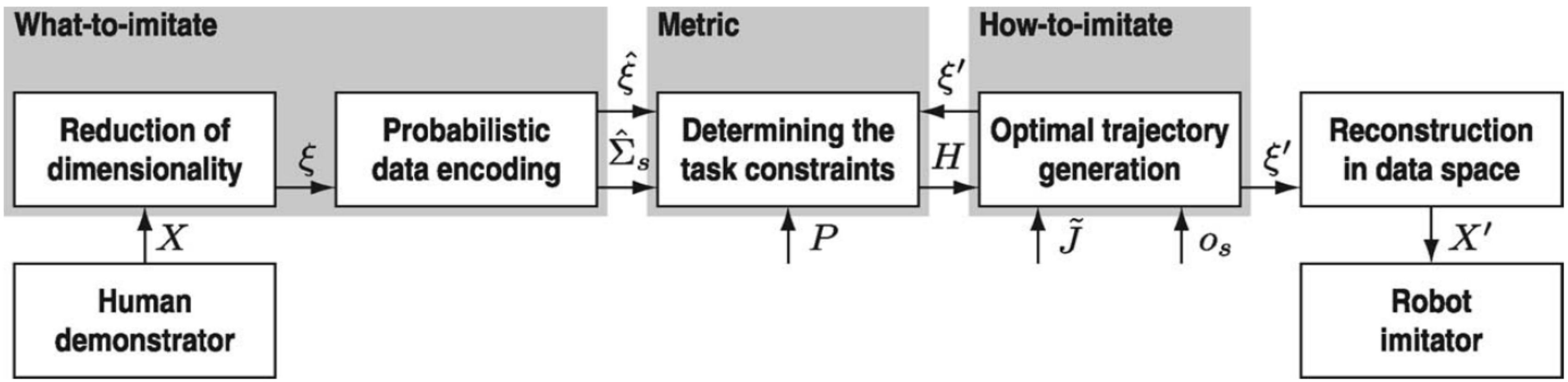
Latent Space

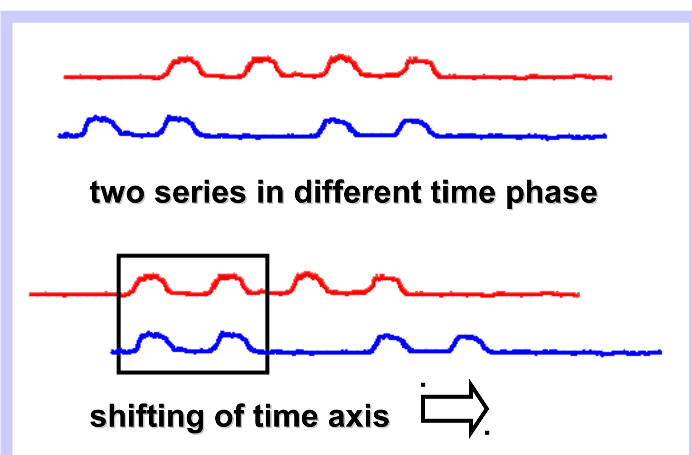


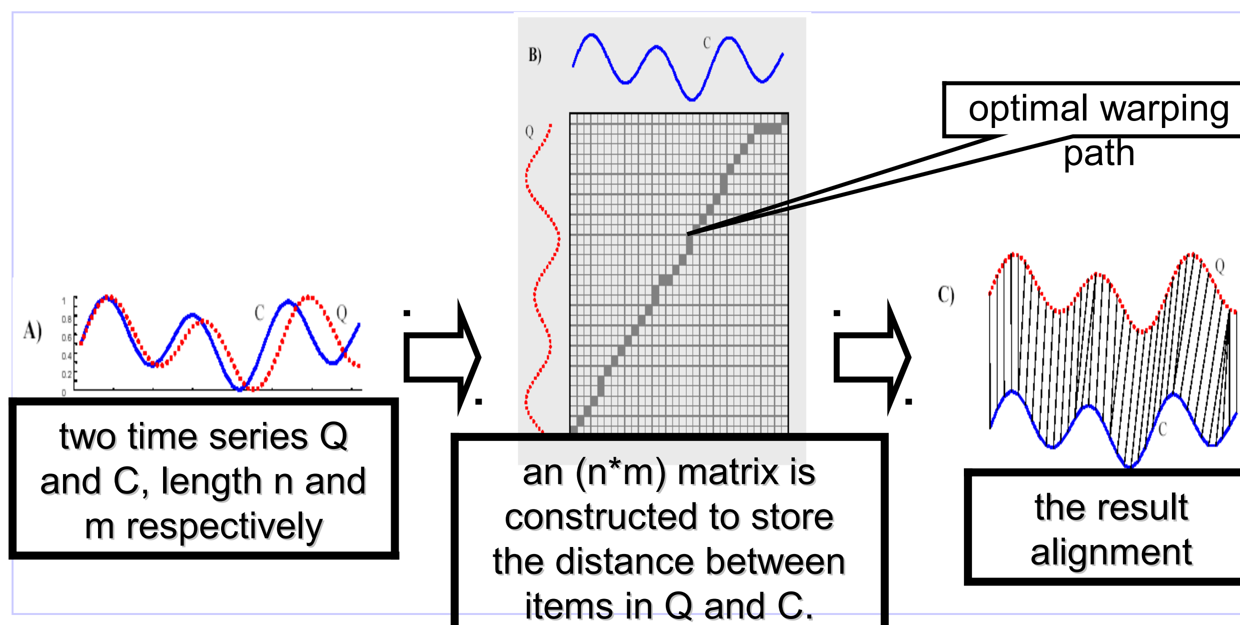
- Project the original dataset to find an optimal representation for the given task.
  - An optimal latent space for a writing task is typically represented as a projectionof the 3-D original Cartesian position of the hand onto a 2-D latent space defined by ?
  - a waving motion is typically represented as a combination of ?
- Principle Component Analysis (PCA) is used



