The European projec GOAL-Robots aiming to build robots that can learn motor skills in an open-ended fashion driven by curiosity

Gianluca Baldassarre

Laboratory of Computational Embodied Neuroscience (LOCEN),
Institute of Cognitive Sciences and Technologies (ISTC),
Italian National Research Council (CNR),
Rome, Italy

Shangai Lecture

23 November 2016

Outline

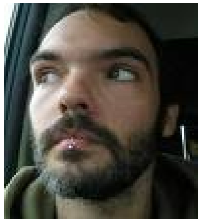
- LOCEN: hint to research methods
- New project “GOAL-Robots” funding this research
- Our problem: “robots able to autonomously learn multiple goals/skills”
- Solution: “intrinsic motivations → goals → skill learning”
- Two robotic models examples
  - Learning parameterised skills
  - A whole architecture of open-ended learning

Senior collaborators of this research

CNR: Italian National Research Council
ISTC: Institute of Cognitive Sciences and Technologies
LOCEN: Laboratory of Computational Embodied Neuroscience

Francesco Mannella, Postdoc,
BA/MA Cognitive Science,
PhD: Modelling Embodied Intelligence

Daniele Caligiore, Researcher,
BA/MA Robotics,
PhD: Bioengineering



Valerio Sperati, PhD student,
BA/MA Cognitive Science,
PhD: Computer Science, AI

Vieri Santucci, Postdoc,
BA/MA Philosophy,
PhD: Computer Science, AI

Emilio Cartoni, PhD student,
BA/MA Neuroscience,
PhD: Bayesian Models

Simona Bosco,
BA/MA Biology,
Admin, projects, e-knowledge

Key elements of our research method

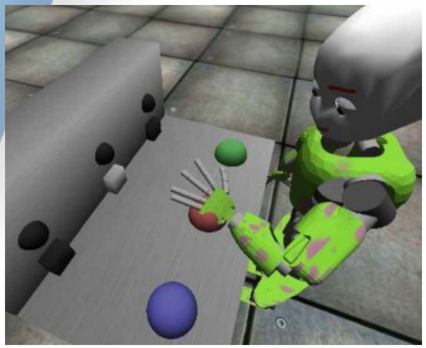
...and behaviour consequences affect brain learning

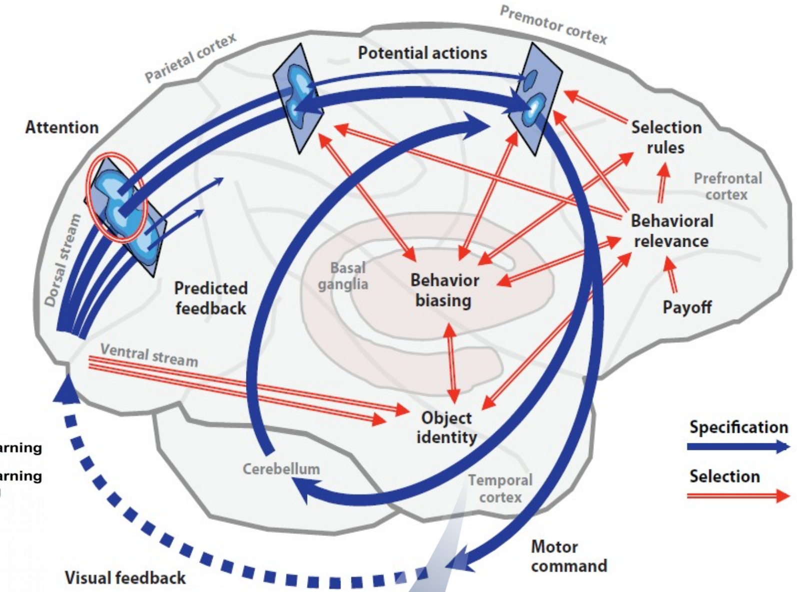
Brain...

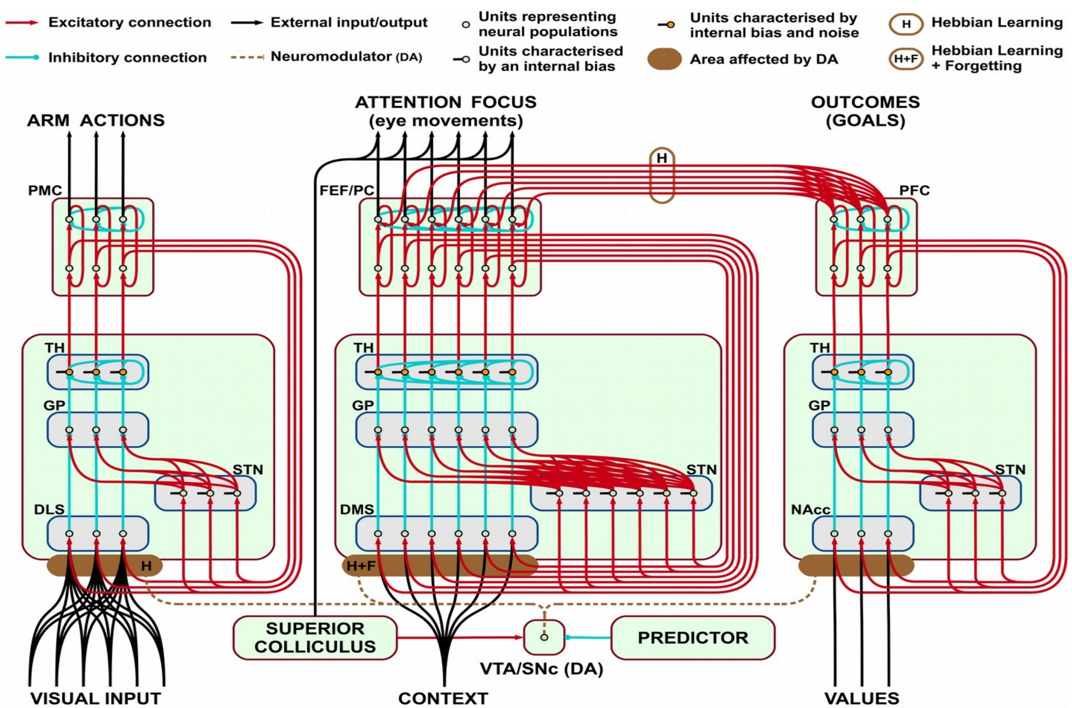
...produces behaviour in the world...

LOCEN: Laboratory Of Computational Embodied Neuroscience

Hint to our bio-constrained models...







System computational neuroscience!

Baldassarre et al. (2013). *Neural Networks*

Key elements of our research method

Today's focus!

...and behaviour consequences affect learning

Brain...

LOCEN

...produces behaviour in the world...

LOCEN: Laboratory Of Computational Embodied Neuroscience

A new EU project on the issues I will discuss

11/2016-10/2020 FET-OPEN project **GOAL-Robots**:
“Goal-based Open-ended Autonomous Learning Robots”

- FET-OPEN call April 2016: **1st of 11 funded projects out of 800 :)**
- **3.5 million euros**
- Four Principal Investigators/Partners:
    1. **Gianluca Baldassarre**, Italian National Research Council, Rome, Italy
    2. **Kevin O'Regan**, Université Paris Descartes, Paris, France
    3. **Jochen Triesh**, Frankfurt Institute for Advanced Studies, Germany
    4. **Jan Peters**, Technische Universitaet Darmstadt, Darmstadt, Germany

GOAL Robots

Info: cordis.europa.eu/project/rcn/203543_en.html

Soon: www.goal-robots.eu

Overall aim of the project and our research

- **Understanding how humans and robots**
can cumulatively acquire multiple sensorimotor skills
by autonomously interacting with the environment.

Overall aim of the project and our research

- **Understanding how humans and robots can cumulatively acquire multiple sensorimotor skills** by autonomously interacting with the environment.
- **Scientifically important to understand human behaviour and development**: discovery of multiple goals and skills related to body, objects, multiple objects, and complex objects interactions

Overall aim of the project and our research

- **Understanding how humans and robots can cumulatively acquire multiple sensorimotor skills** by autonomously interacting with the environment.
- **Technologically important to build future robots**

Autonomous learning in unstructured environment

Overall aim of the project and our research

- **Understanding how humans and robots can cumulatively acquire multiple sensorimotor skills** by autonomously interacting with the environment.
- **Technologically important to build future robots**

Autonomous learning in unstructured environment

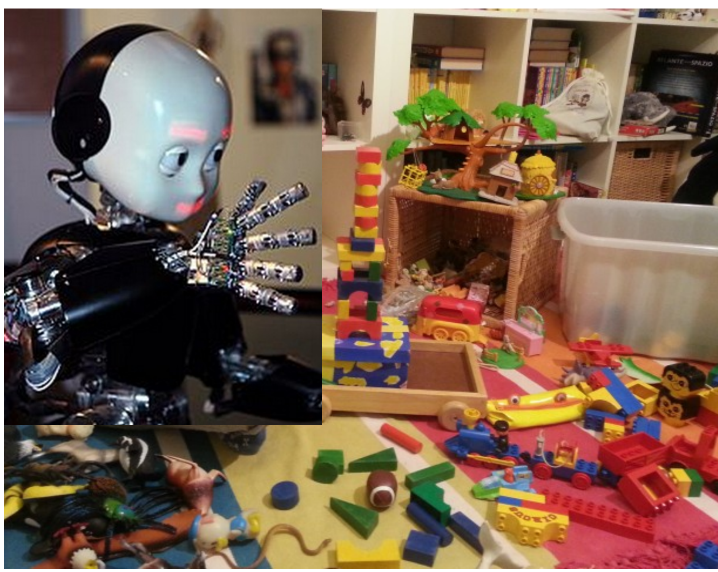
Unstructured environments pose challenges unexpected at design time

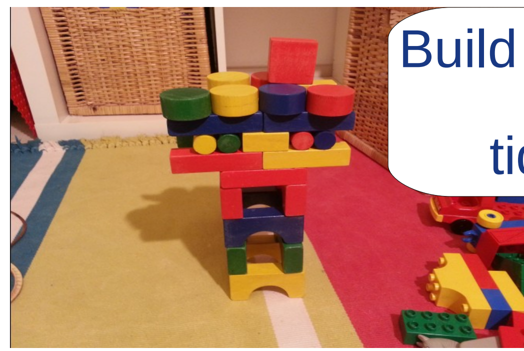
Therefore not possible to pre-program or pre-train them!

Overall aim of the project and our research

- **Understanding how humans and robots can cumulatively acquire multiple sensorimotor skills** by autonomously interacting with the environment.
- **Technologically important to build future robots**

Autonomous learning in unstructured environment





Build structure!
or
tidy-up!

