# MACHINE LEARNING FOR DEVELOPMENTAL ROBOTICS

**Manuel Lopes, Luis Montesano and Francisco S. Melo**
Institute for Systems and Robotics – Instituto Superior Técnico – Lisboa, Portugal
http://vislab.isr.ist.utl.pt

## Goal:

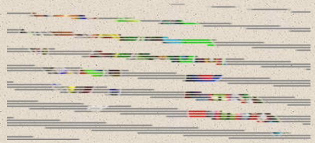
- **Program highly complex robotic systems**

Multimodal sensors: cameras, inertial, touch, sound
Actuation: >40 dof

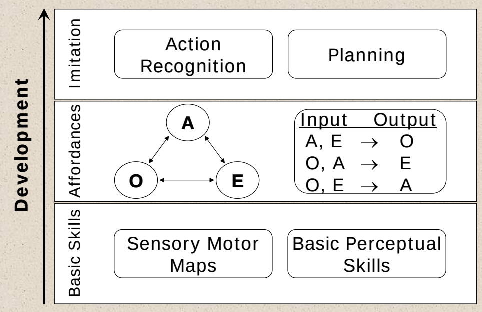
- **Create interactive artificial system**

Able to deal with unexpected situations, interact with people, learn from people

Learning & Cognition



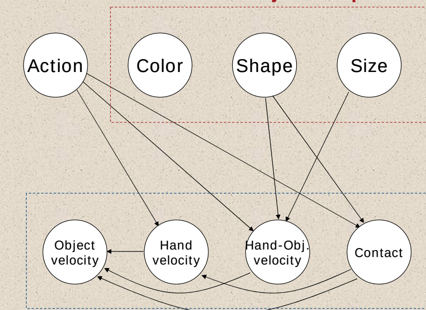
## Developmental framework:




**Copy the natural development of humans:**
Starting from a set of basic skills, construct new ones, reducing the complexity of the overall task

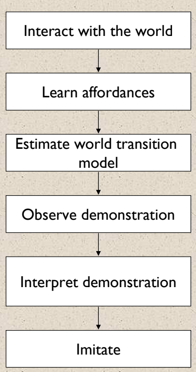
## Affordance learning:

- *Affordances* define the relation of the robot with the surrounding environment in terms of its *motor and perceptual capabilities*. They play a crucial role in core cognitive capabilities such as *prediction* and *planning*:
  - Interpretation of observed actions in terms of its own actions (body-correspondence)
  - Estimation of a dynamic model for the world
- World entities described as actions, object features and observed effects. Descriptors of objects and effects learnt by *unsupervised clustering*
- *Bayesian networks* model Action-Feature-Effect relationships
- The robot autonomously explores the environment and collects a set of data $D$. The posterior $P(G|D)$ over network structures is computed using *MCMC*




## Imitation learning:

### Imitation architecture

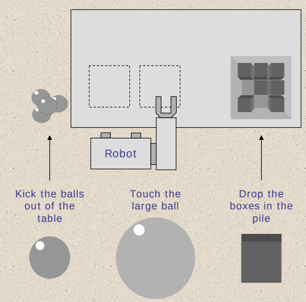


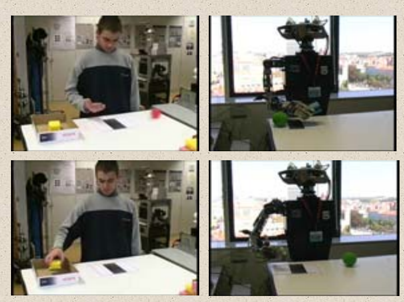

- World described as a controlled Markov chain

$$P_a(i,j) = \mathbb{P}[X_{t+1} = j \mid X_t = i, A_t = a]$$

- *Bayesian inverse RL* computes task description from demo (using MCMC)
- Task description as a *reinforcement function* maximizing likelihood of observed demo
- Imitation policy computed using *DP*

*Here imitation is not a direct replication of the observed actions. The demontrator's goal is inferred by computing a hypothetical reinforcement function to optimize. Task execution is performed idiosyncraticly, using the affordances' knowledge*

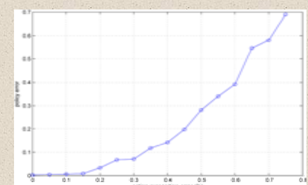


### Task example:

- The goal is to separate different types of objects in to different containers
- The demonstration does not visits all states and is inconsistent
- The system is able to achieve the goal of the task

### Robustness to errors & incomplete demos:



## Future work:

- Generalizing robot-object interaction knowledge requires taking into account groups of objects, sequences of actions and delayed effects
- Active learning strategies should be implemented to deal with huge search spaces
- Interaction between the different learning processes, e.g., the evolution of actions from pure joint positions or velocities to (possibly parameterized) motion primitives
- Evaluate the impact of inaccurate learnt models in successive developmental phases (and their correction)

## References:

- Montesano et al. Learning object affordances: from sensory-motor coordination to imitation. *IEEE Transactions on Robotics*, In press.
- Lopes and Santos-Victor. Developmental Roadmap for Learning by Imitation in Robots. *IEEE Transactions on Syst. Men & Cyber. (B)*, 2007
- Ramachandran and Amir. Bayesian inverse reinforcement learning. *Proc. IJCAI*, 2007.

*Partially supported by the FCT POS_C program that includes FEDER funds and by the EU Projects (IST-004370) RobotCub and (EU-FP6-NEST-5010) Contact and national project PTDC/EEAACR/70174/2006.*