IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Lab-STICC

INTERACTIVE MACHINE LEARNING

PRINCIPLES OF INTERACTIVE LEARNING: IMITATION LEARNING AND ACTIVE REQUESTS

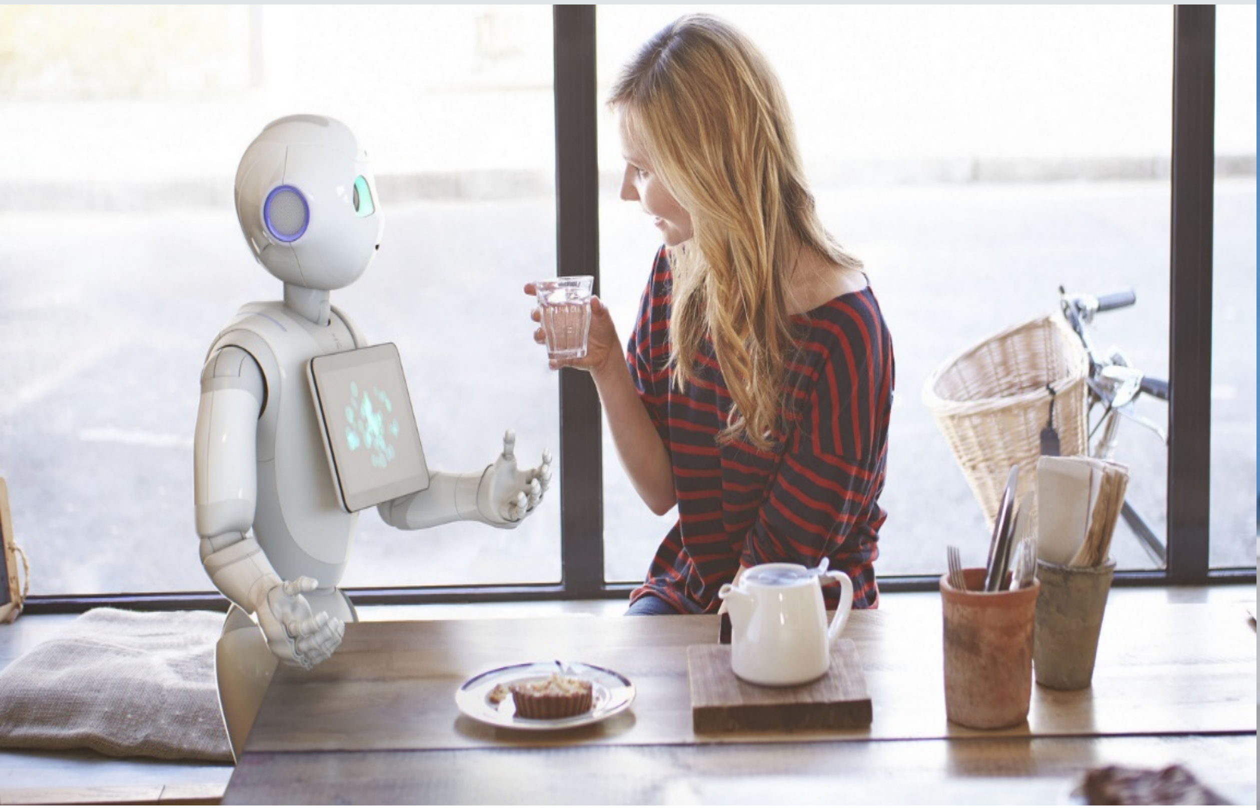
Mai Nguyen
http://nguyensmai.free.fr

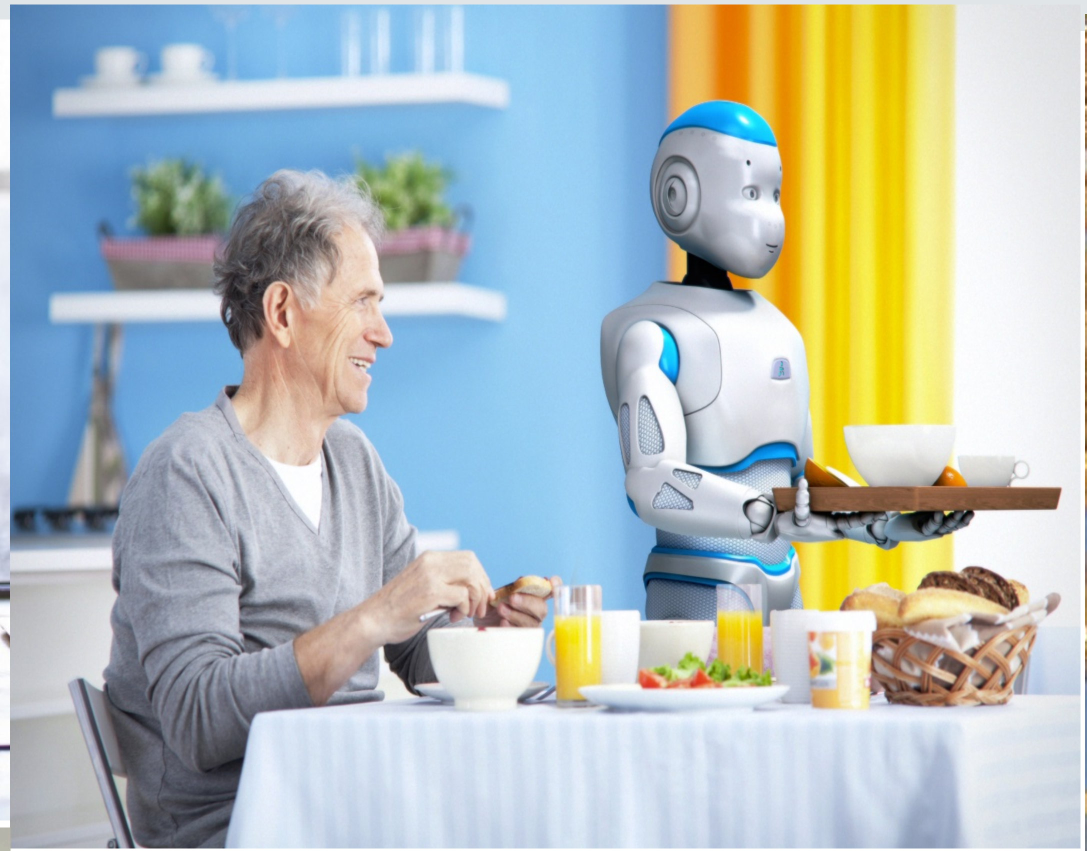
WHAT DOES INTERACTIVE LEARNING MEAN?

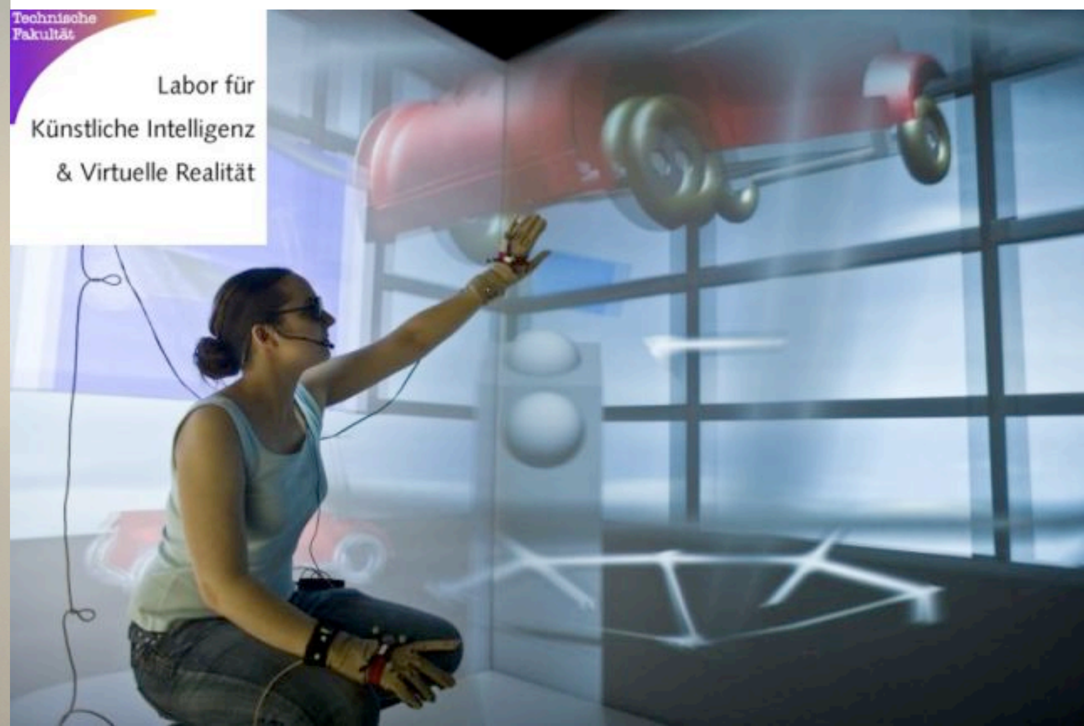
IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

