IN PARTNERSHIP WITH:
Ecole nationale supérieure des techniques avancées

Activity Report 2019

Project-Team FLOWERS

Flowing Epigenetic Robots and Systems

RESEARCH CENTER
Bordeaux - Sud-Ouest

THEME
Robotics and Smart environments

Table of contents

Project-Team FLOWERS

Creation of the Team: 2008 April 01, updated into Project-Team: 2011 January 01

Keywords:

Computer Science and Digital Science:

A5.1.1. - Engineering of interactive systems
A5.1.2. - Evaluation of interactive systems
A5.1.4. - Brain-computer interfaces, physiological computing
A5.1.5. - Body-based interfaces
A5.1.6. - Tangible interfaces
A5.1.7. - Multimodal interfaces
A5.3.3. - Pattern recognition
A5.4.1. - Object recognition
A5.4.2. - Activity recognition
A5.7.3. - Speech
A5.8. - Natural language processing
A5.10.5. - Robot interaction (with the environment, humans, other robots)
A5.10.7. - Learning
A5.10.8. - Cognitive robotics and systems
A5.11.1. - Human activity analysis and recognition
A6.3.1. - Inverse problems
A9. - Artificial intelligence
A9.2. - Machine learning
A9.5. - Robotics
A9.7. - AI algorithmics

Other Research Topics and Application Domains:

B1.2.1. - Understanding and simulation of the brain and the nervous system
B1.2.2. - Cognitive science
B5.6. - Robotic systems
B5.7. - 3D printing
B5.8. - Learning and training
B9. - Society and Knowledge
B9.1. - Education
B9.1.1. - E-learning, MOOC
B9.2. - Art
B9.2.1. - Music, sound
B9.2.4. - Theater
B9.6. - Humanities
B9.6.1. - Psychology
B9.6.8. - Linguistics
B9.7. - Knowledge dissemination

1. Team, Visitors, External Collaborators

Research Scientists

Pierre-Yves Oudeyer [Team leader, Inria, Senior Researcher, HDR]
Clément Moulin-Frier [Inria, Researcher, from Oct 2019]

Faculty Members

David Filliat [École Nationale Supérieure de Techniques Avancées, Professor, HDR]
Helene Sauzeon [Univ de Bordeaux, Professor, from May 2019, HDR]
Natalia Diaz Rodriguez [ENSTA, Associate Professor]

Post-Doctoral Fellows

Cécile Mazon [Univ de Bordeaux, ATER, from Nov 2019]
Adrien Laversanne Finot [Inria, until Mar 2019]
Chris Reinke [Inria]

PhD Students

Florence Carton [CEA]
Hugo Caselles-Dupré [Softbank]
Benjamin Clément [Inria, until Apr 2019, then Technique]
Cédric Colas [Inria]
Thibault Desprez [Inria, until Nov 2019]
Mayalen Etcheverry [Inria, from Feb 2019 until Nov 2019]
Sébastien Forestier [Inria, until Aug 2019]
Tristan Karch [Inria, from Oct 2019]
Timothee Lesort [École Nationale Supérieure de Techniques Avancées]
Vyshakh Palli Thaza [Renault]
Rémy Portelas [Inria]
Alexandr Ten [Inria]
Adrien Bennetot [Segula Technologies]

Technical staff

Grgur Kovac [Inria, Engineer, from Nov 2019]
Alexandre Pere [Inria, Engineer]
Didier Roy [Inria, Engineer]

Interns and Apprentices

Medhi Alaimi [Inria, until Jul 2019]
Timothee Anne [Inria, from Feb 2019 until Jun 2019]
Anouche Banikyan [Inria, from Feb 2019 until Jul 2019]
Lucie Galland [Ecole Normale Supérieure Paris, from Jun 2019 until Aug 2019]
Tallulah Gilliard [Inria, from Feb 2019 until Jul 2019]
Marion Schaeffer [Inria, from Jul 2019 until Sep 2019]
Martin Serret [Inria, from Feb 2019 until Aug 2019]
Maria Teodorescu [Inria, from Sep 2019]

Visiting Scientists

Kevyn Collins Thompson [University of Michigan, Professor, from Oct 2019]
Martina Parizkova [Charles University, from May 2019 until Jun 2019]

2. Overall Objectives

2.1. Overall Objectives

Can a machine learn like a child? Can it learn new skills and new knowledge in an unknown and changing environment? How can an embodied agent, e.g. a robot, discover its body and its relationships with the physical and social environment? How can its cognitive capacities continuously develop without the intervention of an engineer? What can it learn through natural social interactions with humans?

These are the questions that are being investigated in the FLOWERS research team at Inria Bordeaux Sud-Ouest and Ensta ParisTech. Rather than trying to imitate the intelligence of adult humans like in the field of Artificial Intelligence, we believe that trying to reconstruct the processes of development of the child's mind will allow for more adaptive, more robust and more versatile machines. This fundamental approach to the challenge of **autonomous learning** is called developmental robotics, or epigenetic robotics, and integrates concepts and theories from artificial intelligence, machine learning, neuroscience and developmental psychology. As many theories in neuroscience and developmental psychology are not formalized, this implies a crucial computational modeling activity, which in return provides means to assess the internal coherence of theories and sketch new hypothesis about the development of the human child's sensorimotor and cognitive abilities. Such computational modelling is also used as a foundational conceptual basis to build flexible lifelong autonomous machine learning systems.

Our team focuses in particular on the study of developmental constraints that allow for efficient open-ended learning of novel sensorimotor and interaction skills in embodied systems. In particular, we study constraints that guide exploration in large sensorimotor spaces:

- Mechanisms of intrinsically motivated exploration and active learning, including artificial curiosity, allowing to learn diverse skills in the absence of any external rewards, and in particular to self-organize developmental trajectories (also called automated curriculum learning) and collect efficiently learning data;
- Mechanisms of adequately constrained optimization and statistical inference for sample efficient sensorimotor skill acquisition (e.g. for optimizing motor policies in real robots through few interactions with the real world);
- Mechanisms for social learning, e.g. learning by imitation or demonstration, which implies both issues related to machine learning and human-robot interaction;
- Constraints related to embodiment, in particular through the concept of morphological computation, as well as the structure of motor primitives/muscle synergies that can leverage the properties of morphology and physics for simplifying motor control and perception;
- Maturational constraints which, coupled with the other constraints, can allow the progressive release of novel sensorimotor degrees of freedom to be explored;

We also study how these constraints on exploration can allow a machine to bootstrap multimodal perceptual abstractions associated to motor skills, in particular in the context of modelling language acquisition as a developmental process grounded in action.

Among the developmental principles that characterize human infants and can be used in developmental machines, FLOWERS focuses on the following three principles:

- **Exploration is progressive.** The space of skills that can be learnt in real world sensorimotor spaces is so large and complicated that not everything can be learnt at the same time. Simple skills are learnt first, and only when they are mastered, new skills of progressively increasing difficulty become the behavioural focus;
- **Internal representations are (partially) not innate but learnt and adaptive.** For example, the body map, the distinction self/non-self and the concept of "object" are discovered through experience with initially uninterpreted sensors and actuators, guided by experience, the overall pre-determined connection structure of the brain, as well as a small set of simple innate values or preferences.

- **Exploration can be self-guided and/or socially guided.** On the one hand, internal and intrinsic motivation systems regulate and organize spontaneous exploration; on the other hand, exploration can be guided through social learning and interaction with caretakers.

2.1.1. Research axis

The work of FLOWERS is organized around the following axis:

- **Curiosity-driven exploration and sensorimotor learning:** intrinsic motivation are mechanisms that have been identified by developmental psychologists to explain important forms of spontaneous exploration and curiosity. In FLOWERS, we try to develop computational intrinsic motivation systems, and test them on embodied machines, allowing to regulate the growth of complexity in exploratory behaviours. These mechanisms are studied as active learning mechanisms, allowing to learn efficiently in large inhomogeneous sensorimotor spaces and environments;
- **Cumulative learning of sensorimotor skills:** FLOWERS develops machine learning algorithms that can allow embodied machines to acquire cumulatively sensorimotor skills. In particular, we develop optimization and reinforcement learning systems which allow robots to discover and learn dictionaries of motor primitives, and then combine them to form higher-level sensorimotor skills.
- **Natural and intuitive social learning:** FLOWERS develops interaction frameworks and learning mechanisms allowing non-engineer humans to teach a robot naturally. This involves two sub-themes: 1) techniques allowing for natural and intuitive human-robot interaction, including simple ergonomic interfaces for establishing joint attention; 2) learning mechanisms that allow the robot to use the guidance hints provided by the human to teach new skills;
- **Discovering and abstracting the structure of sets of uninterpreted sensors and motors:** FLOWERS studies mechanisms that allow a robot to infer structural information out of sets of sensorimotor channels whose semantics is unknown, for example the topology of the body and the sensorimotor contingencies (propriocetive, visual and acoustic). This process is meant to be open-ended, progressing in continuous operation from initially simple representations to abstract concepts and categories similar to those used by humans.
- **Body design and role of the body in sensorimotor and social development:** We study how the physical properties of the body (geometry, materials, distribution of mass, growth, ...) can impact the acquisition of sensorimotor and interaction skills. This requires to consider the body as an experimental variable, and for this we develop special methodologies for designing and evaluating rapidly new morphologies, especially using rapid prototyping techniques like 3D printing.
- **Intelligent Tutoring Systems:**FLOWERS develops methods for online personalization of teaching sequences for educational software and MOOCs. This work builds on top of online optimization methods and motivational research previously developed.

3. Research Program

3.1. Research Program

Research in artificial intelligence, machine learning and pattern recognition has produced a tremendous amount of results and concepts in the last decades. A blooming number of learning paradigms - supervised, unsupervised, reinforcement, active, associative, symbolic, connectionist, situated, hybrid, distributed learning... - nourished the elaboration of highly sophisticated algorithms for tasks such as visual object recognition, speech recognition, robot walking, grasping or navigation, the prediction of stock prices, the evaluation of risk for insurances, adaptive data routing on the internet, etc... Yet, we are still very far from being able to build machines capable of adapting to the physical and social environment with the flexibility, robustness, and versatility of a one-year-old human child.

Indeed, one striking characteristic of human children is the nearly open-ended diversity of the skills they learn. They not only can improve existing skills, but also continuously learn new ones. If evolution certainly provided them with specific pre-wiring for certain activities such as feeding or visual object tracking, evidence shows that there are also numerous skills that they learn smoothly but could not be "anticipated" by biological evolution, for example learning to drive a tricycle, using an electronic piano toy or using a video game joystick. On the contrary, existing learning machines, and robots in particular, are typically only able to learn a single pre-specified task or a single kind of skill. Once this task is learnt, for example walking with two legs, learning is over. If one wants the robot to learn a second task, for example grasping objects in its visual field, then an engineer needs to re-program manually its learning structures: traditional approaches to task-specific machine/robot learning typically include engineer choices of the relevant sensorimotor channels, specific design of the reward function, choices about when learning begins and ends, and what learning algorithms and associated parameters shall be optimized.

As can be seen, this requires a lot of important choices from the engineer, and one could hardly use the term "autonomous" learning. On the contrary, human children do not learn following anything looking like that process, at least during their very first years. Babies develop and explore the world by themselves, focusing their interest on various activities driven both by internal motives and social guidance from adults who only have a folk understanding of their brains. Adults provide learning opportunities and scaffolding, but eventually young babies always decide for themselves what activity to practice or not. Specific tasks are rarely imposed to them. Yet, they steadily discover and learn how to use their body as well as its relationships with the physical and social environment. Also, the spectrum of skills that they learn continuously expands in an organized manner: they undergo a developmental trajectory in which simple skills are learnt first, and skills of progressively increasing complexity are subsequently learnt.

A link can be made to educational systems where research in several domains have tried to study how to provide a good learning experience to learners. This includes the experiences that allow better learning, and in which sequence they must be experienced. This problem is complementary to that of the learner that tries to learn efficiently, and the teacher here has to use as efficiently the limited time and motivational resources of the learner. Several results from psychology [59] and neuroscience [85] have argued that the human brain feels intrinsic pleasure in practicing activities of optimal difficulty or challenge. A teacher must exploit such activities to create positive psychological states of flow [73].

A grand challenge is thus to be able to build machines that possess this capability to discover, adapt and develop continuously new know-how and new knowledge in unknown and changing environments, like human children. In 1950, Turing wrote that the child's brain would show us the way to intelligence: "Instead of trying to produce a program to simulate the adult mind, why not rather try to produce one which simulates the child's" [154]. Maybe, in opposition to work in the field of Artificial Intelligence who has focused on mechanisms trying to match the capabilities of "intelligent" human adults such as chess playing or natural language dialogue [91], it is time to take the advice of Turing seriously. This is what a new field, called developmental (or epigenetic) robotics, is trying to achieve [109] [158]. The approach of developmental robotics consists in importing and implementing concepts and mechanisms from developmental psychology [117], cognitive linguistics [72], and developmental cognitive neuroscience [96] where there has been a considerable amount of research and theories to understand and explain how children learn and develop. A number of general principles are underlying this research agenda: embodiment [63] [131], grounding [89], situatedness [49], self-organization [150] [132], enaction [156], and incremental learning [67].

Among the many issues and challenges of developmental robotics, two of them are of paramount importance: exploration mechanisms and mechanisms for abstracting and making sense of initially unknown sensorimotor channels. Indeed, the typical space of sensorimotor skills that can be encountered and learnt by a developmental robot, as those encountered by human infants, is immensely vast and inhomogeneous. With a sufficiently rich environment and multimodal set of sensors and effectors, the space of possible sensorimotor activities is simply too large to be explored exhaustively in any robot's life time: it is impossible to learn all possible skills and represent all conceivable sensory percepts. Moreover, some skills are very basic to learn, some other very complicated, and many of them require the mastery of others in order to be learnt. For example, learning to manipulate a piano toy requires first to know how to move one's hand to reach the piano and how to touch specific parts of the toy with the fingers. And knowing how to move the hand might require to know how to track it visually.

Exploring such a space of skills randomly is bound to fail or result at best on very inefficient learning [128]. Thus, exploration needs to be organized and guided. The approach of epigenetic robotics is to take inspiration from the mechanisms that allow human infants to be progressively guided, i.e. to develop. There are two broad classes of guiding mechanisms which control exploration:

1. **internal guiding mechanisms,** and in particular intrinsic motivation, responsible of spontaneous exploration and curiosity in humans, which is one of the central mechanisms investigated in FLOWERS, and technically amounts to achieve online active self-regulation of the growth of complexity in learning situations;
2. **social learning and guidance,** a learning mechanisms that exploits the knowledge of other agents in the environment and/or that is guided by those same agents. These mechanisms exist in many different forms like emotional reinforcement, stimulus enhancement, social motivation, guidance, feedback or imitation, some of which being also investigated in FLOWERS;

3.1.1. Internal guiding mechanisms

In infant development, one observes a progressive increase of the complexity of activities with an associated progressive increase of capabilities [117], children do not learn everything at one time: for example, they first learn to roll over, then to crawl and sit, and only when these skills are operational, they begin to learn how to stand. The perceptual system also gradually develops, increasing children perceptual capabilities other time while they engage in activities like throwing or manipulating objects. This make it possible to learn to identify objects in more and more complex situations and to learn more and more of their physical characteristics.

Development is therefore progressive and incremental, and this might be a crucial feature explaining the efficiency with which children explore and learn so fast. Taking inspiration from these observations, some roboticists and researchers in machine learning have argued that learning a given task could be made much easier for a robot if it followed a developmental sequence and "started simple" [53] [79]. However, in these experiments, the developmental sequence was crafted by hand: roboticists manually build simpler versions of a complex task and put the robot successively in versions of the task of increasing complexity. And when they wanted the robot to learn a new task, they had to design a novel reward function.

Thus, there is a need for mechanisms that allow the autonomous control and generation of the developmental trajectory. Psychologists have proposed that intrinsic motivations play a crucial role. Intrinsic motivations are mechanisms that push humans to explore activities or situations that have intermediate/optimal levels of novelty, cognitive dissonance, or challenge [59] [73] [75]. The role and structure of intrinsic motivation in humans have been made more precise thanks to recent discoveries in neuroscience showing the implication of dopaminergic circuits and in exploration behaviours and curiosity [74] [93] [147]. Based on this, a number of researchers have began in the past few years to build computational implementation of intrinsic motivation [128] [129] [145] [57] [94] [112] [146]. While initial models were developed for simple simulated worlds, a current challenge is to manage to build intrinsic motivation systems that can efficiently drive exploratory behaviour in high-dimensional unprepared real world robotic sensorimotor spaces [129], [128], [130], [143]. Specific and complex problems are posed by real sensorimotor spaces, in particular due to the fact that they are both high-dimensional as well as (usually) deeply inhomogeneous. As an example for the latter issue, some regions of real sensorimotor spaces are often unlearnable due to inherent stochasticity or difficulty, in which case heuristics based on the incentive to explore zones of maximal unpredictability or uncertainty, which are often used in the field of active learning [70] [90] typically lead to catastrophic results. The issue of high dimensionality does not only concern motor spaces, but also sensory spaces, leading to the problem of correctly identifying, among typically thousands of quantities, those latent variables that have links to behavioral choices. In FLOWERS, we aim at developing intrinsically motivated exploration mechanisms that scale in those spaces, by studying suitable abstraction processes in conjunction with exploration strategies.

3.1.2. Socially Guided and Interactive Learning

Social guidance is as important as intrinsic motivation in the cognitive development of human babies [117]. There is a vast literature on learning by demonstration in robots where the actions of humans in the environment are recognized and transferred to robots [52]. Most such approaches are completely passive: the human executes actions and the robot learns from the acquired data. Recently, the notion of interactive learning has been introduced in [151], [60], motivated by the various mechanisms that allow humans to socially guide a robot [139]. In an interactive context the steps of self-exploration and social guidance are not separated and a robot learns by self exploration and by receiving extra feedback from the social context [151], [99], [113].

Social guidance is also particularly important for learning to segment and categorize the perceptual space. Indeed, parents interact a lot with infants, for example teaching them to recognize and name objects or characteristics of these objects. Their role is particularly important in directing the infant attention towards objects of interest that will make it possible to simplify at first the perceptual space by pointing out a segment of the environment that can be isolated, named and acted upon. These interactions will then be complemented by the children own experiments on the objects chosen according to intrinsic motivation in order to improve the knowledge of the object, its physical properties and the actions that could be performed with it.

In FLOWERS, we are aiming at including intrinsic motivation system in the self-exploration part thus combining efficient self-learning with social guidance [122], [123]. We also work on developing perceptual capabilities by gradually segmenting the perceptual space and identifying objects and their characteristics through interaction with the user [110] and robots experiments [95]. Another challenge is to allow for more flexible interaction protocols with the user in terms of what type of feedback is provided and how it is provided [107].

Exploration mechanisms are combined with research in the following directions:

3.1.3. Cumulative learning, reinforcement learning and optimization of autonomous skill learning

FLOWERS develops machine learning algorithms that can allow embodied machines to acquire cumulatively sensorimotor skills. In particular, we develop optimization and reinforcement learning systems which allow robots to discover and learn dictionaries of motor primitives, and then combine them to form higher-level sensorimotor skills.

3.1.4. Autonomous perceptual and representation learning

In order to harness the complexity of perceptual and motor spaces, as well as to pave the way to higher-level cognitive skills, developmental learning requires abstraction mechanisms that can infer structural information out of sets of sensorimotor channels whose semantics is unknown, discovering for example the topology of the body or the sensorimotor contingencies (proprioceptive, visual and acoustic). This process is meant to be open- ended, progressing in continuous operation from initially simple representations towards abstract concepts and categories similar to those used by humans. Our work focuses on the study of various techniques for:

- autonomous multimodal dimensionality reduction and concept discovery;
- incremental discovery and learning of objects using vision and active exploration, as well as of auditory speech invariants;
- learning of dictionaries of motion primitives with combinatorial structures, in combination with linguistic description;
- active learning of visual descriptors useful for action (e.g. grasping);

3.1.5. Embodiment and maturational constraints

FLOWERS studies how adequate morphologies and materials (i.e. morphological computation), associated to relevant dynamical motor primitives, can importantly simplify the acquisition of apparently very complex skills such as full-body dynamic walking in biped. FLOWERS also studies maturational constraints, which are mechanisms that allow for the progressive and controlled release of new degrees of freedoms in the sensorimotor space of robots.

3.1.6. Discovering and abstracting the structure of sets of uninterpreted sensors and motors

FLOWERS studies mechanisms that allow a robot to infer structural information out of sets of sensorimotor channels whose semantics is unknown, for example the topology of the body and the sensorimotor contingencies (proprioceptive, visual and acoustic). This process is meant to be open-ended, progressing in continuous operation from initially simple representations to abstract concepts and categories similar to those used by humans.

4. Application Domains

4.1. Application Domains

Neuroscience, Developmental Psychology and Cognitive Sciences The computational modelling of life-long learning and development mechanisms achieved in the team centrally targets to contribute to our understanding of the processes of sensorimotor, cognitive and social development in humans. In particular, it provides a methodological basis to analyze the dynamics of the interaction across learning and inference processes, embodiment and the social environment, allowing to formalize precise hypotheses and later on test them in experimental paradigms with animals and humans. A paradigmatic example of this activity is the Neurocuriosity project achieved in collaboration with the cognitive neuroscience lab of Jacqueline Gottlieb, where theoretical models of the mechanisms of information seeking, active learning and spontaneous exploration have been developed in coordination with experimental evidence and investigation, see https://flowers.inria.fr/neurocuriosityproject/.

Personal and lifelong learning robotics Many indicators show that the arrival of personal robots in homes and everyday life will be a major fact of the 21st century. These robots will range from purely entertainment or educative applications to social companions that many argue will be of crucial help in our society. Yet, to realize this vision, important obstacles need to be overcome: these robots will have to evolve in unpredictable homes and learn new skills in a lifelong manner while interacting with non-engineer humans after they left factories, which is out of reach of current technology. In this context, the refoundation of intelligent systems that developmental robotics is exploring opens potentially novel horizons to solve these problems. In particular, this application domain requires advances in artificial intelligence that go beyond the current state-of-the-art in fields like deep learning. Currently these techniques require tremendous amounts of data in order to function properly, and they are severely limited in terms of incremental and transfer learning. One of our goals is to drastically reduce the amount of data required in order for this very potent field to work. We try to achieve this by making neural networks aware of their knowledge, i.e. we introduce the concept of uncertainty, and use it as part of intrinsically motivated multitask learning architectures, and combined with techniques of learning by imitation.

Human-Robot Collaboration. Robots play a vital role for industry and ensure the efficient and competitive production of a wide range of goods. They replace humans in many tasks which otherwise would be too difficult, too dangerous, or too expensive to perform. However, the new needs and desires of the society call for manufacturing system centered around personalized products and small series productions. Human-robot collaboration could widen the use of robot in this new situations if robots become cheaper, easier to program and safe to interact with. The most relevant systems for such applications would follow an expert worker and works with (some) autonomy, but being always under supervision of the human and acts based on its task models.

Environment perception in intelligent vehicles. When working in simulated traffic environments, elements of FLOWERS research can be applied to the autonomous acquisition of increasingly abstract representations of both traffic objects and traffic scenes. In particular, the object classes of vehicles and pedestrians are if interest when considering detection tasks in safety systems, as well as scene categories (”scene context”) that have a strong impact on the occurrence of these object classes. As already indicated by several investigations in the field, results from present-day simulation technology can be transferred to the real world with little impact on performance. Therefore, applications of FLOWERS research that is suitably verified by real-world benchmarks has direct applicability in safety-system products for intelligent vehicles.

Automated Tutoring Systems. Optimal teaching and efficient teaching/learning environments can be applied to aid teaching in schools aiming both at increase the achievement levels and the reduce time needed. From a practical perspective, improved models could be saving millions of hours of students' time (and effort) in learning. These models should also predict the achievement levels of students in order to influence teaching practices.

5. Highlights of the Year

5.1. Highlights of the Year

- Clément Moulin-Frier was recruited as CRCN permanent research scientist.
- PY Oudeyer was invited to give plenary keynote talks at several international AI conferences, including ICLR 2019 in New Orleans (https://www.youtube.com/watch?v=7bJ0fnvPLaA) and ACM International Conference on Virtual Agents (ACM IVA 2019), Paris.
- The team published papers in several major machine learning conferences, including ICML [33], Neurips [32], CoRL [38] and IJCNN [36], [37], and one in a major educational technology conference, CHI 2019 [29].
- PY Oudeyer was awarded an individual "Chaire IA" in the context of the national plan on artificial intelligence.
- The Poppy Station association, initiated by the team from the Poppy Education project, and co-directed by Didier Roy, was created and gathers several major national and international educational associations. It aims at scaling up and disseminating the educational robotics kits designed by the Flowers team, and now used in many educational and artistic projects, see https://www.poppy-station.org.
- The work of the PhD of Sébastien Forestier (sup. by PY Oudeyer) on curiosity-driven learning of tool use in robots and children, was integrated as a video interview in the new permanent exhibition on robots at Cité des Sciences et de l'Industrie, Paris, see http://www.cite-sciences.fr/fr/au-programme/expos-permanentes/expos-permanentes-dexplora/robots/lexposition/.

5.1.1. Awards

- Y Oudeyer was awarded the Atos Joseph Fourier prize for his work on curiosity-driven machine learning.

6. New Software and Platforms

6.1. Explauto

an autonomous exploration library

KEYWORD: Exploration

SCIENTIFIC DESCRIPTION: An important challenge in developmental robotics is how robots can be intrinsically motivated to learn efficiently parametrized policies to solve parametrized multi-task reinforcement learning problems, i.e. learn the mappings between the actions and the problem they solve, or sensory effects they produce. This can be a robot learning how arm movements make physical objects move, or how movements of a virtual vocal tract modulates vocalization sounds. The way the robot will collects its own sensorimotor experience have a strong impact on learning efficiency because for most robotic systems the involved spaces are high dimensional, the mapping between them is non-linear and redundant, and there is limited time allowed for learning. If robots explore the world in an unorganized manner, e.g. randomly, learning algorithms will be often ineffective because very sparse data points will be collected. Data are precious due to the high dimensionality and the limited time, whereas data are not equally useful due to non-linearity and redundancy. This is why learning has to be guided using efficient exploration strategies, allowing the robot to actively drive its own interaction with the environment in order to gather maximally informative data to optimize the parametrized policies. In the recent year, work in developmental learning has explored various families of algorithmic principles which allow the efficient guiding of learning and exploration.

Explauto is a framework developed to study, model and simulate curiosity-driven learning and exploration in real and simulated robotic agents. Explauto's scientific roots trace back from Intelligent Adaptive Curiosity algorithmic architecture [127], which has been extended to a more general family of autonomous exploration architectures by [1] and recently expressed as a compact and unified formalism [119]. The library is detailed in [120]. In Explauto, interest models are implementing the strategies of active selection of particular problems / goals in a parametrized multi-task reinforcement learning setup to efficiently learn parametrized policies. The agent can have different available strategies, parametrized problems, models, sources of information, or learning mechanisms (for instance imitate by mimicking vs by emulation, or asking help to one teacher or to another), and chooses between them in order to optimize learning (a processus called strategic learning [124]). Given a set of parametrized problems, a particular exploration strategy is to randomly draw goals/ RL problems to solve in the motor or problem space. More efficient strategies are based on the active choice of learning experiments that maximize learning progress using bandit algorithms, e.g. maximizing improvement of predictions or of competences to solve RL problems [127]. This automatically drives the system to explore and learn first easy skills, and then explore skills of progressively increasing complexity. Both random and learning progress strategies can act either on the motor or on the problem space, resulting in motor babbling or goal babbling strategies.

- Motor babbling consists in sampling commands in the motor space according to a given strategy (random or learning progress), predicting the expected effect, executing the command through the environment and observing the actual effect. Both the parametrized policies and interest models are finally updated according to this experience.
- Goal babbling consists in sampling goals in the problem space and to use the current policies to infer a motor action supposed to solve the problem (inverse prediction). The robot/agent then executes the command through the environment and observes the actual effect. Both the parametrized policies and interest models are finally updated according to this experience.It has been shown that this second strategy allows a progressive solving of problems much more uniformly in the problem space than with a motor babbling strategy, where the agent samples directly in the motor space [1].

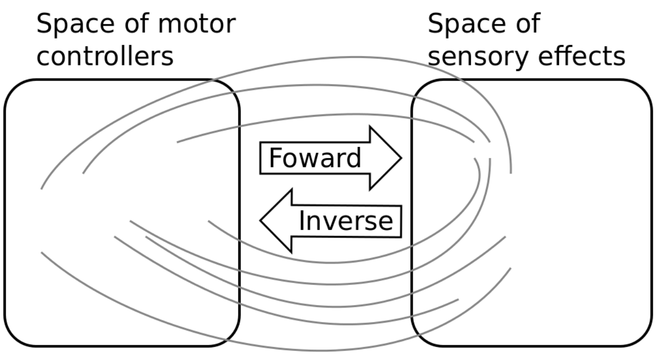


Figure 1. Complex parametrized policies involve high dimensional action and effect spaces. For the sake of visualization, the motor M and sensory S spaces are only 2D each in this example. The relationship between M and S is non-linear, dividing the sensorimotor space into regions of unequal stability: small regions of S can be reached very precisely by large regions of M, or large regions in S can be very sensitive to variations in M.: s as well as a non-linear and redundant relationship. This non-linearity can imply redundancy, where the same sensory effect can be attained using distinct regions in M.

FUNCTIONAL DESCRIPTION: This library provides high-level API for an easy definition of:

- Real and simulated robotic setups (Environment level),
- Incremental learning of parametrized policies (Sensorimotor level),
- Active selection of parametrized RL problems (Interest level).

The library comes with several built-in environments. Two of them corresponds to simulated environments: a multi-DoF arm acting on a 2D plan, and an under-actuated torque-controlled pendulum. The third one allows to control real robots based on Dynamixel actuators using the Pypot library. Learning parametrized policies involves machine learning algorithms, which are typically regression algorithms to learn forward models, from motor controllers to sensory effects, and optimization algorithms to learn inverse models, from sensory effects, or problems, to the motor programs allowing to reach them. We call these sensorimotor learning algorithms sensorimotor models. The library comes with several built-in sensorimotor models: simple nearest-neighbor look-up, non-parametric models combining classical regressions and optimization algorithms, online mixtures of Gaussians, and discrete Lidstone distributions. Explauto sensorimotor models are online learning algorithms, i.e. they are trained iteratively during the interaction of the robot in theenvironment in which it evolves. Explauto provides also a unified interface to define exploration strategies using the InterestModel class. The library comes with two built-in interest models: random sampling as well as sampling maximizing the learning progress in forward or inverse predictions.

Explauto environments now handle actions depending on a current context, as for instance in an environment where a robotic arm is trying to catch a ball: the arm trajectories will depend on the current position of the ball (context). Also, if the dynamic of the environment is changing over time, a new sensorimotor model (Non-Stationary Nearest Neighbor) is able to cope with those changes by taking more into account recent experiences. Those new features are explained in Jupyter notebooks.

