Activity Report 2018

Project-Team FLOWERS

Flowing Epigenetic Robots and Systems

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]

    Olivier Sigaud [Inria, Senior Researcher, until Aug 2018, HDR]

Faculty Members

    David Filliat [École Nationale Supérieure de Techniques Avancées, Professor, HDR]

    Natalia Diaz Rodriguez [École Nationale Supérieure de Techniques Avancées, from march 2018, Associate Professor]

Post-Doctoral Fellows

    Adrien Laversanne Finot [Inria]

    Chris Reinke [Inria, from Nov 2018]

PhD Students

    Florence Carton [CEA]

    Benjamin Clément [Inria]

    Cedric Colas [Inria]

    Alexandra-Alice Delmas-Mitchell [Inria, until Aug 2018]

    Thibault Desprez [Inria]

    Sébastien Forestier [Univ. de Bordeaux]

    Florian Golemo [Inria, until Nov 2018]

    Timothee Lesort [Thales]

    Adrien Matricon [École Nationale Supérieure de Techniques Avancées, until Jun 2018]

    Vyshakh Palli Thaza [Renault]

    Remy Portelas [Inria, from Oct 2018]

    William Schueller [Univ de Bordeaux]

    Alexandr Ten [Inria, from Mar 2018]

    Hugo Caselles-Dupré [Softbank Robotics, from Jun 2018]

Technical staff

    Alexandre Pere [Inria]

    Didier Roy [Inria]

    Théo Segonds [Inria, until Oct 2018]

    Antonin Raffin [ENSTA, until Oct 2018]

    Simon Baudry [Inria, until Apr 2018]

    Stéphanie Noirpoudre [Inria, until June 2018]

    Cyril Sagonero [Inria, until May 2018]

Interns

    Josias Levi Alvares [Inria, from Feb 2018 until Aug 2018]

    Sandy Manolios [Inria, from Mar 2018 until Aug 2018]

    René Traoré [ENSTA, from Mar 2018 until Aug 2018]

    Hill Ashley [ENSTA, from Mar 2018 until Sept 2018]

Administrative Assistant

    Nathalie Robin [Inria]

Visiting Scientist

    Cynthia Liem [University of Delft, The Netherlands, until Apr 2018]

External Collaborator

    Helene Sauzeon [Univ. Bordeaux, Habilite]

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 (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.

- **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 [56] and neuroscience [81] 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 [69].

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" [149]. 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 [86], 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 [103] [153]. The approach of developmental robotics consists in importing and implementing concepts and mechanisms from developmental psychology [110], cognitive linguistics [68], and developmental cognitive neuroscience [91] 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 [60] [126], grounding [84], situatedness [47], self-organization [147] [129], enaction [151], and incremental learning [64].

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 [123]. 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 [110], 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" [51] [73]. 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 [56] [69] [71]. 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 [70] [88] [142]. Based on this, a number of researchers have began in the past few years to build computational implementation of intrinsic motivation [123] [124] [138] [55] [89] [106] [139]. 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 [124], [123], [125], [136]. 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 [66] [85] 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 [110]. 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 [50]. 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 [148], [57], motivated by the various mechanisms that allow humans to socially guide a robot [133]. 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 [148], [95], [107].

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 [115], [116]. We also work on developing perceptual capabilities by gradually segmenting the perceptual space and identifying objects and their characteristics through interaction with the user [104] and robots experiments [90]. 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 [99].

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

- PY Oudeyer was awarded the prize Inria of Académie des Sciences (category young researchers, http://www.academie-sciences.fr/fr/Laureats/laureats-2018-prix-inria.html)
- The Poppy Education ecosystem of educational robotics kits, associated technologies and educational community created by the Flowers team has been transferred to the newly created Poppy Station association (the creation process being coordinated by Didier Roy), gathering large scale national organizations including Ligue de l'enseignement, Hesam, IFE, EPFL, Arts et Métiers ParisTech, CESI, Le Cnam, Generation Robots, Pollen Robotics, Konex inc, see https://www.poppystation.org/
- PY Oudeyer co-authored with his collaborator Jacqueline Gottlieb (Columbia Univ., NY) a review article [22] in the high impact journal Nature Reviews Neuroscience, entitled "Towards a neuroscience of active sampling and curiosity", https://www.nature.com/articles/s41583-018-0078-0
- PY Oudeyer co-organized (with J. Gottlieb, A. Shankar and P. Zurn) the international conference "Curiosity: Emerging Sciences and Educational Innovations" at University of Pennsylvania, US, gathering researchers from multiple disciplines (neuroscience, psychology, artificial intelligence, HCI, robotics, philosophy, education) around the topic of curiosity, learning and education. https://www.sp2.upenn.edu/sp2-event/curiosity-emerging-sciences-and-educational-innovations.

6. New Software and Platforms

6.1. 3rd hand infrastructure

KEYWORDS: Interaction - Robotics - Infrastructure software - Framework - Robot Operating System (ROS)
FUNCTIONAL DESCRIPTION: The infrastructure is predicate-based to handle relational actions and covers perception (scene description generation, human actions recognition), decision making (teleoperated, scripted or learning from demonstrations), interaction with end users (GUI, voice, gestures) and parallel executions of robotic actions (hold, pick, grasp, bring, ...).

- Contact: Yoan Mollard
- URL: https://github.com/3rdHand-project/thr_infrastructure

6.2. Aversive++

FUNCTIONAL DESCRIPTION: Aversive++ is a C++ library that eases micro-controller programming. Its aim is to provide an interface simple enough to be able to create complex applications, and optimized enough to enable small micro-controllers to execute these applications. The other aspect of this library is to be multiplatform. Indeed, it is designed to provide the same API for a simulator (named SASIAE) and for AVR-based and ARM-based micro-controllers.

- Contact: Loïc Dauphin
- URL: https://github.com/AversivePlusPlus

6.3. DMP-BBO

Black-Box Optimization for Dynamic Movement Primitives
KEYWORD: -
FUNCTIONAL DESCRIPTION: The DMP-BBO Matlab library is a direct consequence of the insight that black-box optimization outperforms reinforcement learning when using policies represented as Dynamic Movement Primitives. It implements several variants of the PIBB algorithm for direct policy search. The dmp-bbo C++ library has been extended to include the "unified model for regression". The implementation of several of the function approximators have been made real-time compatible.

- Participant: Freek Stulp
- Partner: ENSTA
- Contact: Freek Stulp
- URL: https://github.com/stulp/dmpbbo

6.4. 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 [122], which has been extended to a more general family of autonomous exploration architectures by [1] and recently expressed as a compact and unified formalism [113]. The library is detailed in [114]. 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 [118]). 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 [122]. 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].
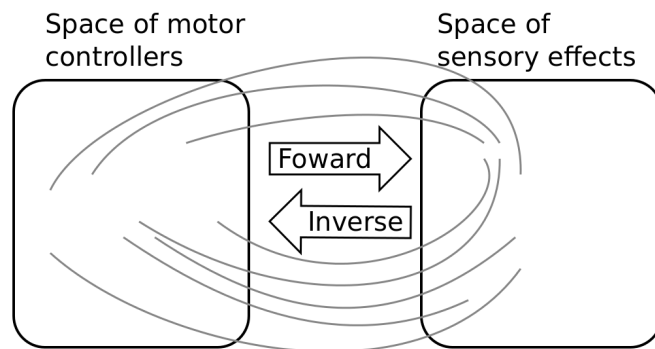
Figure 1. Complex parametrized policies involve high dimensional action and effect spaces. For the sake of visualization, the motor M and sensory S spaces are only 2D each in this example. The relationship between M and S is non-linear, dividing the sensorimotor space into regions of unequal stability: 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: Sébastien Forestier
- URL: https://github.com/flowersteam/explauto

6.5. HiPi Board

FUNCTIONAL DESCRIPTION: Hipi is a board to control robots on Raspberry Pi. It is an extension of the Pixl board with the following features:

- A DC/DC power converter from 12V (motor) to 5V (Raspberry Pi) at 3A.
- A stereo audio amplifier 3W.
- A MPU9250 central motion unit .
- A RS232 and a RS485 bus connected to the Raspberry Pi by SPI for driving MX and RX Dynamixel motor series.

This board will be integrated soon in the new head of the Poppy Humanoid and Poppy Torso.

Using the Raspberry Pi for every Poppy robots will simplify the hardware complexity (we maintain 4 types of embedded boards, with different Linux kernel and configurations) and improve the usage and installation of new robots.

- Contact: Theo Segonds
- URL: https://forum.poppy-project.org/t/poppy-1-1-hipi/2137

6.6. IKPy

Inverse Kinematics Python Library
FUNCTIONAL DESCRIPTION: IKPy is a Python Inverse Kinematics library, designed to be simple to use and extend. It provides Forward and Inverse kinematics functionality, bundled with helper tools such as 3D plotting of the kinematics chains. Being written entirely in Python, IKPy is lightweight and is based on numpy and scipy for fast optimization. IKPy is compatible with many robots, by automatically parsing URDF files. It also supports other (such as DH-parameters) and custom representations. Moreover, it provides a framework to easily implement new Inverse Kinematics strategies. Originally developed for the Poppy project, it can also be used as a standalone library.

- Contact: Pierre Manceron
- URL: https://github.com/Phylliade/ikpy

6.7. KERAS-QR

KERAS with Quick Reset

KEYWORDS: Library - Deep learning
- Participant: Florian Golemo
- Contact: Florian Golemo
- URL: https://github.com/fgolemo/keras

6.8. KidBreath

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

6.9. 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.

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.

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

6.10. Kidlearn: script for Kidbreath use

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.

Github link to explanation about it : https://github.com/flowersteam/kidlearn/.

- Contact: Benjamin Clement
- URL: https://github.com/flowersteam/kidlearn/

6.11. 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/

6.12. Kinect 2 Server

Kinect 2 server

KEYWORDS: Depth Perception - Speech recognition - Gesture recognition - Kinect

FUNCTIONAL DESCRIPTION: The server written in C# uses the Kinect SDK v2 to get the RGBD raw image, skeleton tracking information, recognized speech. It also uses the text-to-speech from Microsoft. Then it streams JSON data over the network using the Publisher/Subscriber pattern from the ZeroMQ network library. A Linux client has been written in Python but it can be written in any other language that is compatible with ZeroMQ. Features are controllable through a Graphical User Interface on Windows, or through the code from any Linux/Windows client. The clients can for instance enable features (speech recognition on, skeleton tracking off, . . .) and parameters (set new speech to recognize, change language, . . .) from remote.

- Contact: Yoan Mollard
- URL: https://github.com/baxter-flowers/kinect_2_server/

6.13. Multimodal

FUNCTIONAL DESCRIPTION: The python code provides a minimum set of tools and associated libraries to reproduce the experiments in [98] , together with the choreography datasets. The code is primarily intended for reproduction of the mulimodal learning experiment mentioned above. It has already been reused in several experimentations by other member of the team and is expected to play an important role in further collaborations. It is also expected that the public availability of the code encourages further experimentation by other scientists with data coming from other domains, thus increasing both the impact of the aforementioned publication and the knowledge on the algorithm behaviors.

- Participant: Olivier Mangin
- Contact: Olivier Mangin
- URL: https://github.com/omangin/multimodal

6.14. OptiTrack

FUNCTIONAL DESCRIPTION: This python library allows you to connect to an OptiTrack from NaturalPoint. This camera permits the tracking of 3D markers efficiently and robustly. With this library, you can connect to the Motive software used by the OptiTrack and retrieve the 3D position and orientation of all your tracked markers directly from python.

- Participant: Pierre Rouanet
- Contact: Pierre Rouanet
- URL: http://www.optitrack.com/

6.15. Pixl Board

FUNCTIONAL DESCRIPTION: Pixl is a tiny board used to create low cost robots based on Raspberry Pi board and Dynamixel XL-320 motors. This board has 2 main features:

- The power part, allowing the user to plug a 7.5V AC/DC converter or a battery directly into the Pixl. This power is distributed to all XL320 motors and is converted to 5V for the Raspberry Pi board.
- The communication part, which converts full duplex to half duplex and vice-versa. The half duplex part switch between RX and TX automatically. Another connector allows the user to connect his XL320 network.

The board is used in the Poppy Ergo Jr robot.

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

6.16. Poppy

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/

6.17. Poppy Ergo Jr

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.
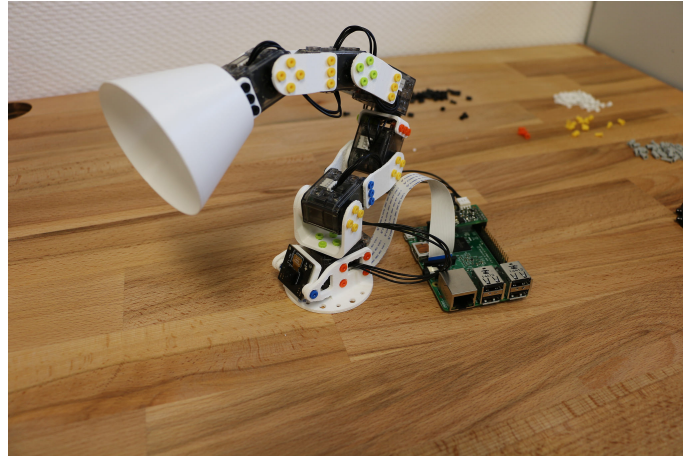
Figure 3. Poppy Ergo Jr, 6-DoFs arm robot for education
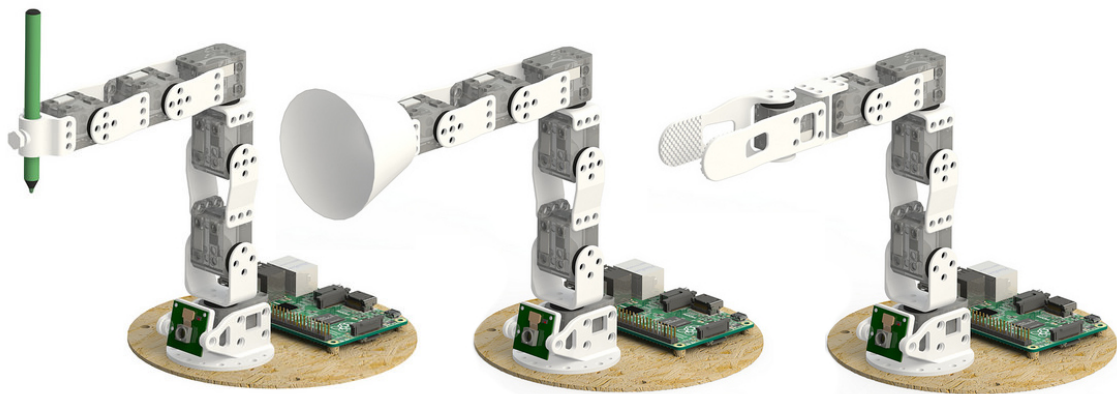
Figure 4. The available Ergo Jr tools: a pen holder, a lampshade and a gripper

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

6.18. Poppy Ergo Jr Installer

FUNCTIONAL DESCRIPTION: An alternative way to install the Ergo Jr robot software is made available using containers.

Users can own their own operating system installation, then add the Ergo Jr required software in a sandboxed environment. This results in a non-intrusive installation on the host system.

Docker containers implementation were used, and image is hosted at Docker Hub.

- Contact: Damien Caselli
- URL: https://hub.docker.com/r/poppycommunity/ergo-jr/

6.19. Poppy Ergo Jr Simulator

FUNCTIONAL DESCRIPTION: Poppy Project, through Poppy Education, wants users to get used to robotics, even without owning a physical robot.

For that purpose, Poppy Project team created a dummy robot in Pypot that is meant to be used in conjunction with a consumer application. We choose to develop a web hosted application using a 3D engine (Threejs) to render the robot.

Our ambition is to have a completely standalone simulated robot with physics. Some prototypes were created to benchmark possible solutions.

- Contact: Damien Caselli
- URL: https://github.com/poppy-project/poppy-simu

6.20. ProMP

Probabilistic Movement Primitives
KEYWORDS: Interaction - Robotics - Probability - Motion model - Robot Operating System (ROS)
FUNCTIONAL DESCRIPTION: Joint-space primitives with a task-space constraint: The primitives are stored in joint-space but demonstrations are provided both in joint space and task space, context. Thanks to this context, task-space goals can be requested to these joint-space primitives. The benefit is that requesting a new task-space goal does not require to call an IK method which would return demonstrations-agnostic joint configurations.

Vocal interactive learning and clustering: This work includes an interactive learning aspect which allows to automatically cluster motor primitives based on the standard deviation of their demonstrations. A new primitive is created automatically if the provided demonstration is out of 2 standard deviation of the existing primitives, otherwise the demonstration is distributed to an existing one.

- Contact: Yoan Mollard
- URL: https://github.com/baxter-flowers/promplib

6.21. PyPot

SCIENTIFIC DESCRIPTION: Pypot is a framework developed to make it easy and fast to control custom robots based on Dynamixel motors. This framework provides different levels of abstraction corresponding to different types of use. Pypot can be used to:

- control Robotis motors through a USB2serial device,
- define the structure of a custom robot and control it through high-level commands,
- define primitives and easily combine them to create complex behavior.