1. WHAT DOES INTERACTIVE LEARNING MEAN?

1.1. The Artificial Agent in Its Environment



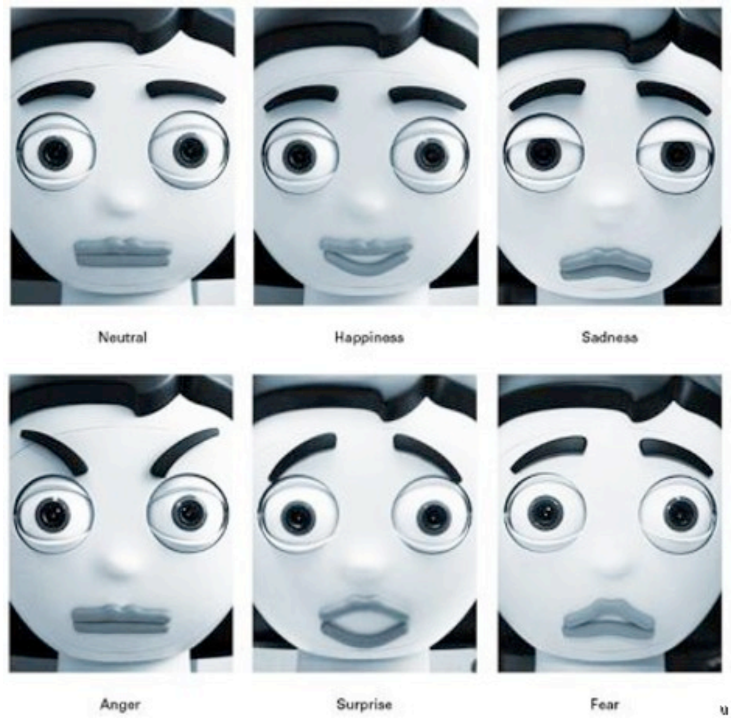




1. WHAT DOES INTERACTIVE LEARNING MEAN?

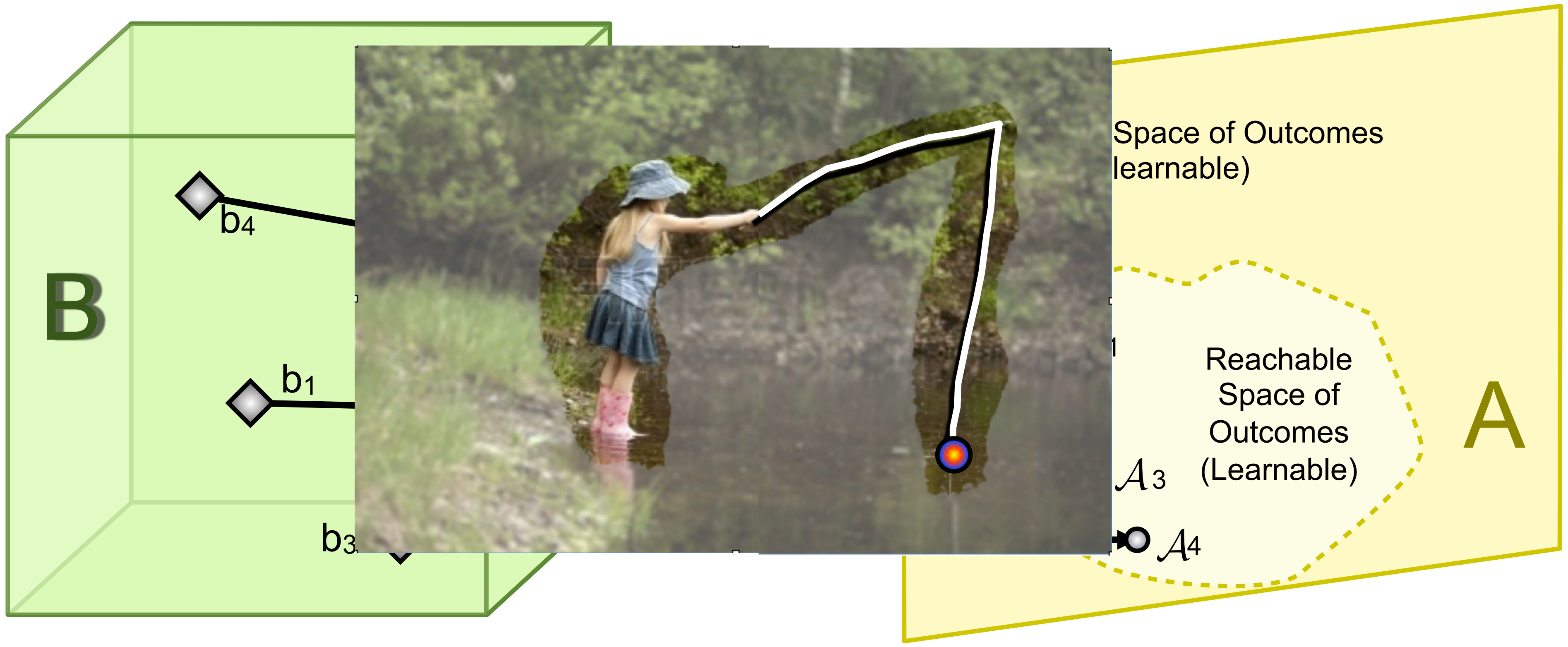
1.2. Challenges

- **Vocal interaction**: speech recognition, speech generation (text to-speech)
- **Natural interaction** : multi-modal, non-verbal interaction, gesture, expressive emotion-based interaction
- **Socio-cognitive skills** : socially acceptable behaviours, turn-taking, coordination, theory of mind
- **Physical interaction** : touch (tactile sensors), grasping, manipulation



1.3. Theoretical approaches

- **Embodiment** : the environment has a physical incarnation, the agent has a physical incarnation => its learning, capacities, behaviour depends on its physical body
- **Enactivism** : Learning of the agent in its environment
- **Life-long learning** : the environment and tasks can change
- **Developmental approaches** : there is an orderly way to learn multiple tasks, the learning is progressive and hierarchical -> Developmental psychology
- **Cognitive approaches** : inspired by cognitive science, neuroscience, neuronal computation models. Decomposes into a task into cognitive skills/ functions



B
b4
b1
b3
Space of Outcomes
learnable)
Reachable
Space of
Outcomes
(Learnable)
$\mathcal{A}_3$
$\mathcal{A}_4$
A

PROGRAMMING BY DEMONSTRATION

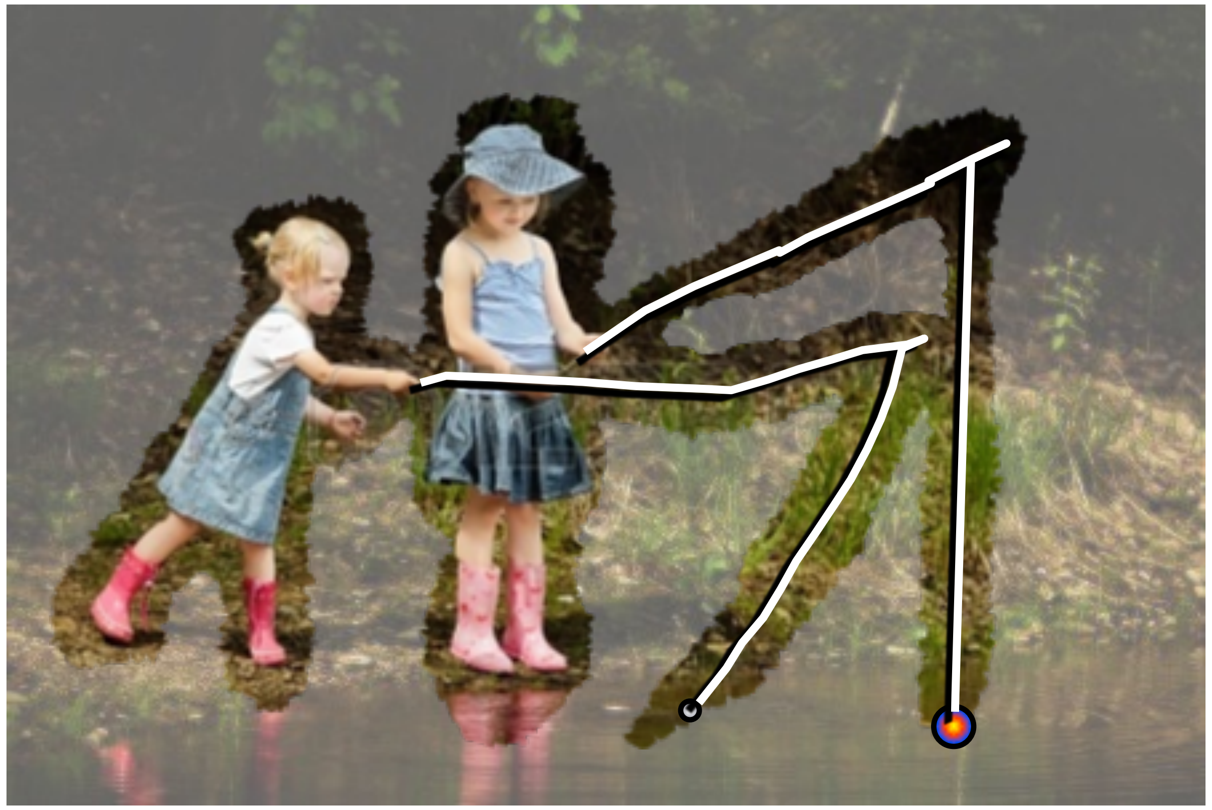
IMITATION LEARNING

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

2. PROGRAMMING BY DEMONSTRATION

2.1. Learning by imitation

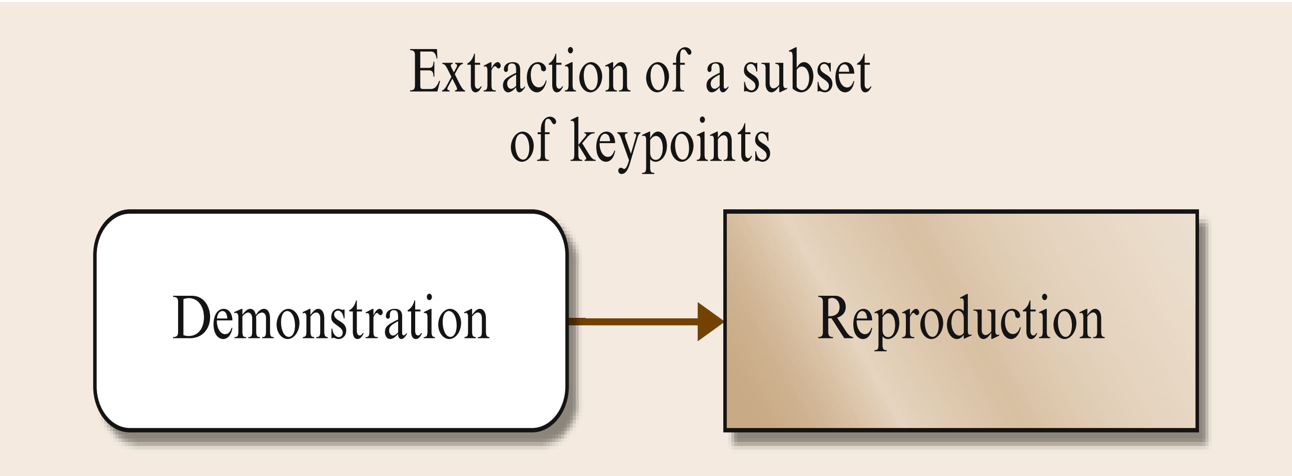
Mimicry



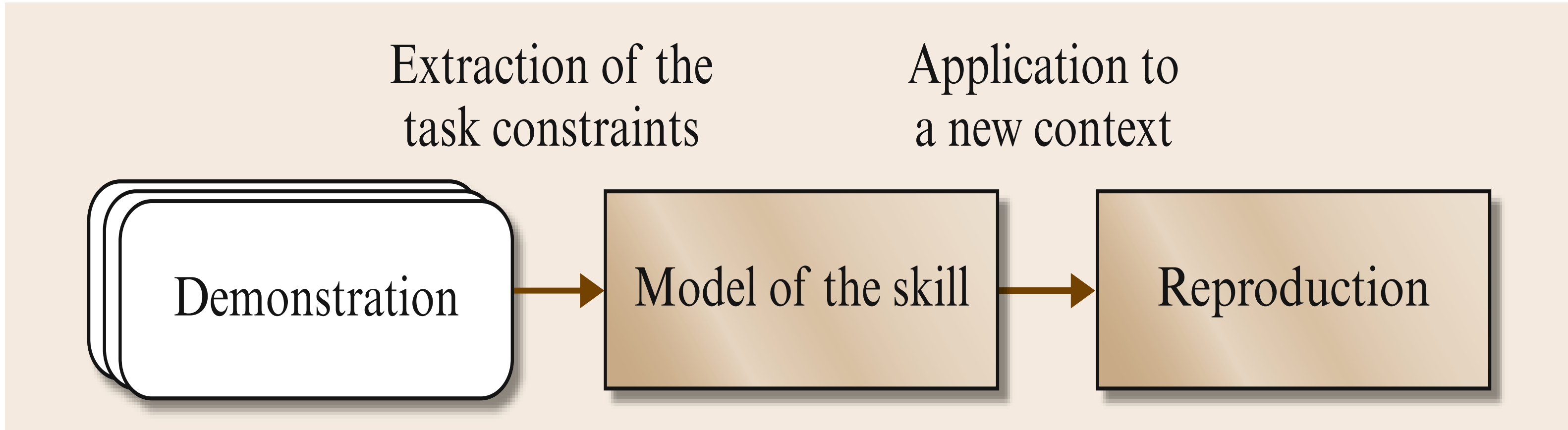
Emulation

- An implicit, *natural* means of training a machine that would be **accessible to lay people**
- A powerful mechanism for **reducing the complexity of search** spaces for learning
- Studying and modeling the **coupling of perception and action**

Copying the demonstrated movements



Generalize across sets of demonstrations.