This library has been used in many experiments including:

- the control of a 2D simulated arm,
- the exploration of the inverse kinematics of a poppy humanoid (both on the real robot and on the simulated version),
- acoustic model of a vocal tract.

Explauto is crossed-platform and has been tested on Linux, Windows and Mac OS. It has been released under the GPLv3 license.

- Contact: Sebastien Forestier
- URL: https://github.com/flowersteam/explauto

6.2. KidBreath

KEYWORD: Machine learning

FUNCTIONAL DESCRIPTION: KidBreath is a web responsive application composed by several interactive contents linked to asthma and displayed to different forms: learning activities with quiz, short games and videos. There are profil creation and personalization, and a part which describes historic and scoring of learning activities, to see evolution of Kidreath use. To test Kidlearn algorithm, it is iadapted and integrated on this platform. Development in PHP, HTML-5, CSS, MySQL, JQuery, Javascript. Hosting in APACHE, LINUX, PHP 5.5, MySQL, OVH.

- Partner: ItWell SAS
- Contact: Alexandra Delmas
- URL: http://www.kidbreath.fr

Figure 2. Four principal regions are defined in the graphical interface. The first is the wallet location where users can pick and drag the money items and drop them on the repository location to compose the correct price. The object and the price are present in the object location. Four different types of exercises exist: M : customer/one object, R : merchant/one object, MM : customer/two objects, RM : merchant/two objects.

6.3. Kidlearn: money game application

FUNCTIONAL DESCRIPTION: The games is instantiated in a browser environment where students are proposed exercises in the form of money/token games (see Figure 2). For an exercise type, one object is presented with a given tagged price and the learner has to choose which combination of bank notes, coins or abstract tokens need to be taken from the wallet to buy the object, with various constraints depending on exercises parameters. The games have been developed using web technologies, HTML5, javascript and Django.

- Contact: Benjamin Clement
- URL: https://flowers.inria.fr/research/kidlearn/

6.4. Kidlearn: script for Kidbreath use

KEYWORD: PHP

FUNCTIONAL DESCRIPTION: A new way to test Kidlearn algorithms is to use them on Kidbreath Plateform. The Kidbreath Plateform use apache/PHP server, so to facilitate the integration of our algorithm, a python script have been made to allow PHP code to use easily the python library already made which include our algorithms.

- Contact: Benjamin Clement
- URL: https://flowers.inria.fr/research/kidlearn/

6.5. KidLearn

KEYWORD: Automatic Learning

FUNCTIONAL DESCRIPTION: KidLearn is a software which adaptively personalize sequences of learning activities to the particularities of each individual student. It aims at proposing to the student the right activity at the right time, maximizing concurrently his learning progress and its motivation.

- Participants: Benjamin Clement, Didier Roy, Manuel Lopes and Pierre Yves Oudeyer
- Contact: Pierre-Yves Oudeyer
- URL: https://flowers.inria.fr/research/kidlearn/

6.6. Poppy

KEYWORDS: Robotics - Education

FUNCTIONAL DESCRIPTION: The Poppy Project team develops open-source 3D printed robots platforms based on robust, flexible, easy-to-use and reproduce hardware and software. In particular, the use of 3D printing and rapid prototyping technologies is a central aspect of this project, and makes it easy and fast not only to reproduce the platform, but also to explore morphological variants. Poppy targets three domains of use: science, education and art.

In the Poppy project we are working on the Poppy System which is a new modular and open-source robotic architecture. It is designed to help people create and build custom robots. It permits, in a similar approach as Lego, building robots or smart objects using standardized elements.

Poppy System is a unified system in which essential robotic components (actuators, sensors...) are independent modules connected with other modules through standardized interfaces:

- Unified mechanical interfaces, simplifying the assembly process and the design of 3D printable parts.
- Unified communication between elements using the same connector and bus for each module.
- Unified software, making it easy to program each module independently.

Our ambition is to create an ecosystem around this system so communities can develop custom modules, following the Poppy System standards, which can be compatible with all other Poppy robots.

- Participants: Jonathan Grizou, Matthieu Lapeyre, Pierre Rouanet and Pierre-Yves Oudeyer
- Contact: Pierre-Yves Oudeyer
- URL: https://www.poppy-project.org/

6.7. Poppy Ergo Jr

KEYWORDS: Robotics - Education

FUNCTIONAL DESCRIPTION: Poppy Ergo Jr is an open hardware robot developed by the Poppy Project to explore the use of robots in classrooms for learning robotic and computer science.

It is available as a 6 or 4 degrees of freedom arm designed to be both expressive and low-cost. This is achieved by the use of FDM 3D printing and low cost Robotis XL-320 actuators. A Raspberry Pi camera is attached to the robot so it can detect object, faces or QR codes.

The Ergo Jr is controlled by the Pypot library and runs on a Raspberry pi 2 or 3 board. Communication between the Raspberry Pi and the actuators is made possible by the Pixl board we have designed.

The Poppy Ergo Jr robot has several 3D printed tools extending its capabilities. There are currently the lampshade, the gripper and a pen holder.

With the release of a new Raspberry Pi board early 2016, the Poppy Ergo Jr disk image was updated to support Raspberry Pi 2 and 3 boards. The disk image can be used seamlessly with a board or the other.

- Contact: Theo Segonds
- URL: https://github.com/poppy-project/poppy-ergo-jr

6.8. S-RL Toolbox

Reinforcement Learning (RL) and State Representation Learning (SRL) for Robotics

KEYWORDS: Machine learning - Robotics

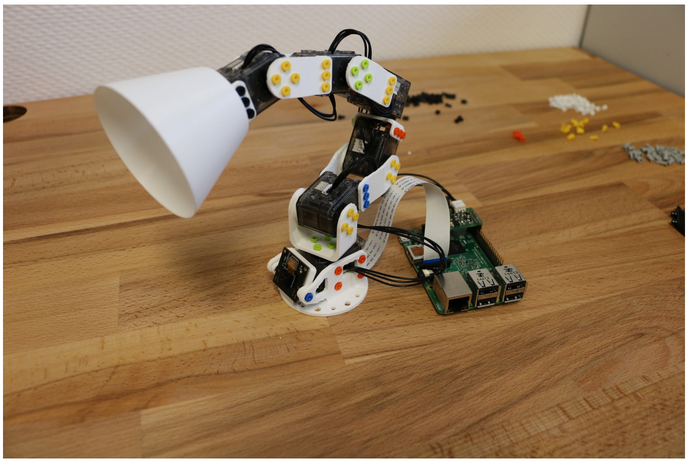


Figure 3. Poppy Ergo Jr, 6-DoFs arm robot for education

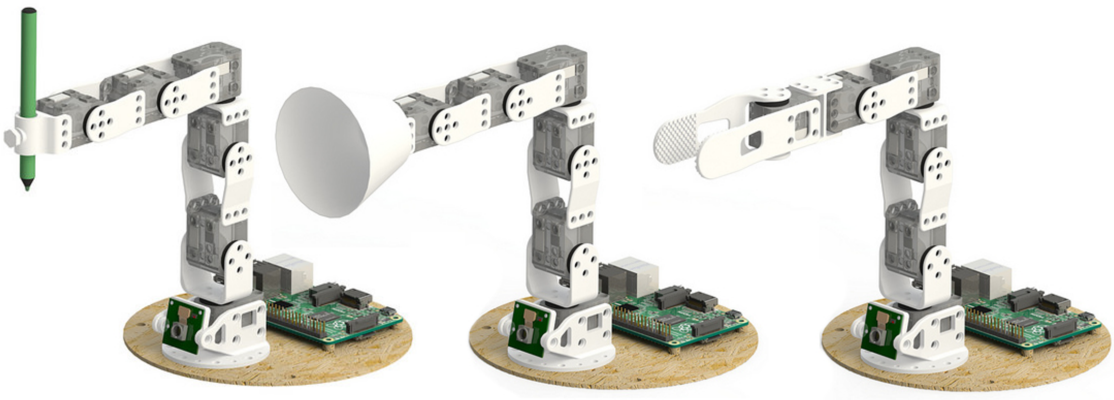


Figure 4. The available Ergo Jr tools: a pen holder, a lampshade and a gripper

FUNCTIONAL DESCRIPTION: This repository was made to evaluate State Representation Learning methods using Reinforcement Learning. It integrates (automatic logging, plotting, saving, loading of trained agent) various RL algorithms (PPO, A2C, ARS, ACKTR, DDPG, DQN, ACER, CMA-ES, SAC, TRPO) along with different SRL methods (see SRL Repo) in an efficient way (1 Million steps in 1 Hour with 8-core cpu and 1 Titan X GPU).

- Partner: ENSTA
- Contact: David Filliat
- URL: https://github.com/araffin/robotics-rl-srl

6.9. Deep-Explauto

KEYWORDS: Deep learning - Unsupervised learning - Learning - Experimentation

FUNCTIONAL DESCRIPTION: Until recently, curiosity driven exploration algorithms were based on classic learning algorithms, unable to handle large dimensional problems (see explauto). Recent advances in the field of deep learning offer new algorithms able to handle such situations.

Deep explauto is an experimental library, containing reference implementations of curiosity driven exploration algorithms. Given the experimental aspect of exploration algorithms, and the low maturity of the libraries and algorithms using deep learning, proposing black-box implementations of those algorithms, enabling a blind use of those, seem unrealistic.

Nevertheless, in order to quickly launch new experiments, this library offers an set of objects, functions and examples, allowing to kickstart new experiments.

- Contact: Alexandre Pere

6.10. Orchestra

KEYWORD: Experimental mechanics

FUNCTIONAL DESCRIPTION: Ochestra is a set of tools meant to help in performing experimental campaigns in computer science. It provides you with simple tools to:

+ Organize a manual experimental workflow, leveraging git and lfs through a simple interface. + Collaborate with other peoples on a single experimental campaign. + Execute pieces of code on remote hosts such as clusters or clouds, in one line. + Automate the execution of batches of experiments and the presentation of the results through a clean web ui.

A lot of advanced tools exists on the net to handle similar situations. Most of them target very complicated workflows, e.g. DAGs of tasks. Those tools are very powerful but lack the simplicity needed by newcomers. Here, we propose a limited but very simple tool to handle one of the most common situation of experimental campaigns: the repeated execution of an experiment on variations of parameters.

In particular, we include three tools: + expegit: a tool to organize your experimental campaign results in a git repository using git-lfs (large file storage). + runaway: a tool to execute code on distant hosts parameterized with easy to use file templates. + orchestra: a tool to automate the use of the two previous tools on large campaigns.

- Contact: Alexandre Pere

6.11. Curious

Curious: Intrinsically Motivated Modular Multi-Goal Reinforcement Learning

KEYWORDS: Exploration - Reinforcement learning - Artificial intelligence

FUNCTIONAL DESCRIPTION: This is an algorithm enabling to learn a controller for an agent in a modular multi-goal environment. In these types of environments, the agent faces multiple goals classified in different types (e.g. reaching goals, grasping goals for a manipulation robot).

- Contact: Cedric Colas

6.12. teachDeepRL

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

KEYWORDS: Machine learning - Git

FUNCTIONAL DESCRIPTION: Codebase from our CoRL2019 paper https://arxiv.org/abs/1910.07224

This github repository provides implementations for the following teacher algorithms: - Absolute Learning Progress-Gaussian Mixture Model (ALP-GMM), our proposed teacher algorithm - Robust Intelligent Adaptive Curiosity (RIAC), from Baranes and Oudeyer, R-IAC: robust intrinsically motivated exploration and active learning. - Covar-GMM, from Moulin-Frier et al., Self-organization of early vocal development in infants and machines: The role of intrinsic motivation.

- Contact: Remy Portelas
- URL: https://github.com/flowersteam/teachDeepRL

6.13. Automated Discovery of Lenia Patterns

KEYWORDS: Exploration - Cellular automaton - Deep learning - Unsupervised learning

SCIENTIFIC DESCRIPTION: In many complex dynamical systems, artificial or natural, one can observe selforganization of patterns emerging from local rules. Cellular automata, like the Game of Life (GOL), have been widely used as abstract models enabling the study of various aspects of self-organization and morphogenesis, such as the emergence of spatially localized patterns. However, findings of self-organized patterns in such models have so far relied on manual tuning of parameters and initial states, and on the human eye to identify “interesting” patterns. In this paper, we formulate the problem of automated discovery of diverse self-organized patterns in such high-dimensional complex dynamical systems, as well as a framework for experimentation and evaluation. Using a continuous GOL as a testbed, we show that recent intrinsically-motivated machine learning algorithms (POP-IMGEPs), initially developed for learning of inverse models in robotics, can be transposed and used in this novel application area. These algorithms combine intrinsically motivated goal exploration and unsupervised learning of goal space representations. Goal space representations describe the “interesting” features of patterns for which diverse variations should be discovered. In particular, we compare various approaches to define and learn goal space representations from the perspective of discovering diverse spatially localized patterns. Moreover, we introduce an extension of a state-of-the-art POP-IMGEP algorithm which incrementally learns a goal representation using a deep auto-encoder, and the use of CPPN primitives for generating initialization parameters. We show that it is more efficient than several baselines and equally efficient as a system pre-trained on a hand-made database of patterns identified by human experts.

FUNCTIONAL DESCRIPTION: Python source code of experiments and data analysis for the paper " Intrinsically Motivated Discovery of Diverse Patterns in Self-Organizing Systems" (Chris Reinke, Mayalen Echeverry, Pierre-Yves Oudeyer in Submitted to ICLR 2020). The software includes: Lenia environment, exploration algorithms (IMGEPs, random search), deep learning algorithms for unsupervised learning of goal spaces, tools and configurations to run experiments, and data analysis tools.

- Contact: Chris Reinke
- URL: https://github.com/flowersteam/automated_discovery_of_lenia_patterns

6.14. ZPDES_ts

ZPDES in typescript

KEYWORDS: Machine learning - Education

FUNCTIONAL DESCRIPTION: ZPDES is a machine learning-based algorithm that allows you to customize the content of training courses for each learner's level. It has already been implemented in the Kidlern software in python with other algorithms. Here, ZPDES is implemented in typescript.

- Authors: Benjamin Clement, Pierre-Yves Oudeyer, Didier Roy and Manuel Lopes
- Contact: Benjamin Clement
- URL: https://flowers.inria.fr/research/kidlearn/

6.15. GEP-PG

Goal Exploration Process - Policy Gradient

KEYWORDS: Machine learning - Deep learning

FUNCTIONAL DESCRIPTION: Reinforcement Learning algorithm working with OpenAI Gym environments. A first phase implements exploration using a Goal Exploration Process (GEP). Samples collected during exploration are then transferred to the memory of a deep reinforcement learning algorithm (deep deterministic policy gradient or DDPG). DDPG then starts learning from a pre-initialized memory so as to maximize the sum of discounted rewards given by the environment.

- Contact: Cedric Colas

7. New Results

7.1. Curiosity-Driven Learning in Humans

7.1.1. Computational Models Of Information-Seeking and Curiosity-Driven Learning in Human Adults

Participants: Pierre-Yves Oudeyer [correspondant], Sébastien Forestier, Alexandr Ten.

This project involves a collaboration between the Flowers team and the Cognitive Neuroscience Lab of J. Gottlieb at Columbia Univ. (NY, US), on the understanding and computational modeling of mechanisms of curiosity, attention and active intrinsically motivated exploration in humans.

It is organized around the study of the hypothesis that subjective meta-cognitive evaluation of information gain (or control gain or learning progress) could generate intrinsic reward in the brain (living or artificial), driving attention and exploration independently from material rewards, and allowing for autonomous lifelong acquisition of open repertoires of skills. The project combines expertise about attention and exploration in the brain and a strong methodological framework for conducting experimentations with monkeys, human adults and children together with computational modeling of curiosity/intrinsic motivation and learning.

Such a collaboration paves the way towards a central objective, which is now a central strategic objective of the Flowers team: designing and conducting experiments in animals and humans informed by computational/mathematical theories of information seeking, and allowing to test the predictions of these computational theories.

7.1.1.1. Context

Curiosity can be understood as a family of mechanisms that evolved to allow agents to maximize their knowledge (or their control) of the useful properties of the world - i.e., the regularities that exist in the world - using active, targeted investigations. In other words, we view curiosity as a decision process that maximizes learning/competence progress (rather than minimizing uncertainty) and assigns value ("interest") to competing tasks based on their epistemic qualities - i.e., their estimated potential allow discovery and learning about the structure of the world.

Because a curiosity-based system acts in conditions of extreme uncertainty (when the distributions of events may be entirely unknown) there is in general no optimal solution to the question of which exploratory action to take [108], [130], [141]. Therefore we hypothesize that, rather than using a single optimization process as it has been the case in most previous theoretical work [86], curiosity is comprised of a family of mechanisms that include simple heuristics related to novelty/surprise and measures of learning progress over longer time scales [128] [56], [118]. These different components are related to the subject's epistemic state (knowledge and beliefs) and may be integrated with fluctuating weights that vary according to the task context. Our aim is to quantitatively characterize this dynamic, multi-dimensional system in a computational framework based on models of intrinsically motivated exploration and learning.

Because of its reliance on epistemic currencies, curiosity is also very likely to be sensitive to individual differences in personality and cognitive functions. Humans show well-documented individual differences in curiosity and exploratory drives [106], [140], and rats show individual variation in learning styles and novelty seeking behaviors [80], but the basis of these differences is not understood. We postulate that an important component of this variation is related to differences in working memory capacity and executive control which, by affecting the encoding and retention of information, will impact the individual's assessment of learning, novelty and surprise and ultimately, the value they place on these factors [133], [149], [50], [155]. To start understanding these relationships, about which nothing is known, we will search for correlations between curiosity and measures of working memory and executive control in the population of children we test in our tasks, analyzed from the point of view of a computational models of the underlying mechanisms.

A final premise guiding our research is that essential elements of curiosity are shared by humans and non-human primates. Human beings have a superior capacity for abstract reasoning and building causal models, which is a prerequisite for sophisticated forms of curiosity such as scientific research. However, if the task is adequately simplified, essential elements of curiosity are also found in monkeys [106], [98] and, with adequate characterization, this species can become a useful model system for understanding the neurophysiological mechanisms.

7.1.1.2. Objectives

Our studies have several highly innovative aspects, both with respect to curiosity and to the traditional research field of each member team.

- Linking curiosity with quantitative theories of learning and decision making: While existing investigations examined curiosity in qualitative, descriptive terms, here we propose a novel approach that integrates quantitative behavioral and neuronal measures with computationally defined theories of learning and decision making.
- Linking curiosity in children and monkeys: While existing investigations examined curiosity in humans, here we propose a novel line of research that coordinates its study in humans and non-human primates. This will address key open questions about differences in curiosity between species, and allow access to its cellular mechanisms.
- Neurophysiology of intrinsic motivation: Whereas virtually all the animal studies of learning and decision making focus on operant tasks (where behavior is shaped by experimenter-determined primary rewards) our studies are among the very first to examine behaviors that are intrinsically motivated by the animals' own learning, beliefs or expectations.
- Neurophysiology of learning and attention: While multiple experiments have explored the single-neuron basis of visual attention in monkeys, all of these studies focused on vision and eye movement control. Our studies are the first to examine the links between attention and learning, which are recognized in psychophysical studies but have been neglected in physiological investigations.
- Computer science: biological basis for artificial exploration: While computer science has proposed and tested many algorithms that can guide intrinsically motivated exploration, our studies are the first to test the biological plausibility of these algorithms.
- Developmental psychology: linking curiosity with development: While it has long been appreciated that children learn selectively from some sources but not others, there has been no systematic investigation of the factors that engender curiosity, or how they depend on cognitive traits.

7.1.1.3. Current results: experiments in Active Categorization

In 2018, we have been occupied by analyzing data of the human adult experiment conducted in 2017. In this experiment we asked whether humans possess, and use, metacognitive abilities to guide task choices in two contexts motivational contexts, in which they could freely choose to learn about 4 competing tasks. Participants (n = 505, recruited via Amazon Mechanical Turk) were asked to play a categorization game with four distinct difficulty levels. Some participants had been explicitly prescribed a goal of maximizing their learning across the difficulty levels (across tasks), while others did not receive any specific instructions regarding the goal of the game. The experiment yielded a rich but complex set of data. The data includes records of participants' classification responses, task choices, reaction times, and post-task self-reports about various subjective evaluations of the competing tasks (e.g. subjective interest, progress, learning potential, etc.). We are now finalizing the results and a computational model of the underlying cognitive and motivational mechanisms in order to prepare them for public dissemination.

The central question going into the study was, how do active learners become interested in specific learning exercises: how do they decide which task to engage with, when none of the tasks provide external rewards. Last year, we identified some of the key behavioral observations that merited further attention. First, we saw a clear effect of an external goal prescription on the participants' overall task selection strategy. People who were explicitly instructed to try to maximize their learning across the 4 tasks challenged themselves more by giving preference towards harder tasks. In contrast, those who were simply familiarized with the rules of the game and not given any explicit suggestions from the experiments did not show this overchallenge bias and had a slight preference for easier tasks (see figure 5). Second, we observed that although strategies varied between the two instruction groups, there was some considerable within-group variability in learning. We found that in both groups, people had varying success in learning the classification task for each task family resulting in four distinct performance based groups: learners of 0, 1, 2, or 3 tasks (task 4 was unlearnable), as shown in figure 6. Importantly, successful 3-task learners in both instruction groups exhibited similar task preferences, suggesting that (1) even in the absence of external instruction, people can be motivated to explore the task space and (2) intrinsically motivated exploration is similar to strategies employed when a learner is trying to maximize her learning.

Assuming that task choice decisions are based on a subjective evaluative process that assigns value to choice candidates, we considered a simple choice model of task selection. In a classic conditional logit model [116], choices are made probabilistically and the choice probabilities are proportional to choice utilities (the inherent subjective value of a choice; also see ref [159]). We elaborated on the utility component of the basic choice model to consider two utility aspects of interest: a relative measure of learning progress (LP) and an absolute measure of proportion correct (PC). Although both measures are based on empirical feedback (correct / incorrect), the LP measure is considered relative, because it captures how performance changes over time by comparing performance estimates across different time scales, while PC is absolute in a sense that it only characterizes performance at a given instance. While PC alone does not differentiate between an unfamiliar (but potentially easy) task on which the performance might be low and a familiar but very hard task, the former can have markedly LP (due to the gradual improvement on that task) than the latter. Only the tasks characterized by high LP are then worthy of time and effort if the goal is to master tasks. The utility component in our model thus includes two principal quantities:

$$u_{i,t} = \alpha LP_{i,t} + \beta PC_{i,t}$$

where α and β are free parameters indexing the model's sensitivity to LP and PC, respectively. Index i designates the task, while t indexes time. Thus, in our model, utility is seen as a dual-component linear computation of both relative and absolute competence quantities. Task utilities enter the decision-making process that assigns relative preference to each task:

$$p(task_i) = \frac{e^{u_i/\tau}}{\sum_j e^{u_j/\tau}}$$

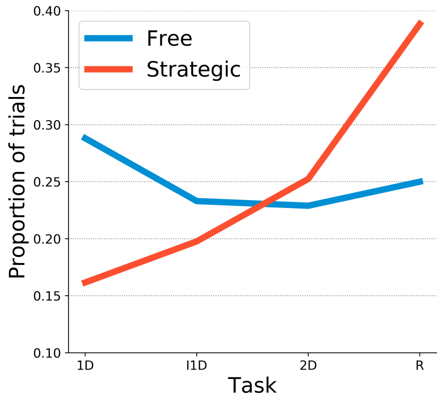


Figure 5. Proportion of trials on each task (1D, I1D, 2D, and R). The group with an externally prescribed learning maximization goal is referred to as "Strategic", and the unconstrained group is referred to as "Free". 1D was the task were categorization was determined by a single variable dimension. In I1D (ignore 1D), the stimuli varied across 2 dimensions, but only one determined the stimulus category. In 2D, there were 2 variable dimensions and both jointly determined the category. Finally in R, there were 2 variable dimensions, but none of them could reliably predict the stimulus class. The top plot shows data aggregated across experimental groups, shown separately in the bottom plot. In hte figure, task difficulty increases from left to right.

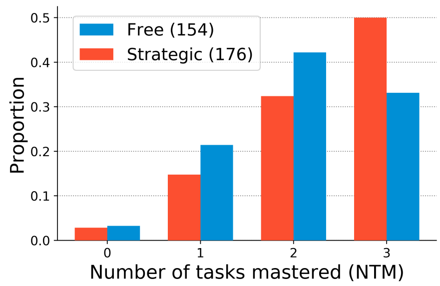


Figure 6. Learning outcomes by group. The group with an externally prescribed learning maximization goal is referred to as "Strategic", and the unconstrained group is referred to as "Free". Number of people in the groups are given in parentheses. The figure shows that some people seem to set and pursue learning maximization goals in the absence of external rewards and instructions. Additionally, learning goals have to be matched with appropriate behavioral strategies in order to be reached, which explains the failure to learn all 3 tasks by some people in the Strategic group.

where τ is another free parameter (known as temperature) that controls the stochasticity of utility-based decision. The sum over j elements, $\sum_j$ constitutes the total exponentiated utility of each task in the task space, thus normalizing each individual exponentiated task utility u_i.

The computation of the utility components is important for the model, because it ultimately determines how well the model can fit to choice data. We started exploring the model with a simple definition for both LP and PC. Both components are based on averaging binary feedback over multiple trials in the past. Since the familiarization stage of our experiment was 15-trials-long, the first free choice was made based on feedback data from 15 trials on each task. Accordingly, we defined PC to be the proportion of correct guesses in 15 trials. LP was defined as the absolute difference between the first 9 and the last 6 trials of the recent most 15 trials in the past. While a participant was engaged with one of the tasks, LP and PC for that task changed according to her dynamic performance, while LPs and PCs for other tasks remained unchanged. We acknowledge that there are probably multiple other components at play when it comes to utility computation, some of which may have little to do with task competence. We also submit that there are multiple ways of defining the PC and LP components that are more biologically rooted and plausible given what we know of memory and metacognition. Finally, we do not rule out the possibility of dynamic changes of free parameters themselves, corresponding to changes in motivation during the learning process. All of these considerations are worthy directions of future research, but in this study we focused on finding some necessary evidence for the sensitivity to learning progress.

We fitted the model to each individual's choice data using maximum likelihood estimation. Assuming that choice probabilities on each trial come from a categorical distribution (also called a generalized Bernoulli distribution), where the probability of choosing item i is given by:

$$P(\mathbf{x} \mid \mathbf{p}) = \prod_{i=1}^{k} p_i^{x_i}$$

where **p** is a vector of probabilities associated with k events, and **p** is a one-hot encoded vector representing discrete items x_i. We add a time index to indicate the dynamic quality of choice probabilities, so that:

$$P_t(\mathbf{x} \mid \mathbf{p}_t) = \prod_{i=1}^{k} p_{i,t}^{x_i}$$

Then, the likelihood of the choice model ($p(task_i \mid \alpha, \beta, \tau)$) at time t is equal to the product of choice probabilities given by that model for that time step:

$$L_t(\alpha, \beta, \tau \mid \mathbf{x}) = \prod_{i}^{k} p_{i,t}^{x_i}$$

and since the empirical choice data can be represented in a one-hot format, the likelihood of the model for a given time point boils down to the predicted probability of the actualy choice:

$$L_t(\alpha, \beta, \tau \mid choice = i) = p_{i,t} = \frac{e^{u_{i,t}/\tau}}{\sum_j e^{u_{j,t}/\tau}}$$

The likelihood of the model across all m trials is obtained by applying the product rule of probability:

$$L_{overall}(\alpha, \beta, \tau \mid choice = i) = \prod_{t}^{m} p_{i,t} = \frac{e^{u_{i,t}/\tau}}{\sum_j e^{u_{j,t}/\tau}}$$

For convenience, we use the negative log transformation to avoid computational precision problems and convert a likelihood maximization objective into negative likelihood minimization problem solvable by publicly accessible optimization tools:

$$-\log L_{overall}(\alpha, \beta, \tau \mid choice = i) = -\sum_{t}^{m} \log p_{i,t}$$

Having formulated the likelihood function, we optimized the free parameters to obtain a model that fits the individual data best. We thus fit an individual-level model to each participant's choice data. The fitted parameters can be interpreted as relative sensitivity to the competence quantities of interest (LP and PC), since these quantities share the same range of values (0 to 1). Finally, we performed some group-level analyses on these individual-level parameter estimates to evaluate certain group-level effects that might influence them (e.g. effect of instruction or learning proficiency).

The group with a learning maximization goal devalued tasks with higher positive feedback expectation. Qualitatively, this matched our prior observations showing their strong preference for harder tasks. However, the best learners (3-task learners) across both instruction groups showed a slight preference for learning progress and a relatively strong aversion to positive feedback. It appears that what separated better and worse learners among the learning maximizers was whether they followed learning progress, and not just the feedback heuristic. On the other hand, while less successful learners in the unconstrained group seemed to choose tasks according to learning progress, they valued positive feedback over it, which prevented them from exploring more challenging learnable tasks. This is reflected in group mean values of the fitted parameters summarized in table 1

Table 1. Fitted parameter values averaged within number of tasks learned and instruction groups. The group with an externally prescribed learning maximization goal is referred to as "Strategic", and the unconstrained group is referred to as "Free". NTM stands for the number of tasks mastered

Group	NTM	Learning progress	Proportion correct	Temperature
	1	0.27	0.56	6.92
	2	0.07	0.32	5.90
	3	0.12	-0.15	6.14
	1	-0.01	-0.56	7.14
	2	-0.15	-0.41	6.42
	3	0.11	-0.44	6.59

We also looked at the relative importance of the utility model parameters by performing model comparisons. We compared 4 models based on combinations of 2 factors: PC and LP. According to the AIC scores (see Table 2), the best model was the one which included both LP and PC factors, followed by the PC-only model, and then by the LP-only model. The random-choice model came in last with the highest AIC score. The results of this model comparison show that both learning progress and positive feedback expectation factors provide substantive improvement to model likelihood compared to when these factors are included alone, or when neither of them is present. This is potentially important, as it provides some evidence for the role of relative competence kind of measure in autonomous exploration. We are planning to submit the work described about to a high impact peer-reviewed journal focusing on computational modeling of human behavior.

7.1.2. Experimental study of the role of intrinsic motivation in developmental psychology experiments and in the development of tool use

Participants: Pierre-Yves Oudeyer, Sébastien Forestier [correspondant], Laurianne Rat-Fisher.

Table 2. Model comparisons.