State-of-the-art of Developmental Robotics

For example:

- **IM-CleVeR (old project)**
- Schmidhuber
- Barto
- Oudeyer
- Merrick
- Baldassarre
- ...see ICDL proceedings...

Intrinsic motivations → Skills learning

State-of-the-art of Developmental Robotics

But: developmental robotics still fails to produce truly open-ended learning

~~Intrinsic motivations~~ → ~~Skills learning~~

Insight from psychology/biology/models

Simon Newell (…, 1972, ...)
Castelfranchi Parisi (1976, ...)
Hommel (…, 2001, ...)
von Hofsten (…, 2004, …)
...

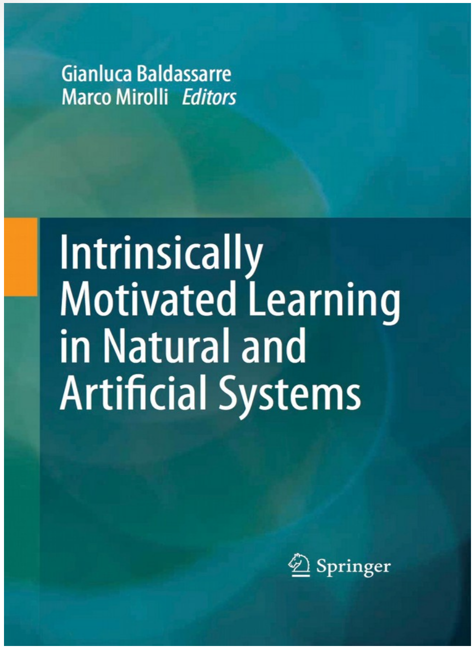
Fuster (…, 1997, …)
Dickinson Balleine (1998, ...)
Passingham Wise (…, 2012, …)
…
Santucci, Mirolli, Baldassarre (2012, ...)
Rolf (2010, ...)
Oudeyer (2010, ...)

Intrinsic motivations → **Self-generated goals** → Skills learning

Paradigm shift of GOAL Robots

Intrinsic motivations → **Self-generated goals** → Skills learning

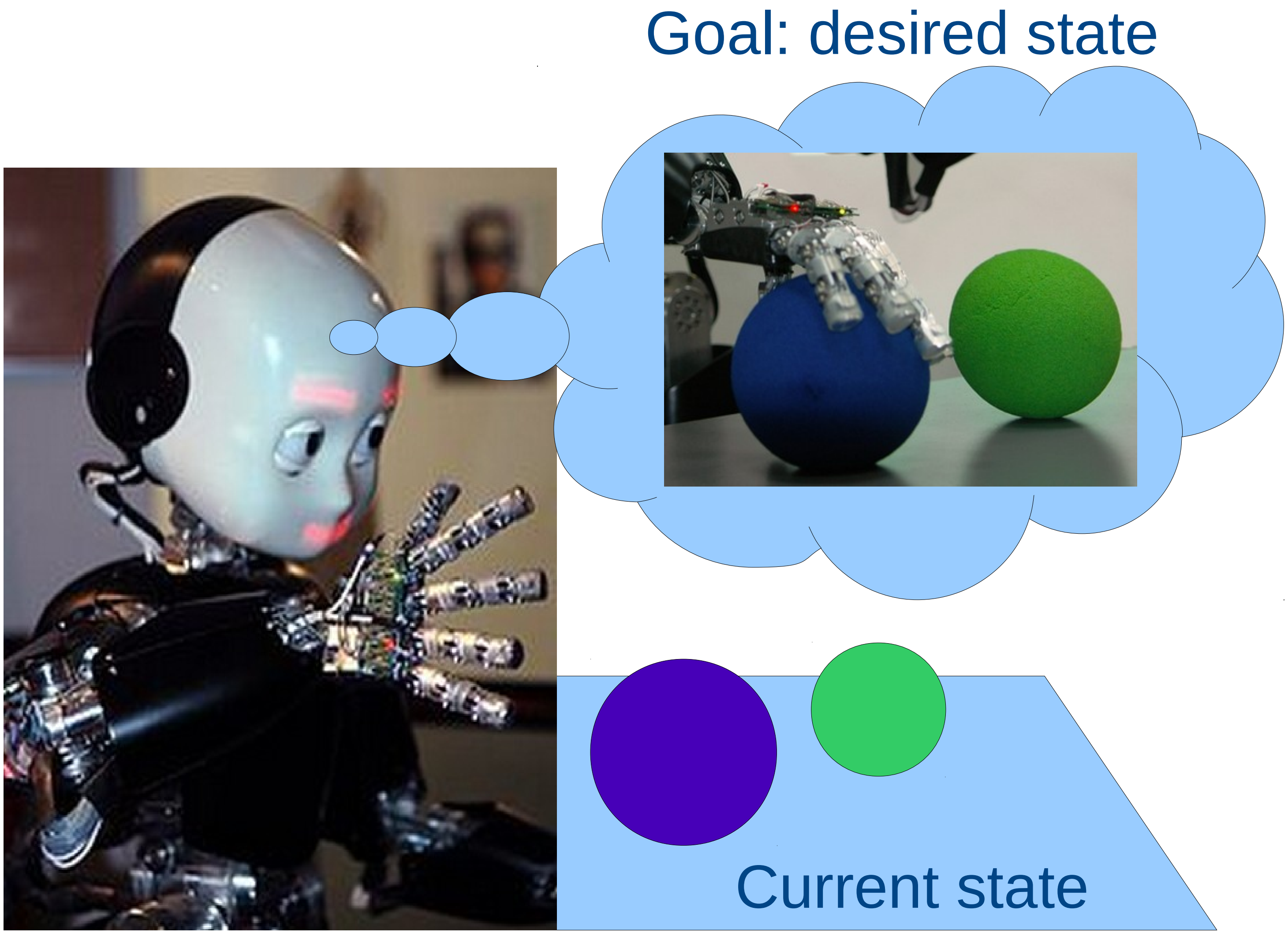
Interim zoom: different intrinsic motivations (mechanisms)



From Baldassarre & Mirolli (2013):

- **Surprise** (e.g., Shmidhuber, 1990; Oudeyer, 2007)
  Based on: violation of predictions
  Measured as: error/rate of improvement of predictions
- **Novelty detection** (e.g., Nehmzow, 2000)
  Based on: lack of information in memory
  Measured as: quality/rate of improvement of memories
- **Competence acquisition** (e.g., Barto, 2004; Schembri, 2007)
  Based on: performance to accomplish a task/goal
  Measured as: probability/probabilty-increase of success

Interim zoom: what is a goal?

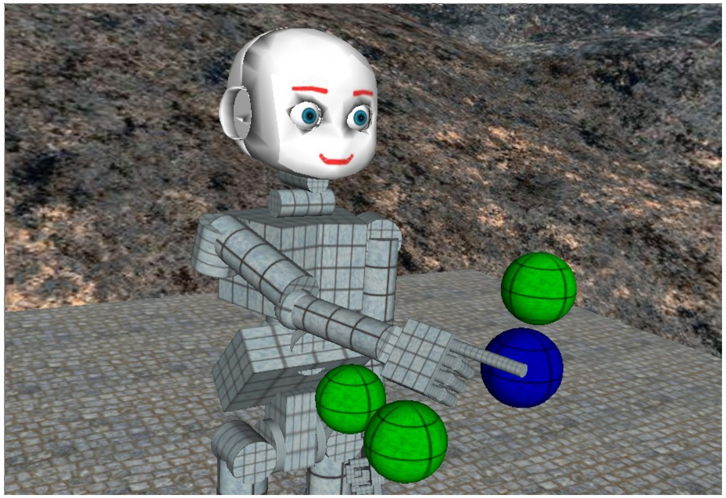


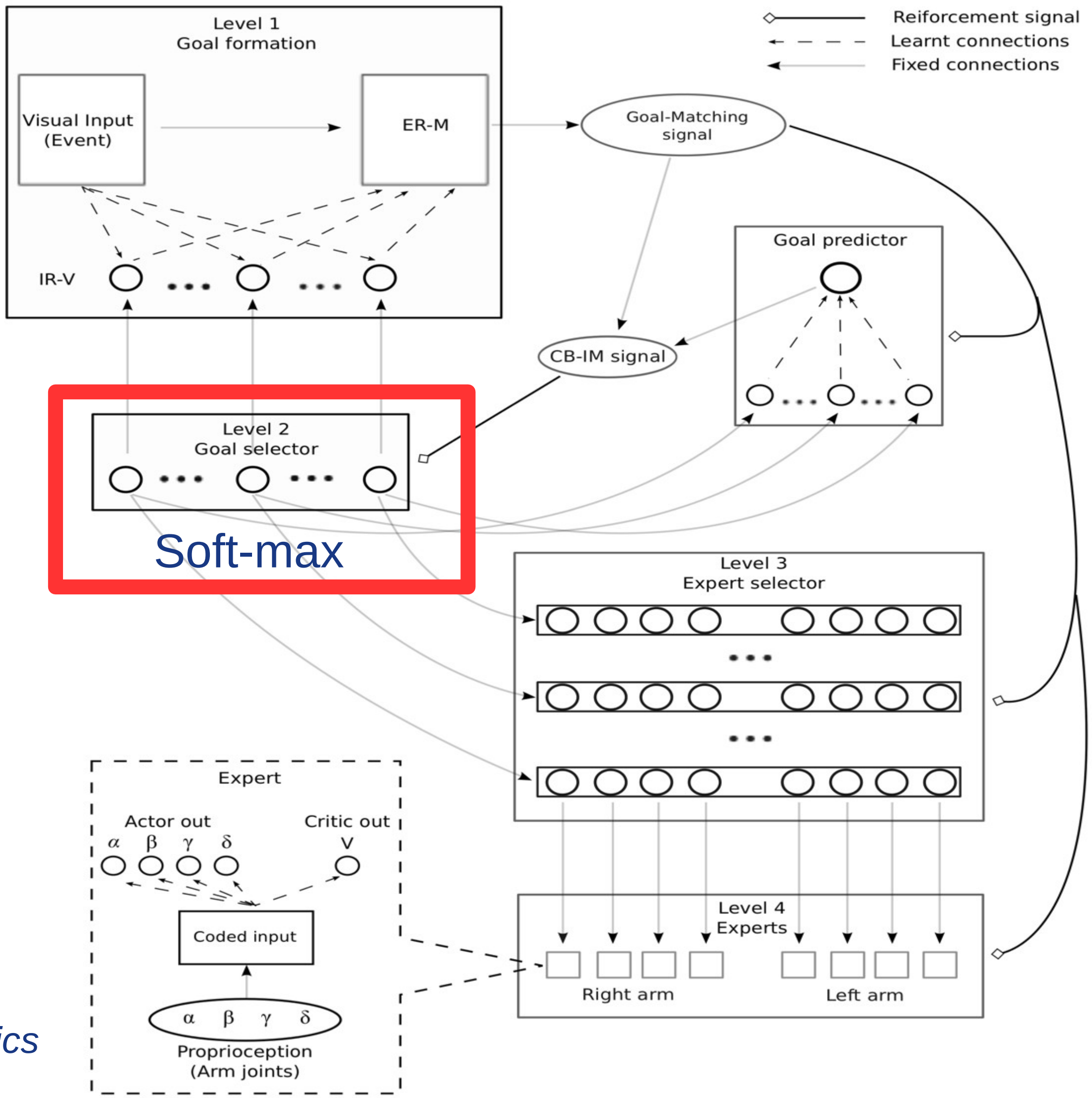
Paradigm shift: → GOAL Robots

Intrinsic motivations → **Self-generated goals** → Skills learning

Model 3: GRAIL: Goal-discovering Robotic Architecture for Intrinsically-motivated Learning