- How to **generalize** a task
- How to **evaluate** a reproduction attempt
- How to better define the role of the **user** during learning

Observation of multiple demonstrations

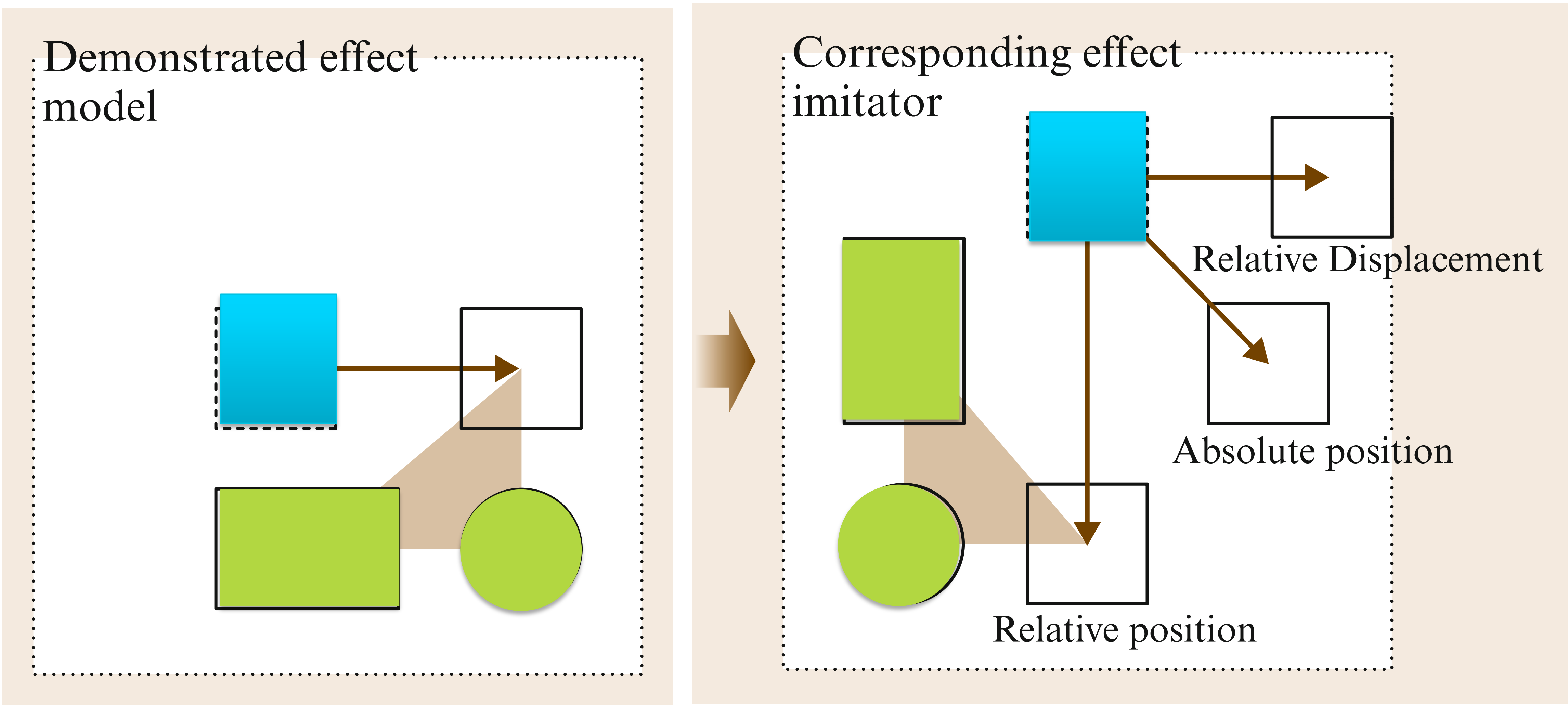
Reproduction of a generalized motion in a different situation

The different types of representation to encode a skill

- **a low-level representation** of the skill, taking the form of a non-linear mapping between sensory and motor information, which we will later refer to as *trajectories encoding*
- **high-level representation** of the skill that decomposes the skill in a sequence of action-perception units, which we will refer to as *symbolic encoding*

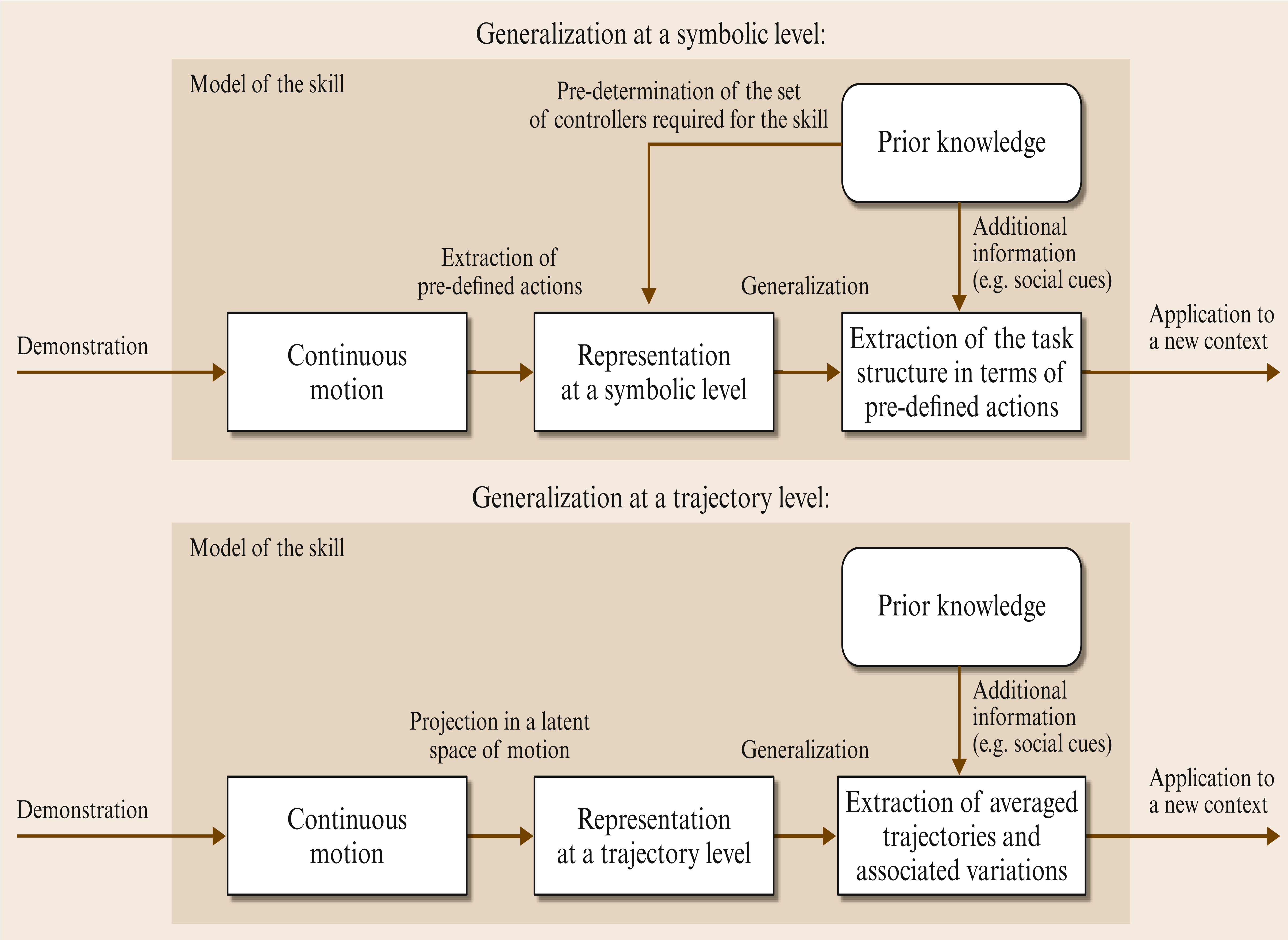
what to imitate, how to imitate, when to imitate and who to imitate : making no assumptions on the type of skills that may be transmitted

- **Metric of imitation performance:** extract the important features characterizing the skill
- An **optimal controller** to imitate by trying to **minimize this metric**



2. PROGRAMMING BY DEMONSTRATION

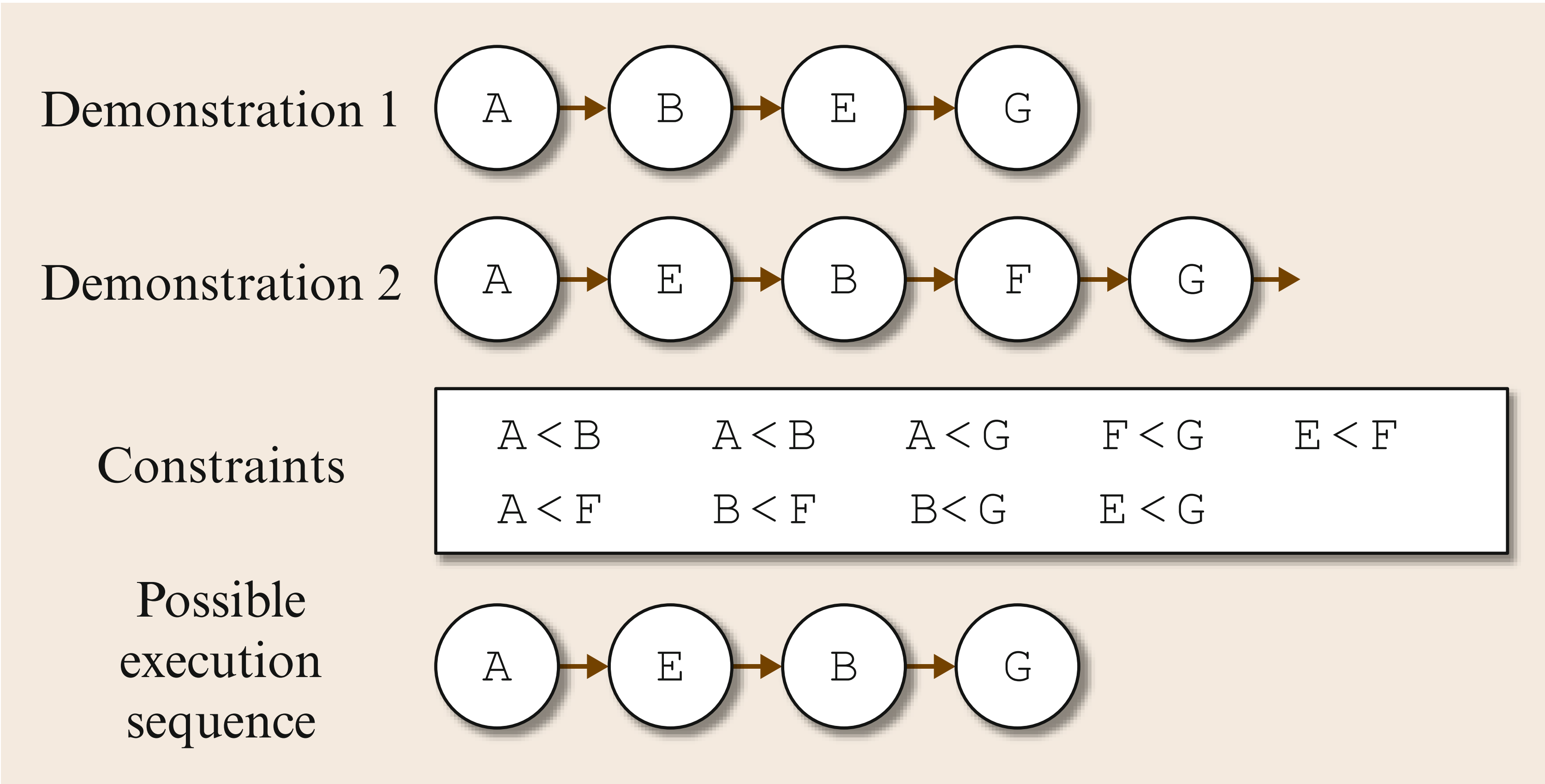
2.4. How to evaluate a reproduction attempt



IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

- **Segment and encode** the task according to sequences of *predefined* actions

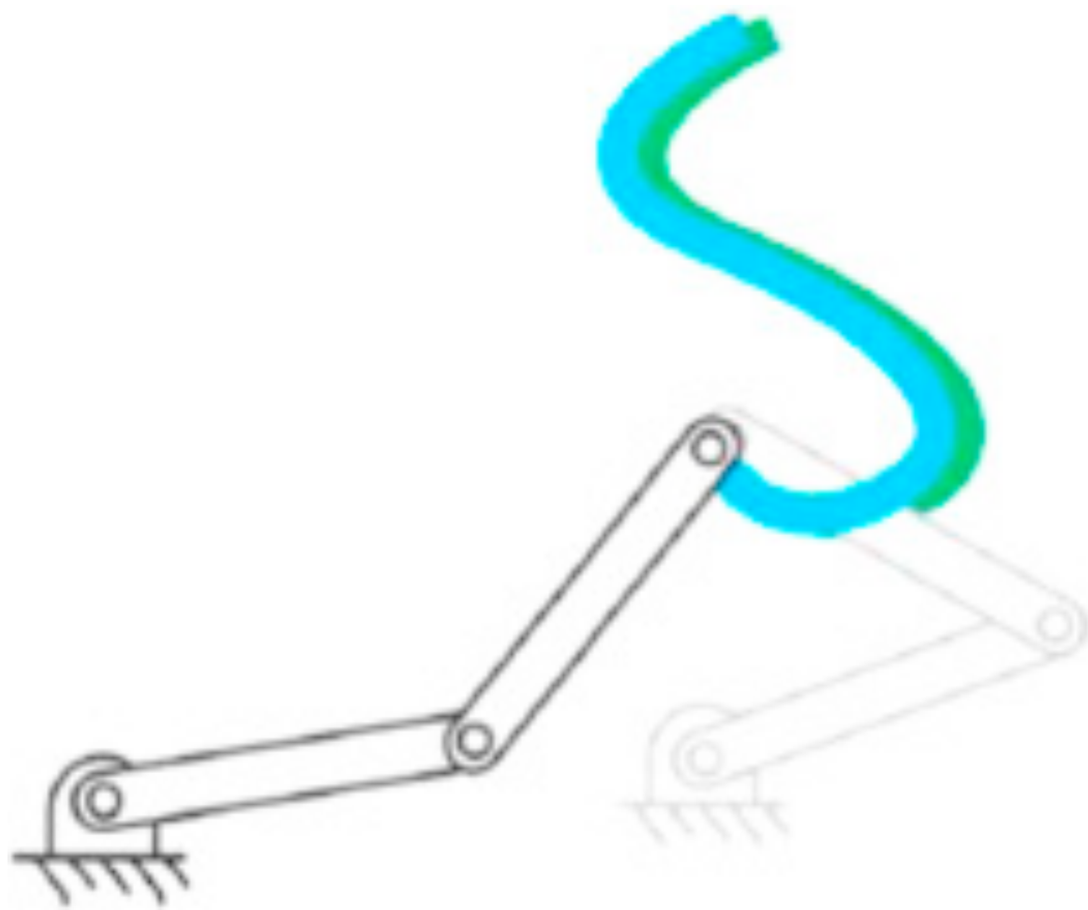
- **Encoding and regenerating** (HMM)



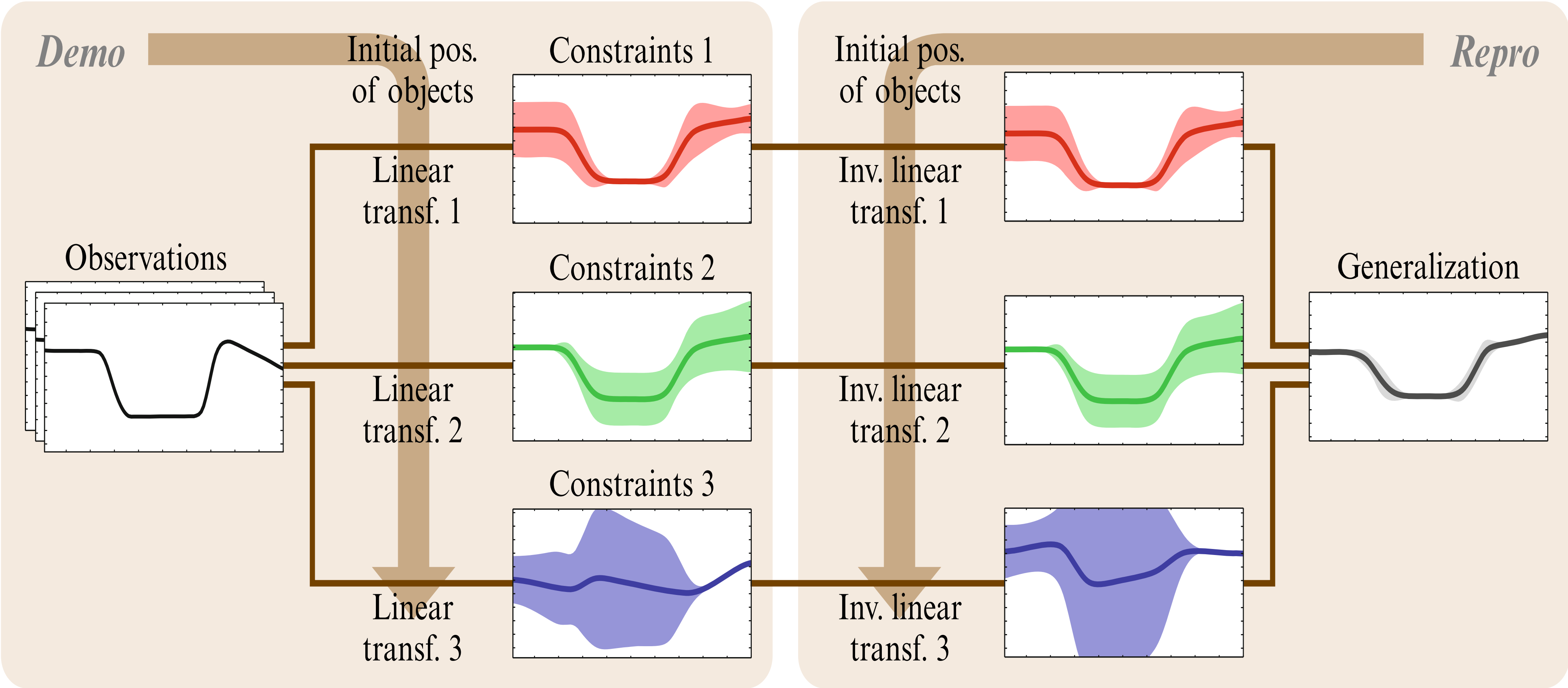
2. PROGRAMMING BY DEMONSTRATION

2.5. Learning and Encoding a Skill at Trajectory-Level

- Choosing the **variables** well to encode a particular movement
- Encode human movements in **joint, task or torque** space
- **Cyclic/discrete** motion
- ***Skill Encoding Based on Statistical Modeling*** *:* how statistical learning techniques deal with the high variability inherent to the demonstrations.



- **Gaussian Mixture Model (GMM)** to encode a set of trajectories
- **Gaussian Mixture Regression (GMR)** to retrieve a smooth generalized version of these trajectories and associated variabilities



A dataset $\xi = \{\xi_j\}_{j=1}^{N}$ is defined by N observations $\xi_j \in \mathbb{R}^D$ of sensory data changing through time (e.g., joint angle trajectories, hand paths), where each datapoint $\xi_j = \{\xi_t, \xi_s\}$ consists of a temporal value $\xi_t \in \mathbb{R}$ and a spatial vector $\xi_s \in \mathbb{R}^{(D-1)}$. The dataset ξ is modelled by a gaussian mixture model (GMM) of K components, defined by the probability density function

$$p(\xi_j) = \sum_{k=1}^{K} \pi_k \mathcal{N}(\xi_j; \mu_k, \Sigma_k),$$

where π_k are prior probabilities and $\mathcal{N}(\xi_j; \mu_k, \Sigma_k)$ are Gaussian distributions defined by *mean* vectors μ_k and *covariance* matrices Σ_k, whose temporal and spatial components can be represented separately as

$$\mu_k = \{\mu_{t,k}, \mu_{s,k}\}, \quad \Sigma_k = \begin{pmatrix} \Sigma_{tt,k} & \Sigma_{ts,k} \\ \Sigma_{st,k} & \Sigma_{ss,k} \end{pmatrix}.$$

For each component k, the expected distribution of ξ_s given the temporal value ξ_t is defined by

$$\begin{aligned} p(\xi_s|\xi_t, k) &= \mathcal{N}(\xi_s; \hat{\xi}_{s,k}, \hat{\Sigma}_{ss,k}), \\ \hat{\xi}_{s,k} &= \mu_{s,k} + \Sigma_{st,k}(\Sigma_{tt,k})^{-1}(\xi_t - \mu_{t,k}), \\ \hat{\Sigma}_{ss,k} &= \Sigma_{ss,k} - \Sigma_{st,k}(\Sigma_{tt,k})^{-1}\Sigma_{ts,k}. \end{aligned}$$

By considering the complete GMM, the expected distribution is defined by

$$p(\xi_s|\xi_t) \ = \sum_{k=1}^{K} \beta_k \, \mathcal{N}(\xi_s; \hat{\xi}_{s,k}, \hat{\Sigma}_{ss,k}),$$

where $\beta_k = p(k|\xi_t)$ is the probability of the component k to be responsible for ξ_t, i. e.,

$$\beta_k = \frac{p(k)p(\xi_t|k)}{\sum_{i=1}^{K} p(i)p(\xi_t|i)} = \frac{\pi_k \mathcal{N}(\xi_t; \mu_{t,k}, \Sigma_{tt,k})}{\sum_{i=1}^{K} \pi_i \mathcal{N}(\xi_t; \mu_{t,i}, \Sigma_{tt,i})}.$$

By using the linear transformation properties of Gaussian distributions, an estimation of the conditional expectation of ξ_s given ξ_t is thus defined by $p(\xi_s|\xi_t) \sim \mathcal{N}(\hat{\xi}_s, \hat{\Sigma}_{ss})$, where the parameters of the Gaussian distribution are defined by

$$\hat{\xi}_s = \sum_{k=1}^{K} \beta_k \, \hat{\xi}_{s,k} \ , \quad \hat{\Sigma}_{ss} = \textstyle\sum_{k=1}^{K} \beta_k^2 \, \hat{\Sigma}_{ss,k}.$$