Model	AIC	$AIC - AIC_{min}$
LP + PC	568.99	-
PC	593.51	24.52
LP	658.46	89.47
Random	693.60	124.61

Children are so curious to explore their environment that it can be hard to focus their attention on one given activity. Many experiments in developmental psychology evaluate particular skills of children by setting up a task that the child is encouraged to solve. However, children may sometimes be following their own motivation to explore the experimental setup or other things in the environment. We suggest that considering the intrinsic motivations of children in those experiments could help understanding their role in the learning of related skills and on long-term child development. To illustrate this idea, we reanalyzed and reinterpretd a typical experiment aiming to evaluate particular skills in infants. In this experiment run by Lauriane Rat-Fischer et al, 32 21-month old infants have to retrieve a toy stuck inside a tube, by inserting several blocks in sequence into the tube. In order to understand the mechanisms of the motivations of babies, we studied in detail their behaviors, goals and strategies in this experiment. We showed that their motivations are diverse and do not always coincide with the target goal expected and made salient by the experimenter. Intrinsically motivated exploration seems to play an important role in the observed behaviors and to interfere with the measured success rates. This new interpretation provides a motivation for studying curiosity and intrinsic motivations in robotic models.

7.2. Intrinsically Motivated Learning in Artificial Intelligence

7.2.1. Intrinsically Motivated Goal Exploration and Goal-Parameterized Reinforcement Learning

Participants: Sébastien Forestier, Pierre-Yves Oudeyer [correspondant], Olivier Sigaud, Cédric Colas, Adrien Laversanne-Finot, Rémy Portelas, Grgur Kovac.

7.2.1.1. Intrinsically Motivated Exploration of Modular Goal Spaces and the Emergence of Tool use

A major challenge in robotics is to learn goal-parametrized policies to solve multi-task reinforcement learning problems in high-dimensional continuous action and effect spaces. Of particular interest is the acquisition of inverse models which map a space of sensorimotor goals to a space of motor programs that solve them. For example, this could be a robot learning which movements of the arm and hand can push or throw an object in each of several target locations, or which arm movements allow to produce which displacements of several objects potentially interacting with each other, e.g. in the case of tool use. Specifically, acquiring such repertoires of skills through incremental exploration of the environment has been argued to be a key target for life-long developmental learning [54].

Recently, we developed a formal framework called "Intrinsically motivated goal exploration processes" (IMGEPs), enabling study agents that generate and sequence their own goals to learn world models and skill repertoirs, that is both more compact and more general than our previous models [82]. We experimented several implementations of these processes in a complex robotic setup with multiple objects (see Fig. 7), associated to multiple spaces of parameterized reinforcement learning problems, and where the robot can learn how to use certain objects as tools to manipulate other objects. We analyzed how curriculum learning is automated in this unsupervised multi-goal exploration process, and compared the trajectory of exploration and learning of these spaces of problems with the one generated by other mechanisms such as hand-designed learning curriculum, or exploration targeting a single space of problems, and random motor exploration. We showed that learning several spaces of diverse problems can be more efficient for learning complex skills than only trying to directly learn these complex skills. We illustrated the computational efficiency of IMGEPs as these robotic experiments use a simple memory-based low-level policy representations and search algorithm, enabling the whole system to learn online and incrementally on a Raspberry Pi 3.

Figure 7. Robotic setup. Left: a Poppy Torso robot (the learning agent) is mounted in front of two joysticks. Right: full setup: a Poppy Ergo robot (seen as a robotic toy) is controlled by the right joystick and can hit a tennis ball in the arena which changes some lights and sounds.

In order to run many systematic scientific experiments in a shorter time, we scaled up this experimental setup to a platform of 6 identical Poppy Torso robots, each of them having the same environment to interact with. Every robot can run a different task with a specific algorithm and parameters each (see Fig. 8). Moreover, each Poppy Torso can also perceives the motion of a second Poppy Ergo robot, than can be used, this time, as a distractor performing random motions to complicate the learning problem. 12 top cameras and 6 head cameras can dump video streams during experiments, in order to record video datasets. This setup is now used to perform more experiments to compare different variants of curiosity-driven learning algorithms.

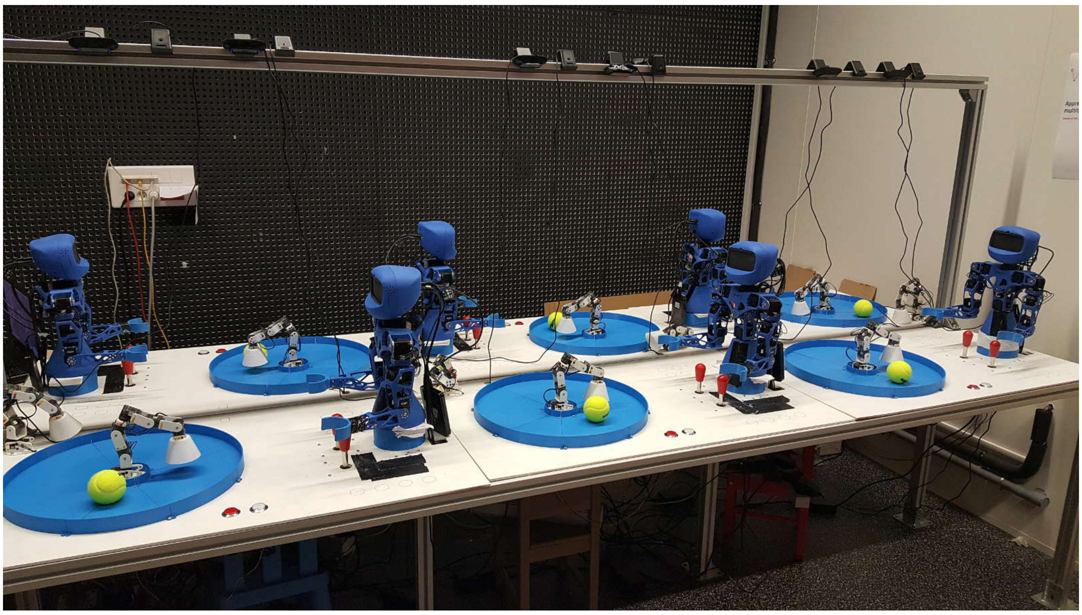


Figure 8. Platform of 6 robots with identical environment: joysticks, Poppy Ergo, ball in an arena, and a distractor. The central bar supports the 12 top cameras.

7.2.1.2. Leveraging the Malmo Minecraft platform to study IMGEP in rich simulations

We continued to leverage the Malmo platform to study curiosity-driven learning applied to multi-goal reinforcement learning tasks (https://github.com/Microsoft/malmo). The first step was to implement an environment called Malmo Mountain Cart (MMC), designed to be well suited to study multi-goal reinforcement learning (see figure [9]). We then showed that IMGEP methods could efficiently explore the MMC environment without any extrinsic rewards. We further showed that, even in the presence of distractors in the goal space, IMGEP methods still managed to discover complex behaviors such as reaching and swinging the cart, especially Active Model Babbling which ignored distractors by monitoring learning progress.

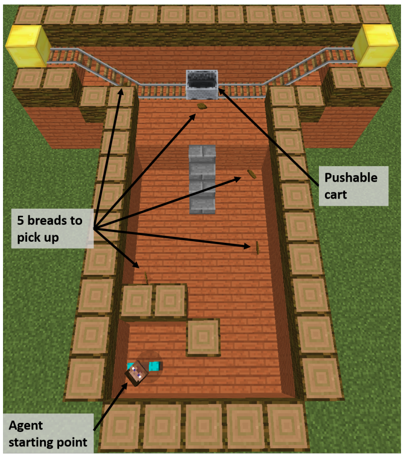


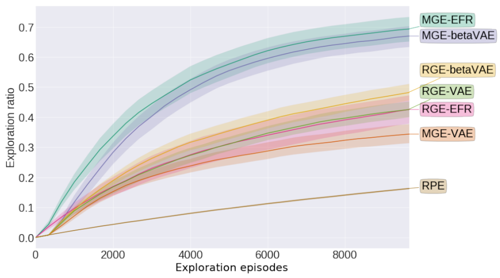
Figure 9. Malmo Mountain Cart. In this episodic environment the agent starts at the bottom left corner of the arena and is able to act on the environment using 2 continuous action commands: move and strafe. If the agent manages to get out of its starting area it may be able to collect items dispatched within the arena. If the agent manages to climb the stairs it may get close enough to the cart to move it along its railroad.

7.2.1.3. Unsupervised Learning of Modular Goal Spaces for Intrinsically Motivated Goal Exploration

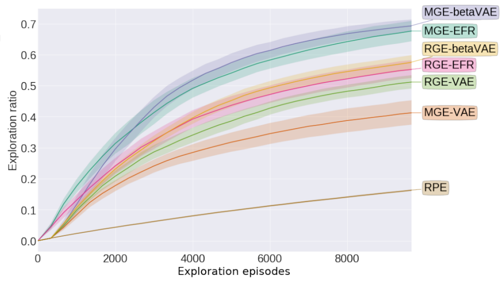
Intrinsically motivated goal exploration algorithms enable machines to discover repertoires of policies that produce a diversity of effects in complex environments. These exploration algorithms have been shown to allow real world robots to acquire skills such as tool use in high-dimensional continuous state and action spaces, as shown in previous sections. However, they have so far assumed that self-generated goals are sampled in a specifically engineered feature space, limiting their autonomy. We have proposed an approach using deep representation learning algorithms to learn an adequate goal space. This is a developmental 2-stage approach: first, in a perceptual learning stage, deep learning algorithms use passive raw sensor observations of world changes to learn a corresponding latent space; then goal exploration happens in a second stage by sampling goals in this latent space. We made experiments with a simulated robot arm interacting with an object, and we show that exploration algorithms using such learned representations can closely match, and even sometimes improve, the performance obtained using engineered representations. This work was presented at ICLR 2018 [136].

However, in the case of more complex environments containing multiple objects or distractors, an efficient exploration requires that the structure of the goal space reflects the one of the environment. We studied how the structure of the learned goal space using a representation learning algorithm impacts the exploration phase. In particular, we studied how disentangled representations compare to their entangled counterparts in a paper published at CoRL 2019 [101], associated with a blog post available at: https://openlab-flowers.inria.fr/t/discovery-of-independently-controllable-features-through-autonomous-goal-setting/494.

Those ideas were evaluated on a simple benchmark where a seven joints robotic arm evolves in an environment containing two balls. One of the ball can be grasped by the arm and moved around whereas the second one acts as a distractor: it cannot be grasped by the robotic arm and moves randomly across the environment.



(a) Small exploration noise ($\sigma = 0.05$)



(b) Large exploration noise ($\sigma = 0.1$)

Figure 10. Exploration ratio during exploration for different exploration noises. Architectures with disentangled representations as a goal space (βVAE) explore more than those with entangled representations (VAE). Similarly modular architectures (MGE) explore more than flat architecture (RGE).

Our results showed that using a disentangled goal space leads to better exploration performances than an entangled goal space: the goal exploration algorithm discovers a wider variety of outcomes in less exploration steps (see Figure 10). We further showed that when the representation is disentangled, one can leverage it by sampling goals that maximize learning progress in a modular manner. Lastly, we have shown that the measure of learning progress, used to drive curiosity-driven exploration, can be used simultaneously to discover abstract independently controllable features of the environment.

Finally, we experimented the applicability of those principles on a real-world robotic setup, where a 6-joint robotic arm learns to manipulate a ball inside an arena, by choosing goals in a space learned from its past experience, presented in this technical report: https://arxiv.org/abs/1906.03967.

7.2.1.4. Monolithic Intrinsically Motivated Modular Multi-Goal Reinforcement Learning

In this project we merged two families of algorithms. The first family is the population-based Intrinsically Motivated Goal Exploration Processes (IMGEP) developped in the team (see [83] for a presentation). In this family, autonomous learning agents sets their own goals and learn to reach them. Intrinsic motivation under the form of absolute learning progress is used to guide the selection of goals to target. In some variations of this framework, goals can be represented as coming from different *modules* or *tasks*. Intrinsic motivations are then used to guide the choice of the next task to target.

The second family encompasses goal-parameterized reinforcement learning algorithms. The first algorithm of this category used an architecture called Universal Value Function Approximators (UVFA), and enabled to train a single policy on an infinite number of goals (continuous goal spaces) [144] by appending the current goal to the input of the neural network used to approximate the value function and the policy. Using a single network allows to share weights among the different goals, which results in faster learning (shared representations). Later, HER [51] introduced a goal replay policy: the actual goal aimed at, could be replaced by a fictive goal when learning. This could be thought of as if the agent were pretending it wanted to reach a goal that it actually reached later on in the trajectory, in place of the true goal. This enables cross-goal learning and speeds up training. Finally, UNICORN [111] proposed to use UVFA to achieve multi-task learning with a discrete task-set.

In this project, we developed CURIOUS [33] (ICML 2019), an intrinsically motivated reinforcement learning algorithm able to achieve both multiple tasks and multiple goals with a single neural policy. It was tested on a custom multi-task, multi-goal environment adapted from the OpenAI Gym Fetch environments [61], see Figure 11. CURIOUS is inspired from the second family as it proposes an extension of the UVFA architecture. Here, the current task is encoded by a one-hot code corresponding to the task id. The goal is of size $\sum_{i=1}^{N} dim(\mathcal{G}_i)$ where $\mathcal{G}_i$ is the goal space corresponding to task T_i. All components are zeroed except the ones corresponding to the current goal g_i of the current task T_i, see Figure 12.

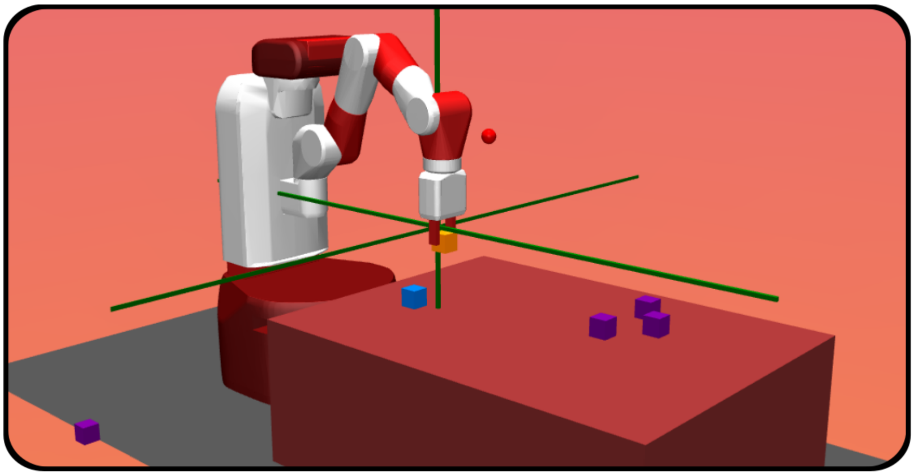


Figure 11. Custom multi-task and multi-goal environment to test the CURIOUS algorithm.

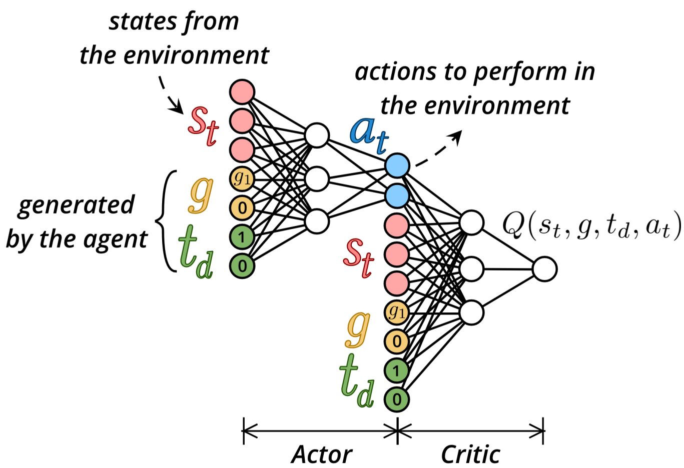


Figure 12. Architecture extended from Universal Value Function Approximators. In this example, the agent is targeting task T_1 among two tasks, each corresponding to a 1 dimension goal.

CURIOUS is also inspired from the first family, as it self-generates its own tasks and goals and uses a measure of learning progress to decide which task to target at any given moment. The learning progress is computed as the absolute value of the difference of non-overlapping window average of the successes or failures

$$LP_i(t) = \frac{|\sum_{\tau=t-2l}^{t-l} S_\tau - \sum_{\tau=t-l}^{t} S_\tau|}{2l},$$

where S_τ describes a success (1) or a failure (0) and l is a time window length. The learning progress is then used in two ways: it guides the selection of the next task to attempt, and it guides the selection of the task to replay. Cross-goal and cross-task learning are achieved by replacing the goal and/or task in the transition by another. When training on one combination of task and goal, the agent can therefore use this sample to learn about other tasks and goals. Here, we decide to replay and learn more on tasks for which the absolute learning progress is high. This helps for several reasons: 1) the agent does not focus on already learned tasks, as the corresponding learning progress is null, 2) the agent does not focus on impossible tasks for the same reason. The agent focuses more on tasks that are being learned (therefore maximizing learning progress), and on tasks that are being forgotten (therefore fighting the problem of forgetting). Indeed, when many tasks are learned in a same network, chances are tasks that are not being attempted often will be forgotten after a while.

In this project, we compare CURIOUS to two baselines: 1) a flat representation algorithm where goals are set from a multi dimensional space including all tasks (equivalent to HER); 2) a task-expert algorithm where a multi-goal UVFA expert policy is trained for each task. The results are shown in Figure 13.

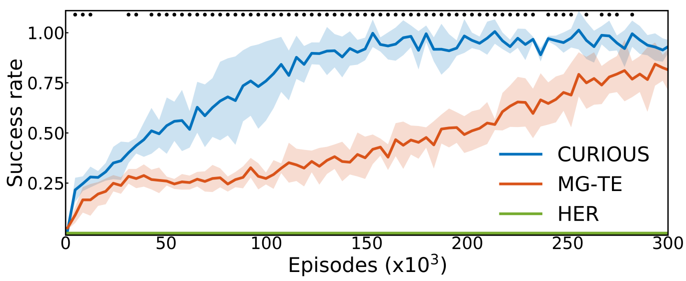


Figure 13. Comparison of CURIOUS to alternative algorithms.

7.2.1.5. Autonomous Multi-Goal Reinforcement Learning with Natural Language

This project follows the CURIOUS project on intrinsically motivated modular multi-goal reinforcement learning [33]. In the CURIOUS algorithm, we presented an agent able to tackle multiple goals of multiple types using a single controller. However, the agent needed to have access to the description of each of the goal types, and their associated reward functions. This represents a considerable amount of prior knowledge for the engineer to encode into the agent. In our new project, the agent builds its own representations of goals, can tackle a growing set of goals, and learns its own reward function, all this through interactions in natural language with a social partner.

The agent does not know any potential goal at first, and act randomly. As it reaches outcomes that are meaningful for the social partner, the social partner provides descriptions of the scene in natural language. The agent stores descriptions and corresponding states for two purposes. First it builds a list of potential goals, reaching back these outcomes that the social partner described. Second, it uses the combination of state and state description to learn a reward function, mapping current state and language descriptions to a binary feedback: 1 if the description is satisfied by the current state, 0 if not.

The agent sets goals to itself from the set of previously discovered descriptions, and is able to learn how to reach them thanks to its learned internal reward function. Concretely, the agent learns a set of 50+ goals from these interactions. We showed co-learning of the reward function and the policy did not produce consequent overhead compared to using an oracle reward function (see Figure 14). This project led to an article accepted at the NeurIPS workshop Visually Grounded Interaction and Language [35]. Current work aims at learning the language model mapping the description into a continuous goal space used as input to the policy and reward function using recurrent networks.

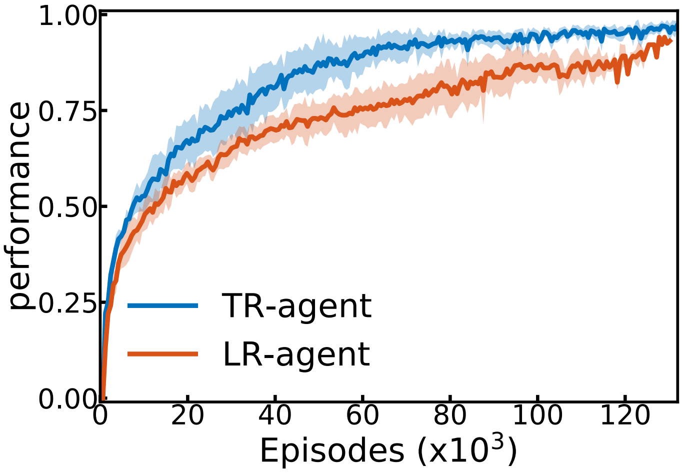


Figure 14. Comparison of performance of agents having access to the true oracle reward function (TR-agent), and agents co-learning policy and reward function (LR-agent).

7.2.1.6. Intrinsically Motivated Exploration and Multi-Goal RL with First-Person Images

The aim of this project is to create an exploration process in first-person 3D environments. Following the work presented in [121], [135], [157] the current algorithm is setup in a similar manner. The agent observes the environment in a first person manner and is given a goal to reach. Furthermore, the goal policy samples goals that encourage the agent to explore the environment.

Currently, there are two ways of representing and sampling goals. Following [157] the goal policy has a buffer of states previously visited by the agent and samples from this buffer the next goals as first person observations. In this setting a goal conditioned reward function is also learned in the form of a reachability network introduced in [142]. On the other hand, following [136], [121] and [135] one can learn a latent representation of a goal using an autoencoder (VAE) and then sample from this generative model or in the latent space. Then, we can use the L2 distance in the latent space as a reward function. The experiments are conducted on a set of Unity environments created in the team.

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

Participants: Remy Portelas [correspondant], Katja Hoffman, Pierre-Yves Oudeyer.

In this work we considered the problem of how a teacher algorithm can enable an unknown Deep Reinforcement Learning (DRL) student to become good at a skill over a wide range of diverse environments. To do so, we studied how a teacher algorithm can learn to generate a learning curriculum, whereby it sequentially samples parameters controlling a stochastic procedural generation of environments. Because it does not initially know the capacities of its student, a key challenge for the teacher is to discover which environments are easy, difficult or unlearnable, and in what order to propose them to maximize the efficiency of learning over the learnable ones. To achieve this, this problem is transformed into a surrogate continuous bandit problem where the teacher samples environments in order to maximize absolute learning progress of its student. We presented ALP-GMM (see figure 15), a new algorithm modeling absolute learning progress with Gaussian mixture models. We also adapted existing algorithms and provided a complete study in the context of DRL. Using parameterized variants of the BipedalWalker environment, we studied their efficiency to personalize a learning curriculum for different learners (embodiments), their robustness to the ratio of learnable/unlearnable environments, and their scalability to non-linear and high-dimensional parameter spaces. Videos and code are available at https://github.com/flowersteam/teachDeepRL.

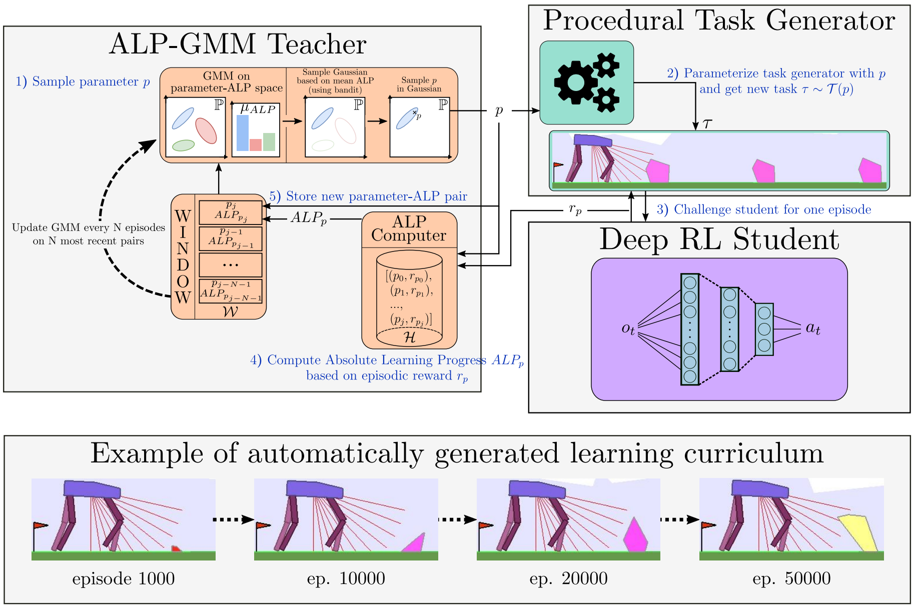


Figure 15. Schematic view of an ALP-GMM teacher's workflow

Overall, this work demonstrated that LP-based teacher algorithms could successfully guide DRL agents to learn in difficult continuously parameterized environments with irrelevant dimensions and large proportions of unfeasible tasks. With no prior knowledge of its student's abilities and only loose boundaries on the task space, ALP-GMM, our proposed teacher, consistently outperformed random heuristics and occasionally even expert-designed curricula (see figure 16). This work was presented at CoRL 2019 [38].

ALP-GMM, which is conceptually simple and has very few crucial hyperparameters, opens-up exciting perspectives inside and outside DRL for curriculum learning problems. Within DRL, it could be applied to previous work on autonomous goal exploration through incremental building of goal spaces [101]. In this case several ALP-GMM instances could scaffold the learning agent in each of its autonomously discovered goal spaces. Another domain of applicability is assisted education, for which current state of the art relies heavily on expert knowledge [68] and is mostly applied to discrete task sets.

7.3. Automated Discovery in Self-Organizing Systems

7.3.1. Curiosity-driven Learning for Automated Discovery of Physico-Chemical Structures

Participants: Chris Reinke [correspondant], Mayalen Etcheverry, Pierre-Yves Oudeyer.

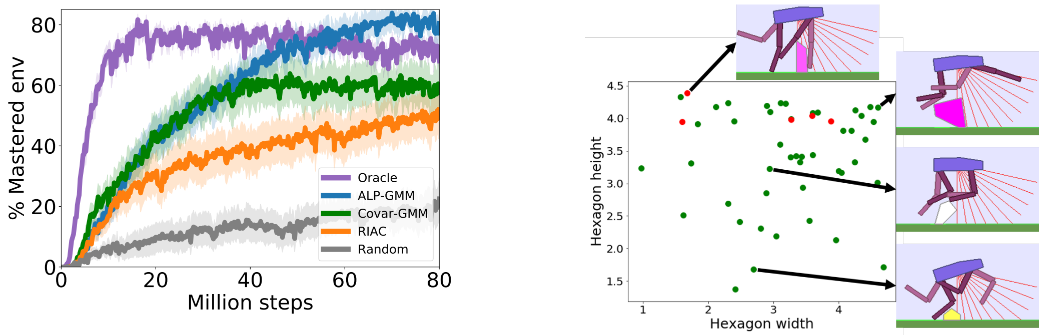


Figure 16. ***Teacher-Student approaches in Hexagon Tracks.*** ***Left:*** *Evolution of mastered tracks for Teacher-Student approaches in Hexagon Tracks. 32 seeded runs (25 for Random) of 80 Millions steps where performed for each condition. The mean performance is plotted with shaded areas representing the standard error of the mean.* ***Right:*** *A visualization of which track distributions of the test-set are mastered (i.e $r_t > 230$, shown by green dots) by an ALP-GMM run after 80 million steps.*

7.3.1.1. Introduction

Intrinsically motivated goal exploration algorithms (IMGEPs) enable machines to discover repertoires of action policies that produce a diversity of effects in complex environments. In robotics, these exploration algorithms have been shown to allow real world robots to acquire skills such as tool use [81] [55]. In other domains such as chemistry and physics, they open the possibility to automate the discovery of novel chemical or physical structures produced by complex dynamical systems [134]. However, they have so far assumed that self-generated goals are sampled in a specifically engineered feature space, limiting their autonomy. Recent work has shown how unsupervised deep learning approaches could be used to learn goal space representations [136] but they have used precollected data to learn the representations. This project studies how IMGEPs can be extended and used for automated discovery of behaviours of dynamical systems in physics or chemistry without using assumptions or knowledge about such systems.

As a first step towards this goal we choose Lenia [66], a simulated high-dimensional complex dynamical system, as a target system. Lenia is a continuous cellular automaton where diverse visual structures can self-organize (Fig.17, c). It consists of a two-dimensional grid of cells $A \in [0,1]^{256\times256}$ where the state of each cell is a real-valued scalar activity $A^t(x) \in [0,1]$. The state of cells evolves over discrete time steps t. The activity change is computed by integrating the activity of neighbouring cells. Lenia's behavior is controlled by its initial pattern $A^{t=1}$ and several settings that control the dynamics of the activity change. Lenia can produce diverse patterns with different dynamics. Most interesting, spatially localized coherent patterns that resemble in their shapes microscopic *animals* can emerge. Our goal was to develop methods that allow to explore a high diversity of such animal patterns.

We could successfully accomplish this goal [30] based on two key contributions of our research: 1) the usage of compositional pattern producing networks (CPPNs) for the generation of initial states for Lenia, and 2) the development of a novel IMGEP algorithm that learns goal representations online during the exploration of the system.

7.3.1.2. 1) CPPNs for the generation of initial states

A key role in the generation of patterns in dynamical systems is their initial state $A^{t=1}$. IMGEPs sample these initial states and apply random perturbations to them during the exploration. For Lenia this state is a two-dimensional grid with 256×256 cells. Performing directly a random sampling of the 256×256 grid cells results in initial patterns that resemble white noise. Such random states result mainly in the emergence of global patterns that spread over the whole state space, complicating the search for spatially localized patterns.

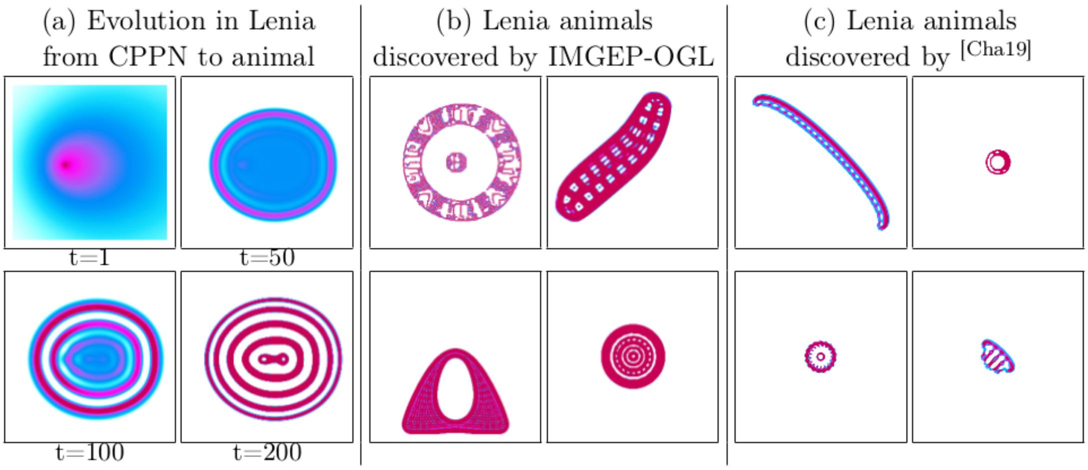


Figure 17. Example patterns produced by the Lenia system. Illustration of the dynamical morphing from an initial CPPN image to an animal (a). The automated discovery (b) is able to find similar complex animals as a human-expert manual search (c) by [66].