Temporally align the signals

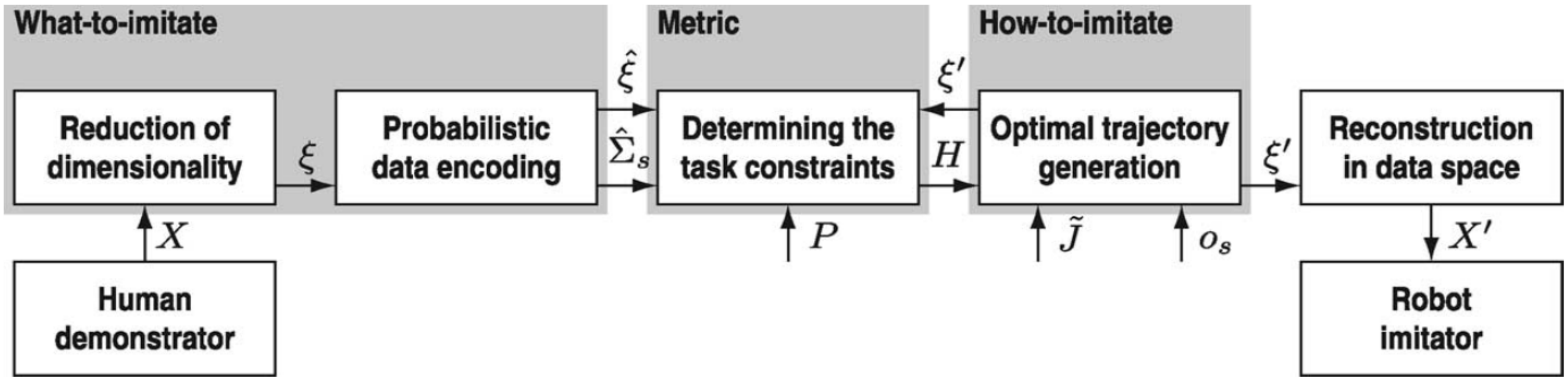






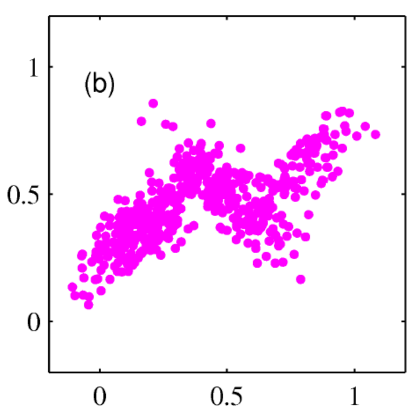
Dr Eamonn Keogh's slides

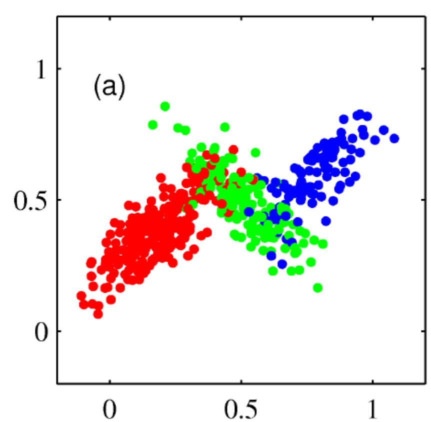
Probabilistic Representation



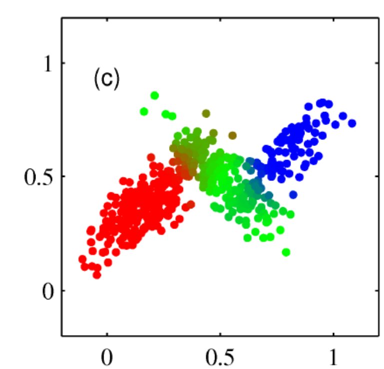
- Gaussian Mixture Model for continuous data
- Bernoulli Mixture Model for discrete data

Model data as mixture of multivariate Gaussians



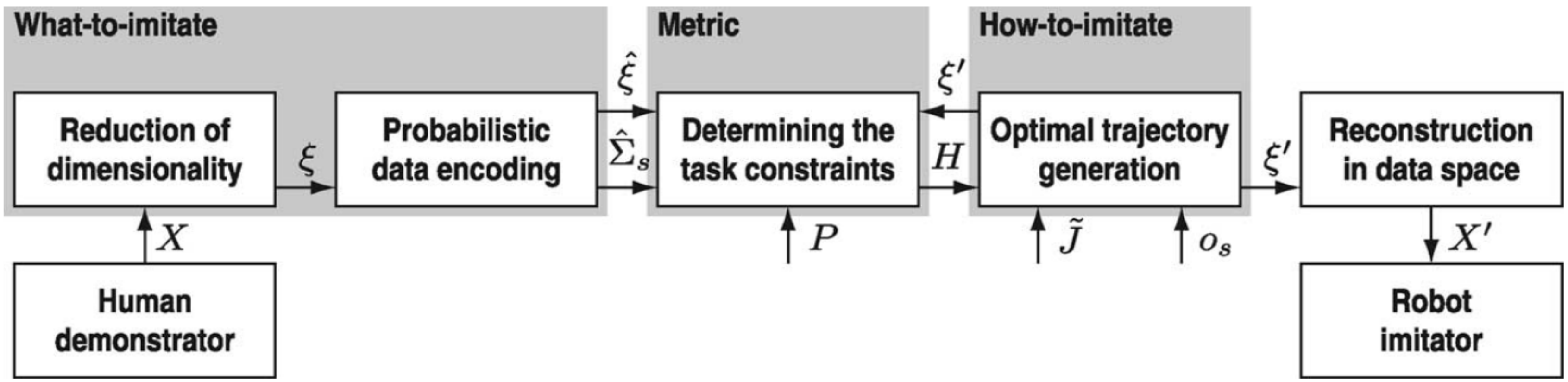


Model data as mixture of multivariate Gaussians



Shown is the *posterior probability* that a point was generated from i^th Gaussian: $\Pr(Y = i \mid x)$

Probabilistic Representation

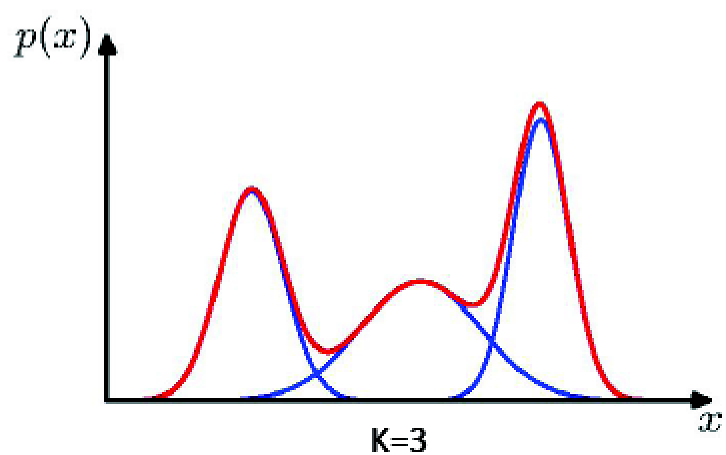


- Gaussian Mixture Model – 1D example

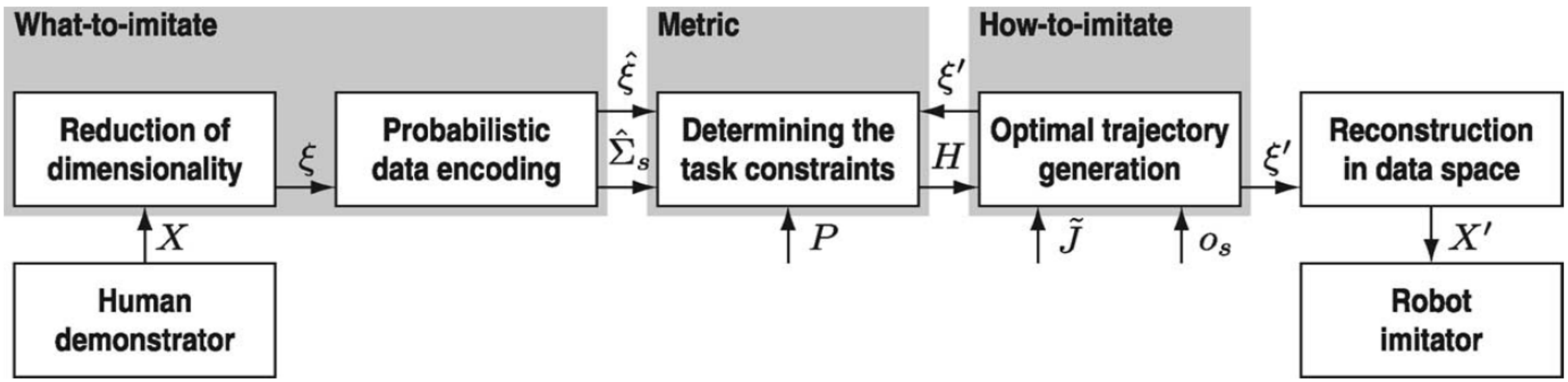
http://localhost:8888/notebooks/w08.ipynb

$$p(\mathbf{x}) = \sum_{k=1}^{K} \underbrace{\pi_k}_{\text{Mixing coefficient}} \underbrace{\mathcal{N}(\mathbf{x}|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)}_{\text{Component}}$$

