

RESEARCH CENTRE

**Inria Centre
at the University of Bordeaux**

IN PARTNERSHIP WITH:

**Ecole nationale supérieure des techniques
avancées**

2023

ACTIVITY REPORT

Team

FLOWERS

Flowing Epigenetic Robots and Systems

Inria teams are typically groups of researchers working on the definition of a common project, and objectives, with the goal to arrive at the creation of a project-team. Such project-teams may include other partners (universities or research institutions)

DOMAIN

Perception, Cognition and Interaction

THEME

Robotics and Smart environments

Inria

Contents

Team FLOWERS	1
1 Team members, visitors, external collaborators	3
2 Overall objectives	4
3 Research program	6
4 Application domains	9
5 Social and environmental responsibility	11
5.1 Footprint of research activities	11
5.2 Impact of research results	11
6 Highlights of the year	12
6.1 Awards	15
6.2 PhD defenses	15
7 New software, platforms, open data	15
7.1 New software	15
7.1.1 SocialAI	15
7.1.2 AutoDisc	16
7.1.3 Kids Ask	16
7.1.4 ToGather	16
7.1.5 mc_training	17
7.1.6 Evolution of adaptation mechanisms in complex environments	17
7.1.7 SAPIENS	18
7.1.8 architect-builder-abig	18
7.1.9 EAGER	19
7.1.10 IMGC-MARL	19
7.1.11 Flow-Lenia	20
7.1.12 Kidlearn: money game application	20
7.1.13 cognitive-testbattery	21
7.1.14 LLM_stability	21
7.1.15 Sensorimotor-lenia	21
7.1.16 MetaIPPO	22
7.1.17 Lamorel	22
7.1.18 GLAM	22
7.1.19 ER-MRL	23
7.1.20 ecoevojax_analysis	23
7.1.21 SBMLtoODEjax	23
7.2 New platforms	24
7.2.1 ToGather application	24
8 New results	24
8.1 Open-ended learning and autotelic AI with large language models	25
8.1.1 Language-Model-Augmented Autotelic Agents, Towards Open-Ended Skill Discovery	25
8.1.2 Code-Generating Autotelic Agents	27
8.1.3 Codeplay: Autotelic Learning through Collaborative Self-Play in Programming Environments	27
8.1.4 ACES: Generating diverse programming puzzles with autotelic language models and semantic descriptors	30
8.1.5 Grounding Large Language Models in Interactive Environments with Online Reinforcement Learning (GLAM)	31
8.1.6 Automatic Curriculum Learning for Language Modeling	32

8.2	Models of cultural evolution in humans and AI systems	34
8.2.1	The effect of social network structure on collective innovation	34
8.2.2	Cultural evolution in population with heterogeneous and variable preferences	36
8.2.3	Interactions between intrinsically motivated goal-exploration processes and cumulative cultural evolution	37
8.2.4	Learning and Self-organization of Cultural Conventions Between Artificial Agents	37
8.2.5	Autotelic Reinforcement Learning in Multi-Agent Environments	38
8.2.6	The SocialAI School: Insights from Developmental Psychology Towards Artificial Socio-Cultural Agents	40
8.2.7	Value stability in Large Language Models	42
8.3	Ecological Artificial Intelligence	43
8.3.1	Research perspective: The Ecology of Open-Ended skill Acquisition	43
8.3.2	Evolution of plasticity and evolvability in variable environments	44
8.3.3	Open-ended recipe crafting through meta reinforcement learning	45
8.3.4	Evolving Reservoirs for Meta Reinforcement Learning	46
8.3.5	Emergence of Collective Open-Ended Exploration from Decentralized Meta-Reinforcement Learning	47
8.3.6	Cooperative control of environmental extremes by artificial intelligent agents	49
8.3.7	Eco-evolutionary Dynamics of Non-episodic Neuroevolution in Large Multi-agent Environments	49
8.4	Generative AI and educational technologies	52
8.4.1	Fostering curiosity and meta-cognition in children using LLM-based conversational agents	52
8.4.2	Supporting Qualitative Analysis with Large Language Models: Combining Codebook with GPT-3 for Deductive Coding	56
8.5	Curiosity-driven educational technologies	57
8.5.1	Machine Learning for Adaptive Personalization in Intelligent Tutoring Systems	57
8.5.2	Machine learning for adaptive cognitive training	58
8.5.3	ToGather : Interactive website to foster collaboration among stakeholders of school inclusion for pupils with neurodevelopmental disorders	60
8.5.4	Curious and therefore not overloaded : Study of the links between curiosity and cognitive load in learning mediated by immersive technologies	61
8.6	Curiosity-driven AI for assisted scientific discovery	64
8.6.1	Design of an Interactive Software for Automated Discovery in Complex Systems	64
8.6.2	Learning Sensorimotor Agency in Cellular Automata	64
8.6.3	Flow lenia: Mass conservation for the study of virtual creatures in continuous cellular automata	67
8.6.4	Exploration of Gene Regulatory Network Behaviors Using Automated Discovery Tools	68
9	Bilateral contracts and grants with industry	69
9.1	Bilateral contracts with industry	69
9.2	Bilateral contracts with industry	71
9.3	Bilateral Grants with Foundation	71
10	Partnerships and cooperations	71
10.1	International initiatives	71
10.1.1	Inria associate team not involved in an IIL or an international program	72
10.2	International research visitors	72
10.2.1	Visits of international scientists	72
10.2.2	Visits to international teams	73
10.2.3	Horizon Europe	73
10.3	National initiatives	74
10.3.1	Adaptiv'Math	75