Pypot is part of the Poppy project. It is the core library used by the Poppy robots. This abstraction layer allows to seamlessly switch from a given Poppy robot to another. It also provides a common set of tools, such as forward and inverse kinematics, simple computer vision, recording and replaying moves, or easy access to the autonomous exploration library Explauto.

To extend pypot application domains and connection to outside world, it also provides an HTTP API. On top of providing an easy way to connect to smart sensors or connected devices, it is notably used to connect to Snap!, a variant of the well-known Scratch visual programming language.
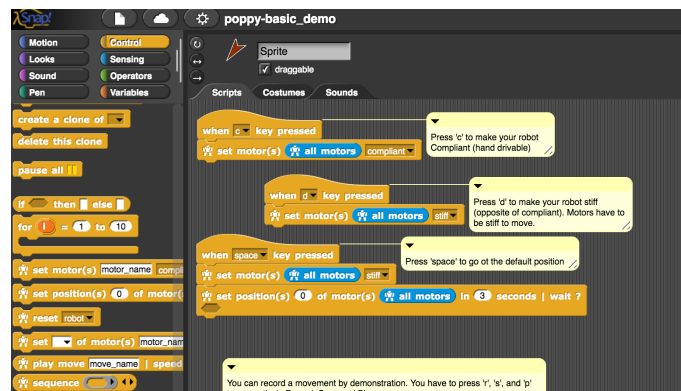
Figure 5. Example of using pypot to program a robot to reproduce a drawn shape

FUNCTIONAL DESCRIPTION: Pypot is entirely written in Python to allow for fast development, easy deployment and quick scripting by non-expert developers. It can also benefit from the scientific and machine learning libraries existing in Python. The serial communication is handled through the standard library and offers high performance (10ms sensorimotor loop) for common Poppy uses. It is cross-platform and has been tested on Linux, Windows and Mac OS.

Pypot is also compatible with the V-REP simulator. This allows the transparent switch from a real robot to its simulated equivalent with a single code base.

Finally, it has been developed to be easily and quickly extended for other types of motors and sensors.

It works with Python 2.7 or Python 3.3 or later, and has also been adapted to the Raspberry Pi board.

Pypot has been connected to Snap!, a variant of the famous Scratch visual language, developed to teach computer science to children. It is based on a drag-and-drop blocks interface to write scripts by assembling those blocks.

Thanks to the Snap! HTTP block, a connection can be made to pypot allowing users to directly control robots through their visual interfaces. A set of dedicated Snap! blocks have been designed, such as *set motor position* or *get motor temperature*. Thanks to the Snap! HTTP block, users can control robots through this visual interfaces connecting to Pypot. A set of dedicated Snap! blocks has been designed, such as *set motor position* or *get motor temperature*.

Snap! is also used as a tool to program the robot by demonstration. Using the *record* and *play* blocks, users can easily trigger kinesthetic recording of the whole robot or only a specific subpart, such as an arm. These records can then be played or "mixed" - either played in sequence or simultaneously - with other recordings to compose complex choreographies. The moves are encoded as a model of mixture of gaussians (GMM) which allows the definition of clean mathematical operators for combining them.

Figure 6. Using Snap! to program a robot by demonstration and create complex choreographies

This recording tool has been developed and used in collaboration with artists who show interest in the concept of robotic moves.
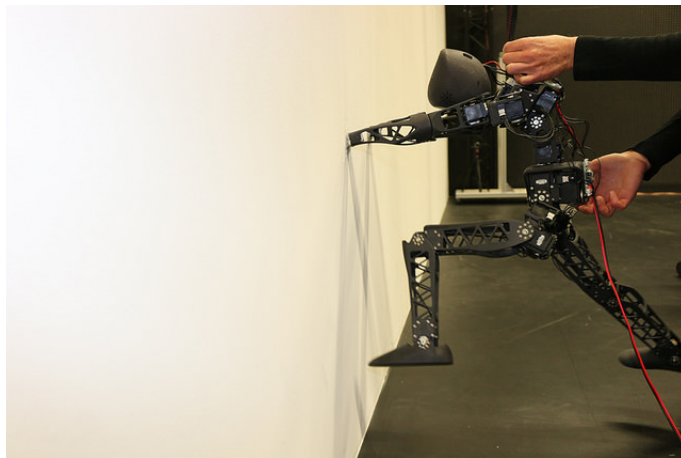
Figure 7. Artistic project exploring the concept of robotic move.

- Participants: Damien Caselli, Matthieu Lapeyre, Pierre Rouanet, Steve Nguyen and Theo Segonds
- Contact: Theo Segonds
- URL: https://github.com/poppy-project/pypot

6.22. PyQMC

Python library for Quasi-Metric Control

FUNCTIONAL DESCRIPTION: PyQMC is a python library implementing the control method described in http://dx.doi.org/10.1371/journal.pone.0083411 It allows to solve discrete markovian decision processes by computing a Quasi-Metric on the state space. This model based method has the advantage to be goal independant and thus can produce a policy for any goal with relatively few recomputation. New addition to this method is the possibility of online learning of the transition model and the Quasi-Metric.

- Participant: Steve Nguyen
- Contact: Steve Nguyen
- URL: https://github.com/SteveNguyen/pyqmc

6.23. ROS Optitrack Publisher

KEYWORDS: Target tracking - Robot Operating System (ROS)
FUNCTIONAL DESCRIPTION: This package allows to publish optitrack markers declared as rigid bodies as TF transforms. Data is gathered through the embedded VRPN server of Motive/Arena. Only rigid bodies are requested to the server, thus single points in 2D/3D are ignored. VRPN server can be enable in View > Data streaming in Motive.

- Contact: Yoan Mollard
- URL: https://github.com/baxter-flowers/optitrack_publisher

6.24. ThifloNet

KEYWORDS: Deep learning - Policy Learning
SCIENTIFIC DESCRIPTION: We created a software architecture that combines a state-of-the-art computer vision system with a policy learning framework. This system is able to perceive a visual scene, given by a still image, extract facts ("predicates"), and propose an optimal action to achieve a given goal. Both systems are chained into a pipeline that is trained by presenting images and demonstrating an optimal action. By providing this information, both the predicate recognition model and the policy learning model are updated.

Our architecture is based on the recent works of Lerer, A., Gross, S., & Fergus, R., 2016 ("Learning Physical Intuition of Block Towers by Example"). They created a large network able to identify physical properties of stacked blocks. Analogously our vision system utilizes the same network layout (without the image prediction auxiliary output), with an added output layer for predicates, based on the expected number and arity of predicates. The vision subsystem is not trained with a common cross-entropy or MSE loss function, but instead receives its loss form the policy learning subsystem. The policy learning module calculates the loss as optimal combination of predicates for the given expert action.

By using this combination of systems, the architecture as a whole requires significantly fewer data samples than other systems (which exclusively utilize neural networks). This makes the approach more feasible to real-life applciation with actual live demonstration.
FUNCTIONAL DESCRIPTION: The neural network consists of ResNet-50 (the currently best-performing computer vision system), with 50 layers, 2 layers for converting the output of ResNet to predicates and a varying amount of output neurons, corresponding to the estimated number of n-arity predicates. The network was pretrained on the ImageNet dataset. The policy learning module incorporates the ACE tree learning tool and a wrapper in Prolog.

Our example domain consists of 2-4 cubes colored in red, blue, green, and yellow and randomly stacked on top of each other in a virtual 3D environment. The dataset used for training and testing contains a total of 30000 elements, each with an image of the scene, the correct predicates, a list of blocks that are present and the corresponding expert action, that would lead to stacking the blocks to a tower.

- Participants: Florian Golemo, Manuel Lopes and Thibaut Munzer
- Contact: Florian Golemo

6.25. S-RL Toolbox

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

KEYWORDS: Machine learning - Robotics

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

6.26. Sets

KEYWORD: Data structures

FUNCTIONAL DESCRIPTION: The sets library allow to manipulate and operate on sets.

Those can be simple sets: + Empty: + Singleton: 'a' + Finite: 'a', 'b' + Integer subset: |[1, 10]| + Reals subset: [1, inf[Or, they can be cartesian product of sets: + 'a', 'b' x |[0, 10]| x]-inf, inf[Or, they can be incomplete unions of sets: + 'a','b' x |[0, 5]| U 'c' x -|[0, 10]|

In particular, every set is hashable. This operation is non-trivial in the case of an incomplete union (equivalent to an orthogonal polyhedron). An extreme vertices representations, corresponding to the state-of-the-art, is used to implement it.

Various operations are available: + Product (Cartesian) + Measure + Partition + Belonging test + Subset test (Proper) + Equality test + Union + Intersection + Exclusion

- Contact: Alexandre Pere

6.27. 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

6.28. 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

7. New Results

7.1. Computational Models Of Human Learning and Development

7.1.1. *Computational Models Of Information-Seeking and Curiosity-Driven Learning in Humans and Animals*

Participants: Pierre-Yves Oudeyer [correspondant], William Schueller, 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 [100], [125], [135]. Therefore we hypothesize that, rather than using a single optimization process as it has been the case in most previous theoretical work [82], 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 [123] [54], [111]. 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 [98], [134], and rats show individual variation in learning styles and novelty seeking behaviors [74], 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 [130], [146], [48], [150]. 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 [98], [93] 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 performance-based or LP-based exploration in two contexts in which they could freely choose to learn about 4 competing tasks. Participants (n = 505, recruited via Amazon Mechanical Turk) were tested on a paradigm in which they could freely choose to engage with one of four different classification tasks. 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 currently analyzing the results and working on a computational models of the underlying cognitive and motivational mechanisms.

The central question going into the study was, how active learners become interested in specific learning exercises: how do they decide which task to be interested in – i.e., allocate "study time" - given that the underlying rewards or patterns are sparse and unknown? Using a family of statistical (multinomial logit), subjective-utility-based models of discrete choice behavior [109] we performed an exploratory all-subsets model selection exercise [61] to see if we can identify important and/or interesting variables that could reliably predict task choices. The initial set of variables included, among other things, various performance-based competence heuristics (e.g. current hit rate, likelihood of current hit rate). Model selection and multimodel inference pointed to a handful of variables that had *relatively* high influence on task choices (including the likelihood of current hit rate and relative amount of time spent on a task), but their absolute effects were small, leaving most of the variation in task choices unexplained. This exercise also pointed out the potential limitations of our approach, either in operationalization of competence as a purely performance-based statistic, or in the potential lack of behavioral constraints in design of the experiment (participants may have been basing their choices on unanticipated variables). This latter limitation is tricky, since we are interested in exploratory behavior in unconstrained settings. What could have alleviated this challenge is a more diverse set of measurements that could include, for example, online records of participants' subjective feelings of interest, competence, liking, or learning potential. At this point, results concerning the LP hypothesis still have not revealed themselves, but we have gained valuable clues on how to find them. The next important step is to use cognitive models with transparent knowledge representations (e.g. Bayesian classifiers or neural networks) as an alternative way to operationalize subjective feelings of competence. The cognitive modeling approach emphasizes the idealistic assumptions made about the mind and examines their implicated behavioral outcomes. By doing that, cognitive models of learning and subjective competence can show whether our assumptions about the cognitive processes involved lead to the same behavioral patterns as the ones humans actually produce.

Although, the results of the Active Categorization study are still inconclusive, we found some interesting interim behavioral trends that are worth replicating and investigating. Participants showed preference for tasks of what we intended to be extreme complexity (i.e. too easy or too difficult) by spending more time on them (see figure 8). The group that was instructed to explore freely allocated their time more evenly, but showed a slight preference towards the easiest task where classification was based on a single dimension. The group that was instructed to try to maximize their learning during the experiment and expected a test at the end spent the majority of their time on the hardest (in fact, impossible) task to learn, where class assignment was independent of the two dimensions of variability. This suggests that active sampling strategies are subjected to extraneous constraints, and specifically, that some constraints may lead to inefficient exploration. It also potentially challenges the LP hypothesis, but it is to early to come to any strong conclusions about that, since we do not know how difficulty of the tasks was ranked subjectively by the participants.

Another puzzling observation comes from self-reported meta-cognitive judgments about the tasks. Figure 9 shows the average (min-max normalized) ratings of future learning potential and sensing the existence of a rule of each task. It is not clear why the learning potential for the hardest task (R) was reported to be high, despite the fact that it was believed to have no rule for classification. On the one hand, it is possible that while participants had not discovered the rule yet, they might have still believed there was a rule to be discovered. On the other hand, participants could really believe tha there was no rule to be discovered, but were not confident in that judgment, so high learning potential relates not to classification per se, but to discovering an interesting aspect of the task itself. There are other competing interpretations. Again, these observations compel us to better understand the contents of knowledge and knowledge-dependent processes used in the task, which we hope to achieve by applying and examining computational cognitive models of learning and meta-cognition.

7.1.2. *Computational Models Of Tool Use and Speech Development: the Roles of Active Learning, Curiosity and Self-Organization*

Participants: Pierre-Yves Oudeyer [correspondant], Sébastien Forestier, Rémy Portelas.
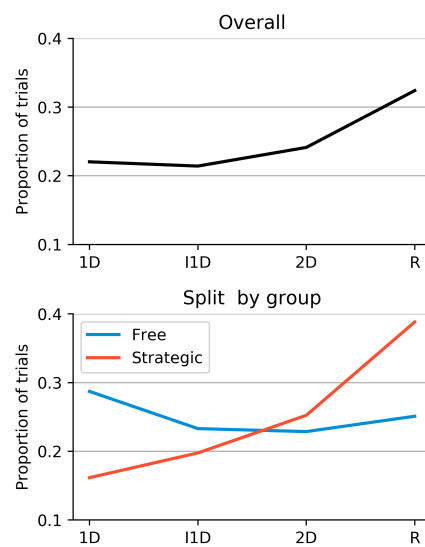
Figure 8. Proportion of trials on each task (1D, I1D, 2D, and R). 1D was the task were categorization was determined by a single variable dimension. In I1D (ignore 1D), the stimuli varied across 2 dimensions, but only one determined the stimulus category. In 2D, there were 2 variable dimensions and both 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.
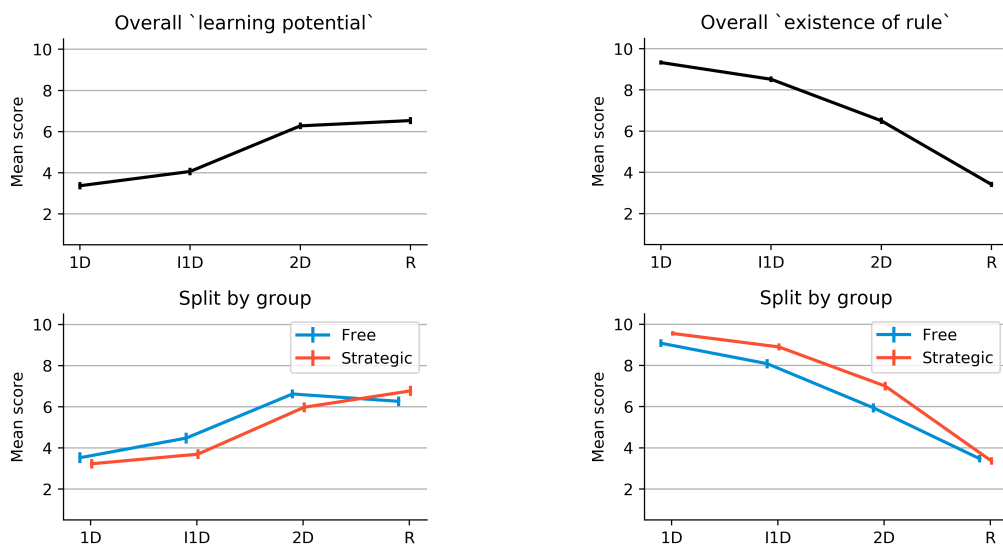
Figure 9. Average self-reported ratings of learning potential and existence of a rule for each task (1D, I1D, 2D, R; see figure 8 for disambiguation). The learning potential ratings ("Rate each monster family based on how much more you think you could learn if you had more time to play with it") were reported on a 10-point scale ([1] Definitely Could Not Learn More – [10] Definitely Could Learn More). The existence of rule ratings ("Rate each monster family based on how likely you think it had a rule for food preferences") were similarly reported ([1] Definitely No Rule – [10] Definitely a Rule. The error bars represent standard errors. The top plots show data aggregated across experimental groups, shown separately in the bottom plots.

7.1.2.1. Modeling Speech and Tool Use Development in Infants

A scientific challenge in developmental and social robotics is to model how autonomous organisms can develop and learn open repertoires of skills in high-dimensional sensorimotor spaces, given limited resources of time and energy. This challenge is important both from the fundamental and application perspectives. First, recent work in robotic modeling of development has shown that it could make decisive contributions to improve our understanding of development in human children, within cognitive sciences [82]. Second, these models are key for enabling future robots to learn new skills through lifelong natural interaction with human users, for example in assistive robotics [127].

In recent years, two strands of work have shown significant advances in the scientific community. On the one hand, algorithmic models of active learning and imitation learning combined with adequately designed properties of robotic bodies have allowed robots to learn how to control an initially unknown high-dimensional body (for example locomotion with a soft material body [53]). On the other hand, other algorithmic models have shown how several social learning mechanisms could allow robots to acquire elements of speech and language [62], allowing them to interact with humans. Yet, these two strands of models have so far mostly remained disconnected, where models of sensorimotor learning were too "low-level" to reach capabilities for language, and models of language acquisition assumed strong language specific machinery limiting their flexibility. Preliminary work has been showing that strong connections are underlying mechanisms of hierarchical sensorimotor learning, artificial curiosity, and language acquisition [128].