Vieri Santucci

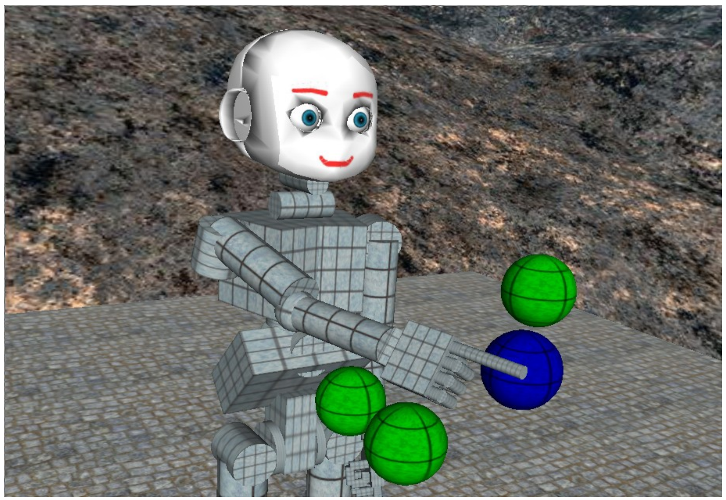


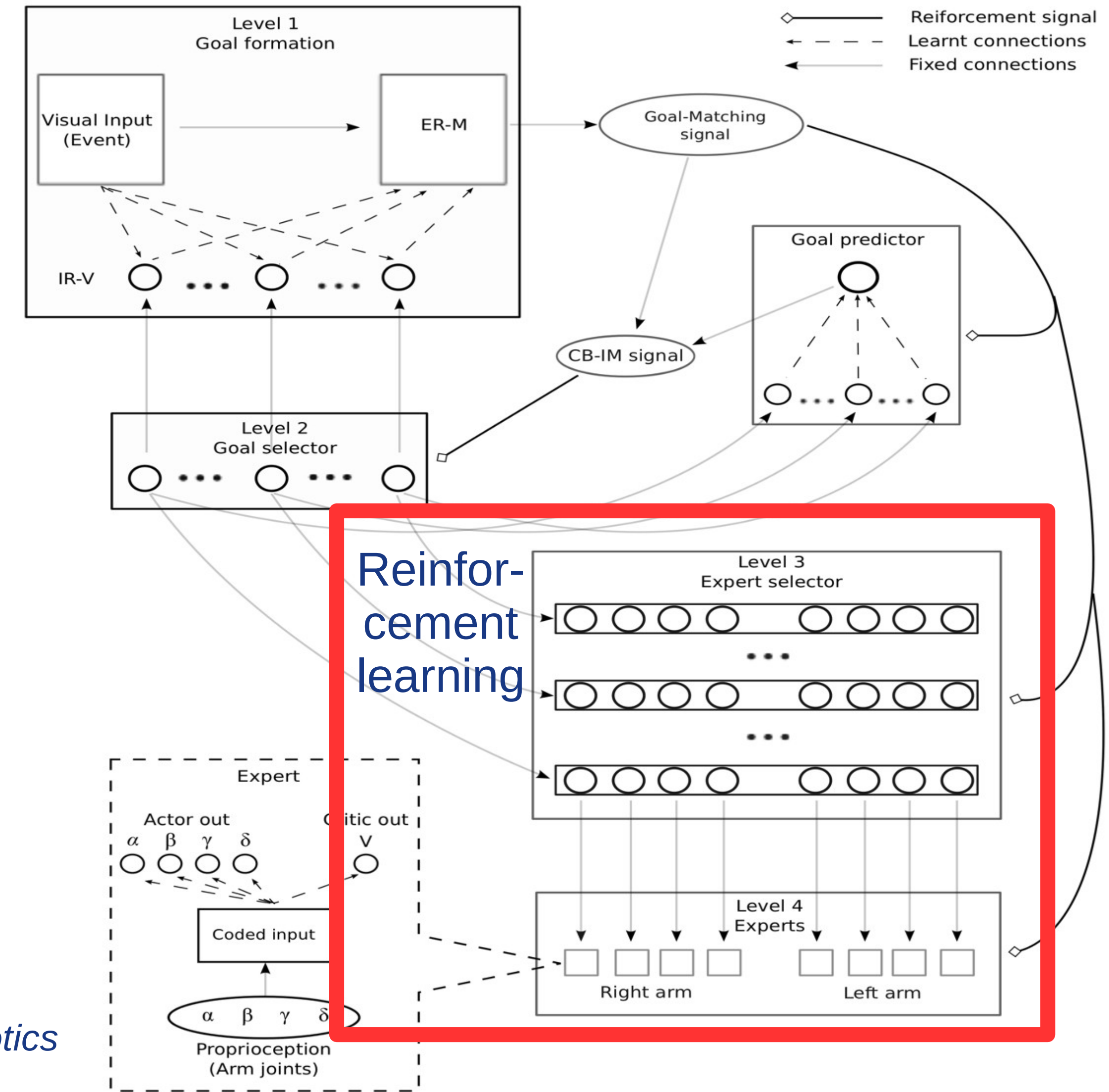


Santucci et al. (2013) *Frontiers in Neurorobotics*
Santucci et al. (2016) *IEEE TAMD*

Model 3: GRAIL: Goal-discovering Robotic Architecture for Intrinsically-motivated Learning

Vieri Santucci

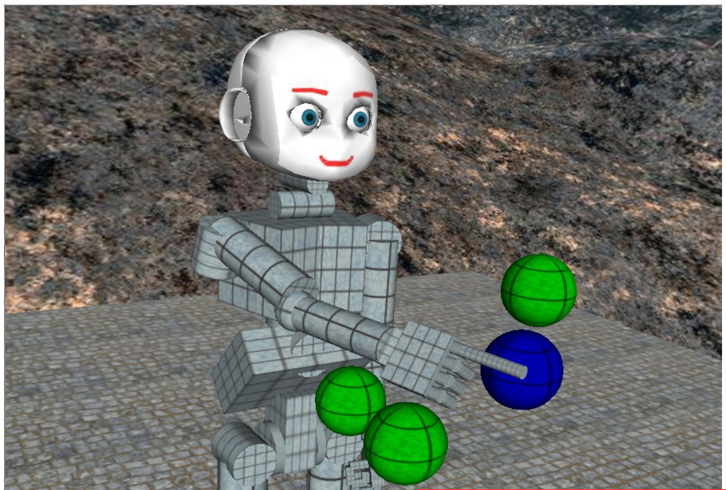


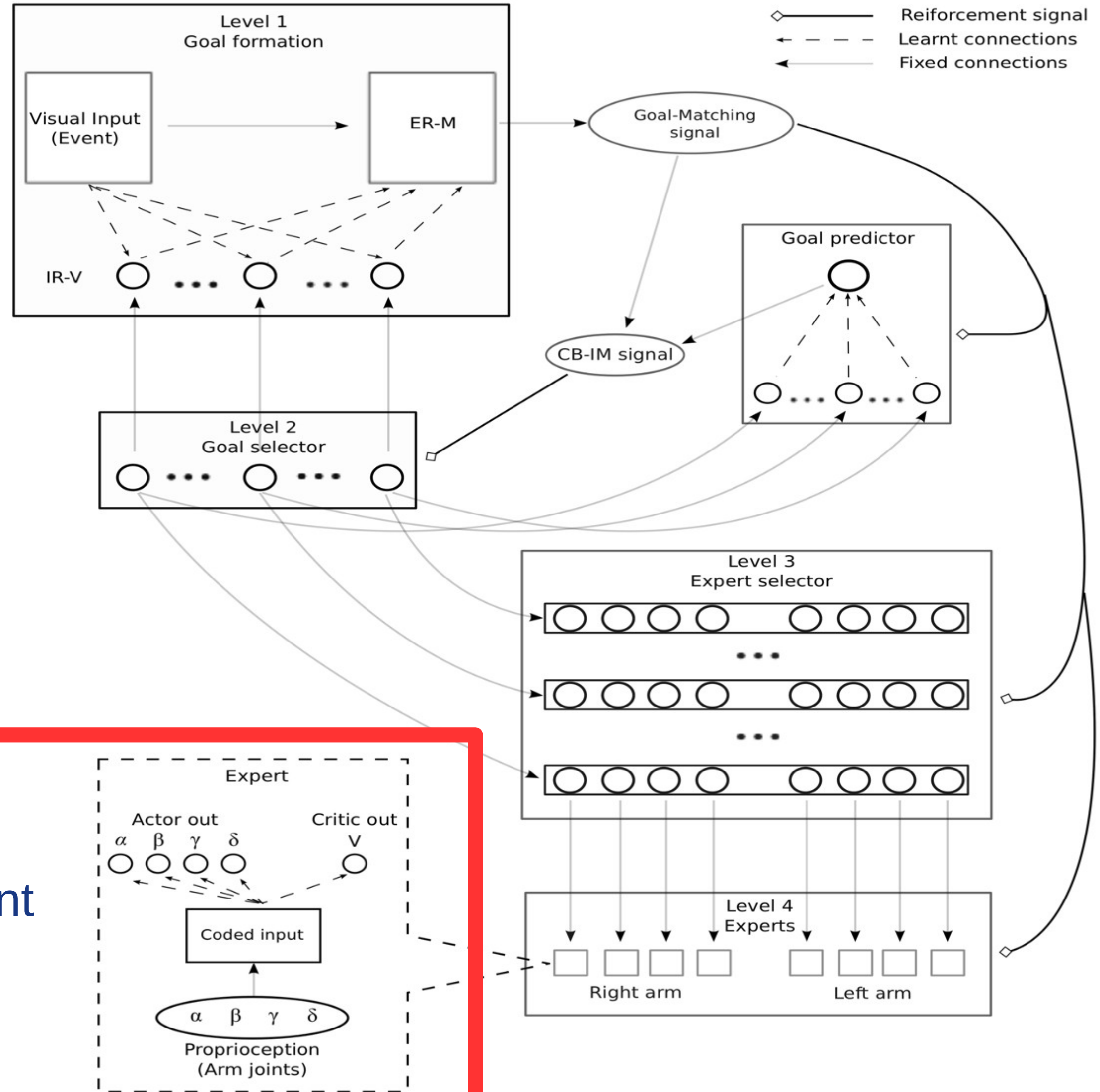


Santucci et al. (2013) *Frontiers in Neurorobotics*
Santucci et al. (2016) *IEEE TAMD*