11 Dissemination	75
11.1 Promoting scientific activities	75
11.1.1 Scientific events: organisation	75
11.1.2 Scientific events: selection	76
11.1.3 Journal	76
11.1.4 Invited talks	76
11.1.5 Leadership within the scientific community	77
11.1.6 Scientific expertise	77
11.1.7 Research administration	78
11.2 Teaching - Supervision - Juries	78
11.2.1 Teaching	78
11.2.2 Supervision	79
11.2.3 Juries	80
11.3 Popularization	80
11.3.1 Internal or external Inria responsibilities	80
11.3.2 Articles and contents	81
11.3.3 Education	81
11.3.4 Interventions	82
12 Scientific production	82
12.1 Major publications	82
12.2 Publications of the year	84
12.3 Other	89
12.4 Cited publications	89

Team FLOWERS

Creation of the Team: 2023 July 01

Keywords

Computer sciences and digital sciences

- 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
- A9.2. – Machine learning
- A9.4. – Natural language processing
- 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 members, visitors, external collaborators

Research Scientists

- Pierre-Yves Oudeyer [Team leader, INRIA, Senior Researcher, HDR]
- Clément Moulin-Frier [INRIA, Researcher]
- Eleni Nisioti [INRIA, Starting Research Position, until Aug 2023]
- H el ene Sauz eon [INRIA, Professor Detachement, HDR]

Faculty Members

- David Filliat [ENSTA, Professor, HDR]
- C ecile Mazon [UNIV BORDEAUX, Associate Professor]
- Mai Nguyen [ENSTA]

Post-Doctoral Fellows

- Louis Annabi [ENSTA, until Aug 2023]
- Cedric Colas [INRIA, Post-Doctoral Fellow]
- Eric Meyer [INRIA, Post-Doctoral Fellow, until Aug 2023]
- Marion Pech [INRIA, Post-Doctoral Fellow, from Apr 2023]
- Remy Portelas [INRIA, Post-Doctoral Fellow, until Feb 2023]

PhD Students

- Rania Abdelghani [EVIDENCEB]
- Maxime Adolphe [ONEPOINT]
- Thomas Carta [UNIV BORDEAUX]
- Marie-Sarah Desvaux [UNIV BORDEAUX, from Oct 2023]
- Mayalen Etcheverry [POIETIS, until Oct 2023]
- Gautier Hamon [INRIA]
- Tristan Karch [INRIA, until Apr 2023]
- Grgur Kovac [INRIA]
- Jeremy Perez [UNIV BORDEAUX, from Oct 2023]
- Matisse Poupard [CATIE, CIFRE]
- Julien Pourcel [INRIA, from Nov 2023]
- Thomas Rojat [GROUPE RENAULT, until Mar 2023]
- Cl ement Romic [HUGGING FACE SAS, CIFRE]
- Isabeau Saint-Supery [UNIV BORDEAUX]
- Maria Teodorescu [INRIA]
- Nicolas Yax [ENS Paris, from May 2023]

Technical Staff

- Jesse Lin [INRIA, Engineer, until Oct 2023]

Interns and Apprentices

- Richard Bornemann [INRIA, Intern, from Apr 2023 until Sep 2023]
- Marie-Sarah Desvoux [INRIA, Intern, from Feb 2023 until Jun 2023]
- Theo Goix [ENS Paris, Intern, from Jun 2023 until Aug 2023]
- Corentin Leger [BORDEAUX INP, from Mar 2023 until Aug 2023]
- Stephanie Mortemousque [INRIA, Intern, from Feb 2023 until Jun 2023]
- Mathieu Perie [INRIA, Apprentice, until Feb 2023]
- Julien Pourcel [INRIA, Intern, from Apr 2023 until Oct 2023]
- Valentin Strahm [INRIA, Intern, from Feb 2023 until Jun 2023]
- Alexandre Torres-Leguët [CENTRALE59, Intern, until Feb 2023]

Administrative Assistant

- Nathalie Robin [INRIA]

External Collaborator

- Didier Roy [N/A, until Jan 2023, Associate research scientist]

2 Overall objectives

Abstract: The Flowers project-team studies models of open-ended development and learning. These models are used as tools to help us understand better how children learn, as well as to build machines that learn like children, i.e. developmental artificial intelligence, with applications in educational technologies, assisted scientific discovery, video games, robotics and human-computer interaction.