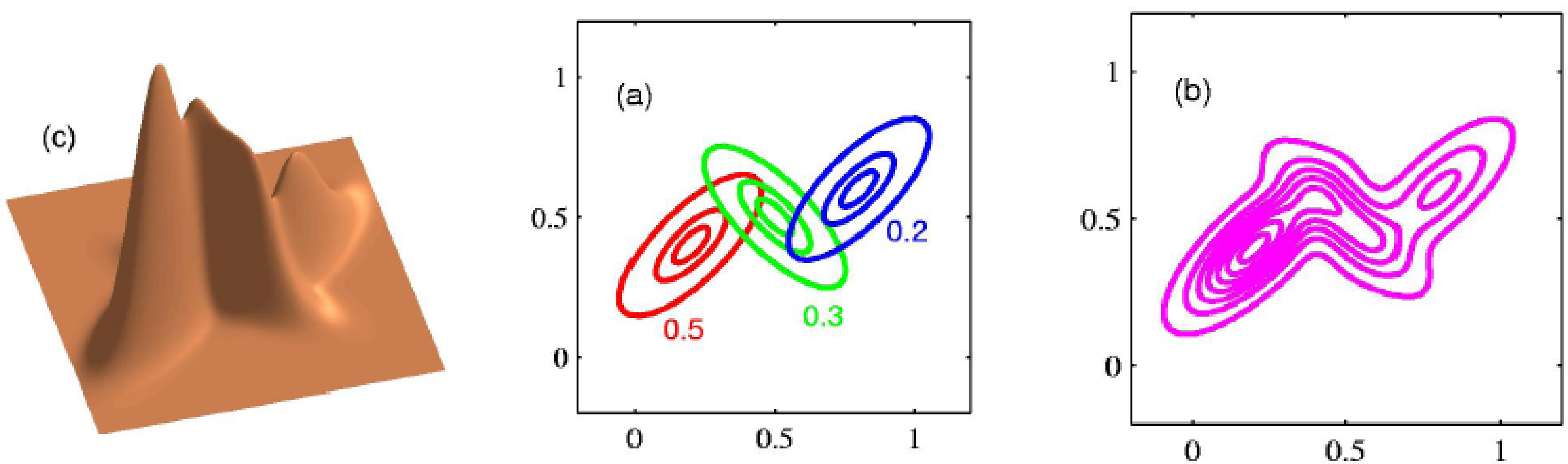
$$\forall k : \pi_k \geqslant 0 \qquad \sum_{k=1}^{K} \pi_k = 1$$