Model 3: GRAIL: Goal-discovering Robotic Architecture for Intrinsically-motivated Learning

Vieri Santucci

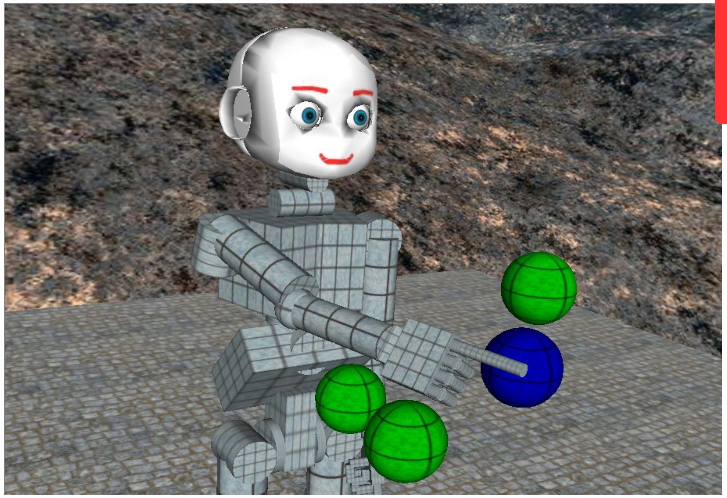


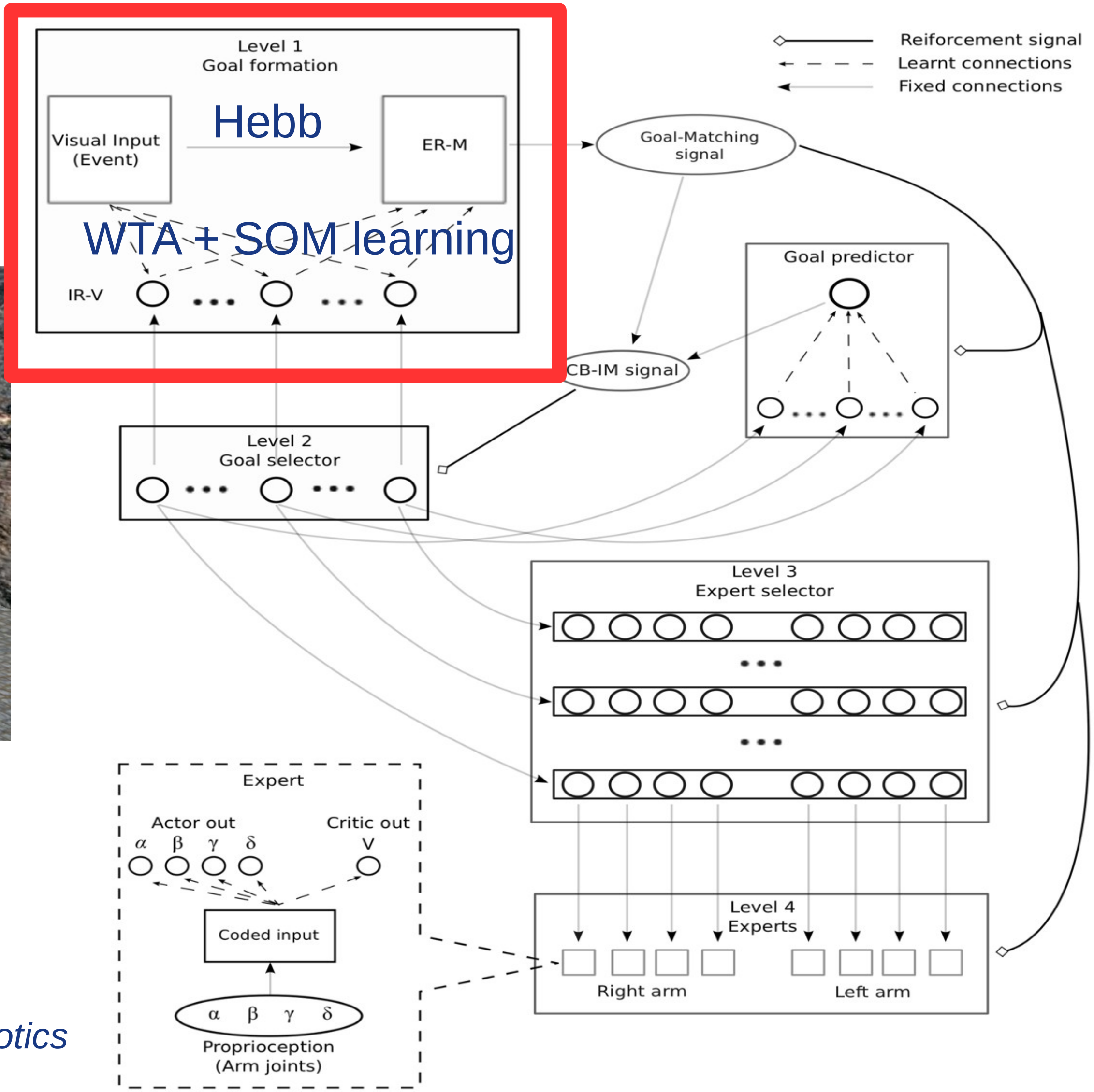


Actor-critic reinforcement learning

Model 3: GRAIL: Goal-discovering Robotic Architecture for Intrinsically-motivated Learning

Vieri Santucci

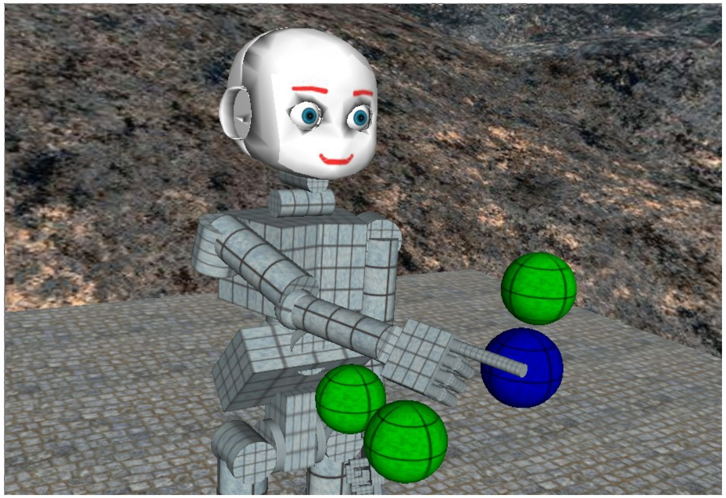


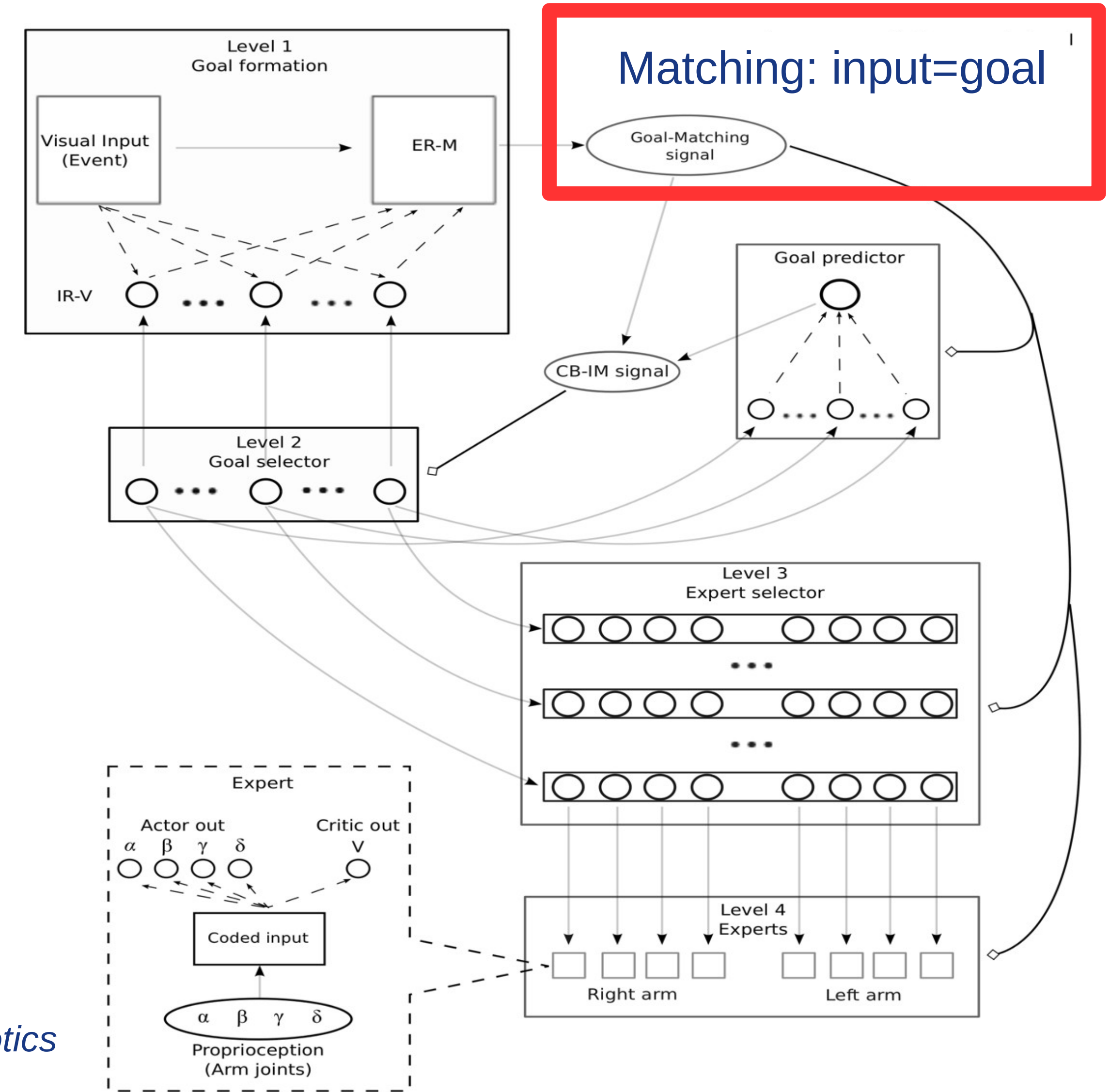


Santucci et al. (2013) *Frontiers in Neurorobotics*
Santucci et al. (2016) *IEEE TAMD*