Recent robotic modeling work in this direction has shown how mechanisms of active curiosity-driven learning could progressively self-organize developmental stages of increasing complexity in vocal skills sharing many properties with the vocal development of infants [112]. Interestingly, these mechanisms were shown to be exactly the same as those that can allow a robot to discover other parts of its body, and how to interact with external physical objects [122].

In such current models, the vocal agents do not associate sounds to meaning, and do not link vocal production to other forms of action. In other models of language acquisition, one assumes that vocal production is mastered, and hand code the meta-knowledge that sounds should be associated to referents or actions [62]. But understanding what kind of algorithmic mechanisms can explain the smooth transition between the learning of vocal sound production and their use as tools to affect the world is still largely an open question.

The goal of this work is to elaborate and study computational models of curiosity-driven learning that allow flexible learning of skill hierarchies, in particular for learning how to use tools and how to engage in social interaction, following those presented in [122], [53], [117], [112]. The aim is to make steps towards addressing the fundamental question of how speech communication is acquired through embodied interaction, and how it is linked to tool discovery and learning.

We take two approaches to study those questions. One approach is to develop robotic models of infant development by looking at the developmental psychology literature about tool use and speech and trying to implement and test the psychologists' hypotheses about the learning mechanisms underlying infant development. Our second approach is to directly collaborate with developmental psychologists to analyze together the data of their experiments and develop other experimental setup that are well suited to answering modeling questions about the underlying exploration and learning mechanisms. We thus started last year a collaboration with Lauriane Rat-Fischer, a developmental psychologist working on the emergence of tool use in the first years of human life (now in Université Paris-Nanterre). We are currently analyzing together the behaviour of 22 month old infants in a tool use task where the infants have to retrieve a toy put in the middle of a tube by inserting sticks into the tube and pushing the toy out. We are looking at the different actions of the infant with tools and toys but also its looking behaviour, towards the tool, toys or the experimenter, and we are studying the multiple goals and exploration strategies of the babies other than the salient goal that the experimenter is pushing the baby to solve (retrieving a toy inside a tube).

In our recent robotic modeling work, we showed that the Model Babbling learning architecture allows the development of tool use in a robotic setup, through several fundamental ideas. First, goal babbling is a powerful form of exploration to produce a diversity of effects by self-generating goals in a task space. Second, the

possible movements of each object define a task space in which to choose goals, and the different task spaces form an object-based representation that facilitates prediction and generalization. Also, cross-learning between tasks updates all skills while exploring one in particular. A novel insight was that early development of tool use could happen without a combinatorial action planning mechanism: modular goal babbling in itself allowed the emergence of nested tool use behaviors.

Last year we extended this architecture so that the agent can imitate caregiver's sounds in addition to exploring autonomously [78]. We hypothesized that these same algorithmic ingredients could allow a joint unified development of speech and tool use. Our learning agent is situated in a simulated environment where a vocal tract and a robotic arm are to be explored with the help of a caregiver. The environment is composed of three toys, one stick that can be used as a tool to move toys, and a caregiver moving around. The caregiver helps in two ways. If the agent touches a toy, the caregiver produces this toy's name, but otherwise produces a distractor word as if it was talking to another adult. If the agent produces a sound close to a toy's name, the caregiver moves this toy within agent reach (see Fig. 10).
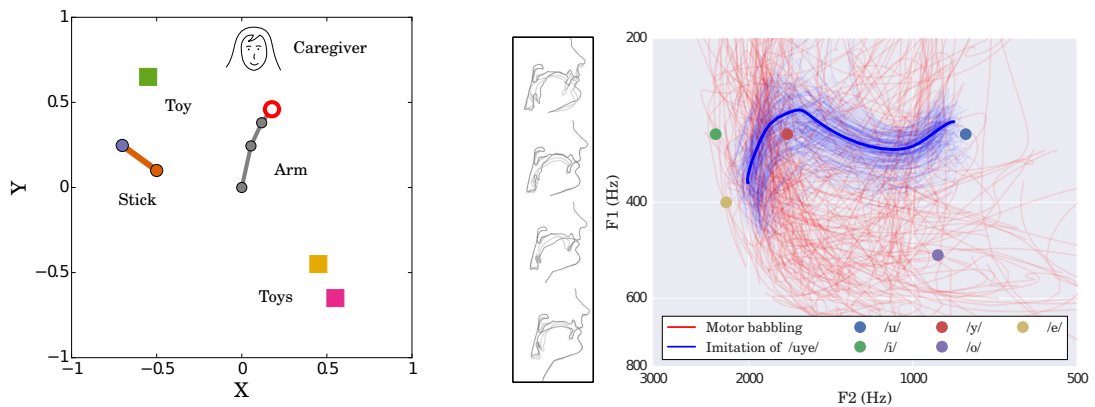
Figure 10. Agent's robotic and vocal environment. Left: Agent's 3 DOF arm, controlled with 21 parameters, grabs toys with its hand, or uses the stick to reach toys. Caregiver brings a toy within reach if the agent says its name. Right: Agent's vocal environment representing sounds as trajectories in the two first formants space. Agent's simulated vocal tract produces sounds given 28 parameters. When agent touches a toy, caregiver says toy's name. Some sounds corresponding to random parameters are plotted in red, and some sounds produced when imitating caregiver's /uye/ word in blue (best imitation in bold, error 0.3).

We showed that our learning architecture based on Model Babbling allows agents to learn how to 1) use the robotic arm to grab a toy or a stick, 2) use the stick as a tool to get a toy, 3) learn to produce toy names with the vocal tract, 4) use these vocal skills to get the caregiver to bring a specific toy within reach, and 5) choose the most relevant of those strategies to retrieve a toy that can be out-of-reach. Also, the grounded exploration of toys accelerates the learning of the production of accurate sounds for toy names once the caregiver is able to recognize them and react by bringing them within reach, with respect to distractor sounds without any meaning in the environment. Our model is the first to allow the study of the early development of tool use and speech in a unified framework. It predicts that infants learn to vocalize the name of toys in a natural play scenario faster than learning other words because they often choose goals related to those toys and engage caregiver's help by trying to vocalize those toys' names.

This year, we extended that model and we are currently studying on the one hand the impact of a partially contingent caregiver on agent's learning, and on the other hand the impact of attentional windows in agent's

sensory perception, to see if and how an attentional window that do not match the time structure of the interaction with the caregiver could impair learning.

We also transposed this experiment to a real robotic setting during a 6-months research internship dedicated to study how IMGEP approaches scale to a real-world robotic setup. This work is related to ongoing research on simulating human babies' curiosity-driven learning mechanisms, which objectives are to test psychologists' hypotheses on human learning and to leverage these models to increase efficiency in reinforcement learning applied to robotics. Previous experiments [78] showed in simulation that intrinsically motivated reinforcement learning could be successfully applied to the early developments of speech and tool-use. The main goal of this internship was to extend this work by designing a real-world robotic experiment using a Poppy-Torso robot and a Baxter. The contributions made during this internship were 1) The design of the Poppy-Baxter robotic playground (see figure [11] including the implementation of the communication architecture using ROS and the modeling of a 3D-printed toy, 2) Tuning of the experiment's parameters and learning process and 3) Analysis of the results in terms of exploration. Using this setup, we showed that the intrinsically motivated approach to model the early developments of speech and tool use developed in simulation can successfully scale to such a real-world experiment. Our curiosity-driven agents efficiently learned to vocalize the toy's name and to handle it in various and complex ways.
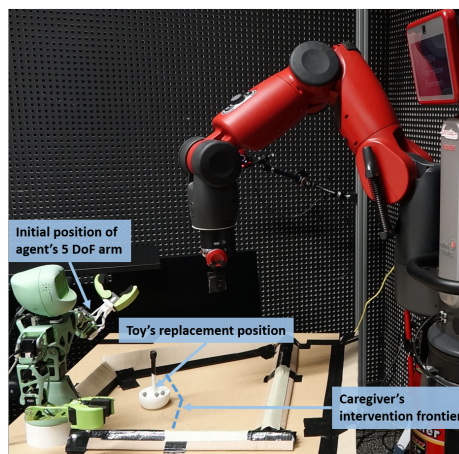
Figure 11. The POBAX Playground, in which the learning agent (Poppy-Torso) is able to interact with its arm and a simulated vocal tract. For each episode, if the agent touches the toy, the Caregiver (Baxter robot) says its name, otherwise the caregiver gives one of 3 distractor names. If the agent says the toys' name, the caregiver replaces it within the agent's arm-reach.

7.1.3. Models of Self-organization of lexical conventions: the role of Active Learning and Active Teaching in Naming Games

Participants: William Schueller [correspondant], Pierre-Yves Oudeyer.

How does language emerge, evolve and gets transmitted between individuals? What mechanisms underly the formation and evolution of linguistic conventions, and what are their dynamics? Computational linguistic studies have shown that local interactions within groups of individuals (e.g. humans or robots) can lead to self-organization of lexica associating semantic categories to words [143]. However, it still doesn't scale well to complex meaning spaces and a large number of possible word-meaning associations (or lexical conventions), suggesting high competition among those conventions.

In statistical machine learning and in developmental sciences, it has been argued that an active control of the complexity of learning situations can have a significant impact on the global dynamics of the learning process [82], [92], [101]. This approach has been mostly studied for single robotic agents learning sensorimotor affordances [123], [113]. However active learning might represent an evolutionary advantage for language formation at the population level as well [128], [145].

Naming Games are a computational framework, elaborated to simulate the self-organization of lexical conventions in the form of a multi-agent model [144]. Through repeated local interactions between random couples of agents (designated *speaker* and *hearer*), shared conventions emerge. Interactions consist of uttering a word – or an abstract signal – referring to a topic, and evaluating the success or failure of communication.

However, in existing works processes involved in these interactions are typically random choices, especially the choice of a communication topic.

The introduction of active learning algorithms in these models produces significant improvement of the convergence process towards a shared vocabulary, with the speaker [121], [140], [67] or the hearer [141] actively controlling vocabulary growth.
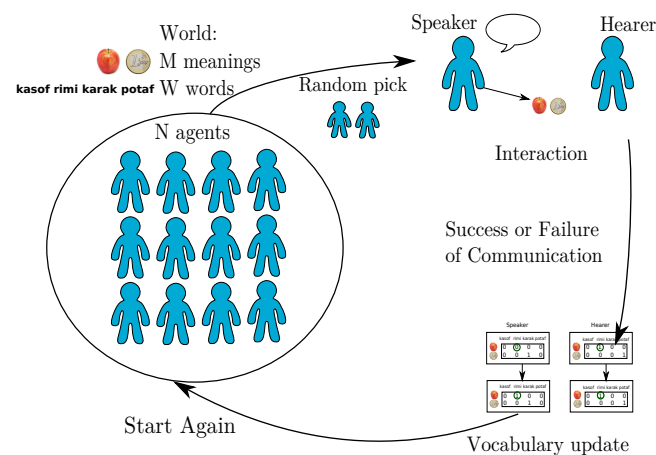
Figure 12. Illustration of the Naming Game model. Through repeated pairwise interactions, a population of individuals agrees on a shared lexicon mapping words to meanings.

7.1.3.1. Active topic choice strategies

Usually, the topic used in an interaction of the Naming Game is picked randomly. A first way of introducing active control of complexity growth is through the mechanism of topic choice: choosing it according to past memory. It allows each agent to balance reinforcement of known associations and invention of new ones, which can be seen as an exploitation vs. exploration problem. This can speed up convergence processes, and even lower significantly local and global complexity: for example in [140], [141], where heuristics based on the number of past successful interactions were used.

We defined new strategies in [31], [14] based on a maximization of an internal measure called LAPS, or Local Approximated Probability of Success. The derived strategies are called LAPSmax (exact measure but heuristical optimization algorithm) and Coherence (simplified measure but exact optimization).

Those strategies can speed up convergence the convergence process, but also diminish significantly the local complexity – i.e. the maximum number of distinct word-meaning association present in the population. See figure 13.
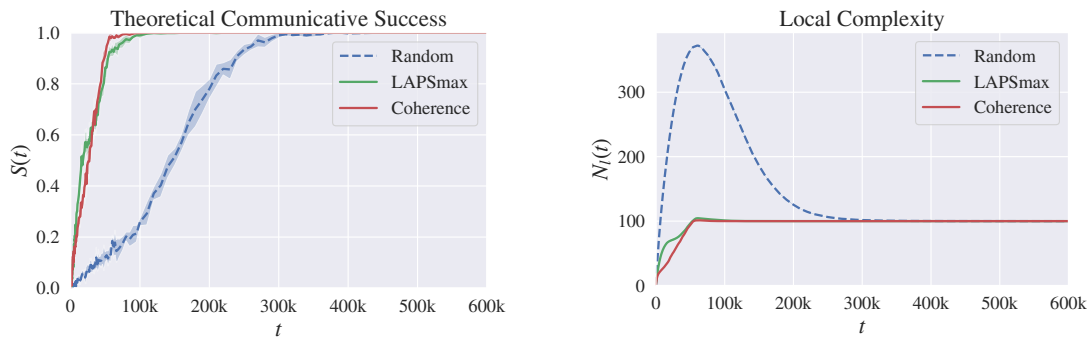
Figure 13. Evolution through time (number of interactions) of the measures of convergence (global probability of success) and global complexity (number of distinct word-meaning association present in the population) for simulations using Random Topic Choice vs. LAPSmax and Coherence Topic Choice strategies. The active topic choice strategy yields faster convergence, with less complexity. $N = 100$, $M = 100$, W is unbounded.

7.1.3.2. Statistical lower bounds and performance measures

We showed that the time needed to converge to a shared lexicon admits a statistical lower bound [14]:

$$t_{conv} \geq N \cdot M \cdot \ln M \tag{1}$$

Where M is the number of meanings and N the population size.

Using this lower bound, we can define performance measures (between 0 and 1, best value being 1) to classify strategies and compare behavior for different values of the parameters (like population size). We distinguish in particular performance measures for convergence time, convergence speed, and maximum lexicon size. Using this, we can show that LAPSmax and Coherence yield good performance measures, which are stable with population size (cf fig. 14), and significantly better than previous strategies.
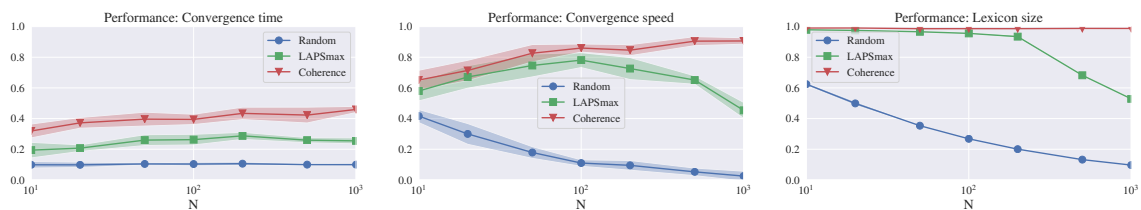
Figure 14. Performance measures for LAPSmax and Coherence strategies, compared with Random Topic Choice, for different values of N (population size). $M = 100$, W is unbounded.

7.1.3.3. Interactive application for collaborative creation of a language: Experimenting how humans actively negotiate new linguistic conventions

How do humans agree and negotiate linguistic conventions? This question is at the root of the domain of experimental semiotics [80], which is the context of our experiment/application. Typically, the experiments of

this field consist in making human subjects play a game where they have to learn how to interact/collaborate through a new unknown communication medium (such as abstract symbols). In recent years, such experiments allowed to see how new conventions could be formed and evolve in population of individuals, shading light on the origins and evolution of languages [94], [79].

We consider a version of the Naming Game [152], [102], focusing on the influence of active learning/teaching mechanisms on the global dynamics. In particular, agreement is reached sooner when agents actively choose the topic of each interaction [121], [140], [141].

Through this experiment, we confront existing topic choice algorithms to actual human behavior. Participants interact through the mediation of a controlled communication system – a web application – by choosing words to refer to objects. Similar experiments have been conducted in previous work to study the agreement dynamics on a name for a single picture [63]. Here, we make several pictures or interaction topics available, and quantify the extent to which participants actively choose topics in their interactions.

Global description: Each user interacts for about 3-4 min (<30 interactions) with a brand new population of 4 simulated agents. They take the role of one designated agent, and play the Naming Game as this agent. Each time they interact as speakers, they can select the topics of conversation from a set of 5 objects, and are offered 6 possible words to refer to them. Their choices influence the global emergence of a common lexical convention, reached when communications are successful. The goal is to maximize a score based on the number of successful interactions (among the 50 in total for each run). They can see a list of the past interactions, with chosen topic, chosen word, and whether the interaction was successful or not. This experiment allows us to directly measure if there is a bias in the choice of topics, compared to random choice, based on memory of past interactions. Performance can then be compared to existing topic choice algorithms [121], [140], [141] and [31].

First version: A first version was developed for the Kreyon Conference in Rome, in September 2017. The experiment was however too close to the theoretical model, and users were not motivated to play and finish the experiment. Provided feedback was often perceived as frustrating.