We solved the sampling problem for the initial states by using compositional pattern producing networks (CPPNs) [148]. CPPNs are recurrent neural networks that allow the generation of structured initial states (Fig.17, a). The CPPNs are used as part of the system parameters which are explored by the algorithms. They are defined by their network structure (number of neurons, connections between neurons) and their connection weights. They include a mechanism for random mutation of the weights and structure.

7.3.1.3. 2) IMGEP for Online Learning of Goal Space Representations

We proposed an online goal space learning IMGEP (IMGEP-OGL), which learns the goal space incrementally during the exploration process. A variational autoencoder (VAE) is used to encode Lenia patterns into a 8-dimensional latent representation used as goal space. The training procedure of the VAE is integrated in the goal sampling exploration process by first initializing the VAE with random weights. The VAE network is then trained every K explorations for E epochs on the previously idetnfied patterns during the exploration.

7.3.1.4. Experiments

We evaluated the performance of the novel IMGEP-OGL to other exploration algorithms by comparing the diversity of their identified patterns. Diversity is measured by the spread of the exploration in an *analytic behavior space*. This space is defined by a latent representation space that was build through the training of a VAE to learn the important features over a very large dataset of Lenia patterns identified during the many experiments over all evaluated algorithms. We then augmented that space by concatenating hand-defined features. Each identified Lenia pattern is represented by a specific point in this space. The space was then discretized in a fixed number of areas/bins of equal size. The final diversity measure of each algorithm is the number of areas/bins in which at least one explored pattern exists.

We compared different exploration algorithms to the novel IMGEP-OGL: 1) Random exploration of system parameters, 2) IMGEP-HGS: IMGEP with a hand-defined goal space, 3) IMGEP-PGL: IMGEP with a learned goal space via an VAE by a precollected dataset of Lenia patterns, and 4) IMGEP-RGS: IMGEP with a VAE with random weights that defines the goal space.

The system parameters θ consisted of a CPPN that generates the initial state $A^{t=1}$ for Lenia and 6 further settings defining Lenia's dynamics: $\theta = [\mathrm{CPPN} \rightarrow A^{t=1}, R, T, \mu, \sigma, \beta_1, \beta_2, \beta_3]$. The CPPNs were initialized and mutated by a random process that defines their structure and connection weights as done. The random initialization of the other Lenia settings was done by an uniform distribution and their mutation by a Gaussian distribution around the original values.

7.3.1.5. Results

The diversity of identified patterns in the analytic behavior space show that IMGEP approaches with learned goal spaces via VAEs (PGL, OGL) could identify the highest diversity of patterns overall (Fig. 18, a). They were followed by the IMGEP with a hand-defined goal space (HGS). The lowest performance had the random exploration and the IMGEP with a random goal space (RGS). The advantage of learned goals space approaches (PGL, OGL) over all other approaches was even stronger for the diversity of animal patterns, i.e. the main goal of our exploration (Fig. 18, b).

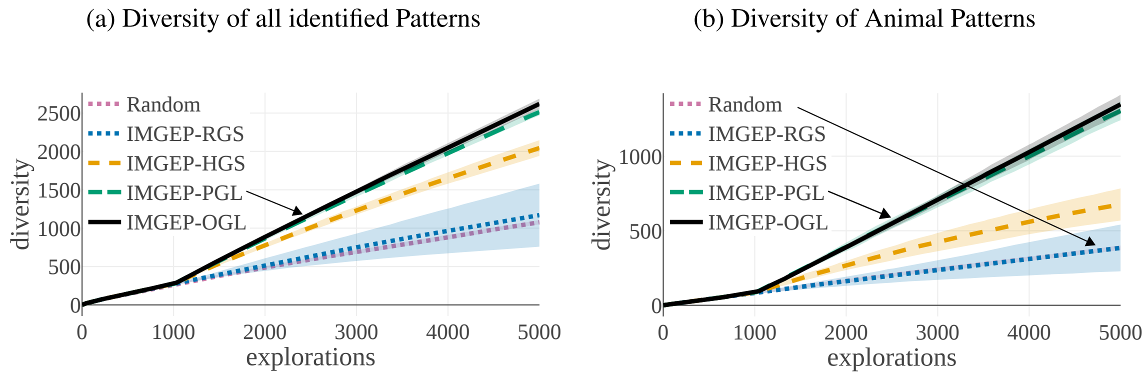


Figure 18. (a) All IMGEPs reach a higher diversity in the analytic behavior space over all patterns than random search. (b) IMGEPs with a learned goal space are especially successful in identifying a diversity of animal patterns. Depicted is the average diversity ($n = 10$) with the standard deviation as shaded area (for some not visible because it is too small).

7.3.1.6. Conclusion

Our goal was to investigate new techniques based on intrinsically motivated goal exploration for the automated discovery of patterns and behaviors in complex dynamical systems. We introduced a new algorithm (IMGEP-OGL) which is capable of learning unsupervised goal space representations during the exploration of an unknown system. Our results for Lenia, a high-dimensional complex dynamical system, show its superior performance over hand-defined goal spaces or random exploration. It shows the same performance as a learned goal space based on precollected data, showing that such a precollection of data is not necessary. We furthermore introduced the usage of CPPNs for the successful initialization of the intial states of the dynamical systems. Both advances allowed us to explore an unknown and high-dimensional dynamical system which shares many similarities with different physical or chemical systems.

7.4. Representation Learning

7.4.1. State Representation Learning in the Context of Robotics

Participants: David Filliat [correspondant], Natalia Diaz Rodriguez, Timothee Lesort, Antonin Raffin, René Traoré, Ashley Hill, Te Sun, Lu Lin, Guanghang Cai, Bunthet Say.

During the DREAM project, we participated in the development of a conceptual framework of open-ended lifelong learning [77] based on the idea of representational re-description that can discover and adapt the states, actions and skills across unbounded sequences of tasks.

In this context, State Representation Learning (SRL) is the process of learning without explicit supervision a representation that is sufficient to support policy learning for a robot. We have finalized and published a large state-of-the-art survey analyzing the existing strategies in robotics control [103], and we developed unsupervised methods to build representations with the objective to be minimal, sufficient, and that encode the relevant information to solve the task. More concretely, we used the developed and open sourced[1] the S-RL toolbox [137] containing baseline algorithms, data generating environments, metrics and visualization tools for assessing SRL methods. Part of this study is the [105] where we present a robustness analysis on Deep unsupervised state representation learning with robotic priors loss functions.

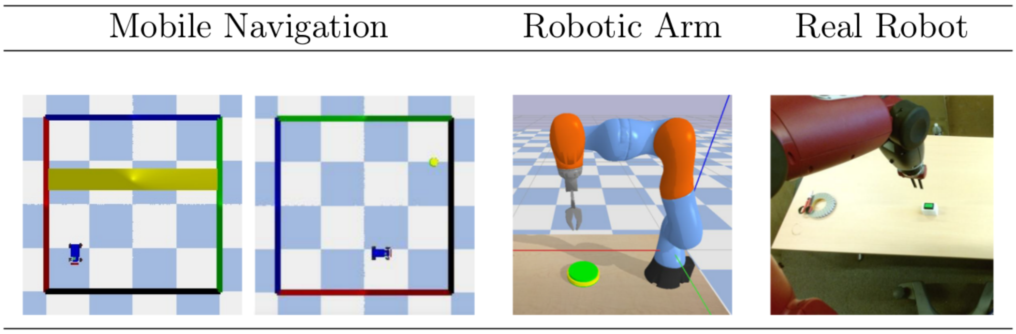


Figure 19. Environments and datasets for state representation learning.

The environments proposed in Fig. 19 are variations of two environments: a 2D environment with a mobile robot and a 3D environment with a robotic arm. In all settings, there is a controlled robot and one or more targets (that can be static, randomly initialized or moving). Each environment can either have a continuous or discrete action space, and the reward can be sparse or shaped, allowing us to cover many different situations.

The evaluation and visualization tools are presented in Fig. 20 and make it possible to qualitatively verify the learned state space behavior (e.g., the state representation of the robotic arm dataset is expected to have a continuous and correlated change with respect to the arm tip position).

We also proposed a new approach that consists of learning a state representation that is split into several parts where each optimizes a fraction of the objectives. In order to encode both target and robot positions, auto-encoders, reward and inverse model losses are used.

The latest work on decoupling feature extraction from policy learning, was presented at the SPIRL workshop at ICLR2019 in New Orleans, LA [138]. We assessed the benefits of state representation learning in goal based robotic tasks, using different self-supervised objectives.

Because combining objectives into a single embedding is not the only option to have features that are *sufficient* to solve the tasks, by stacking representations, we favor *disentanglement* of the representation and prevent objectives that can be opposed from cancelling out. This allows a more stable optimization. Fig. 21 shows the split model where each loss is only applied to part of the state representation.

As using the learned state representations in a Reinforcement Learning setting is the most relevant approach to evaluate the SRL methods, we use the developed S-RL framework integrated algorithms (A2C, ACKTR, ACER, DQN, DDPG, PPO1, PPO2, TRPO) from Stable-Baselines [92], Augmented Random Search (ARS), Covariance Matrix Adaptation Evolutionary Strategy (CMA-ES) and Soft Actor Critic (SAC). Due to its stability, we perform extensive experiments on the proposed datasets using PPO and states learned with the approaches described in [137] along with ground truth (GT).

[1] https://github.com/araffin/robotics-rl-srl

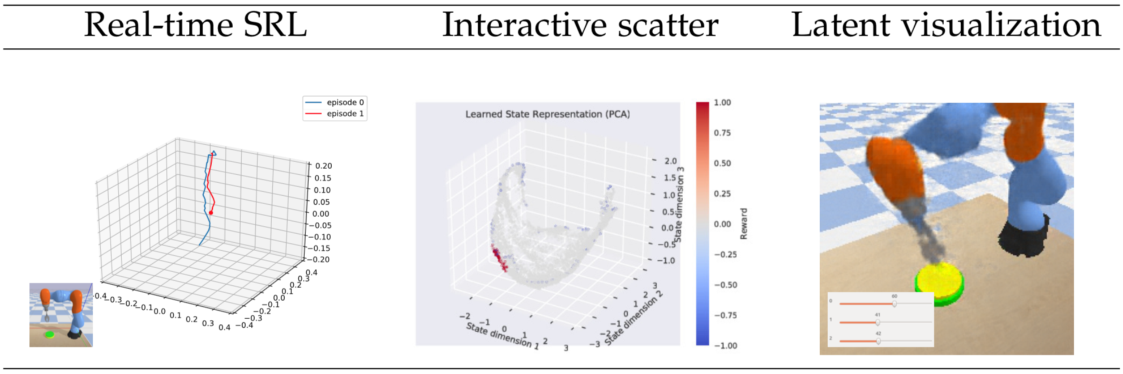


Figure 20. Visual tools for analysing SRL; Left: Live trajectory of the robot in the state space. Center: 3D scatter plot of a state space; clicking on any point displays the corresponding observation. Right: reconstruction of the point in the state space defined by the sliders.

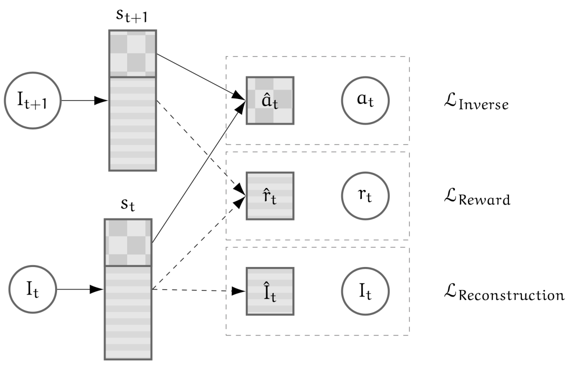


Figure 21. SRL Splits model: combines a reconstruction of an image I, a reward (r) prediction and an inverse dynamic models losses, using two splits of the state representation s. Arrows represent model learning and inference, dashed frames represent losses computation, rectangles are state representations, circles are real observed data, and squares are model predictions.

Ground Truth States	Learned States	RL Performance

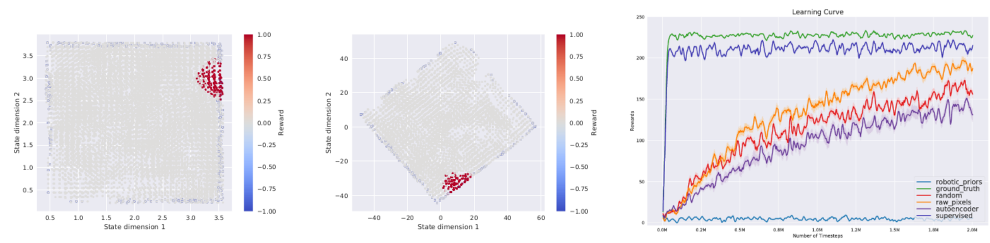


Figure 22. Ground truth states (left), states learned (Inverse and Forward) (center), and RL performance evaluation (PPO) (right) for different baselines in the mobile robot environment. Colour denotes the reward, red for positive, blue for negative and grey for null reward (left and center).

Table 22 illustrates the qualitative evaluation of a state space learned by combining forward and inverse models on the mobile robot environment. It also shows the performance of PPO algorithm based on the states learned by several baseline approaches.

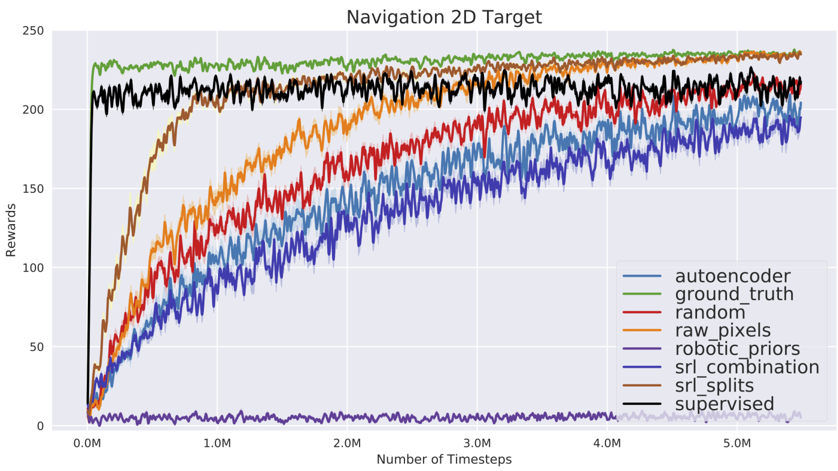


Figure 23. Performance (mean and standard error for 10 runs) for PPO algorithm for different state representations learned in Navigation 2D random target environment.

We verified that our new approach (described in Task 2.1) makes it possible for reinforcement learning to converge faster towards the optimal performance in both environments with the same amount of budget timesteps. Learning curve in Fig. 23 shows that our unsupervised state representation learned with the split model even improves on the supervised case.

7.4.2. Continual learning

Participants: David Filliat [correspondant], Natalia Díaz Rodríguez, Timothee Lesort, Hugo Caselles-Dupré.

Continual Learning (CL) algorithms learn from a stream of data/tasks continuously and adaptively through time to better enable the incremental development of ever more complex knowledge and skills. The main problem that CL aims at tackling is catastrophic forgetting [115], i.e., the well-known phenomenon of a neural network experiencing a rapid overriding of previously learned knowledge when trained sequentially on new data. This is an important objective quantified for assessing the quality of CL approaches, however, the almost exclusive focus on catastrophic forgetting by continual learning strategies, lead us to propose a set of comprehensive, implementation independent metrics accounting for factors we believe have practical implications worth considering with respect to the deployment of real AI systems that learn continually, and in "Non-static" machine learning settings. In this context we developed a framework and a set of comprehensive metrics [78] to tame the lack of consensus in evaluating CL algorithms. They measure Accuracy (A), Forward and Backward (*/remembering*) knowledge transfer (FWT, BWT, REM), Memory Size (MS) efficiency, Samples Storage Size (SSS), and Computational Efficiency (CE). Results on iCIFAR-100 classification sequential class learning is in Table 24.

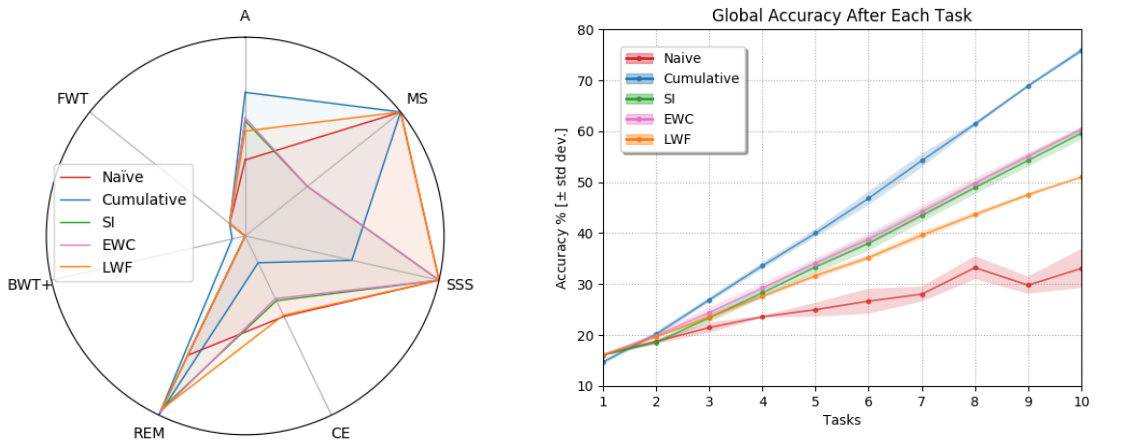


Figure 24. (left) Spider chart: CL metrics per strategy (larger area is better) and (right) Accuracy per CL strategy computed over the fixed test set.

Generative models can also be evaluated from the perspective of Continual learning which we investigated in our work [102]. This work aims at evaluating and comparing generative models on disjoint sequential image generation tasks. We study the ability of Generative Adversarial Networks (GANS) and Variational Auto-Encoders (VAEs) and many of their variants to learn sequentially in continual learning tasks. We investigate how these models learn and forget, considering various strategies: rehearsal, regularization, generative replay and fine-tuning. We used two quantitative metrics to estimate the generation quality and memory ability. We experiment with sequential tasks on three commonly used benchmarks for Continual Learning (MNIST, Fashion MNIST and CIFAR10). We found (see Figure 26) that among all models, the original GAN performs best and among Continual Learning strategies, generative replay outperforms all other methods. Even if we found satisfactory combinations on MNIST and Fashion MNIST, training generative models sequentially on CIFAR10 is particularly instable, and remains a challenge. This work has been published at the NIPS workshop on Continual Learning 2018.

Another extension of previous section on state representation learning (SRL) to the continual learning setting is in our paper [65]. This work proposes a method to avoid catastrophic forgetting when the environment changes using generative replay, i.e., using generated samples to maintain past knowledge. State representations are learned with variational autoencoders and automatic environment change is detected through VAE

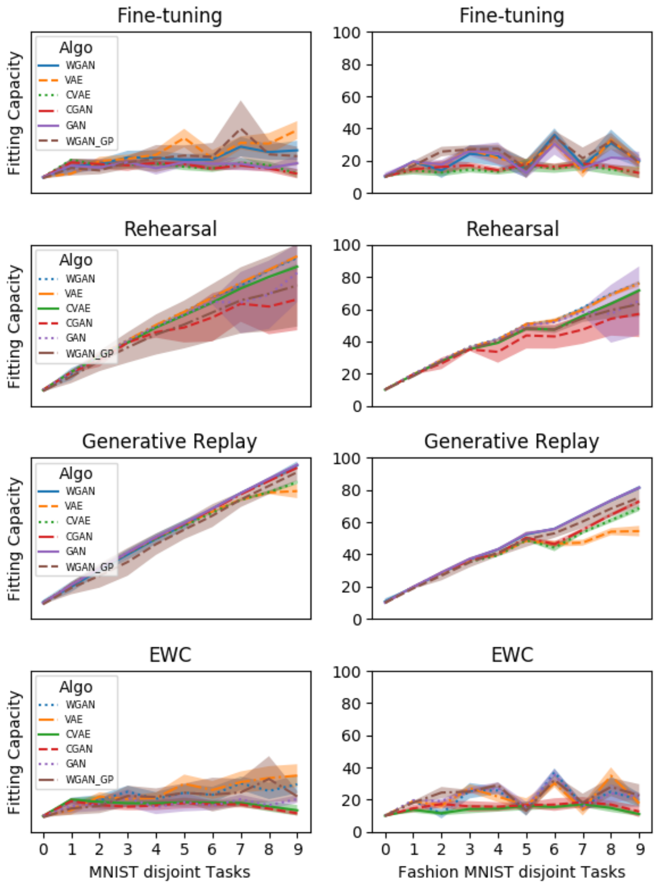


Figure 25. Means and standard deviations over 8 seeds of Fitting Capacity metric evaluation of VAE, CVAE, GAN, CGAN and WGAN. The four considered CL strategies are: Fine Tuning, Generative Replay, Rehearsal and EWC. The setting is 10 disjoint tasks on MNIST and Fashion MNIST.

reconstruction error. Results show that using a state representation model learned continually for RL experiments is beneficial in terms of sample efficiency and final performance, as seen in Figure 26. This work has been published at the NIPS workshop on Continual Learning 2018 and is currently being extended.

The experiments were conducted in an environment built in the lab, called Flatland [64]. This is a lightweight first-person 2-D environment for Reinforcement Learning (RL), designed especially to be convenient for Continual Learning experiments. Agents perceive the world through 1D images, act with 3 discrete actions, and the goal is to learn to collect edible items with RL. This work has been published at the ICDL-Epirob workshop on Continual Unsupervised Sensorimotor Learning 2018, and was accepted as oral presentation.

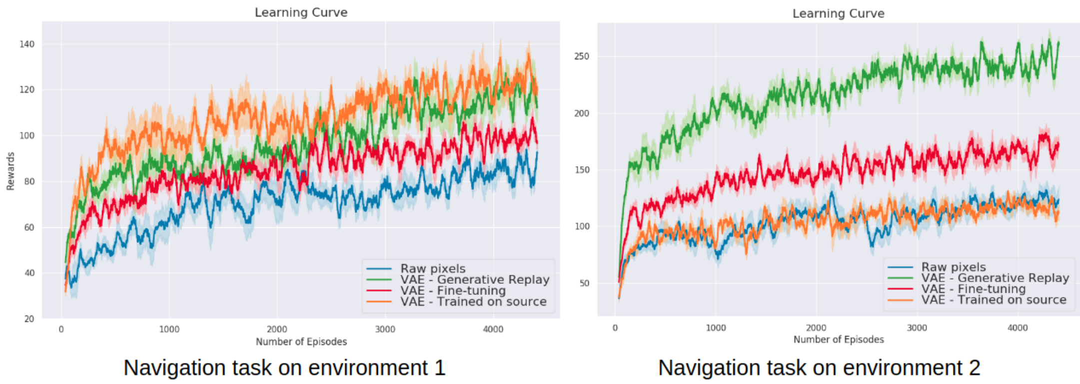


Figure 26. Mean reward and standard error over 5 runs of RL evaluation using PPO with different types of inputs. Fine-tuning and Generative Replay models are trained sequentially on the first and second environment, and then used to train a policy for both tasks. Generative Replay outperforms all other methods. It shows the need for continually learning features in State Representation Learning in settings where the environment changes.

In the last year, we published a survey on continual learning models, metrics and contributed a CL framework to categorize the approaches on this area [104]. Figure 27 shows the different approaches cited and the strategies proposed and a small subset of examples analyzed.

We also worked on validating a distillation approach for multitask learning in a continual learning reinforcement learning setting [152], [153].

Applying State Representation Learning (SRL) into a continual learning setting of reinforcement learning was possible by learning a compact and efficient representation of data that facilitates learning a policy. The proposed a CL algorithm based on distillation does not manually need to be given a task indicator at test time, but learns to infer the task from observations only. This allows to successfully apply the learned policy on a real robot.

We present 3 different 2D navigation tasks to a 3 wheel omni-directional robot to be learned to be solved sequentially. The robot has first access to task 1 only, and then to task 2 only, and so on. It should learn a single policy that solves all tasks and be applicable in a real life scenario. The robot can perform 4 high level discrete actions (move left/right, move up/down). The tasks where the method was validated are in Fig. 28:

Task 1: Target Reaching (TR): Reaching a red target randomly positioned.

Task 2: Target Circling (TC): Circling around a fixed blue target.

Task 3: Target Escaping (TE): Escaping a moving robot.

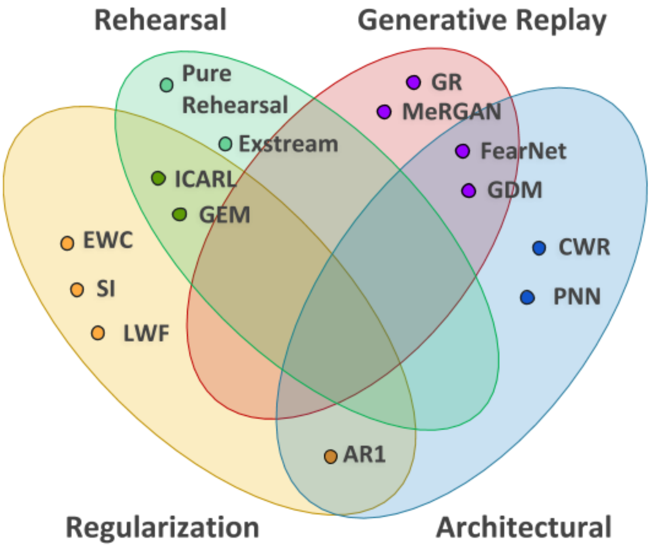


Figure 27. Venn diagram of some of the most popular CL strategies w.r.t the main approaches in the literature (CWR,PNN EWC, SI, LWF, ICARL, GEM, FearNet, GDM, ExStream, GR, MeRGAN, and AR1. Rehearsal and Generative Replay upper categories can be seen as a subset of replay strategies. Better viewed in color [104].

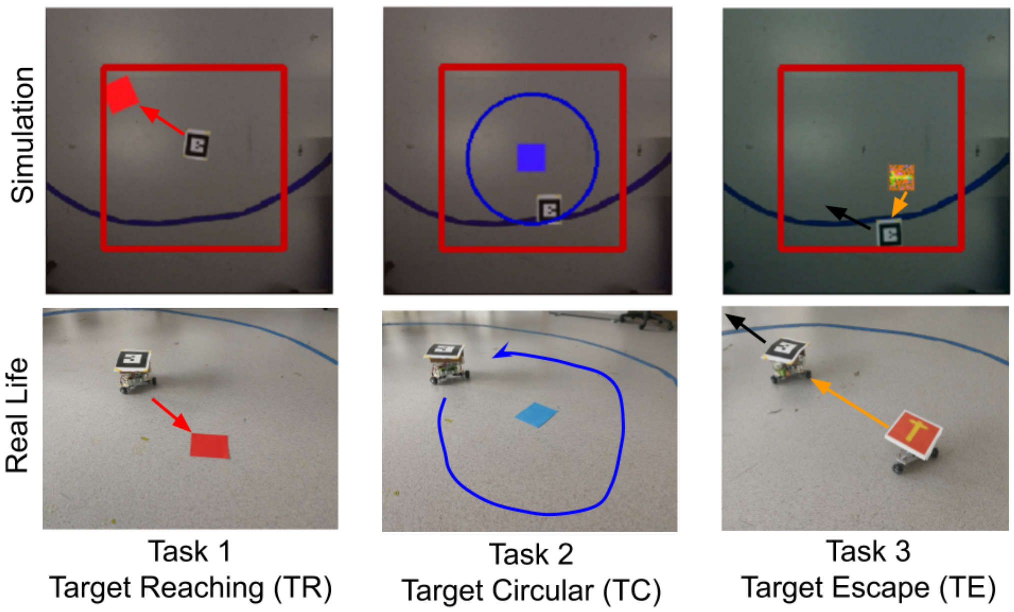


Figure 28. The three tasks, in simulation (top) and in real life (bottom), sequentially experienced. Learning is performed in simulation, the real life setting is only used at test time.

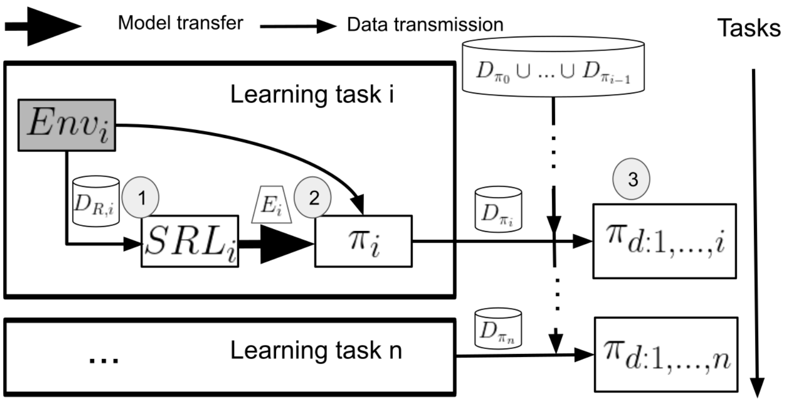


Figure 29. White cylinders are for datasets, gray squares for environments, and white squares for learning algorithms, whose name correspond to the model trained. Each task i is learned sequentially and independently by first generating a dataset $D_{R,i}$ with a random policy to learn a state representation with an encoder E_i with an SRL method (1), then we use E_i and the environment to learn a policy π_i in the state space (2). Once trained, π_i is used to create a distillation dataset D_{π_i} that acts as a memory of the learned behaviour. All policies are finally compressed into a single policy $\pi_{d:1,...,i}$ by merging the current dataset D_{π_i} with datasets from previous tasks $D_{\pi_1} \cup ... \cup D_{\pi_{i-1}}$ and using distillation (3).

DisCoRL (Distillation for Continual Reinforcement learning) is a modular, effective and scalable pipeline for continual RL. This pipeline uses policy distillation for learning without forgetting, without access to previous environments, and without task labels in order to transfer policies into real life scenarios [152]. It was presented as an approach for continual reinforcement learning that sequentially summarizes different learned policies into a dataset to distill them into a student model. Some loss in performance may occur while transferring knowledge from teacher to student, or while transferring a policy from simulation to real life. Nevertheless, the experiments show promising results when learning tasks sequentially, in simulated environments and real life settings.

The overview of DisCoRL full pipeline for Continual Reinforcement Learning is in Fig. 29.

7.4.3. Disentangled Representation Learning for agents

Participants: Hugo Caselles-Dupré [correspondant], David Filliat.