By evaluating $\{\hat{\xi}_s, \hat{\Sigma}_{ss}\}$ at different time steps ξ_t, a generalized form of the motions $\hat{\xi} = \{\xi_t, \hat{\xi}_s\}$ and associated covariance matrices $\hat{\Sigma}_{ss}$ describing the constraints are computed. If multiple constraints are considered (e.g., considering actions $\xi^{(1)}$ and $\xi^{(2)}$ on two different objects), the resulting constraints are computed by first estimating $p(\xi_s|\xi_t) = p(\xi_s^{(1)}|\xi_t) \cdot p(\xi_s^{(2)}|\xi_t)$ and then computing $\mathbb{E}[p(\xi_s|\xi_t)]$ to reproduce the skill. See Fig. 59.14 for an illustration of this method to learning continuous constraints in a set of trajectories. (After [59.66])

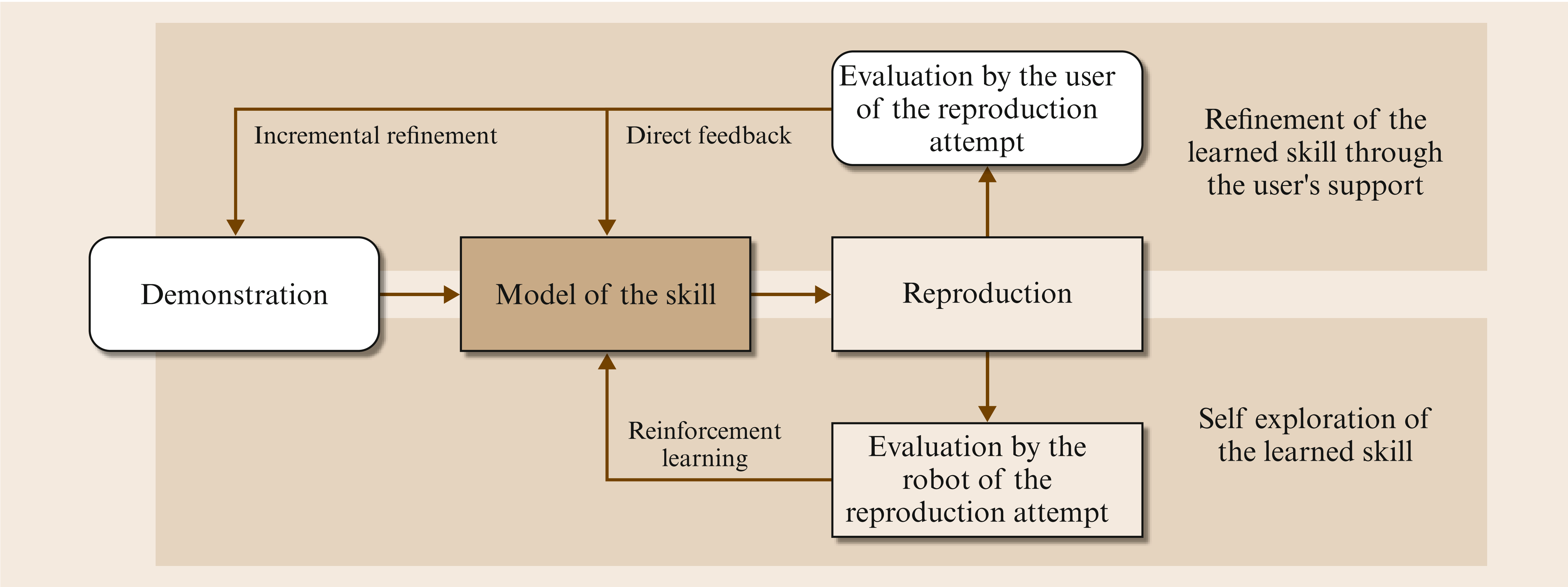
These early works highlighted the importance of providing a set of examples that the robot can use:

- by constraining the demonstrations to **modalities** that the robot can understand
- by providing a sufficient **number of examples** to achieve a desired generality.
- by providing **examples representative** enough of the all the situations
- By limiting the **correspondence problems**

3.1. Limitations of Programming by Demonstration

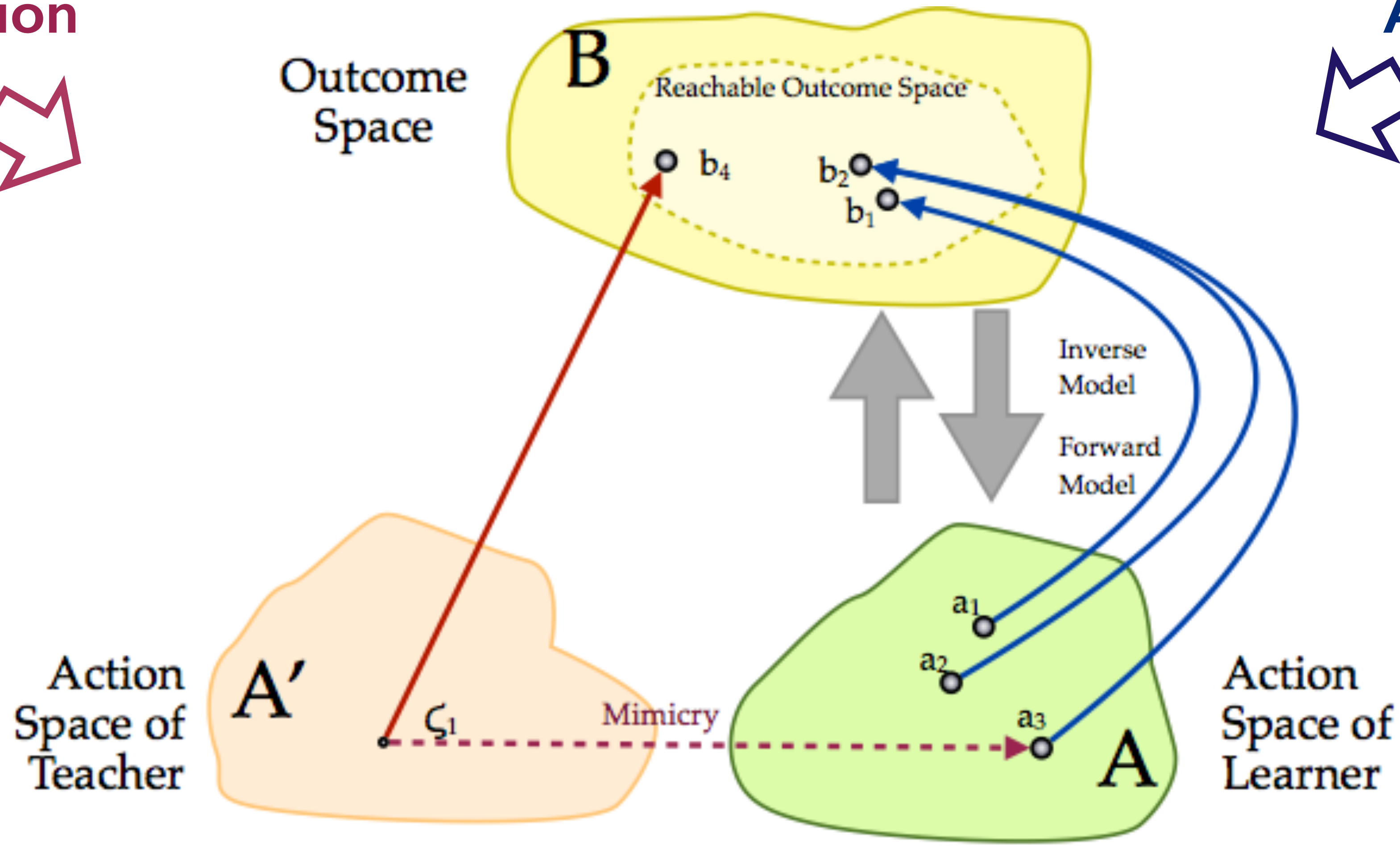
- give the **teacher an active role** during learning
- the interaction aspect of the **transfer process**
- *Social cues*
- Pointing and gazing
- Vocal speech recognition
- Prosody of the speech

PbD can be jointly used with other learning strategies to overcome some limitations of PbD