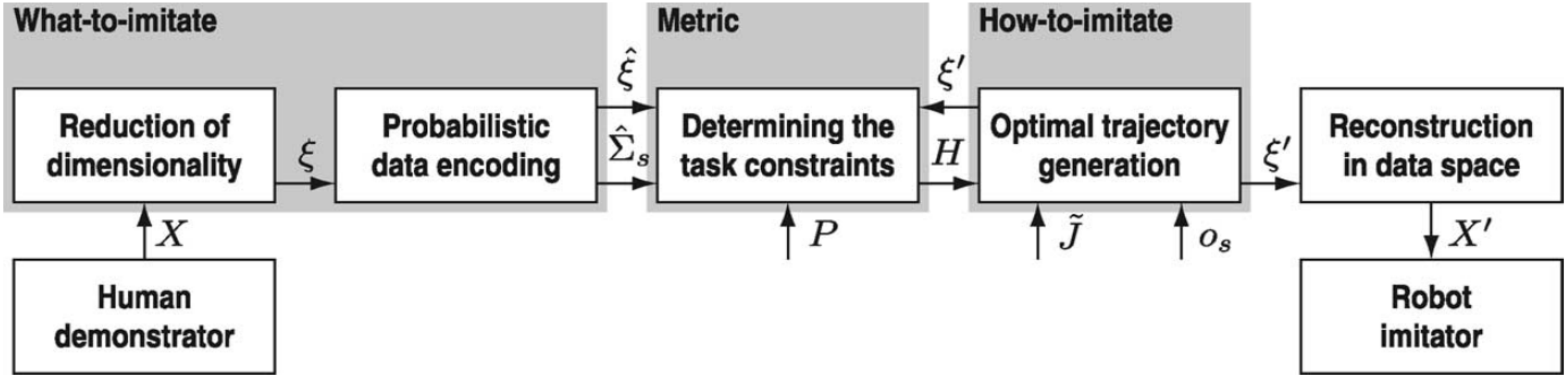
Probabilistic Representation



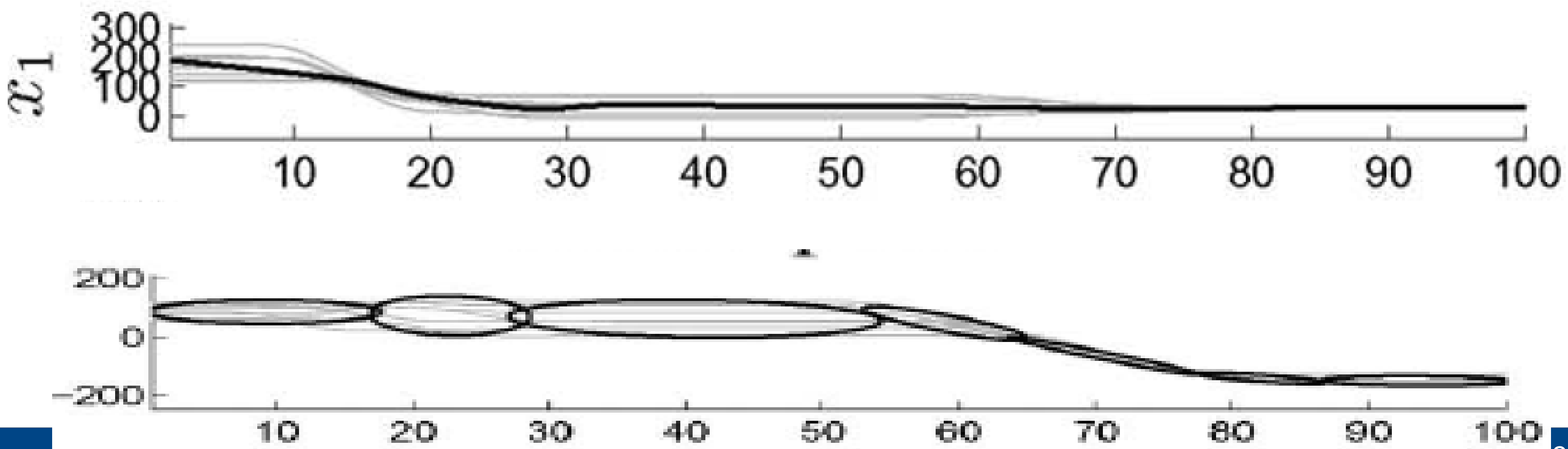
- Gaussian Mixture Model – 2D example



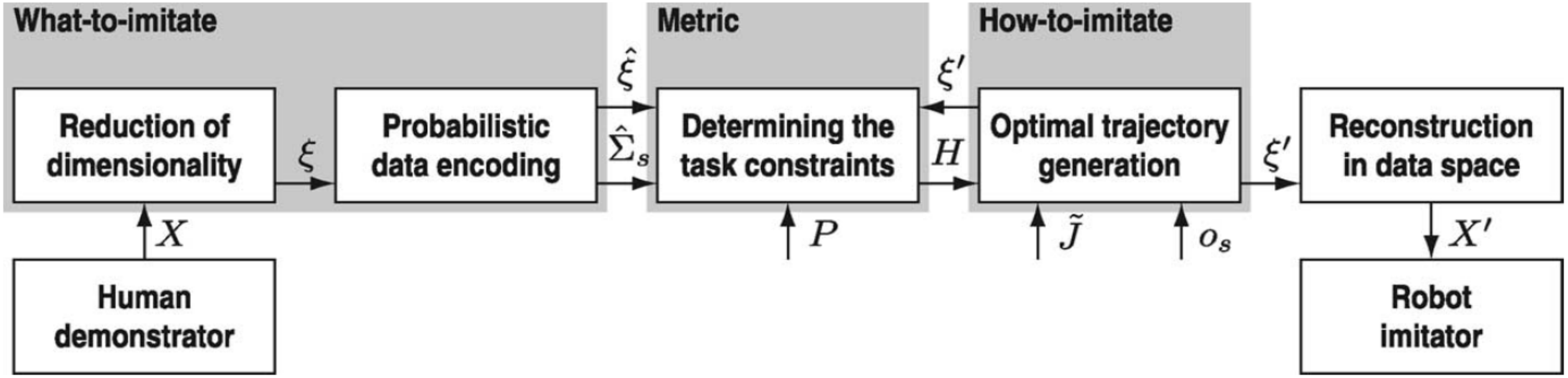
Probabilistic Representation



- Gaussian Mixture Model – 2D example



Reconstruct signals – Gaussian Mixture Regression

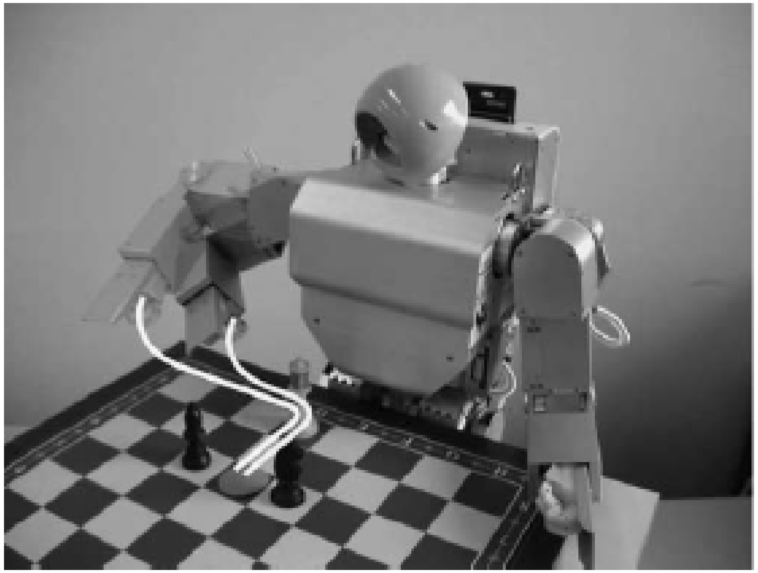


How to imitate

Optimization problem:

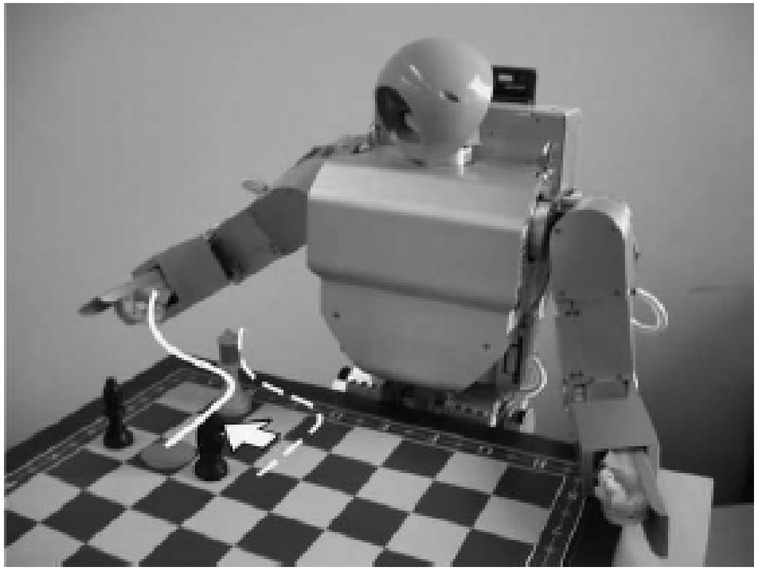
- Metric of imitation: measure evaluates the reproduction performance of a task.
  - A time-dependent similarity measure:
  - the relative importance of each variable and
  - the dependences across the variables
- How to imitate: We then compute the trajectory which optimizes the metric for a certain context, given
  - the robot's body constraints (encapsulated in a Jacobian matrix),
  - and the position of the object(s) in the scene.

What to imitate?