Finding a generally accepted formal definition of a disentangled representation in the context of an agent behaving in an environment is an important challenge towards the construction of data-efficient autonomous agents. Higgins et al. (2018) recently proposed Symmetry-Based Disentangled Representation Learning, a definition based on a characterization of symmetries in the environment using group theory. We build on their work and make observations, theoretical and empirical, that lead us to argue that Symmetry-Based Disentangled Representation Learning cannot only be based on static observations: agents should interact with the environment to discover its symmetries.

Our research was published in NeuRIPS 2019 [32] at Vancouver, Canada.

7.5. Tools for Understanding Deep Learning Systems

7.5.1. Explainable Deep Learning

Participants: Natalia Díaz Rodríguez [correspondant], Adrien Bennetot.

Together with Segula Technologies and Sorbonne Université, ENSTA Paris has been working on eXplainable Artificial Intelligence (XAI) in order to make machine learning more interpretable. While opaque decision systems such as Deep Neural Networks have great generalization and prediction skills, their functioning does not allow obtaining detailed explanations of their behaviour. The objective is to fight the trade-off between performance and explainability by combining connectionist and symbolic paradigms [47].

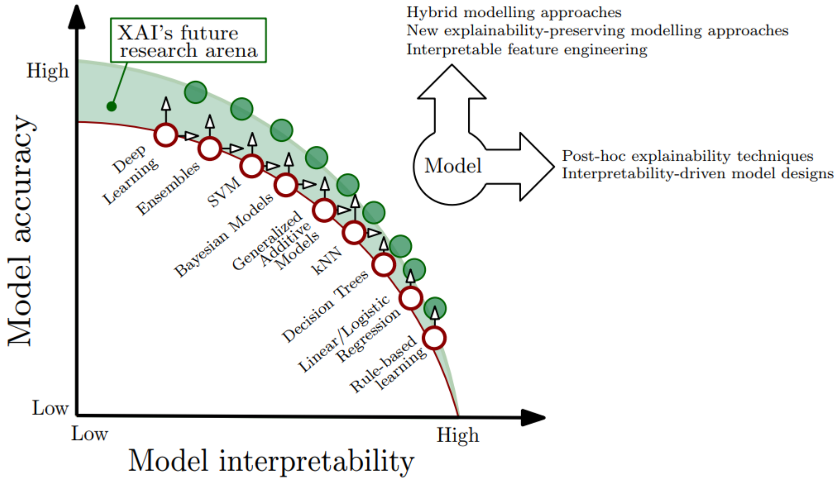


Figure 30. Trade-off between model interpretability and performance, and a representation of the area of improvement where the potential of XAI techniques and tools resides [46].

Broad consensus exists on the importance of interpretability for AI models. However, since the domain has only recently become popular, there is no collective agreement on the different definitions and challenges that constitute XAI. The first step is therefore to summarize previous efforts made in this field. We presented a taxonomy of XAI techniques in [46] and we are currently working on a prediction model that generates itself an explanation of its rationale in natural language while keeping performance as close as possible to the the state of the art [47].

7.5.2. Methods for Statistical Comparison of RL Algorithms

Participants: Cédric Colas [correspondant], Pierre-Yves Oudeyer, Olivier Sigaud.

Following a first article in 2018 [71], we pursued the objective of providing key tools to robustly compare reinforcement learning (RL) algorithms to practioners and researchers. In this year's extension, we compiled a hitchhicker's guide for statistical comparisons of RL algorithms. In particular, we provide a list of statistical tests adapted to compare RL algorithms and compare them in terms of false positive rate and statistial power. In particular, we study the robustness of these tests when their assumptions are violated (non-normal distributions of performances, different distributions, unknown variance, unequal variances etc). We provided an extended study using data from synthetic performance distributions, as well as empirical distributions obtained from running state-of-the-art RL algorithms (TD3 [84] and SAC [88]). From these results we draw a selection of advice for researchers. This study led to an article accepted at the ICLR conference workshop on Reproducibility in Machine Learning [34], to be sumitted to the Neural Networks journal.

7.5.3. Knowledge engineering tools for neural-symbolic learning

Participants: Natalia Díaz Rodríguez [correspondant], Adrien Bennetot.

Symbolic artificial intelligence methods are experiencing a come-back in order to provide deep representation methods the explainability they lack. In this area, a survey on RDF stores to handle ontology-based triple databases has been contributed [97], as well as the use of neural-symbolic tools that aim at integrating both neural and symbolic representations [58].

7.6. Applications in Educational Technologies

7.6.1. Machine Learning for Adaptive Personalization in Intelligent Tutoring Systems

Participants: Pierre-Yves Oudeyer [correspondant], Benjamin Clément, Didier Roy, Helene Sauzeon.

7.6.1.1. The Kidlearn project

Kidlearn is a research project studying how machine learning can be applied to intelligent tutoring systems. It aims at developing methodologies and software which adaptively personalize sequences of learning activities to the particularities of each individual student. Our systems aim at proposing to the student the right activity at the right time, maximizing concurrently his learning progress and its motivation. In addition to contributing to the efficiency of learning and motivation, the approach is also made to reduce the time needed to design ITS systems.

We continued to develop an approach to Intelligent Tutoring Systems which adaptively personalizes sequences of learning activities to maximize skills acquired by students, taking into account the limited time and motivational resources. At a given point in time, the system proposes to the students the activity which makes them progress faster. We introduced two algorithms that rely on the empirical estimation of the learning progress, **RiARiT** that uses information about the difficulty of each exercise and **ZPDES** that uses much less knowledge about the problem.

The system is based on the combination of three approaches. First, it leverages recent models of intrinsically motivated learning by transposing them to active teaching, relying on empirical estimation of learning progress provided by specific activities to particular students. Second, it uses state-of-the-art Multi-Arm Bandit (MAB) techniques to efficiently manage the exploration/exploitation challenge of this optimization process. Third, it leverages expert knowledge to constrain and bootstrap initial exploration of the MAB, while requiring only coarse guidance information of the expert and allowing the system to deal with didactic gaps in its knowledge. The system was evaluated in several large-scale experiments relying on a scenario where 7-8 year old schoolchildren learn how to decompose numbers while manipulating money [68]. Systematic experiments were also presented with simulated students.

7.6.1.2. Kidlearn Experiments 2018-2019: Evaluating the impact of ZPDES and choice on learning efficiency and motivation

An experiment was held between mars 2018 and July 2019 in order to test the Kidlearn framework in classrooms in Bordeaux Metropole. 600 students from Bordeaux Metropole participated in the experiment. This study had several goals. The first goal was to evaluate the impact of the Kidlearn framework on motivation and learning compared to an Expert Sequence without machine learning. The second goal was to observe the impact of using learning progress to select exercise types within the ZPDES algorithm compared to a random policy. The third goal was to observe the impact of combining ZPDES with the ability to let children make different kinds of choices during the use of the ITS. The last goal was to use the psychological and contextual data measures to see if correlation can be observed between the students psychological state evolution, their profile, their motivation and their learning. The different observations showed that generally, algorithms based on ZPDES provided a better learning experience than an expert sequence. In particular, they provide a better motivating and enriching experience to self-determined students. Details of these new results, as well as the overall results of this project, are presented in Benjamin Clément PhD thesis [69] and are currently being processed to be published.

7.6.1.3. Kidlearn and Adaptiv'Math

The algorithms developed during the Kidlearn project and Benjamin Clement thesis [69] are being used in an innovation partnership for the development of a pedagogical assistant based on artificial intelligence intended for teachers and students of cycle 2. The algorithms are being written in typescript for the need of the project. The expertise of the team in creating the pedagogical graph and defining the graph parameters used for the algorithms is also a crucial part of the role of the team for the project. One of the main goal of the team here is to transfer technologies developed in the team in a project with the perspective of industrial scaling and see the impact and the feasibility of such scaling.

7.6.1.4. Kidlearn for numeracy skills with individuals with autism spectrum disorders

Few digital interventions targeting numeracy skills have been evaluated with individuals with autism spectrum disorder (ASD) [114]. Yet, some children and adolescents with ASD have learning difficulties and/or a significant academic delay in mathematics. While ITS are successfully developed for typically developed students to personnalize learning curriculum and then to foster the motivation-learning coupling, they are not or fewly proposed today to student with specific needs. The objective of this pilot study is to test the feasibility of a digital intervention using an STI with high school students with ASD and/or intellectual disability. This application (KidLearn) provides calculation training through currency exchange activities, with a dynamic exercise sequence selection algorithm (ZPDES). 24 students with ASD and/or DI enrolled in specialized classrooms were recruited and divided into two groups: 14 students used the KidLearn application, and 10 students received a control application. Pre-post evaluations show that students using KidLearn improved their calaculation performance, and had a higher level of motivation at the end of the intervention than the control group. These results encourage the use of an STI with students with specific needs to teach numeracy skills, , but need to be replicated on a larger scale. Suggestions for adjusting the interface and teaching method are suggested to improve the impact of the application on students with autism. (Paper is in progress).

7.6.2. Curiosity-driven interaction systems for education

Participants: Pierre-Yves Oudeyer, Hélène Sauzéon [correspondant], Mehdi Alami, Didier Roy, Edith Law.

Three studies have been developed and conducted to newly design curiosity-driven interaction systems aiming to foster learning performance across lifespan : the first two studies include children and the last one includes the older adults.

The first study regards a new interactive robotic system to foster curiosity-driven learning. This led to an article in CHI 2019 [29]. In this work, we explored whether a social peer robot's verbal expression of curiosity can be perceived by participants, produce emotional or behavioural contagion effects, and impact learning. In a between-subject experiment involving 30 participants, a peer robot was manipulated to verbally express: curiosity, curiosity plus rationale, or no curiosity (neutral), within the context of LinkedIt!, a cooperative game we designed for teaching students how to classify rocks. Results show that participants were able to reliably recognize curiosity in the robot and curious robots can be used to elicit significantly more curiosity-driven behaviours among participants.

The second study regards a new interactive educational application to foster curiosity-driven question-asking in children. This study has been performed during the Master 2 internship of Mehdi Alaimi co-supervised by H. Sauzéon, E. Law and PY Oudeyer. The paper submission to CHI'20 is just accepted in december 2019 (« Pedagogical Agents for Fostering Question-Asking Skills in Children »). It addresses a key challenge for 21st-century schools, i.e., teaching diverse students with varied abilities and motivations for learning, such as curiosity within educational settings. Among variables eliciting curiosity state, one is known as « knowledge gap », which is a motor for curiosity-driven exploration and learning. It leads to question-asking which is an important factor in the curiosity process and the construction of academic knowledge. However, children questions in classroom are not really frequent and don't really necessitate deep reasoning. Determined to improve children's curiosity, we developed a digital application aiming to foster curiosity-related question-asking from texts and their perception of curiosity. To assess its efficiency, we conducted a study with 95 fifth grade students of Bordeaux elementary schools. Two types of interventions were designed, one trying to focus children on the construction of low-level question (i.e. convergent) and one focusing them on high-level questions (i.e. divergent) with the help of prompts or questions starters models. We observed that both interventions increased the number of divergent questions, the question fluency performance, while they did not significantly improve the curiosity perception despite high intrinsic motivation scores they have elicited in children. The curiosity-trait score positively impacted the divergent question score under divergent condition, but not under convergent condition. The overall results supported the efficiency and usefulness of digital applications for fostering children's curiosity that we need to explore further.

Finaly, the third study investigates the role of intrinsic motivation in spatial learning in late adulthood [25]. We investigated age differences in memory for spatial routes that were either actively (i.e., intinsic motivation condition) or passively (i.e., control condition) encoded. A series of virtual environments were created and presented to 20 younger (Mean age = 19.71) and 20 older (Mean age = 74.55) adults, through a cardboard viewer. During encoding, participants explored routes presented within city, park, and mall virtual environments, and were later asked to re-trace their travelled routes. Critically, participants encoded half the virtual environments by passively viewing a guided tour along a pre-selected route, and half through active exploration with volitional control of their movements by using a button press on the viewer. During retrieval, participants were placed in the same starting location and asked to retrace the previously traveled route. We calculated the percentage overlap in the paths travelled at encoding and retrieval, as an indicator of spatial memory accuracy, and examined various measures indexing individual differences in their cognitive approach and visuo-spatial processing abilities. Results showed that active navigation, compared to passive viewing during encoding, resulted in a higher accuracy in spatial memory, with the magnitude of this memory enhancement being significantly larger in older than in younger adults. Results suggest that age-related deficits in spatial memory can be reduced by active encoding. In other words, this means that conditions where intrinsic motivation is involved, reduce negative effects of aging on spatial learning.

7.6.3. Poppy Education: Designing and Evaluating Educational Robotics Kits

Participants: Pierre-Yves Oudeyer, Didier Roy [correspondant], Thibault Desprez.

The Poppy Education project aims to create, evaluate and disseminate all-inclusive pedagogical kits, open-source and low cost, for teaching computer science and robotics in secondary education and higher education, scientific literacy centers and Fablabs.

It is designed to help young people to take ownership with concepts and technologies of the digital world, and provide the tools they need to allow them to become actors of this world, with a considerable socio-economic potential. It is carried out in collaboration with teachers and several official french structures (French National Education, High schools, engineering schools, ...).

Poppy Education is based on the robotic platform poppy (open-source platform for the creation, use and sharing of interactive 3D printed robots), including:

- web interface connection (see figure 31)
- Poppy Humanoid, a robust and complete robotics platform designed for genuine experiments in the real world and that can be adapted to specific user needs.
- Poppy Torso, a variant of Poppy Humanoid that can be easily installed on any flat support.
- Ergo Jr, a robotic arm. Durable and inexpensive, it is perfect to be used in class. It can be programmed in Python, directly from a web browser, using Ipython notebooks (an interactive terminal, in a web interface for the Python Programming Language).
- Snap. The visual programming system Snap (see figure 32), which is a variant of Scratch. Its features allow a thorough introduction of information technology. Several specific "blocks" have been developed for this.
- C++, Java, Matlab, Ruby, Javascript, etc. thanks to a REST API that allows you to send commands and receive information from the robot with simple HTTP requests.
- Virtual robots (Poppy Humanoid, Torso and Ergo) can be simulated with the free simulator V-REP (see figure 33). It is possible in the classroom to work on the simulated model and then allow students to run their program on the physical robot.
- Virtual robots (Poppy Ergo) can also be simulated with a 3D web viewer (see figure 34).

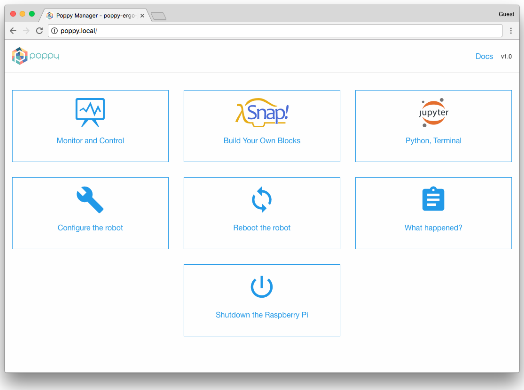


Figure 31. Home page on http://poppy.local

Figure 32. The visual programming system Snap

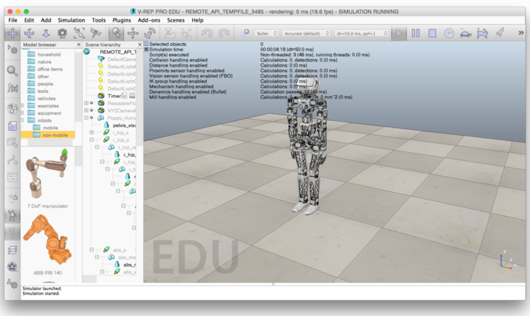


Figure 33. V-rep

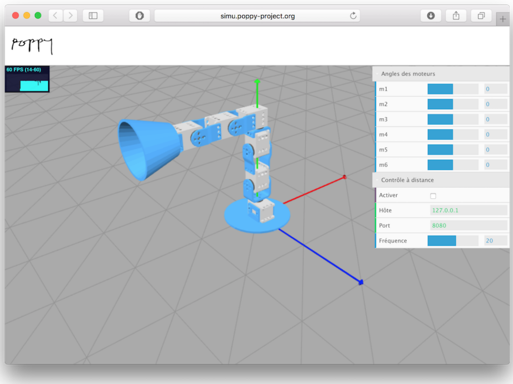


Figure 34. 3D viewer

7.6.3.1. Pedagogical experimentations : Design and experiment robots and the pedagogical activities in classroom.

The robots are designed with the final users in mind. The pedagogical tools of the project (robots and resources) are being created directly with the users and evaluated in real life by experiments. So teachers and researchers co-create activities, test them with students in class-room, share their experience and develop the platform as needed [126].
The activities were designed mainly with Snap! and Python. Most activities use Poppy Ergo Jr, but some use Poppy Torso (mostly in higher school due to its cost).
The pedagogical experiments in classroom carried out during the first year of the project notably allowed to create and experiment many robotic activities. These activities are designed as pedagogical resources introducing robotics. The main objective of the second year was to make all the activities and resources reusable (with description, documentation and illustration) easily and accessible while continuing the experiments and the diffusion of the robotic kits.

Figure 35. Experiment robots and pedagogical activities in classroom

- Pedagogical working group : the teacher partners continued to use the robots in the classroom and to create and test new classroom activities. We organized some training to help them to discover and learn how to use the robotics platform. Also, an engineer of the Poppy Education team went to visit the teachers in their school to see and to evaluate the pedagogical tools (robots and activities) in a real context of use.
  Five meetings have been organized during the year including all teachers part of the project as well as the Poppy Education team in order to exchange about their experience using the robots as a pedagogical tool, to understand their need and to get some feedback from them. This is helping us to understand better the educational needs, to create and improve the pedagogical tools.
  You can see the videos of pedagogical robotics activities here:
  https://www.youtube.com/playlist?list=PLdX8RO6QsgB7hM_7SQNLvyp2QjDAkkzLn

7.6.3.2. Pedagogical documents and resources

- We continued to improve the documentation of the robotic platform Poppy (https://docs.poppy-project.org/en/) and the documentation has been translated into French (https://docs.poppy-project.org/fr/.
  We configured a professional platform to manage the translation of the documentation (https://crowdin.com/project/poppy-docs. This platforms allows anybody to participate in the translation of the documentation to the language of their choice.
- To complete the pedagogical booklet [125] that provides guided activities and small challenges to become familiar with Poppy Ergo Jr robot and the Programming language Snap! (https://hal.inria.fr/hal-01384649/document) we provided a list of Education projects. Educational projects have been written for each activity carried out and tested in class. Each project has its own web page including resources allowing any teacher to carry out the activity (description, pedagogical sheet, photos / videos, pupil's sheet, teacher's sheet with correction etc.).
  The activities are available here:
  https://www.poppy-education.org/activites/activites-lycee
  The pedagogical activities are also available on the Poppy project forum where everyone is invited to comment and create new ones:
  https://forum.poppy-project.org/t/liste-dactivites-pedagogiques-avec-les-robots-poppy/2305

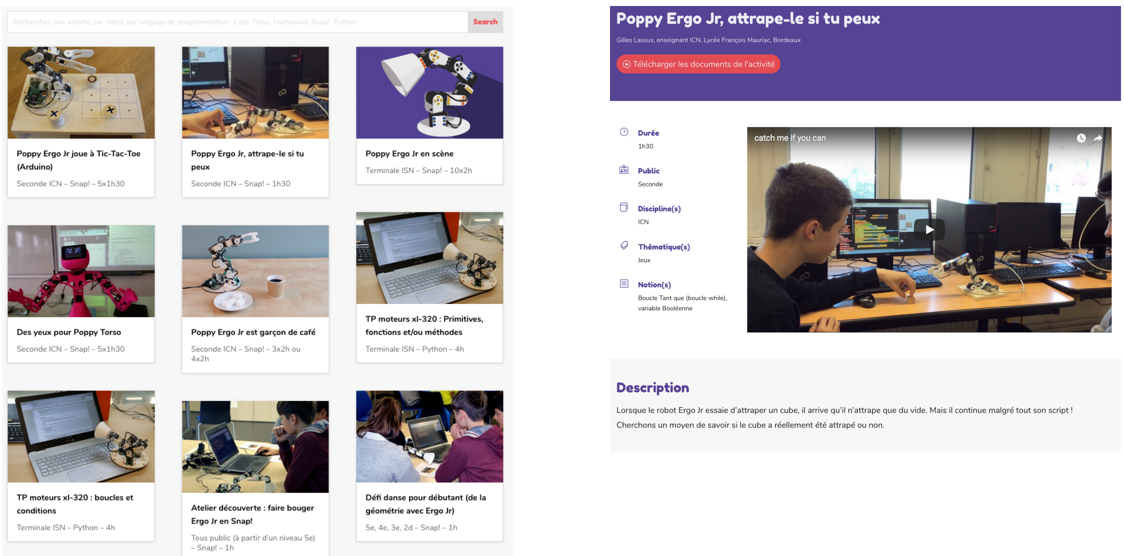


Figure 36. Open-source educational activities with Poppy robots are available on Poppy-Education.org

- A FAQ have been written with the most frequents questions to help the users: https://www.poppy-education.org/aide/
- A website has been created to present the project and to share all resources and activities. https://www.poppy-education.org/

7.6.3.3. Evaluation of the pedagogical kits

The impact of educational tools created in the lab and experimented in class had to be evaluated qualitatively and quantitatively. First, the usability, efficiency and user satisfaction must be evaluated. We must therefore assess, at first, if these tools offer good usability (i.e. effectiveness, efficiency, satisfaction). Then, in a second step, select items that can be influenced by the use of these tools. For example, students' representations of robotics, their motivation to perform this type of activity, or the evolution of their skills in these areas. In 2017 we conducted experiments to evaluate the usability of kits. We also collected data on students' perceptions of robotics.

- Population

  Our sample is made up of 28 teachers and 146 students from the region Nouvelle Aquitaine. Each subject completed an online survey in June 2017. Here, we study several groups of individuals: teachers and students. Among the students we are interested in those who practiced classroom activities with the Ergo Jr kit during the school year 2016 - 2017 (N = 68) (age = 16, std = 2.44). Among these students, 37 where High School students following the "Computer Science and Digital Sciences" stream (BAC S option ISN), 12 followed the stream "Computer and Digital Creation" (BAC S option ICN) and 18 where in Middle School.

  Among the 68 students, 13 declared having used the educational booklet provided in the kit and 16 declared having used other robotic kits. Concerning the time resource dedicated to activities with the robot, 30 students declared having spent less than 6 hours, 22 declared between 6 and 25 hours, and 16 declared having spent more than 25 hours.

  have practiced less than 6 hours of activity with the robot (N = 30), between 6 and 25 hours (N = 22) or more than 25 hours (N = 16); having built the robot (N = 12); have used the visual programming language Snap! (N = 46), the language of Python textual programming (N = 21), both (N = 8) or none (N = 9), it should be noted that these two languages are directly accessible via the main interface of the robot.
- Evaluation of the tool

  We have selected two standardized surveys dealing with this issue: SUS (The Systeme Usability Scales) [62] and The AttrakDiff [100]. These two surveys are complementary and allow to identify the design problems and to account for the perception of the user during the activities. The results of these surveys are available in the article (in French) [76] published at the conference Didapro (Lausanne Feb, 2018). Figures 37 and 38 show the averages of the 96 respondents (68 students + 28 teachers) for each of the 10 statements from the SUS and 28 pairs of antonyms to be scored on a scale of 1 to 5 and a 7-point scale, respectively.

- Evaluation of impact on learner

  One of the objectives of the integration of digital sciences in school is to allow students to have a better understanding of the technological tools that surround them daily (i.e. web, data, algorithm, connected object, etc.). So, we wanted to measure how the practice of activities with ErgoJr robot had changed this apprehension; especially towards robots. For that, we used a standardized survey: "attitude towards robot" *EuroBarometer 382* originally distributed in 2012 to more than 1000 people in each country of the European Union. On the one hand, we sought to establish whether there had been a change in response between 2012 and 2017, and secondly whether there was an impact on the responses of 2017 according to the participation, or not, in educational activities with ErgoJr robot. The analysis of the results is in progress and will be published in 2019.
- Web page for the experimentations

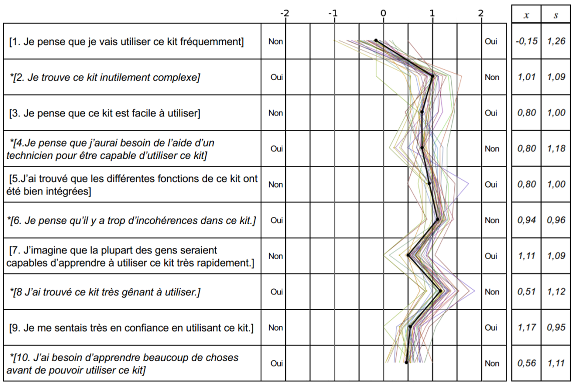


Figure 37. Result of SUS survey

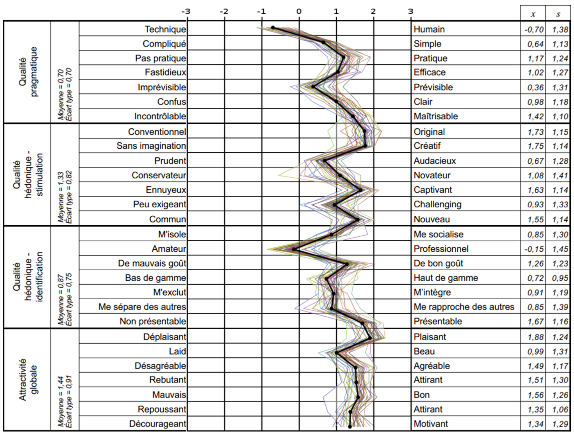


Figure 38. Result of AttrakDiff survey

To facilitate the storage of documents, their availability, and to highlight some information and news, a page dedicated to the experimentations is now available on the website. https://www.poppy-education.org/evaluation/

7.6.3.4. Partnership on education projects

- Ensam

  The Arts and Métiers campus at Bordeaux-Talence in partnership with Inria wishes to contribute to its educational and scientific expertise to the development of new teaching methods and tools. The objective is to develop teaching sequences based on a project approach, relying on an attractive multidisciplinary technological system: the humanoid Inria Poppy robot.

  The humanoid Inria Poppy robot offers an open platform capable of providing an unifying thread for the different subjects covered during the 3-years of the Bachelor training: mechanics, manufacturing (3D printing), electrical, mecha-tronics, computer sciences, design.

- Poppy entre dans la danse (Poppy enters the dance)

  The project "Poppy enters the dance" (Canope 33) took place for the second year. It uses the humanoid robot Poppy. This robot is able to move and experience the dance. The purpose of this project is to allow children to understand the interactions between science and choreography, to play with the random and programmable, to experience movement in dialogue with the machine. At the beginning of the project they attended two days of training on the humanoid robot (Inria - Poppy Education). During the project, they met the choreographer Eric Minh Cuong Castaing and the engineer Segonds Theo (Inria - Poppy Education).

  You can see a description and an overview of the project here:
  https://www.youtube.com/watch?v=XfxXaq899kY

- DANE

  The Academic Delegation for Digital Educational is in charge of supporting the development of digital uses for pedagogy. It implements the educational digital policy of the academy in partnership with local authorities. She accompanies institutions daily, encourages innovations and participates in their dissemination.

- RobotCup Junior

  RoboCupJunior OnStage invites teams to develop a creative stage performance using autonomous robots that they have designed, built and programmed. The objective is to create a robotic performance between 1 to 2 minutes that uses technology to engage an audience. The challenge is intended to be open-ended. This includes a whole range of possible performances, for example dance, storytelling, theatre or an art installation. The performance may involve music but this is optional. Teams are encouraged to be as creative, innovative and entertaining, in both the design of the robots and in the design of the overall performance.

7.7. Other applications

7.7.1. Applications in Robotic myoelectric prostheses

Participants: Pierre-Yves Oudeyer [correspondant], Aymar de Rugy, Daniel Cattaert, Mick Sebastien.

Together with the Hybrid team at INCIA, CNRS (Sébastien Mick, Daniel Cattaert, Florent Paclet, Aymar de Rugy) and Pollen Robotics (Matthieu Lapeyre, Pierre Rouanet), the Flowers team continued to work on a project related to the design and study of myoelectric robotic prosthesis. The ultimate goal of this project is to enable an amputee to produce natural movements with a robotic prosthetic arm (open-source, cheap, easily reconfigurable, and that can learn the particularities/preferences of each user). This will be achieved by 1) using the natural mapping between neural (muscle) activity and limb movements in healthy users, 2) developing a low-cost, modular robotic prosthetic arm and 3) enabling the user and the prosthesis to co-adapt to each other, using machine learning and error signals from the brain, with incremental learning algorithms inspired from the field of developmental and human-robot interaction.

7.7.1.1. Reachy, a 3D-printed Human-like Robotic Arm as a Test Bed for Prosthesis Control Strategies

To this day, despite the increasing motor capability of robotic prostheses, elaborating efficient control strategies is still a key challenge for their design. To provide an amputee with efficient ways to drive a prosthesis, this task requires thorough testing prior to integration into finished products. To preserve consistency with prosthetic applications, employing an actual robot for such testing requires it to show human-like features. To fulfill this need for a biomimetic test platform, we developed the Reachy robotic platform, a seven-joint human-like robotic arm that can emulate a prosthesis. Although it does not include an articulated hand and is therefore more suitable for studying reaching than manipulation, a robotic hand from available research prototypes could be integrated to Reachy. Its 3D-printed structure and off-the-shelf actuators make it inexpensive relatively to the price of a genuine prosthesis. Using an open-source architecture, its design makes it broadly connectable and customizable, so it can be integrated into many applications. To illustrate how Reachy can connect to external devices, we developed several proofs of concept where it is operated with various control strategies, such as tele- operation or vision-driven control. In this way, Reachy can help researchers to develop and test innovative control strategies on a human-like robot.

7.7.2. Ship Motion estimation from sea wave vision

Participants: David Filliat [correspondant], Natalia Díaz Rodríguez, Zhi Zhou, Manuel Cortés-Batet, Nazar-Mykola Kaminskyi.

Together with Naval Group, ENSTA Paris has been working on a set of software tools for simulating sea waves and motion estimation from images. The objective is predicting variables of interest in order to compensate the position and inclination of large boats at deep sea, seconds ahead of time to preserve stability. Work being currently done in partnership with Abo Akademi University (Turku, Finland) will validate the soon to be published Blender wave generator and machine learning algorithms, with real data gathered from the Baltic Sea archipelago.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

8.1.1. Autonomous Driving Commuter Car

Participants: David Filliat [correspondant], Emmanuel Battesti.

We developed planning algorithms for a autonomous electric car for Renault SAS in the continuation of the previous ADCC project. We improved our planning algorithm in order to go toward navigation on open roads, in particular with the ability to reach higher speed than previously possible, deal with more road intersection case (roundabouts), and with multiple lane roads (overtake, insertion...).

