


Developmental machine learning

Machines that learn like children …
and help children learn better

Pierre-Yves Oudeyer

Project-Team INRIA-ENSTA-ParisTech FLOWERS

http://www.pyoudeyer.com

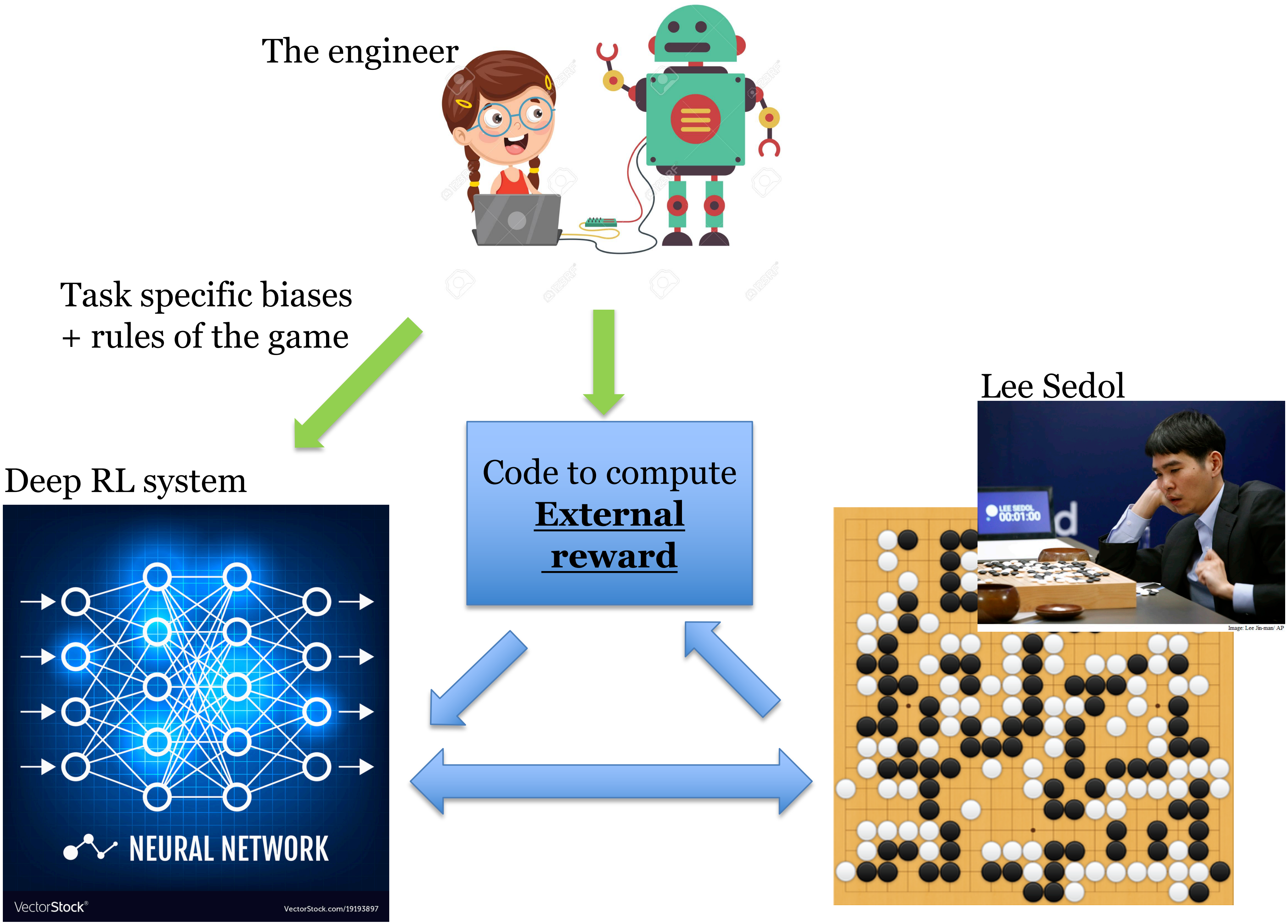
https://flowers.inria.fr

Twitter: @pyoudeyer

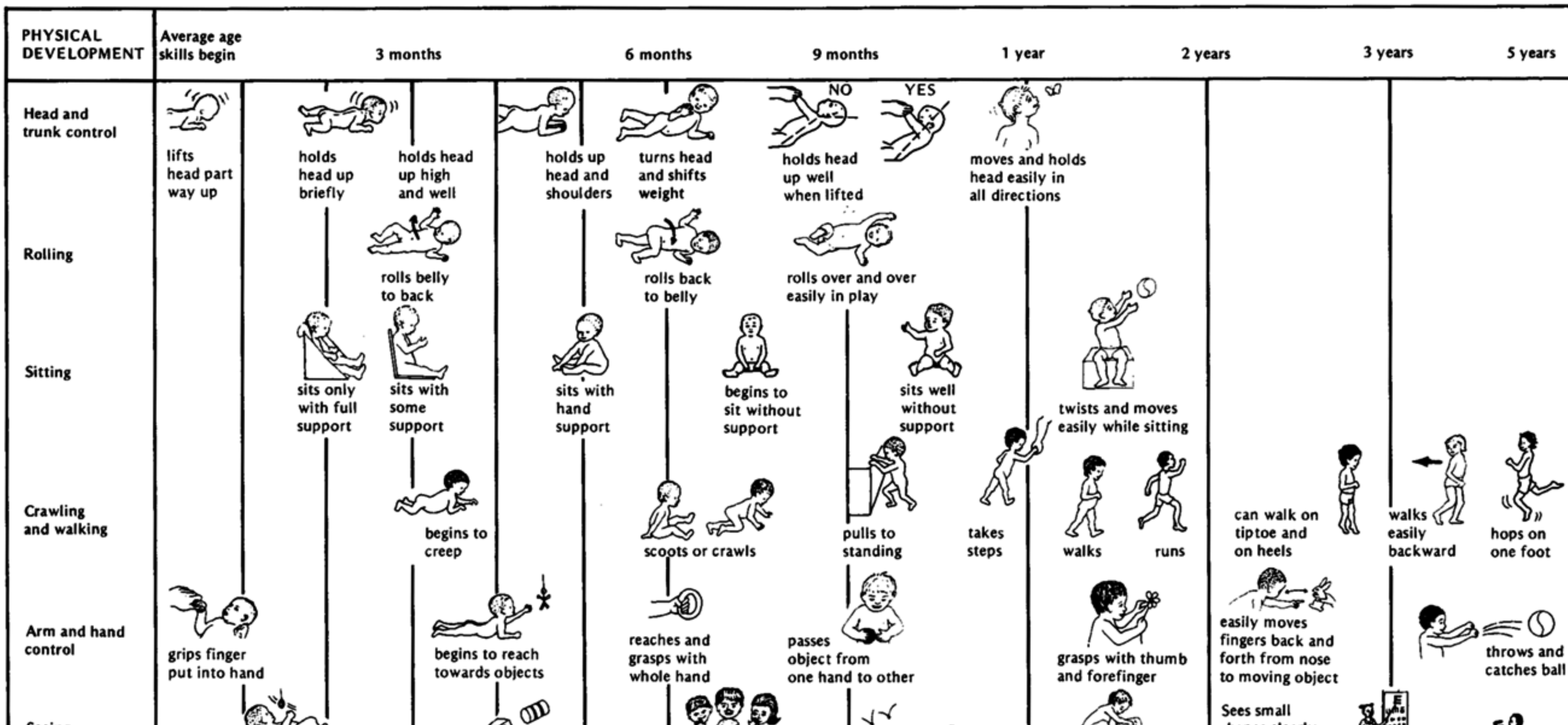
informatics mathematics Inria

erc European Research Council

ENSTA ParisTech


The engineer
Task specific biases
+ rules of the game
Code to compute
External
reward
Lee Sedol
Deep RL system
NEURAL NETWORK
VectorStock®
VectorStock.com/19193897
LEE SEDOL
00:01:00
Image: Lee Jin-man/ AP

Autonomous Learning and Development in Human Infants



- **How do developmental structures form?**
- **What is their role?**

Many collaborations with researchers in
- Developmental psychology
- Neuroscience
- Robotics and AI
- Educational sciences

Cognitive sciences
models to understand better human development

Flowers lab
Inria and Ensta ParisTech
France

Lifelong autonomous learning in **robotics and AI**

Applications in **educational technologies**

Families of developmental « forces »

Body morphology and growth :

- Morphology, body growth and maturation
- Motor and perceptual primitives

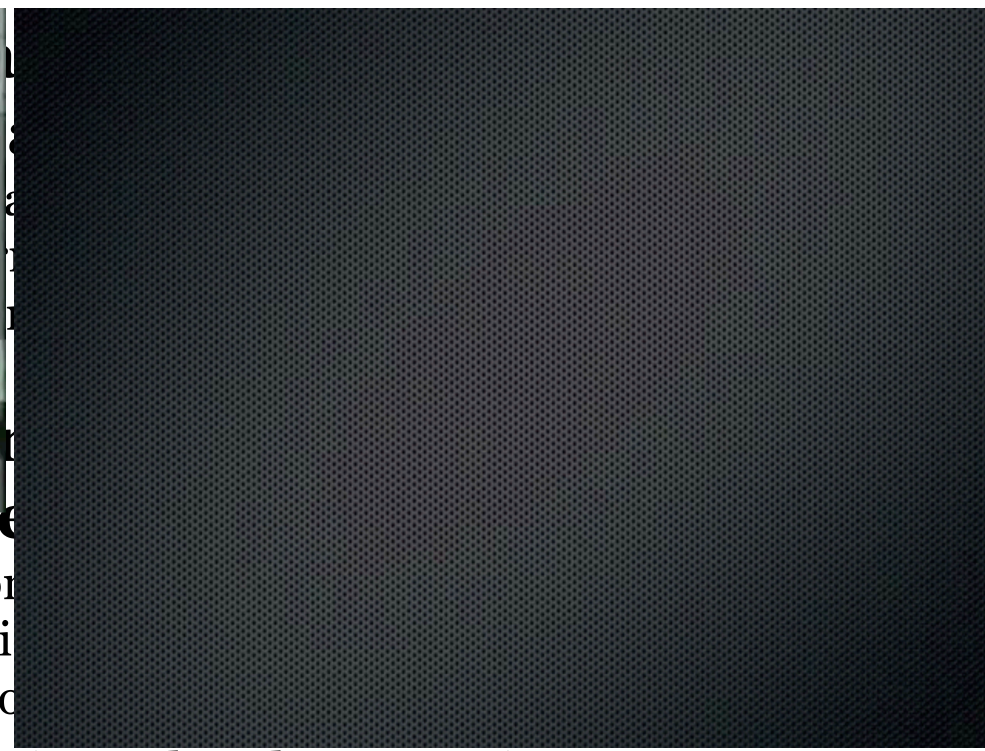
Cognitive biases:

- Affordances
- Perceptual/linguistic categories grounded in action
- Hierarchies of actions, states, objectives

(McGeer, 1991)

activ

- Autor
- Effici
- Self-c



(Ly and Oudeyer, 2010)

Spontaneous active exploration

(Francis Vachon)

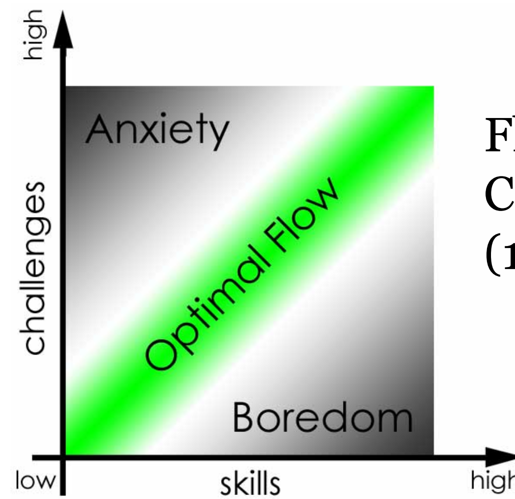
With Lauriane Rat-Fischer, S. Forestier and Alex Kacelnik

With Lauriane Rat-Fischer, S. Forestier and Alex Kacelnik

With Lauriane Rat-Fischer, S. Forestier and Alex Kacelnik

With Lauriane Rat-Fischer, S. Forestier and Alex Kacelnik

Intrinsic motivation, curiosity and active learning

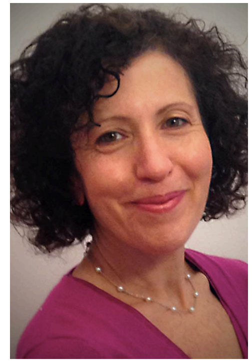


Flow theory
Csikszentmihalyi
(1996)

➔ Intrinsic drive to reduce uncertainty, and to experiencing novelty, surprise, cognitive dissonance, challenge, incongruences, …

➔ Optimal interest = optimal difficulty = neither trivial nor too difficult challenges

Berlyne (1960), White (1960), Kagan (1972), Csikszentmihalyi (1996), (Kidd et al., 2012), …



J.Gottlieb
(Columbia, NY)

L. Smith
(Indiana Univ.)

C. Kidd
(Stanford)

Towards a neuroscience of active sampling and curiosity

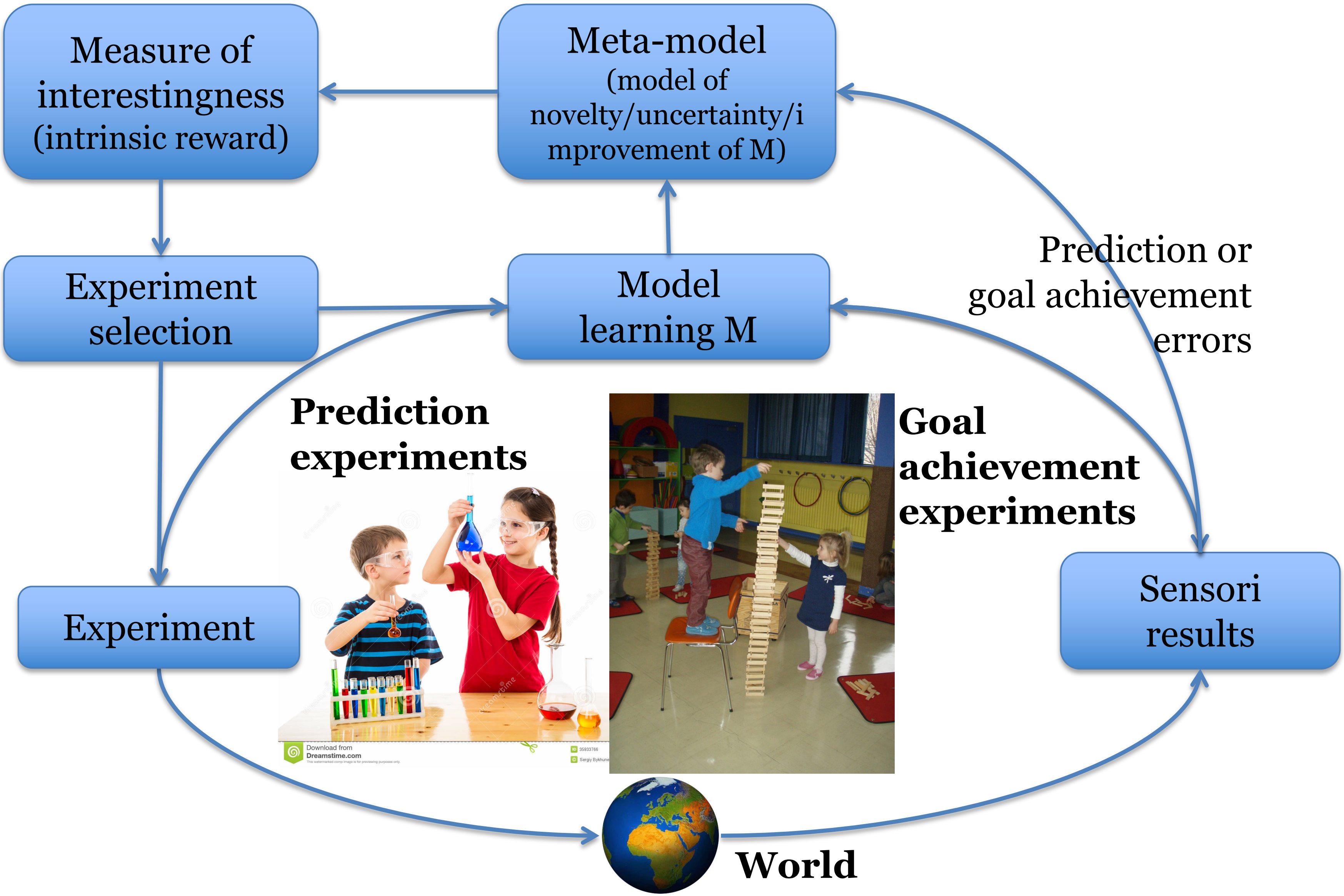
Jacqueline Gottlieb[1,2,3]* *and Pierre-Yves Oudeyer*[4,5]

Development of **a unified formal and theoretical framework** in psychology and neuroscience

(Frontiers in Neuroscience 2007; IEEE TEC 2007; Trends in Cognitive Science, Nov. 2013; Progress in Brain Research, 2016; Frontiers in Neuroscience, 2014; Scientific Reports, 2016; PNAS, 2016; Nature Reviews Neuro. 2018)

The child as a sense-making organism:

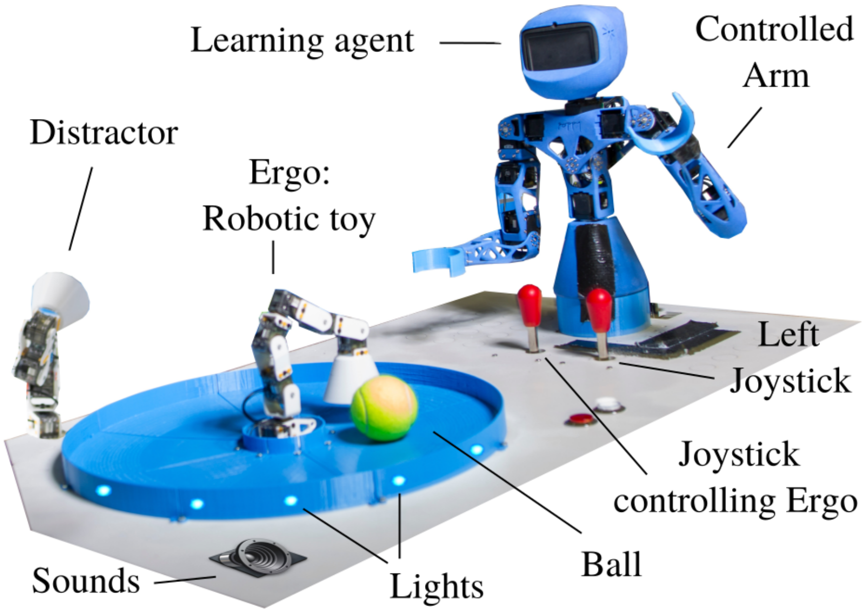
Exploring to make good predictive models of the world and control it!



Robotic Playgrounds

(Oudeyer et al., 2004; 2007)

(Forestier et al., 2016, 2017)



Discovery of sensorimotor affordances

Discovery of **speech communication**

Discovery of **nested tool use**

Essential ingredients:

- Dynamic movement primitives (Schaal, Ijspeert et al, 2003, 2007)
- Object-based perceptual primitives (like infants, builds on prior perceptual learning)
- Self-supervised learning forward/inverse models with hindsight learning and episodic memory
- Curiosity-driven self-organization of learning curriculum through goal exploration

What is an « interesting » learning experiment?

(verbal) hypotheses from psychology and/or developmental biology:

- Cognitive homeostasis/auto-poiesis, high predictability (Varela and Maturana, explo. due to external perturbations)
- High novelty/high uncertainty? (many)
- Knowledge gap, cognitive dissonance? (Kagan, Festinger, Lowenstein)
- Intermediate novelty, intermediate complexity? (Berlyne, Kidd)
- Intermediate challenge? (White, Csikszentmihalyi)

Technical ideas from cognitive modeling or ML:

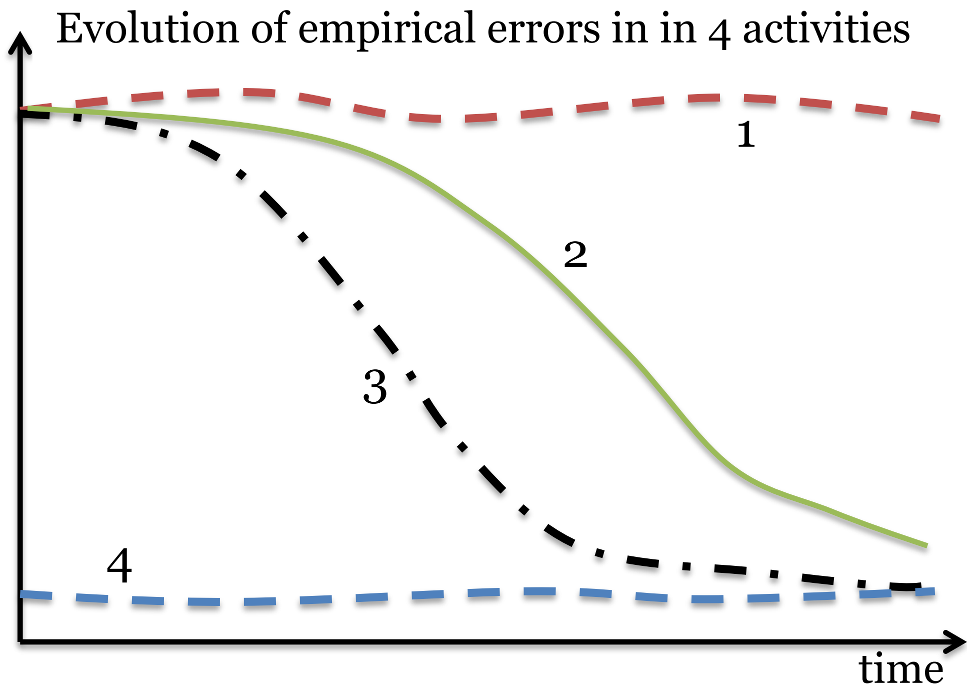
- High novelty/high uncertainty? (many)
- Surprise? (Itti and Baldi)
- Free energy? (Friston)
- Different forms of information gain/learning progress, e.g.:
    - KL-divergence between prior and posterior probabilistic model
    - Predictive information (Martius), predictive information gain (Little & Sommer)
    - Compression progress (Schmidhuber)
    - **Empirical improvement of prediction or control (Oudeyer et al.)**

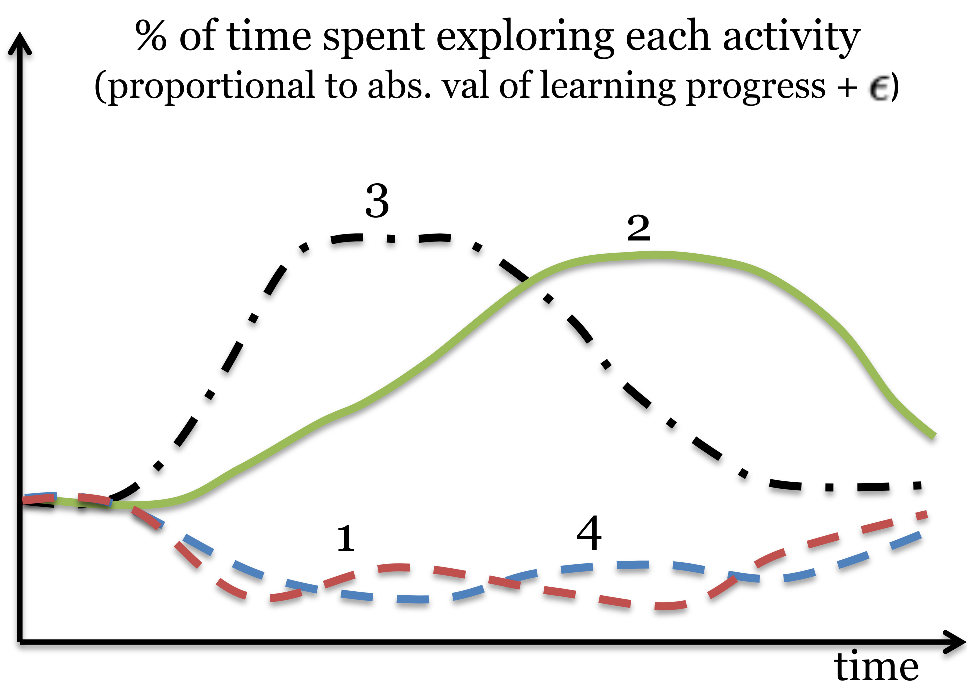
The (absolute) Learning Progress hypothesis

Interestingness
=
proportional to empirical absolute learning progress
(absolute value of derivative)

➔ Automated Curriculum Learning

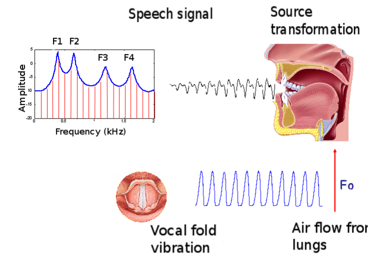
(Oudeyer and Kaplan, 2003; 2007; Gottlieb et al., 2013; Oudeyer et al., 2016)
➔ Few assumptions on underlying learning machinery and on match between biases and real world (as opposed to measures of learning progress based on KL-divergence measures)





Intrinsically Motivated Goal Exploration Processes



Context $s(t)$

Trajectory:

$t + \Delta_t$ time

$$\underbrace{\tau = \{s(t), a(t), \ldots, s(t+\Delta_t)\}}$$

Parameters of motor program π_θ *(DMP, RNN)*

Behavioural descriptors over full trajectory
(can be cost function measuring achievement of a complex property)

$$\varphi = [\varphi_1(\tau), \varphi_2(\tau), \ldots, \varphi_i(\tau)]$$

Mean speed of object C

Vector of params of Bézier curve fitting traj. of obj. A

Classifier counts of events encountered over traj.