- The spatio-temporal variations and correlations across the variables
- Weak correlations at the beginning of the motion
- strong spatio-temporal correlation for grabbing the piece and pushing it toward the desired location without hitting the other pieces on the chessboard

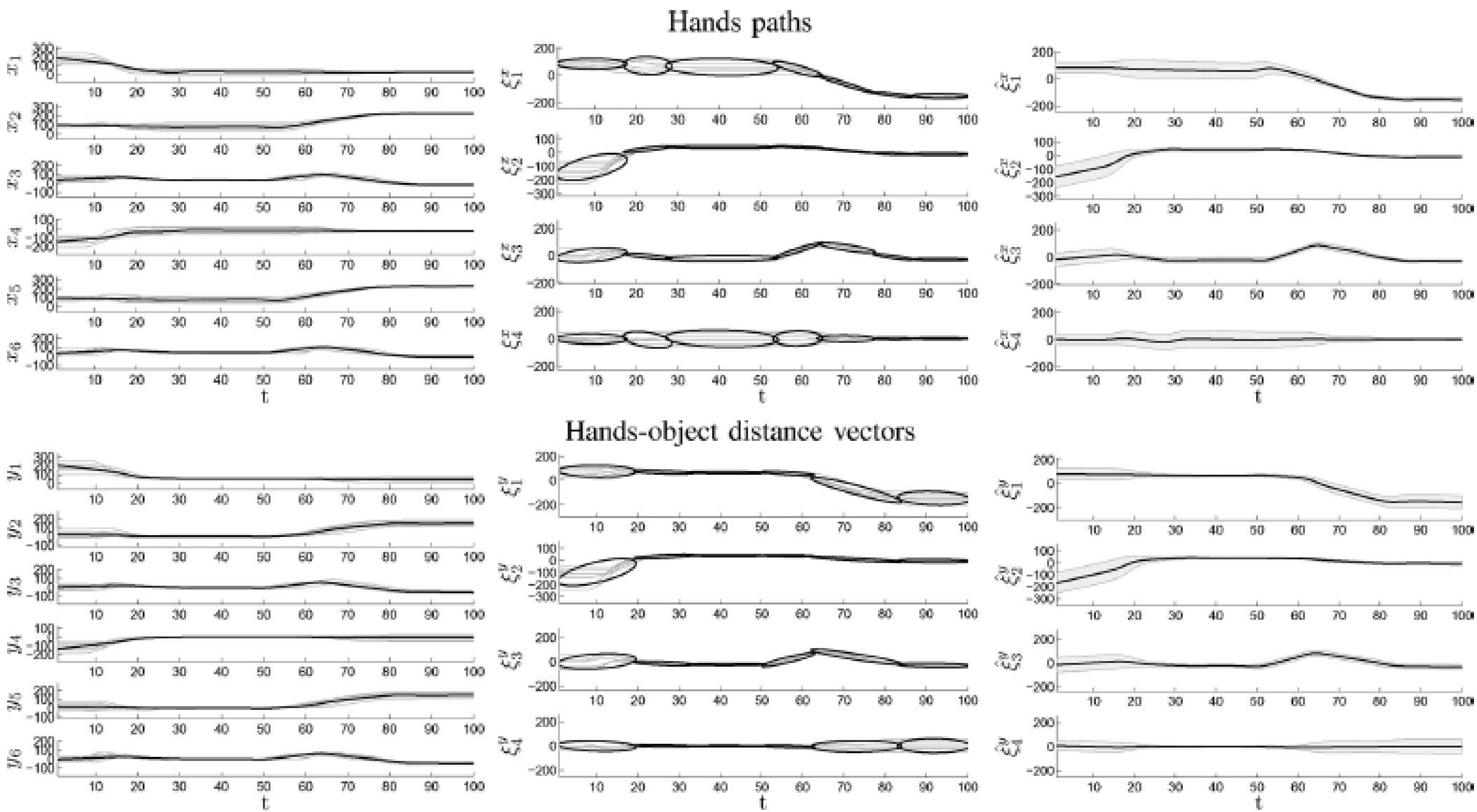
How to imitate?



- the demonstrated joint angles and hand path can be mutually exclusive in the imitator space, it is not possible to fulfill both constraints at the same time
- the trajectory which gives the optimal tradeoff between satisfying the constraints of the task (spatio-temporal correlations across the variables) and its own body constraints.

Experiments

- Control affected only the eight DOFs of the arms, the one DOF of the torso, and the two binary commands to open and close the robot's hands
- the task four to seven times by an expert user.
  - exploring as much as possible the variations allowed by the task,
- Once trained, the robot was required to reproduce each task under different constraints by placing the object at different locations in the robot's workspace



- hands–object relationships are highly constrained when the user is grabbing the object at time steps 30–50, i.e.,
- hands' paths are highly constrained at the end of the motion (the bucket is always placed at a specific location after being grabbed).

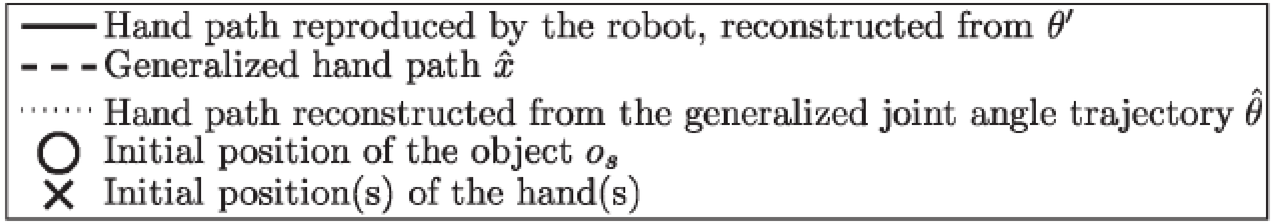


Fig. 7. Legend for Figs. 8–13.