8.2. Bilateral Grants with Industry

8.2.1. Perception Techniques and Sensor Fusion for Level 4 Autonomous Vehicles

Participants: David Filliat [correspondant], Vyshakh Palli-Thazha.

Financing of the CIFRE PhD grant of Vyshakh Palli-Thazha by Renault.

8.2.2. Incremental Methods of Deep Learning for detection and classification in an robotics environment

Participants: David Filliat [correspondant], Timothée Lesort.

Financing of the CIFRE PhD grant of Timothée Lesort by Thales.

8.2.3. Exploration of reinforcement learning algorithms for drone visual perception and control

Participants: David Filliat [correspondant], Florence Carton.

Financing of the CIFRE PhD grant of Florence Carton by CEA.

8.2.4. Incremental learning for sensori-motor control

Participants: David Filliat [correspondant], Hugo Caselles Dupré.

Financing of the CIFRE PhD grant of Hugo Caselles-Dupré by Softbank Robotics.

8.2.5. Curiosity-driven Learning Algorithms for Exploration of Video Game Environments

Participant: Pierre-Yves Oudeyer [correspondant].

Financing of a postdoc grant for a 2 year project with Ubisoft and Région Aquitaine.

8.2.6. Intrinsically Motivated Exploration for Lifelong Deep Reinforcement Learning in the Malmo Environment

Participants: Pierre-Yves Oudeyer [correspondant], Remy Portelas.

Financing of the PhD grant of Rémy Portelas by Microsoft Research.

8.2.7. Explainable continual learning for autonomous driving

Participants: Natalia Díaz Rodríguez [correspondant], Adrien Bennetot.

Financing of the CIFRE PhD grant of Adrien Bennetot by Segula Technologies.

9. Partnerships and Cooperations

9.1. Regional Initiatives

9.1.1. Perseverons

Perseverons

Program: eFran

Duration: January 2016 - December 2019

Coordinator: PY Oudeyer, Inria Flowers

Partners: Inria Flowers

Funding: 140 keuros

The Perseverons project (Perseverance with / by digital objects), coordinated by the university via the ESPE (Higher School of Teaching and Education) of Aquitaine, and by the Rectorat of Bordeaux via the DANE (Academic Delegation digital education), aims to measure the real effectiveness of digital techniques in education to improve school motivation and perseverance, and, in the long term, reduce dropout. The project proposes to analyze the real effects of the use of two types of objects, robots, tablets, by comparing the school and non-school contexts of the *fablabs*. It is one of the 22 winners http://www.gouvernement.fr/efran-les-22-laureats of the "E-Fran" call for projects (training, research and digital animation spaces), following the Monteil mission on digital education, as part of the Investissement d'Avenir 2 program http://ecolenumerique.education.gouv.fr/2016/09/23/1244/. Formed of 12 sub-projects, "perseverons" has many partnerships, especially with the Poppy Education project of Inria Flowers. It is funding the PhD of Thibault Desprez.

Attachement	Type	Name	Adresse	Tel	Web
Poppy Éducation	High School	Alfred Kastler	14 Avenue de l'Université,33402 Talence, France	+33 5 57 35 40 70	http://www.lyceekastler.fr/
Poppy Éducation	Middle School	Anatole France	28 Rue des Micocouliers,33410 Cadillac, France	+33 5 56 62 98 42	http://www.afcadillac.net/
PERSEVERONS	High School	André Malraux	3 Rue du 8 Mai 1945,64200 Biarritz, France	+33 5 59 01 20 40	http://lycee-malraux-biarritz.fr/
Poppy Éducation	High School	Camille Jullian	29 Rue de la Croix Blanche,33000 Bordeaux, France	+33 5 56 01 47 47	http://www.camillejullian.com/
Poppy Éducation	Middle School	de France	Rue du Cimetière Saint-Benoist,75005 Paris, France	+33 1 44 27 12 11	http://www.college-de-france.fr/
Poppy Éducation	High School	des Graves	238 Cours du Général de Gaulle,33170 Gradignan, France	+33 5 56 75 77 56	http://www.grandlebrun.com/
PERSEVERONS	High School	Élie Faure	63 Avenue de la Libération,33310 Lormont, France	+33 5 56 38 23 23	http://www.lyc-eliefaure.fr/
PERSEVERONS	High School	Elisée Reclus	7 Avenue de Verdun,33220 Pineuilh, France	+33 5 57 41 92 50	http://lycee-foyen.fr/
Poppy Éducation	High School	François Mauriac	1 Rue Henri Dunant,33000 Bordeaux, France	+33 5 56 38 52 82	http://lyceemauriac.fr/
PERSEVERONS	High School	Gaston Febus	20 Avenue Georges Moutet,64300 Orthez, France	+33 5 59 67 07 26	http://webetab.ac-bordeaux.fr/cite-gaston-febus-orthez/
PERSEVERONS	Middle School	Giraud de Borneil	10 Boulevard André Dupuy,24160 Excideuil, France	+33 5 53 62 21 16	http://www.gdeborneil.fr/
PERSEVERONS	High School	Grand Air	Avenue du Docteur Lorentz Monod,33120 Arcachon, France	+33 5 56 22 38 00	http://webetab.ac-bordeaux.fr/lycee-grand-air/
PERSEVERONS	High School	Gustave Eiffel	143 Rue Ferbos,33000 Bordeaux, France	+33 5 56 33 83 00	http://www.eiffel-bordeaux.org/
PERSEVERONS	High School	Jacques Monod	10 Rue du Parvis,64230 Lescar, France	+33 5 59 77 92 00	http://lyceejacquesmonod.fr/
Poppy Éducation	High School	Jean Moulin	Avenue de la République,33210 Langon, France	+33 5 56 63 62 30	http://webetab.ac-bordeaux.fr/lycee-jean-moulin-langon/
Poppy Éducation	Middle School	Jean Zay	41 Rue Henri Cochet,33380 Biganos, France	+33 5 57 17 01 70	http://collegebiganos.fr/
Poppy Éducation	High School	La Morlette	62 Rue du Docteur Roux,33150 Cenon, France	+33 5 57 80 37 00	http://lycee-lamorlette.fr/
PERSEVERONS	High School	Les Iris	13 Rue Sourbès,33310 Lormont, France	+33 5 57 80 10 60	http://www.lyceelesiris.fr/
PERSEVERONS	High School	Louis Barthou	2 Boulevard Barbanègre,64000 Pau, France	+33 5 59 98 98 00	http://www.cyberlycee.fr/
PERSEVERONS	High School	Louis de Foix	4 Avenue Jean Rostand,64100 Bayonne/Bayona/Baiona, France	+33 5 59 63 31 10	http://www.louisdefoix.com/
PERSEVERONS	High School	Maine de Biran	108 Rue Valette,24100 Bergerac, France	+33 5 53 74 50 00	http://webetab.ac-bordeaux.fr/lycee-maine-de-biran/
Poppy Éducation	Middle School	Mios	Route du Pujeau,33380 Mios, France	+33 5 56 03 00 77	http://www.villemios.fr/enfance-jeunesse/college/
PERSEVERONS	High School	Nord Bassin	128 Avenue de Bordeaux,33510 Andernos-les-Bains, France	+33 5 56 82 20 77	http://www.lyceenordbassin.com/
Forum Poppy	Primary School	Notre-Dame du Mur	19 Rue de Kermadiou,29600 Morlaix, France	+33 2 98 88 18 69	http://lycee.ecmorlaix.fr/
PERSEVERONS	High School	Pape Clément	1 Rue Léo Lagrange,33600 Pessac, France	+33 5 57 26 63 00	http://lyceepapeclement.fr/
PERSEVERONS	High School	Pays de Soule	Avenue Jean Monnet,64130 Chéraute, France	+33 5 59 28 22 28	http://www.lyceedupaysdesoule.fr/index.php
PERSEVERONS	High School	Pré De Cordy	5 Avenue Joséphine Baker,24200 Sarlat-la-Canéda, France	+33 5 53 31 70 70	http://lycee-predecordy-sarlat.com/
Poppy Éducation	High School	Raoul Follereau	9 Boulevard Saint-Exupéry,58000 Nevers, France	+33 3 86 60 36 00	http://lyc58-renardfollereau.ac-dijon.fr/
PERSEVERONS	High School	René Cassin	2 Rue de Lasseguette,64100 Bayonne/Bayona/Baiona, France	+33 5 59 58 42 00	http://webetab.ac-bordeaux.fr/lycee-rene-cassin/
PERSEVERONS	High School	Saint-Cricq	4 Piste Cyclable,64000 Pau, France	+33 5 59 30 50 55	http://www.lycee-saint-cricq.org/
Poppy Éducation	High School	Saint-Genès	160 Rue de Saint-Genès,33000 Bordeaux, France	+33 5 56 33 84 84	http://www.saint-genes.com/
PERSEVERONS	High School	Saint-John Perse	2 Chemin de Barincou,64000 Pau, France	+33 5 59 62 73 11	http://www.lycee-saint-john-perse.fr/
Poppy Éducation	High School	Sainte-Marie Grand Lebrun	164 Rue François Mauriac,33200 Bordeaux, France	+33 5 56 08 32 13	http://www.grandlebrun.com/
inria	High School	Sainte-Saintonge	12 Rue de Saintonge,33000 Bordeaux, France	+33 5 56 99 39 29	http://www.lyceesaintefamille.com/
Poppy Éducation	High School	Sud-Médoc	Piste du Médoc Bleu,33320 Le Taillan-Médoc, France	+33 5 56 70 10 10	http://www.lyceesudmedoc.fr/
Poppy Éducation	High School	Victor Louis	2 Rue de Mégret,33400 Talence, France	+33 5 56 80 76 40	http://lyceevictorlouis.fr/

Figure 39. List of partner schools

9.1.1.1. Partner schools

In 2018, we have 36 partner schools (show Fig 39). 15 directly from the Poppy Education project. 19 new establishments were equipped in September 2017 by the Perseverons project. 21 of these establishments are located in Gironde. We have 27 high schools, 5 middle school.

9.2. National Initiatives

9.2.1. Myoelectric prosthesis - PEPS CNRS

PY Oudeyer collaborated with Aymar de Rugy, Daniel Cattaert, Mathilde Couraud, Sébastien Mick and Florent Paclet (INCIA, CNRS/Univ. Bordeaux) about the design of myoelectric robotic prostheses based on the Poppy platform, and on the design of algorithms for co-adaptation learning between the human user and the prosthesis. This was funded by a PEPS CNRS grant.

9.2.2. Poppy Station structure

- Since 1 september 2017 until february 2019, PerPoppy and Poppy Station Projects : D. Roy, P.-Y. Oudeyer. These projects aim to perpetuate the Poppy robot ecosystem by creating an external structure from outside Inria, with various partners. After the Poppy Robot Project, the Poppy Education Project has ended and Poppy Station structure is born. PerPoppy is the project which is building the new structure, and Poppy Station is the name of the new structure. Poppy Station, which includes Poppy robot ecosystem (hardware, software, community) from the beginning, is a place of excellence to build future educational robots and to design pedagogical activities to teach computer science, robotics and Artificial Intelligence. https://www.poppy-station.org
- Partners of Poppy Station : Inria, La Ligue de l'Enseignement, HESAM Université, SNCF Développement, IFÉ-ENS Lyon, MOBOTS – EPFL, Génération Robots, Pollen Robotics, KONEX-Inc, Mobsya, CERN Microclub, LINE Lab (Université Nice), Stripes, Canopé Martinique, Rights Tech Women, Editions Nathan.

9.2.3. Adaptiv'Math

Adaptiv'Math

Program: PIA

Duration: 2019 - 2020

Coordinator: EvidenceB

Partners:

- EvidenceB
- Nathan
- APMEP
- LIP6
- Inria
- ISOGRAD
- Daesign
- Schoolab
- BlueFrog

The solution Adaptiv'Math comes from an innovation partnership for the development of a pedagogical assistant based on artificial intelligence. This partnership is realized in the context of a call for projects from the Ministry of Education to develop a pedagogical plateform to propose and manage mathematical activities intended for teachers and students of cycle 2. The role of Flowers team is to work on the AI of the proposed solution to personalize the pedagogical content to each student. This contribution is based on the work done during the Kidlearn Project and the thesis of Benjamin Clement [69], in which algorithms have been developed to manage and personalize sequence of pedagogical activities. One of the main goal of the team here is to transfer technologies developed in the team in a project with the perspective of industrial scaling.

9.3. European Initiatives

9.3.1. Collaborations in European Programs, except FP7 & H2020

9.3.1.1. IGLU

Title: Interactive Grounded Language Understanding (IGLU)

Programm: CHIST-ERA

Duration: October 2015 - September 2018

Coordinator: University of Sherbrooke, Canada

Partners:

University of Sherbrooke, Canada

Inria Bordeaux, France

University of Mons, Belgium

KTH Royal Institute of Technology, Sweden

University of Zaragoza, Spain

University of Lille 1 , France

University of Montreal, Canada

Inria contact: Pierre-Yves Oudeyer

Language is an ability that develops in young children through joint interaction with their caretakers and their physical environment. At this level, human language understanding could be referred as interpreting and expressing semantic concepts (e.g. objects, actions and relations) through what can be perceived (or inferred) from current context in the environment. Previous work in the field of artificial intelligence has failed to address the acquisition of such perceptually-grounded knowledge in virtual agents (avatars), mainly because of the lack of physical embodiment (ability to interact physically) and dialogue, communication skills (ability to interact verbally). We believe that robotic agents are more appropriate for this task, and that interaction is a so important aspect of human language learning and understanding that pragmatic knowledge (identifying or conveying intention) must be present to complement semantic knowledge. Through a developmental approach where knowledge grows in complexity while driven by multimodal experience and language interaction with a human, we propose an agent that will incorporate models of dialogues, human emotions and intentions as part of its decision-making process. This will lead anticipation and reaction not only based on its internal state (own goal and intention, perception of the environment), but also on the perceived state and intention of the human interactant. This will be possible through the development of advanced machine learning methods (combining developmental, deep and reinforcement learning) to handle large-scale multimodal inputs, besides leveraging state-of-the-art technological components involved in a language-based dialog system available within the consortium. Evaluations of learned skills and knowledge will be performed using an integrated architecture in a culinary use-case, and novel databases enabling research in grounded human language understanding will be released. IGLU will gather an interdisciplinary consortium composed of committed and experienced researchers in machine learning, neurosciences and cognitive sciences, developmental robotics, speech and language technologies, and multimodal/multimedia signal processing. We expect to have key impacts in the development of more interactive and adaptable systems sharing our environment in everyday life. http://iglu-chistera.github.io/

9.4. International Initiatives

9.4.1. Inria Associate Teams Not Involved in an Inria International Labs

9.4.1.1. NEUROCURIOSITY

Title: NeuroCuriosity

International Partner (Institution - Laboratory - Researcher):

Columbia Neuroscience (United States) - Cognitive Neuroscience - JACQUELINE GOTTLIEB

Start year: 2016

See also: https://flowers.inria.fr/neurocuriosityproject/

Curiosity can be understood as a family of mechanisms that evolved to allow agents to maximize their knowledge of the useful properties of the world. In this project we will study how different internal drives of an animal, e.g. for novelty, for action, for liking, are combined to generate the rich variety of behaviors found in nature. We will approach such challenge by studying monkeys, children and by developing new computational tools.

9.4.1.2. Idex Bordeaux-Univ. Waterloo collaborative project on curiosity in HCI

Title: Curiosity

International Partner (Institution - Laboratory - Researcher):

University of Waterloo (Canada), Edith Law's HCI Lab and Dana Kulic's Robotics lab.

Start year: 2018

Pierre-Yves Oudeyer collaborated with Edith Law's HCI research group at University of Waterloo on the topic of "Curiosity in HCI system". They obtained a grant from Univ. Bordeaux to set up a project with Inria Potioc team and with Dana Kulic, Robotics lab, Univ. Waterloo. They organized several cross visits and collaborated on the design and experimentation of an educational interactive robotic system to foster curiosity-driven learning. This led to two articles accepted at CHI 2019 and CHI2020 (see new results section).

To continue this collaborative research, a new proposal on « Curiosity-driven learning and personalized (re-)education technologies across the lifespan » have been successfully submitted to UB-UW IDEX call regarding the projects in the field of AI and health sciences (PI: E. Law, PY Oudeyer ; co-PI : M. Fernandes, H. Sauzéon & F. Lotte)

9.4.1.3. Idex Bordeaux-Univ. Waterloo collaborative project on Virtual realty-based study on spatial learning in aging

Title: Spatial learning with aging

International Partner (Institution - Laboratory - Researcher):

University of Waterloo (Canada), Myra Fernandes, Cognitive neurosciences Lab.

Start year: 2016 (end year 2019)

Helene Sauzéon collaborated with Myra Fernandes's cognitive neuroscience Lab at University of Waterloo on the topic of "VR based study of spatial learning in older adults". They obtained a grant from Univ. Bordeaux to set up a project with Quincy Almeida, head of Movement Disorders Research and Rehabilitation Centre, Laurier University. They organized several cross visits and collaborated on the design and experimentation of a virtual reality application allowing to investigate intrinsic motivation (i.e., Active exploration) as cognitive support for older adults' spatial learning. This led to an article published in Brain Science in 2019 (see new results section).

9.4.1.4. Informal International Partners

Pierre-Yves Oudeyer and Didier Roy have created a collaboration with LSRO EPFL and Pr Francesco Mondada, about Robotics and education. The two teams co-organize the annual conference "Robotics and Education"in Bordeaux. Didier Roy teaches "Robotics and Education" in EPFL several times a year.

Didier Roy has created a collaboration with HEP Vaud (Teachers High School) and Bernard Baumberger and Morgane Chevalier, about Robotics and education. Scientific discussions and shared professional training.

Didier Roy has created a collaboration with Biorob - EPFL, LEARN - EPFL, and Canton de Vaud, about Robotics and Computer Science education. Scientific discussions and shared professional training.

Didier Roy has created a collaboration with Mauritius Research Council, Mauritius Education Institute and AUF, about Robotics, AI and Computer Science projects, teaching and learning. Scientific discussions and shared professional training. With Gérard Giraudon (Advisor to the President of Inria, with in particular a mission on "Digital & Training").

A collaboration with Johan Lilius and Sebastien Lafond from Abo Akademi University, Turku (Finland) is ongoing to sign an Erasmus contract for researchers and students visits on the topic of autonomous boats.

Funding applications have been submitted jointly with Davide Maltoni and Vincenzo Lomonaco from University of Bologna (Italy) on the topic of continual learning. Also the project https://www.continualai.org/ is being further developed jointly and on the way to become a non-profit organization.

9.4.2. Participation in Other International Programs

David Filliat participates in the ITEA3 DANGUN project with Renault S.A.S. in france and partners in Korea. The purpose of the DANGUN project is to develop a Traffic Jam Pilot function with autonomous capabilities using low-cost automotive components operating in France and Korea. By incorporating low-cost advanced sensors and simplifying the vehicle designs as well as testing in different scenarios (France & Korea), a solution that is the result of technical cooperation between both countries should lead to more affordable propositions to respond to client needs in the fast moving market of intelligent mobility.

Natalia Díaz Rodríguez collaborates with the Abo Akademi University in Turku, Finland on the autonomous navigation systems project, involving the sailing schools of Novia and Naval Group (France). She also collaborates with the Andalusian Research Institute in Data Science and Computational Intelligence https://dasci.es (DaSCI) and the University of Granada (Spain) on explainable AI.

9.5. International Research Visitors

9.5.1. Visits of International Scientists

- Kevvyn Collins-Thompson, Univ. Michigan (sept.-dec. 2019)
- Franck Guerin, Univ. Aberyswith (dec 2019)
- Justus Piater, Univ. Innsbrucj (dec 2019)
- Verena Hafner, Univ. Berlin (dec 2019)
- Jochen Triesch, Univ. Frankfurt (dec 2019)
- Nivedita Mani, Univ. Gottingen (dec 2019)
- Oksana Hagen, Plymouth University (Oct. 2019)

9.5.2. Internships

- Medhi Alaimi [Inria, until Jul 2019]
- Timothee Anne [Inria, from Feb 2019 until Jun 2019]
- Anouche Banikyan [Inria, from Feb 2019 until Jul 2019]
- Lucie Galland [Ecole Normale Supérieure Paris, from Jun 2019 until Aug 2019]
- Tallulah Gilliard [Inria, from Feb 2019 until Jul 2019]
- Marion Schaeffer [Inria, from Jul 2019 until Sep 2019]
- Martin Serret [Inria, from Feb 2019 until Aug 2019]
- Maria Teodorescu [Inria, from Sep 2019]

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events: Organisation

10.1.1.1. Member of the Organizing Committees

PY Oudeyer was co-organizer of the Workshop on Exploration in Reinforcement Learning (ICML 19), https://sites.google.com/view/erl-2019/organizers.

10.1.2. Scientific Events: Selection

10.1.2.1. Member of the Conference Program Committees

PY Oudeyer was member of the program committee of IEEE International Conference on Developmental Learning and Epigenetic Robotics (ICDL-Epirob).

Natalia Díaz Rodríguez, steering committee and reviewer of CONTLEARN Workshop at CVPR2020 Workshop on Continual Learning in Computer Vision, Seattle, US.

10.1.2.2. Reviewer

PY Oudeyer was a reviewer for: Workshop on Exploration in Reinforcement Learning (ICML 19).

Hélène Sauzéon reviewed a conference paper for CHI 2020.

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

PY Oudeyer was editor of the IEEE Newsletter on Cognitive and Developmental Systems.

PY Oudeyer was associate editor of: IEEE Transactions on Cognitive and Developmental Systems and Frontiers in Neurorobotics.

10.1.3.2. Reviewer - Reviewing Activities

Clément Moulin-Frier reviewed an article for *Journal of Artificial Intelligence Research (JAIR)*.

PY Oudeyer reviewed for the journals: Journal of Artificial Intelligence Research, Cognitive Science, Nature Science Reports.

Hélène Sauzéon reviewed 6 articles for Memory & Cognition (M&C), Annals of Physical and Rehabilitative Medecine (APRM), Psychonomic Bulletin, Motivation & Cognition (MC); Cognition & Emotion (CE)

Natalia Díaz Rodríguez reviewed at Frontiers in Robotics and AI, Transactions on Emerging Telecommunications Technologies, IEEE Robotics & Automation Magazine, Neurocomputing, Robotics and Autonomous Systems.

Cédric Colas reviewed an article for the Robotics Science and Systems (RSS) conference.

10.1.4. Invited Talks

Clément Moulin-Frier gave invited talks at the *Journées Nationales de Recherche en Robotics (JNRR)* in Vittel (France) and in the *Journée GT8 "Robotique et Neurosciences"* du GDR Robotique in Bordeaux (France).

Clément Moulin-Frier and Natalia Díaz-Rodríguez gave a talk and participated, resp. in the workshop *Open Artificial Intelligence: From big data to smart data* at the *Centre de Recherche Interdisciplinaire (CRI)* in Paris (France) 13-15 Nov. 2019.

PY Oudeyer gave a keynote talk at International Conference on Learning Representations (ICLR 2019), New-Orleans, May 2019, https://iclr.cc/Conferences/2019/Schedule?showEvent=1141.

PY Oudeyer gave a keynote talk at ACM International Conference on Virtual Agents (ACM IVA 2019), Paris, on developmental machine learning https://iva2019.sciencesconf.org/.

PY Oudeyer gave an invited talk at the Deep Learning Summit (ReWork 2019), London, on curiosity-driven artificial intelligence, https://www.re-work.co/events/deep-learning-summit-london-2019/speakers/.

PY Oudeyer gave an invited talk at the conference on Reinforcement Learning and Decision Making (RLDM 2019), Montreal, on curiosity-driven developmental learning, http://rldm.org/invited-speakers/.

PY Oudeyer gave an invited talk at the Workshop on Curiosity for Decision Making at RLDM 2019, Montreal, https://sites.google.com/view/rldm-curiosity.

PY Oudeyer gave an invited talk at the HUMAINT conference in Séville, on "Developmental Autonomous Learning: AI, Cognitive Sciences and EdTech", 2019.

PY Oudeyer gave an invited talk at the workshop "AI and cognitive systems in Aquitaine", ENSC/Chaire STAH, Bordeaux, 2019.

PY Oudeyer gave an invited talk at the workshop "Task-agnostic reinforcement learning" at ICLR 2019, https://tarl2019.github.io, on developmental autonomous learning.

PY Oudeyer gave an invited talk at the workshop on Curiosity, Explanation and Exploration at Princeton University, 2019, on models of curiosity-driven learning in humans and machines.

PY Oudeyer gave an invited talk at the workshop on AI and Machine Learning, Telecom ParisTech, 2019, on developmental machine learning, https://workshopmlai.wp.imt.fr.

PY Oudeyer gave an invited talk at Académie de Médecine, Paris, 2019, on models of curiosity-driven learning and educational technologies.

PY Oudeyer gave an invited talk at workshop on Self and Sense of Agency, University Paris Descartes, nov. 2019, on models of curiosity-driven learning in humans and machines.

Hélène Sauzéon gave three invited talks at : 1) the conference on « Neuropsychologie clinique et technologies » - SNLF 2019, Paris, dec., 2-5 ; 2) Annual cycle of conference of Hôpital Salpêtrière, dec., 3; 3) Computer science Dpt. Of the University of Beijing, Beijing, July, 22-23.

Chris Reinke gave three invited talks at: 1) Okinawa Institute of Science and Technology, May 2019, on machine learning methods for the automated discovery of behaviors in physical and chemical systems; 2) LOMA Theory Day at Bordeaux, July 2019, on machine learning in physics; 3) Inria Grenoble, December 2019, on using cognitive models for artificial intelligence

N. Díaz Rodríguez gave an invited keynote (Continual Learning and Robotics: An Overview) and was a panelist at the ICLR workshop on multi-task and lifelong reinforcement learning https://sites.google.com/view/mtlrl

N. Díaz Rodríguez gave a guest talk at the Network of Young Researchers - Neuro Day NYR, Continual learning and Robotics, Paris 7th June 2019[2].

N. Díaz Rodríguez moderated a round table tomorrow in Paris: *WHEN AI AND BIG DATA MEET LIFE SCIENCES: ADVANCES IN RESEARCH AND ETHICAL QUESTIONS* https://yrls.fr/roundtable/ at Institut Imagine, Paris.

10.1.5. Leadership within the Scientific Community

Hélène Sauzéon was member of the ANR committe of « Technologie & Health »

10.1.6. Scientific Expertise

PY Oudeyer was a reviewer for the European Commission (FET program).

Hélène Sauzéon was a reviewer for the ANR call on International and european and scientific networks.

10.1.7. Research Administration

PY Oudeyer has been head of the Flowers team and member of piloting committees of consortium projects Adaptiv'Maths and Perseverons (eFran) on educational technologies.

Hélène Sauzéon was head of HACS team (BPH Lab, Inserm-UB), and thus member of directory committe of « Handicap » IFR (Inserm).

Helène Sauzéon is member of directory committe of the « centre d'excellence BIND », and she managed the Industrial Innovation and transfer sub-committee.

[2] https://ifm-institute.fr/fr/accueil/actualites/470-network-of-young-researchers-neuro-day-7th-june-20?fbclid=IwAR03u8T-opoede5hnCQx_QSphsJPC77YCrmnuNZH8c5XiBoRH9kgotLvYyw

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

PY Oudeyer gave a course on developmental reinforcement learning at ENSEIRB master on AI and machine learning (3h), nov. 2019.

PY Oudeyer gave a course on developmental learning at CogMaster cognitive science master (3h), nov. 2019.

PY Oudeyer gave a course on developmental learning at ENSC/ENSEIRB "option robot" master (3h), dec. 2019.

During the latest academic year, Hélène Sauzéon teached 96h in the BS. and master degrees in cognitive science (Department of Mathematics & interaction, University of Bordeaux). She was (co-)responsible of 9 teaching units (3 in BS et 6 in Master).

N Díaz Rodríguez taught, at ENSTA, a total of 3.25 h in ROB313, 27h at IN104, 10.5 at IN102, 21h at IA301. She also gave 42h at IG.2410 at the engineering school ISEP, and 3h course on Continual Learning and State Representation Learning at the reinforcement learning course at ENSEIRB master on AI and machine learning (3h), nov. 2019.

Didier Roy gave courses on computer science basics, and on computer science, robotics and AI activities for education at Canton de Vaud teachers.

10.2.2. Supervision

- PhD defended: Sébastien Forestier, "Intrinsically Motivated Goal Exploration in Child Development and Artificial Intelligence: Learning and Development of Speech and Tool Use", University of Bordeaux (supervised by PY Oudeyer).
- PhD defended: Thibault Desprez, "Conception et évaluation de kits robotiques pédagogiques", University of Bordeaux (supervisors: PY. Oudeyer and D. Roy)
- PhD in progress : Rémy Portelas, "Teacher algorithms for curriculum learning in Deep RL", beg. in sept. 2018 (supervisors: PY Oudeyer and K Hoffmann)
- PhD in progress: Cédric Colas, "Intrinsically Motivated Deep RL", beg. in sept. 2017 (supervisors: PY Oudeyer and O Sigaud)
- PhD in progress: Tristan Karch, "Language acquisition in curiosity-driven Deep RL", beg. in sept. 2019 (supervisors: PY Oudeyer and C Moulin-Frier)
- PhD in progress: Alexandr Ten, "Models of human curiosity-driven learning and exploration", beg. in sept. 2018 (supervisor: PY Oudeyer)
- PhD defended: Cécile Mazon, "Des Technologies Numériques Pour L'inclusion Scolaire Des Collégiens Avec TSA : des approches individuelles aux approches écosystémiques pour soutenir l'individu et ses aidants ", University of Bordeaux (supervised by H. Sauzéon).
- PhD defended: Pierre-Antoine Cinquin, "Conception, intégration et validation de systèmes numériques d'enseignement accessibles aux personnes en situation de handicap cognitif ", University of Bordeaux (supervised by H. Sauzéon & P. Guitton).
- PhD in progress: Adrien Bennetot, "Explainable continual learning for autonomous driving", Sorbonne University and ENSTA Paris (supervised by N Díaz Rodríguez & R Chatila).
- Master thesis defended: Anouche Banikyian "Curiosity, intrinsic motivation and spatial learning in children", University of Bordeaux (supervised by H. Sauzéon).
- Master thesis defended: Mehdi Alaimi "New educational application for fostering curiosity-related question-asking in children", University of Bordeaux (supervised by H. Sauzéon & PY Oudeyer).
- Master thesis defended: Juewan Wang "Can an accessible MOOC player improve the retention of disabled students? A MOOC accessibility assessment based on analytic method ", University of Bordeaux (supervised by H. Sauzéon & P. Guitton).

10.2.3. Juries

PY Oudeyer was reviewer of the PhD of Simon Hangl, entitled « Autonomous Robotics: an Integrated Approach from Controllers to Cognitive Capabilities », University of Innsbruck, Austria, 2019.

PY Oudeyer was reviewer of the PhD of Yannick Bourrier, titled "Diagnostic et prise de décision pédagogique pour la construction de compétences non- techniques en situation critique", from university Paris-Sorbonne, 2019.