Learned RNN embedding

Forward model(s) $F_i : s, \theta$

Inverse model(s) $I_i : s, \varphi_i$

Prediction progress

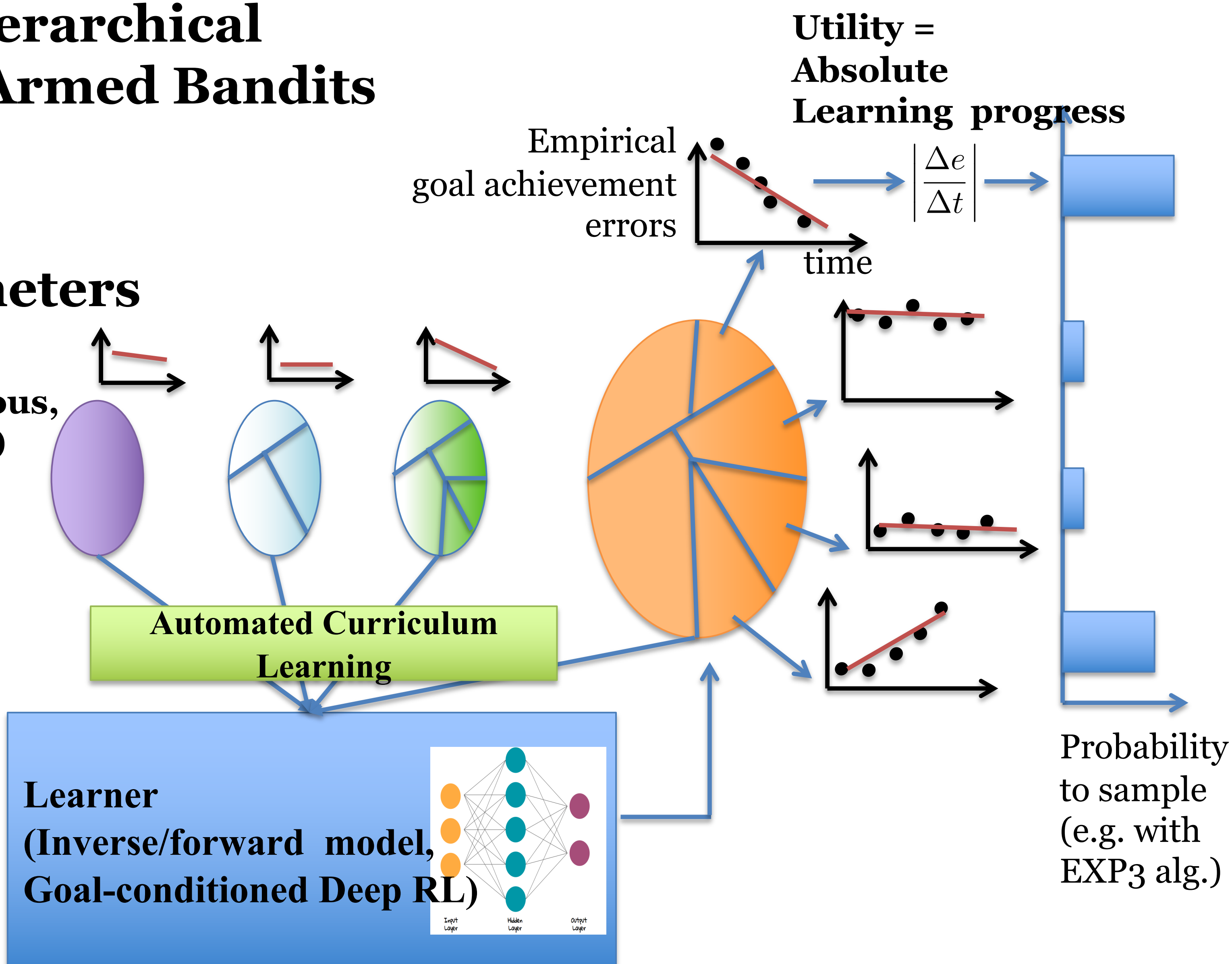
Competence progress

Intrinsically Motivated Goal Exploration (IMGEPs)

(Oudeyer and Kaplan, 2007; Oudeyer and Baranes, 2013)

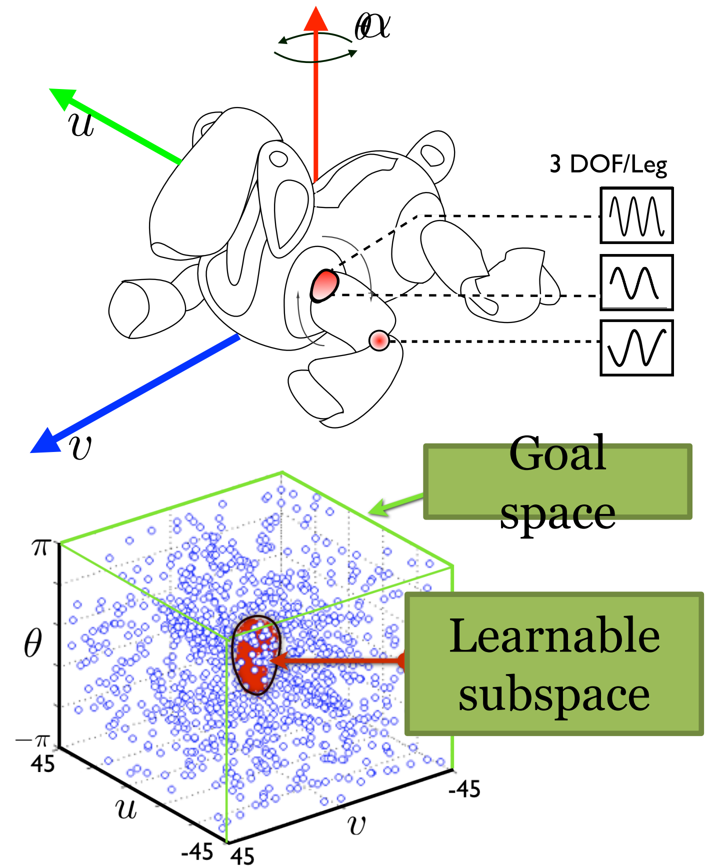
Goal sampling with Hierarchical Multi-Armed Bandits

Goal parameters space
(continuous, high dim)



(Oudeyer and Kaplan, IEEE TEC 2007; Baranes and Oudeyer, RAS 2013; Colas et al., ICML 2019)

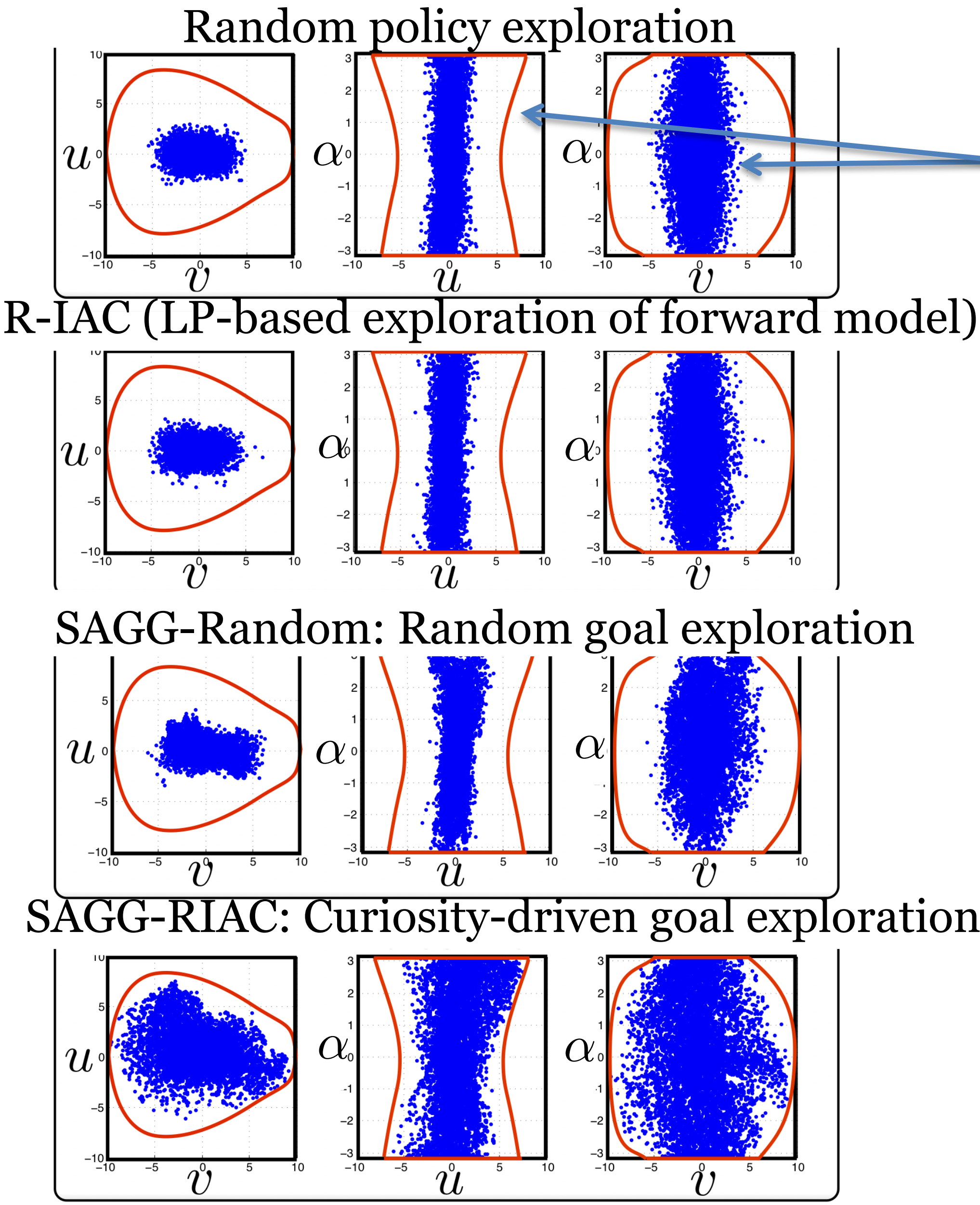
Exploring omni-directional locomotion



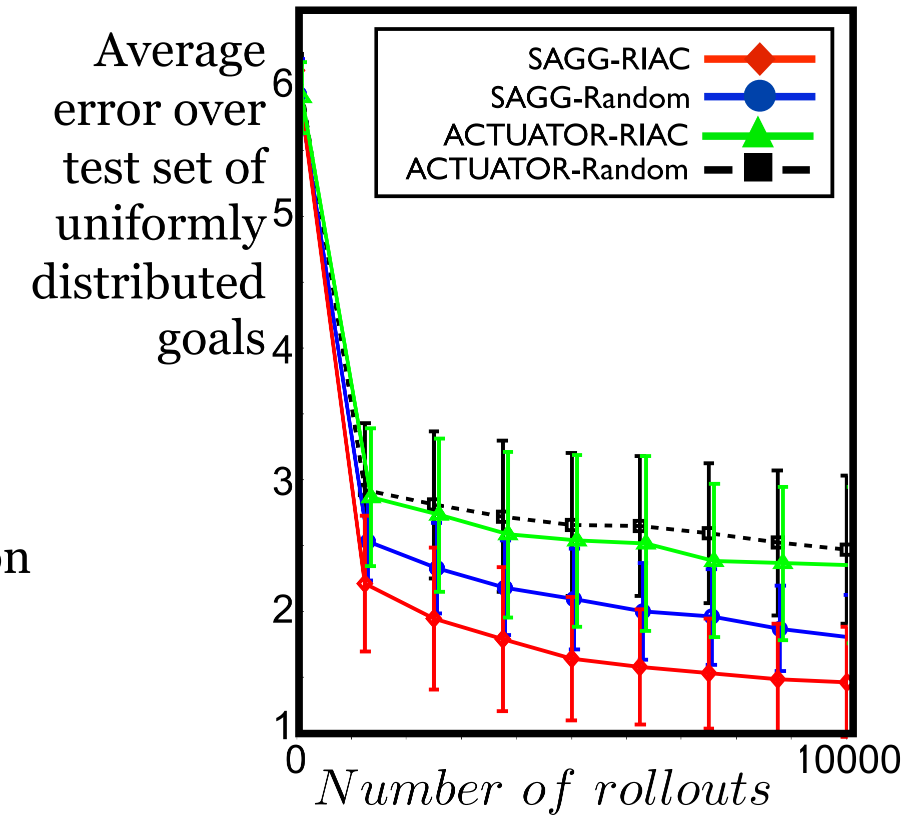
Policies: π_θ : oscillators in 8 motors, $\theta \in [-1, 1]^{24}$

Behavioral descriptors: $\varphi \in \mathcal{R}^3$: translation and rotation over 3s

(Baranes and Oudeyer, IROS 2010, RAS 2013)

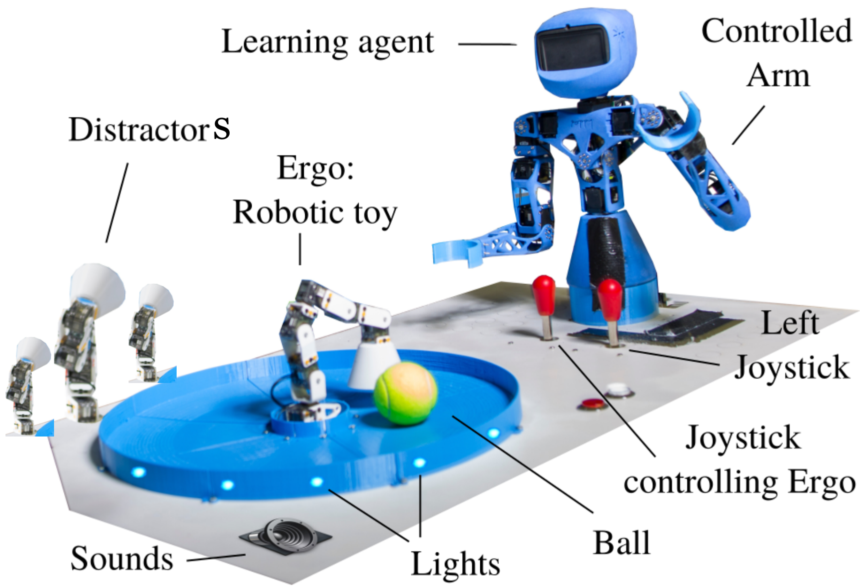


Distribution of discovered goal solutions projected in several planes of (u, v, α)

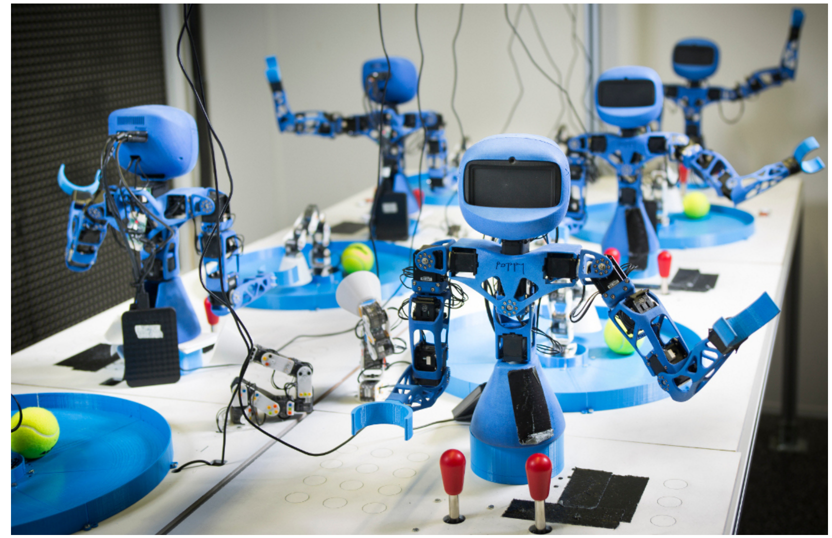


(Baranes and Oudeyer, Robotics and Autonomous Systems, 2013)

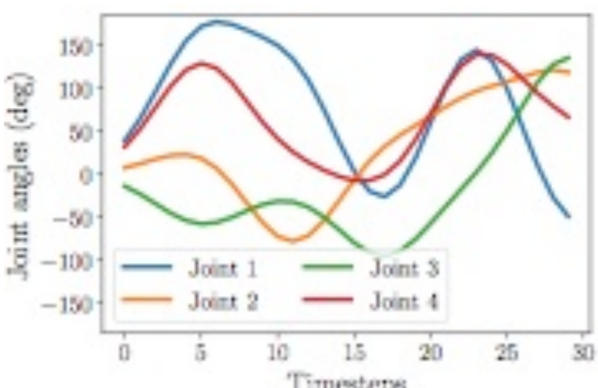
MACOB: Modular population-based IMGEPs



Poppy open-source robots: http://www.poppy-project.org



π_θ 32 dim. Dynamic Motion Primitive



(b) Example Joints Trajectory

(Forestier et al., 2017)

Behavioral descriptors

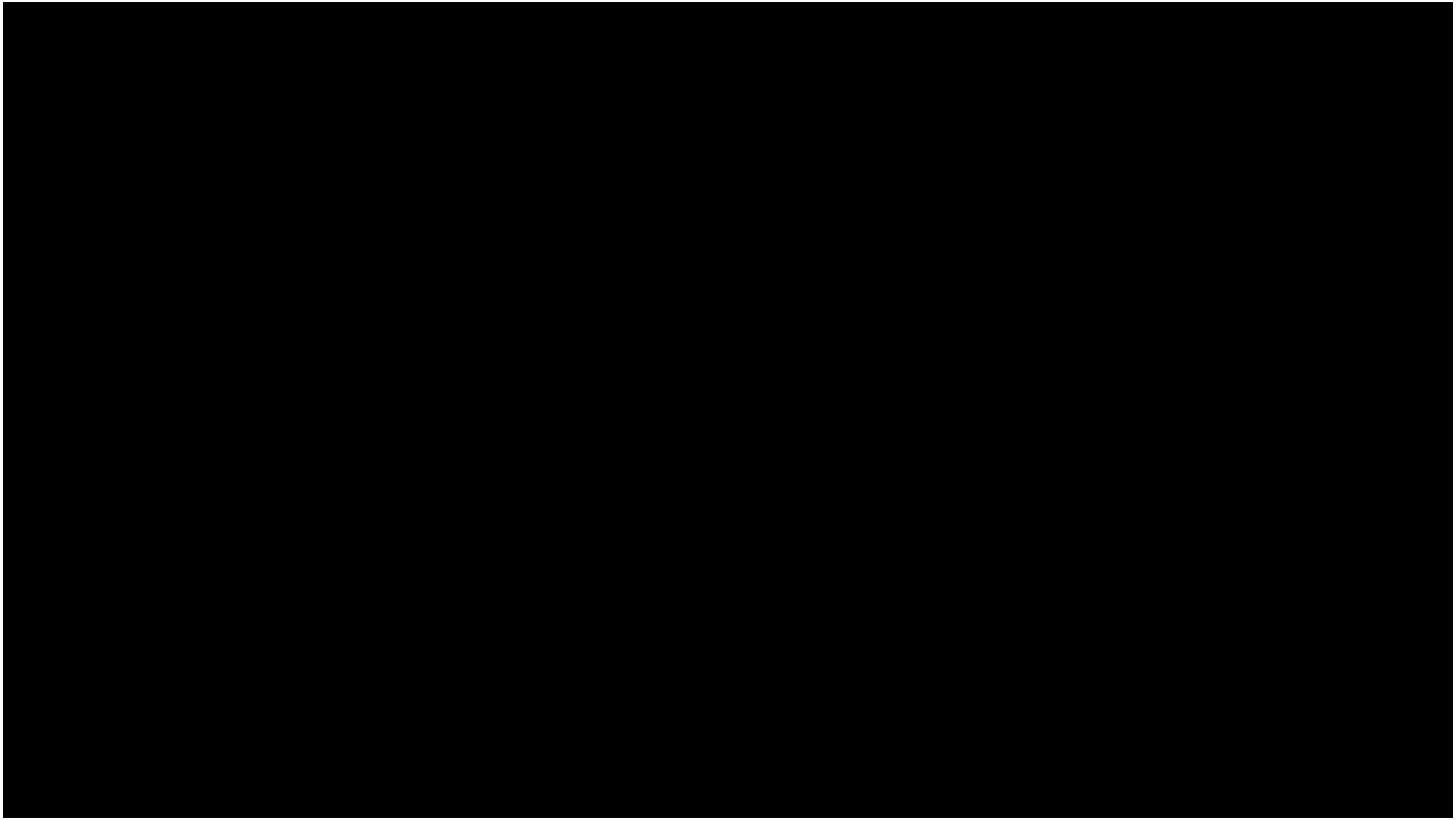
$$\varphi = [\varphi_i] \in \mathcal{R}^{310}$$

Traj. Params. of object positions/sound/light

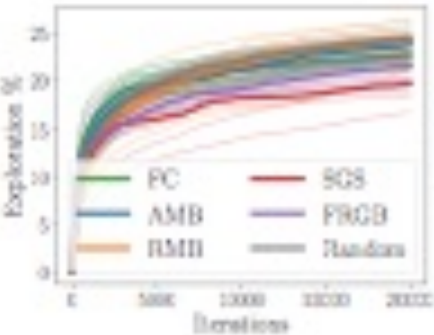
φ_{hand}
$\varphi_{JoystickR}$
$\varphi_{JoystickL}$
$\varphi_{WhiteToy}$
$\varphi_{Ball} \cdots$

$\varphi_{distractor1}$
$\varphi_{distractor2}$
$\vdots$
$\varphi_{distractor8}$

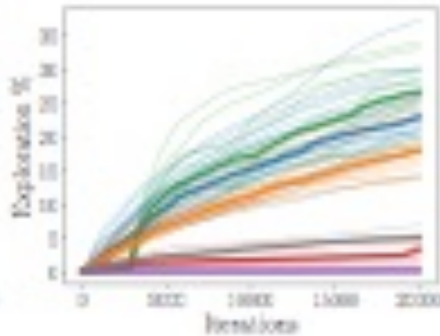
Curiosity-driven discovery of tool use



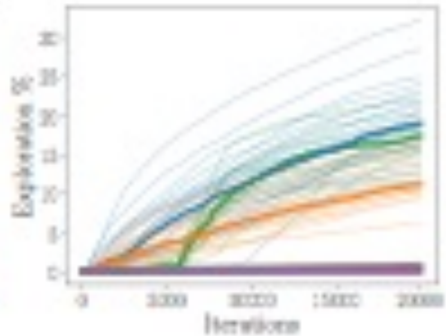
https://www.youtube.com/watch?v=NOLAwD4ZTW0
(Forestier et al., 2017)



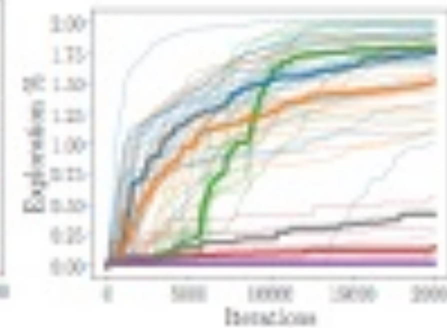
(a) Hand



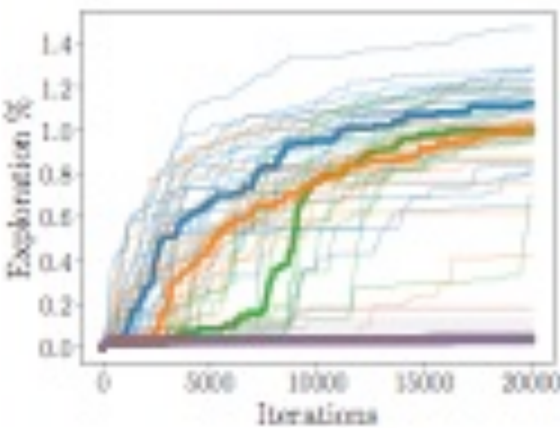
(b) Joystick Left



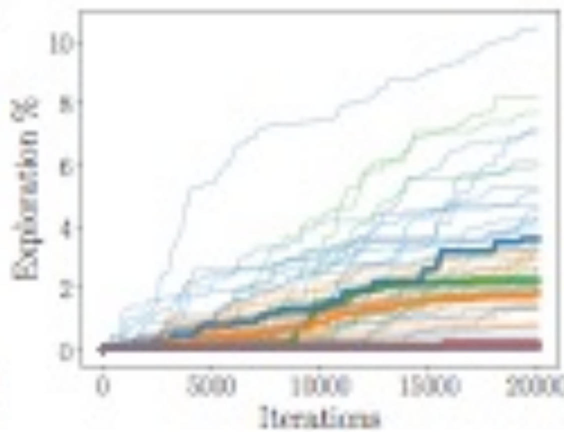
(c) Joystick Right



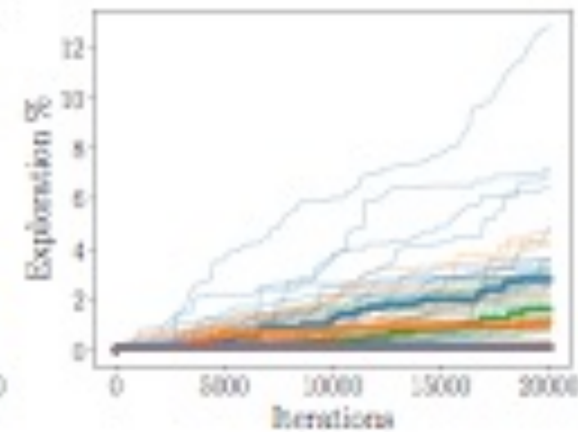
(d) Ergo



(e) Ball

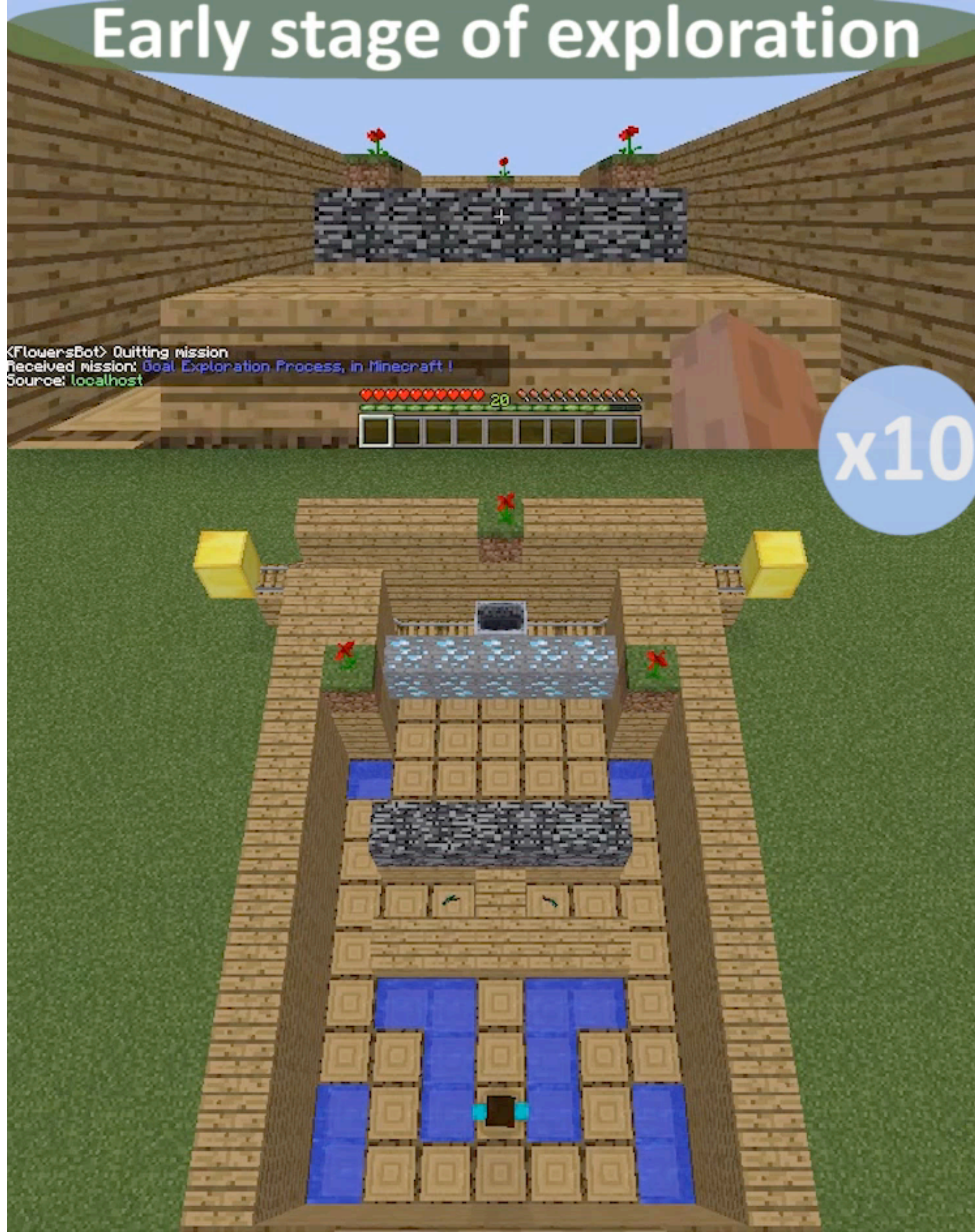


(f) Light



(g) Sound

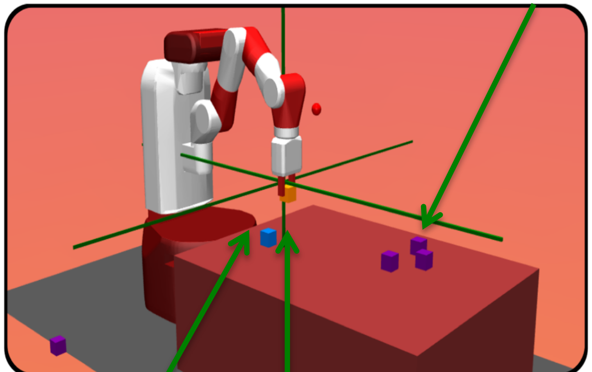
Project Malmo (Minecraft) with neural net controllers



Rémy Portelas
(Microsoft-Inria grant)

CURIOUS: intrinsically motivated modular multi-goal Deep RL

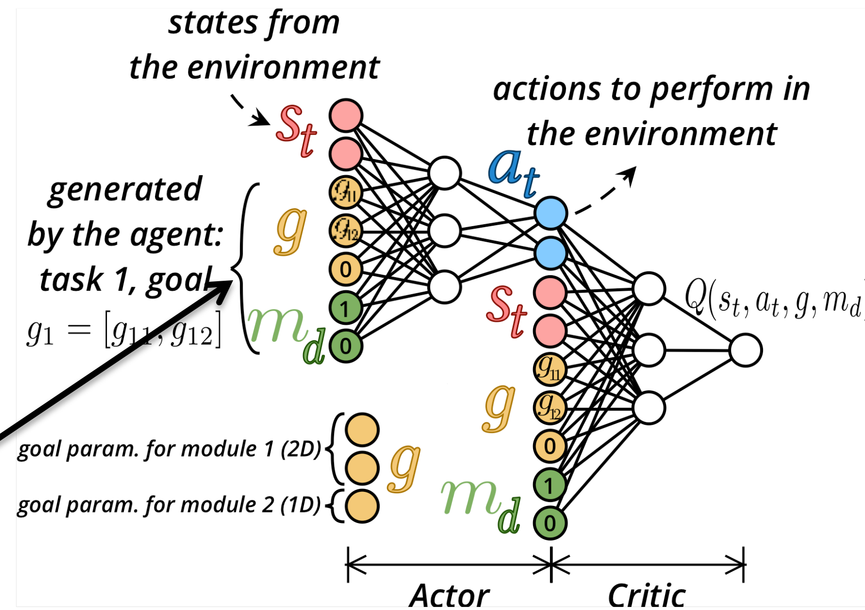
Distractors



Controllable objects

External world

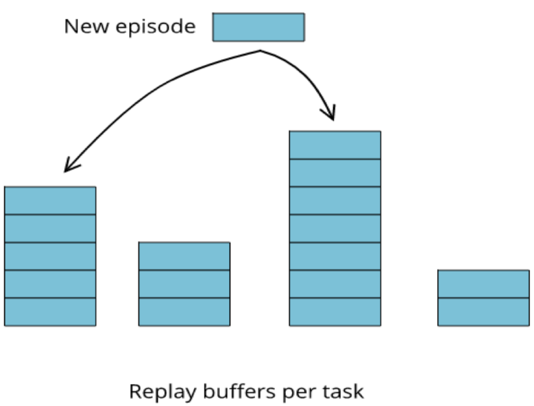
LP-based sampling of modules and goals



Modular UVFA (extended-UVFA)

Goal types and goal values:

Move gripper to (x,y,z)
Pickandplace cube2 at (x,y,z)
Push(cube1) at position (x,y)
Stack cube1 over cube3 ...