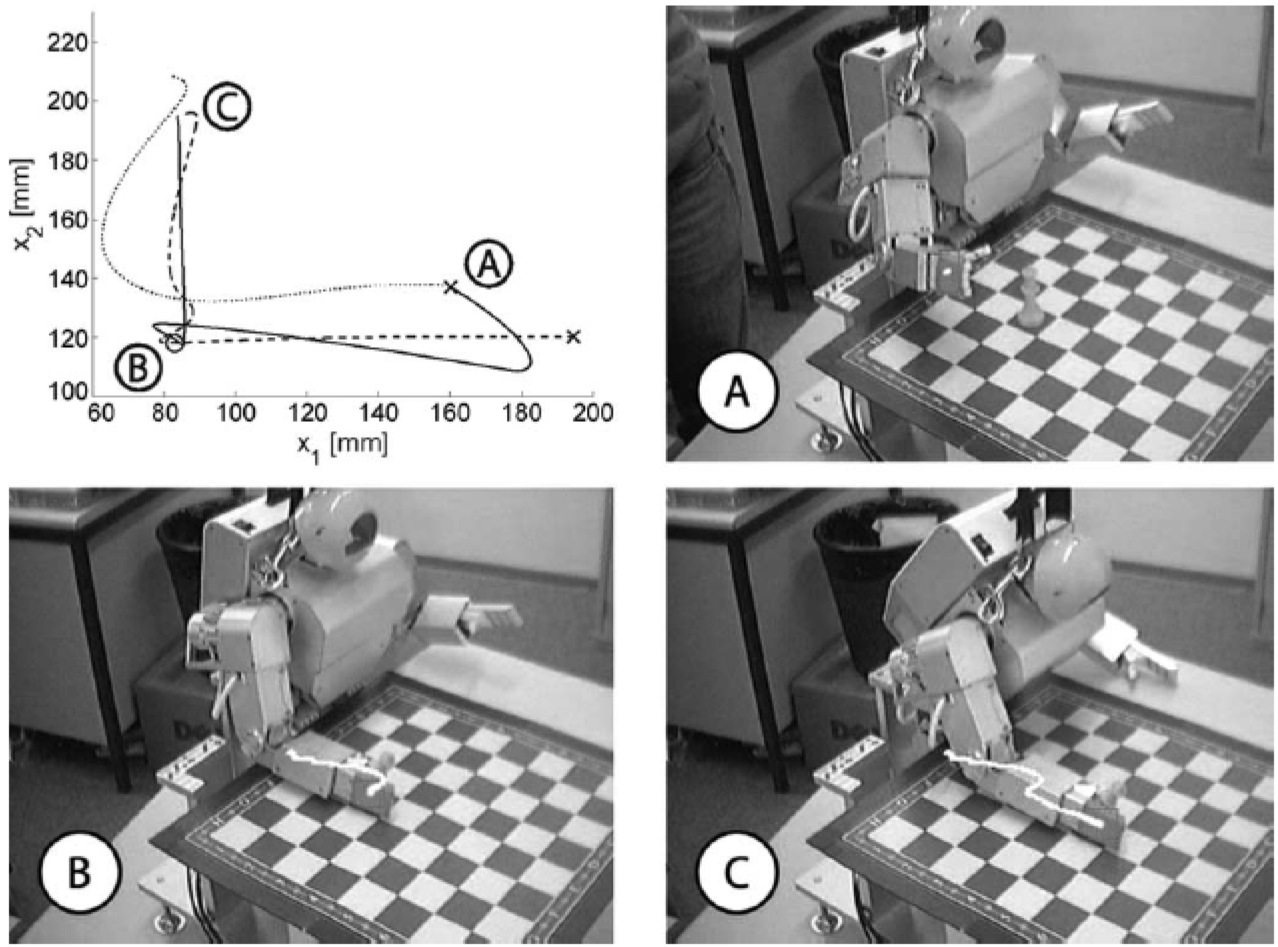
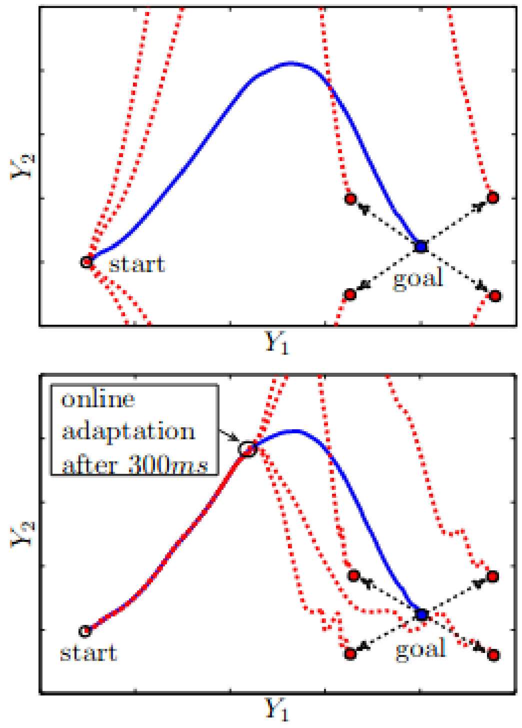


Fig. 8. Decomposition of the chess task when reproducing the task with an initial position of the object which is close to the generalized trajectories. The hands' paths have been tracked by a stereoscopic vision system.

Problems of the original DMPs

The original DMP formulation 3 drawbacks

- if start and goal pos are same
  - non linear term in cannot drive the system away from its initial state
- If g−x0 is close to zero;
  - the scaling g−x0 is problematic; a small change in g may lead to hug e accelerations, which can break the limits of the robot.
- Whenever a movement adapts to a new goal g new such that (g_{new}−x0) changes its sign compared to ($g_{original}$−x0)
  - the resulting generalization is mirrored.

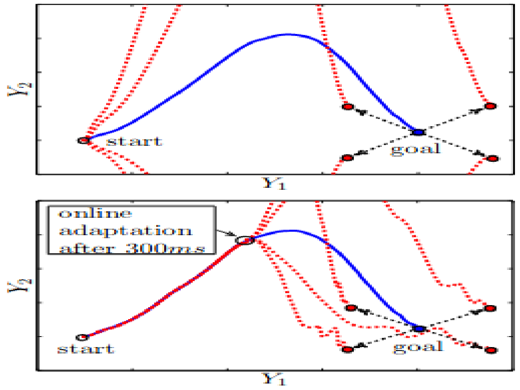


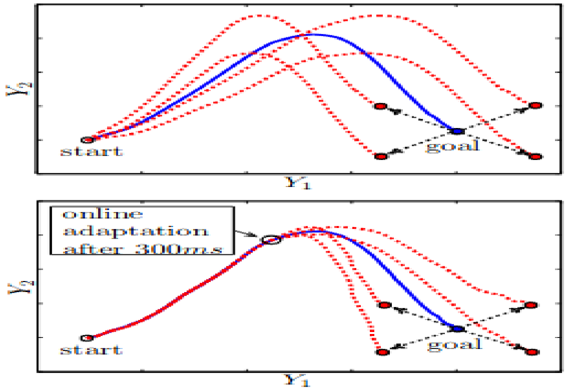
$$
\begin{aligned}
\tau \dot{v} &= K(g-x) - Dv + (g-x_0)\,f \\
\tau \dot{x} &= v\ ,
\end{aligned}
$$

Problems of the original DMPs

The original DMP formulation 3 drawbacks

- if start and goal pos are same
  - non linear term in cannot drive the system away from its initial state
- If g−x0 is close to zero;
  - the scaling g−x0 is problematic; a small change in g may lead to hug e accelerations, which can break the limits of the robot.
- Whenever a movement adapts to a new goal g new such that (g_{new}−x0) changes its sign compared to ($g_{original}$−x0)
  - the resulting generalization is mirrored.





$$\begin{aligned} \tau\dot{v} &= K(g-x) - Dv + (g-x_0)\,f \\ \tau\dot{x} &= v\ , \end{aligned}$$

$$\begin{aligned} \tau\dot{v} &= K(g-x) - Dv - K(g-x_0)s + Kf(s) \\ \tau\dot{x} &= v\ , \end{aligned}$$

Extending original formulation: obstacle avoidance

A major feature of using dynamic systems for movement representation is robustness against perturbations.

$$\tau\dot{\mathbf{v}} = \mathbf{K}(\mathbf{g}-\mathbf{x})-\mathbf{D}\mathbf{v}-\mathbf{K}(\mathbf{g}-\mathbf{x}_0)\,s+\mathbf{K}\mathbf{f}(s)+\mathbf{p}(\mathbf{x},\mathbf{v})$$

$$\mathbf{p}(\mathbf{x},\mathbf{v}) = \gamma\,\mathbf{R}\,\mathbf{v}\,\varphi\exp(-\beta\,\varphi) \quad ,$$

$\mathbf{R}$ is a rotational matrix with axis $\mathbf{r} = (\mathbf{x}-\mathbf{o})\times\mathbf{v}$

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

2009 IEEE International Conference on Robotics and Automation
Kobe International Conference Center
Kobe, Japan, May 12-17, 2009

Learning and Generalization of Motor Skills by Learning from Demonstration

Peter Pastor, Heiko Hoffmann, Tamim Asfour, and Stefan Schaal

2012 12th IEEE-RAS International Conference on Humanoid Robots
Nov.29-Dec.1, 2012. Business Innovation Center Osaka, Japan

Towards Associative Skill Memories

Peter Pastor*, Mrinal Kalakrishnan*, Ludovic Righetti*†, Stefan Schaal*†
*Computational Learning and Motor Control Lab, University of Southern California, Los Angeles, CA 90089, USA.
†Autonomous Motion Department, Max-Planck-Institute for Intelligent Systems, 72076 Tübingen, Germany.
pastorsa@usc.edu, kalakris@usc.edu, ludovic.righetti@a3.epfl.ch, sschaal@usc.edu

Associative Skill Memories

https://www.youtube.com/watch?v=lL4-onuLDy0

- Stereotypical motions generate stereotypical sensory feedbacks
  - e.g., in terms of kinesthetic variables, haptic variables, or, if processed appropriately, visual variables
- a movement primitive executed towards a particular object in the environment will associate a large number of sensory variables that are typical for this manipulation skill.
- These association can be used to increase robustness towards perturbations,and they also allow failure detection and switching towards other behaviors.