Second version: A second version was developed with the help of Sandy Manolios. This second version is more entertaining, includes a motivating context, a backstory, more adapted feedback, and a more user-friendly visual interface.

Results: Users in both experiments seem to actively control their rate of invention of new conventions, by selecting more often (than random) objects that they already have a word for. See figure 16.

Figure 15. Examples with the interface of respectively the first and the second version. Play the game here:
http://naming-game.space
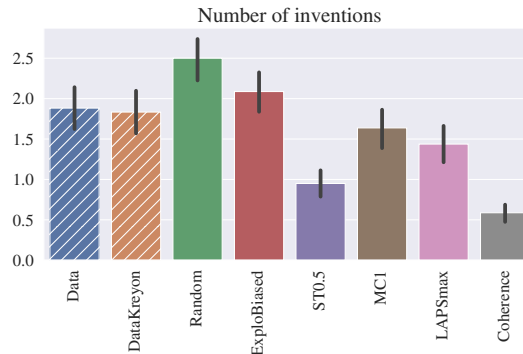
Figure 16. Average number of inventions for both experiments. DataKreyon: data collected at the Kreyon Conference in 2017 (first version); Data: data collected during summer 2018 (second version version); Other: simulated data for various strategies.

7.2. Autonomous Learning in Artificial Intelligence

7.2.1. *Intrinsically Motivated Goal Exploration and Multi-Task Reinforcement Learning*

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

7.2.1.1. Intrinsically Motivated Exploration of Spaces of Parameterized Goals and Application to Robot Tool Learning

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 [52].

Last year we developed a formal framework called "Unsupervised Multi-Goal Reinforcement Learning", as well as a formalization of intrinsically motivated goal exploration processes (IMGEPs), that is both more compact and more general than our previous models [76]. We experimented several implementations of these processes in a complex robotic setup with multiple objects (see Fig. 17), 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 17. 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. 18). 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.
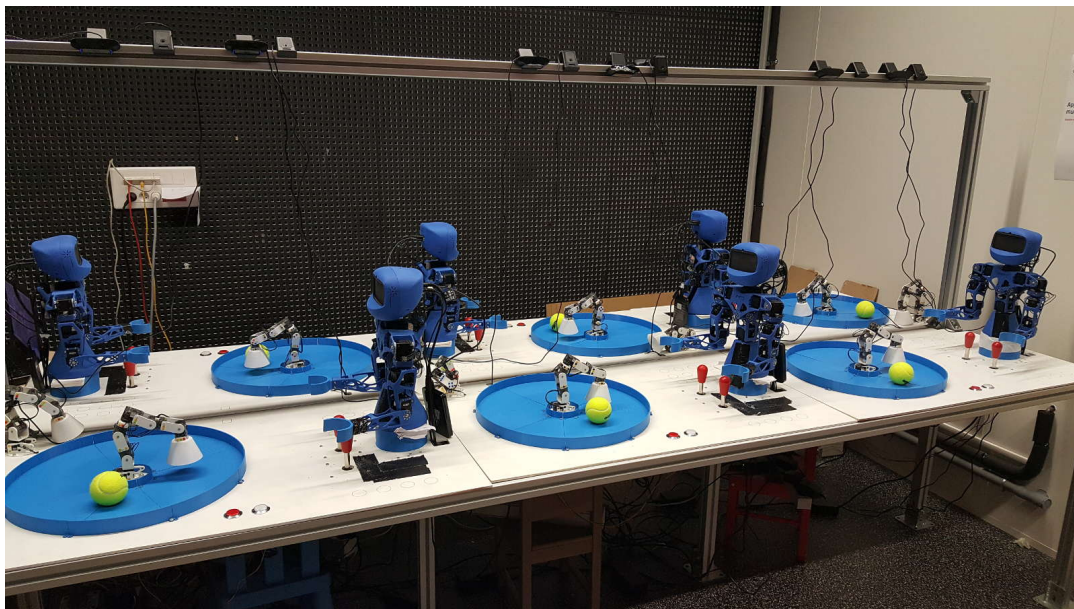
Figure 18. Platform of 6 robots with identical environment: joysticks, Poppy Ergo, ball in an arena, and a distractor. The central bar supports the 12 top cameras.

In 2018 we started 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 [19]). 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.
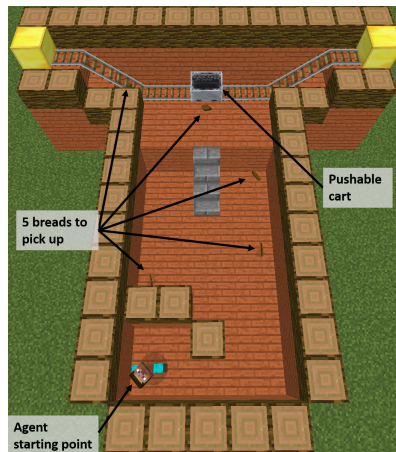
Figure 19. Malmo Mountain Cart. In this episodic environment the agent starts at the bottom left corner of the arena and is able to act on the environment using 2 continuous action commands: move and strafe. If the agent manages to get out of its starting area it may be able to collect items dispatched 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 Deep Learning of 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. 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 [38].

7.2.1.4. Curiosity Driven Exploration of Learned Disentangled Goal Spaces

As described in the previous paragraph, we have shown in [38] that one can use deep representation learning algorithms to learn an adequate goal space in simple environments. 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 [28].

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 ($\sigma = 0.1$)
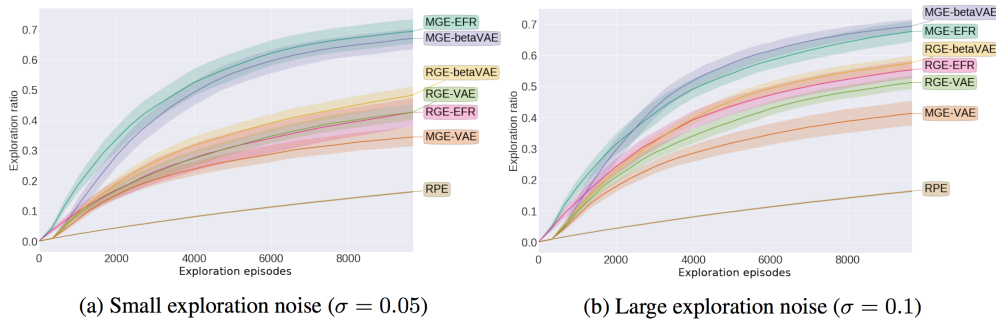
Figure 20. Exploration ratio during exploration for different exploration noises. Architectures with disentangled representations as a goal space (βVAE) explore more than those with entangled representations (VAE). Similarly modular architectures (MGE) explore more than flat architecture (RGE).

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 20). 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.

7.2.1.5. Combining deep reinforcement learning and curiosity-driven exploration

A major challenge of autonomous robot learning is to design efficient algorithms to learn sensorimotor skills in complex and high-dimensional continuous spaces. Deep reinforcement learning (RL) algorithms are natural candidates in this context, because they can be adapted to the problem of learning continuous control policies with low sample complexity. However, these algorithms, such as DDPG [97] suffer from exploration issues in the context of sparse or deceptive reward signals.

In this project, we investigate how to integrate deep reinforcement learning algorithms with curiosity-driven exploration methods. A key idea consists in decorrelating the exploration stage from the policy learning stage by using a memory structure used in deep RL called a replay buffer. Curiosity-driven exploration algorithms, also called Goal Exploration Processes (GEPs) are used in a first stage to efficiently explore the state and action space of the problem, and the corresponding data is stored into a replay buffer. Then a DDPG learns a control policy from the content of this replay buffer.

Last year, an internship has been dedicated to trying this methodology in practice but did not reach conclusions because of instability issues related to DDPG. The project was restarted this year, which led to an ICML publication [25]. We used an open-source implementations [72], and compared the combination GEP-PG (GEP + DDPG) to the traditional DDPG [97] as well as the former state-of-the art DDPG implementation endowed with parameter-based exploration [131]. Results were presented on the popular OpenAI Gym benchmarks Continuous Mountain Car (CMC) and Half-Cheetah (HC) [58], where state-of-the-art results were demonstrated.

7.2.1.6. Monolithic Intrinsically Motivated Multi-Goal and Multi-Task Reinforcement Learning
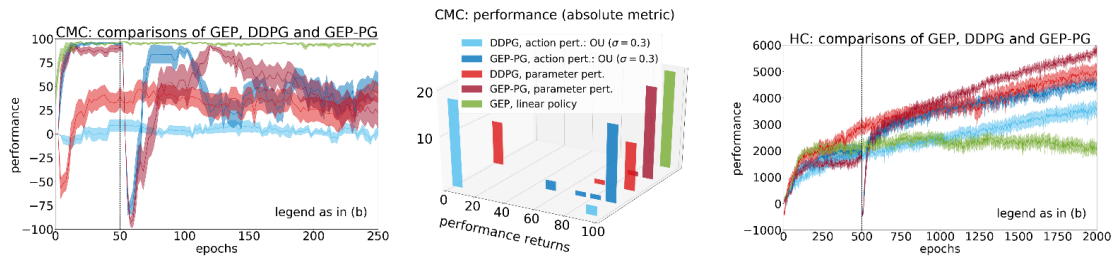
Figure 21. Left: learning curves on Continuous Mountain Car (20 seeds, mean +/- sem). Middle, best performances reached across learning on CMC. Right, learning curves on Half-Cheetah (20 seeds, mean +/- sem).

In this project we merged two families of algorithms. The first family is the Intrinsically Motivated Goal Exploration Processes (IMGEP) developed in the team (see [77] for a description of the framework). 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) [137] 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 [49] 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 [105] proposed to use UVFA to achieve multi-task learning with a discrete task-set.

In this project, we developed CURIOUS [43], 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 [58], see Figure 22. 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 23.

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

Figure 22. Custom multi-task and multi-goal environment to test the CURIOUS algorithm.
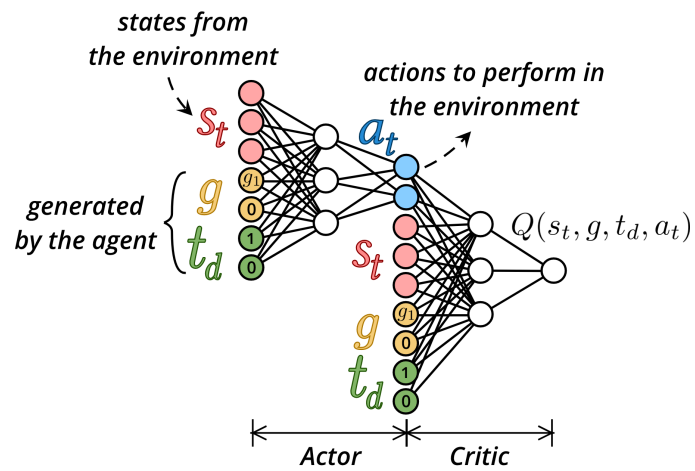
Figure 23. Architecture extended from Universal Value Function Approximators. In this example, the agent is targeting task T_1 among two tasks, each corresponding to a 1 dimension goal.

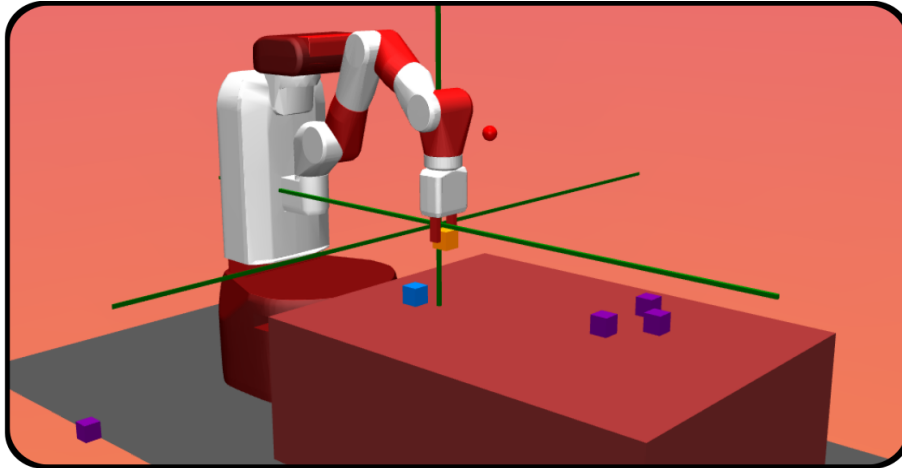progress is high. This helps for several reasons: 1) the agent does not focus on already learned tasks, as the corresponding learning progress is null, 2) the agent does not focus on impossible tasks for the same reason. The agent focuses more on tasks that are being learned (therefore maximizing learning 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 24.
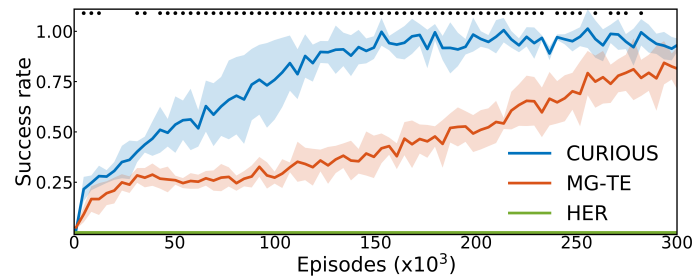
Figure 24. Comparison of CURIOUS to alternative algorithms.

7.2.2. *Transfer Learning from Simulated to Real World Robotic Setups with Neural-Augmented Simulators*

Participants: Florian Golemo [correspondant], Pierre-Yves Oudeyer.

This work was made in collaboration with Adrien Ali Taiga and Aaron Courville, and presented at CoRL 2018 [27]. Reinforcement learning with function approximation has demonstrated remarkable performance in recent years. Prominent examples include playing Atari games from raw pixels, learning complex policies for continuous control, or surpassing human performance on the game of Go. However most of these successes were achieved in non-physical environments (simulations, video games, etc.). Learning complex policies on physical systems remains an open challenge. Typical reinforcement learning methods require a lot of data which makes it unsuitable to learn a policy on a physical system like a robot, especially for dynamic tasks like throwing or catching a ball. One approach to this problem is to use simulation to learn control policies before applying them in the real world. This raises new problems as the discrepancies between simulation and real world environments ("reality gap") prevent policies trained in simulation from performing well when transfered to the real world. This is an instance of *domain adaption* where the input distribution of a model changes between training (in simulation) and testing (in real environment). The focus of this work is in settings where resetting the environment frequently in order to learn a policy directly in the real environment is highly impractical. In these settings the policy has to be learned entirely in simulation but is evaluated in the real environment, as *zero-shot transfer*.

In simulation there are differences in physical properties (like torques, link weights, noise, or friction) and in control of the agent, specifically joint control in robots. We propose to compensate for both of these source of issues with a generative model to bridge the gap between the source and target domain. By using data collected in the target domain through task-independent exploration we train our model to map state transitions from the source domain to state transition in the target domain. This allows us to improve the quality of our simulated robot by grounding its trajectories in realistic ones. With this learned transformation of simulated trajectories we are able to run an arbitrary RL algorithm on this augmented simulator and transfer the learned policy directly to the target task. We evaluated our approach in several OpenAI gym environments that were modified to allow for drastic torque and link length differences.

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

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

Intrinsically motivated goal exploration algorithms 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 in high-dimensional continuous state and action spaces (e.g. [75], [53]). 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 (e.g. [132]). 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 (e.g. [38]), but they have focused on goals represented as static target configurations of the environment in robotics sensorimotor spaces. This project studies how these new families of machine learning algorithms can be extended and used for automated discovery of behaviours of dynamical systems in physics/chemistry.

The work on the project started in November 2018 and is currently in the state of exploring potential systems and algorithms.

7.2.4. *Statistical Comparison of RL Algorithms.*

In this project [42], we review the statistical tests recommended by [87] to compare RL algorithms (T-test, bootstrap test, Kolmogorov-Smirnov). Kolmogorov-test is discarded as it only allows to compare distributions and not mean or median performance. We wrote a tutorial in the form of an arxiv paper [42] to present the use of the Welch's t-test and the bootstrap test to compare algorithm performances. In the last section of that paper, initial assumptions of the test are described. In particular, two limiting points are discussed. First, the computation of the required sample size to satisfy requirements in type-II error (false negative) is highly sensitive to the estimations of the empirical standard deviations of the algorithms performance distributions. It is showed that for small sample sizes (< 20) the empirical standard deviation of a Gaussian distribution is biased negatively in average. Using an empirical standard deviation smaller than the true one results in under-estimations of the type-II error and therefore of the required sample size to meet requirement on that error. Second we propose empirical estimations of the type-I error (false positive) as a function of the sample size for the Welch's t-test and the bootstrap test for a particular example (Half-Cheetah environment from OpenAI Gym [58] using DDPG algorithm [97]). It is showed that the type-I error is largely underestimated by the bootstrap test for small sample size (x6 for $n = 2$, x2 for $n = 5$, x1.5 for $n = 20$). The Welch's t-test offers a satisfying estimation of the type-I error for all sample size. In conclusion, the bootstrap test should not be used. The Welch's t-test should be used with a sufficient number of seeds to obtain a reasonable estimation of the standard deviation so as to get an accurate measure of type-II error (N>20).