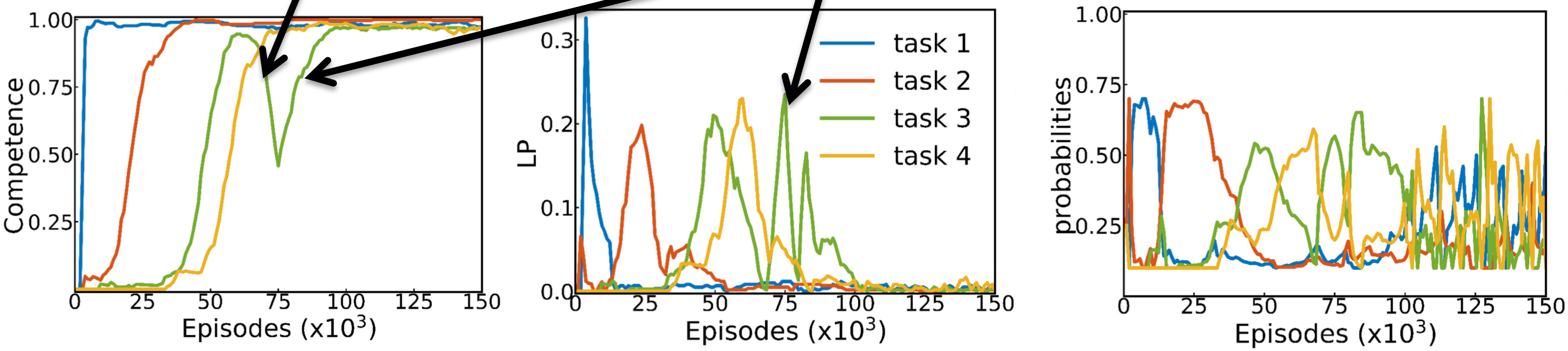
Modular replay buffer

with hinsight learning
(module and goal substitution)

(Colas et al., ICML19)

Forgetting due to interferences among modules/goals

Mitigated thanks to LP-based re-exploration



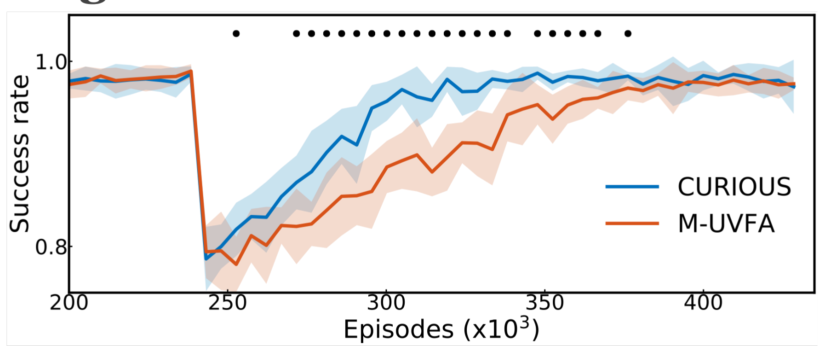
Competence

Absolute Learning Progress

Selection Probabilities

Recovery following a sensory failure.

CURIOUS recovers 95 % of its original performance twice as fast as M-UVFA+HER.



Deep RL based IMGEPs (Curious) vs. Population-based IMGEPs:
+ better generalization
- Slower initial discoveries

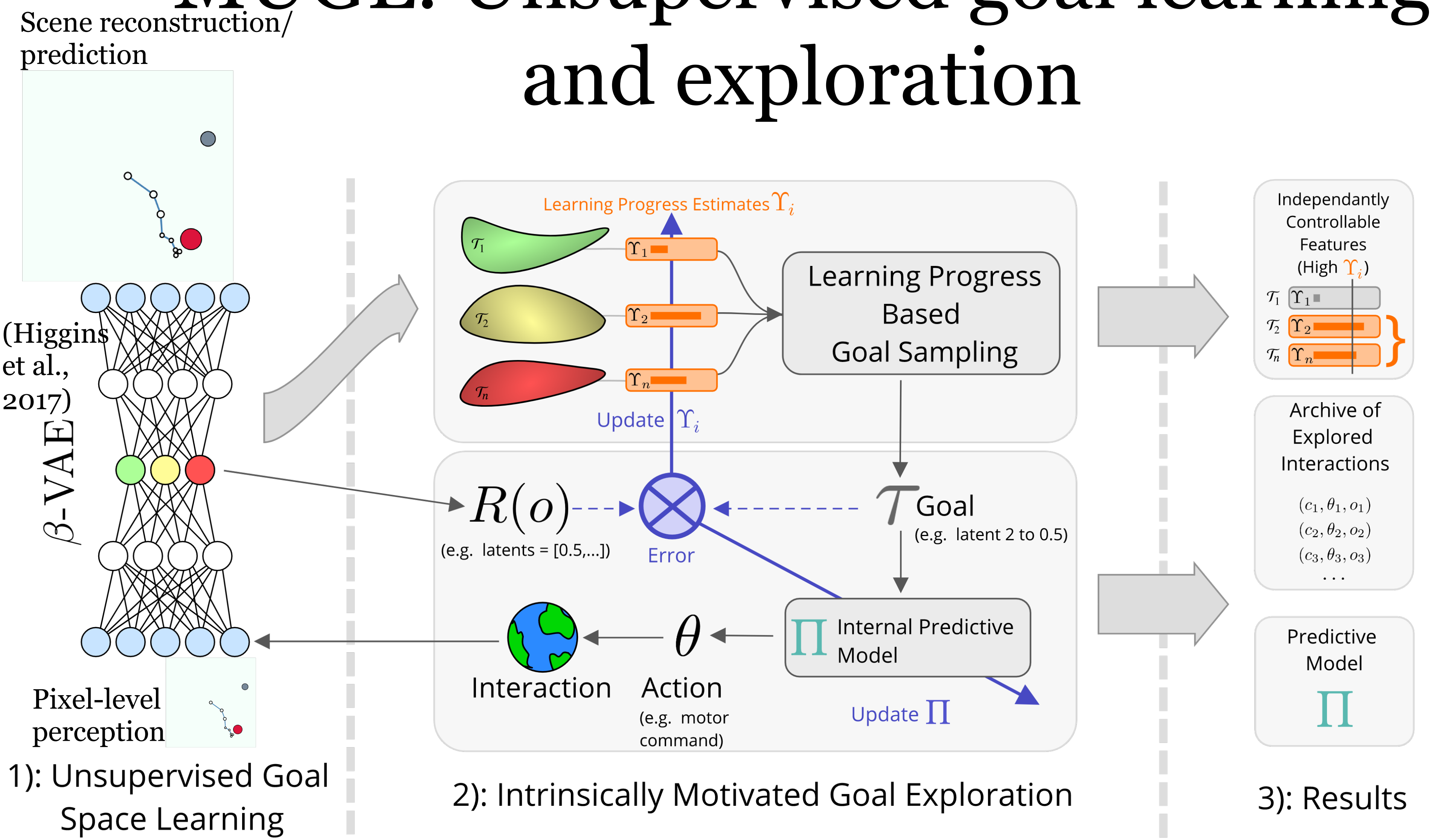
How to learn (modular) representation of goals?

2 approaches:

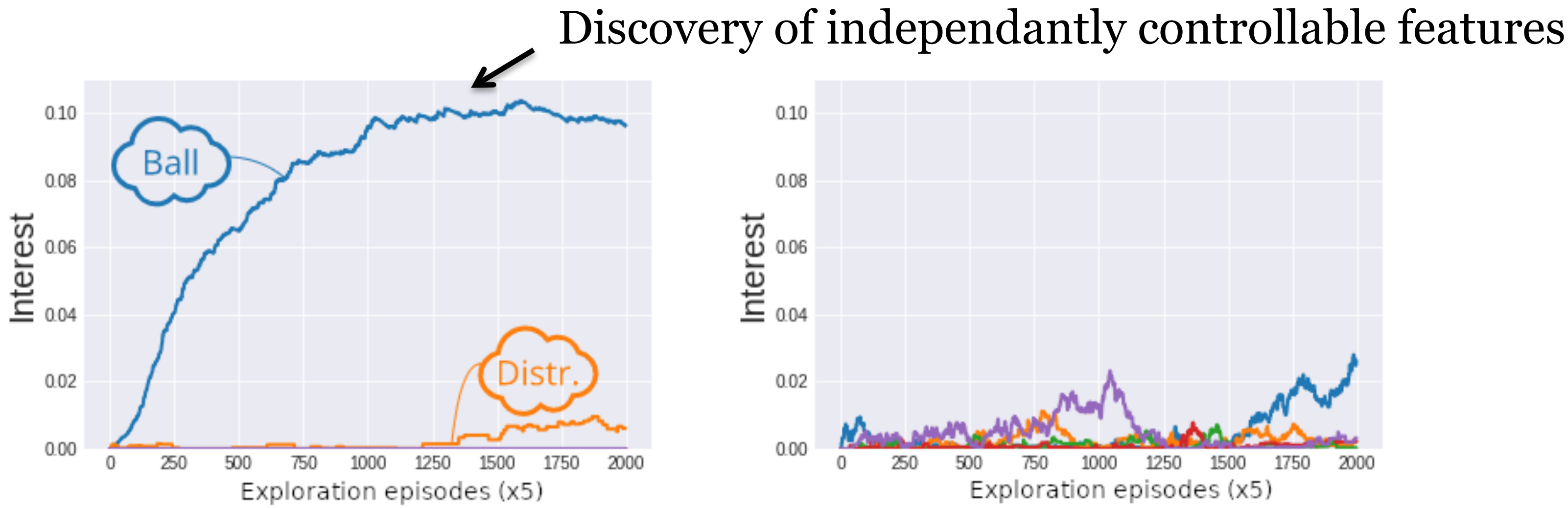
1) **Unsupervised learning** (beta-VAEs)
(Laversanne-Finot et al, CoRL 2018)

2) Leveraging **language** and its compositionality
(Lair et al., Vigil workshop at Neurips 2019)

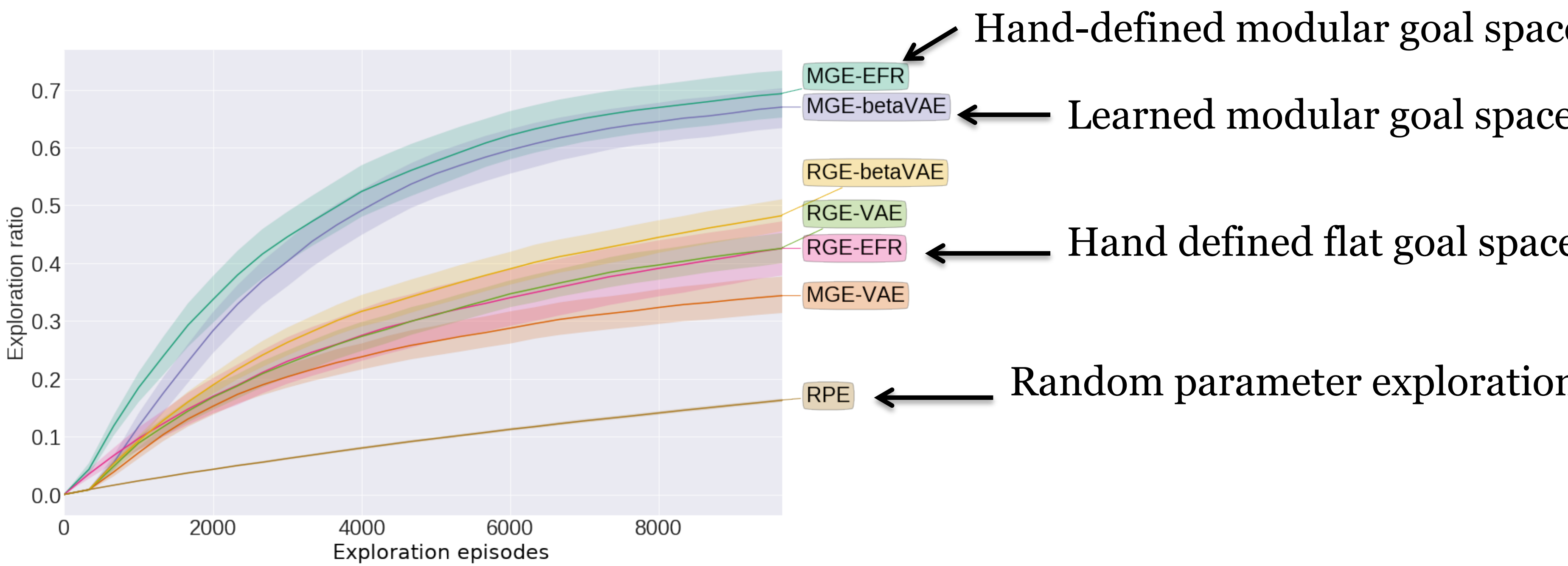
MUGL: Unsupervised goal learning and exploration



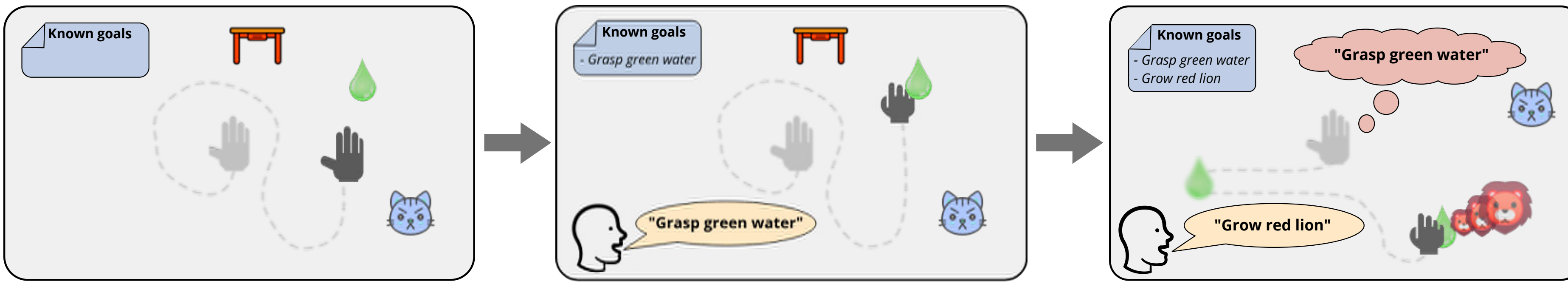
(Péré et al., ICLR 2018; Laversanne-Finot et al., CoRL 2018)



(a) Disentangled representation (βVAE) (b) Entangled representation (VAE)

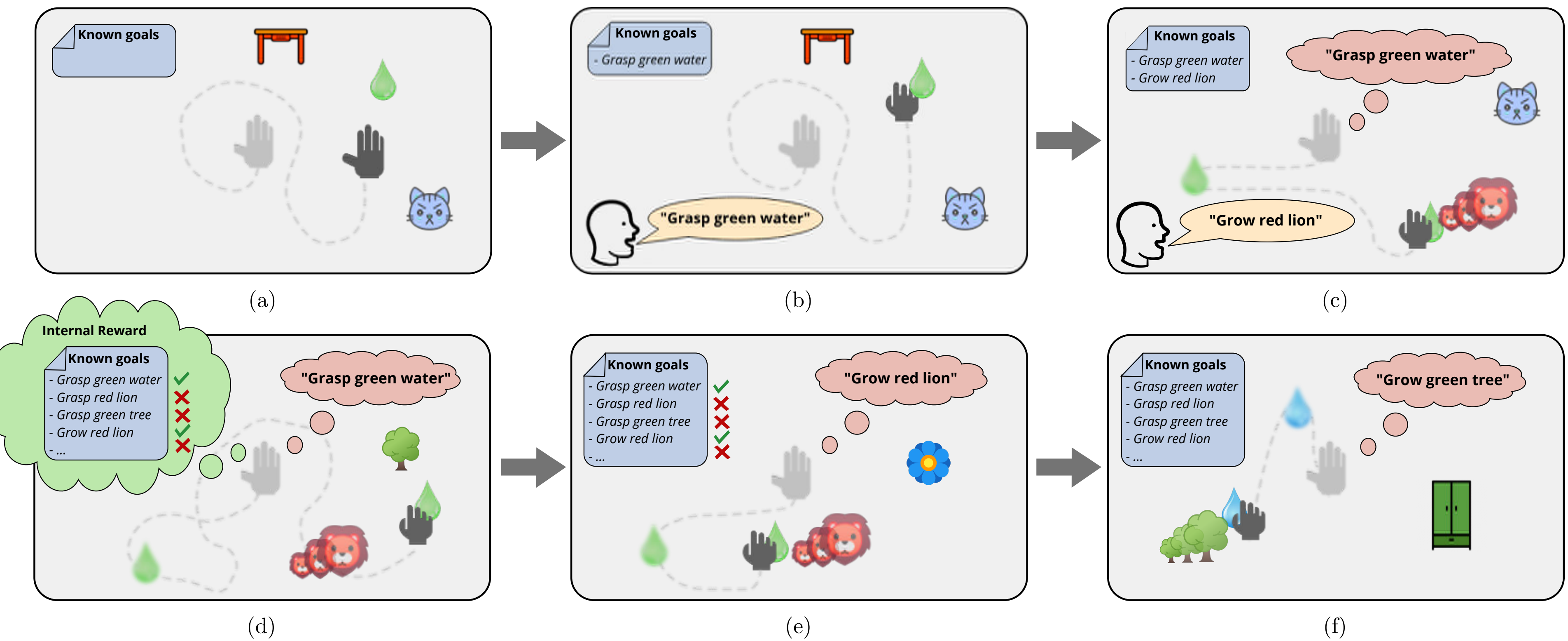


Using language as a cognitive tool to imagine new goals in curiosity-driven exploration



Language Grounding through Social Interactions and Curiosity-Driven Multi-Goal Learning
Nicolas Lair, Cédric Colas, Rémy Portelas, Jean-Michel Dussoux, Peter Ford Dominey, Pierre-Yves Oudeyer
Vigil Workshop at Neurips 19

Using language as a cognitive tool to imagine new goals in curiosity-driven exploration



(a) (b) (c) (d) (e) (f)

Language Grounding through Social Interactions and Curiosity-Driven Multi-Goal Learning

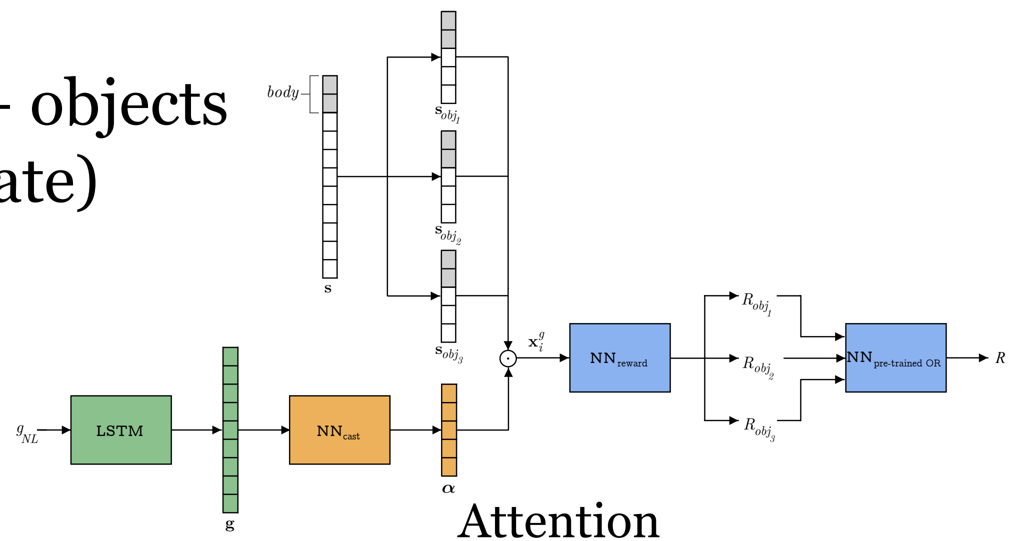
[Nicolas Lair](), [Cédric Colas](), [Rémy Portelas](), [Jean-Michel Dussoux](), [Peter Ford Dominey](), [Pierre-Yves Oudeyer]()

Vigil Workshop at Neurips 19

Understanding sentences by learning a reward function that predicts when it becomes true

State of body + objects
+ delta(state)

Sentence



Learned sentence embedding

Attention mechanism

Predicts whether sentence is true in that situation

➔ Can be used as an internal reward function to measure whether an internally generated goal (= a sentence) is achieved by the goal-parameterized policy being learnt

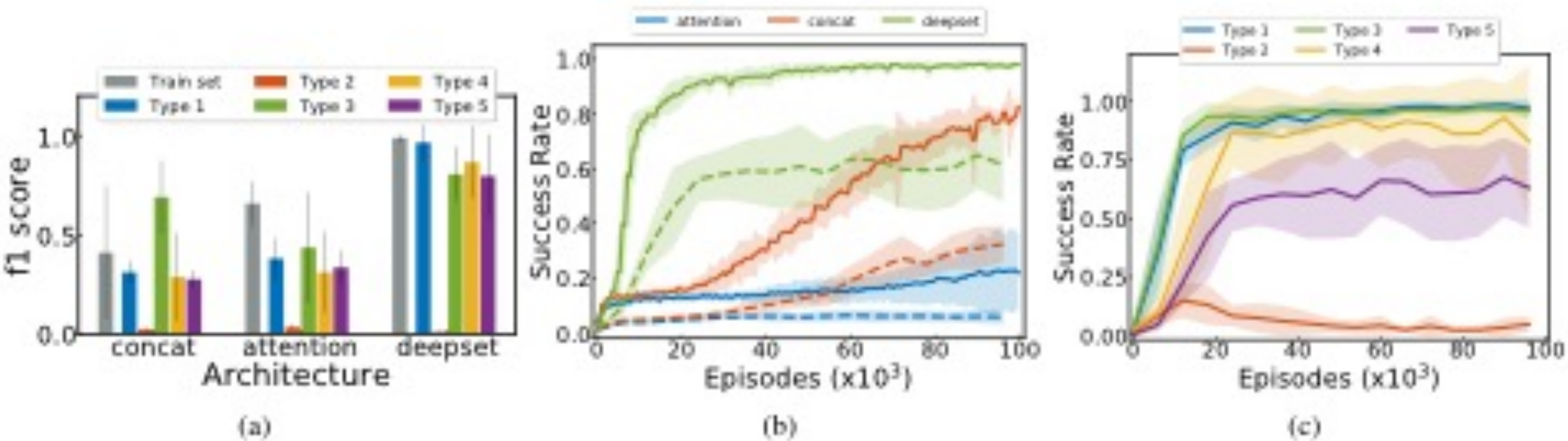
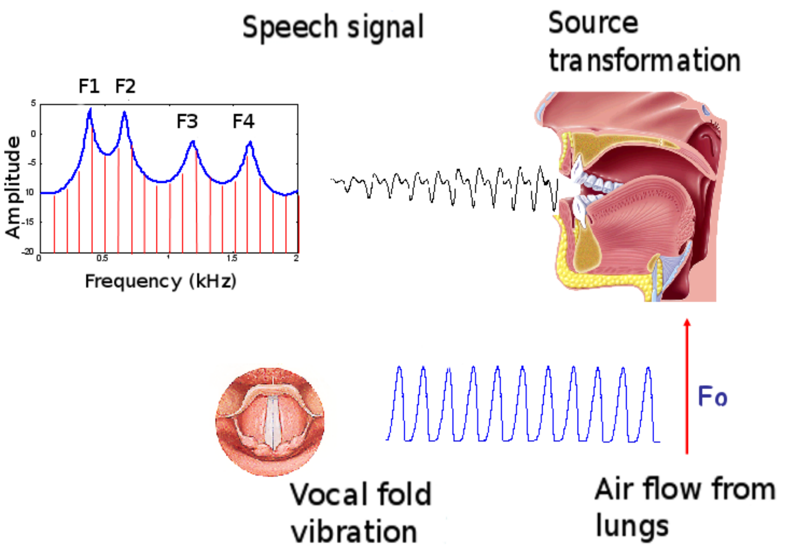


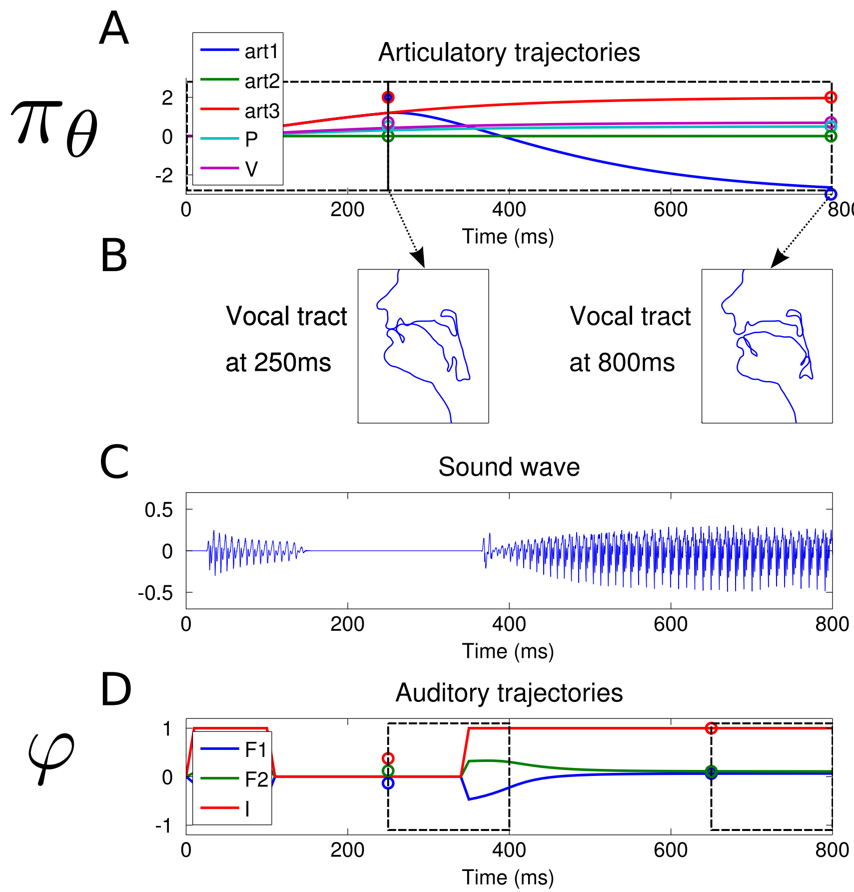
Figure 3. **Reward function and policy learning.** a: F1-score of the reward function after convergence on $\mathcal{G}^{train}$ (grey) and on the 5 types of goals from $\mathcal{G}^{test}$ (colors). b: Average success rates on $\mathcal{G}^{train}$ (plain) and $\mathcal{G}^{test}$ (dashed) for various policy/critic architectures. c: Average success rates on $\mathcal{G}^{test}$, split into the 5 generalization types. Mean +/- std over 10 seeds for all figures.