Associative Skill Memories

- Learn the stereotypical sensory feedback: $F_{des}(t)$

$$\tau \dot{v} = K(g - x) - D\,v - K(g - x_0)s + K\,f(s)$$
$$\tau \dot{x} = v \ ,$$
$$f(s) = \frac{\sum_i \psi_i(s)\,\theta_i\,s}{\sum_i \psi_i(s)} \ ,$$
$$\tau\,\dot{s} = -\alpha\,s \ ,$$

Online trajectory generation using sensory feedback:

$$\begin{aligned} \tau \dot{v} &= K(g - x) - D\,v - K(g - x_0)s + K\,f(s) + \zeta \\ \tau \dot{x} &= v \ . \end{aligned}$$

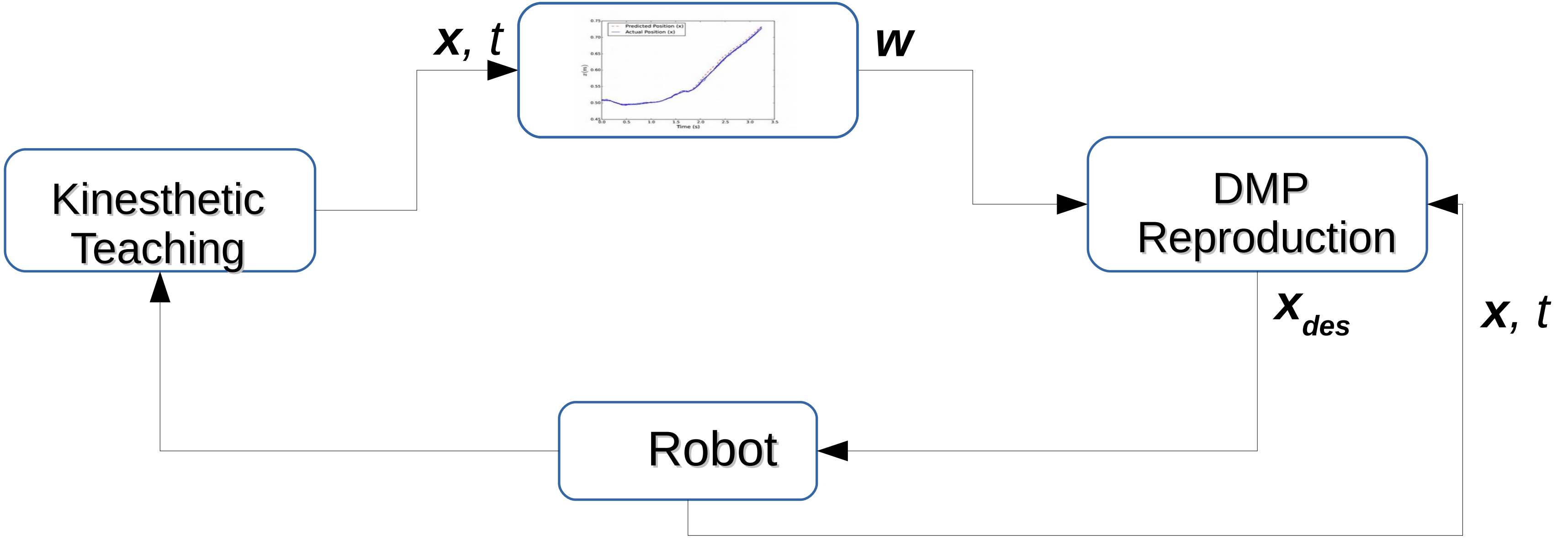
$$\zeta = \mathbf{K}_1 \mathbf{J}^T_{sensor} \mathbf{K}_2 (\mathbf{F} - \mathbf{F}_{des}) \ ,$$

Exploit predictions

- The planned action trajectories might fail due to noise/perturbations
- **Failure** can be detected if **predictions do not hold**.
- The trajectories can be **corrected** through **minimizing the prediction error**.
  - Make **corrective movements** so that the difference between predicted and actual sensory feedback is minimized.
  - i.e. (predicted-actual=0)

Exploit force-feedback predictions

Assume a movement is learned by demonstration and encoded as DMP.



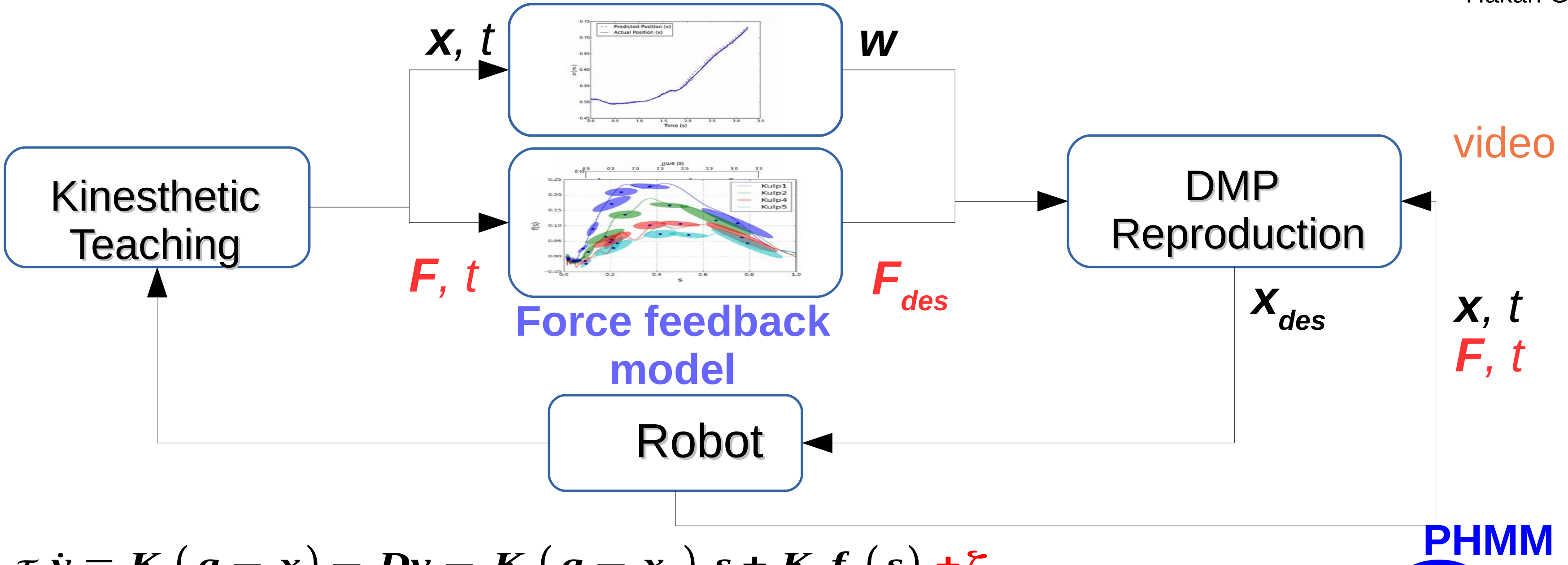
$$\tau \dot{v} = K(g - x) - Dv - K(g - x_0)s + Kf(s)$$