Model 3: GRAIL: Goal-discovering Robotic Architecture for Intrinsically-motivated Learning

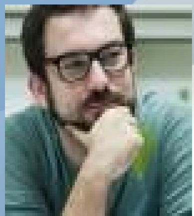
Vieri Santucci



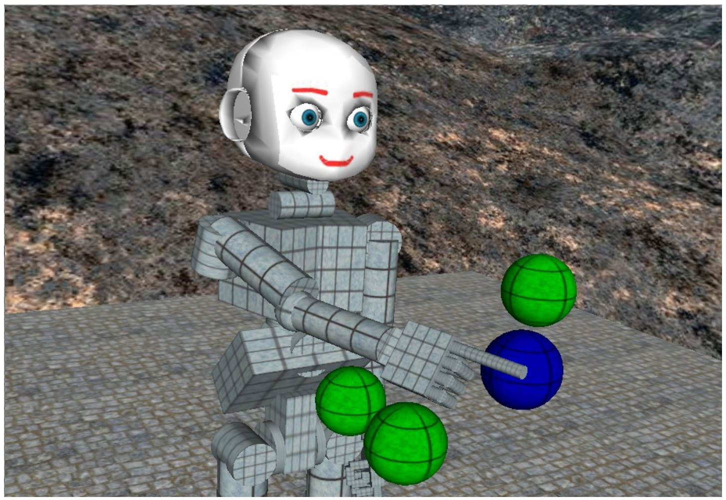


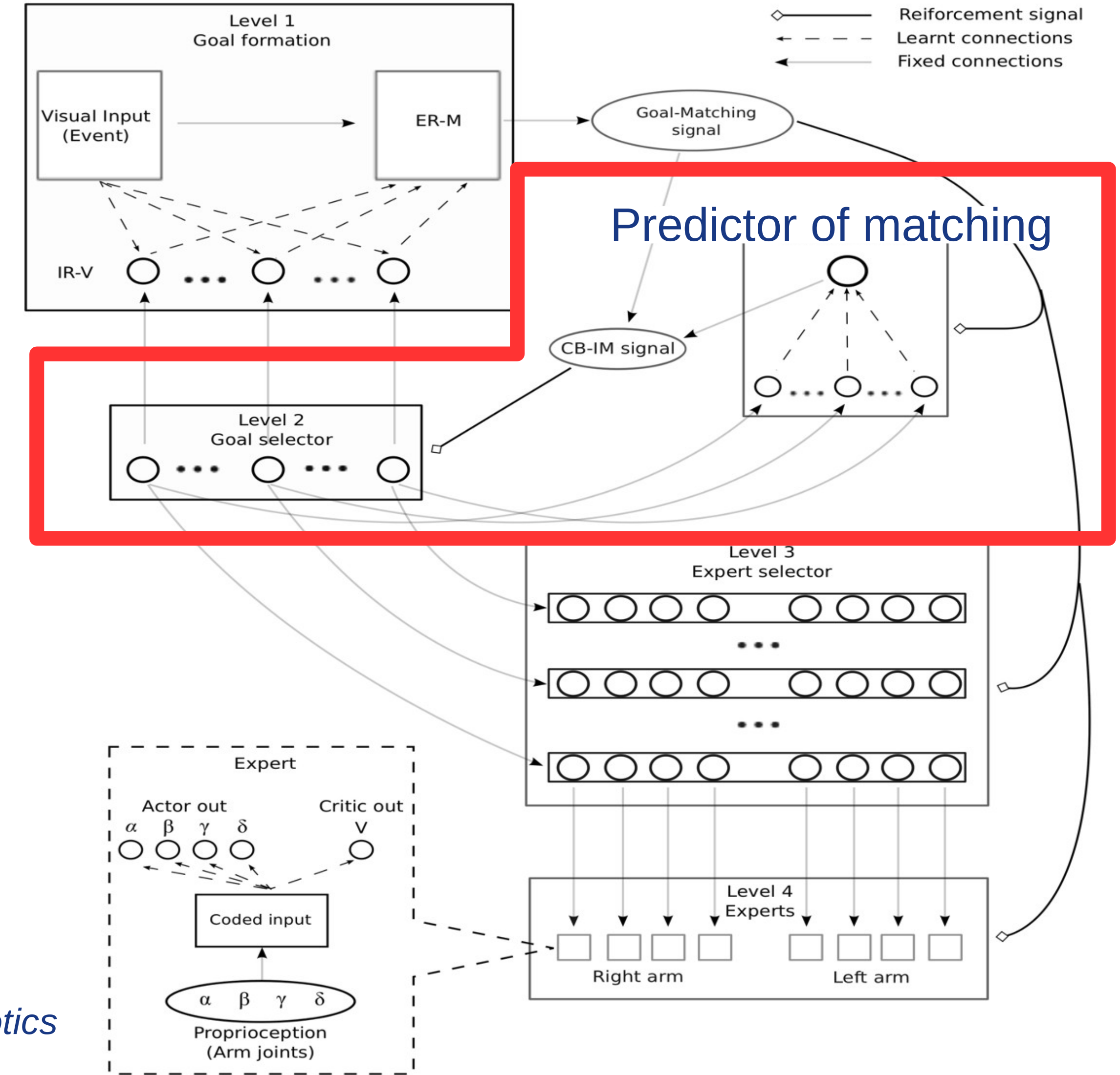
Santucci et al. (2013) *Frontiers in Neurorobotics*
Santucci et al. (2016) *IEEE TAMD*

Model 3: GRAIL: Goal-discovering Robotic Architecture for Intrinsically-motivated Learning



Vieri Santucci

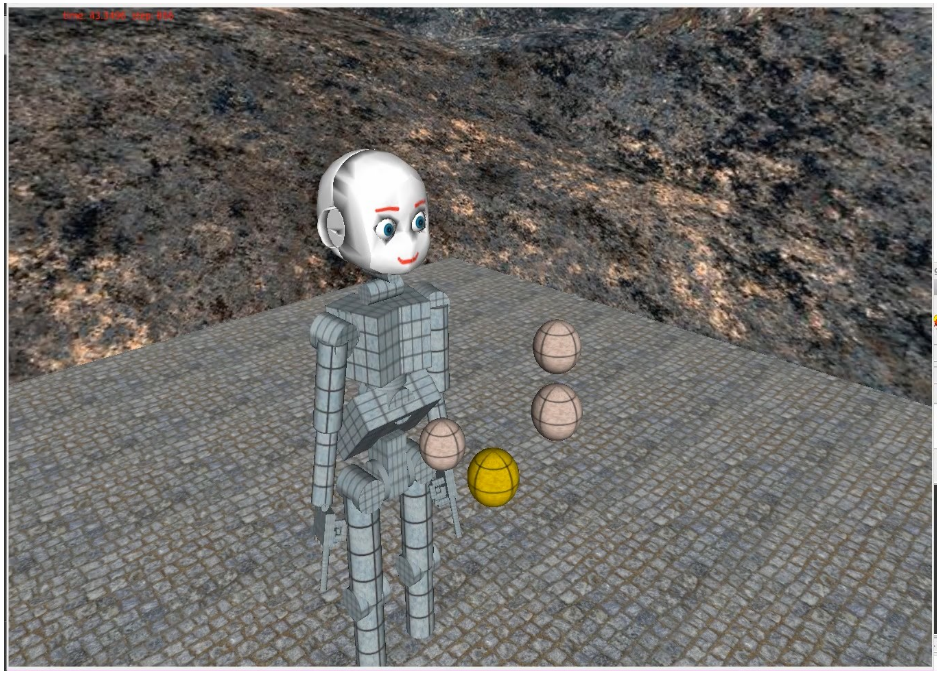




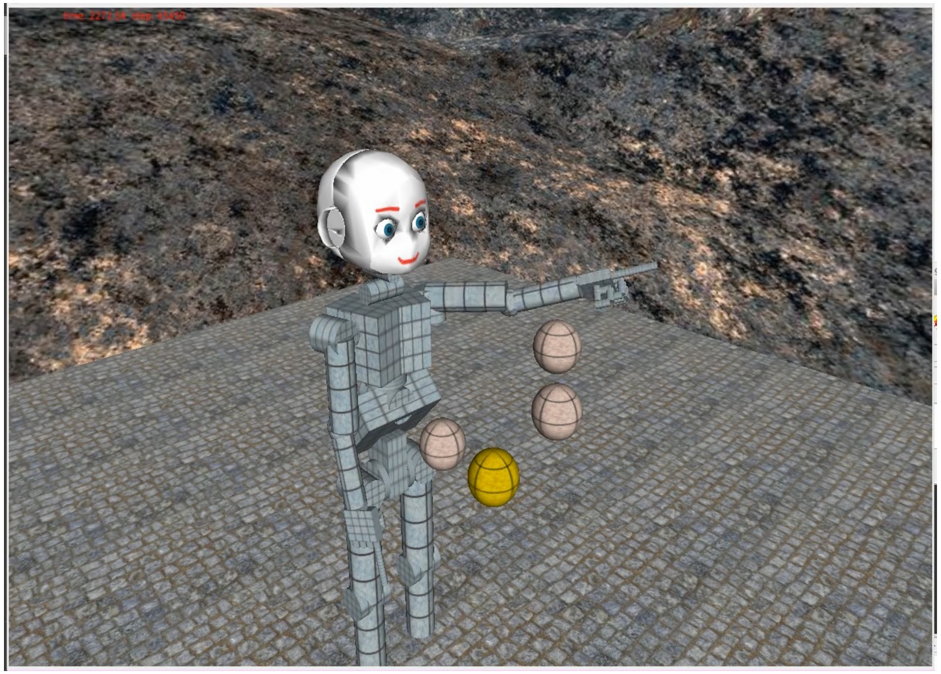
Santucci et al. (2013) *Frontiers in Neurorobotics*
Santucci et al. (2016) *IEEE TAMD*