Models of
child development data

Self-organization of vocal development



DIVA Vocal tract model (Guenther et al.)

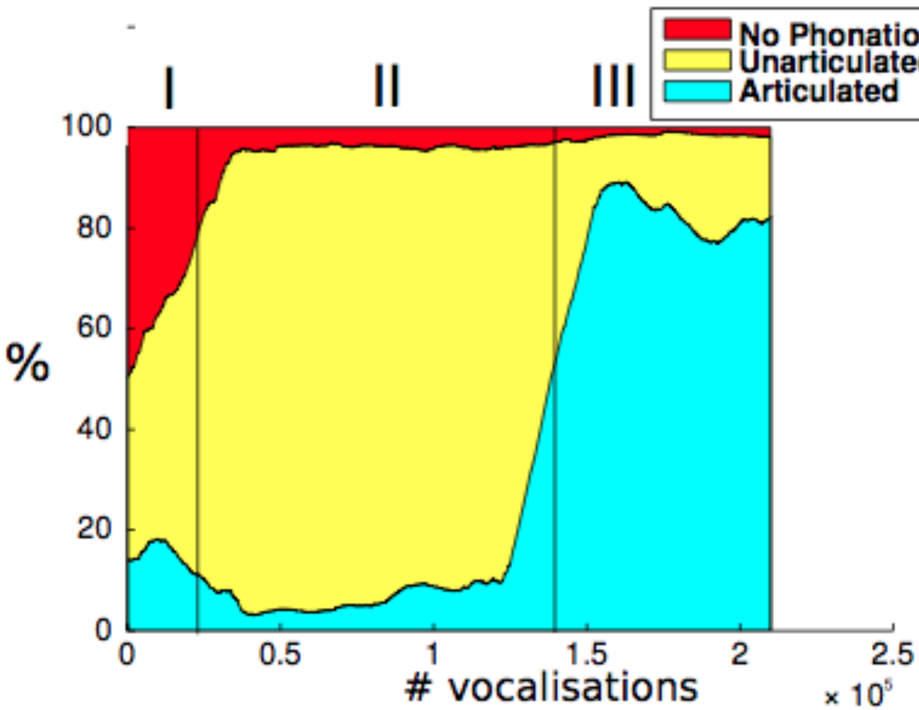


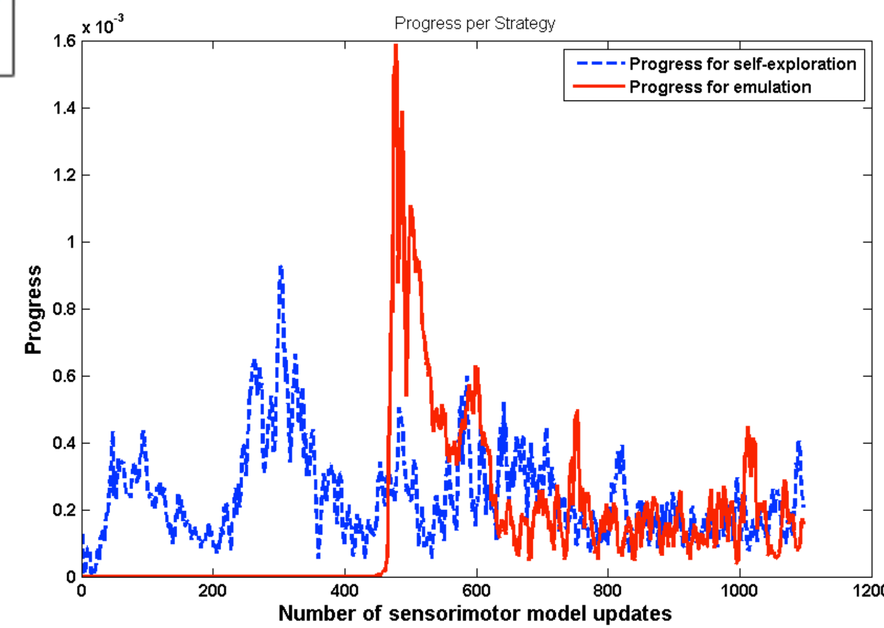
Two-layers of LP—based intrinsically motivated learning:

1) **Active choice self-exploration vs. imitation**
2) **If self-exploration: active goal selection**

(Moulin-Frier, Nguyen and Oudeyer, Frontiers in Cognitive Science, 2014)

Emergent developmental stages





0-3 mo:
squeals, growls,
yeals ...

3-7 mo:
quasi-vowels

7-10 mo:
language-independent
proto-syllables

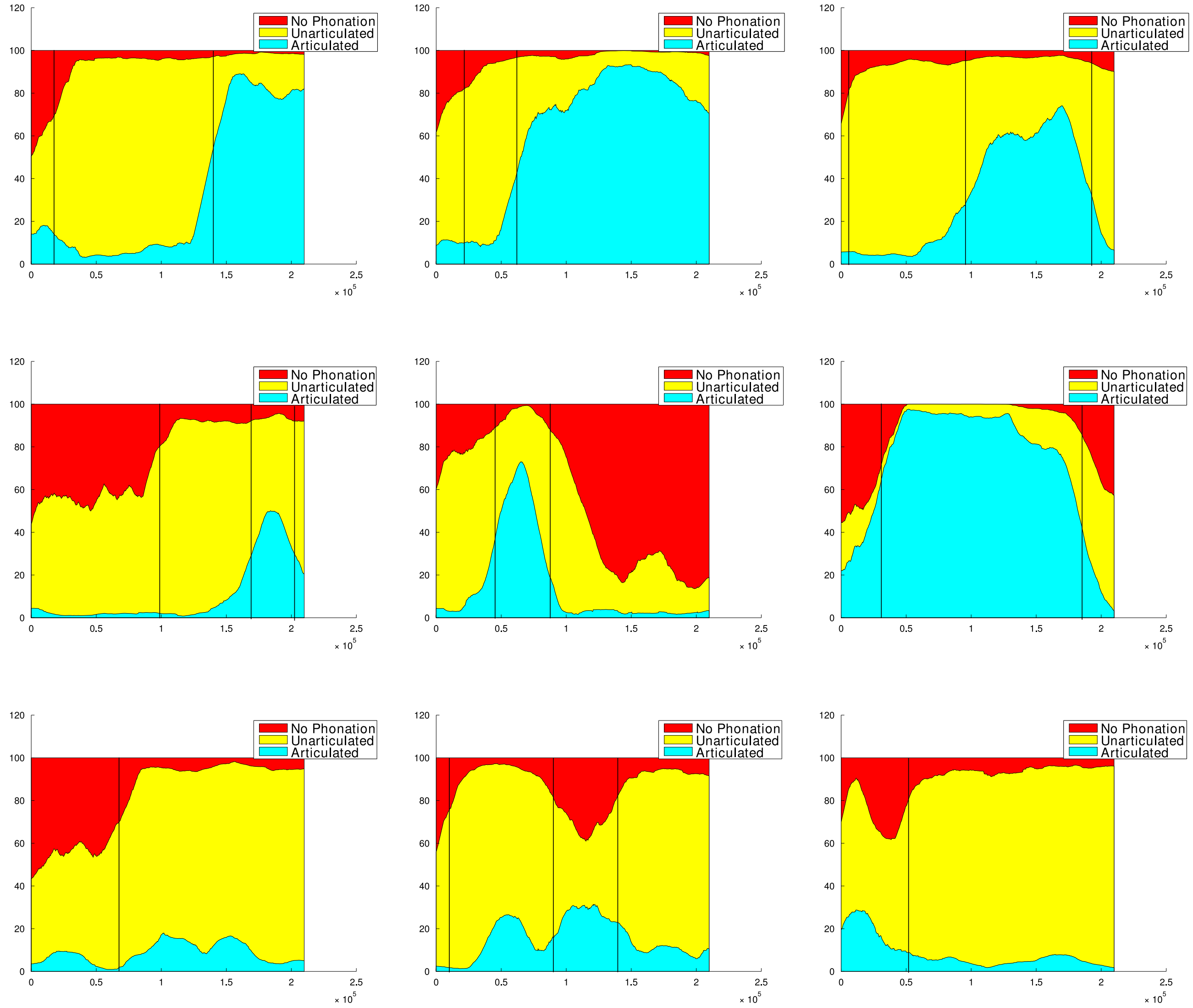
10 mo:
influence by
ambient language

12 mo:
first words

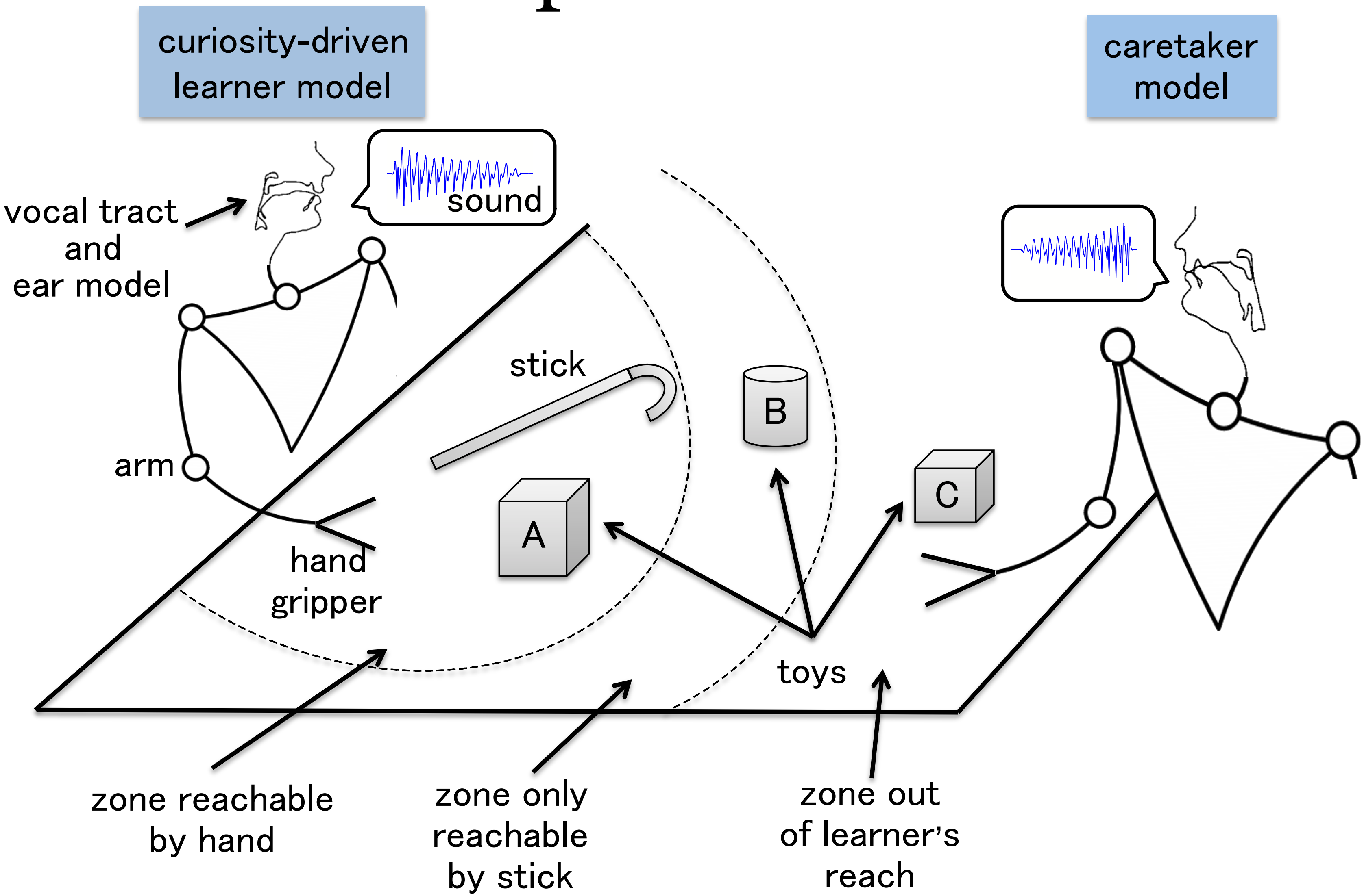
Approximate age

(Oller, 2000)

Regularities and diversity of individual developmental trajectories



Curiosity-driven discovery of language as a tool to manipulate the environment



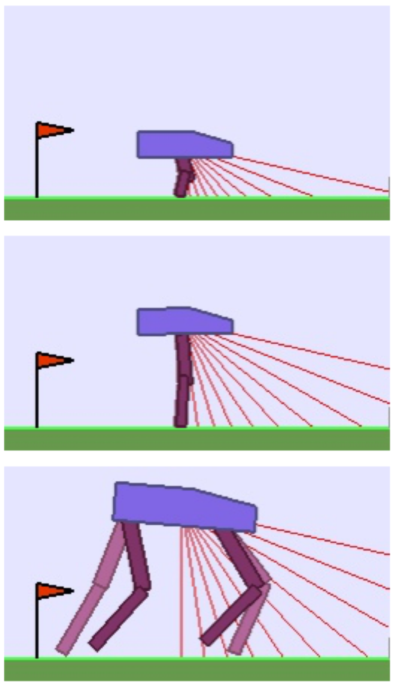
(Forestier and Oudeyer, CogSci 2017)

Teacher algorithms for curriculum learning of Deep RL
in continuously parameterized environments

Rémy Portelas[1], Cédric Colas[1], Katja Hofmann[2], Pierre-Yves Oudeyer[1]
[1]Inria (FR) [2]Microsoft Research (UK)
CoRL 2019

Microsoft Research - Inria
JOINT CENTRE

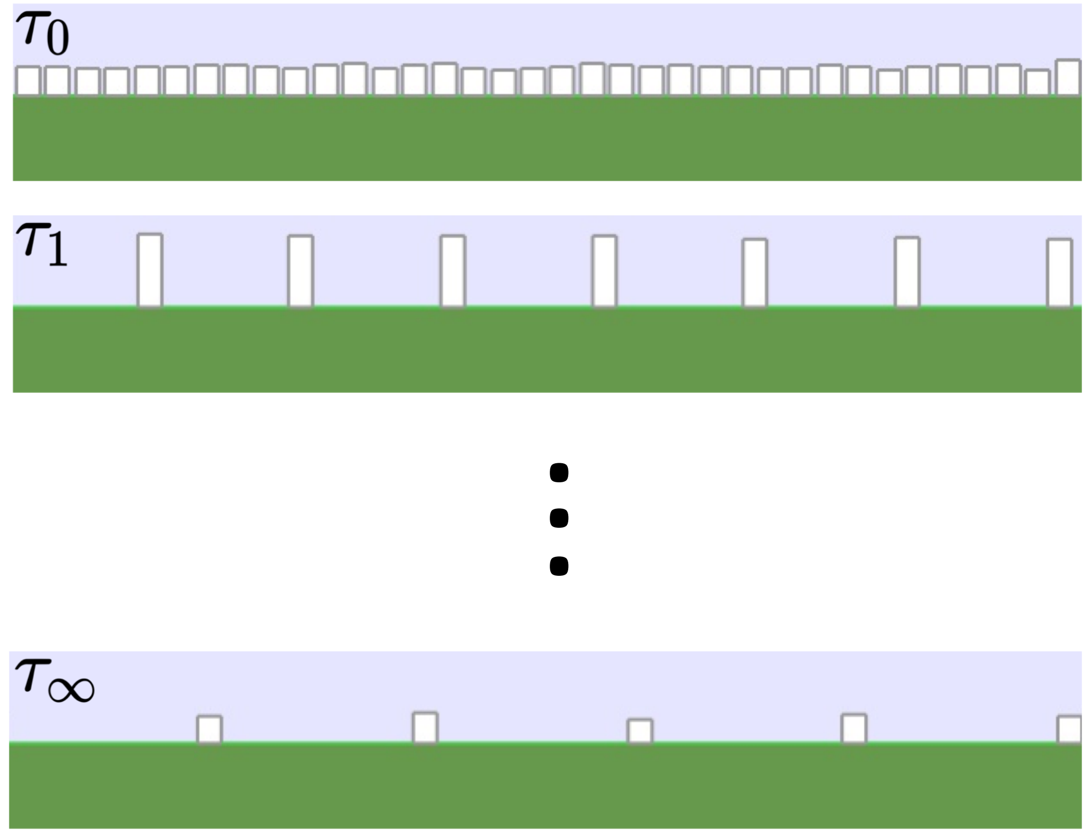
Learners



VS

Continuous set of tasks/envs

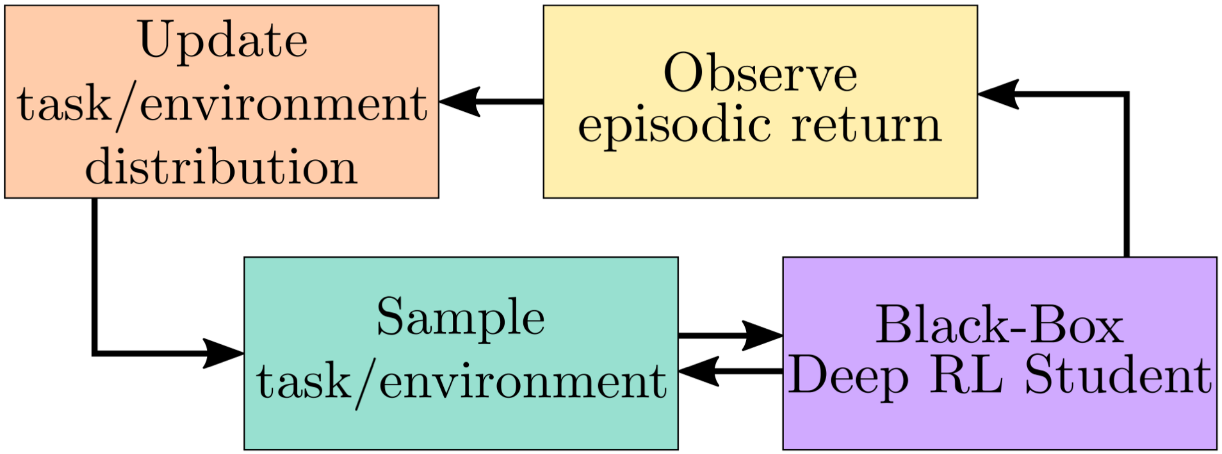
(through procedural generation)



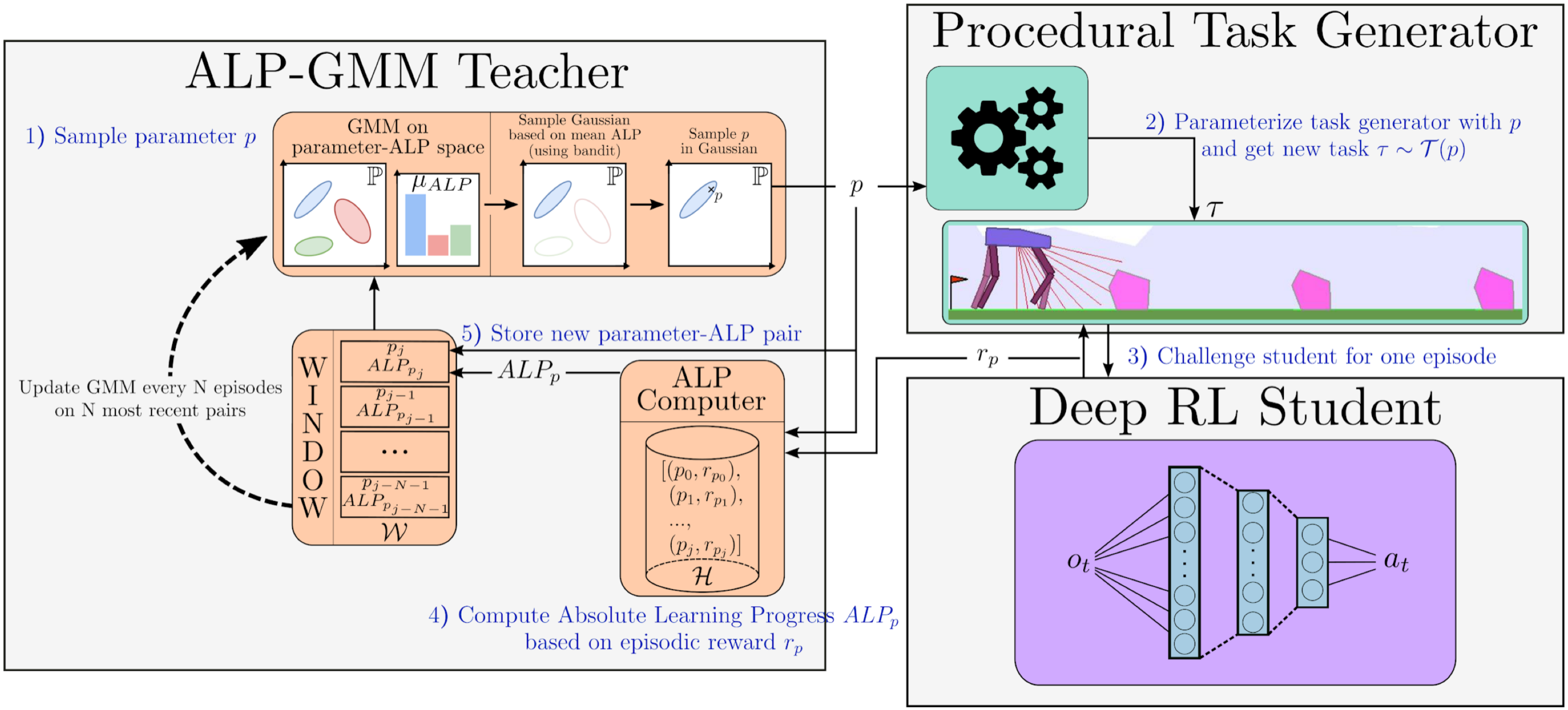
Methods - The CTS Framework

(CTS: **C**ontinuous **T**eacher-**S**tudent)

- The teacher samples parameters mapping to *distributions* of tasks/envs
  ⇒ creates ***a curriculum where tasks/envs distributions evolve***
- The Deep RL Student is a black-box
- The parameter space may contain:
  - unfeasible subspaces
  - irrelevant dimensions
  - non-linear difficulty

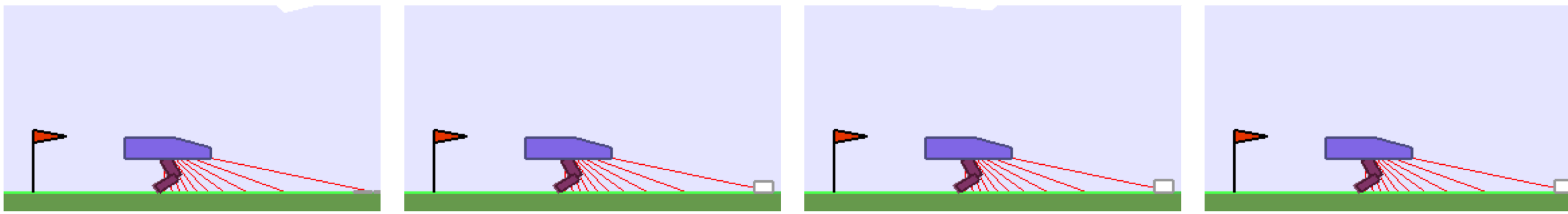


ALP-GMM: sample tasks/envs distributions that maximize absolute learning progress