Socially Guided Exploration

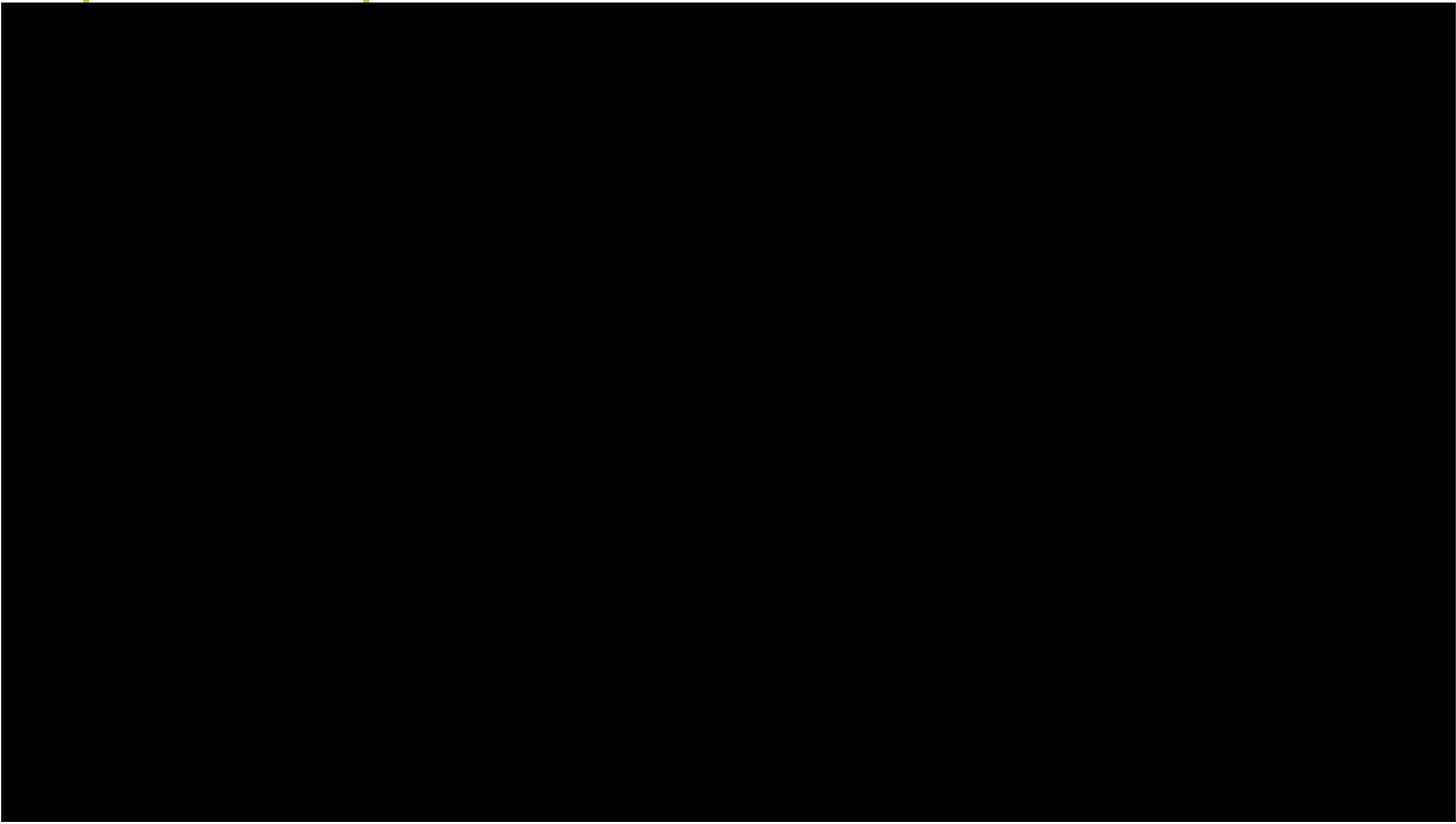
Autonomous Exploration



3. BEYOND PROGRAMMING BY DEMONSTRATION

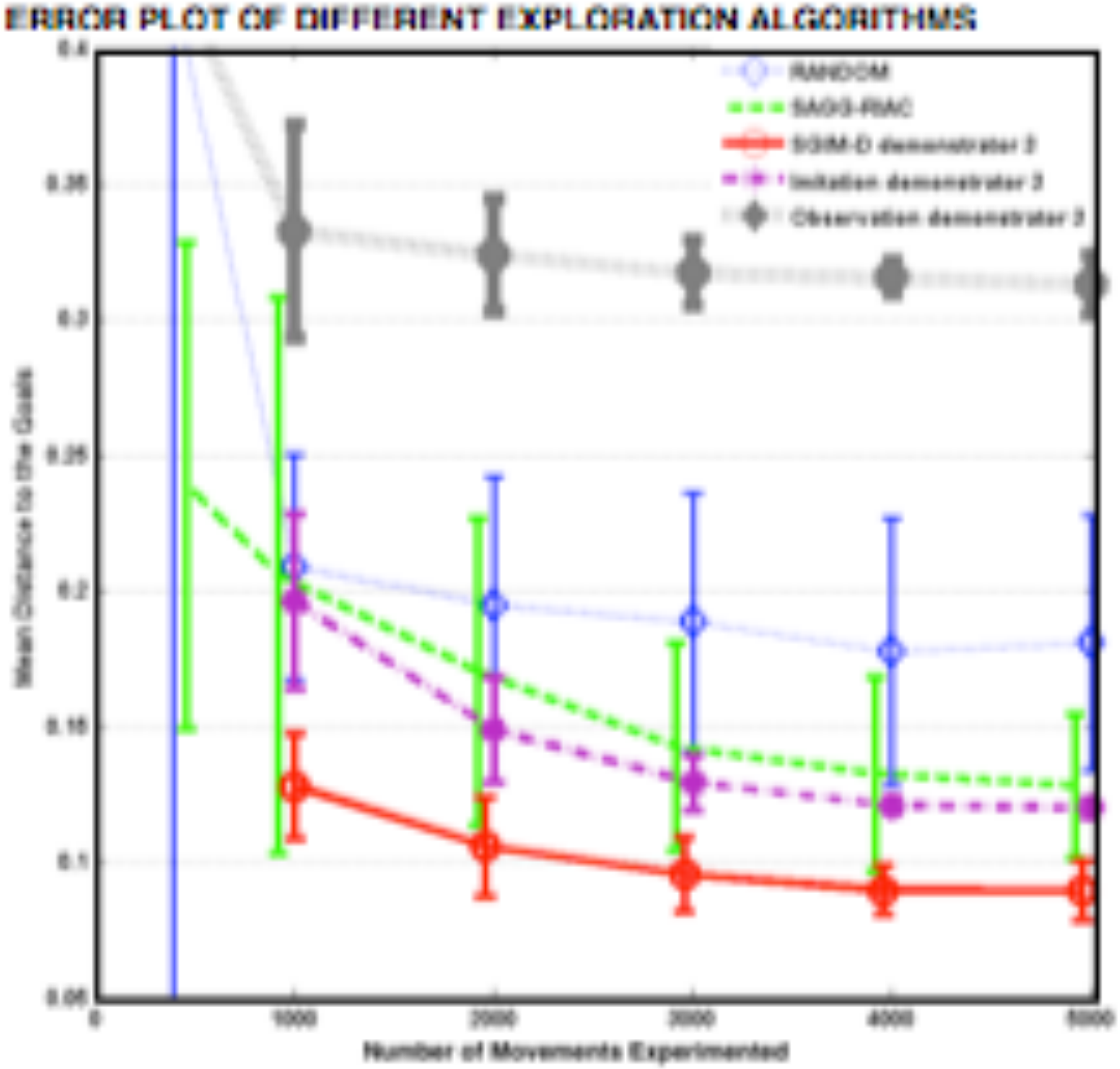
3.3. Combination of several learning strategies - Example

Experimental Setup



3. BEYOND PROGRAMMING BY DEMONSTRATION

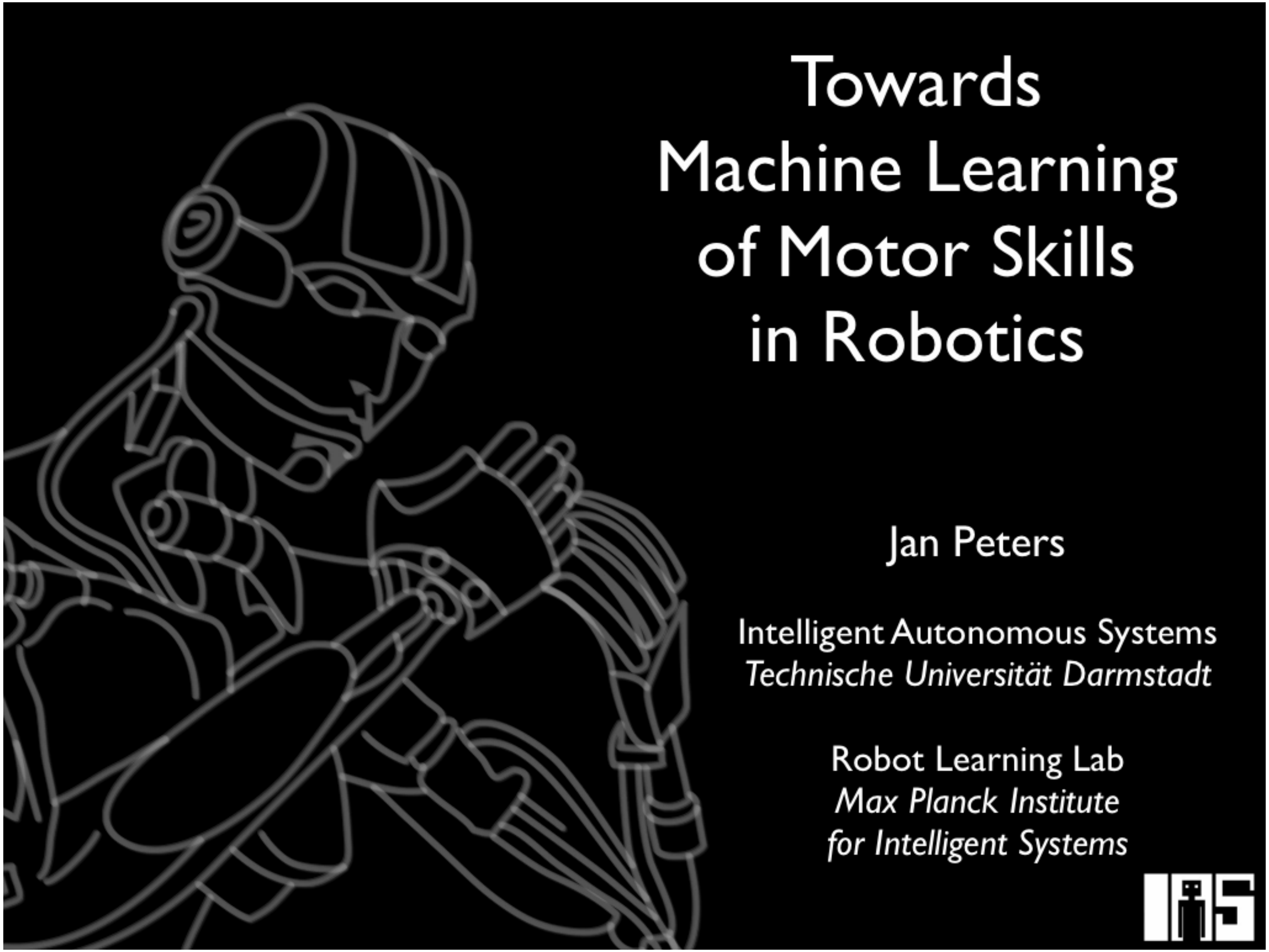
3.3. Combination of several learning strategies - Example



1) SGIM-D learns with **better precision**: error in SGIM-D is lower (t-test $p < 0.05$)
2) SGIM-D learns more **reliably**
3) SGIM-D learns **faster**

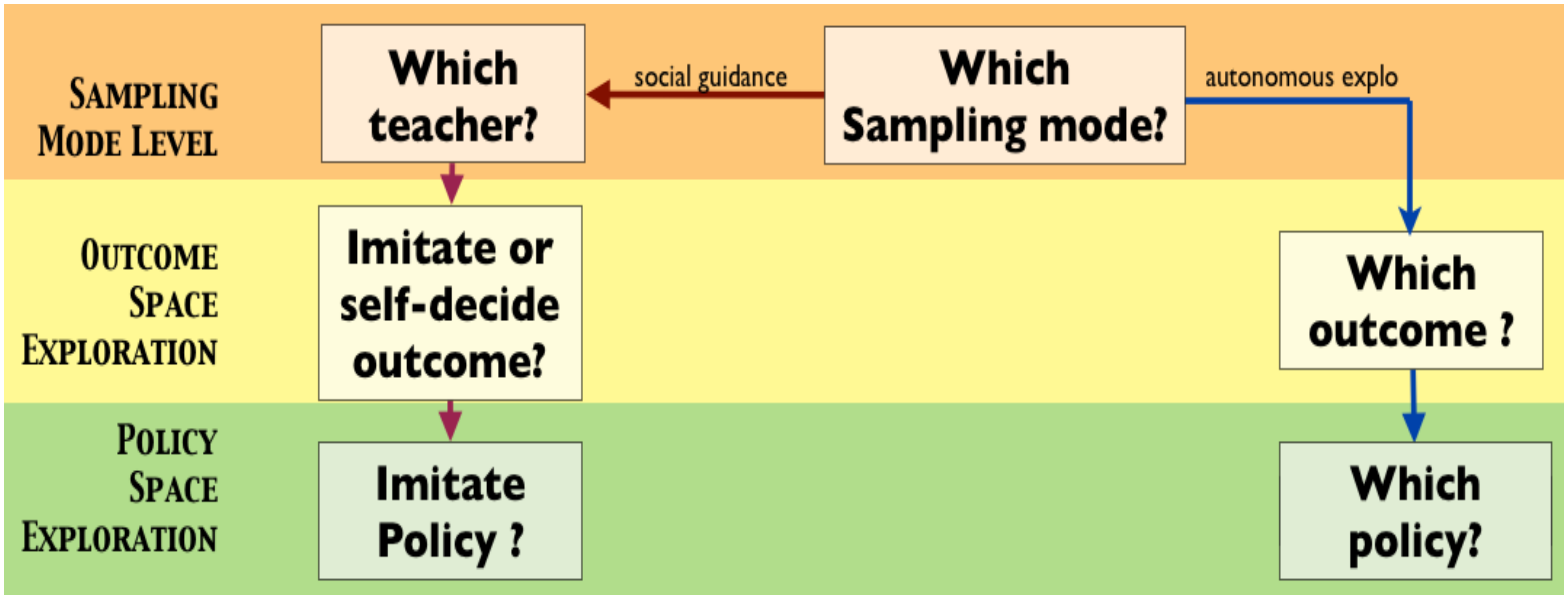
Nguyen and Oudeyer, Autonomous Robots, 2014.