Model 3: Video of GRAIL learning and functioning

While learning

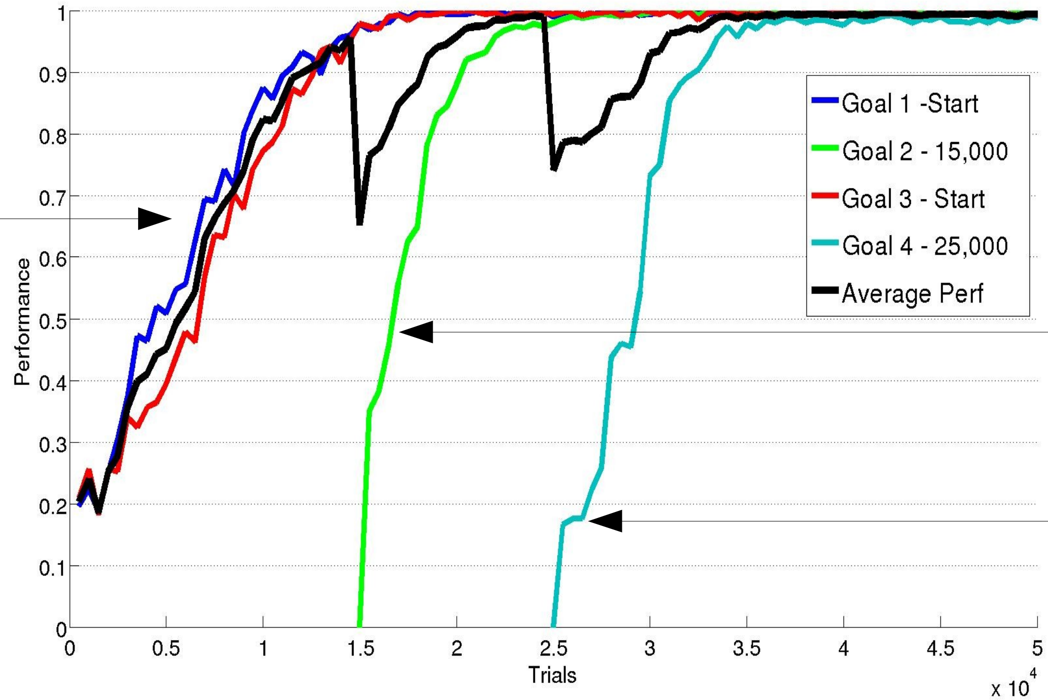


After learning



Model 3: GRAIL learning of four skills

Two spheres from beginning

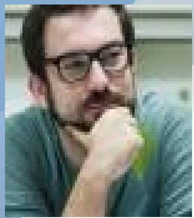


Activation of sphere contingency

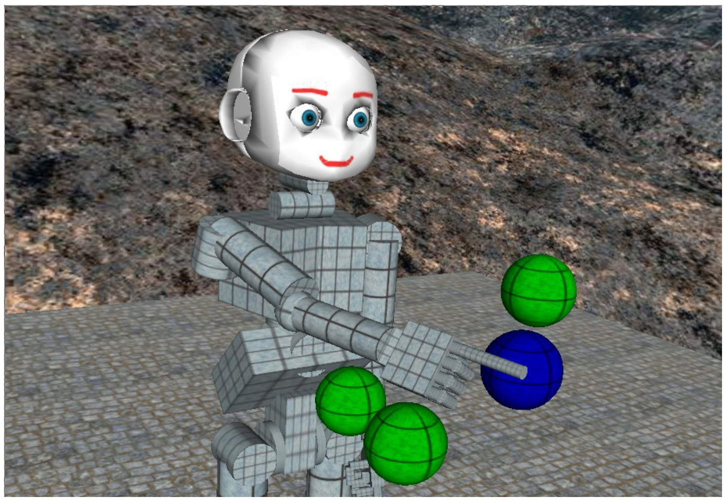
Introduction of new sphere

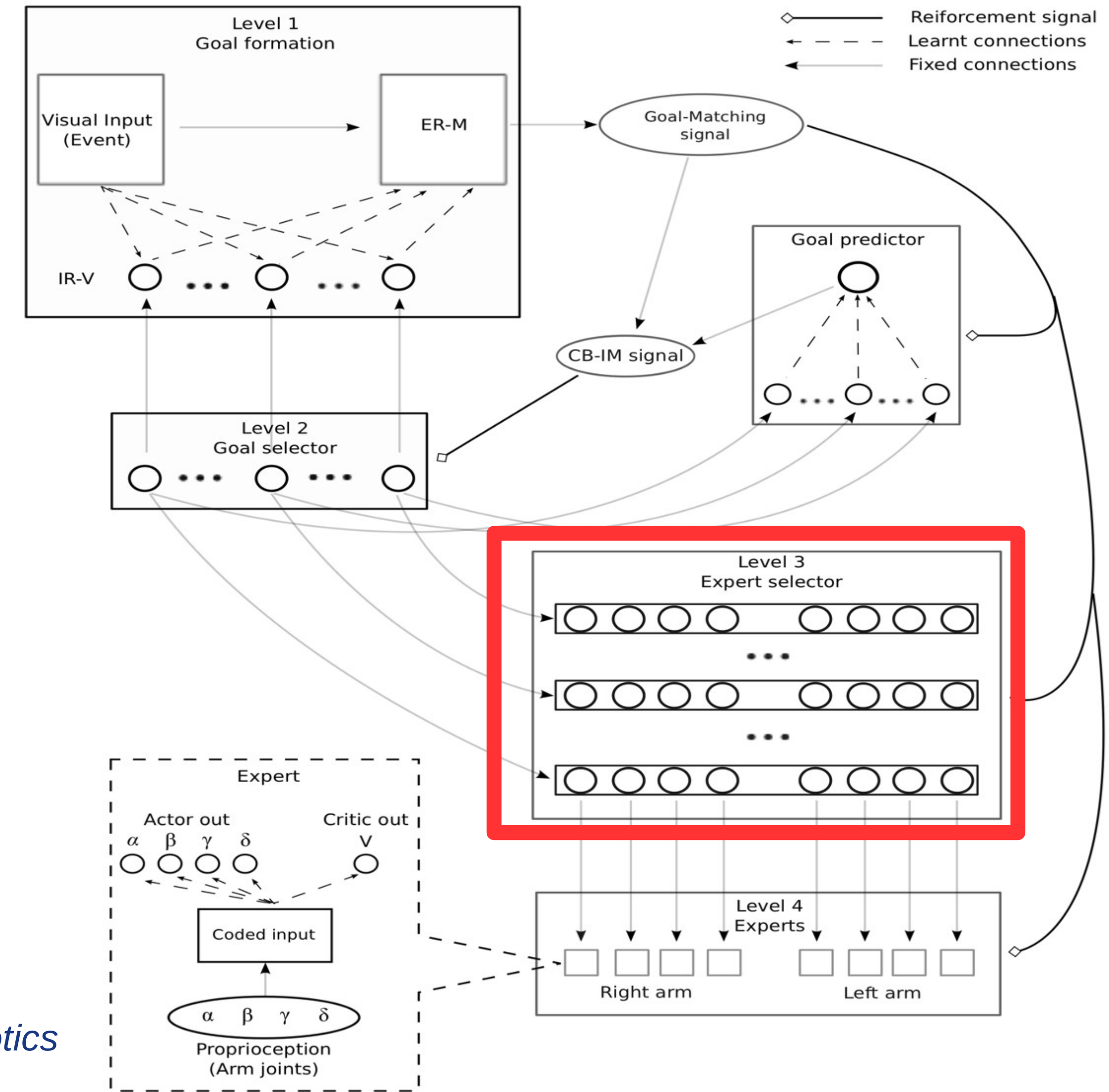
Success in reaching different spheres

Model 3: GRAIL: Goal-discovering Robotic Architecture for Intrinsically-motivated Learning



Vieri Santucci





Santucci et al. (2013) *Frontiers in Neurorobotics*
Santucci et al. (2016) *IEEE TAMD*

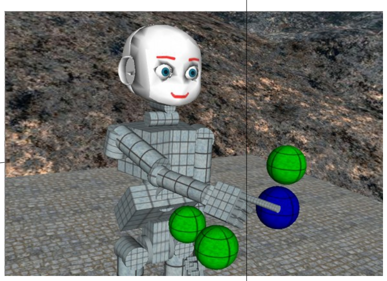
Model 3: learning of goal-skill link

- Systems learns to associate the best expert to each goal:
  - Here: suitable arm (output)
  - In general: best expert for input/internal_resources/output

right Y left

	Obj 1	Obj 2	Obj 3	Obj 4
Right Arm	0	0	20	0
Left Arm	20	20	0	20

Sagittal central line



Dynamic Movement Primitives (DMPs)

- DMPs: Dynamic Movement Primitives
  (Ijspeert, Nakanishi, Shaal, 2002)

Dynamic Movement Primitives

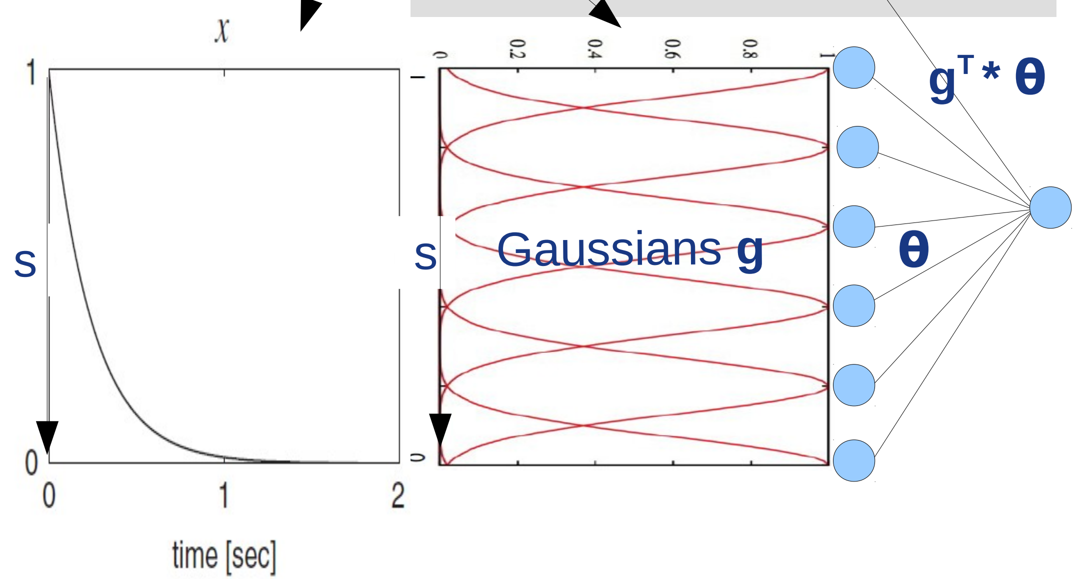
Similar to a PD

$$\frac{1}{\tau}\ddot{x}_t = \alpha(\beta(g - x_t) - \dot{x}_t) + \mathbf{g}_t^{\intercal}\boldsymbol{\theta} \qquad \text{Transform. system} \qquad (1)$$

$$[\mathbf{g}_t]_j = \frac{w_j(s_t)\cdot s_t}{\sum_{k=1}^{p} w_k(s_t)}(g - x_0) \qquad \text{Basis functions} \qquad (2)$$

$$w_j = \exp\left(-0.5h_j(s_t - c_j)^2\right) \qquad \text{Gaussian kernel} \qquad (3)$$

$$\frac{1}{\tau}\dot{s}_t = -\alpha s_t \qquad \text{Canonical. system} \qquad (4)$$