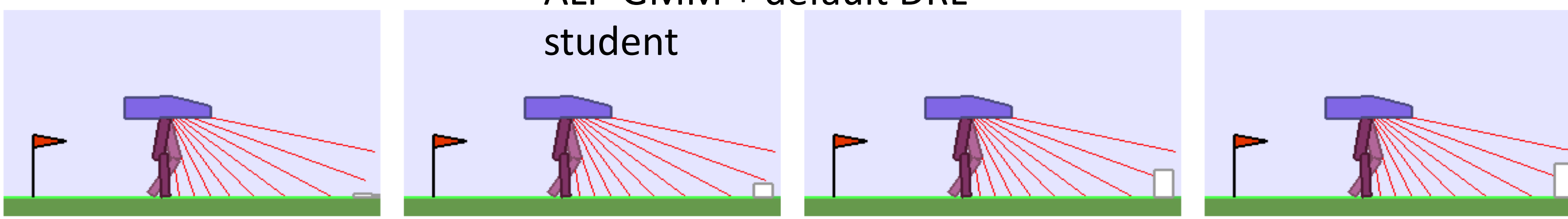


Example of mastered tasks after training

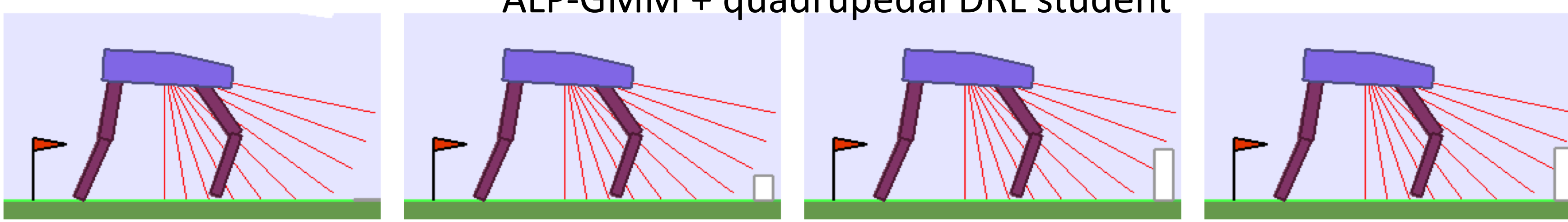
ALP-GMM + short DRL student



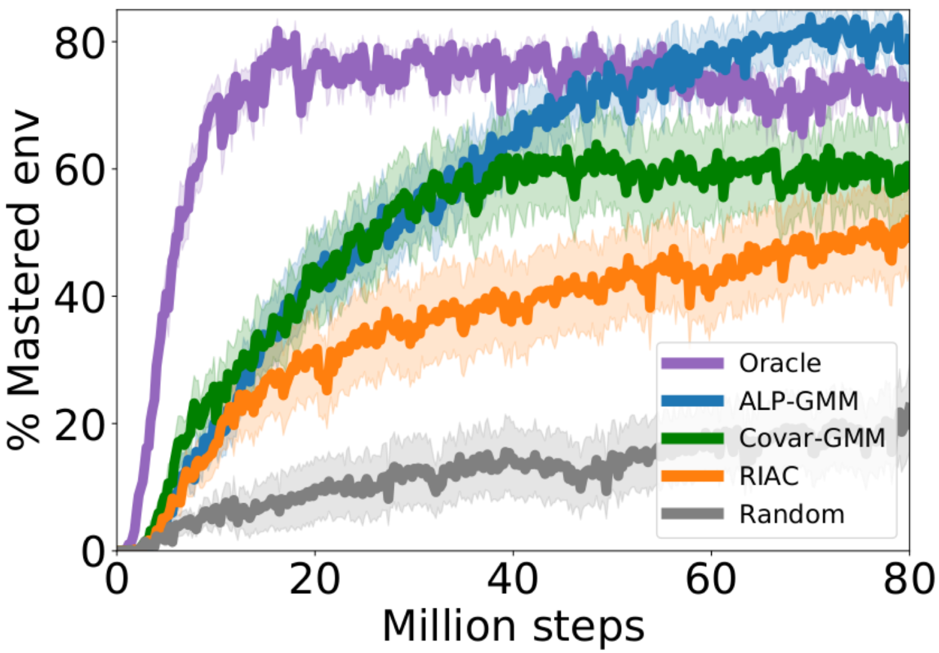
ALP-GMM + default DRL
student

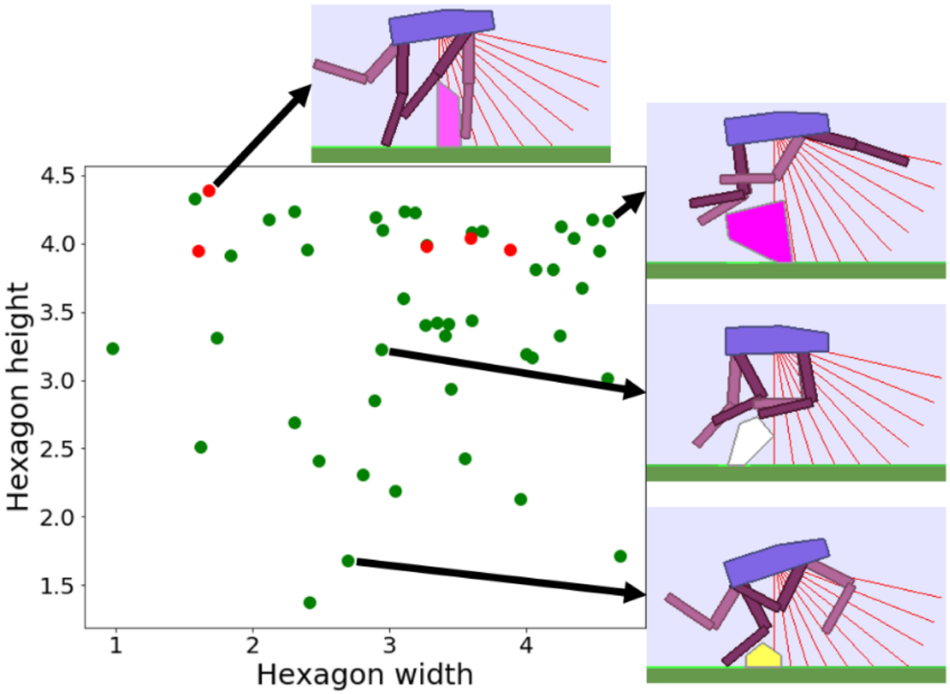


ALP-GMM + quadrupedal DRL student



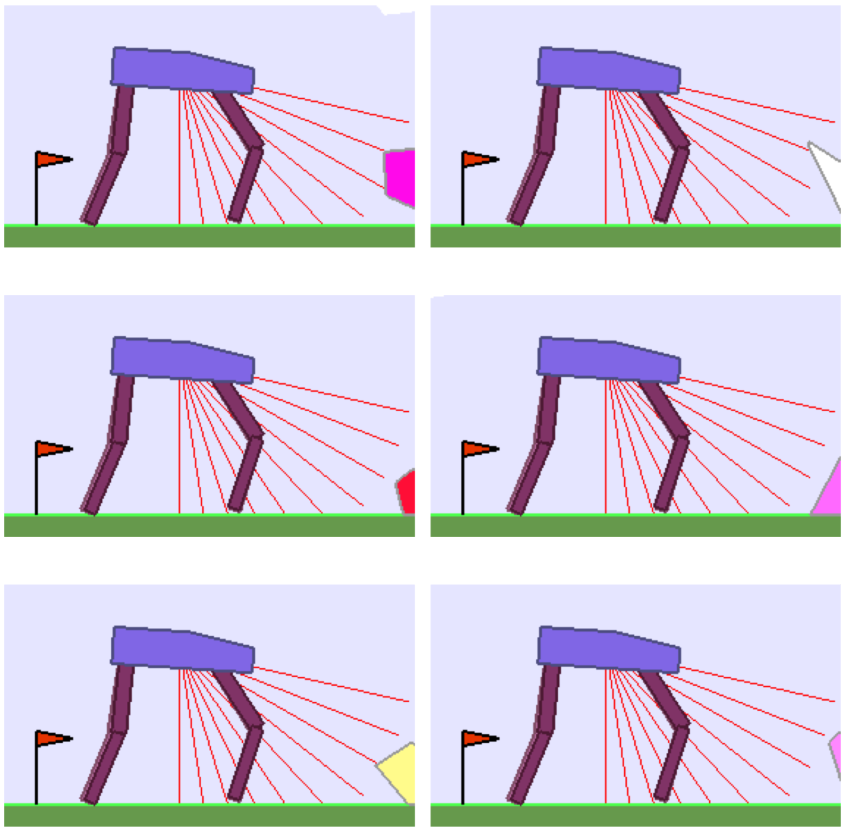
Performance analysis on Hexagon Tracks





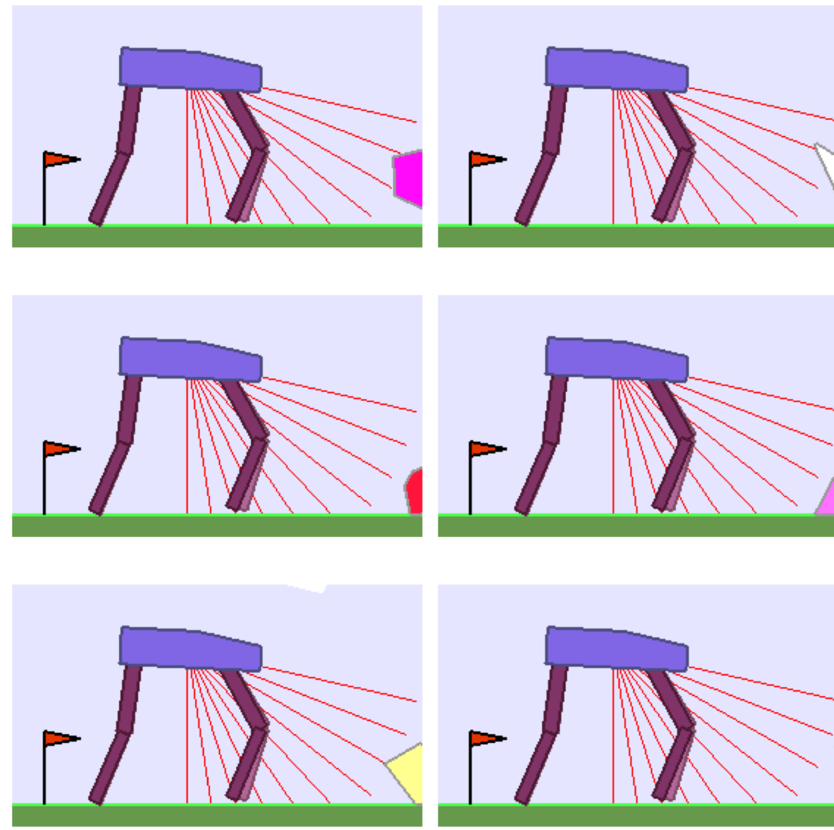
ALP-GMM

Good **generalization** to **diverse** obstacles



Random

Poor learning and generalization



https://github.com/flowersteam/teachDeepRL

Applications in educational technologies

Technologies for fostering efficient learning and intrinsic motivation

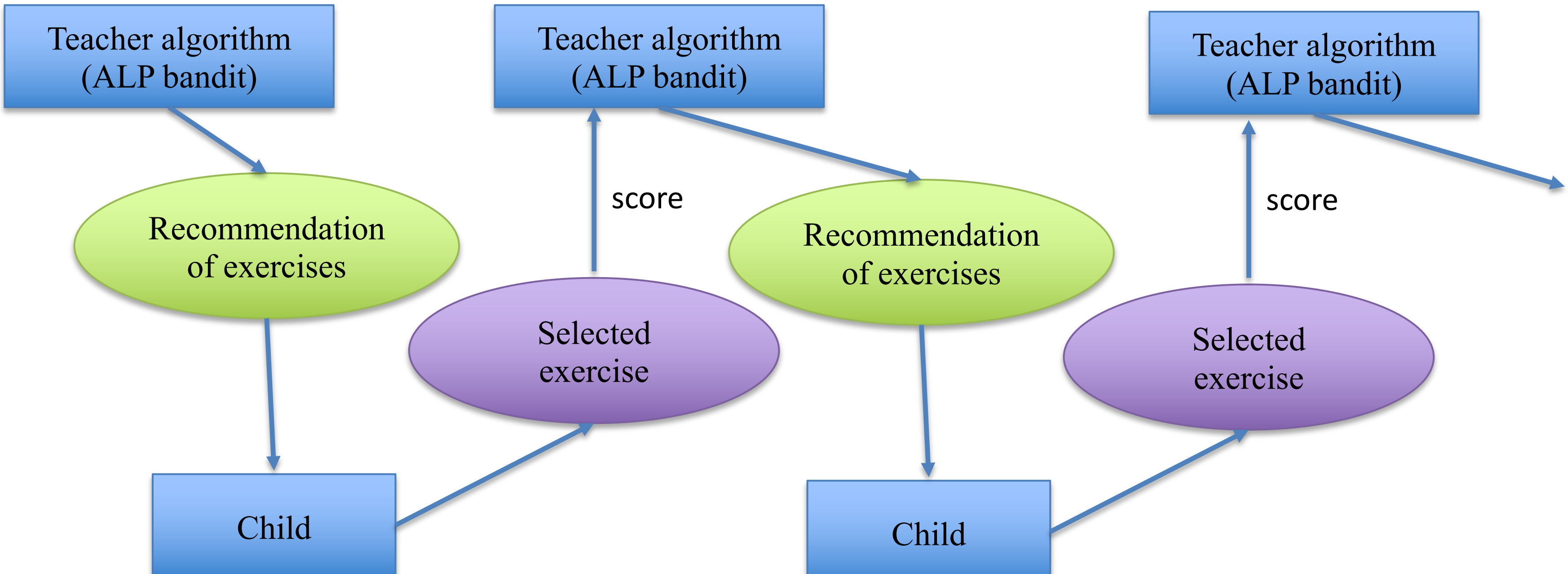
KidLearn project:

Personalization of teaching sequences (curriculum) in Intelligent Tutoring Systems

(Clement et al., Journal of Educational Data Mining, 2015; in prep.)

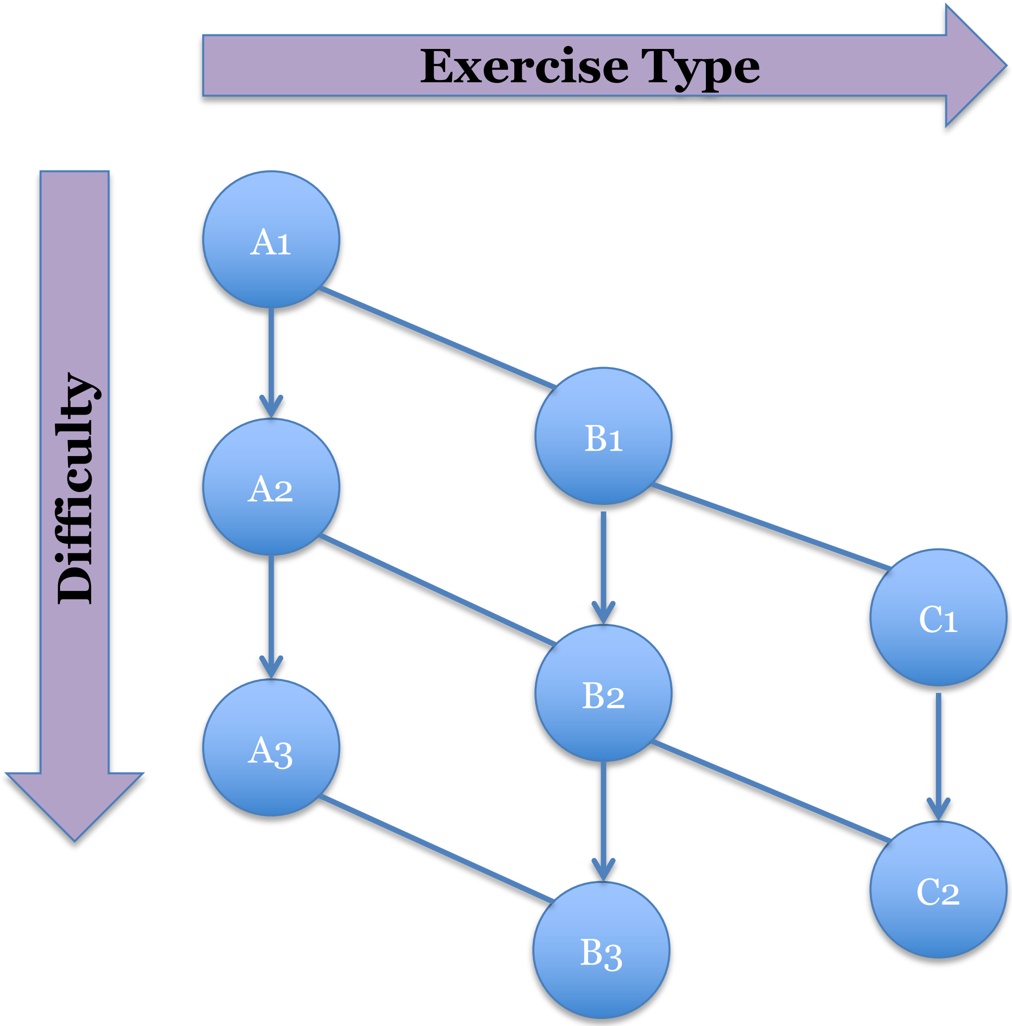
- Experiments with > 1000 children in more than 30 schools in Aquitaine

ZPDES-CO algorithm:
ALP + warm-start graph + final choice by child



(Clement et al., Journal of Educational Data Mining, 2015; in prep.)

ZPDES-CO algorithm

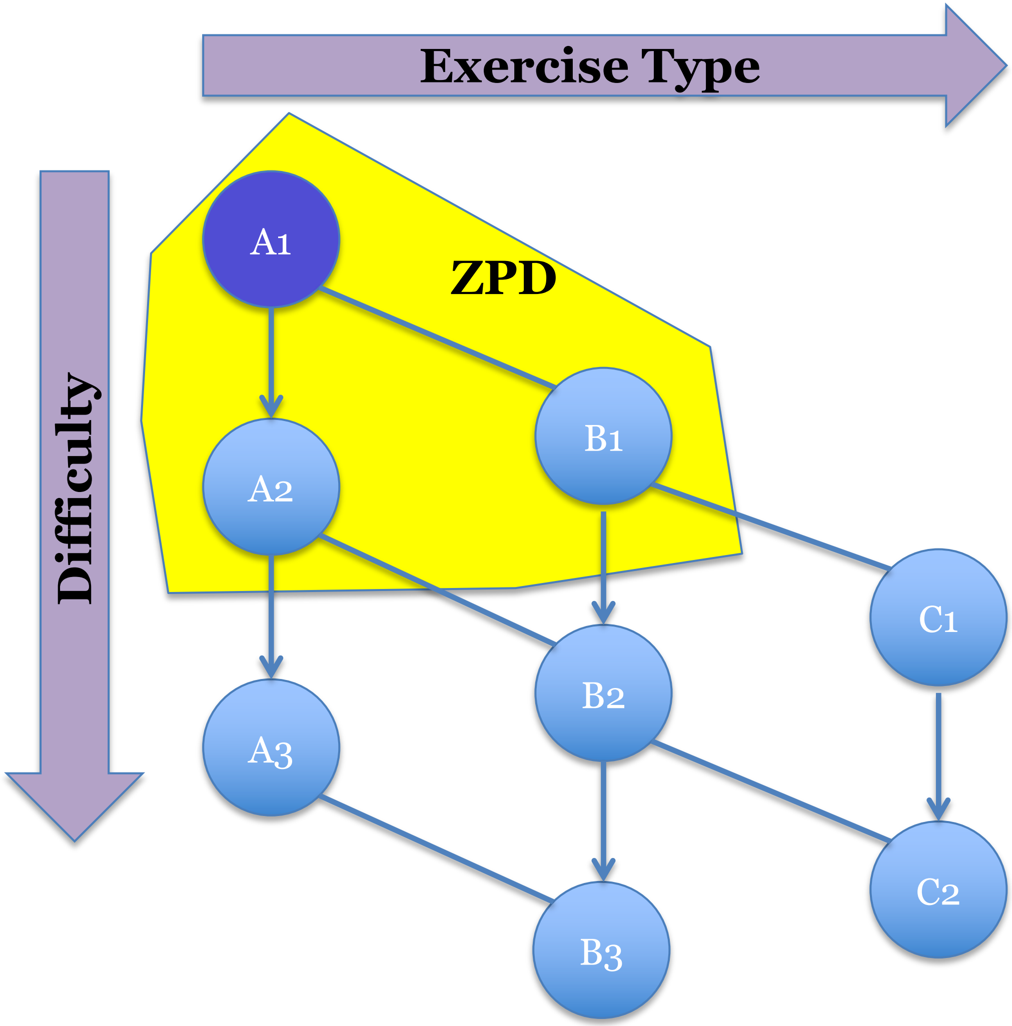


Inside **the zone of proximal development** choose exercises **<u>stochastically</u>** according to the learning progress

Exercise Type:
- aiming at different KC;
- or presented in a different modality;

(Clement, Roy, Oudeyer and Lopes, 2015, Journal of Educational Data Mining)

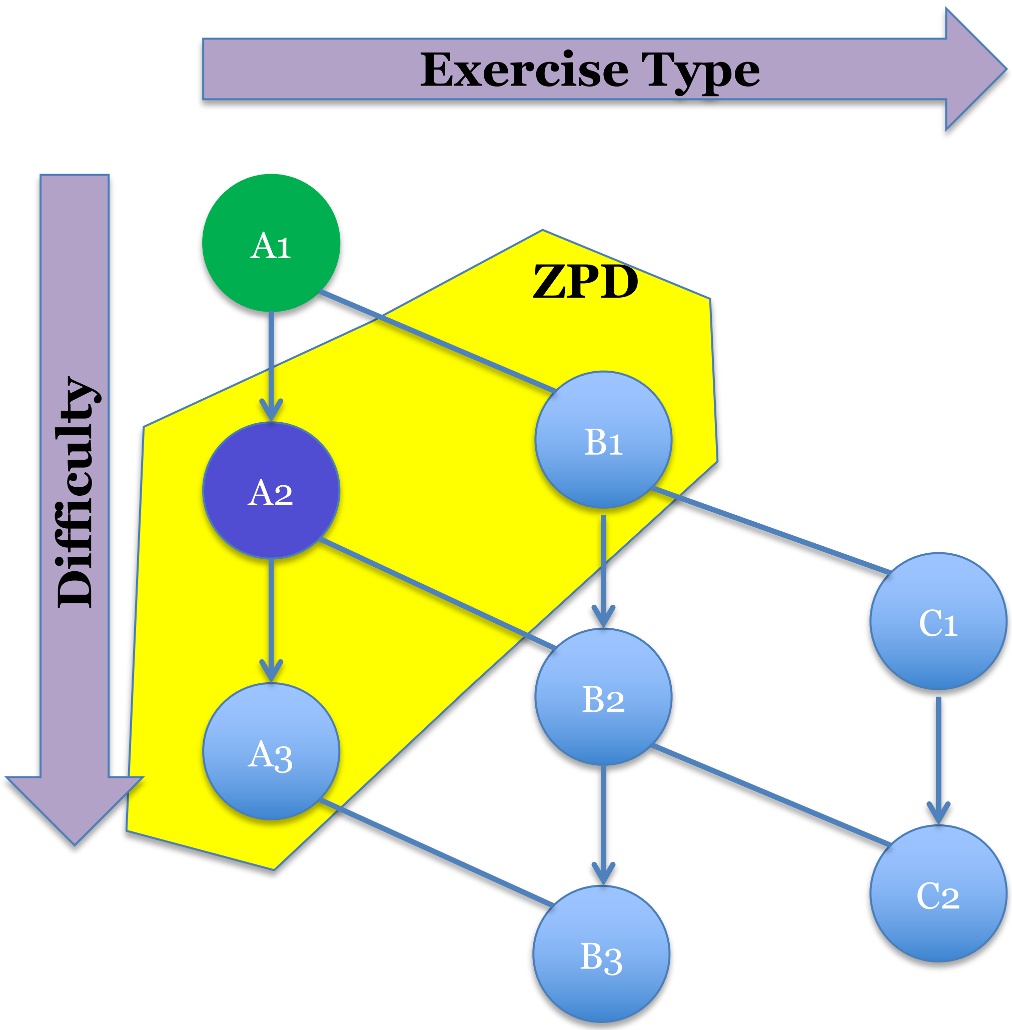
ZPDES-CO algorithm



Inside **the zone of proximal development** choose exercises **<u>stochastically</u>** according to the learning progress

A1 A2 B1

ZPDES-CO algorithm

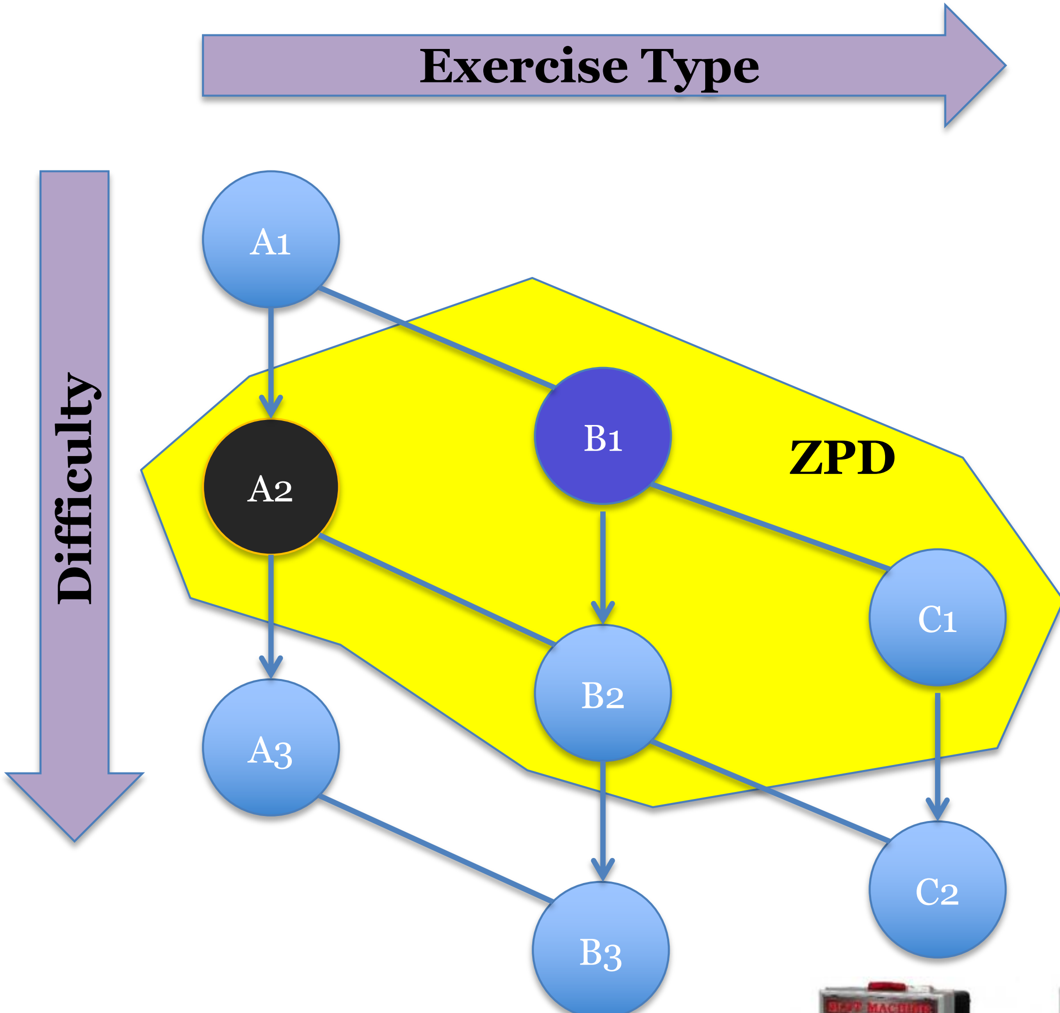


Inside **the zone of proximal development** choose exercises **stochastically** according to the learning progress

After being able to solve A1, extend the ZPD

A2 A3 B1

ZPDES-CO algorithm



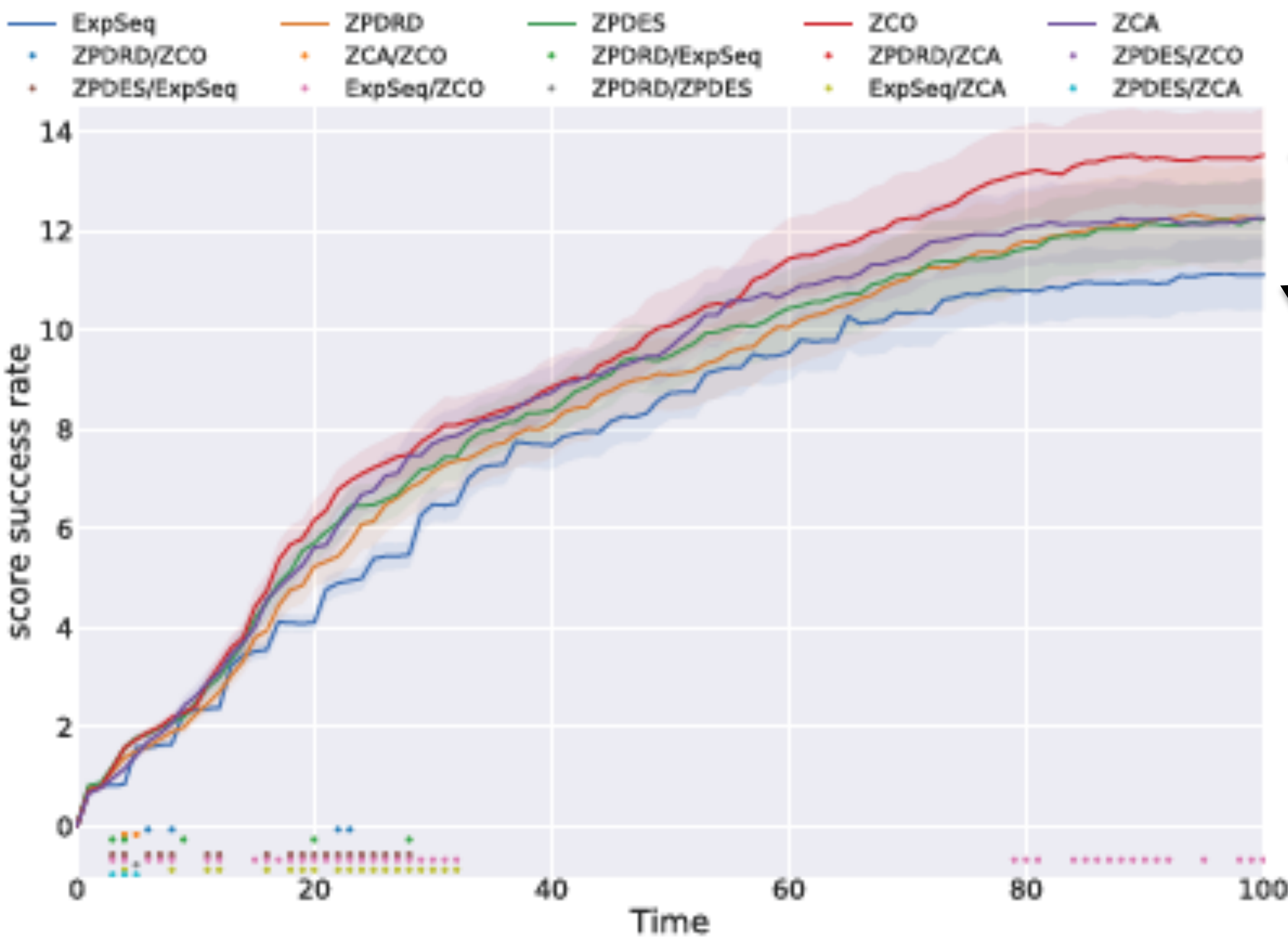
Inside **the zone of proximal development** choose exercises **<u>stochastically</u>** according to the learning progress

Always a probability of choosing other exercise types due to:

- individual characteristics
- problems in the knowledge graph

A2 B1 B2 C1

Learning impact

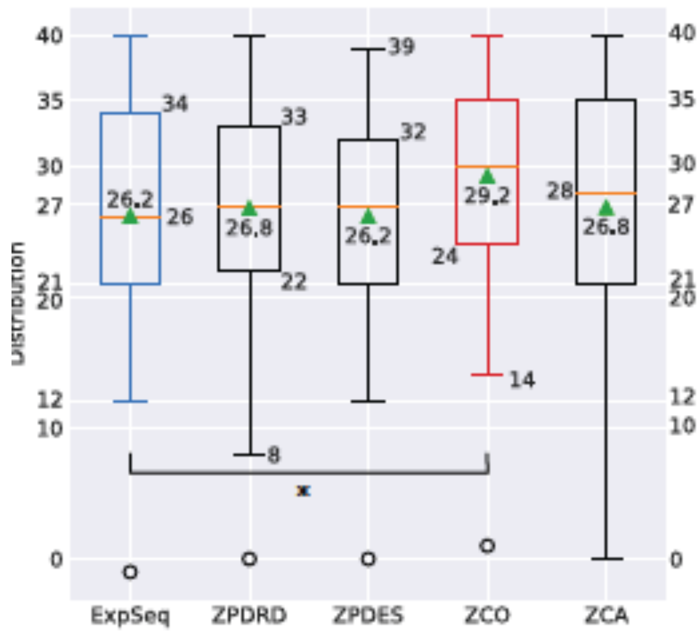


ZPDES-CO algorithm

Oracle algorithm (Pedagogical expert)

Motivational impact during learning sessions

Intrinsic motivation score session S3 (IMI questionnaire)



Oracle algorithm
(Pedagogical expert)

ZPDES-CO algorithm

Take away

<u>Fundamental role of spontaneous developmental exploration</u>

- Autonomous goal exploration
- Driven by empirical learning progress measured at various scales of time and space

Scales to
real world
(high-dimensions,
limited time,
distractors)

Organizes
developmental
trajectories
+
Enables discoveries
(tool use, language)

Can be used to guide
human edTech design

(Gottlieb and Oudeyer, Nature Rev. Neurosc., 2018.; Oudeyer et al., 2016;)

Developmental autonomous learning

Thanks to:

PhDs/Postdocs/engineers: A. Baranes, F. Benureau, B. Clément, C. Colas, S. Forestier, P. Fournier, M. Lapeyre, A. Laversanne-Finot, Y. Mollard, C. Moulin-Frier, M. Nguyen, A. Péré, R. Portelas, P. Rouanet.