7.3. Representation Learning

7.3.1. *State Representation Learning in the Context of Robotics*

Participants: David Filliat [correspondant], Natalia Diaz Rodriguez, Timothee Lesort, Antonin Raffin, René Traoré, Ashley Hill.

During the DREAM project, we participated in the development of a conceptual framework of open-ended lifelong learning [20] 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 [23], and we have 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 have developed and open sourced[1] the S-RL

---

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

toolbox [39] containing baseline algorithms, data generating environments, metrics and visualization tools for assessing SRL methods. The framework has been published at the NIPS workshop on Deep Reinforcement Learning 2018.

Mobile Navigation	Robotic Arm	Real Robot
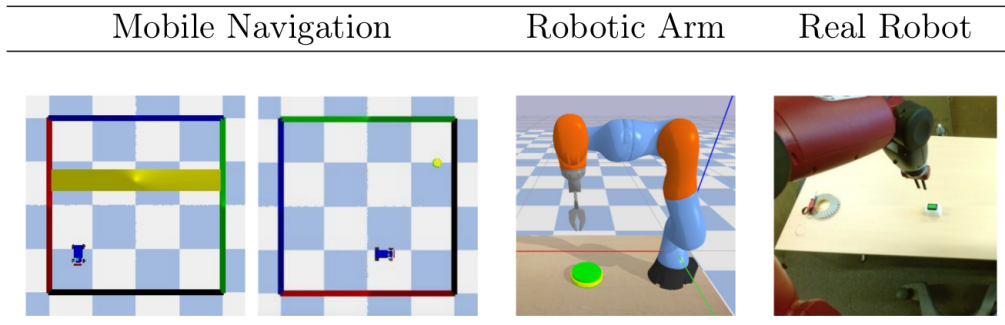
Figure 25. Environments and datasets for state representation learning.

The environments proposed in Fig. 25 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. 26 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).

Real-time SRL	Interactive scatter	Latent visualization
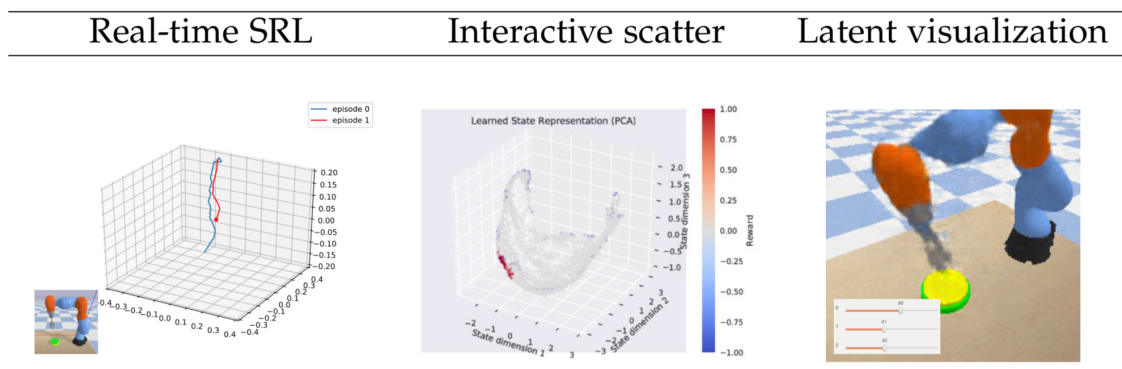
Figure 26. Visual tools for analysing SRL; Left: Live trajectory of the robot in the state space. Center: 3D scatter plot of a state space; clicking on any point displays the corresponding observation. Right: reconstruction of the point in the state space defined by the sliders.

We also proposed a new approach that consists in 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.
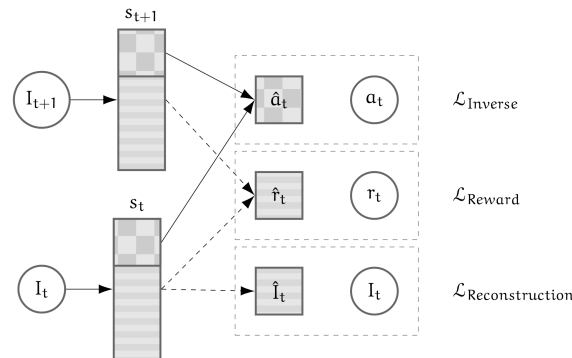
Figure 27. SRL Splits model: combines a reconstruction of an image I, a reward (r) prediction and an inverse dynamic models losses, using two splits of the state representation s. Arrows represent model learning and inference, dashed frames represent losses computation, rectangles are state representations, circles are real observed data, and squares are model predictions.

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. 27 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 [72], 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 [39] along with ground truth (GT).
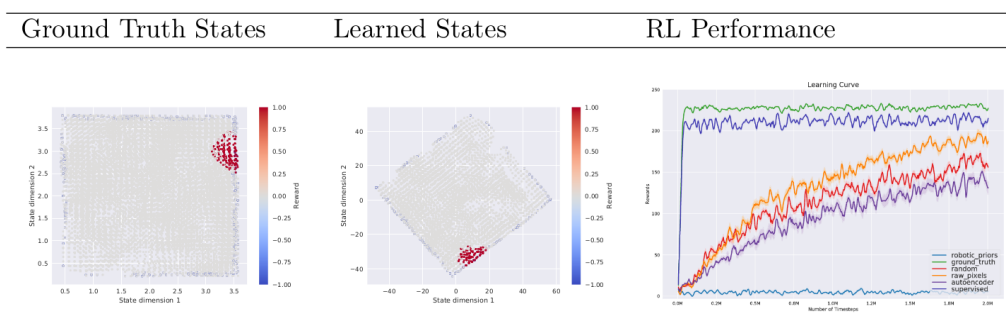
Figure 28. Ground truth states (left), states learned (Inverse and Forward) (center), and RL performance evaluation (PPO) (right) for different baselines in the mobile robot environment. Colour denotes the reward, red for positive, blue for negative and grey for null reward (left and center).

Table 28 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 [39].
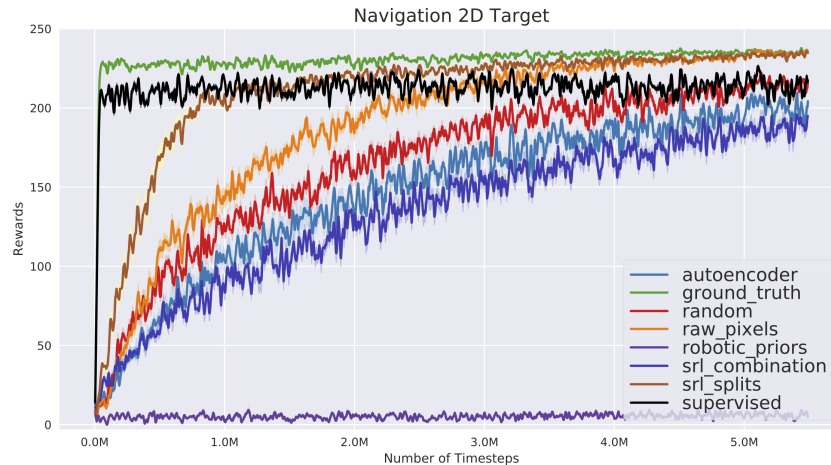
Figure 29. Performance (mean and standard error for 10 runs) for PPO algorithm for different state representations learned in Navigation 2D random target environment.

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. 29 shows that our unsupervised state representation learned with the split model even improves on the supervised case.

7.3.2. *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 [108], 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 [34] 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 30.

Generative models can also be evaluated from the perspective of Continual learning.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
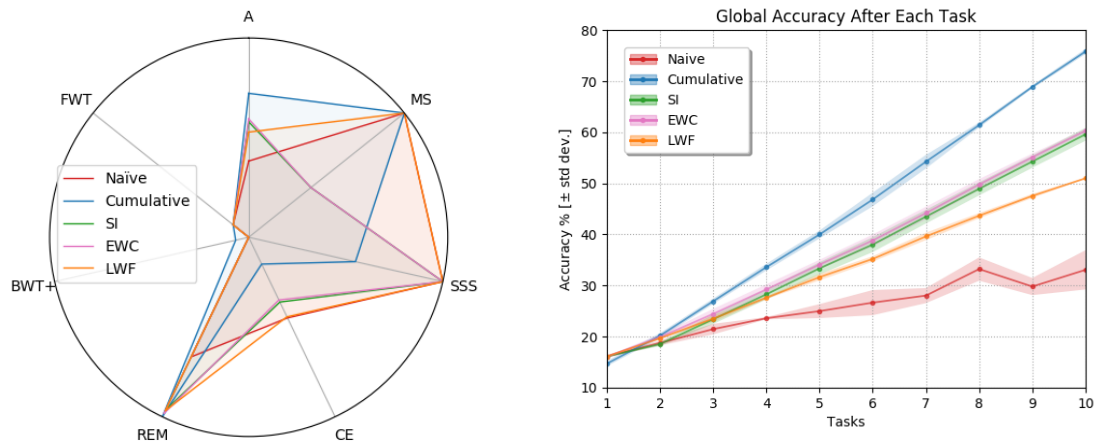
Figure 30. (left) Spider chart: CL metrics per strategy (larger area is better) and (right) Accuracy per CL strategy computed over the fixed test set.

sequential tasks on three commonly used benchmarks for Continual Learning (MNIST, Fashion MNIST and CIFAR10). We found (see Figure 32) 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 [33]. 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 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 32. 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 [32]. 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.

Real life examples of applications envisioned for continual learning include learning on the edge, real time embedded systems, and applications such as the project proposal at the NeurIPS workshop on AI for Good on intelligent drone swarms for search and rescue operations at sea [36].

7.3.3. *Knowledge engineering tools for neural-symbolic learning*

Participant: Natalia Díaz Rodríguez [correspondant].

This section includes diverse partners distributed world wide and is result of former stablished collaborations and includes work in the context of knowledge engineering for neural-symbolic learning and reasoning systems. In [35] we presented *Datil*, a tool for learning fuzzy ontology datatypes based on clustering techniques and fuzzyDL reasoner. Ontologies for modelling healthcare data aggregation as well as knowledge graphs for modelling influence in the fashion domain are concrete ontological proposals for knowledge
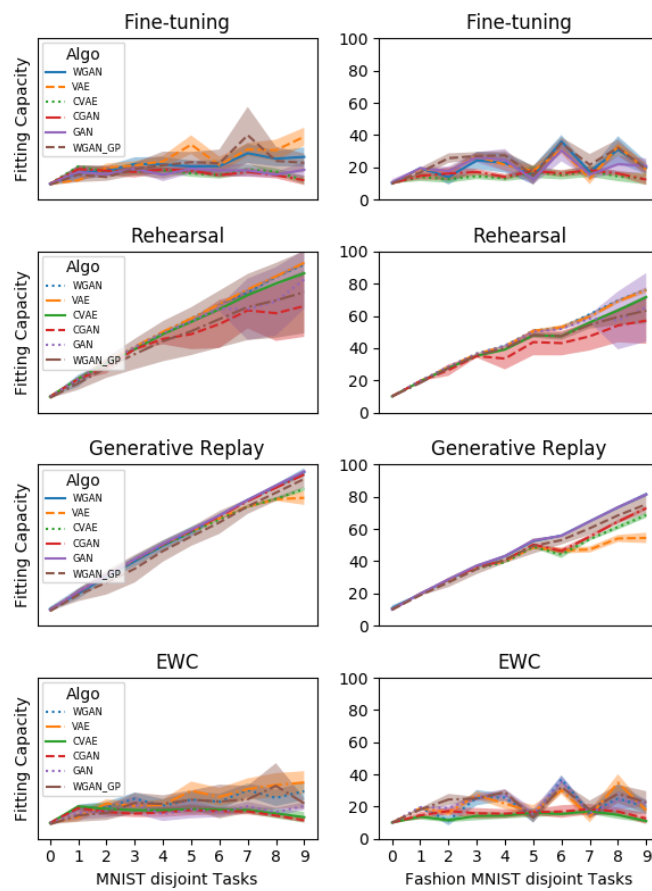
Figure 31. Means and standard deviations over 8 seeds of Fitting Capacity metric evaluation of VAE, CVAE, GAN, CGAN and WGAN. The four considered CL strategies are: Fine Tuning, Generative Replay, Rehearsal and EWC. The setting is 10 disjoint tasks on MNIST and Fashion MNIST.

Navigation task on environment 1   Navigation task on environment 2
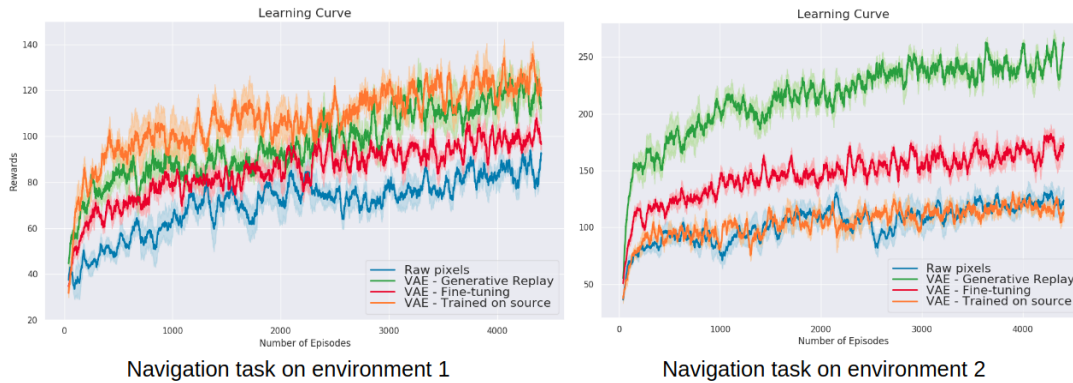
Figure 32. Mean reward and standard error over 5 runs of RL evaluation using PPO with different types of inputs. Fine-tuning and Generative Replay models are trained sequentially on the first and second environment, and then used to train a policy for both tasks. Generative Replay outperforms all other methods. It shows the need for continually learning features in State Representation Learning in settings where the environment changes.

modelling. The former looks at wearables data interoperability for Ambient Assisted Living application development, including concepts such as height, weight, locations, activities, activity levels, activity energy expenditure, heart rate, or stress levels, among others [41]. The second proposal, considers the intrinsic subjectivity needed to effectively model subjective domains such as fashion in recommendations systems. Subjective influence networks are proposed to better quantify influence and novelty in networks. A set of use cases this approach is intended to address is discussed, as well as possible classes of prediction questions and machine learning experiments that could be executed to validate or refute the model [40].

7.4. Applications in Robotic myoelectric prostheses

Participant: Pierre-Yves Oudeyer [correspondant].

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.

7.4.1. 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.

7.5. Applications in Educational Technologies

7.5.1. *Machine Learning for Adaptive Personalization in Intelligent Tutoring Systems*

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

7.5.1.1. 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 [65]. Systematic experiments were also presented with simulated students.

7.5.1.2. Kidlearn Experiments in 2018: Evaluating the impact of ZPDES and choice on learning efficiency and motivation

An experiment was held between mars 2018 and July 2018 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. The details of these new results, as well as the overall results of this project, were presented during the PhD defense of Benjamin Clement in decembre 2018.

7.5.1.3. Fostering Health Education With a Serious Game in Children With Asthma: Pilot Studies for Assessing Learning Efficacy and Automatized Learning Personalization

Coupled with Health Education programs, an e-learning platform—KidBreath—was participatory designed [19] and assessed in situ (Study 1) and was augmented and tested with an Intelligent Tutoring System (ITS) based on Multi-Armed Bandit Methods (Study 2). For each study, the impact of KidBreath practice was assessed in children with asthma in terms of pedagogical efficacy (knowledge of the illness), pedagogical efficiency (usability, type of motivation and level of interest elicited), and therapeutic effect (illness perception, system's expectation and judgement in disease self-management, child's implication in study). For the Study 1, asthma children aged 8 to 11 years used the tool at home without time pressure for 2 months according to a predefined learning sequence defined by the research team. Results supported pedagogical efficacy of KidBreath, with a significant increase of general knowledge about asthma after use. It also featured a greater learning gain for children knowing the least about the illness before use. Results on pedagogical efficiency revealed a great intrinsic motivation elicited by KidBreath showing a deep level of interest in the edutainment activities. Study 2 explored an augmented version of KidBreath with learning optimization algorithm (called ZPDES) after its use during 1 month. Pedagogical efficacy was less conclusive than Study 1 because less content was displayed due to algorithm parameters. However, the ITS-augmented KidBreath use showed a strong impact in pedagogical efficiency and therapeutic adherence features. Even if implementation improvements must be done in future works, this preliminary study highlighted the viability of our methods to design an ITS as serious game in health education context for all chronic diseases.

- Journée EdTech, Talence, mai 2018

7.5.2. *Poppy Education: Designing and Evaluating Educational Robotics Kits*

Participants: Pierre-Yves Oudeyer [correspondant], Didier Roy, Thibault Desprez, Théo Segonds, Stéphanie Noirpoudre.