PY Oudeyer was reviewer of the PhD of Leni Kenneth Le Goff, titled "Bootstrapping robotic ecological perceptin with exploration and interactions", from University Pierre et Marie Curie, 2019.

PY Oudeyer was examiner of the PhD of Jelena Mladenovic, titled "Modélisation computationnelle des compétences et des états de l'utilisateur pour optimiser les taches d'entraînement aux Interfaces Cerveaux-Ordinateur", University of Bordeaux, 2019.

PY Oudeyer was examiner of the PhD of Pierre Fournier, titled "Intrinsically Motivated and Interactive Reinforcement Learning: a Developmental Approach", University Pierre and Marie Curie, 2019.

PY Oudeyer was examiner of the PhD of Lisa Jacquey, titled "La sensibilité aux contingences sensorimotrices chez le bébé et son rôle dans le développement du savoir-faire corporel", from University of Paris, 2019.

PY Oudeyer was examiner in the PhD jury of Svetlana Meyer, on the topic "Entraînement de l'attention visuelle pour l'apprentissage de la lecture : l'apport du jeu vidéo d'action", University of Grenoble (feb. 2019).

Hélène Sauzéon organized a selection commitee for recruitment of Assistant professor in Rehabilitative science (University of de Bordeaux).

Hélène Sauzéon was external member of a selection commitee for recruitment of Assistant professor in cognitive psychology (University of Toulouse – LeMirail).

Hélène Sauzéon performed several scientific expertises for application requests such as HDR (ED SP2, University of Bordeaux) or local careers advancement (University of Bordeaux).

N. Díaz Rodríguez was invited jury (President) of the PhD thesis "Deep Learning for Abnormal Movement Detection using Wearable Sensors: Case Studies on Stereotypical Motor Movements in Autism and Freezing of Gait in Parkinson's Disease" in the University of Trento, Italy May 2019.

10.3. Popularization

10.3.1. Internal or external Inria responsibilities

Pierre-Yves Oudeyer has been member of the scientific committee for the permanent exhibition "Robots" at Cité des Sciences et de l'Industrie, Paris: http://www.cite-sciences.fr/fr/au-programme/expos-permanentes/expos-permanentes-dexplora/robots/lexposition/. See also https://www.inria.fr/fr/la-robotique-a-la-cite-des-sciences.

Pierre-Yves Oudeyer has been member of the scientific committee for the exhibition "Robots" at Cap Sciences, http://www.cap-sciences.net/au-programme/exposition/robots.

Pierre-Yves Oudeyer was member of the jury of the popular science competition Creathon in Poitiers.

Didier Roy is member of the team "Livre blanc Inria EdTech", with Gérard Giraudon, Pascal Guitton, Thierry Viéville and Margarida Romero.

Didier Roy is member of the Class'code and Class'code AI teams.

Didier Roy has been member of the organization committee for the public workshop/exhibition "Robots" at Cité des Sciences et de l'Industrie, Paris.

Didier Roy is vice-president of Poppy Station Structure. Poppy Station, which includes Poppy robot ecosystem (hardware, software, community) from the beginning, is a place of excellence to build future educational robots and to design pedagogical activities to teach computer science, robotics and Artificial Intelligence. https://www.poppy-station.org. Partners of Poppy Station : Inria, La Ligue de l'Enseignement, HESAM Université, SNCF Développement, IFÉ-ENS Lyon, MOBOTS – EPFL, Génération Robots, Pollen Robotics, KONEXInc, Mobsya, CERN Microclub, LINE Lab (Université Nice), Stripes, Canopé Martinique, Rights Tech Women, Editions Nathan.

10.3.2. Articles and contents

Sébastien Forestier and Pierre-Yves Oudeyer were interviewed for a video documentary on curiosity-driven learning in robots, now permanently displayed in the exhibition "Robots" at Cité des Sciences et de l'Industrie, Paris: http://www.cite-sciences.fr/fr/au-programme/expos-permanentes/expos-permanentes-dexplora/robots/lexposition/. The video is available here: https://www.youtube.com/watch?v=Kw724djJpUs.

PY Oudeyer was interviewed in RFI radio program "Autour de la question" on artificial intelligence and society (feb. 2019), http://www.rfi.fr/emission/20190211-comment-vivre-bonne-intelligence-toutes-intelligences.

Didier Roy and PY Oudeyer wrote an illustrated book on artificial intelligence and robotics for 7-8 years old children, to be published by Nathan.

10.3.3. Education

Hélène Sauzéon, as co-designer (with P. Guitton) of "Digital Accessibility" MOOC displayed on FUN platform, has interacted with the learners during the permanent session in 2019.

N Díaz Rodríguez participated on the Future Fest organized by the director and deans of School of AI with the objective of creating and self fund the first free degree on AI (April 2019, Granada, Spain. theSchool.ai).

Didier Roy is organizing an International R2T2 Robotics Mission Event for the EduCamp (RoboCup Junior 2020 at Bordeaux).

10.3.4. Interventions

Pierre-Yves Oudeyer gave a popular science conference at Lycée Montaigne, Bordeaux, on machine learning and AI (17th jan. 2019).

PY Oudeyer participated to a popular science debate on science-fiction movies and artificial intelligence at cinema Utopia, Bordeaux (jan. 2019).

PY Oudeyer gave an invited popular science talk at Collège des Bernardins, Paris, 2019, on affective machine learning and educational technologies, https://www.collegedesbernardins.fr/content/affects-numeriques-et-education and video at https://www.youtube.com/watch?v=WYd6RzaCQDc.

PY Oudeyer gave a popular science conference to 2nde students from Lycée La Sauque, Bordeaux.

PY Oudeyer gave a popular science presentation on artificial intelligence in the event "AI and education" organized at Inria Bordeaux by Didier Roy and Nicolas Rougier.

Hélène Sauzéon participated to several talks targetted disability-related professionals, students or industries : 1) Technologies educatives & Handicap Conference, organised by ESPE of Academy of Nantes, Mar., 20 ; 2) Journée d'étude de l'Observatoire des ressources numériques adaptées (Orna) « Autisme et outils numériques : de la recherche aux applications », May, 3at INSHE , Paris. ; 3) Colloque « Augmentation de l'humain » STAH Industrial Chair of Nouvelle Aquitaine, Mar., 28

Didier Roy gave an invited talk at CERN, Geneve, Switzerland, november 2019, on Educational Power of International R2T2 Robotics Missions

Didier Roy gave an invited talk at LEARN DAY, Bern, Switzerland, november 2019, on Strategies for introducing Robotics and Computer science activities at school.

Didier Roy organized the APEIA DAY (Artificial Intelligence & Education), at Inria BSO, Inria Sophia and open video-conference. Talks and workshop about unplugged and plugged activities for teaching AI from primary school to high education. With teachers, animators, education experts and AI researchers.

10.3.5. Internal action

Hélène Sauzéon has presented the « Collège + » and Kidlearn projects to Sophie Cluzel, Secretary of State for Disability, during the launch day of the Disability Plan of Inria (November, 16).

10.3.6. Creation of media or tools for science outreach

10.3.6.1. AIANA: an accessible multimedia player

Hélène Sauzéon is co-designer with P. Guitton of the accessible AIANA multimedia player for MOOCs. The V1 was developed by Learning Lab Inria in 2016 and the V2 by Damien Caseli (ADT engineer) as part of the POTIOC-Inria team in 2019. This player offers a flexible interface, configurable in terms of sensory and motor alternatives for the communication of people with physical disabilities, and of cognitive alternatives (attention load, reading aid, note-taking, etc.) for people with cognitive disabilities. The most convincing result is that it has increased the number of disabled learners completing a MOOC course (16%) on FUN platform compared to the available benchmark (10

10.3.6.2. Poppy Station: Robotics and AI for Education, Arts and Research with the Poppy platform

Didier Roy continued to promote and co-manage the Poppy Station Association.

Among the recent actions of poppy station, there were the training provided at the 2nd chance school created in Saintes by SNCF Development, the training given at the Institut Pasteur in Paris, on IT, robotics and artificial intelligence. Lasting collaborations have taken hold.

There is also a new artistic collaboration with the artist Clara Maïda and her project "(a)utom@ton" of contemporary musical and robotic creation (work for electroacoustic device and two humanoid robots).

Poppy Station has also participated in various events concerning educational robotics and artificial intelligence : APEIA Day at Inria Bordeaux, National Educational Robotics Colloquium at French Institute of education (ENS Lyon), Educatec Educatice exhibition in Paris, Ludovia Summer university in Ax-les-thermes.

A new open source wheeled robot printed in 3d was created for poppy Station by Pollen robotics, partner of poppy Station.

Poppy Station is also involved in European calls for projects with the Institut Pasteur around the continuous training of scientific staff on IT, robotics and AI.

Poppy Station Association brings together players from the world of business, research, training, culture and education and aims to develop and preserve robotic ecosystems and associated open source or free technologies, in all areas where their use can allow this development and preservation. The association pays particular attention to the fields of education, training, arts and research. Poppy Station is the result of a transfer of Inria research from its open-source robotics ecosystem Poppy to an external multi-partners structure. The Poppy ecosystem includes software and hardware tools to create and program robots, as well as educational content for education and training, and a large interdisciplinary community of users. This ecosystem was created and developed by the Inria Flowers team, with the aim of facilitating the experimentation and creation of innovative robotic tools in the fields of education, research and the arts. http://www.poppy-station.org.

10.3.6.3. IniRobot: Educational Robotics in Primary Schools

Didier Roy and PY Oudeyer continued to promote and disseminate the IniRobot pedagogical toolkit. Irirobot, a project done in collaboration with EPFL/Mobsya, aims to create, evaluate and disseminate a pedagogical kit which uses Thymio robot, an open-source and low cost robot, for teaching computer science and robotics.

IniRobot Project aims to produce and diffuse a pedagogical kit for teachers and animators, to help them and to train them directly or by the way of external structures. The aim of the kit is to initiate children to computer science and robotics. The kit provides a micro-world for learning, and takes an inquiry-based educational approach, where kids are led to construct their understanding through practicing an active investigation methodology within teams. See https://dm1r.inria.fr/c/kits-pedagogiques/inirobot or http://www.inirobot.fr.

Deployment: After 4 years of activity, IniRobot is used by more than 3000 adults, 30 000 children in France. Inirobot is also used in higher education, for example in Master 2 "Neurosciences, human and animal cognition" at the Paul Sabatier University in Toulouse. Inirobot is additionally used to train the management and elected officials of the Bordeaux metropolitan area (20 people). The digital mediators of the 8 Inria centers are trained to Inirobot and use it in their activities.

10.3.6.3.1. Partnership

The project continues to be carried out in main collaboration with the LSRO Laboratory from EPFL (Lausanne) and others collaborations such as the French National Education/Rectorat d'Aquitaine, the Canopé Educational Network, the ESPE (teacher's school) Aquitaine, the ESPE Martinique, the ESPE Poitiers and the National Directorate of Digital Education.

10.3.6.3.2. Created pedagogical documents and resources

- The inirobot pedagogical kit [87]: This pedagogical booklet provides activities scenarized as missions to do. An updated version of the Inirobot pedagogical kit is available at: https://dm1r.inria.fr/uploads/default/original/1X/70037bdd5c290e48c7ec4cb4f26f0e426a4b4cf6.pdf. Another pedagogical booklet has been also created by three pedagogical advisers for primary school, with pedagogical instructions and aims, under our supervision. The new pedagogical kit,"Inirobot Scolaire, Langages et robotique", which extends Inirobot to a full primary school approach is available at https://blogacabdx.ac-bordeaux.fr/numerique33/2018/10/04/robotique-sequence-inirobot-scolaire/
- Inirobot website and forum: https://dm1r.inria.fr/c/kits-pedagogiques/inirobot or http://www.inirobot.fr On this website, teachers, animators and general public can download documents, exchange about their use of inirobot's kit.

10.3.6.3.3. Scientific mediation

Inirobot is very popular and often presented in events (conferences, workshops, ...) by us and others.

10.3.6.3.4. Spread of Inirobot activities

Inirobot activities are used by several projects: Dossier 123 codez from Main à la Pâte Fundation, Classcode project, ...

10.3.6.3.5. MOOC Thymio

The MOOC Thymio, released in october 2018, in collaboration with Inria Learning Lab and EPFL (Lausanne, Switzerland), on FUN platform and edX EPFL Platform), use Inirobot activities to teach how to use Thymio robot in education.

11. Bibliography

Major publications by the team in recent years

[1] A. BARANES, P.-Y. OUDEYER. *Active Learning of Inverse Models with Intrinsically Motivated Goal Exploration in Robots*, in "Robotics and Autonomous Systems", January 2013, vol. 61, n^{o} 1, pp. 69-73 [*DOI :* 10.1016/J.ROBOT.2012.05.008], https://hal.inria.fr/hal-00788440

[2] H. CASELLES-DUPRÉ, M. GARCIA-ORTIZ, D. FILLIAT. *Symmetry-Based Disentangled Representation Learning requires Interaction with Environments*, in "NeurIPS 2019", Vancouver, Canada, December 2019, https://hal.archives-ouvertes.fr/hal-02379399

[3] C. COLAS, P. FOURNIER, O. SIGAUD, M. CHETOUANI, P.-Y. OUDEYER. *CURIOUS: Intrinsically Motivated Modular Multi-Goal Reinforcement Learning*, in "International Conference on Machine Learning", Long Beach, France, June 2019, https://hal.archives-ouvertes.fr/hal-01934921

[4] C. COLAS, O. SIGAUD, P.-Y. OUDEYER. *GEP-PG: Decoupling Exploration and Exploitation in Deep Reinforcement Learning Algorithms*, in "International Conference on Machine Learning (ICML)", Stockholm, Sweden, July 2018, https://hal.inria.fr/hal-01890151

[5] C. CRAYE, T. LESORT, D. FILLIAT, J.-F. GOUDOU. *Exploring to learn visual saliency: The RL-IAC approach*, in "Robotics and Autonomous Systems", February 2019, vol. 112, pp. 244-259, https://hal.archives-ouvertes.fr/hal-01959882

[6] S. FORESTIER, Y. MOLLARD, P.-Y. OUDEYER. *Intrinsically Motivated Goal Exploration Processes with Automatic Curriculum Learning*, November 2017, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01651233

[7] S. FORESTIER, P.-Y. OUDEYER. *A Unified Model of Speech and Tool Use Early Development*, in "39th Annual Conference of the Cognitive Science Society (CogSci 2017)", London, United Kingdom, Proceedings of the 39th Annual Conference of the Cognitive Science Society, July 2017, https://hal.archives-ouvertes.fr/hal-01583301

[8] J. GOTTLIEB, P.-Y. OUDEYER. *Towards a neuroscience of active sampling and curiosity*, in "Nature Reviews Neuroscience", December 2018, vol. 19, n^{o} 12, pp. 758-770, https://hal.inria.fr/hal-01965608

[9] A. LAVERSANNE-FINOT, A. PÉRÉ, P.-Y. OUDEYER. *Curiosity Driven Exploration of Learned Disentangled Goal Spaces*, in "CoRL 2018 - Conference on Robot Learning", Zürich, Switzerland, October 2018, https://hal.inria.fr/hal-01891598

[10] T. LESORT, N. DÍAZ-RODRÍGUEZ, J.-F. GOUDOU, D. FILLIAT. *State Representation Learning for Control: An Overview*, in "Neural Networks", December 2018, vol. 108, pp. 379-392 [*DOI :* 10.1016/J.NEUNET.2018.07.006], https://hal.archives-ouvertes.fr/hal-01858558

[11] M. E. MEADE, J. G. MEADE, H. SAUZÉON, M. A. FERNANDES. *Active Navigation in Virtual Environments Benefits Spatial Memory in Older Adults*, in "Brain Sciences", 2019, vol. 9 [*DOI :* 10.3390/BRAINSCI9030047], https://hal.inria.fr/hal-02049031

[12] C. MOULIN-FRIER, J. BROCHARD, F. STULP, P.-Y. OUDEYER. *Emergent Jaw Predominance in Vocal Development through Stochastic Optimization*, in "IEEE Transactions on Cognitive and Developmental Systems", 2017, n^{o} 99, pp. 1-12 [*DOI :* 10.1109/TCDS.2017.2704912], https://hal.inria.fr/hal-01578075

[13] R. PORTELAS, C. COLAS, K. HOFMANN, P.-Y. OUDEYER. *Teacher algorithms for curriculum learning of Deep RL in continuously parameterized environments*, in "CoRL 2019 - Conference on Robot Learning", Osaka, Japan, October 2019, https://arxiv.org/abs/1910.07224 , https://hal.archives-ouvertes.fr/hal-02370165

[14] A. PÉRÉ, S. FORESTIER, O. SIGAUD, P.-Y. OUDEYER. *Unsupervised Learning of Goal Spaces for Intrinsically Motivated Goal Exploration*, in "ICLR2018 - 6th International Conference on Learning Representations", Vancouver, Canada, April 2018, https://hal.archives-ouvertes.fr/hal-01891758

[15] C. REINKE, M. ETCHEVERRY, P.-Y. OUDEYER. *Intrinsically Motivated Discovery of Diverse Patterns in Self-Organizing Systems*, in "International Conference on Learning Representations (ICLR)", Addis Ababa, Ethiopia, April 2020, Source code and videos athttps://automated-discovery.github.io/, https://hal.inria.fr/hal-02370003

Publications of the year

Doctoral Dissertations and Habilitation Theses

[16] C. MAZON. *Digital technologies for the school inclusion of children with ASD in middle school : from individual to ecosystemic approaches in supporting the individuals and their caregivers*, Université de Bordeaux, November 2019, https://hal.inria.fr/tel-02398226

Articles in International Peer-Reviewed Journals

[17] L. CAROUX, C. CONSEL, M. MERCIOL, H. SAUZÉON. *Acceptability of notifications delivered to older adults by technology-based assisted living services*, in "Universal Access in the Information Society", July 2019 [*DOI :* 10.1007/S10209-019-00665-Y], https://hal.inria.fr/hal-02179319

[18] P.-A. CINQUIN, P. GUITTON, H. SAUZÉON. *Online e-learning and cognitive disabilities: A systematic review*, in "Computers and Education", March 2019, vol. 130, pp. 152-167 [*DOI :* 10.1016/J.COMPEDU.2018.12.004], https://hal.archives-ouvertes.fr/hal-01954983

[19] C. CRAYE, T. LESORT, D. FILLIAT, J.-F. GOUDOU. *Exploring to learn visual saliency: The RL-IAC approach*, in "Robotics and Autonomous Systems", February 2019, vol. 112, pp. 244-259 [*DOI :* 10.1016/J.ROBOT.2018.11.012], https://hal.archives-ouvertes.fr/hal-01959882

[20] L. DUPUY, B. N'KAOUA, P. DEHAIL, H. SAUZÉON. *Role of cognitive resources on everyday functioning among oldest-old physically frail*, in "Aging Clinical and Experimental Research", October 2019 [*DOI :* 10.1007/S40520-019-01384-3], https://hal.inria.fr/hal-02353741

[21] C. FAGE, C. CONSEL, K. ETCHEGOYHEN, A. AMESTOY, M. BOUVARD, C. MAZON, H. SAUZÉON. *An emotion regulation app for school inclusion of children with ASD: Design principles and evaluation*, in "Computers and Education", April 2019, vol. 131, pp. 1-21 [*DOI :* 10.1016/J.COMPEDU.2018.12.003], https://hal.inria.fr/hal-02124850

[22] P. FOURNIER, C. COLAS, M. CHETOUANI, O. SIGAUD. *CLIC: Curriculum Learning and Imitation for object Control in non-rewarding environments*, in "IEEE Transactions on Cognitive and Developmental Systems", 2019, 1 p. , forthcoming [*DOI :* 10.1109/TCDS.2019.2933371], https://hal.archives-ouvertes.fr/hal-02370859

[23] T. LESORT, V. LOMONACO, A. STOIAN, D. MALTONI, D. FILLIAT, N. DÍAZ-RODRÍGUEZ. *Continual Learning for Robotics: Definition, Framework, Learning Strategies, Opportunities and Challenges*, in "Information Fusion", December 2019, https://arxiv.org/abs/1907.00182 [*DOI :* 10.1016/J.INFFUS.2019.12.004], https://hal.archives-ouvertes.fr/hal-02381343

[24] C. MAZON, C. FAGE, C. CONSEL, A. AMESTOY, I. HESLING, M. BOUVARD, K. ETCHEGOYHEN, H. SAUZÉON. *Cognitive Mediators of School-Related Socio- Adaptive Behaviors in ASD and Intellectual Disability Pre-and Adolescents: A Pilot-Study in French Special Education Classrooms*, in "Brain Sciences", 2019, vol. 9 [*DOI :* 10.3390/BRAINSCI9120334], https://hal.inria.fr/hal-02374929

[25] M. E. MEADE, J. G. MEADE, H. SAUZÉON, M. A. FERNANDES. *Active Navigation in Virtual Environments Benefits Spatial Memory in Older Adults*, in "Brain Sciences", 2019, vol. 9 [*DOI :* 10.3390/BRAINSCI9030047], https://hal.inria.fr/hal-02049031

[26] S. MICK, M. LAPEYRE, P. ROUANET, C. HALGAND, J. BENOIS-PINEAU, F. PACLET, D. CATTAERT, P.-Y. OUDEYER, A. DE RUGY. *Reachy, a 3D-Printed Human-Like Robotic Arm as a*

Testbed for Human-Robot Control Strategies, in "Frontiers in Neurorobotics", August 2019, vol. 13 [*DOI :* 10.3389/FNBOT.2019.00065], https://hal.archives-ouvertes.fr/hal-02326321

Articles in National Peer-Reviewed Journals

[27] C. ATLAN, J.-P. ARCHAMBAULT, O. BANUS, F. BARDEAU, A. BLANDEAU, A. COIS, M. COURBIN-COULAUD, G. GIRAUDON, S.-C. LEFÈVRE, V. LETARD, B. MASSE, F. MASSEGLIA, B. NINASSI, S. DE QUATREBARBES, M. ROMERO, D. ROY, T. VIÉVILLE. *Apprentissage de la pensée informatique : de la formation des enseignant·e·s à la formation de tou·te·s les citoyen·ne·s*, in "Revue de l'EPI (Enseignement Public et Informatique)", June 2019, https://arxiv.org/abs/1906.00647 , https://hal.inria.fr/hal-02145478

Invited Conferences

[28] H. SAUZÉON. *Assistances numériques pour la cognition quotidienne à tous les âges de la vie : Rôle de la motivation intrinsèque*, in "Colloque - Augmentation de l'humain : vers des systèmes cognitivement augmentés (chaire industrielle « Systèmes Technologiques pour l'Augmentation de l'Humain »)", Bordeaux, France, March 2019, https://hal.inria.fr/hal-02375475

International Conferences with Proceedings

[29] J. CEHA, N. CHHIBBER, J. GOH, C. MCDONALD, P.-Y. OUDEYER, D. KULIĆ, E. LAW. *Expression of Curiosity in Social Robots: Design, Perception, and Effects on Behaviour*, in "CHI 2019 - The ACM CHI Conference on Human Factors in Computing Systems", Glasgow, United Kingdom, May 2019, https://hal.inria.fr/hal-02371252

[30] C. REINKE, M. ETCHEVERRY, P.-Y. OUDEYER. *Intrinsically Motivated Discovery of Diverse Patterns in Self-Organizing Systems*, in "International Conference on Learning Representations (ICLR)", Addis Ababa, Ethiopia, April 2020, https://arxiv.org/abs/1908.06663 - Source code and videos athttps://automated-discovery.github.io/, https://hal.inria.fr/hal-02370003

Conferences without Proceedings

[31] C. ATLAN, J.-P. ARCHAMBAULT, O. BANUS, F. BARDEAU, A. BLANDEAU, A. COIS, M. COURBIN-COULAUD, G. GIRAUDON, S.-C. LEFÈVRE, V. LETARD, B. MASSE, F. MASSEGLIA, B. NINASSI, S. DE QUATREBARBES, M. ROMERO, D. ROY, T. VIÉVILLE. *Apprentissage de la pensée informatique : de la formation des enseignant·e·s à la formation de tou·te·s les citoyen·ne·s*, in "EIAH'19 Wokshop - Apprentissage de la pensée informatique de la maternelle à l'Université : retours d'expériences et passage à l'échelle", Paris, France, June 2019, https://hal.inria.fr/hal-02145480

[32] H. CASELLES-DUPRÉ, M. GARCIA-ORTIZ, D. FILLIAT. *Symmetry-Based Disentangled Representation Learning requires Interaction with Environments*, in "NeurIPS 2019 6 Neural Information Processing Conference", Vancouver, Canada, December 2019, https://arxiv.org/abs/1904.00243 , https://hal.archives-ouvertes.fr/hal-02379399

[33] C. COLAS, P. FOURNIER, O. SIGAUD, M. CHETOUANI, P.-Y. OUDEYER. *CURIOUS: Intrinsically Motivated Modular Multi-Goal Reinforcement Learning*, in "ICML 2019 - Thirty-sixth International Conference on Machine Learning", Long Beach, United States, June 2019, https://hal.archives-ouvertes.fr/hal-01934921

[34] C. COLAS, O. SIGAUD, P.-Y. OUDEYER. *A Hitchhiker's Guide to Statistical Comparisons of Reinforcement Learning Algorithms*, in "ICLR Worskhop on Reproducibility", Nouvelle-Orléans, United States, May 2019, https://arxiv.org/abs/1904.06979 , https://hal.archives-ouvertes.fr/hal-02369859

[35] N. LAIR, C. COLAS, R. PORTELAS, J.-M. DUSSOUX, P. DOMINEY, P.-Y. OUDEYER. *Language Grounding through Social Interactions and Curiosity-Driven Multi-Goal Learning*, in "NeurIPS Workshop on Visually Grounded Interaction and Language", Vancouver, Canada, December 2019, https://arxiv.org/abs/1911.03219 , https://hal.archives-ouvertes.fr/hal-02369866

[36] T. LESORT, H. CASELLES-DUPRÉ, M. GARCIA-ORTIZ, J.-F. GOUDOU, D. FILLIAT. *Generative Models from the perspective of Continual Learning*, in "IJCNN - International Joint Conference on Neural Networks", Budapest, Hungary, July 2019, https://hal.archives-ouvertes.fr/hal-01951954

[37] T. LESORT, M. SEURIN, X. LI, N. DÍAZ-RODRÍGUEZ, D. FILLIAT. *Deep unsupervised state representation learning with robotic priors: a robustness analysis*, in "IJCNN 2019 - International Joint Conference on Neural Networks", Budapest, Hungary, IEEE, July 2019, pp. 1-8 [*DOI :* 10.1109/IJCNN.2019.8852042], https://hal.archives-ouvertes.fr/hal-02381375

[38] R. PORTELAS, C. COLAS, K. HOFMANN, P.-Y. OUDEYER. *Teacher algorithms for curriculum learning of Deep RL in continuously parameterized environments*, in "CoRL 2019 - Conference on Robot Learning", Osaka, Japan, October 2019, https://arxiv.org/abs/1910.07224 , https://hal.archives-ouvertes.fr/hal-02370165

[39] A. RAFFIN, A. HILL, R. TRAORÉ, T. LESORT, N. DÍAZ-RODRÍGUEZ, D. FILLIAT. *Decoupling feature extraction from policy learning: assessing benefits of state representation learning in goal based robotics*, in "SPiRL 2019 : Workshop on Structure and Priors in Reinforcement Learning at ICLR 2019", Nouvelle Orléans, United States, May 2019, https://arxiv.org/abs/1901.08651 - Github repo: https://github.com/araffin/srl-zoo Documentation: https://srl-zoo.readthedocs.io/en/latest/, As part of SRL-Toolbox: https://s-rl-toolbox.readthedocs.io/en/latest/. Accepted to the Workshop on Structure & Priors in Reinforcement Learning at ICLR 2019, https://hal.archives-ouvertes.fr/hal-02285831

[40] R. TRAORÉ, H. CASELLES-DUPRÉ, T. LESORT, T. SUN, G. CAI, D. FILLIAT, N. DÍAZ-RODRÍGUEZ. *DISCORL: Continual reinforcement learning via policy distillation : A preprint*, in "NeurIPS workshop on Deep Reinforcement Learning", Vancouver, Canada, December 2019, https://hal.archives-ouvertes.fr/hal-02381494

[41] R. TRAORÉ, H. CASELLES-DUPRÉ, T. LESORT, T. SUN, N. DÍAZ-RODRÍGUEZ, D. FILLIAT. *Continual Reinforcement Learning deployed in Real-life using Policy Distillation and Sim2Real Transfer*, in "ICML Workshop on “Multi-Task and Lifelong Reinforcement Learning”", Long Beach, United States, June 2019, https://arxiv.org/abs/1906.04452 - accepted to the Workshop on Multi-Task and Lifelong Reinforcement Learning, ICML 2019, https://hal.archives-ouvertes.fr/hal-02285839

Scientific Books (or Scientific Book chapters)

[42] P.-A. CINQUIN, P. GUITTON, H. SAUZÉON. *Accessibilité numérique des systèmes d’enseignement en ligne pour des personnes en situation de handicap d’origine cognitif*, in "Handicaps et recherches - Regards pluridiciplinaires", E. DUGAS (editor), Editions CNRS, 2019, https://hal.inria.fr/hal-02433430

[43] P. KARVINEN, N. DÍAZ-RODRÍGUEZ, S. GRÖNROOS, J. LILIUS. *RDF Stores for Enhanced Living Environments: An Overview*, in "Enhanced Living Environments: Algorithms, Architectures, Platforms, and Systems", I. GANCHEV, N. M. GARCIA, C. DOBRE, C. X. MAVROMOUSTAKIS, R. GOLEVA (editors), Springer, January 2019, pp. 19-52 [*DOI :* 10.1007/978-3-030-10752-9_2], https://hal.archives-ouvertes.fr/hal-02381354

[44] P.-Y. OUDEYER, G. KACHERGIS, W. SCHUELLER. *Computational and Robotic Models of Early Language Development: A Review*, in "International Handbook of Language Acquisition", May 2019, https://hal.inria.fr/hal-02371233

[45] H. SAUZÉON, L. DUPUY, C. FAGE, C. MAZON. *Assistances numériques pour la cognition quotidienne à tous les âges de la vie*, in "Handicap et Recherches : Regards pluridisciplinaires", CNRS Editions, May 2019, https://hal.inria.fr/hal-02375456

Other Publications

[46] A. B. ARRIETA, N. DÍAZ-RODRÍGUEZ, J. DEL SER, A. BENNETOT, S. TABIK, A. BARBADO, S. GARCÍA, S. GIL-LÓPEZ, D. MOLINA, R. BENJAMINS, R. CHATILA, F. HERRERA. *Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI*, November 2019, https://arxiv.org/abs/1910.10045 - 67 pages, 13 figures, under review in the Information Fusion journal [*DOI :* 10.10045], https://hal.archives-ouvertes.fr/hal-02381211