Senior colleagues: F. Kaplan, M. Lopes, O. Sigaud, J. Gottlieb, L. Smith, V. Hafner, H. Sauzéon, M. Chetouani, C. Kidd, L. Rat-Fisher.

Funding/Sponsors:

Inria informatics mathematics

erc European Research Council

ENSTA ParisTech

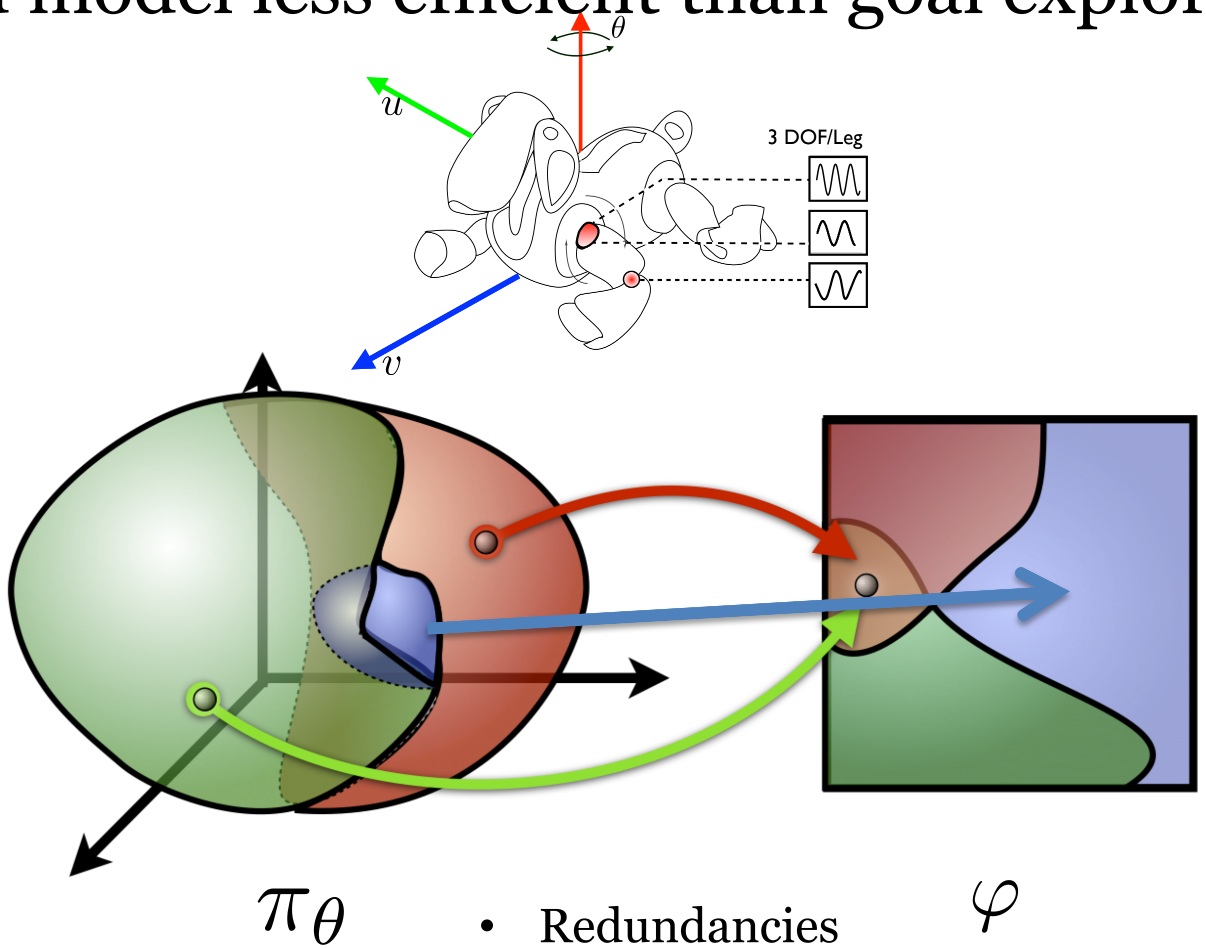
HFSP

Région Nouvelle-Aquitaine

Microsoft Research - Inria JOINT CENTRE

Microsoft Research

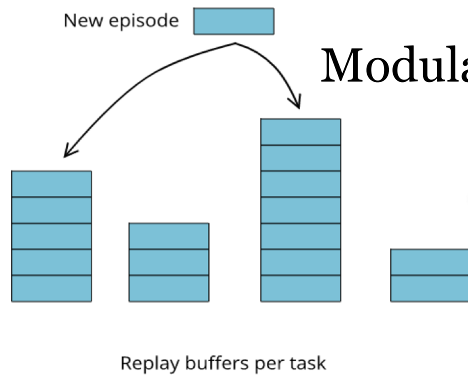
Why is (curiosity-driven) exploration of forward model less efficient than goal exploration?



- Redundancies
- Inhomogeneities

Forward model exploration: Knowing ***many*** ways to produce a ***few*** effects
Goal exploration: Knowing a ***few*** ways to produce ***many*** effects

Combining population-based and Deep-RL based IMGEPs

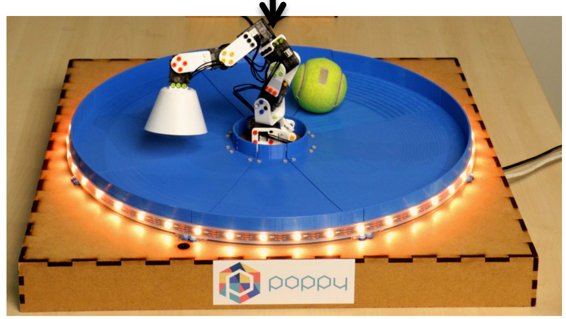


Modular replay buffers

Population-based IMGEP
(fast discoveries, episodic memory;
Forestier et al., 2016)

Unsupervised modular goal space learning
(Laversanne-Finot et al, 2018)

Monolithic Deep RL IMGEP
(good generalization; Colas et al., 2019)



(Extension of Colas et al., GEP-PG: Decoupling exploration and exploitation in Deep RL, ICML 2018)

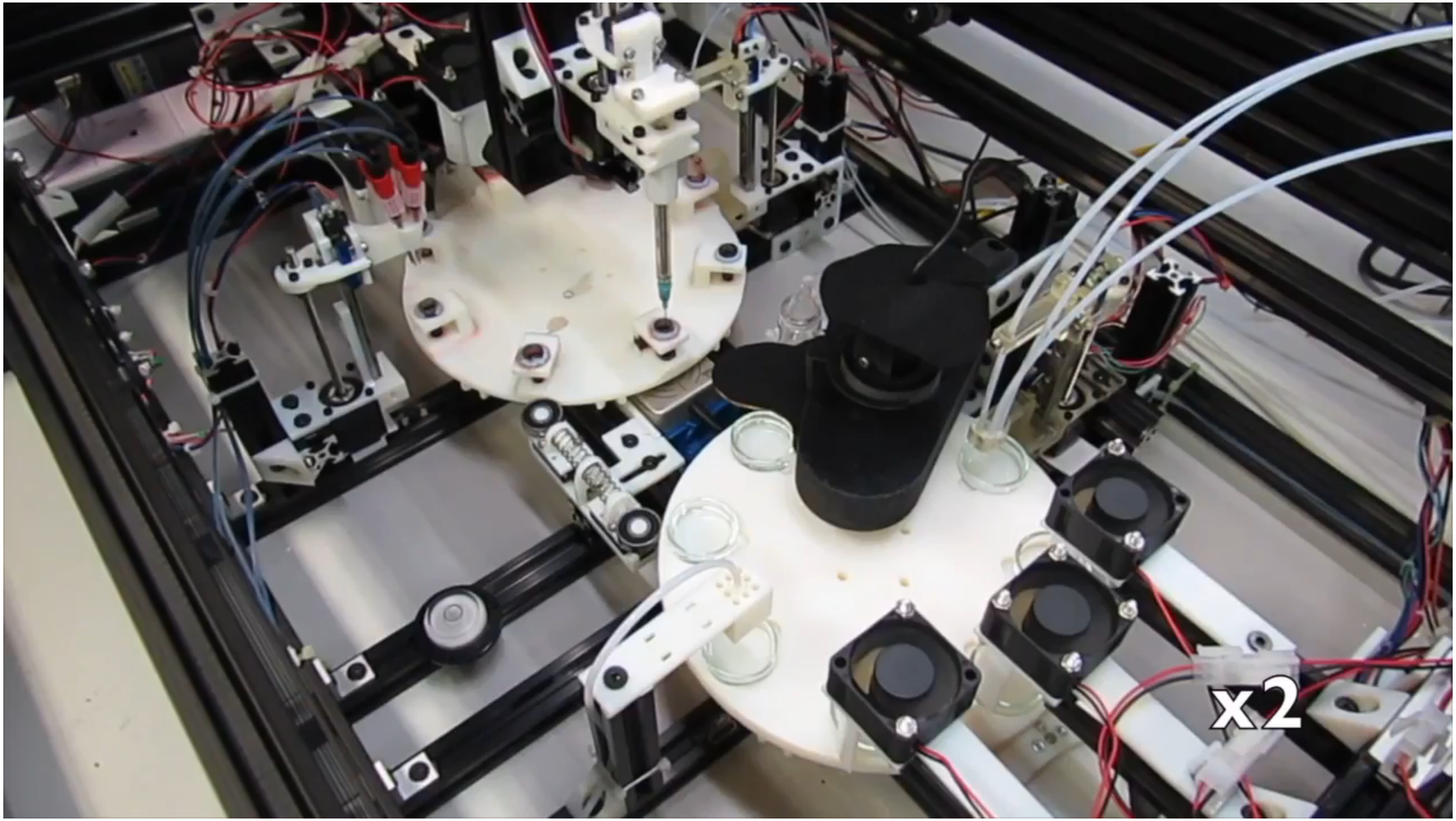
Curiosity applications beyond video games and robots:

Automated scientific discovery

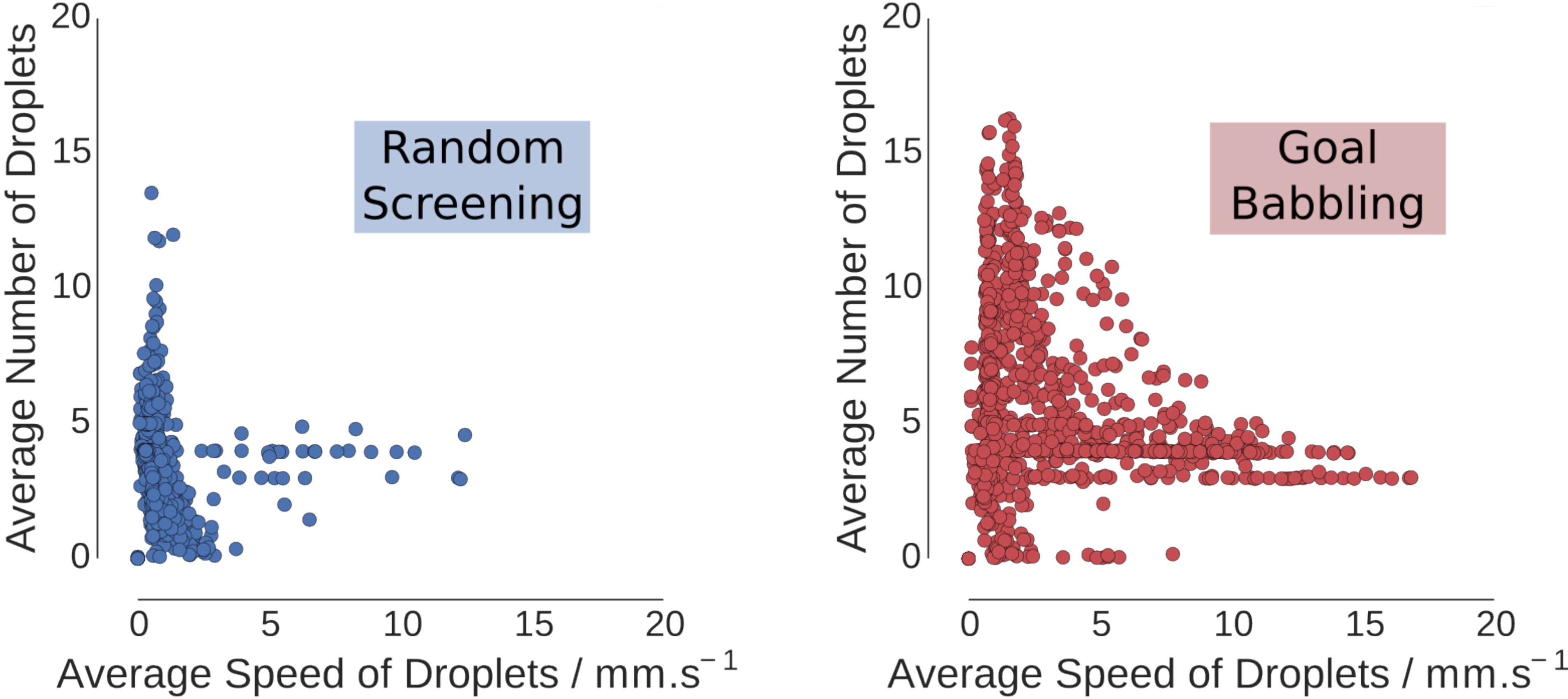
Oil-in-Water Droplets Self-Organization

Grizou et al. (2018) Exploration of Self-Propelling Droplets Using a Curiosity Driven Robotic Assistant, Arxiv/1904.12635, Cronin Lab, Univ. Glasgow.

Automatized robot experiments

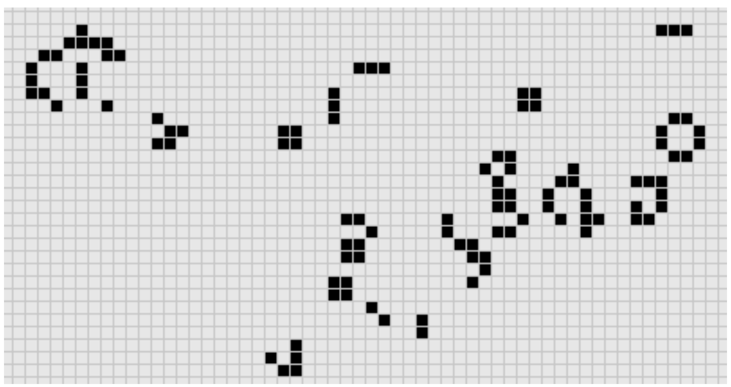


- 8 experiments running in parallel
- Specialized and stationary working stations
- Oils and surfactant handled separately

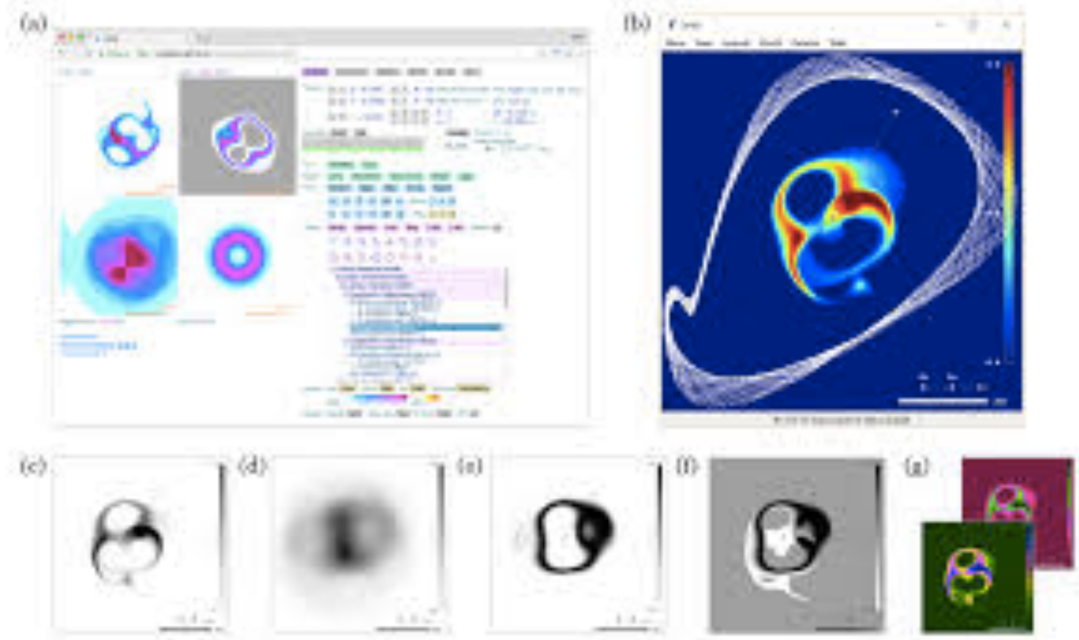


Random Screening
Goal Babbling
Average Number of Droplets
Average Speed of Droplets / mm.s⁻¹
0
5
10
15
20

Intrinsically motivated goal exploration in a continuous game of life



Discrete Game of Life

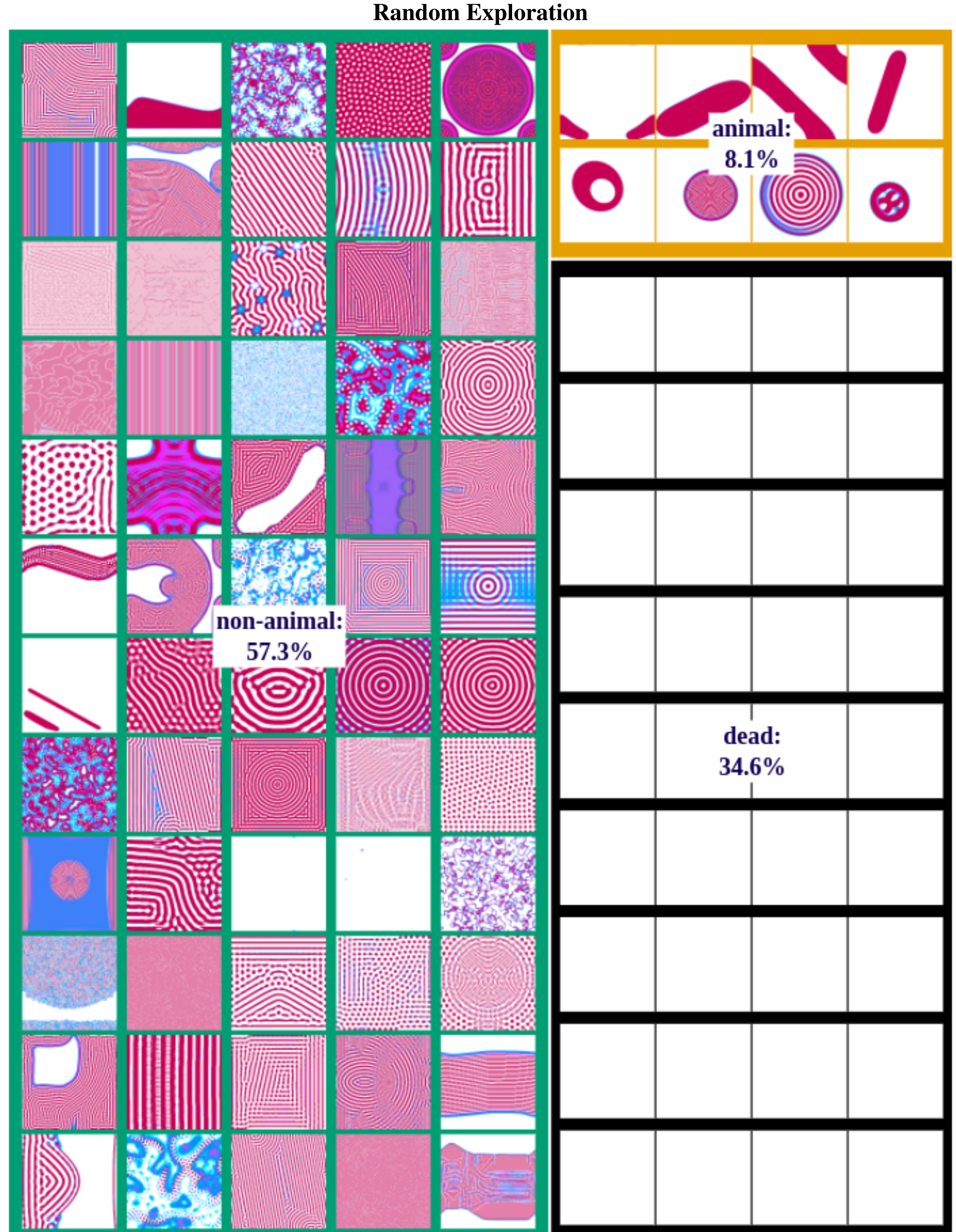


Continuous Game of Life
Lenia, Bert Chan (2018)

Reinke, C., Etcheverry, M., Oudeyer, P-Y. (in prep) Intrinsically Motivated Exploration for Automated Discovery of Patterns in Morphogenetic Systems

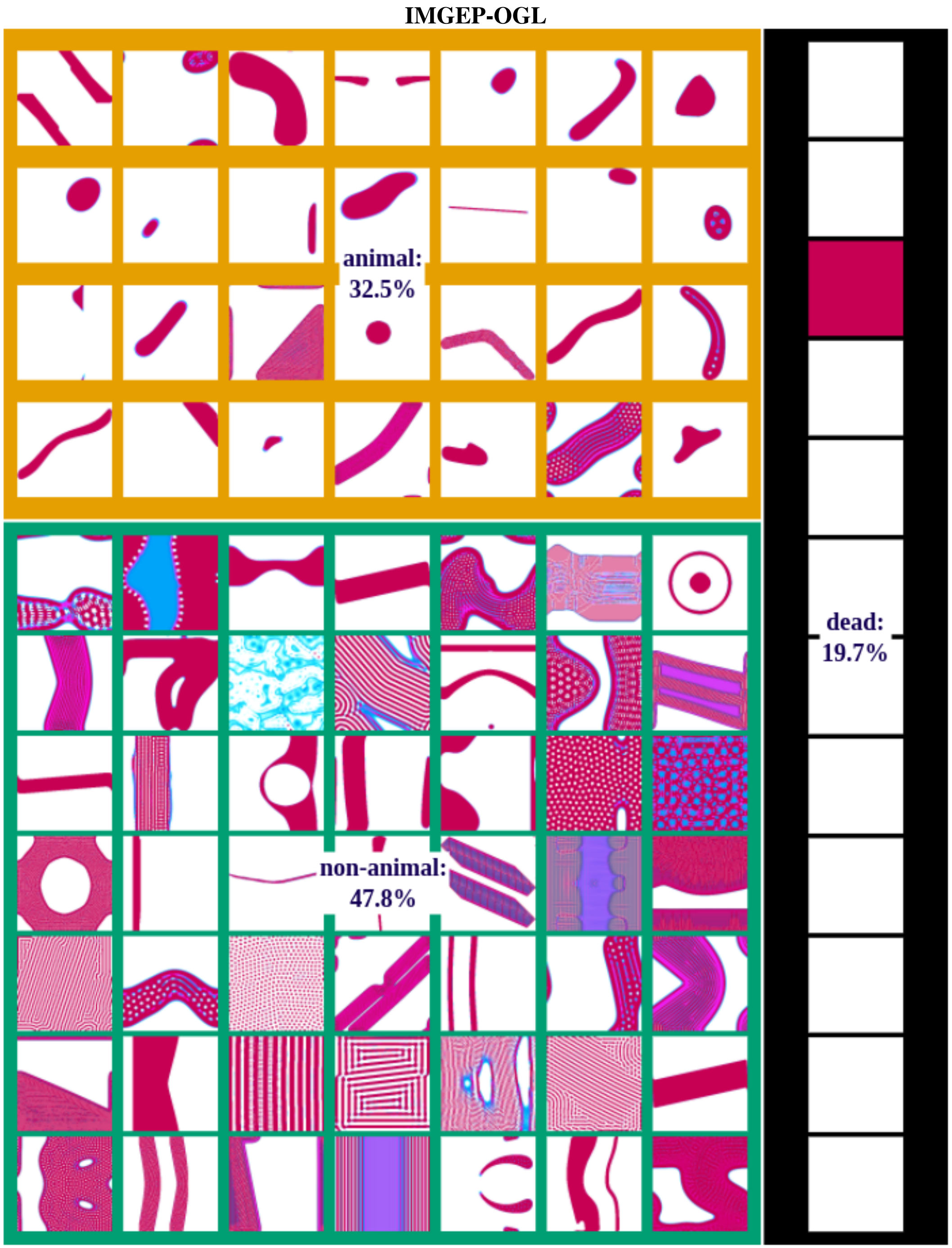
Random Experiments

(mostly dead
or uniform patterns)



Intrinsically Motivated Goal Exploration

32% spatially localized patterns (« animals »)



Goal exploration process

Intrinsically Motivated Exploration

for Automated Discovery of Patterns

in Morphogenetic Systems

Take away

<u>Fundamental role of spontaneous developmental exploration</u>

- Autonomous goal exploration
- Driven by empirical learning progress measured at various scales of time and space

Scales to
real world
(high-dimensions,
limited time,
distractors)

Organizes
developmental
trajectories
+
Enables discoveries
(tool use, language)

Can be used to guide
human edTech design

(Gottlieb and Oudeyer, Nature Rev. Neurosc., 2018.; Oudeyer et al., 2016;)

Developmental autonomous learning

Thanks to:

PhDs/Postdocs/engineers: A. Baranes, F. Benureau, B. Clément, C. Colas, S. Forestier, P. Fournier, M. Lapeyre, A. Laversanne-Finot, Y. Mollard, C. Moulin-Frier, M. Nguyen, A. Péré, R. Portelas, P. Rouanet.

Senior colleagues: F. Kaplan, M. Lopes, O. Sigaud, J. Gottlieb, L. Smith, V. Hafner, H. Sauzéon, M. Chetouani, C. Kidd, L. Rat-Fisher.