Dynamic Movement Primitives (DMPs)

- DMPs: Dynamic Movement Primitives
  (Ijspeert, Nakanishi, Shaal, 2002)
- RL Policy Search: PI^{BB}
  (Stulp Sigaud, 2012)

Dynamic Movement Primitives

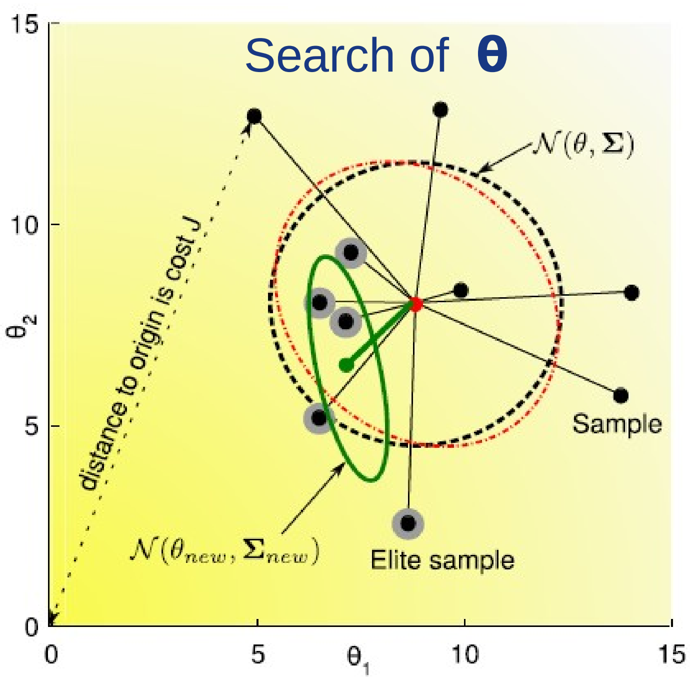
Similar to a PD

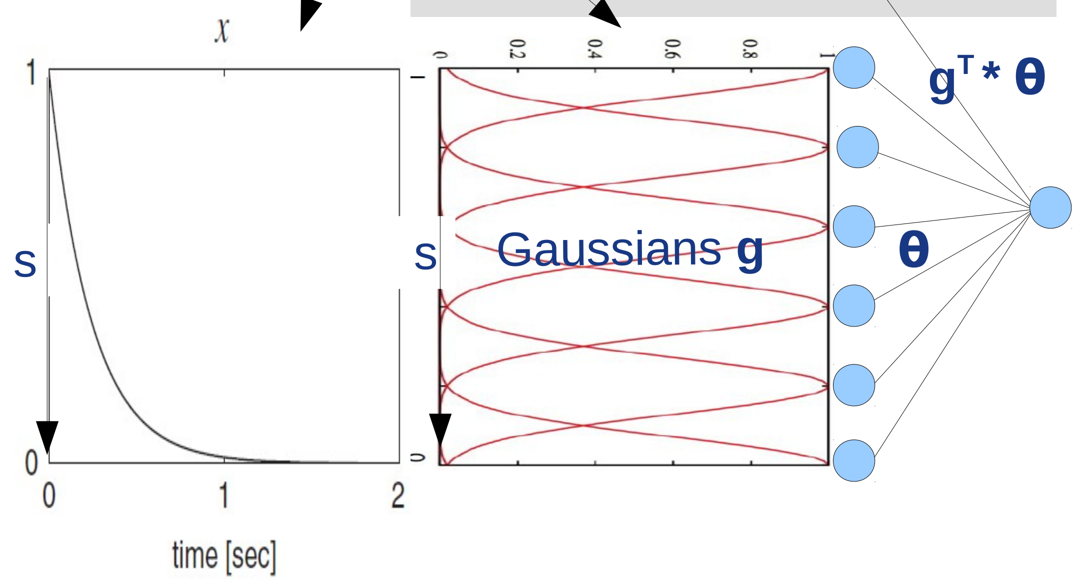
$$\frac{1}{\tau}\ddot{x}_t = \underbrace{\alpha(\beta(g - x_t) - \dot{x}_t)} + \mathbf{g}_t^{\intercal}\boldsymbol{\theta} \qquad \text{Transform. system} \qquad (1)$$

$$[\mathbf{g}_t]_j = \frac{w_j(s_t)\cdot s_t}{\sum_{k=1}^{p} w_k(s_t)}(g - x_0) \qquad \text{Basis functions} \qquad (2)$$

$$w_j = \exp\left(-0.5 h_j (s_t - c_j)^2\right) \qquad \text{Gaussian kernel} \qquad (3)$$

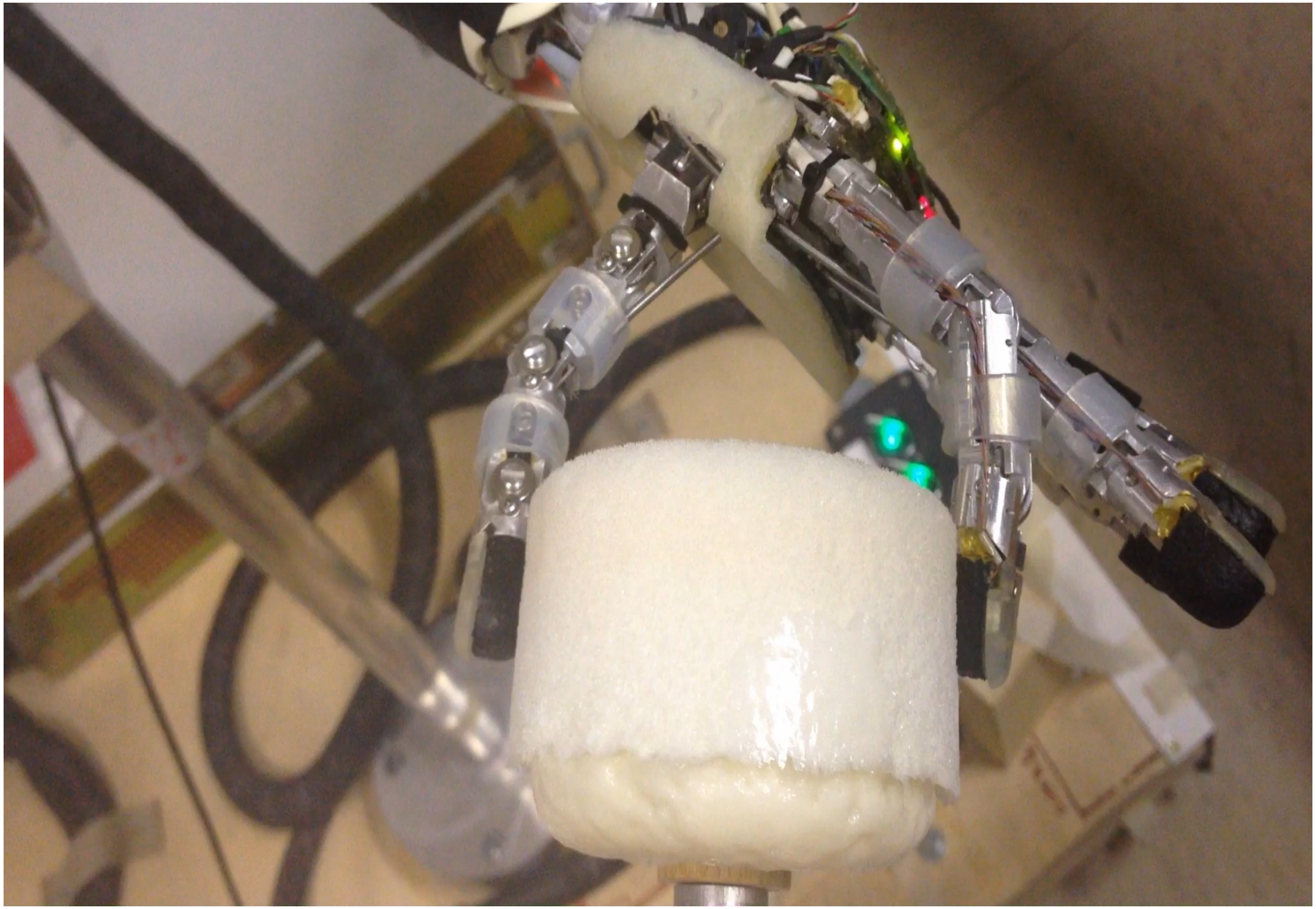
$$\frac{1}{\tau}\dot{s}_t = -\alpha s_t \qquad \text{Canonical. system} \qquad (4)$$





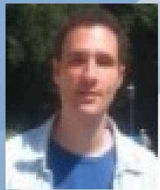
DMPs and policy search RL

Directly learned on the real robot in 10x35 trials

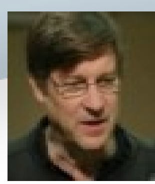


Meola et al. Baldassarre (2016), *IEEE Transact. Cognitive Developmental Syst.*

Model 5: transfer by generalisation

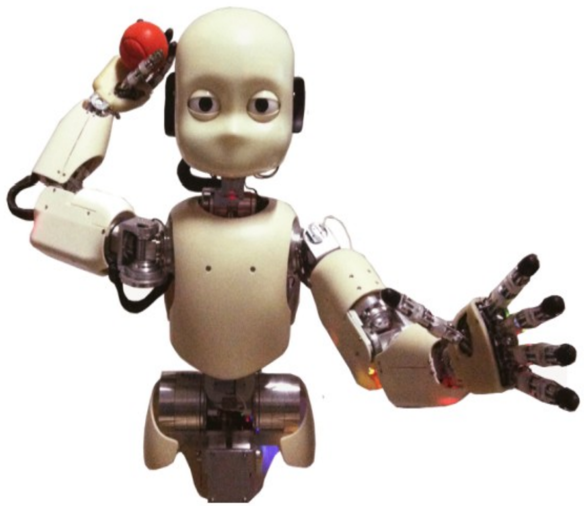


Bruno Castro da Silva



Andrew Barto

Policy parameters: control of 7 DOFs



Different goal parameters: x,y bottle position

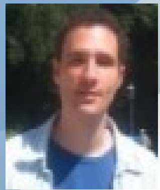
Mapping (e.g., neural net 1)