$$\tau \dot{x} = v$$

Exploit force-feedback predictions

Hakan Girgin

video



$$\tau \dot{v} = K(g - x) - Dv - K(g - x_0)s + Kf(s) + \zeta$$

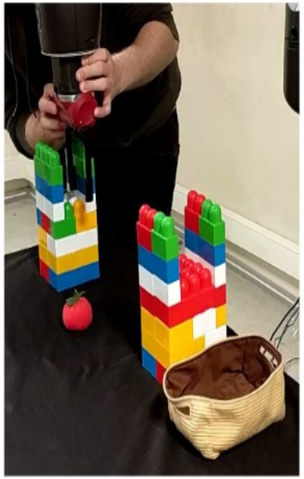
$$\tau \dot{x} = v$$

$$\zeta = K_1 J^T_{sensor} K_2 (F - F_{des})$$

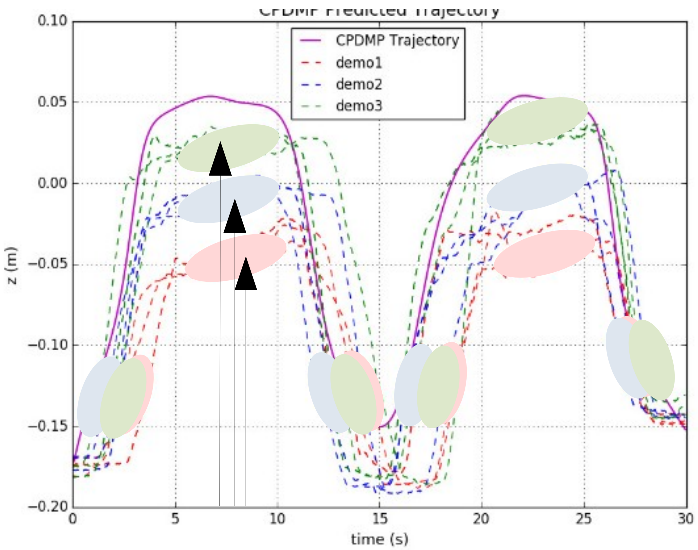
H. Girgin, E. Ugur, Associative Skill Memory Models, IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 6043-6048, 2018.

Trajectory encoding with probabilistic models

- Multiple trajectories encoded as Task Parameterized Hidden Markov Models (TP-HMM)

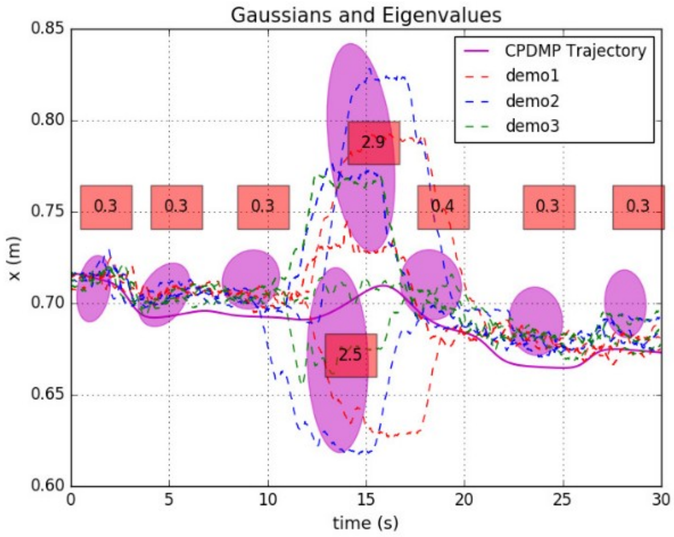


(a) Block height: 14.5 cm.



Exploit force-feedback predictions

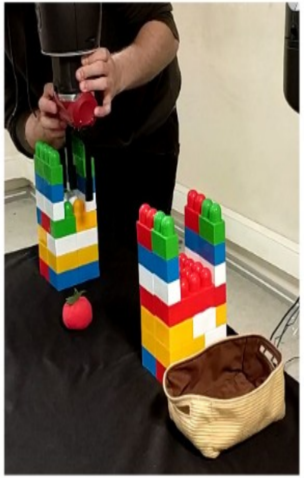
- Learn expected force feedback
  - Use expected feedback to enable compliance
- Learn motion trajectory distribution
  - Enable compliance if large variance in parts of the demonstrations
  - Disable compliance if no variance in the demonstrations



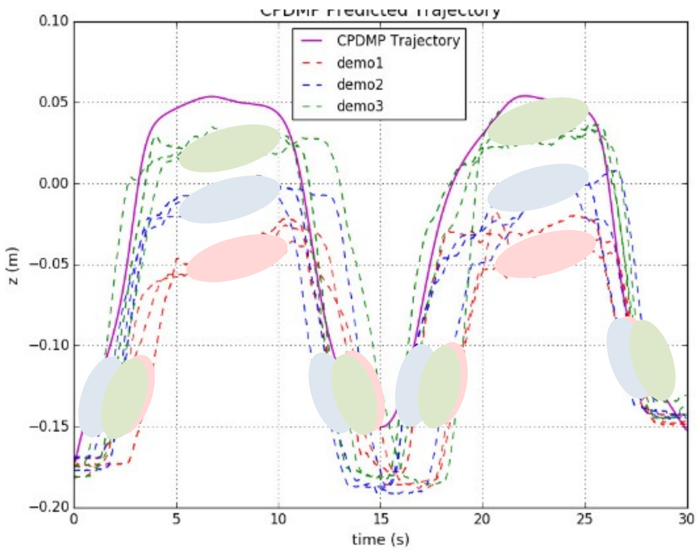
Ugur et al. Robotica, Video

Scale up in prediction

- Multiple trajectories encoded as Task Parameterized Hidden Markov Models (TP-HMM)



(a) Block height: 14.5 cm.



- TP-GMM, TP-HMM, ProMP can only learn linear relationship between parameters and motion trajectories

Challenges to scale up

- Learn distribution from multiple demonstrations
  - with possibly multiple modes of operations.
- Discover task-related features
  - embedded in the multi-modal sensorimotor trajectories
- External parameters/goals ↔ complex movements
  - from few demonstrations.
- High-dimensional SM spaces ↔ complex motions
  - using large number of demonstrations.
- Respond to external perturbations on-the-fly