Funding/Sponsors:

informatics mathematics Inria

erc European Research Council

ENSTA ParisTech

HFSP

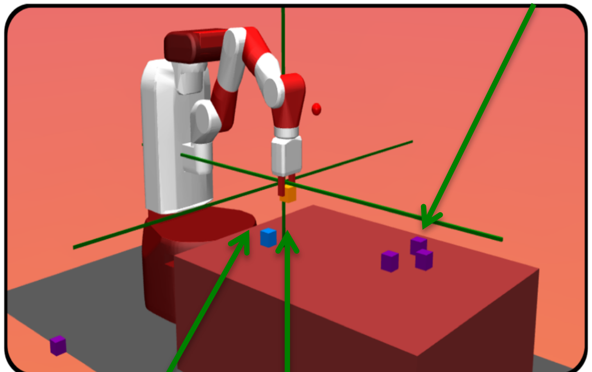
Région Nouvelle-Aquitaine

Microsoft Research - Inria JOINT CENTRE

Microsoft Research

CURIOUS: intrinsically motivated modular multi-goal Deep RL

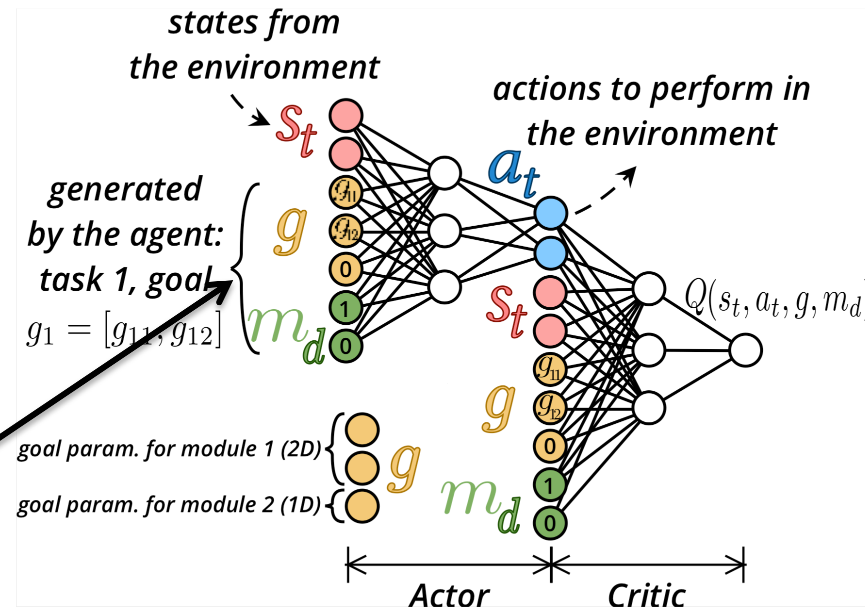
Distractors



Controllable objects

External world

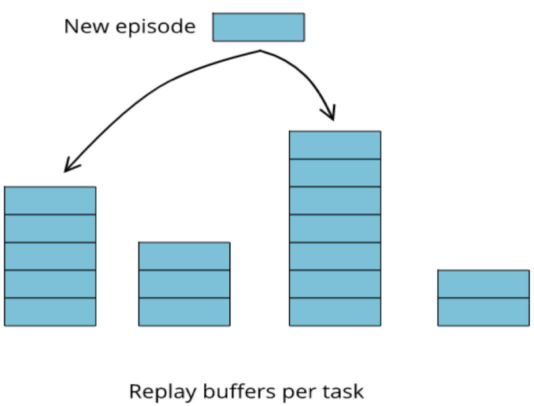
LP-based sampling of modules and goals



Modular UVFA

E.g. of modular goals:

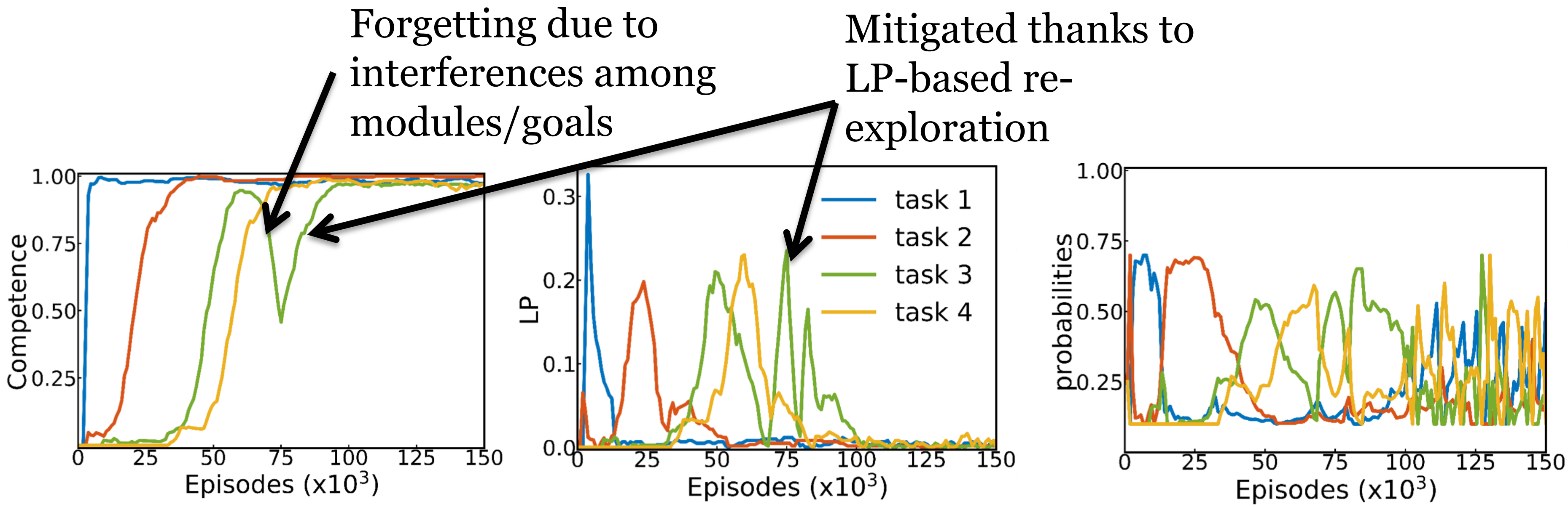
Move gripper to (x,y,z)
Pickandplace cube2 at (x,y,z)
Push(cube1) at position (x,y)
Stack cube1 over cube3 ...



Modular replay buffer

with hinsight learning
(module and goal substitution)

(Colas et al., to appear at ICML19)



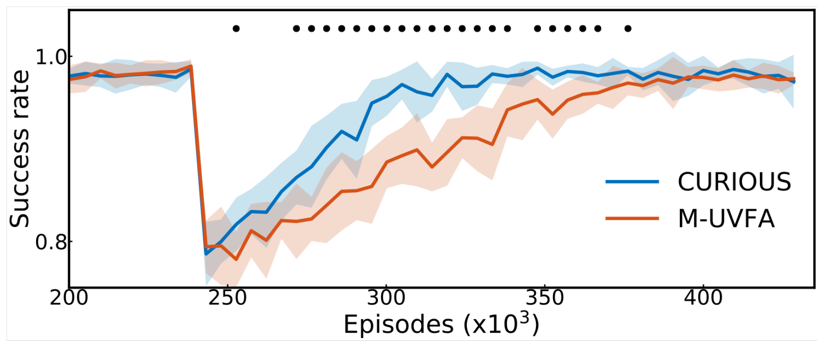
Competence

Absolute Learning Progress

Selection Probabilities

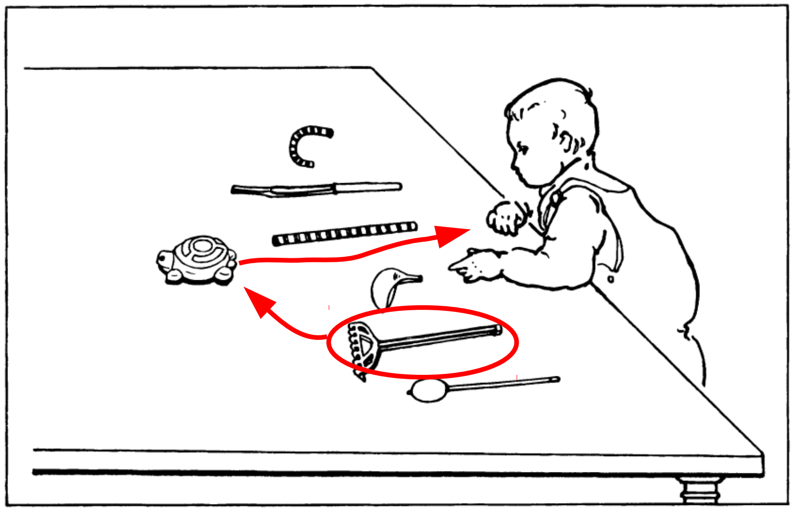
Recovery following a sensory failure.

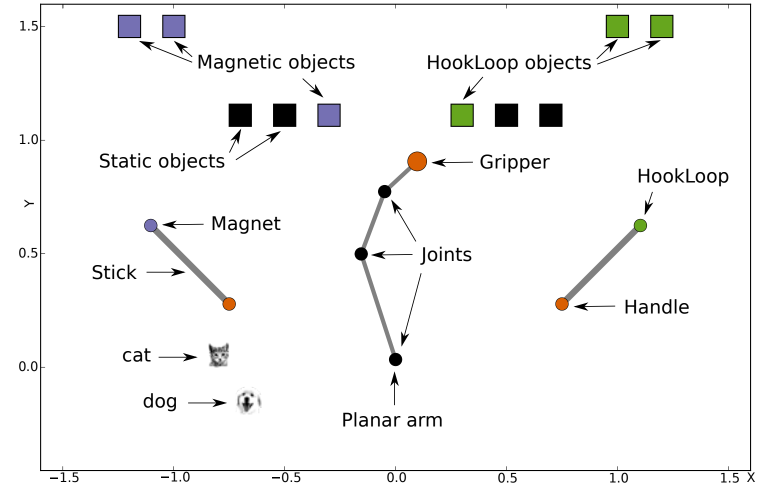
CURIOUS recovers 95 % of its original performance twice as fast as M-UVFA+HER.



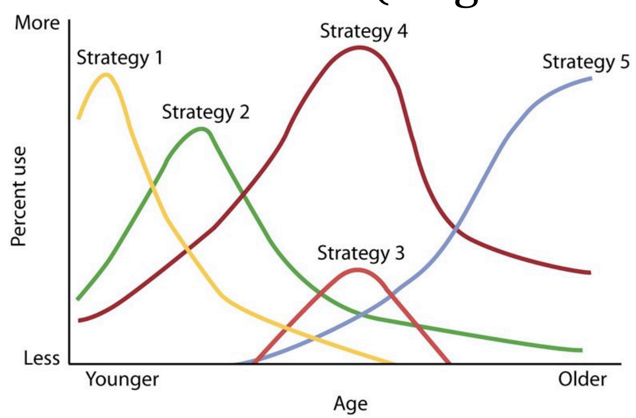
Deep RL based IMGEPs (Curious) vs. Population-based IMGEPs:
+ better generalization
- Slower initial discoveries

Modeling overlapping waves of tool use development

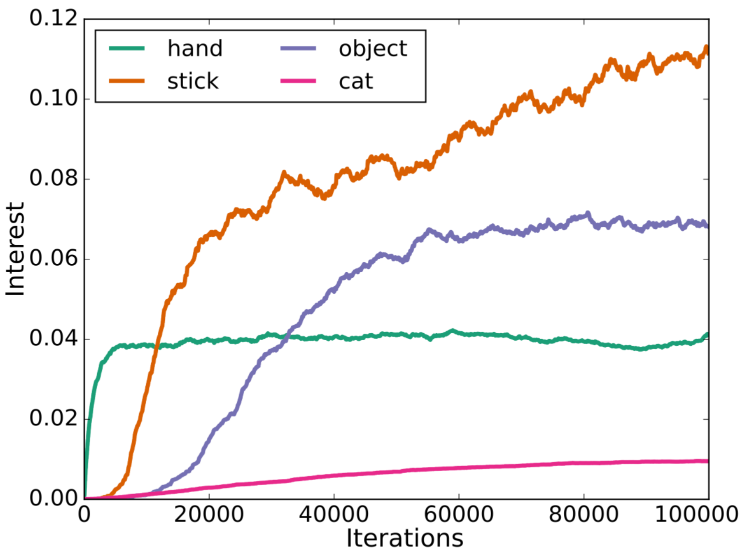




The overlapping-waves model (Siegler et al., 1996)



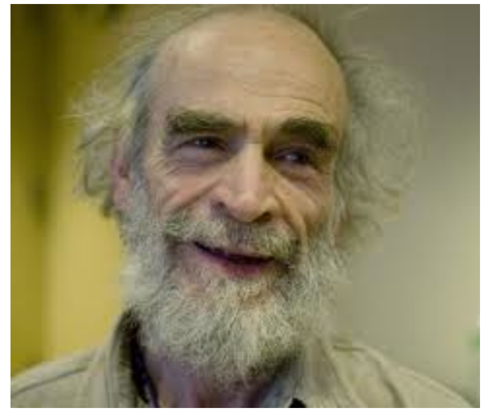
The overlapping waves model proposes that at any one age, children use multiple strategies; that with age and experience, they rely increasingly on more strategies (the ones with the higher numbers); and that development involves changes in use of existing strategies as well as discovery of new approaches.



(Forestier and Oudeyer, CogSci, 2016; ICDL-Epirob, 2016)

The Ergo-Robots (2012)

« Mathematics, a beautiful Elsewhere »
Fondation Cartier for Contemporary Art, Paris

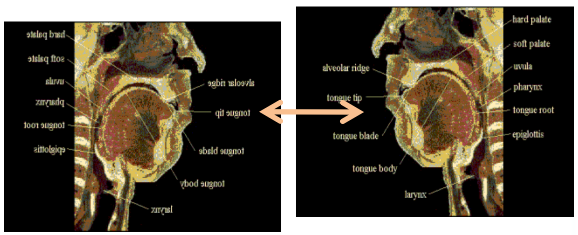


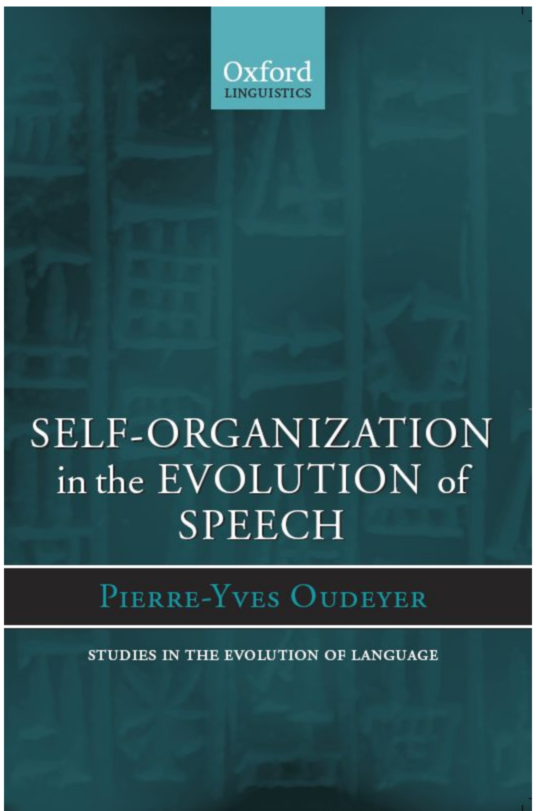
with
Mikhail Gromov
Mathematician

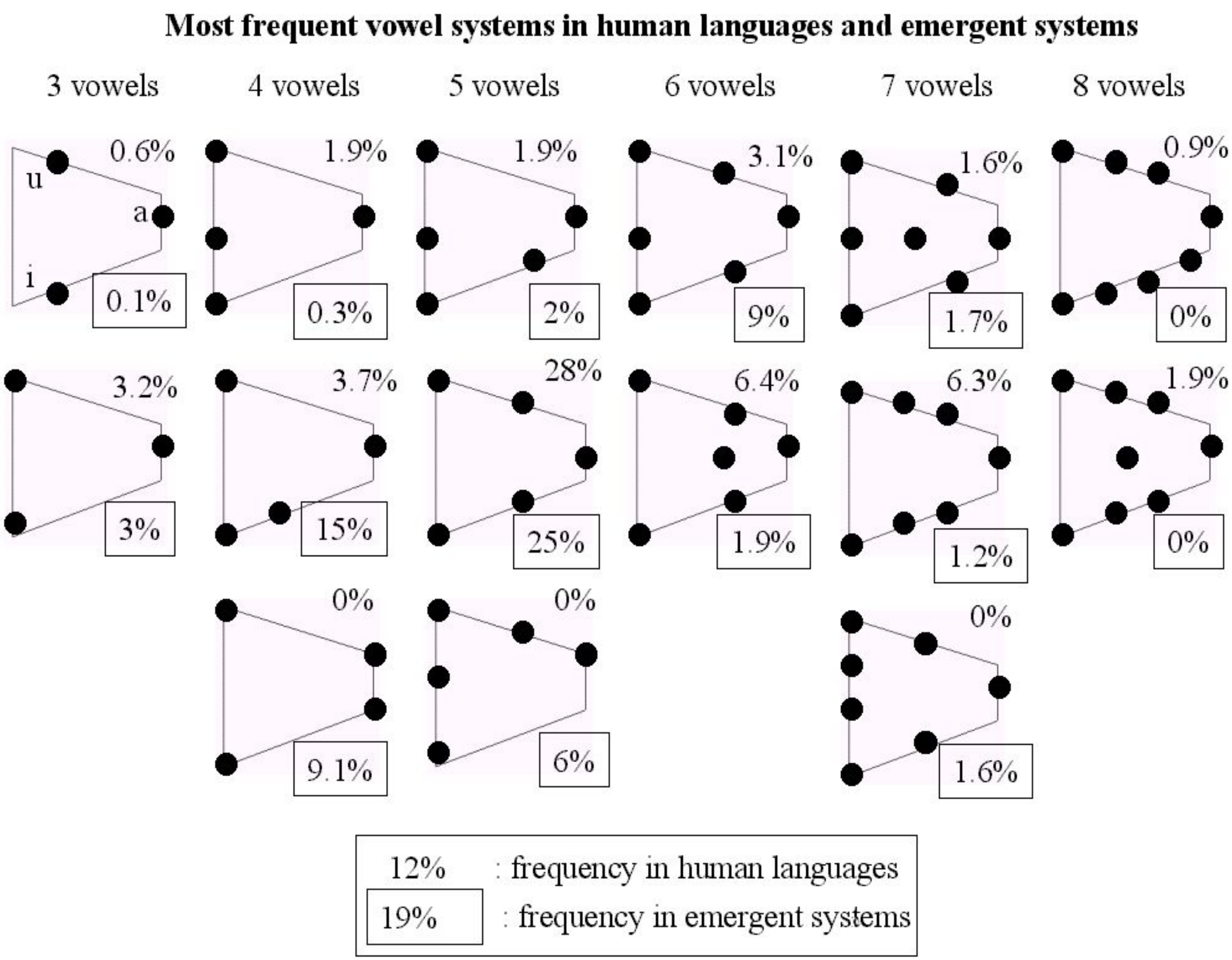
and
David Lynch
Film maker, artist

http://flowers.inria.fr/ergo-robots.php

Self-organization of culturally shared speech sounds



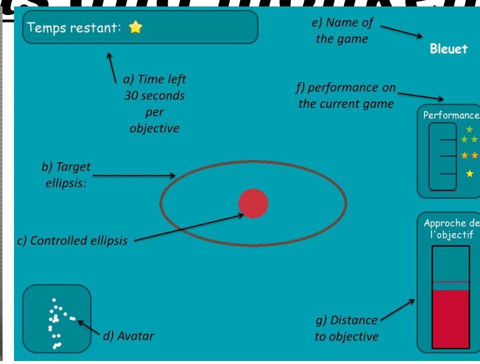


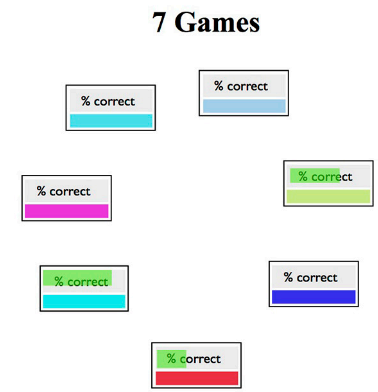


Models of the formation of speech sound systems in populations of individuals (de Boer,2001; Oudeyer, 2006/19; Moulin-Frier et al., 2011)

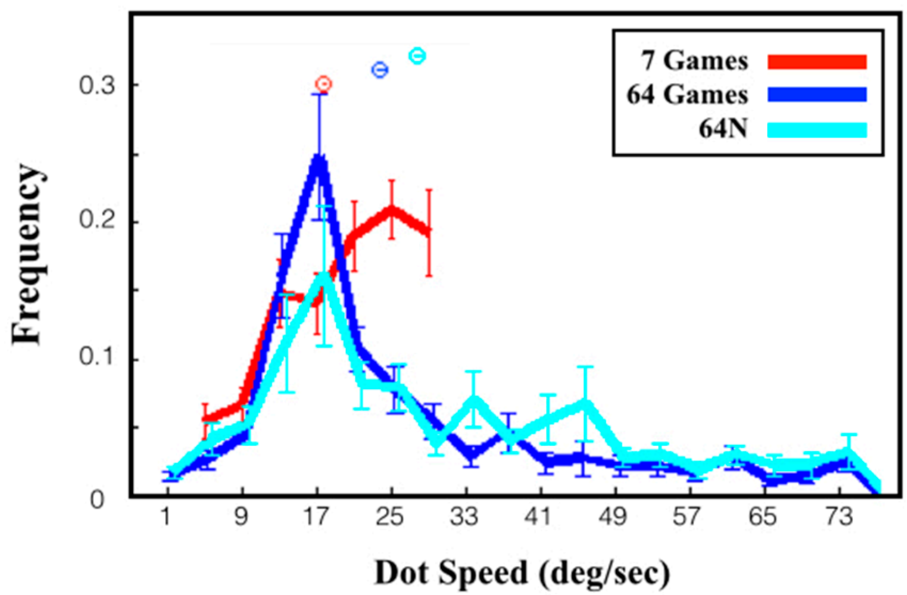
Future research: learning to represent experiments

Studying the structure of free exploration in humans and monkeys



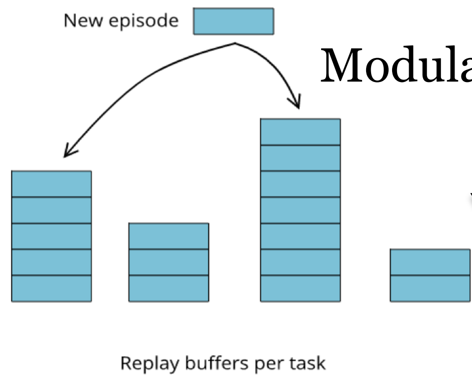


e examples of Kidd
rken/language



(Frontiers in Neuroscience, 2014; ICDL-Epirob 2014;
See also Scientific Reports, 2016; PNAS, 2017; Nature Reviews Neuroscience, in press)

Combining population-based and Deep-RL based IMGEPs

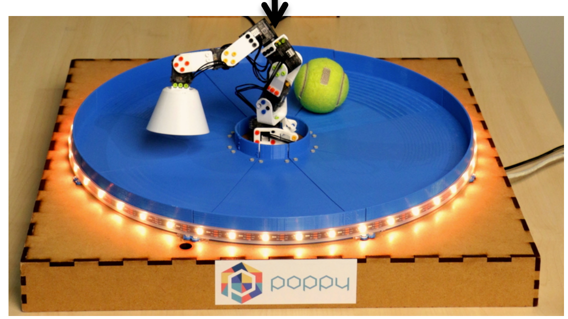


Modular replay buffers

Population-based IMGEP
(fast discoveries, episodic memory;
Forestier et al., 2016)

Unsupervised modular goal space learning
(Laversanne-Finot et al, 2018)

Monolithic Deep RL IMGEP
(good generalization; Colas et al., 2019)



(Extension of Colas et al., GEP-PG: Decoupling exploration and exploitation in Deep RL, ICML 2018)

Related to various research lines

Psychology (1940-60)
(Berlyne, White, Kaga, Festinger, …)

Theoretical biology and cognitive modeling

Varela, Maturana
(autopoïesis, 1974)

Oudeyer, Kaplan et al.
(2003)

Theoretical machine learning and RL

Fedorov et al.
(active learning,
Optimal exp.
Design, 1972)

Andreae et al.
(novelty search
with RL, 1978)

Schmidhuber
(LP based RL, 1991)

Barto, Singh et al.
(IMRL, 2004)

Evolutionary computing

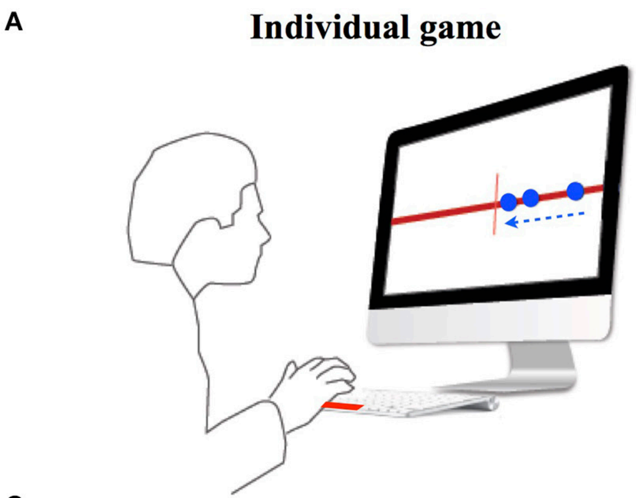
Stanley et al., 2008; Mouret, Doncieux et al.
(novelty search with GA/ES)

➔ Focus on modeling spontaneous curiosity-driven exploraition in humans

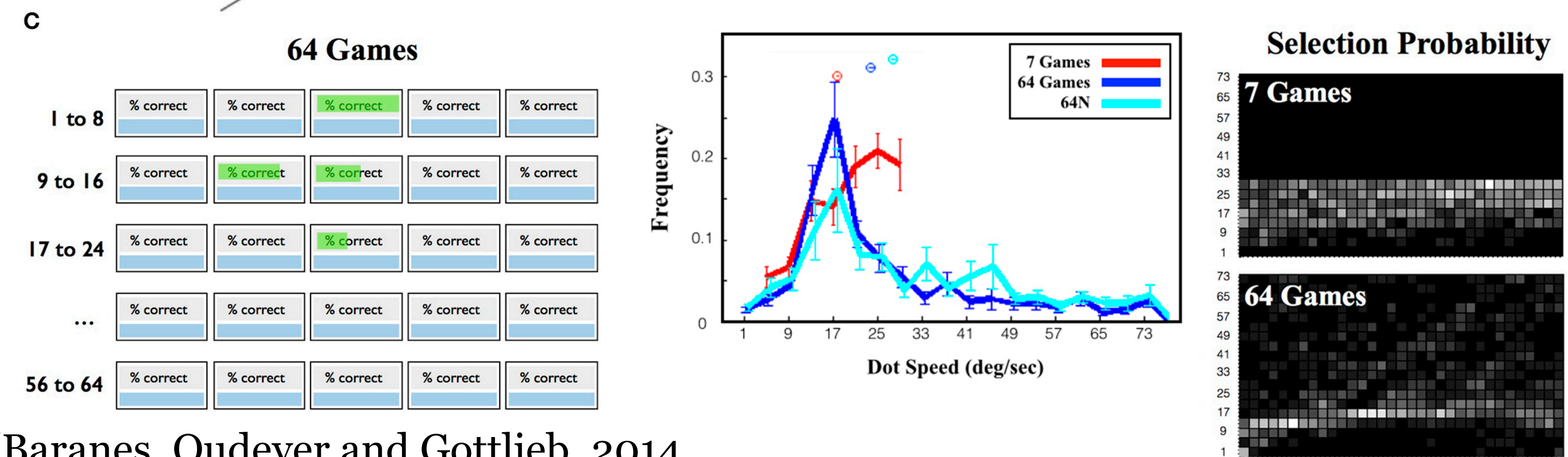
- ➔ Understanding how it can be made to work for acquisition of motor skills in high-dimensional real world (robotic) bodies (***Developmental robotics)***
- ➔ Understanding how it links with developmental organization

Back to human experiments

How spontaneous exploration is structured during free play



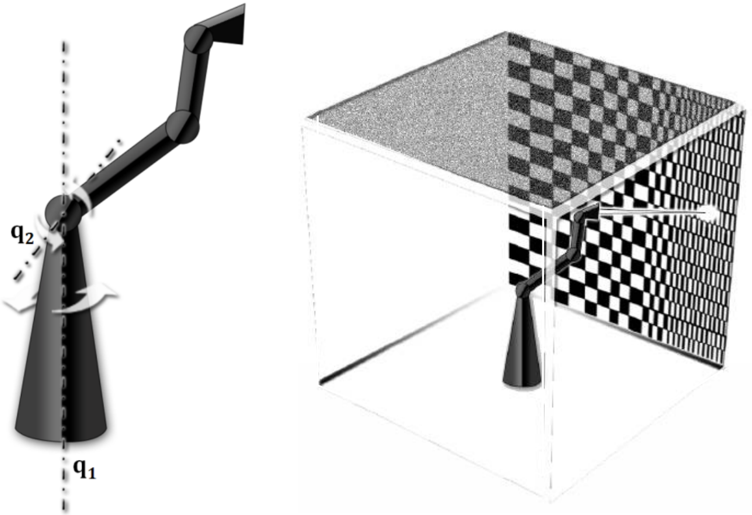
- Exploration follows a growth in complexity actively controlled as predicted by models
- Factors influence exploration patterns: task difficulty, novelty, size of the choice space