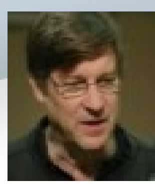
Functioning: **goal params → policy params**

DMP: Input → Output

Controller (e.g., neural net 2)

Castro da Silva et al. (2014) *IEEE ICRA*



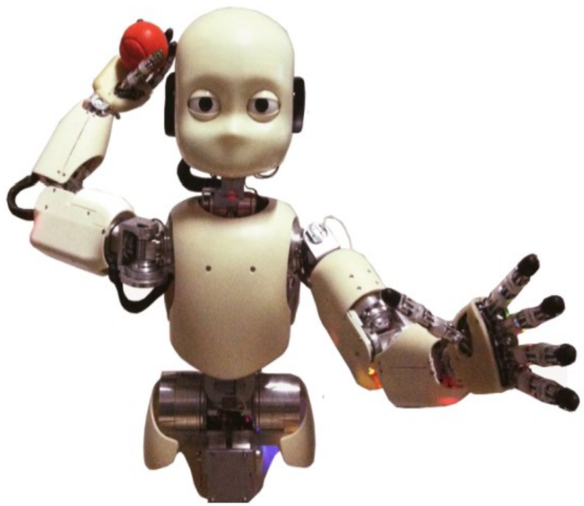


Bruno Castro da Silva

Andrew Barto

Model 5: transfer by generalisation

Policy parameters: control of 7 DOFs



Different goal parameters: x,y bottle position

Supervised learning

Learning: **goal params → policy params**

↑

DMP: Input → Output

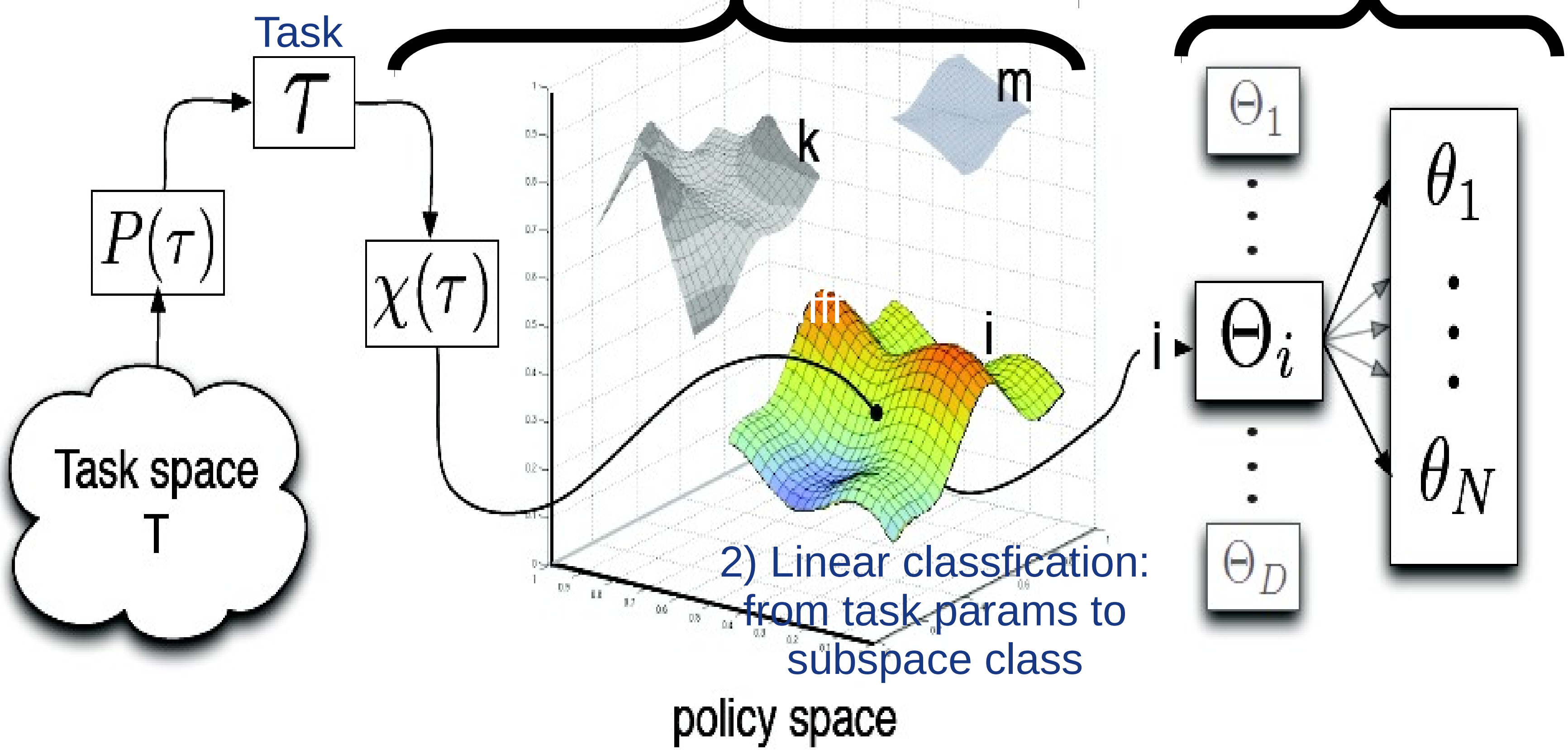
RL

Castro da Silva et al. (2014) *IEEE ICRA*

Model 5: transfer by generalisation

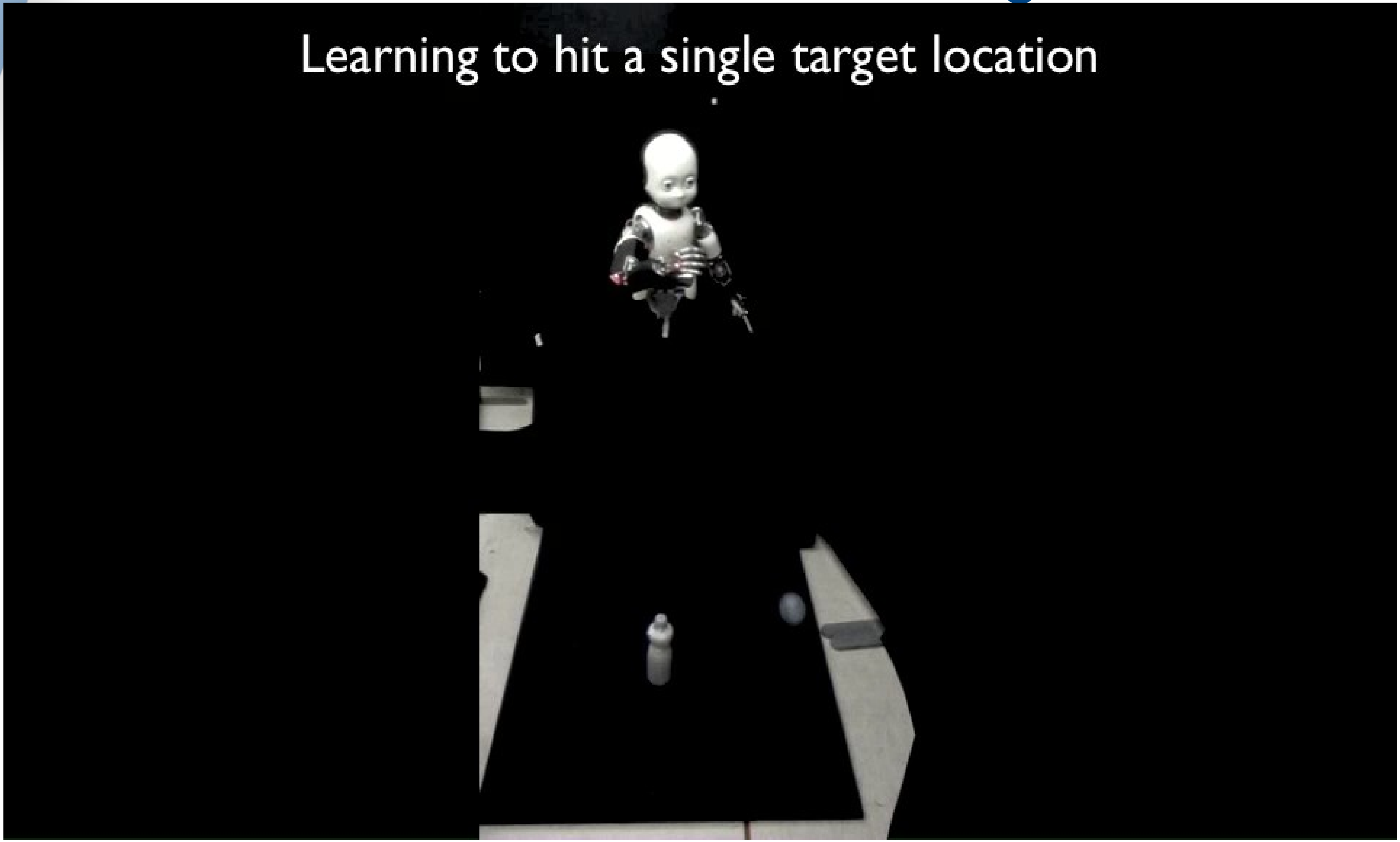
1) Manifold search with ISOMAP, Tenenbaum et al.. (2000) Identification of sub-spaces

3) Many regressions: from task params to policy params



Castro da Silva et al. Baldassarre (2014) *IEEE ICRA*

Model 5, video: robot learns to hit bottle with balls, and generalises



Conclusions

- GOAL-Robots: a novel hypothesis for open-ended learning: IMs → goal self-generation → skill learning
- This solution needs sophisticated architectures (as brain!)
- Key principles to build such architectures:
  - Different IM mechanisms for different key functions
  - Goals as pivot of architectures: learn skills, recall skills, match,...
  - Self-generation of goals as engine of open-ended learning
  - Dynamic models are key for motor exploration, knowledge transfer, catastrophic forgetting avoidance