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 33)
- 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 34), 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 35). 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 36).
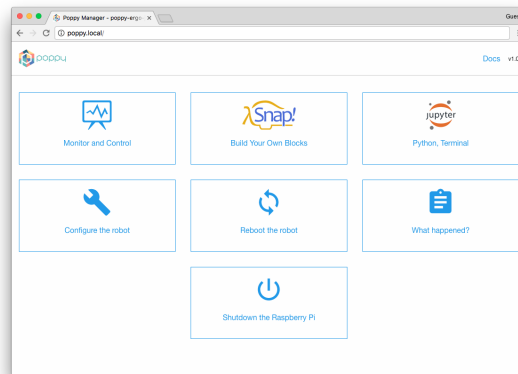
Figure 33. Home page on http://poppy.local
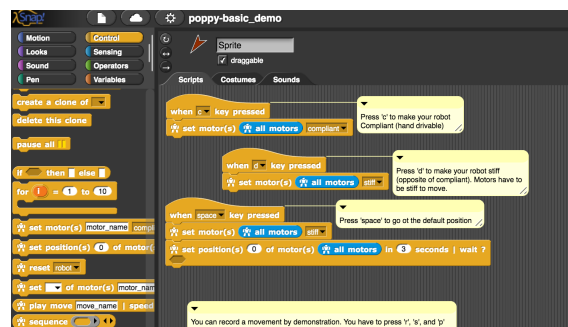
Figure 34. The visual programming system Snap
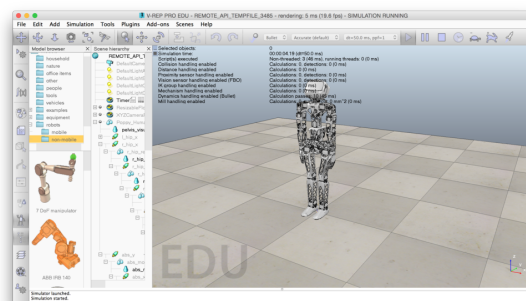
Figure 35. V-rep
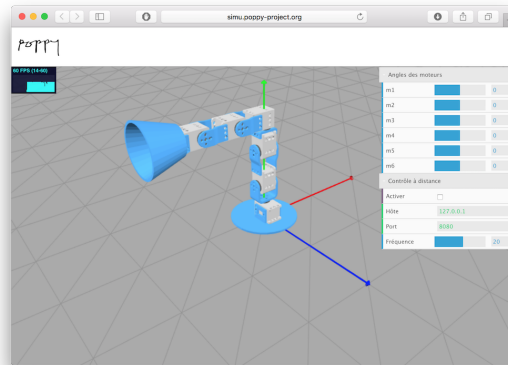
Figure 36. 3D viewer

7.5.2.1. 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 [120].

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 37. 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

7.5.2.2. 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 [119] 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 38. 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/

7.5.2.3. 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) [59] and The AttrakDiff [96]. 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) [26] published at the conference Didapro (Lausanne Feb, 2018). Figures 39 and 40 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

	-2	-1	0	1	2		x	s
[1. Je pense que je vais utiliser ce kit fréquemment]	Non					Oui	-0,15	1,26
*[2. Je trouve ce kit inutilement complexe]	Oui					Non	1,01	1,09
[3. Je pense que ce kit est facile à utiliser]	Non					Oui	0,80	1,00
*[4.Je pense que j'aurai besoin de l'aide d'un technicien pour être capable d'utiliser ce kit]	Oui					Non	0,80	1,18
[5.J'ai trouvé que les différentes fonctions de ce kit ont été bien intégrées]	Non					Oui	0,80	1,00
*[6. Je pense qu'il y a trop d'incohérences dans ce kit.]	Oui					Non	0,94	0,96
[7. J'imagine que la plupart des gens seraient capables d'apprendre à utiliser ce kit très rapidement.]	Non					Oui	1,11	1,09
*[8 J'ai trouvé ce kit très gênant à utiliser.]	Oui					Non	0,51	1,12
[9. Je me sentais très en confiance en utilisant ce kit.]	Non					Oui	1,17	0,95
*[10. J'ai besoin d'apprendre beaucoup de choses avant de pouvoir utiliser ce kit]	Oui					Non	0,56	1,11

Figure 39. Result of SUS survey

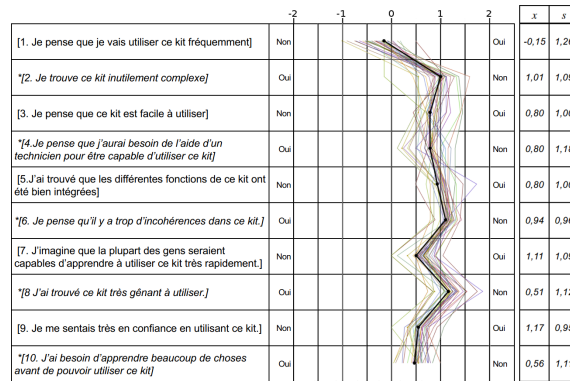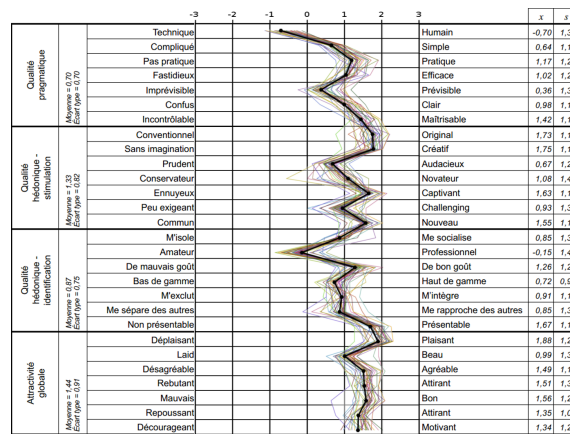|			-3	-2	-1	0	1	2	3		x	s
Qualité pragmatique	Moyenne = 0,70 Ecart type = 0,70	Technique								Humain	-0,70	1,38
		Compliqué								Simple	0,64	1,13
		Pas pratique								Pratique	1,17	1,24
		Fastidieux								Efficace	1,02	1,27
		Imprévisible								Prévisible	0,36	1,31
		Confus								Clair	0,98	1,18
		Incontrôlable								Maîtrisable	1,42	1,10
Qualité hédonique - stimulation	Moyenne = 1,33 Ecart type = 0,82	Conventionnel								Original	1,73	1,15
		Sans imagination								Créatif	1,75	1,14
		Prudent								Audacieux	0,67	1,28
		Conservateur								Novateur	1,08	1,41
		Ennuyeux								Captivant	1,63	1,14
		Peu exigeant								Challenging	0,93	1,33
		Commun								Nouveau	1,55	1,14
Qualité hédonique - identification	Moyenne = 0,87 Ecart type = 0,75	M'isole								Me socialise	0,85	1,30
		Amateur								Professionnel	-0,15	1,45
		De mauvais goût								De bon goût	1,26	1,23
		Bas de gamme								Haut de gamme	0,72	0,95
		M'exclut								M'intègre	0,91	1,19
		Me sépare des autres								Me rapproche des autres	0,85	1,39
		Non présentable								Présentable	1,67	1,16
Attractivité globale	Moyenne = 1,44 Ecart type = 0,91	Déplaisant								Plaisant	1,88	1,24
		Laid								Beau	0,99	1,31
		Désagréable								Agréable	1,49	1,17
		Rebutant								Attirant	1,51	1,30
		Mauvais								Bon	1,56	1,26
		Repoussant								Attirant	1,35	1,06
		Décourageant								Motivant	1,34	1,29

Figure 40. 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/

7.5.2.4. 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.

7.5.3. IniRobot: Educational Robotics in Primary Schools

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

Reminder : IniRobot (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.

7.5.3.1. 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.

7.5.3.2. Created pedagogical documents and resources

- The inirobot pedagogical kit [83]: 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 http://tice33.ac-bordeaux.fr/Ecolien/ASTEP/tabid/5953/language/fr-FR/Default.aspx

- 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.

7.5.3.3. Scientific mediation

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

7.5.3.4. Spread of Inirobot activities

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

7.5.3.5. 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.

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. *Adaptive device for disease awareness and treatment adherence of asthma in children*

Participants: Manuel Lopes [correspondant], Alexandra Delmas, Pierre-Yves Oudeyer.

Financing of the CIFRE PhD grant of Alexandra Delmas by Itwell with the goal of developing a tool for self-learning for patients to improve their compliance to treatment.

8.2.2. *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.

8.2.3. *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.

8.2.4. *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.

8.2.5. *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.

8.2.6. *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.

8.2.7. *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.

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.

9.1.1.1. Partner schools

In 2018, we have 36 partner schools (show Fig 41). 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.

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 41. List of partner schools

9.1.2. KidLearn and Region Aquitaine

KidLearn

Program: Région Aquitaine research grant

Duration: 2016 - 2018

Coordinator: PY Oudeyer and M Lopes, Inria Flowers

Partners: Inria Flowers

Funding: 140 keuros (PhD grant of Benjamin Clément)

We propose here a research project that aims at elaborating algorithms and software systems to help humans learn efficiently, at school, at home or at work, by adapting and personalizing sequences of learning activities to the particularities of each individual student. This project leverages recent innovative algorithmic models of human learning (curiosity in particular, developed as a result of ERC European project of the Flowers team), and combines it with state-of-the-art optimization algorithms and an original integration with existing expert knowledge (human teachers). Given a knowledge domain and a set of possible learning activities, it will be able to propose the right activity at the right time to maximize learning progress. It can be applied to many learning situations and potential users: children learning basic knowledge in schools and with the support of their teachers, older kids using educational software at home, of adults needing to acquire new skills through professional training ("formation professionnelle"). Because it combines innovations in computational sciences (machine learning and optimization) with theories of human cognition (theories of human learning and of education), this project is also implementing a strong cross-fertilization between technology and human sciences (SHS).

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.poppystation.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.

9.3. European Initiatives

9.3.1. *DREAM*

Title: Deferred Restructuring of Experience in Autonomous Machines

Programm: H2020

Duration: January 2015 - December 2018

Coordinator: UPMC

Partners:

Armines (ENSTA ParisTech)

Edimbourgh (Scotland)

University of A Coruna (Spain)

Vrije University Amsterdam (Holland)

Contact: David Filliat

Abstract: A holy grail in robotics and artificial intelligence is to design a machine that can accumulate adaptations on developmental time scales of months and years. From infancy through adult- hood, such a system must continually consolidate and bootstrap its knowledge, to ensure that the learned knowledge and skills are compositional, and organized into meaningful hierarchies. Consolidation of previous experience and knowledge appears to be one of the main purposes of sleep and dreams for humans, that serve to tidy the brain by removing excess information, to recombine concepts to improve information processing, and to consolidate memory. Our approach – Deferred Restructuring of Experience in Autonomous Machines (DREAM) – incorporates sleep and dream-like processes within a cognitive architecture. This enables an individual robot or groups of robots to consolidate their experience into more useful and generic formats, thus improving their future ability to learn and adapt. DREAM relies on Evo- lutionary Neurodynamic ensemble methods (Fernando et al, 2012 Frontiers in Comp Neuro; Bellas et al., IEEE-TAMD, 2010) as a unifying principle for discovery, optimization, re- structuring and consolidation of knowledge. This new paradigm will make the robot more autonomous in its acquisition, organization and use of knowledge and skills just as long as they comply with the satisfaction of pre-established basic motivations. DREAM will enable robots to cope with the complexity of being an information-processing entity in domains that are open-ended both in terms of space and time. It paves the way for a new generation of robots whose existence and purpose goes far beyond the mere execution of dull tasks. http://www.robotsthatdream.eu

9.3.2. Collaborations in European Programs, except FP7 & H2020

9.3.2.1. 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 GOT-TLIEB

Start year: 2016

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

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 an article accepted at CHI 2019.

9.4.1.3. Informal International Partners

Pierre-Yves Oudeyer and Didier Roy have create 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.

Florian Golemo and PY Oudeyer have had an active collaboration with Aaron Courville from MILA Montreal to work on the IGLU project together.

William Schueller and PY Oudeyer continued to collaborate with Vittorio Loreto (CNR Rome and Sony CSL Paris).

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.

9.5. International Research Visitors

9.5.1. Visits of International Scientists

- Bart de Boer, VUB Brussels (Dec 2018)
- Dan Dediu, Univ. Lyon (Dec 2018)
- Kenny Smith, Univ. Edinburgh (Dec 2018)
- Jochen Triesch, Univ. Frankfurt (Nov 2018)
- Vincenzo Lomonaco, University of Bologna, Italy (Aug. 2018)

9.5.2. Internships

- Ashley Hill, Univ. Paris-Sud, Paris Saclay.
- René Traoré, UPMC, Sorbonne Universite, Paris.
- Josias Lévi Alvarès, ENSC, Bordeaux.
- Sandy Manolios, Univ. de Lyon, Lyon.
- Rémy Portelas, UPMC, Paris.
- Chuan Qin, ENSTA, Paris Saclay.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. General Chair, Scientific Chair

- PY. Oudeyer has been general co-chair (with J. Gottlieb, A. Shankar and P. Zurn) of the international conference "Curiosity: Emerging Sciences and Educational Innovations" at University of Pennsylvania, US, gathering researchers from multiple disciplines (neuroscience, psychology, artificial intelligence, HCI, robotics, philosophy, education) around the topic of curiosity, learning and education. https://www.sp2.upenn.edu/sp2-event/curiosity-emerging-sciences-and-educational-innovations.
- PY Oudeyer has been vice-chair of the IEEE CIS Technical Committee on Cognitive and Developmental Systems.

10.1.2. Scientific Events Selection

10.1.2.1. Conference Reviewer

- N. Diaz Rodriguez was reviewer for ICSC18 (International Conf. on Semantic Computing) and Artificial Intelligence and Knowledge Engineering Conference (IEEE AIKE 2018)
- O. Sigaud was reviewer for NIPS, ICLR and ICML 2018

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards
- PY. Oudeyer was associate editor of IEEE Transactions on CDS and Frontiers in Neurorobotics.

10.1.3.2. Reviewer - Reviewing Activities
- David Filliat was reviewer for Frontiers in Robotics and AI
- N. Diaz Rodriguez was reviewer for: Transactions on Emerging Telecommunications Technologies, Knowledge-Based Systems, Robotics and Autonomous Systems, IEEE Robotics & Automation Magazine, IEEE Transactions on Cognitive and Developmental Systems.
- PY Oudeyer was reviewer for Cognitive Science, Child Development Perspectives and Nature Scientific Reports.

10.1.4. Invited Talks

- David Filliat gave an invited presentation at 'Journées Robotique et IA' in PFIA18, on July 5th, 2018.
- PY Oudeyer, "Apprentissage autonome développemental et modèles du développement cognitif chez l'enfant", Paris, Académie des Sciences, Oct. 2018.
- PY Oudeyer, "Developmental Autonomous Learning: Can a Machine Learn like a Child?", Académie des Technologies, Paris, Oct. 2018
- PY Oudeyer, "Computational Theories of Curiosity-driven Learning: Cognitive Science and AI", Exploring Curiosity conference, Amsterdam, Nov. 2018.
- PY Oudeyer, "Developmental Autonomous Learning: Artificial Intelligence and Cognitive Sciences", Univ. Columbia, NY, Oct. 2018.
- PY Oudeyer, "Le projet KidLearn: motivations intrinsèques, apprentissage et edTech", Rectorat de l'Académie de Bordeaux, Bordeaux, Jan. 2018.
- PY Oudeyer, "Computational Theories of Curiosity-driven Development", Multidisciplinary Developmental Dynamics conference (ETF 18), University of East Anglia, UK, June 2018.
- PY Oudeyer, "Developmental Autonomous Learning: AI, Cognitive Science and Educational Technologies", joint lab Inria and Microsoft Research event, Paris, june 2018.
- PY Oudeyer, "Developmental Autonomous Learning: AI, Cognitive Science and Educational Technologies", Ubisoft, Paris, June 2018.
- PY Oudeyer, "Developmental Autonomous Learning: AI, Cognitive Science and Educational Technologies", Summer School on Human-Robot Interaction, Animatas project, Paris, sept. 2018.
- PY. Oudeyer, "Developmental exploration and active learning", Cloud Temple event on AI and machine learning, Paris, january 2018.
- PY Oudeyer, "From models of curiosity-driven learning to applications in Personalized eLearning technologies", international conference "Curiosity: Emerging Sciences and Educational Innovations" at University of Pennsylvania, US, Dec. 2018.
- PY Oudeyer, "Computational Theories of Curiosity-driven Learning", Symposium on the Biology of Decision Making (SBDM 2018), Ecole Normale Supérieure, Paris, june 2018.
- N. Diaz Rodriguez gave an invited talk at Inria Flowers Deep RL workshop 4/4/2018, Research seminar on State representation learning for robotics control at JRC Sevilla European Commission, Spain, 23 April 2018 [2]
- N. Diaz Rodriguez gave an invited talk at Satellite workshop 24 May 2018 @ Sorbonne Université on Learning and decision-making at the interface between Neuroscience, Artificial Intelligence and Robotics. http://sbdm2018.isir.upmc.fr/index.php?perma=1520611011

---

[2]https://ec.europa.eu/jrc/communities/community/event/research-seminar-natalia-diaz-rodriguez-ensta-paris-state-representation-learning

10.1.5. Leadership within the Scientific Community

- PY. Oudeyer has been editor of the IEEE CIS Newsletter on Cognitive and Developmental systems, organizing two interdisciplinary dialogs, see https://openlab-flowers.inria.fr/t/ieee-cis-newsletter-on-cognitive-and-developmental-systems/129.