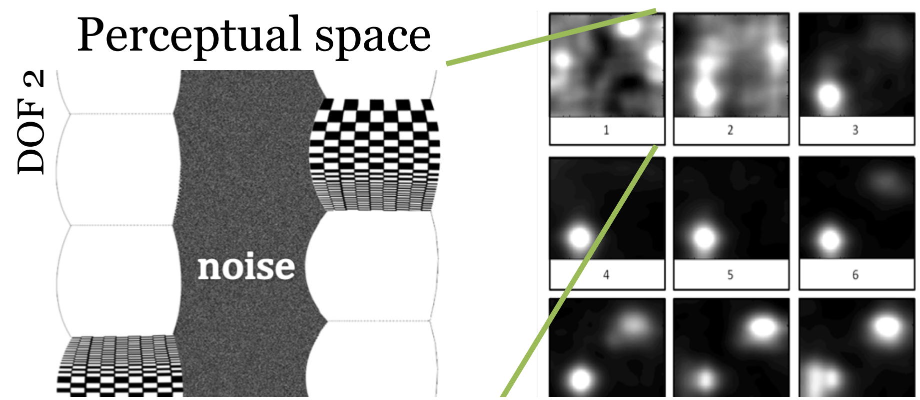
(Baranes, Oudeyer and Gottlieb, 2014 Frontiers in Neuroscience)

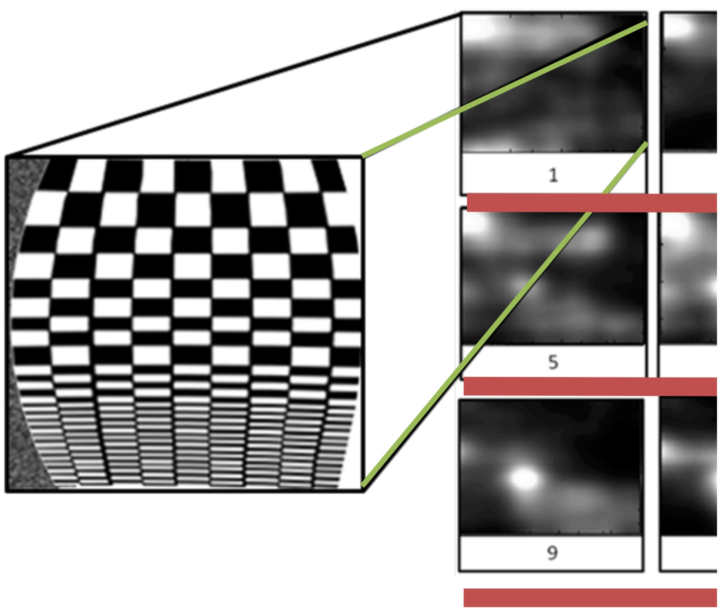
(Baranes and Oudeyer, TAMD 2009)

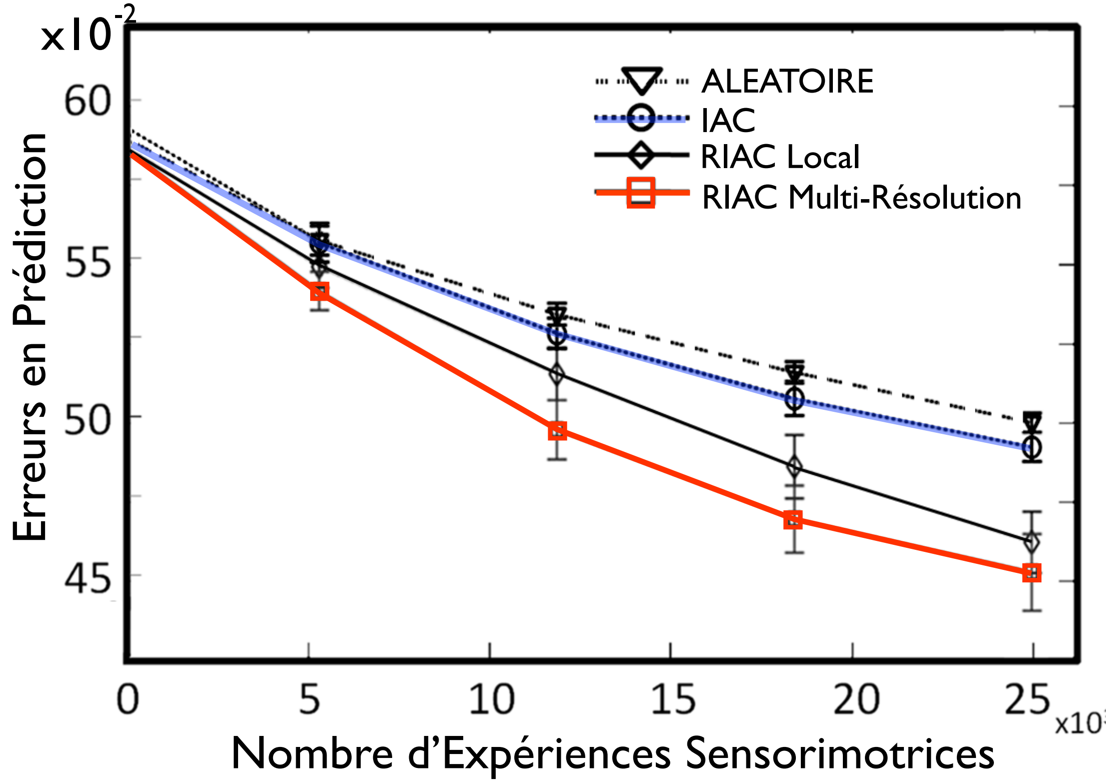
Simple example

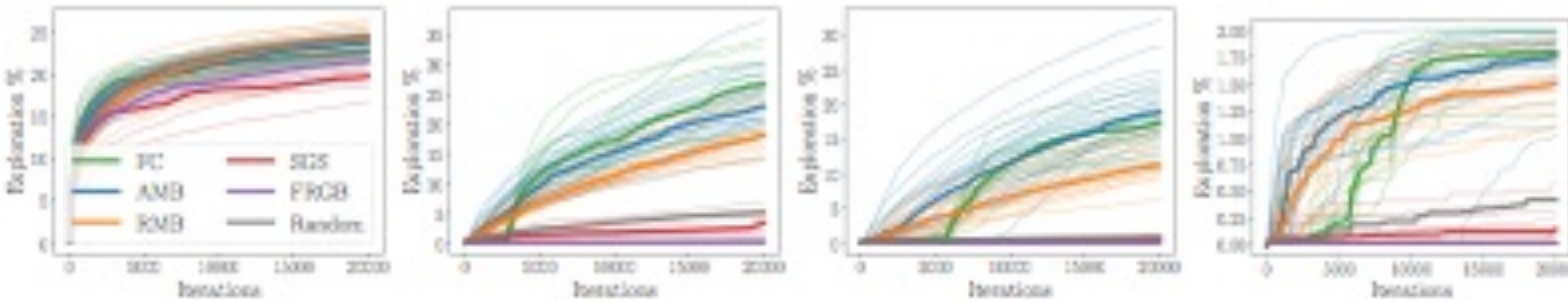


(motor 1, motor2) → intensity (1 pix

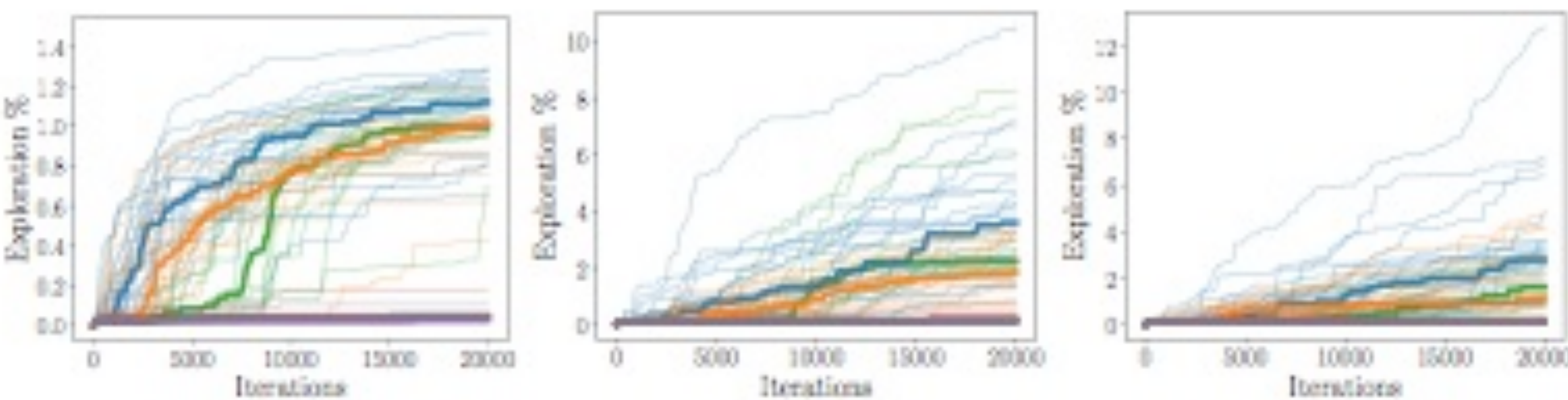






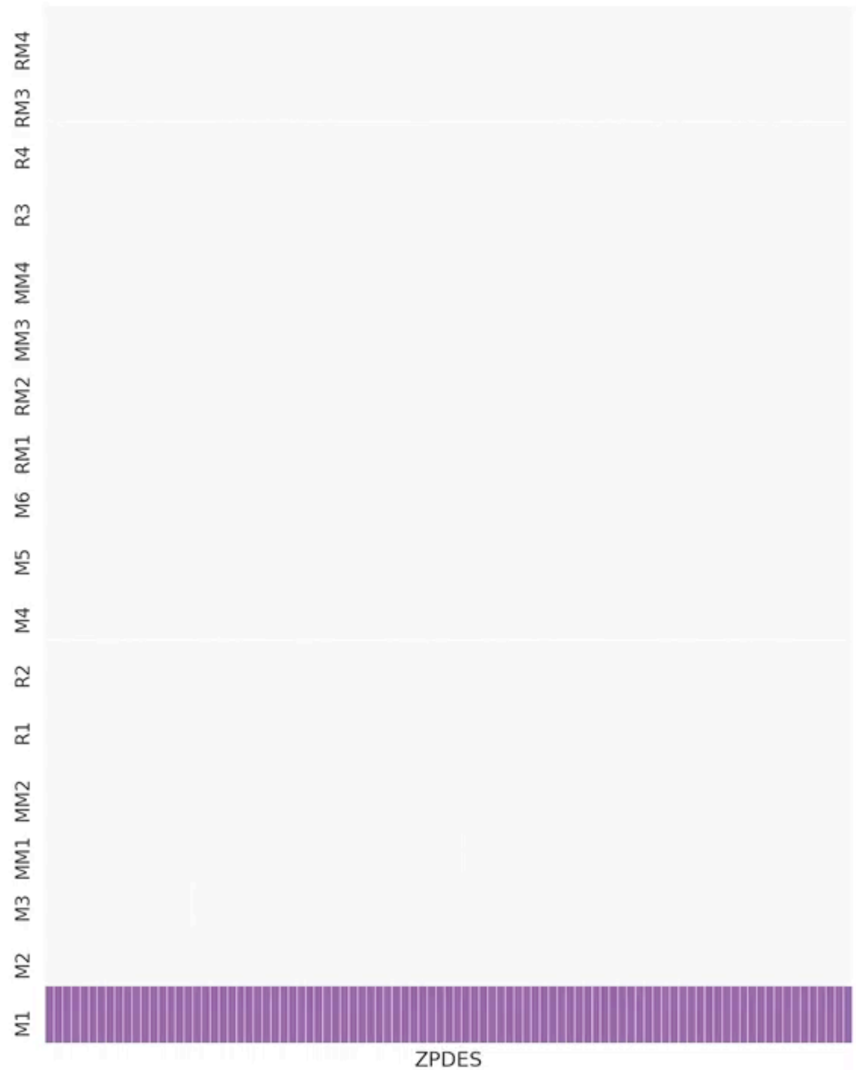


(a) Hand (b) Joystick Left (c) Joystick Right (d) Ergo



(e) Ball (f) Light (g) Sound

Learning and exploration dynamics



TARL: combining population-based and multi-goal Deep RL based IMGEPs (GEP-PG)

DDPG/TD3/SAC, …

Population-based IMGEP

Replay buffer

CURIOUS, UVFA+HER

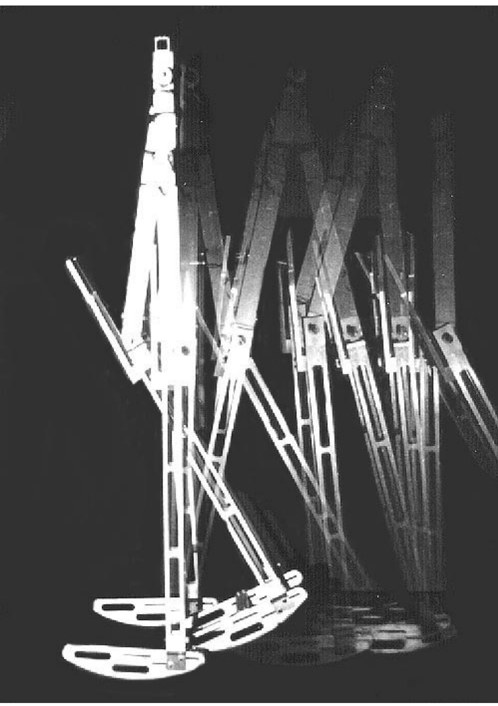
Fast exploration

Good generalization

(1) Robots are useful to better conceptualize the impact of the **body**

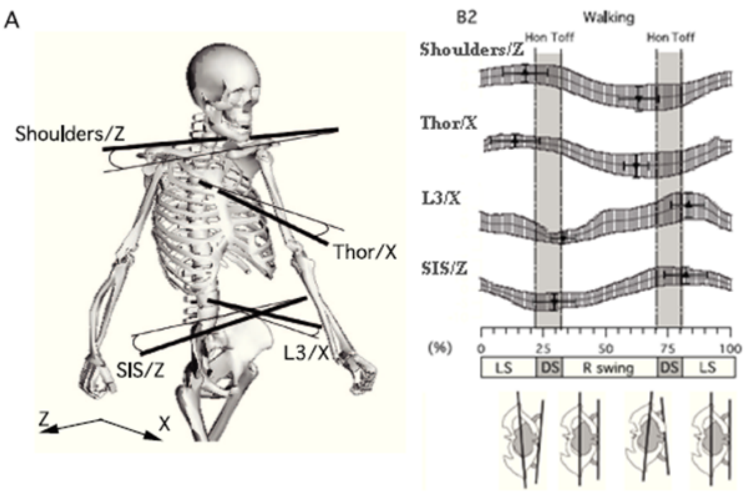
The example of walking

Morphology and self-organization of biped locomotion



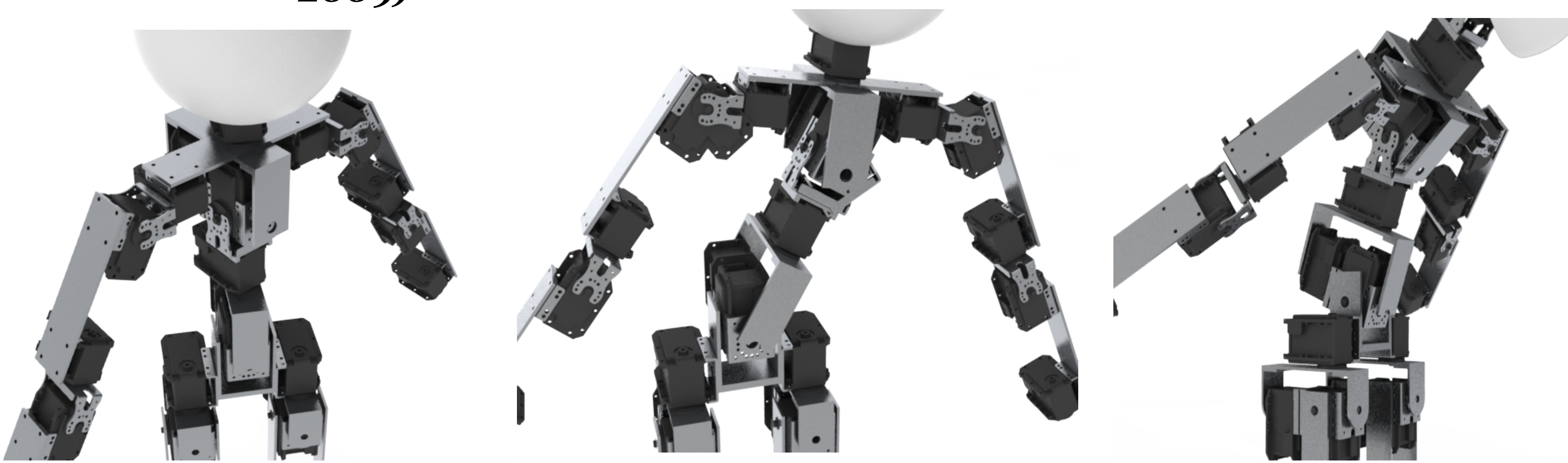
Tad McGeer (McGeer, 1990), Nagoya Univ. (2005)

Morphological computation



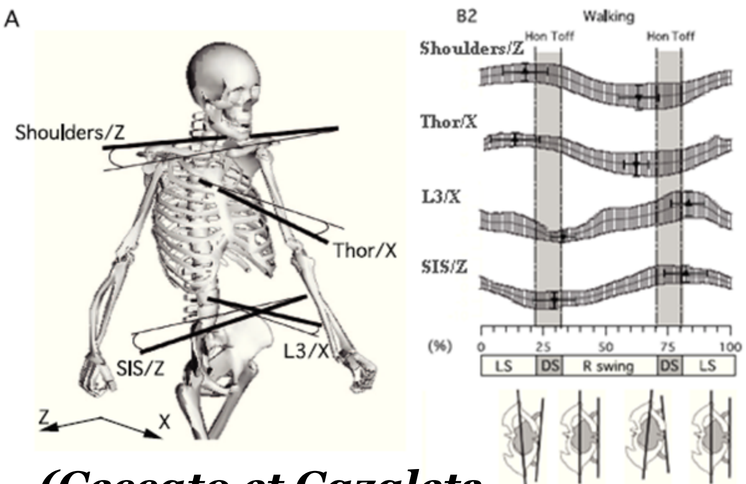
(Ceccato et Cazalets, 2009)

- **Collaboration with Labri/Univ. Bordeaux I**
- **Collaboration with J-R. Cazalets, Integrative Neuroscience Institute, Bordeaux**

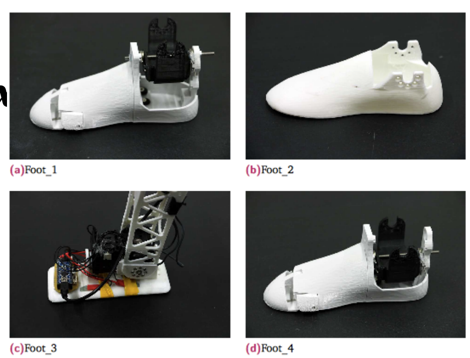


The Acroban humanoid *(Ly, Lapeyre, Oudeyer, 2011, IROS)*

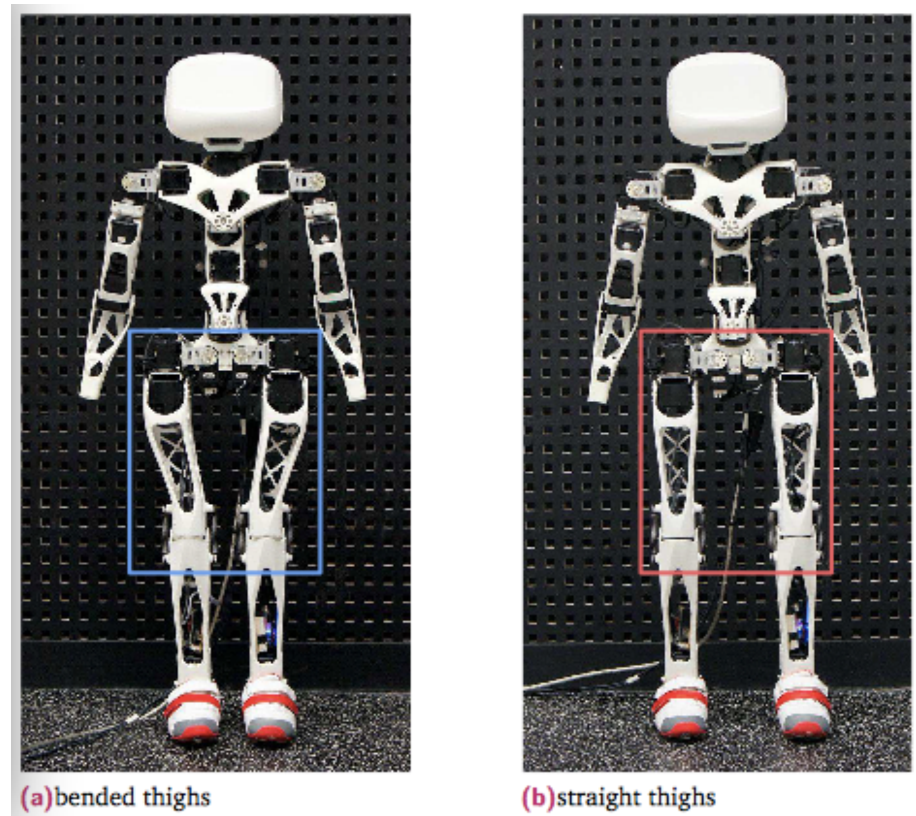
Body:
morphology, synergies and self-organization



(Ceccato et Cazalets, 2009) Neuroscience, Univ. Bordeaux



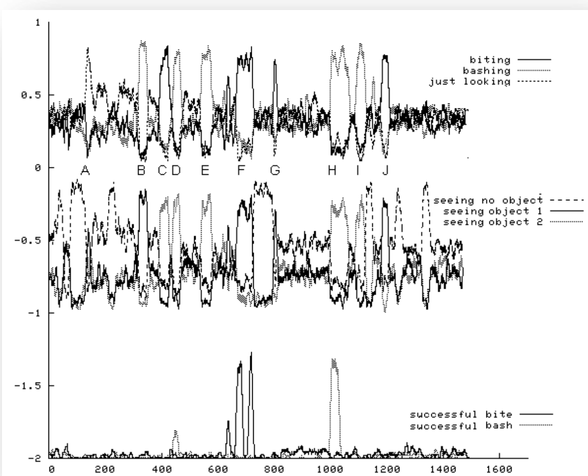
dy of properties of various feet, including passive spring loaded articulations (Humanoids 2014)



A human-like bended leg shape reduces the motion amplitude on the upper body by 45% and increases the head stability by 30% (Humanoids 2013; IROS 2013)

From affordances to vocal interaction

Playground Experiments

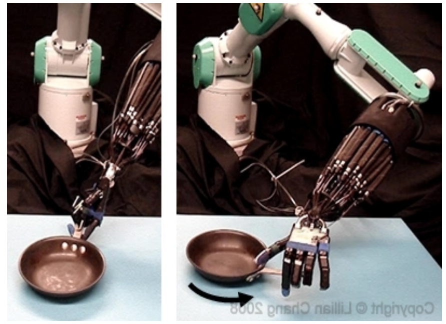


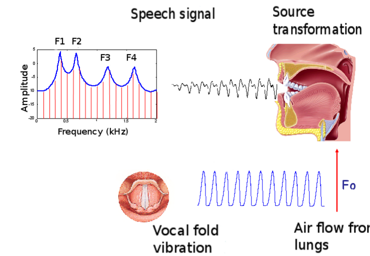
- Autonomous learning of novel affordances and and skills, e.g. object manipulation
- Self-organization of developmental trajectories, bootstrapping of communication
- Automatic formation of internal distinctive concepts for « self » vs « objects » vs « others »
- Regularities/diversity

➔ New hypotheses for understanding information seeking and curiosity in infant development

(Oudeyer et al., 2007 IEEE TEC)
(Kaplan and Oudeyer, Front. Neuroscience, 2007)

Development of sensorimotor skills





Context $\boxed{s(t)}$

$t + \Delta_t$ time

Trajectory:

$$\underbrace{\tau = \{s(t), a(t), \ldots, s(t+\Delta_t)\}}$$

Parameters of motor program π_θ *(DMP, RNN)*

Behavioural descriptors over full trajectory
(can be cost function measuring achievement of a complex property)

$$\varphi = [\varphi_1(\tau), \varphi_2(\tau), \ldots, \varphi_i(\tau)]$$

Mean speed of object C

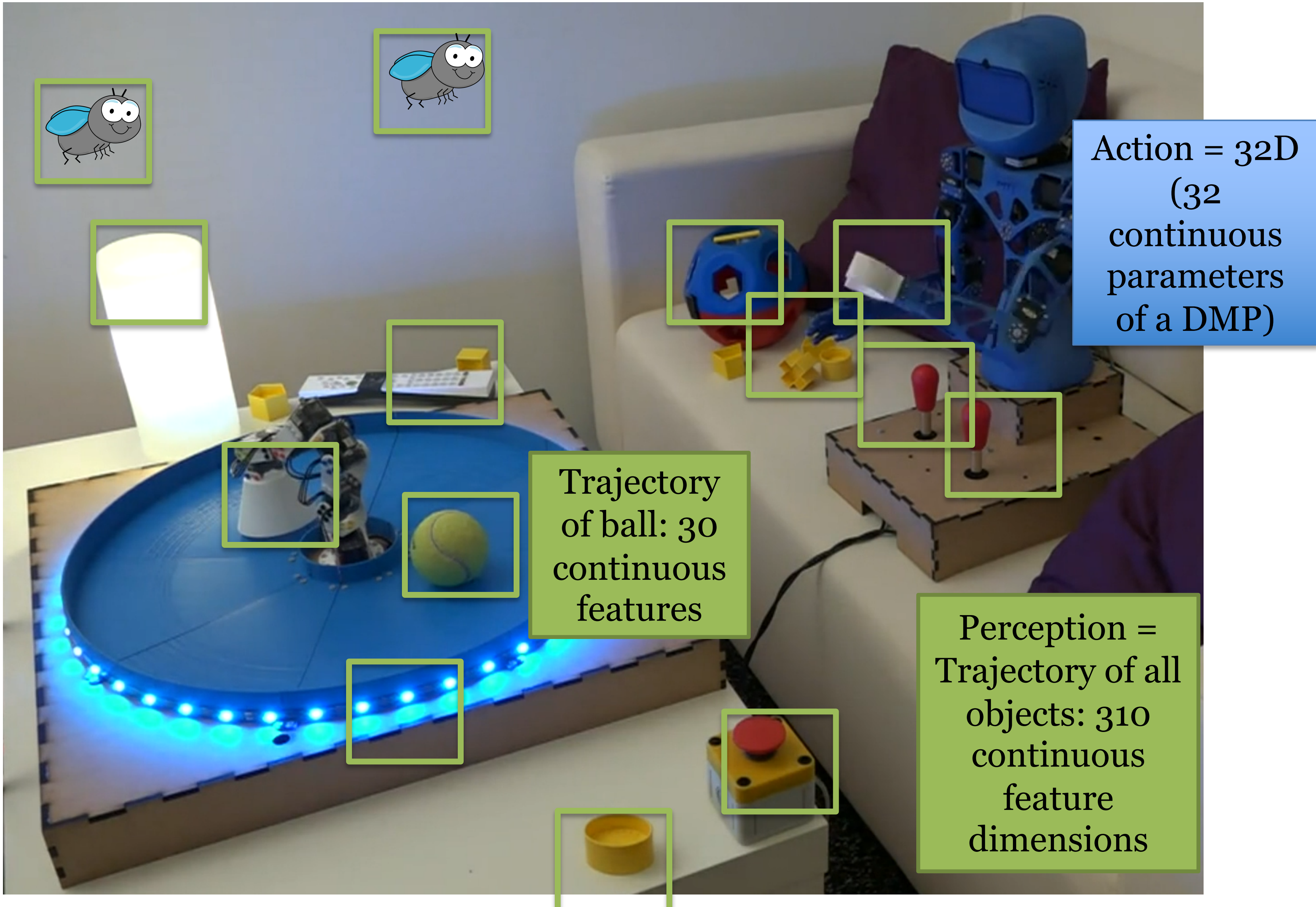
Vector of params of Bezier curve fitting traj. of obj. A

Classifier counts of events encountered over traj.

Learned RNN embedding

(Oudeyer and Kaplan, 2007)

MACOB: Modular population-based IMGEPs



(Forestier et al., 2017)