[47] A. BENNETOT, J.-L. LAURENT, R. CHATILA, N. DÍAZ-RODRÍGUEZ. *Towards Explainable Neural-Symbolic Visual Reasoning*, November 2019, https://arxiv.org/abs/1909.09065 - Accepted at IJCAI19 Neural-Symbolic Learning and Reasoning Workshop (https://sites.google.com/view/nesy2019/home), https://hal.archives-ouvertes.fr/hal-02379596

[48] T. GILLIARD, T. DESPREZ, P.-Y. OUDEYER. *Conception and testing of modular robotic kits based on Poppy Ergo Jr for educational purposes*, March 2019, Colloque des Jeunes Chercheurs en Sciences Cognitives (CJC2019), Poster, https://hal.inria.fr/hal-02154848

References in notes

[49] L. STEELS, R. BROOKS (editors). *The Artificial Life Route to Artificial Intelligence: Building Embodied, Situated Agents*, L. Erlbaum Associates Inc., Hillsdale, NJ, USA, 1995

[50] B. A. ANDERSON, P. A. LAURENT, S. YANTIS. *Value-driven attentional capture*, in "Proceedings of the National Academy of Sciences", 2011, vol. 108, n^{o} 25, pp. 10367–10371

[51] M. ANDRYCHOWICZ, F. WOLSKI, A. RAY, J. SCHNEIDER, R. FONG, P. WELINDER, B. MCGREW, J. TOBIN, O. P. ABBEEL, W. ZAREMBA. *Hindsight experience replay*, in "Advances in Neural Information Processing Systems", 2017, pp. 5048–5058

[52] B. ARGALL, S. CHERNOVA, M. VELOSO. *A Survey of Robot Learning from Demonstration*, in "Robotics and Autonomous Systems", 2009, vol. 57, n^{o} 5, pp. 469–483

[53] M. ASADA, S. NODA, S. TAWARATSUMIDA, K. HOSODA. *Purposive Behavior Acquisition On A Real Robot By Vision-Based Reinforcement Learning*, in "Machine Learning", 1996, vol. 23, pp. 279-303

[54] G. BALDASSARRE, M. MIROLLI. *Intrinsically Motivated Learning in Natural and Artificial Systems*, Springer, 2013

[55] A. BARANES, P.-Y. OUDEYER. *Active Learning of Inverse Models with Intrinsically Motivated Goal Exploration in Robots*, in "Robotics and Autonomous Systems", January 2013, vol. 61, n^{o} 1, pp. 69-73 [*DOI :* 10.1016/J.ROBOT.2012.05.008], https://hal.inria.fr/hal-00788440

[56] A. BARTO, M. MIROLLI, G. BALDASSARRE. *Novelty or surprise?*, in "Frontiers in psychology", 2013, vol. 4

[57] A. BARTO, S. SINGH, N. CHENTANEZ. *Intrinsically Motivated Learning of Hierarchical Collections of Skills*, in "Proceedings of the 3rd International Conference on Development and Learning (ICDL 2004)", Salk Institute, San Diego, 2004

[58] A. BENNETOT, J.-L. LAURENT, R. CHATILA, N. DÍAZ-RODRÍGUEZ. *Towards Explainable Neural-Symbolic Visual Reasoning*, November 2019, Accepted at IJCAI19 Neural-Symbolic Learning and Reasoning Workshop (https://sites.google.com/view/nesy2019/home), https://hal.archives-ouvertes.fr/hal-02379596

[59] D. BERLYNE. *Conflict, Arousal and Curiosity*, McGraw-Hill, 1960

[60] C. BREAZEAL. *Designing sociable robots*, The MIT Press, 2004

[61] G. BROCKMAN, V. CHEUNG, L. PETTERSSON, J. SCHNEIDER, J. SCHULMAN, J. TANG, W. ZAREMBA. *Openai gym*, in "arXiv preprint arXiv:1606.01540", 2016

[62] J. BROOKE. *SUS-A quick and dirty usability scale*, in "Usability evaluation in industry", 1996, vol. 189, n^{o} 194, pp. 4–7

[63] R. BROOKS, C. BREAZEAL, R. IRIE, C. C. KEMP, B. SCASSELLATI, M. WILLIAMSON. *Alternative essences of intelligence*, in "Proceedings of 15th National Conference on Artificial Intelligence (AAAI-98)", AAAI Press, 1998, pp. 961–968

[64] H. CASELLES-DUPRÉ, L. ANNABI, O. HAGEN, M. GARCIA-ORTIZ, D. FILLIAT. *Flatland: a Lightweight First-Person 2-D Environment for Reinforcement Learning*, in "Workshop on Continual Unsupervised Sensorimotor Learning, ICDL-EpiRob 2018", Tokyo, Japan, September 2018, https://hal.archives-ouvertes.fr/hal-01951945

[65] H. CASELLES-DUPRÉ, M. GARCIA-ORTIZ, D. FILLIAT. *Continual State Representation Learning for Reinforcement Learning using Generative Replay*, in "Workshop on Continual Learning, NeurIPS 2018 (Neural Information Processing Systems)", Montreal, Canada, December 2018, https://hal.archives-ouvertes.fr/hal-01951951

[66] B. W.-C. CHAN. *Lenia: Biology of Artificial Life*, in "Complex Systems", 2019, vol. 28, n^{o} 3, pp. 251—286

[67] A. CLARK. *Mindware: An Introduction to the Philosophy of Cognitive Science*, Oxford University Press, 2001

[68] B. CLÉMENT, D. ROY, P.-Y. OUDEYER, M. LOPES. *Multi-Armed Bandits for Intelligent Tutoring Systems*, in "Journal of Educational Data Mining (JEDM)", June 2015, vol. 7, n^{o} 2, pp. 20–48, https://hal.inria.fr/hal-00913669

[69] B. CLÉMENT. *Adaptive Personalization of Pedagogical Sequences using Machine Learning*, Université de Bordeaux, December 2018, https://hal.inria.fr/tel-01968241

[70] D. COHN, Z. GHAHRAMANI, M. JORDAN. *Active learning with statistical models*, in "Journal of artificial intelligence research", 1996, vol. 4, pp. 129–145

[71] C. COLAS, O. SIGAUD, P.-Y. OUDEYER. *How Many Random Seeds? Statistical Power Analysis in Deep Reinforcement Learning Experiments*, October 2018, working paper or preprint, https://hal.inria.fr/hal-01890154

[72] W. CROFT, D. CRUSE. *Cognitive Linguistics*, Cambridge Textbooks in Linguistics, Cambridge University Press, 2004

[73] M. CSIKSZENTHMIHALYI. *Flow-the psychology of optimal experience*, Harper Perennial, 1991

[74] P. DAYAN, W. BELLEINE. *Reward, motivation and reinforcement learning*, in "Neuron", 2002, vol. 36, pp. 285–298

[75] E. DECI, R. RYAN. *Intrinsic Motivation and Self-Determination in Human Behavior*, Plenum Press, 1985

[76] T. DESPREZ, S. NOIRPOUDRE, T. SEGONDS, D. CASELLI, D. ROY, P.-Y. OUDEYER. *Poppy Ergo Jr : un kit robotique au coeur du dispositif Poppy éducation*, in "Didapro 7 2018 - DidaSTIC Colloque de didactique de l'informatique", Lausanne, Switzerland, February 2018, pp. 1-6, https://hal.inria.fr/hal-01753111

[77] S. DONCIEUX, D. FILLIAT, N. DÍAZ-RODRÍGUEZ, T. HOSPEDALES, R. DURO, A. CONINX, D. M. ROIJERS, B. GIRARD, N. PERRIN, O. SIGAUD. *Open-Ended Learning: A Conceptual Framework Based on Representational Redescription*, in "Frontiers in Neurorobotics", 2018, vol. 12, 59 p. [*DOI :* 10.3389/FNBOT.2018.00059], https://hal.sorbonne-universite.fr/hal-01889947

[78] N. DÍAZ-RODRÍGUEZ, V. LOMONACO, D. FILLIAT, D. MALTONI. *Don't forget, there is more than forgetting: new metrics for Continual Learning*, in "Workshop on Continual Learning, NeurIPS 2018 (Neural Information Processing Systems", Montreal, Canada, December 2018, https://hal.archives-ouvertes.fr/hal-01951488

[79] J. ELMAN. *Learning and development in neural networks: The importance of starting small*, in "Cognition", 1993, vol. 48, pp. 71–99

[80] S. B. FLAGEL, H. AKIL, T. E. ROBINSON. *Individual differences in the attribution of incentive salience to reward-related cues: Implications for addiction*, in "Neuropharmacology", 2009, vol. 56, pp. 139–148

[81] S. FORESTIER, Y. MOLLARD, D. CASELLI, P.-Y. OUDEYER. *Autonomous exploration, active learning and human guidance with open-source Poppy humanoid robot platform and Explauto library*, in "The Thirtieth Annual Conference on Neural Information Processing Systems (NIPS 2016)", 2016

[82] S. FORESTIER, Y. MOLLARD, P.-Y. OUDEYER. *Intrinsically Motivated Goal Exploration Processes with Automatic Curriculum Learning*, November 2017, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01651233

[83] S. FORESTIER, Y. MOLLARD, P.-Y. OUDEYER. *Intrinsically motivated goal exploration processes with automatic curriculum learning*, in "arXiv preprint arXiv:1708.02190", 2017

[84] S. FUJIMOTO, H. VAN HOOF, D. MEGER. *Addressing function approximation error in actor-critic methods*, in "arXiv preprint arXiv:1802.09477", 2018

[85] J. GOTTLIEB, P.-Y. OUDEYER, M. LOPES, A. BARANES. *Information-seeking, curiosity, and attention: computational and neural mechanisms*, in "Trends in Cognitive Sciences", November 2013, vol. 17, n^{o} 11, pp. 585-93 [*DOI :* 10.1016/J.TICS.2013.09.001], https://hal.inria.fr/hal-00913646

[86] J. GOTTLIEB, P.-Y. OUDEYER, M. LOPES, A. BARANES. *Information-seeking, curiosity, and attention: computational and neural mechanisms*, in "Trends in cognitive sciences", 2013, vol. 17, n^{o} 11, pp. 585–593

[87] T. GUITARD, D. ROY, P.-Y. OUDEYER, M. CHEVALIER. *IniRobot*, January 2016, Des activités robotiques pour l'initiation aux sciences du numérique, https://hal.inria.fr/hal-01412928

[88] T. HAARNOJA, A. ZHOU, P. ABBEEL, S. LEVINE. *Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor*, in "arXiv preprint arXiv:1801.01290", 2018

[89] S. HARNAD. *The symbol grounding problem*, in "Physica D", 1990, vol. 40, pp. 335–346

[90] M. HASENJAGER, H. RITTER. *Active learning in neural networks*, Physica-Verlag GmbH, Heidelberg, Germany, Germany, 2002, pp. 137–169

[91] J. HAUGELAND. *Artificial Intelligence: the very idea*, The MIT Press, Cambridge, MA, USA, 1985

[92] A. HILL, A. RAFFIN, M. ERNESTUS, R. TRAORE, P. DHARIWAL, C. HESSE, O. KLIMOV, A. NICHOL, M. PLAPPERT, A. RADFORD, J. SCHULMAN, S. SIDOR, Y. WU. *Stable Baselines*, GitHub, 2018, https://github.com/hill-a/stable-baselines

[93] J.-C. HORVITZ. *Mesolimbocortical and nigrostriatal dopamine responses to salient non-reward events*, in "Neuroscience", 2000, vol. 96, n^{o} 4, pp. 651-656

[94] X. HUANG, J. WENG. *Novelty and reinforcement learning in the value system of developmental robots*, in "Proceedings of the 2nd international workshop on Epigenetic Robotics : Modeling cognitive development in robotic systems", C. PRINCE, Y. DEMIRIS, Y. MAROM, H. KOZIMA, C. BALKENIUS (editors), Lund University Cognitive Studies 94, 2002, pp. 47–55

[95] S. IVALDI, N. LYUBOVA, D. GÉRARDEAUX-VIRET, A. DRONIOU, S. ANZALONE, M. CHETOUANI, D. FILLIAT, O. SIGAUD. *Perception and human interaction for developmental learning of objects and affordances*, in "Proc. of the 12th IEEE-RAS International Conference on Humanoid Robots - HUMANOIDS", Japan, 2012, forthcoming, http://hal.inria.fr/hal-00755297

[96] M. JOHNSON. *Developmental Cognitive Neuroscience*, 2nd, Blackwell publishing, 2005

[97] P. KARVINEN, N. DÍAZ-RODRÍGUEZ, S. GRÖNROOS, J. LILIUS. *RDF Stores for Enhanced Living Environments: An Overview*, in "Enhanced Living Environments: Algorithms, Architectures, Platforms, and Systems", I. GANCHEV, N. M. GARCIA, C. DOBRE, C. X. MAVROMOUSTAKIS, R. GOLEVA (editors), Springer, January 2019, pp. 19-52 [*DOI :* 10.1007/978-3-030-10752-9_2], https://hal.archives-ouvertes.fr/hal-02381354

[98] C. KIDD, B. HAYDEN. *The psychology and neuroscience of curiosity*, in "Neuron (in press)", 2015

[99] W. B. KNOX, P. STONE. *Combining manual feedback with subsequent MDP reward signals for reinforcement learning*, in "Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems (AAMAS'10)", 2010, pp. 5–12

[100] C. LALLEMAND, V. KOENIG, G. GRONIER, R. MARTIN. *Création et validation d'une version française du questionnaire AttrakDiff pour l'évaluation de l'expérience utilisateur des systèmes interactifs*, in "Revue Européenne de Psychologie Appliquée/European Review of Applied Psychology", 2015, vol. 65, n^{o} 5, pp. 239–252

[101] A. LAVERSANNE-FINOT, A. PÉRÉ, P.-Y. OUDEYER. *Curiosity Driven Exploration of Learned Disentangled Goal Spaces*, in "CoRL 2018 - Conference on Robot Learning", Zürich, Switzerland, October 2018, https://hal.inria.fr/hal-01891598

[102] T. LESORT, H. CASELLES-DUPRÉ, M. GARCIA-ORTIZ, J.-F. GOUDOU, D. FILLIAT. *Generative Models from the perspective of Continual Learning*, in "IJCNN - International Joint Conference on Neural Networks", Budapest, Hungary, July 2019, https://hal.archives-ouvertes.fr/hal-01951954

[103] T. LESORT, N. DÍAZ-RODRÍGUEZ, J.-F. GOUDOU, D. FILLIAT. *State Representation Learning for Control: An Overview*, in "Neural Networks", December 2018, vol. 108, pp. 379-392 [*DOI :* 10.1016/J.NEUNET.2018.07.006], https://hal.archives-ouvertes.fr/hal-01858558

[104] T. LESORT, V. LOMONACO, A. STOIAN, D. MALTONI, D. FILLIAT, N. DÍAZ-RODRÍGUEZ. *Continual Learning for Robotics: Definition, Framework, Learning Strategies, Opportunities and Challenges*, November 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02381343

[105] T. LESORT, M. SEURIN, X. LI, N. DÍAZ-RODRÍGUEZ, D. FILLIAT. *Deep unsupervised state representation learning with robotic priors: a robustness analysis*, in "IJCNN 2019 - International Joint Conference on Neural Networks", Budapest, Hungary, IEEE, July 2019, pp. 1-8 [*DOI :* 10.1109/IJCNN.2019.8852042], https://hal.archives-ouvertes.fr/hal-02381375

[106] G. LOEWENSTEIN. *The psychology of curiosity: A review and reinterpretation*, in "Psychological bulletin", 1994, vol. 116, n^{o} 1, 75 p.

[107] M. LOPES, T. CEDERBORG, P.-Y. OUDEYER. *Simultaneous Acquisition of Task and Feedback Models*, in "Development and Learning (ICDL), 2011 IEEE International Conference on", Germany, 2011, pp. 1 - 7 [*DOI :* 10.1109/DEVLRN.2011.6037359], http://hal.inria.fr/hal-00636166/en

[108] M. LOPES, T. LANG, M. TOUSSAINT, P.-Y. OUDEYER. *Exploration in Model-based Reinforcement Learning by Empirically Estimating Learning Progress*, in "Neural Information Processing Systems (NIPS)", Lake Tahoe, United States, December 2012, http://hal.inria.fr/hal-00755248

[109] M. LUNGARELLA, G. METTA, R. PFEIFER, G. SANDINI. *Developmental Robotics: A Survey*, in "Connection Science", 2003, vol. 15, n^{o} 4, pp. 151-190

[110] N. LYUBOVA, D. FILLIAT. *Developmental Approach for Interactive Object Discovery*, in "Neural Networks (IJCNN), The 2012 International Joint Conference on", Australia, June 2012, pp. 1-7 [*DOI :* 10.1109/IJCNN.2012.6252606], https://hal.archives-ouvertes.fr/hal-00755298

[111] D. J. MANKOWITZ, A. ŽÍDEK, A. BARRETO, D. HORGAN, M. HESSEL, J. QUAN, J. OH, H. VAN HASSELT, D. SILVER, T. SCHAUL. *Unicorn: Continual Learning with a Universal, Off-policy Agent*, in "arXiv preprint arXiv:1802.08294", 2018

[112] J. MARSHALL, D. BLANK, L. MEEDEN. *An Emergent Framework for Self-Motivation in Developmental Robotics*, in "Proceedings of the 3rd International Conference on Development and Learning (ICDL 2004)", Salk Institute, San Diego, 2004

[113] M. MASON, M. LOPES. *Robot Self-Initiative and Personalization by Learning through Repeated Interactions*, in "6th ACM/IEEE International Conference on Human-Robot", Switzerland, 2011 [*DOI :* 10.1145/1957656.1957814], http://hal.inria.fr/hal-00636164/en

[114] C. MAZON, C. FAGE, H. SAUZÉON. *Effectiveness and usability of technology-based interventions for children and adolescents with ASD: A systematic review of reliability, consistency, generalization and durability related to the effects of intervention*, in "Computers in Human Behavior", 2019, vol. 93, pp. 235–251

[115] M. MCCLOSKEY, N. J. COHEN. *Catastrophic interference in connectionist networks: The sequential learning problem*, in "Psychology of learning and motivation", Elsevier, 1989, vol. 24, pp. 109–165

[116] D. MCFADDEN. *Conditional logit analysis of qualitative choice behavior*, in "Frintiers in Econometrics", New York, P. ZAREMBKA (editor), Academic Press, 1973, chap. 4, pp. 105-142

[117] P. MILLER. *Theories of developmental psychology*, 4th, New York: Worth, 2001

[118] M. MIROLLI, G. BALDASSARRE. *Functions and mechanisms of intrinsic motivations*, in "Intrinsically Motivated Learning in Natural and Artificial Systems", Springer, 2013, pp. 49–72

[119] C. MOULIN-FRIER, P.-Y. OUDEYER. *Exploration strategies in developmental robotics: a unified probabilistic framework*, in "ICDL-Epirob - International Conference on Development and Learning, Epirob", Osaka, Japan, August 2013, https://hal.inria.fr/hal-00860641

[120] C. MOULIN-FRIER, P. ROUANET, P.-Y. OUDEYER. *Explauto: an open-source Python library to study autonomous exploration in developmental robotics*, in "ICDL-Epirob - International Conference on Development and Learning, Epirob", Genoa, Italy, October 2014, https://hal.inria.fr/hal-01061708

[121] A. NAIR, V. PONG, M. DALAL, S. BAHL, S. LIN, S. LEVINE. *Visual Reinforcement Learning with Imagined Goals*, in "CoRR", 2018, vol. abs/1807.04742, http://arxiv.org/abs/1807.04742

[122] S. M. NGUYEN, A. BARANES, P.-Y. OUDEYER. *Bootstrapping Intrinsically Motivated Learning with Human Demonstrations*, in "IEEE International Conference on Development and Learning", Frankfurt, Germany, 2011, http://hal.inria.fr/hal-00645986/en

[123] S. M. NGUYEN, A. BARANES, P.-Y. OUDEYER. *Constraining the Size Growth of the Task Space with Socially Guided Intrinsic Motivation using Demonstrations.*, in "IJCAI Workshop on Agents Learning Interactively from Human Teachers (ALIHT)", Barcelona, Spain, 2011, http://hal.inria.fr/hal-00645995/en

[124] S. M. NGUYEN, P.-Y. OUDEYER. *Socially Guided Intrinsic Motivation for Robot Learning of Motor Skills*, in "Autonomous Robots", March 2014, vol. 36, n° 3, pp. 273-294 [*DOI :* 10.1007/S10514-013-9339-Y], https://hal.inria.fr/hal-00936938

[125] S. NOIRPOUDRE, D. ROY, M. DEMANGEAT, T. DESPREZ, T. SEGONDS, P. ROUANET, D. CASELLI, N. RABAULT, M. LAPEYRE, P.-Y. OUDEYER. *Livret pédagogique : Apprendre à programmer Poppy Ergo Jr en Snap!*, June 2016, 50 p. , Un livret composé d'activités pédagogiques pour apprendre les bases de la programmation (programmation séquentielles, boucles, conditions, variables etc.) et des idées de défis et de projets pour appliquer les connaissances, https://hal.inria.fr/hal-01384649

[126] S. NOIRPOUDRE, D. ROY, T. DESPREZ, T. SEGONDS, D. CASELLI, P.-Y. OUDEYER. *Poppy Education: un dispositif robotique open source pour l'enseignement de l'informatique et de la robotique*, in "EIAH 2017-Environnements Informatiques pour l'Apprentissage Humain", 2017, 8 p.

[127] P.-Y. OUDEYER, F. KAPLAN, V. HAFNER. *Intrinsic Motivation for Autonomous Mental Development*, in "IEEE Transactions on Evolutionary Computation", January 2007, vol. 11, n° 2, pp. 265-286 [*DOI :* 10.1109/TEVC.2006.890271], https://hal.inria.fr/hal-00793610

[128] P.-Y. OUDEYER, F. KAPLAN, V. HAFNER. *Intrinsic Motivation Systems for Autonomous Mental Development*, in "IEEE Transactions on Evolutionary Computation", 2007, vol. 11, n° 1, pp. 265–286, http://www.pyoudeyer.com/ims.pdf

[129] P.-Y. OUDEYER, F. KAPLAN. *Intelligent adaptive curiosity: a source of self-development*, in "Proceedings of the 4th International Workshop on Epigenetic Robotics", L. BERTHOUZE, H. KOZIMA, C. PRINCE, G. SANDINI, G. STOJANOV, G. METTA, C. BALKENIUS (editors), Lund University Cognitive Studies, 2004, vol. 117, pp. 127–130

[130] P.-Y. OUDEYER, F. KAPLAN. *What is intrinsic motivation? A typology of computational approaches*, in "Frontiers in Neurorobotics", 2007, vol. 1, n° 1

[131] P.-Y. OUDEYER. *Sur les interactions entre la robotique et les sciences de l'esprit et du comportement*, in "Informatique et Sciences Cognitives : influences ou confluences ?", C. GARBAY, D. KAISER (editors), Presses Universitaires de France, 2009, http://hal.inria.fr/inria-00420309/en/

[132] P.-Y. OUDEYER. *L'auto-organisation dans l'évolution de la parole*, in "Parole et Musique: Aux origines du dialogue humain, Colloque annuel du Collège de France", S. DEHAENE, C. PETIT (editors), Odile Jacob, 2009, pp. 83-112, http://hal.inria.fr/inria-00446908/en/

[133] M. PELZ, S. T. PIANTADOSI, C. KIDD. *The dynamics of idealized attention in complex learning environments*, in "IEEE International Conference on Development and Learning and on Epigenetic Robotics", 2015

[134] L. J. POINTS, J. W. TAYLOR, J. GRIZOU, K. DONKERS, L. CRONIN. *Artificial intelligence exploration of unstable protocells leads to predictable properties and discovery of collective behavior*, in "Proceedings of the National Academy of Sciences", 2018, 201711089 p.

[135] V. H. PONG, M. DALAL, S. LIN, A. NAIR, S. BAHL, S. LEVINE. *Skew-Fit: State-Covering Self-Supervised Reinforcement Learning*, in "CoRR", 2019, vol. abs/1903.03698, http://arxiv.org/abs/1903.03698

[136] A. PÉRÉ, S. FORESTIER, O. SIGAUD, P.-Y. OUDEYER. *Unsupervised Learning of Goal Spaces for Intrinsically Motivated Goal Exploration*, in "ICLR2018 - 6th International Conference on Learning Representations", Vancouver, Canada, April 2018, https://hal.archives-ouvertes.fr/hal-01891758

[137] A. RAFFIN, A. HILL, R. TRAORÉ, T. LESORT, N. DÍAZ-RODRÍGUEZ, D. FILLIAT. *S-RL Toolbox: Environments, Datasets and Evaluation Metrics for State Representation Learning*, in "NeurIPS 2018 Workshop on “Deep Reinforcement Learning”", Montreal, Canada, December 2018, https://hal.archives-ouvertes.fr/hal-01931713

[138] A. RAFFIN, A. HILL, R. TRAORÉ, T. LESORT, N. DÍAZ-RODRÍGUEZ, D. FILLIAT. *Decoupling feature extraction from policy learning: assessing benefits of state representation learning in goal based robotics*, in "SPiRL 2019 : Workshop on Structure and Priors in Reinforcement Learning at ICLR 2019", Nouvelle Orléans, United States, May 2019, Github repo: https://github.com/araffin/srl-zoo Documentation: https://srl-zoo.readthedocs.io/en/latest/, As part of SRL-Toolbox: https://s-rl-toolbox.readthedocs.io/en/latest/. Accepted to the Workshop on Structure & Priors in Reinforcement Learning at ICLR 2019, https://hal.archives-ouvertes.fr/hal-02285831

[139] A. REVEL, J. NADEL. *How to build an imitator?*, in "Imitation and Social Learning in Robots, Humans and Animals: Behavioural, Social and Communicative Dimensions", K. DAUTENHAHN, C. NEHANIV (editors), Cambridge University Press, 2004

[140] E. F. RISKO, N. C. ANDERSON, S. LANTHIER, A. KINGSTONE. *Curious eyes: Individual differences in personality predict eye movement behavior in scene-viewing*, in "Cognition", 2012, vol. 122, n^o 1, pp. 86–90

[141] V. G. SANTUCCI, G. BALDASSARRE, M. MIROLLI. *Which is the best intrinsic motivation signal for learning multiple skills?*, in "Frontiers in neurorobotics", 2013, vol. 7

[142] N. SAVINOV, A. RAICHUK, R. MARINIER, D. VINCENT, M. POLLEFEYS, T. P. LILLICRAP, S. GELLY. *Episodic Curiosity through Reachability*, in "CoRR", 2018, vol. abs/1810.02274, http://arxiv.org/abs/1810.02274

[143] T. SCHATZ, P.-Y. OUDEYER. *Learning motor dependent Crutchfield's information distance to anticipate changes in the topology of sensory body maps*, in "IEEE International Conference on Learning and Development", Chine Shangai, 2009, http://hal.inria.fr/inria-00420186/en/

[144] T. SCHAUL, D. HORGAN, K. GREGOR, D. SILVER. *Universal value function approximators*, in "International Conference on Machine Learning", 2015, pp. 1312–1320

[145] M. SCHEMBRI, M. MIROLLI, G. BALDASSARRE. *Evolving internal reinforcers for an intrinsically motivated reinforcement-learning robot*, in "IEEE 6th International Conference on Development and Learning, 2007. ICDL 2007.", July 2007, pp. 282-287, http://dx.doi.org/10.1109/DEVLRN.2007.4354052

[146] J. SCHMIDHUBER. *Curious Model-Building Control Systems*, in "Proceedings of the International Joint Conference on Neural Networks, Singapore", IEEE press, 1991, vol. 2, pp. 1458–1463

[147] W. SCHULTZ, P. DAYAN, P. MONTAGUE. *A neural substrate of prediction and reward*, in "Science", 1997, vol. 275, pp. 1593-1599

[148] K. O. STANLEY. *Exploiting regularity without development*, in "Proceedings of the AAAI Fall Symposium on Developmental Systems", AAAI Press Menlo Park, CA, 2006, 37 p.

[149] E. SUMNER, E. DEANGELIS, M. HYATT, N. GOODMAN, C. KIDD. *Toddlers Always Get the Last Word: Recency biases in early verbal behavior*, in "Proceedings of the 37th Annual Meeting of the Cognitive Science Society", 2015

[150] E. THELEN, L. B. SMITH. *A dynamic systems approach to the development of cognition and action*, MIT Press, Cambridge, MA, 1994

[151] A. L. THOMAZ, C. BREAZEAL. *Teachable robots: Understanding human teaching behavior to build more effective robot learners*, in "Artificial Intelligence Journal", 2008, vol. 172, pp. 716-737

[152] R. TRAORÉ, H. CASELLES-DUPRÉ, T. LESORT, T. SUN, G. CAI, D. FILLIAT, N. DÍAZ-RODRÍGUEZ. *DISCORL: Continual reinforcement learning via policy distillation*, in "NeurIPS workshop on Deep Reinforcement Learning", Vancouver, Canada, December 2019, https://hal.archives-ouvertes.fr/hal-02381494

[153] R. TRAORÉ, H. CASELLES-DUPRÉ, T. LESORT, T. SUN, N. DÍAZ-RODRÍGUEZ, D. FILLIAT. *Continual Reinforcement Learning deployed in Real-life using Policy Distillation and Sim2Real Transfer*, in "ICML Workshop on “Multi-Task and Lifelong Reinforcement Learning”", Long Beach, United States, June 2019, accepted to the Workshop on Multi-Task and Lifelong Reinforcement Learning, ICML 2019, https://hal.archives-ouvertes.fr/hal-02285839

[154] A. TURING. *Computing machinery and intelligence*, in "Mind", 1950, vol. 59, pp. 433-460

[155] M. R. UNCAPHER, M. K. THIEU, A. D. WAGNER. *Media multitasking and memory: Differences in working memory and long-term memory*, in "Psychonomic bulletin & review", 2015, pp. 1–8

[156] F. VARELA, E. THOMPSON, E. ROSCH. *The embodied mind : Cognitive science and human experience*, MIT Press, Cambridge, MA, 1991

[157] S. VENKATTARAMANUJAM, E. CRAWFORD, T. DOAN, D. PRECUP. *Self-supervised Learning of Distance Functions for Goal-Conditioned Reinforcement Learning*, in "CoRR", 2019, vol. abs/1907.02998, http://arxiv.org/abs/1907.02998

[158] J. WENG, J. MCCLELLAND, A. PENTLAND, O. SPORNS, I. STOCKMAN, M. SUR, E. THELEN. *Autonomous mental development by robots and animals*, in "Science", 2001, vol. 291, pp. 599-600

[159] R. C. WILSON, A. G. COLLINS. *Ten simple rules for the computational modeling of behavioral data*, in "eLife", 2019, vol. 8, e49547 p.