10.1.6. Scientific Expertise

- PY. Oudeyer has been a reviewer for the European Commission (FET program).
- D. Filliat has been a member of the ANR ASTRID evaluation comittee.
- N. Diaz Rodriguez has been external expert reviewer for ANR-JST CREST IS 2018 program.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

ENSEIRB, 12h, Robotics project (Thibault Desprez)

IUT Informatique, 64h, IUT Informatique Bordeaux (Rémy Portelas).

Master: Robotique Mobile, 21 heures. M2, ENSTA - ParisTech (David Filliat).

Master: Perception pour la Robotique, 6 heures. M2, ENSTA - ParisTech (David Filliat).

Master: Perception pour la robotique, 12 heures. M2 Systemes Avances et Robotique, Sorbonnes University (David Filliat)

Master: Perception pour la Robotique Développementale, 3 hours, CogMaster (David Filliat)

Master: IN104 Projet Informatique, 20 h. TD (N. Diaz Rodriguez).

Master: IA301 (Telecom ParisTech): Logics and Symbolic Artificial Intelligence, 9h (N. Diaz Rodriguez)

Master: ROB313: Computer vision for autonomous systems, 8.5 h TD (Perception pour les Systèmes Autonomes, N. Diaz Rodriguez)

Master: Cours de robotique développementale, option robot, ENSEIRB (2h), PY. Oudeyer

10.2.2. Supervision

PhD in progress: Rémy Portelas, Intrinsically Motivated Goal Exploration in Open-Ended Worlds (Minecraft) (superv. P-Y. Oudeyer)

PhD in progress: Cédric Colas, Algorithms for Intrinsically Motivated Goal Exploration (superv. P-Y. Oudeyer)

PhD in progress: Sébastien Forestier, Models of curiosity-driven learning of tool use and speech development, started in sept. 2015 (superv. P-Y. Oudeyer)

PhD in progress: Thibault Desprez, Design and study of the impact of educational robotic kits in computer science education, started in dec. 2016 (superv. P-Y. Oudeyer)

PhD completed in 2018: Benjamin Clement, Intelligent Tutoring Systems, started oct 2015 (superv. Manuel Lopes and Pierre-Yves Oudeyer).

PhD completed in 2018: William Schueller, Study of the impact of active learning and teaching in naming games dynamics, started in sept. 2015 (superv. P-Y. Oudeyer)

PhD completed in 2018: Florian Golemo, Design and study of policy learning and Sim2Real transfer algorithms for robotics (superv. Pierre-Yves Oudeyer and Aaron Courville)

PhD completed: Baptiste Busch, Interactive Learning, started oct 2014 (superv. Manuel Lopes).

PhD completed in 2018: Alexandra Delmas, Auto-Apprentissage Auto-Adaptable pour la compliance au traitement, started oct 2014 (superv. Hélène Sauzéon and Pierre-Yves Oudeyer).

PhD completed in 2018: Adrien Matricon : Task dependent visual feature selection for optimising and generalizing robotics skills (superv. David Filliat, Pierre-Yves Oudeyer), defended June, 11th, 2018.

PhD in progress: José Magno Mendes Filho, Planning and control of an autonomous AGV in environment shared with humans, started Oct. 2015 (superv. David Filliat and Eric Lucet (CEA))

PhD in progress: Timothée Lesort, Incremental Deep Learning for Detection and Classification in a Robotic Context. started june 2017 (superv. David Filliat and Jean-Francois Goudou (THALES)).

PhD in progress: Vyshakh Palli Thazha, Data fusion for autonomous vehicles. started sept 2017 (superv. David Filliat and Hervé Illy (Renault)).

PhD in progress: Florence Carton, Exploration of reinforcement learning algorithms for drone visual perception and control started dec 2017 (superv. David Filliat and Jaonary Rabarisoa (CEA)).

PhD in progress: Hugo Caselles-Dupré, Incremental learning for sensori-motor control started june 2018 (superv. David Filliat and Michael Garcia-Ortiz (Softbank Robotics)).

10.2.3. *Juries*

David Filliat was in the PhD jury of François de la Bourdonnaye (18/12/18, Rapporteur), Arnaud Tanguy (28/11/18, Rapporteur), Dinesh Atchuthan (23/10/18, Examinateur), Zhan Wang (19/10/18, Examinateur), Quentin Bateux (12/02/18, Examinateur), Clément Delgrange (reviewer).

David Filliat was in the Hdr jury of Sylvain Argentieri (06/12/18, Rapporteur)

PY Oudeyer was in the PhD jury of Héloise Thero (ENS Paris, examiner), Matthieu Geisert (Univ. Toulouse, reviewer), Konstantinos Chatzilygeroudis (Univ. Lorraine, reviewer), Clément Delgrange (Univ. Dijon, examiner).

Thibault Desprez was in the internship jury at Enseirb Bordeaux in October 2018.

10.3. Popularization

10.3.1. *Teaching and Education*

10.3.1.1. *Inirobot*

IniRobot (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/ 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.

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.

Created pedagogical documents and resources:

- The inirobot pedagogical kit [83]: 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 http://tice33.ac-bordeaux.fr/Ecolien/ASTEP/tabid/5953/language/fr-FR/Default.aspx

- Inirobot website and forum: https://dm1r.inria.fr/ or http://www.inirobot.fr On this website, teachers, animators and general public can download documents, exchange about their use of inirobot's kit.

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

10.3.1.2. MOOC Thymio

Didier Roy played a central role in the design and making of 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. Web: https://www.fun-mooc.fr/courses/course-v1:inria+41017+session01/about

10.3.1.3. Poppy Education

As part of the Poppy Education project, thanks the robotic platform Poppy we created pedagogical kits opensource and low cost for teaching computer science and robotics. It is designed to help young people to take ownership with concepts and technologies of the digital world.

The Pedagogical kits includes robots and pedagogical resources. They have been co-created directly with users (mainly high schools teachers) and evaluated in real life by experiments in classrooms [120].
The activities were designed with the visual programming language Snap! (Scratch like) and Python, but some are in Java / Processing (thanks the robot API you can use the language of your choice).

Most activities are using the robot Poppy Ergo Jr, but some use Poppy Torso (mostly in higher school because of its cost) and Poppy Humanoid (in kinder-garden for dance projects) :

- The Poppy Ergo Jr robot is a small and low cost 6-degree-of-freedom robot arm. It consists of simple shapes which can be easily 3D printed. It has several 3D printed tools extending its capabilities (there are currently the lampshade, the gripper and a pen holder but you can design new ones). They are assembled via rivets which can be removed and added very quickly with the OLLO tool. Each motor has LEDs on (8 different color can be activated). The electronic card (raspberry Pi) is visible next to the robot, that allow to manipulate, and plug extra sensors.

- The Poppy Torso robot is an open-source humanoid robot torso which can be installed easily on tabletops. More affordable than the robot Poppy Humanoid, it is an ideal medium to learn science, technology, engineering and mathematics.

We continued to improve the robots functionalities and you can see below the resources we created :

- A website have been created to present the project and to share all resources and activities. https://www.poppy-education.org/

- To complete the pedagogical booklet [119] that provides guided activities and small challenges to become familiar with Poppy Ergo Jr robot and the Programming language Snap! (https://drive.google.com/file/d/0B2jV8VX-lQHwTUxXZjF3OGxHVGM/view) we provided a list of Education projects. Educational projects have been written for each activity carried out and tested in class. So each projects have its own web page including resources allowing any other teacher to carry out the activity (description, pedagogical sheet, photos / videos, pupil's sheet, teacher's sheet with correction etc.). Their is now 32 activities documented available on Poppy Education website.

You can see the activities on this links (in french):

- Introduction to Ergo Jr and Snap! :
  https://www.poppy-education.org/activites/initiation-ergo-jr-et-snap

- Ergo Jr and Python tutorials :
  https://www.poppy-education.org/activites/tuto-python-robot-ergojr

- High-school levels :
  www.poppy-education.org/activites/activites-lycee

- Middle-school level :
  www.poppy-education.org/activites/activites-college

- Primary Schools level :
  https://www.poppy-education.org/activites/activites-primaire/

- Demonstrations (just videos to show the possibilities) :
  https://www.poppy-education.org/activites/demos/

- 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/.

- A FAQ have been written with the most frequents questions to help users: https://www.poppy-education.org/aide/

- New activities on Poppy Education website and forum.

- New section : Activities with Python.

- Improvements in the Resources page of the Poppy Education website.

10.3.2. Talks and Hands-on

- Thibault Desprez, December 2018 at Inria Bordeaux, welcomed three students from middle-school during two days to discover the working environment and to introduce them to robotics.

- Thibault Desprez, Inria Bordeaux open day, November 2018, exhibition stand to present Poppy Education and Poppy robots

- Thibault Desprez, National Meeting of Educational Robotics, October 2018 at ifé ENS Lyon, two talk to present Poppy robots kits in school.

- Thibault Desprez, National Meeting of Educational Robotics, October 2018 at ifé ENS Lyon, exhibition stand to present Poppy Station and Poppy robots.

- Thibault Desprez, Théo Segonds, Fête de la science (Inria Bordeaux Sud-Ouest), October 2018, 4 programming workshop in 2 days (with middle school students) using Snap! and the robot Poppy Ergo Jr.

- Thibault Desprez, Meet-up & Educate, October 2018 at INP Bordeaux, exhibition stand to student recruitment for a project on Poppy robots.

- Thibault Desprez, PI space inauguration, July 2018 at ESPE Mérignac, exhibition stand to present Poppy Station and Poppy robots.

- Thibault Desprez, Bordeaux Geek Festival, May 2018, Parc expo, Talk about societal problem on robotics.

- Thiabult Desprez, Usine Végétale inauguration, May 2018 at Le Fieu, exhibition stand to present Poppy Education and Poppy robots in rural zone.

- Thibault Desprez, Connect'thouars, April 2018 at Talence, Workshops to initiate in programming.

- Thibault Desprez, Didapro 7, February 2018 at HEP Vaud, Lausanne, talk to present the article : "Poppy Ergo Jr : un kit robotique au coeur du dispositif Poppy Éducation"

- Thibault Desprez, Théo Segonds, Fondation Main à la pate, February 2018 at Paris, Two days to train a group of teachers to robotics and programmation with Poppy Ergo Jr robot.
- Thibault Desprez, e-Fran seminar, January 2018 at Minister of Higher Education, Research and Innovation, poster to present my thesis.
- Théo Segonds. Poppy Ergo Jr Workshop at CERN (Geneve). Construction and programming of the robotic arm Poppy Ergo Jr.
- Théo Segonds. Scientae Robotica, Lausanne. Poppy Ergo Jr Workshop. Construction and programming of the robotic arm Poppy Ergo Jr.
- Benjamin Clement and Alexandra Delmas, EdTech days, may 2018. Presentation of kidlearn and kidbreath projects.
- Theo Segonds, Didier Roy. PLAIRE Festival in Evian with Poppy exhibition during 2 days.
- Alexandra Delmas, Didier Roy. Forum Educavox in Bordeaux. Presentation of kidlearn and kidbreath projects.
- Didier Roy. R2T2 Richter event, remote robotics programming, in caribbean islands, in collaboration with EPFL.
- Didier Roy. Inria Scientific Days, presentation of educational projects in Flowers Team.
- Stephanie Noirpoudre. Poppy Education présent à la journée EIDOS 65 : Le forum des pratiques numériques pour l'éducation. Description and feedback of the 9th edition of the EIDOS 65 day (the digital practice forum for education).
- PY Oudeyer mentored students from College de Cadillac for their robotics project (2 days), march 2018.
- PY Oudeyer gave a talk "Intelligence artificielle: un outil pour nous aider à mieux comprendre l'intelligence naturelle?" at Collège Cadillac, Gironde, may 2018.
- PY Oudeyer gave a talk "Intelligence artificielle: apprentissage automatique et sciences cognitives" at a training event for members of Bordeaux Metropole political and decision staff, Nov. 2018.
- PY Oudeyer gave a talk "Intelligence artificielle: apprentissage automatique et sciences cognitives" at Université de Tous les Savoirs, Arcachon, janv. 2018.

10.3.3. Popularizing inside Inria

- Théo Segonds and Thibault Desprez. Poppy Ergo Jr training for Inria Scientific Mediation members.
- Inria National Scientific Mediation Seminar: Presentation by Stéphanie Noirpoudre and Théo Segonds of Poppy Ergo Jr, and workshop.
- Sébastien Forestier made a presentation on models of curiosity-driven development at Unithé ou Café.

10.3.4. Innovation and transfer

- 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. Many exchanges have already taken place with potential partners such as the EPFL, the ENSAM network, the «Ligue de l'Enseignement», Génération Robots, the French Institute of Education, several academies, the direction of digital education of the Ministry of Education, ... 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.poppystation.org

10.3.5. Internal or external Inria responsibilities

D. Roy is member of the Class'code team (Inria is member of the consortium of this project) https://pixees.fr/classcode/accueil/. Class'code is a blended formation for teachers and animators who aim to initate young people to computer science and robotics. D. Roy has in charge the robotics module of the project.

D. Roy is adviser of the organization of computer science exhibition in "Palais de la découverte" which has begun on 2018 March. He helps for robotics part.

D. Roy is member of the team "Education en Scène" which organize educational activities with robotics in Bordeaux Digital City.

D. Roy is member of the scientific committee of "Learning Computer Science at School" project in Canton de Vaud (Switzerland).

D. Roy is member of the Robocup Junior French committee, an international robotics challenge http://rcj. robocup.org/.

D. Roy is member of the scientific committee of "Ludovia CH" Conference which will be held in Yverdon (Switzerland) on 2019 March.

D. Roy is project co-leader of MOOC Thymio, in collaboration with EPFL and Inria Learning Lab. The aim of this MOOC is to propose to teachers a training on basics of computer science, using the robotic platform Thymio.

D. Roy is associate member of the EPFL "LEARN" center.

PY. Oudeyer continued to be the PI of the Poppy Education project.

PY. Oudeyer was scientific mentor for stiudents of College de Cadillac, within the program "Main à la pâte" of Maison des Sciences.

10.3.6. Articles and contents

- Adrien Laversanne-Finot wrote a blog post on "Discovery of independently controllable features through autonomous goal setting", https://openlab-flowers.inria.fr/t/discovery-of-independently-controllable-features-through-autonomous-goal-setting/494

- Cédric Colas wrote a blog post on "How Many Random Seeds Should I Use? Statistical Power Analysis in (Deep) Reinforcement Learning Experiments", https://openlab-flowers.inria.fr/t/how-many-random-seeds-should-i-use-statistical-power-analysis-in-deep-reinforcement-learning-experiments/457

- Cédric Colas wrote a blog post on "Bootstrapping Deep RL with population-based diversity search"

- PY Oudeyer was interviewed for an article of **The Economist** on curiosity-driven learning, http://www.pyoudeyer.com/TheEconomist2018.pdf

- PY Oudeyer was interviewed for an article of **Scientific American** on curiosity-driven learning, http://www.pyoudeyer.com/IntelligentMachinesThatLearnLikeChildren-Scientific%20American2018.pdf

- PY Oudeyer was interviewed for an article of **MIT Technology Review** on curiosity-driven learning, http://www.pyoudeyer.com/may17MITTechnology%20Review.pdf

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, n⁰ 1, pp. 69-73 [*DOI :* 10.1016/J.ROBOT.2012.05.008], https://hal.inria.fr/hal-00788440

[2] 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, n⁰ 2, pp. 20–48, https://hal.inria.fr/hal-00913669

[3] 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

[4] S. FORESTIER, P.-Y. OUDEYER. *Modular Active Curiosity-Driven Discovery of Tool Use*, in "IEEE/RSJ International Conference on Intelligent Robots and Systems", Daejeon, South Korea, Proceedings of the 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2016, https://hal.archives-ouvertes.fr/hal-01384566

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

[6] S. IVALDI, S. M. NGUYEN, N. LYUBOVA, A. DRONIOU, V. PADOIS, D. FILLIAT, P.-Y. OUDEYER, O. SIGAUD. *Object learning through active exploration*, in "IEEE Transactions on Autonomous Mental Development", 2013, pp. 1-18 [*DOI :* 10.1109/TAMD.2013.2280614], https://hal.archives-ouvertes.fr/hal-00919694

[7] 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

[8] 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

[9] O. MANGIN, D. FILLIAT, L. TEN BOSCH, P.-Y. OUDEYER. *MCA-NMF: Multimodal Concept Acquisition with Non-Negative Matrix Factorization*, in "PLoS ONE", October 2015, vol. 10, n° 10, e0140732 p. [*DOI :* 10.1371/JOURNAL.PONE.0140732.T005], https://hal.archives-ouvertes.fr/hal-01137529

[10] C. MOULIN-FRIER, S. M. NGUYEN, P.-Y. OUDEYER. *Self-Organization of Early Vocal Development in Infants and Machines: The Role of Intrinsic Motivation*, in "Frontiers in Psychology", 2013, vol. 4, n° 1006 [*DOI :* 10.3389/FPSYG.2013.01006], https://hal.inria.fr/hal-00927940