3.3. Combination of several learning strategies - Example



3:13

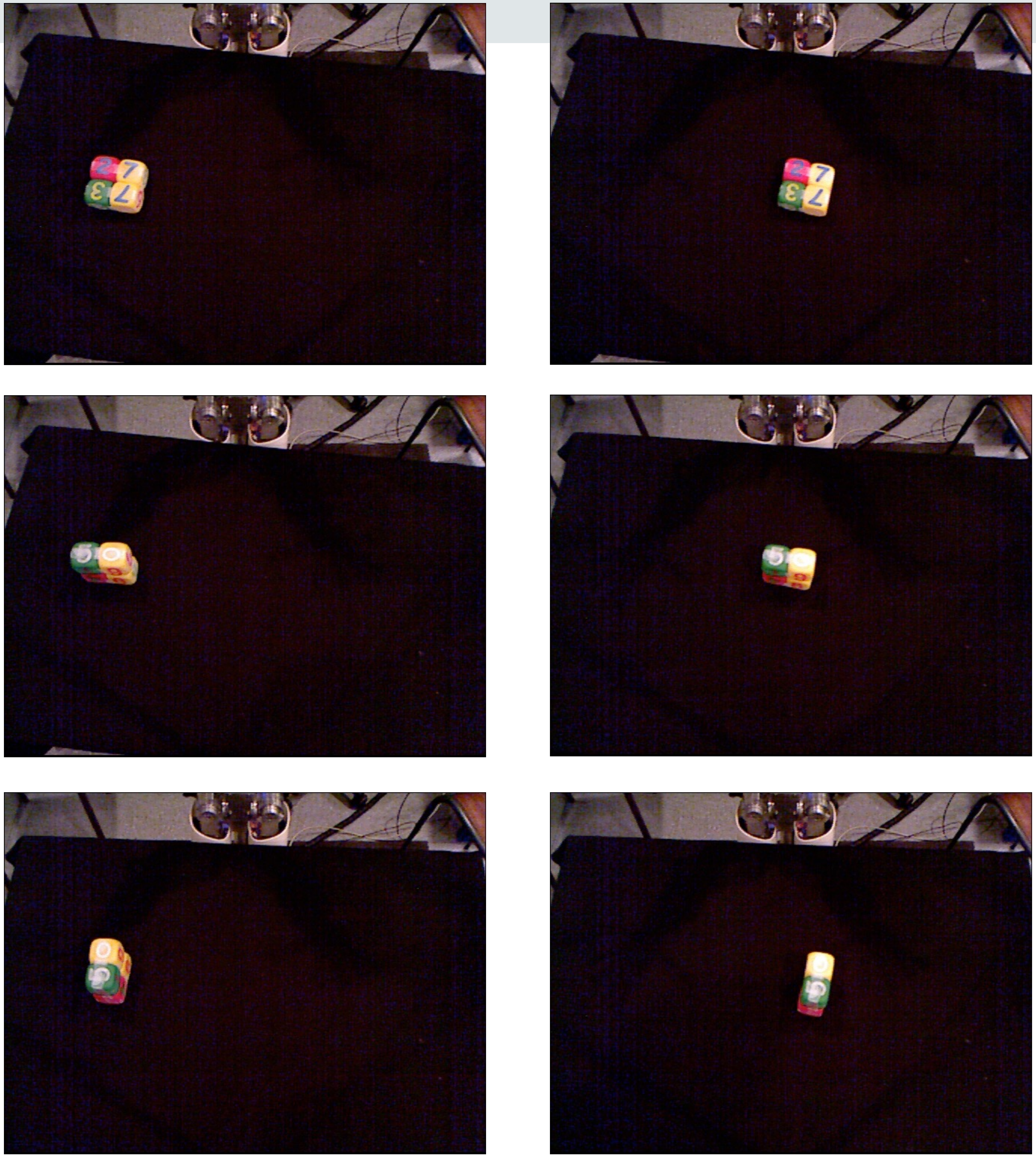
what to imitate, how to imitate, when to imitate and who to imitate



3. BEYOND PROGRAMMING BY DEMONSTRATION

3.5. Example of a strategic learner

How can we learn to recognise 3d objects with all its different views?



ANR MACSi Project in collaboration with ISIR/ENSTA/INRIA
Nguyen et al, TAMD 2013

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

3.4. Example 1

How can we learn to recognise 3d objects with all its different views?

- Manipulation: which **manipulation** will bring you more useful information about the object?
  - push, lift&drop, ask human

ANR MACSi Project in collaboration with ISIR/ENSTA/INRIA
Nguyen et al., TAMD 2013

3. BEYOND PROGRAMMING BY DEMONSTRATION

3.4. Example 1

How can we learn to recognise 3d objects with all its different views?

- Manipulation: which **manipulation** will bring you more useful information about the object?
  - push, lift&drop, ask human
- Several objects: which **object** should you choose to manipulate?

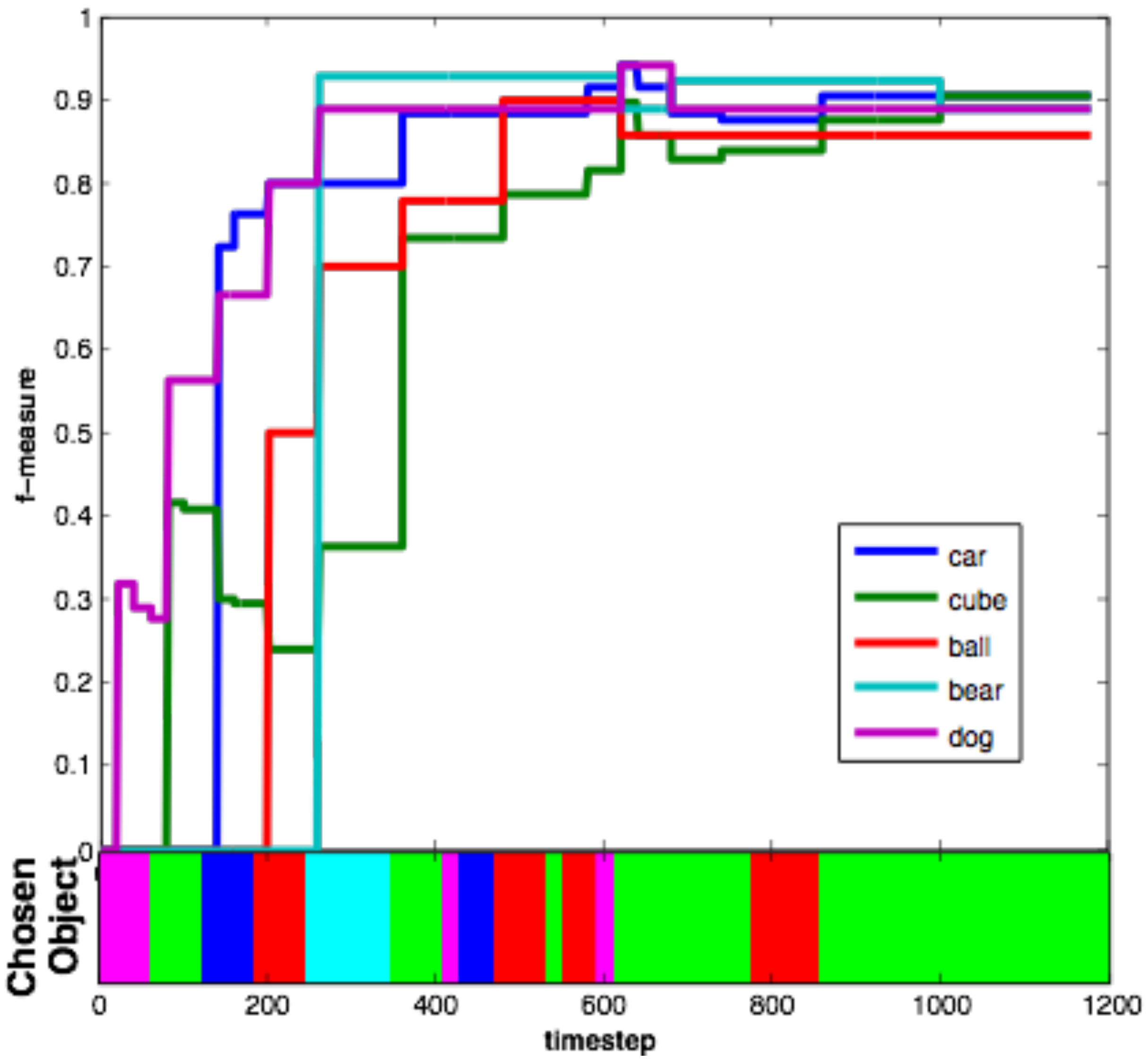
ANR MACSi Project in collaboration with ISIR/ENSTA/INRIA
Nguyen et al, TAMD 2013

3. BEYOND PROGRAMMING BY DEMONSTRATION

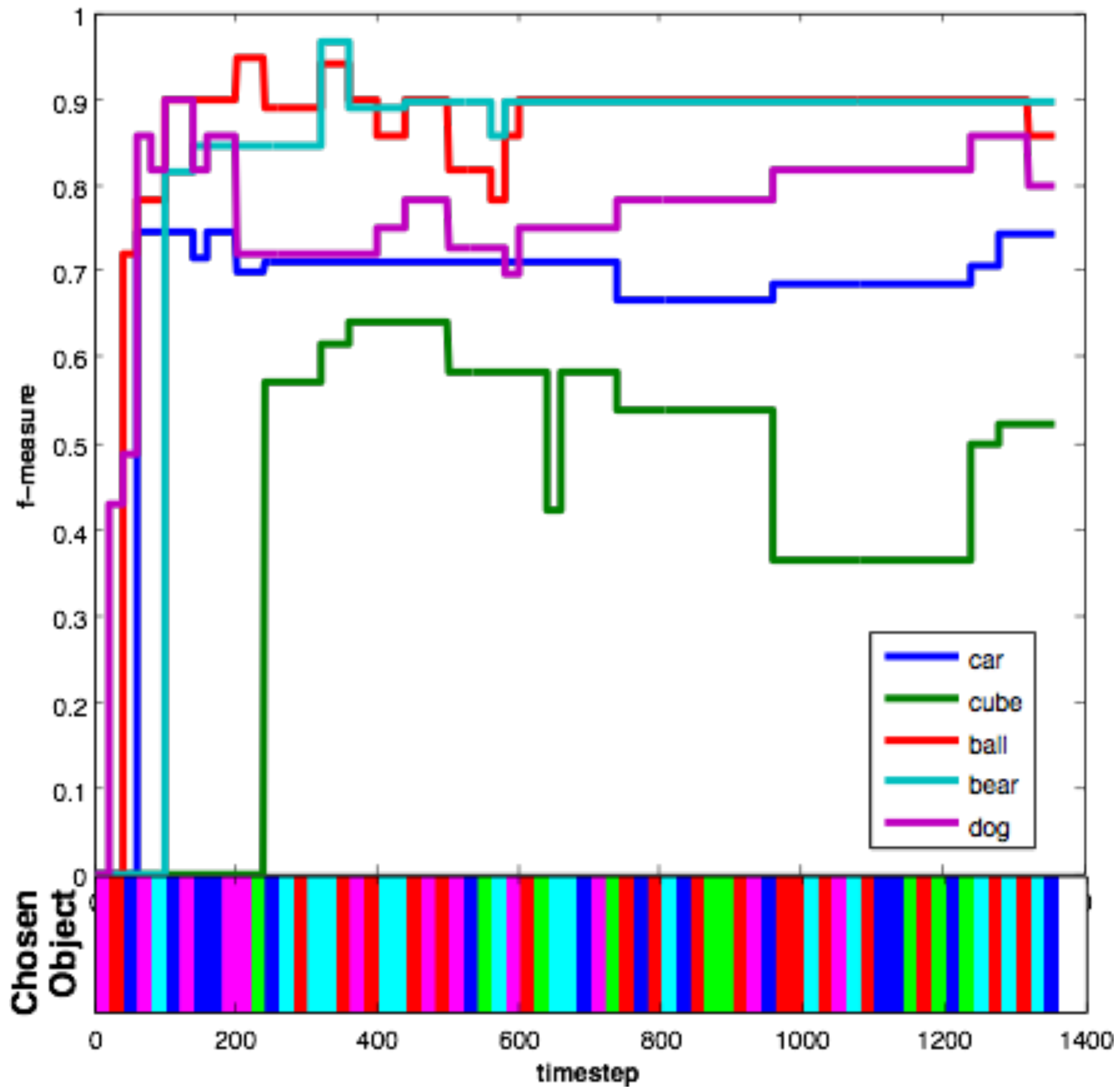
3.4. Example 1

Results 1: with a normal teacher

Recognition level (f-measure)



SGIM-ACTS



Random sampling

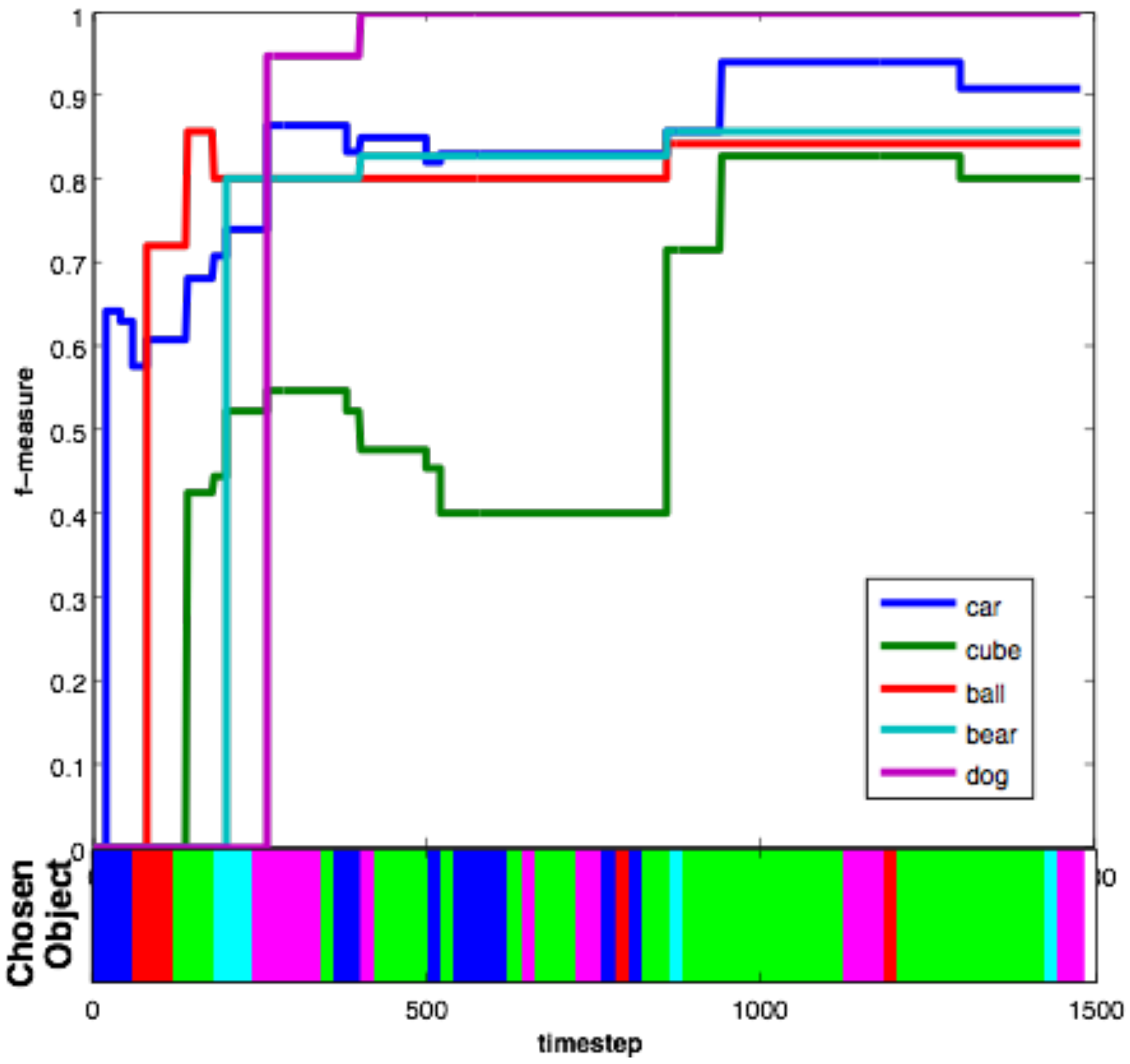
SGIM-ACTS learns better than a Random sampling.

Active data collection improves performance

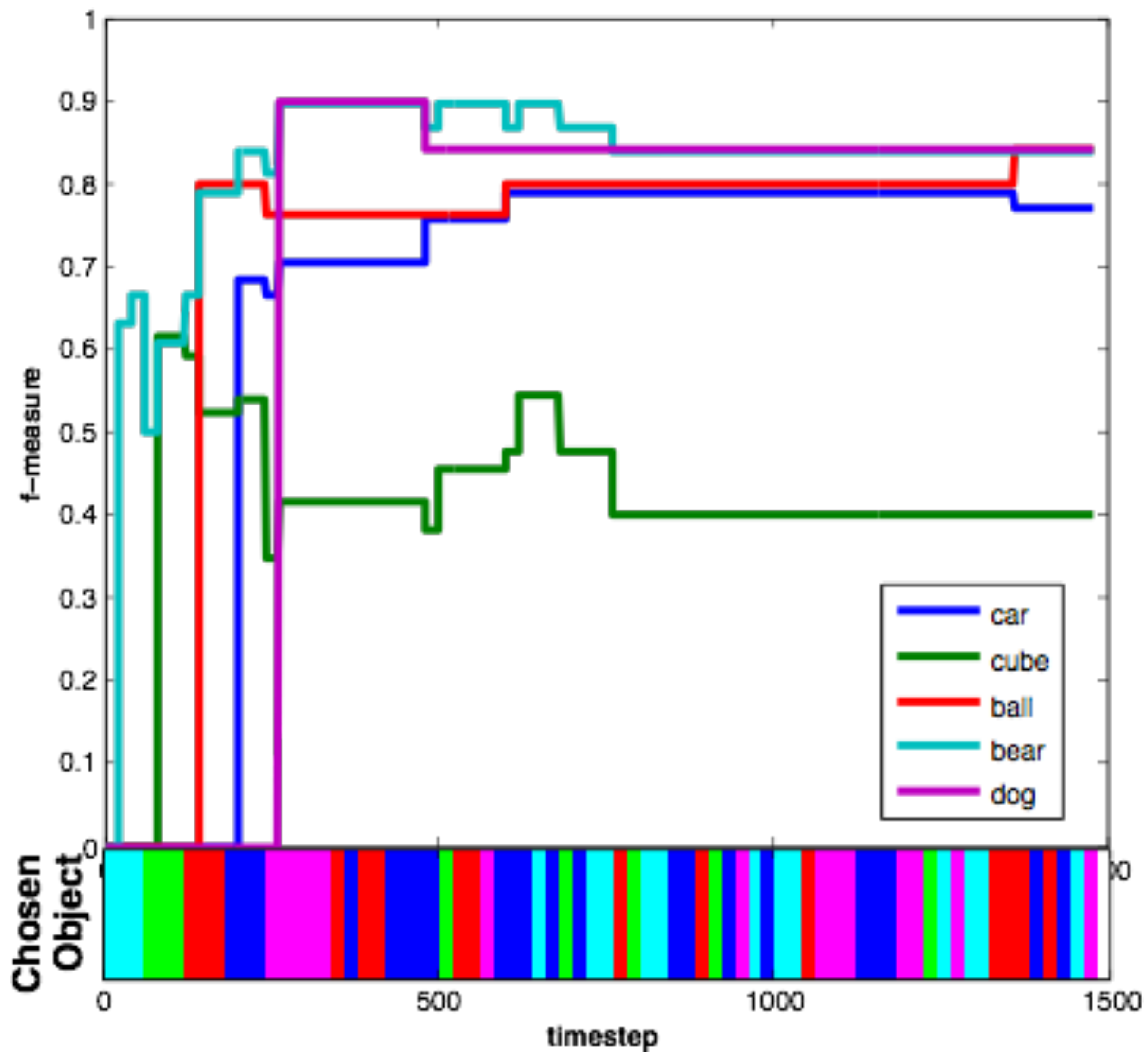
ANR MACSi Project in collaboration with ISIR/ENSTA/INRIA
Nguyen et al, TAMD 2013

Results 2: with a bad teacher

Recognition level (f-measure)



SGIM-ACTS



Random sampling

A bad teacher can affect the performance of a passive learner

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

ANR MACSi Project in collaboration with ISIR/ENSTA/INRIA
Nguyen et al, TAMD 2013

- Aude Billard, Sylvain Calinon, Rüdiger Dillmann, Stefan Schaal, Ch 59 Robot Programming by Demonstration in : Siciliano, Bruno, and Oussama Khatib, eds. *Springer handbook of robotics*. Springer, 2016.
- S. Calinon, A. Billard: What is the Teacher's Role in Robot Programming by Demonstration? - Toward Benchmarks for Improved Learning, Interact. Stud. 8(3), 441–464 (2007), Special Issue on Psychological Benchmarks in Human-Robot Interaction
- S. Calinon, F. Guenter, A. Billard: On Learning Representing and Generalizing a Task in a Humanoid Robot, IEEE Trans. Syst. Man Cybernet. 37(2), 286– 298 (2007), Special issue on robot learning by observation, demonstration and imitation
- S. Ivaldi, S. M. Nguyen, N. Lyubova, A. Droniou, V. Padois, D. Filliat, P.-Y. Oudeyer, and O. Sigaud. Object learning through active exploration. Transactions on Autonomous Mental Development, PP(99):1–1, 2013.
- S. M. Nguyen and P.-Y. Oudeyer. Socially guided intrinsic motivation for robot learning of motor skills. Autonomous Robots, 36(3):273–294, 2014.