Publications of the year

Doctoral Dissertations and Habilitation Theses

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

[12] F. GOLEMO. *How to Train Your Robot - New Environments for Robotic Training and New Methods for Transferring Policies from the Simulator to the Real Robot*, Université de Bordeaux, December 2018, https://hal.inria.fr/tel-01974203

[13] A. MATRICON. *Merging robotic skills into more general skills*, Université de Bordeaux, June 2018, https://tel.archives-ouvertes.fr/tel-01895789

[14] W. SCHUELLER. *Active Control of Complexity Growth in Language Games*, Université de Bordeaux, December 2018, https://hal.inria.fr/tel-01966815

Articles in International Peer-Reviewed Journals

[15] Y. CHEN, J.-B. BORDES, D. FILLIAT. *Comparison studies on active cross-situational object-word learning using Non-Negative Matrix Factorization and Latent Dirichlet Allocation*, in "IEEE Transactions on Cognitive and Developmental Systems", 2018 [*DOI :* 10.1109/TCDS.2017.2725304], https://hal.archives-ouvertes.fr/hal-01561168

[16] 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, https://hal.archives-ouvertes.fr/hal-01954983

[17] C. CRAYE, D. FILLIAT, J.-F. GOUDOU. *BioVision: a Biomimetics Platform for Intrinsically Motivated Visual Saliency Learning*, in "IEEE Transactions on Cognitive and Developmental Systems", 2018 [*DOI :* 10.1109/TCDS.2018.2806227], https://hal.archives-ouvertes.fr/hal-01728340

[18] 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

[19] A. DELMAS, B. CLÉMENT, P.-Y. OUDEYER, H. SAUZÉON. *Fostering Health Education With a Serious Game in Children With Asthma: Pilot Studies for Assessing Learning Efficacy and Automatized Learning Personalization*, in "Frontiers in Education ", November 2018, vol. 3 [*DOI :* 10.3389/FEDUC.2018.00099], https://hal.archives-ouvertes.fr/hal-01922316

[20] 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

[21] C. FAGE, C. CONSEL, E. BALLAND, K. ETCHEGOYHEN, A. AMESTOY, M. BOUVARD, H. SAUZÉON. *Tablet Apps to Support First School Inclusion of Children With Autism Spectrum Disorders (ASD) in Mainstream Classrooms: A Pilot Study*, in "Frontiers in Psychology", October 2018, vol. 9 [*DOI :* 10.3389/FPSYG.2018.02020], https://hal.inria.fr/hal-01904791

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

[23] 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

[24] 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", December 2018, https://hal.inria.fr/hal-01950078

International Conferences with Proceedings

[25] 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

[26] 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

[27] F. GOLEMO, A. A. TAÏGA, P.-Y. OUDEYER, A. COURVILLE. *Sim-to-Real Transfer with Neural-Augmented Robot Simulation*, in "Conference on Robot Learning (CoRL) 2018", Zurich, Switzerland, October 2018, https://hal.inria.fr/hal-01911978

[28] 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

[29] C. MAZON, C. FAGE, A. AMESTOY, I. HESLING, M. BOUVARD, K. ETCHEGOYHEN, H. SAUZÉON. *Cognitive mediators of school-related socio-adaptive behaviors in children and adolescents with ASD: A pilot study*, in "4th International Congress of Clinical and Health Pscyhology on Children and Adolescents", Palma de Mallorca, Spain, Book of Abstracts of the 4th International Congress of Clinical and Health Pscyhology on Children and Adolescents, Aitana Investigacíon, November 2018, https://hal.inria.fr/hal-01939740

[30] P. PAPADAKIS, D. FILLIAT. *Generic Object Discrimination for Mobile Assistive Robots using Projective Light Diffusion*, in "WACV 2018 - IEEE Winter Conference on Applications of Computer Vision, Workshop CV-AAL - Computer Vision for Active and Assisted Living", Reno, United States, March 2018, pp. 1-9, https://hal.inria.fr/hal-01699842

[31] W. SCHUELLER, V. LORETO, P.-Y. OUDEYER. *Complexity Reduction in the Negotiation of New Lexical Conventions*, in "40th Annual Conference of the Cognitive Science Society (CogSci 2018)", Madison, WI, United States, July 2018, https://hal.inria.fr/hal-01891762

Conferences without Proceedings

[32] 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

[33] 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 - Thirty-second Conference on Neural Information Processing Systems", Montréal, Canada, December 2018, https://hal.archives-ouvertes.fr/hal-01951399

[34] 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

[35] I. HUITZIL, U. STRACCIA, N. DÍAZ-RODRÍGUEZ, F. BOBILLO. *Datil: Learning Fuzzy Ontology Datatypes*, in "IPMU 2018: 17th Information Processing and Management of Uncertainty in Knowledge-Based Systems Conference", Cádiz, Spain, June 2018, https://hal.archives-ouvertes.fr/hal-01951785

[36] V. LOMONACO, A. TROTTA, M. ZIOSI, J. DE DIOS YÁÑEZ ÁVILA, N. DÍAZ-RODRÍGUEZ. *Intelligent Drone Swarm for Search and Rescue Operations at Sea*, in "Workshop on AI for Good, NeurIPS 2018 (Neural Information Processing Systems)", Montreal, Canada, December 2018, https://hal.archives-ouvertes.fr/hal-01951515

[37] J. M. MENDES FILHO, E. LUCET, D. FILLIAT. *Experimental Validation of a Multirobot Distributed Receding Horizon Motion Planning Approach*, in "ICARCV 2018 - 15th International Conference on Control, Automation, Robotics and Vision", Singapour, Singapore, November 2018, https://hal.archives-ouvertes.fr/hal-01935322

[38] 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

[39] 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 "NIPS 2018 Deep RL workshop", Montreal, Canada, December 2018, https://hal.archives-ouvertes.fr/hal-01931713

Scientific Books (or Scientific Book chapters)

[40] K. BOLLACKER, N. DÍAZ-RODRÍGUEZ, X. LI. *Extending Knowledge Graphs with Subjective Influence Networks for personalized fashion*, in "Designing Cognitive Cities", September 2018, https://hal.archives-ouvertes.fr/hal-01952205

[41] N. DÍAZ-RODRÍGUEZ, S. GRÖNROOS, F. WICKSTRÖM, J. LILIUS, H. EERTINK, A. BRAUN, P. DILLEN, J. CROWLEY, J. ALEXANDERSSON. *An Ontology for Wearables Data Interoperability and Ambient Assisted Living Application Development*, in "Recent Developments and the New Direction in Soft-Computing Foundations and Applications", L. A. ZADEH, R. R. YAGE, S. N. SHAHBAZOVA, M. Z. REFORMA, V. KREINOVICH (editors), Studies in Fuzziness and Soft Computing, Springer, 2018, vol. 361, pp. 559-568, Selected Paper from the 6th World Conference on Soft Computing, May 22-25, 2016, Berkeley, USA [*DOI :* 10.1007/978-3-319-75408-6_43], https://hal.archives-ouvertes.fr/hal-01951531

Other Publications

[42] 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

[43] C. COLAS, O. SIGAUD, P.-Y. OUDEYER, P. FOURNIER. *CURIOUS: Intrinsically Motivated Multi-Task Multi-Goal Reinforcement Learning*, November 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01934921

[44] T. DESPREZ, S. NOIRPOUDRE, T. SEGONDS, D. CASELLI, D. ROY, P.-Y. OUDEYER. *Conception et évaluation de kits robotiques pédagogiques: Analyse écologique et expérimentale des utilisations de la robotique à l'école en termes de connaissances et de représentations* , March 2018, Journée de l'EDMI, Poster, https://hal.archives-ouvertes.fr/hal-01780511

[45] T. DESPREZ, S. NOIRPOUDRE, T. SEGONDS, D. CASELLI, D. ROY, P.-Y. OUDEYER. *Design and Evaluation of Pedagogical Robotic Kits : Ecological and experimental analysis of the uses of robotics in schools in terms of knowledge and representation*, January 2018, 1 p. , Colloque e-Fran, Territoires éducatifs d'innovation numérique, Poster, https://hal.archives-ouvertes.fr/hal-01688310

[46] C. MAZON, C. FAGE, H. SAUZÉON. *Effectiveness and Usability of Technology-based Interventions with children and adolescents with ASD: a systematic review*, July 2018, EuroScience Open Forum (ESOF 2018), Poster, https://hal.inria.fr/hal-01939765

References in notes

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

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

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

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

[51] 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

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

[53] 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° 1, pp. 69-73 [*DOI :* 10.1016/J.ROBOT.2012.05.008], https://hal.inria.fr/hal-00788440

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

[55] 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

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

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

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

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

[60] 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

[61] K. P. BURNHAM, D. R. ANDERSON. *Model selection and multimodel inference: A practical information-theoretic approach*, Springer Science & Business Media, 2003

[62] A. CANGELOSI, G. METTA, G. SAGERER, S. NOLFI, C. NEHANIV, K. FISCHER, J. TANI, T. BELPAEME, G. SANDINI, F. NORI. *Integration of action and language knowledge: A roadmap for developmental robotics*, in "Autonomous Mental Development, IEEE Transactions on", 2010, vol. 2, n° 3, pp. 167–195

[63] D. CENTOLA, A. BARONCHELLI. *The spontaneous emergence of conventions: An experimental study of cultural evolution*, in "Proceedings of the National Academy of Sciences", 2015, vol. 112, n° 7, pp. 1989–1994

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

[65] 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° 2, pp. 20–48, https://hal.inria.fr/hal-00913669

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

[67] M. CORNUDELLA, P. VAN EECKE, R. VAN TRIJP. *How Intrinsic Motivation can Speed Up Language Emergence*, in "Proceedings of the European Conference on Artificial Life", 2015, pp. 571–578

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

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

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

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

[72] P. DHARIWAL, C. HESSE, O. KLIMOV, A. NICHOL, M. PLAPPERT, A. RADFORD, J. SCHULMAN, S. SIDOR, Y. WU, P. ZHOKHOV. *OpenAI Baselines*, GitHub, 2017, https://github.com/openai/baselines

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

[74] 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

[75] 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

[76] 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

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

[78] 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

[79] M. C. FRANK, N. D. GOODMAN. *Predicting pragmatic reasoning in language games*, in "Science", 2012, vol. 336, n⁰ 6084, pp. 998–998

[80] B. GALANTUCCI, S. GARROD. *Experimental semiotics: a review*, in "Frontiers in human neuroscience", 2011, vol. 5, 11 p.

[81] 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, n⁰ 11, pp. 585-93 [*DOI :* 10.1016/J.TICS.2013.09.001], https://hal.inria.fr/hal-00913646

[82] 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, n⁰ 11, pp. 585–593

[83] 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

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

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

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

[87] P. HENDERSON, R. ISLAM, P. BACHMAN, J. PINEAU, D. PRECUP, D. MEGER. *Deep reinforcement learning that matters*, in "arXiv preprint arXiv:1709.06560", 2017

[88] J.-C. HORVITZ. *Mesolimbocortical and nigrostriatal dopamine responses to salient non-reward events*, in "Neuroscience", 2000, vol. 96, n⁰ 4, pp. 651-656

[89] 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

[90] 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

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

[92] F. KAPLAN, P.-Y. OUDEYER, B. BERGEN. *Computational models in the debate over language learnability*, in "Infant and Child Development", 2008, vol. 17, n° 1, pp. 55–80

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

[94] S. KIRBY, T. GRIFFITHS, K. SMITH. *Iterated learning and the evolution of language*, in "Current opinion in neurobiology", 2014, vol. 28, pp. 108–114

[95] 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

[96] 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° 5, pp. 239–252

[97] T. P. LILLICRAP, J. J. HUNT, A. PRITZEL, N. HEESS, T. EREZ, Y. TASSA, D. SILVER, D. WIERSTRA. *Continuous control with deep reinforcement learning*, in "arXiv preprint arXiv:1509.02971", 2015

[98] G. LOEWENSTEIN. *The psychology of curiosity: A review and reinterpretation*, in "Psychological bulletin", 1994, vol. 116, n° 1, 75 p.

[99] 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

[100] 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

[101] M. LOPES, L. MONTESANO. *Active Learning for Autonomous Intelligent Agents: Exploration, Curiosity, and Interaction*, in "CoRR", 2014, vol. abs/1403.1, 40 p.

[102] V. LORETO, A. BARONCHELLI, A. MUKHERJEE, A. PUGLISI, F. TRIA. *Statistical physics of language dynamics*, in "Journal of Statistical Mechanics: Theory and Experiment", 2011, vol. 2011, n° 04, P04006 p.

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

[104] 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

[105] 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

[106] 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

[107] 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

[108] 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

[109] 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

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

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

[112] C. MOULIN-FRIER, S. M. NGUYEN, P.-Y. OUDEYER. *Self-Organization of Early Vocal Development in Infants and Machines: The Role of Intrinsic Motivation*, in "Frontiers in Psychology", 2013, vol. 4, n° 1006 [*DOI :* 10.3389/FPSYG.2013.01006], https://hal.inria.fr/hal-00927940

[113] 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

[114] 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

[115] 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

[116] 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

[117] S. M. NGUYEN, P.-Y. OUDEYER. *Active Choice of Teachers, Learning Strategies and Goals for a Socially Guided Intrinsic Motivation Learner*, in "Paladyn", September 2012, vol. 3, n⁰ 3, pp. 136-146 [*DOI :* 10.2478/S13230-013-0110-Z], https://hal.inria.fr/hal-00936932

[118] S. M. NGUYEN, P.-Y. OUDEYER. *Socially Guided Intrinsic Motivation for Robot Learning of Motor Skills*, in "Autonomous Robots", March 2014, vol. 36, n⁰ 3, pp. 273-294 [*DOI :* 10.1007/S10514-013-9339-Y], https://hal.inria.fr/hal-00936938

[119] 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

[120] 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.

[121] P.-Y. OUDEYER, F. DELAUNAY. *Developmental exploration in the cultural evolution of lexical conventions*, in "8th International Conference on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems", Brighton, United Kingdom, 2008, https://hal.inria.fr/inria-00420303

[122] P.-Y. OUDEYER, F. KAPLAN, V. HAFNER. *Intrinsic Motivation for Autonomous Mental Development*, in "IEEE Transactions on Evolutionary Computation", January 2007, vol. 11, n⁰ 2, pp. 265-286 [*DOI :* 10.1109/TEVC.2006.890271], https://hal.inria.fr/hal-00793610

[123] P.-Y. OUDEYER, F. KAPLAN, V. HAFNER. *Intrinsic Motivation Systems for Autonomous Mental Development*, in "IEEE Transactions on Evolutionary Computation", 2007, vol. 11, n⁰ 1, pp. 265–286, http://www.pyoudeyer.com/ims.pdf

[124] 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

[125] P.-Y. OUDEYER, F. KAPLAN. *What is intrinsic motivation? A typology of computational approaches*, in "Frontiers in Neurorobotics", 2007, vol. 1, n⁰ 1

[126] 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/

[127] P.-Y. OUDEYER. *Developmental Learning of Sensorimotor Models for Control in Robotics*, in "SIAM News", September 2014, vol. 47, n⁰ 7, https://hal.inria.fr/hal-01061633

[128] P.-Y. OUDEYER, L. SMITH. *How Evolution May Work Through Curiosity-Driven Developmental Process*, in "Topics in cognitive science", February 2016, vol. 8 [*DOI :* 10.1111/TOPS.12196], https://hal.inria.fr/hal-01404334

[129] 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/

[130] 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

[131] M. PLAPPERT, R. HOUTHOOFT, P. DHARIWAL, S. SIDOR, R. Y. CHEN, X. CHEN, T. ASFOUR, P. ABBEEL, M. ANDRYCHOWICZ. *Parameter space noise for exploration*, in "arXiv preprint arXiv:1706.01905", 2017

[132] 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.

[133] 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

[134] 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° 1, pp. 86–90

[135] 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

[136] 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/

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

[138] 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

[139] 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

[140] W. SCHUELLER, P.-Y. OUDEYER. *Active Learning Strategies and Active Control of Complexity Growth in Naming Games*, in "the 5th International Conference on Development and Learning and on Epigenetic Robotics", Providence, RI, United States, August 2015, https://hal.inria.fr/hal-01202654

[141] W. SCHUELLER, P.-Y. OUDEYER. *Active Control Of Complexity Growth In Naming Games: Hearer's Choice*, in "The Evolution of Language: Proceedings of the 11th International Conference (EVOLANGX11)", 2016

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

[143] L. STEELS, F. KAPLAN, A. MCINTYRE, J. VAN LOOVEREN. *Crucial factors in the origins of word-meaning*, in "The transition to language", 2002, vol. 12, pp. 252–271

[144] L. STEELS. *Language games for autonomous robots*, in "Intelligent Systems, IEEE", 2001, vol. 16, n⁰ 5, pp. 16–22

[145] L. STEELS. *The Autotelic Principle*, in "Science", 2004, vol. 3139, pp. 1–16

[146] 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

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

[148] 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

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

[150] 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

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

[152] P. WELLENS. *Adaptive Strategies in the Emergence of Lexical Systems*, Vrije Universiteit Brussel, Brussels, 2012, http://ai.vub.ac.be/publications/918

[153] 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