Context: Great advances have been made recently in artificial intelligence concerning the topic of how autonomous agents can learn to act in uncertain and complex environments, thanks to the development of advanced Deep Reinforcement Learning techniques. These advances have for example led to impressive results with AlphaGo [186] or algorithms that learn to play video games from scratch [153, 133]. However, these techniques are still far away from solving the ambitious goal of lifelong autonomous machine learning of repertoires of skills in real-world, large and open environments. They are also very far from the capabilities of human learning and cognition. Indeed, developmental processes allow humans, and especially infants, to continuously acquire novel skills and adapt to their environment over their entire lifetime. They do so autonomously, i.e. through a combination of self-exploration and linguistic/social interaction with their social peers, sampling their own goals while benefiting from the natural language guidance of their peers, and without the need for an “engineer” to open and retune the brain and the environment specifically for each new task (e.g. for providing a task-specific external reward channel). Furthermore, humans are extremely efficient at learning fast (few interactions with their environment) skills that are very high-dimensional both in perception and action, while being embedded in open changing environments with limited resources of time, energy and computation.

Thus, a major scientific challenge in artificial intelligence and cognitive sciences is to understand how humans and machines can efficiently acquire world models, as well as open and cumulative repertoires of skills over an extended time span. Processes of sensorimotor, cognitive and social development are organized along ordered phases of increasing complexity, and result from the complex interaction between the brain/body with its physical and social environment. Making progress towards these

fundamental scientific challenges is also crucial for many downstream applications. Indeed, autonomous lifelong learning capabilities similar to those shown by humans are key requirements for developing virtual or physical agents that need to continuously explore and adapt skills for interacting with new or changing tasks, environments, or people. This is crucial for applications like assistive technologies with non-engineer users, such as robots or virtual agents that need to explore and adapt autonomously to new environments, adapt robustly to potential damages of their body, or help humans to learn or discover new knowledge in education settings, and need to communicate through natural language with human users, grounding the meaning of sentences into their sensorimotor representations.

The Developmental AI approach: Human and biological sciences have identified various families of developmental mechanisms that are key to explain how infants can acquire so robustly a wide diversity of skills [135, 152], in spite of the complexity and high-dimensionality of the body [93] and the open-endedness of its potential interactions with the physical and social environment. To advance the fundamental understanding of these mechanisms of development as well as their transposition in machines, the FLOWERS team has been developing an approach called *Developmental artificial intelligence*, leveraging and integrating ideas and techniques from developmental robotics ([202, 144, 100, 163], Deep (Reinforcement) Learning and developmental psychology. This approach consists in developing computational models that leverage advanced machine learning techniques such as intrinsically motivated Deep Reinforcement Learning, in strong collaboration with developmental psychology and neuroscience. In particular, the team focuses on models of intrinsically motivated learning and exploration (also called curiosity-driven learning), with mechanisms enabling agents to learn to represent and generate their own goals, self-organizing a learning curriculum for efficient learning of world models and skill repertoire under limited resources of time, energy and compute. The team also studies how autonomous learning mechanisms can enable humans and machines to acquire and develop grounded and culturally shared language skills, using neuro-symbolic architectures for learning structured representations and handling systematic compositionality and generalization.

Our fundamental research is organized along three strands:

Strand 1: Lifelong autonomous learning in machines. Understanding how developmental mechanisms can be functionally formalized/transposed in machines and explore how they can allow these machines to acquire efficiently open-ended repertoires of skills through self-exploration and social interaction.

Strand 2: Computational models as tools to understand human development in cognitive sciences. The computational modelling of lifelong 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 interactions 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 [18, 31].

Strand 3: Applications. Beyond leading to new theories and new experimental paradigms to understand human development in cognitive science, as well as new fundamental approaches to developmental machine learning, the team explores how such models can find applications in robotics, human-computer interaction, multi-agent systems, automated discovery and educational technologies. In robotics, the team studies how artificial curiosity combined with imitation learning can provide essential building blocks allowing robots to acquire multiple tasks through natural interaction with naive human users, for example in the context of assistive robotics. The team also studies how models of curiosity-driven learning can be transposed in algorithms for intelligent tutoring systems, allowing educational software to incrementally and dynamically adapt to the particularities of each human learner, and proposing personalized sequences of teaching activities.

3 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 or training 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 who tries to progress efficiently, and the teacher here has to use as efficiently the limited time and motivational resources of the learner. Several results from psychology [92] and neuroscience [122] 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 [114] for fostering the individual engagement in learning activities. A such view is also relevant for reeducation issues where inter-individual variability, and thus intervention personalization are challenges of the same magnitude as those for education of children.

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” [197]. 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 [129], 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 [144] [202]. The approach of developmental robotics consists in importing and implementing concepts and mechanisms from developmental psychology [152], cognitive linguistics [113], and developmental cognitive neuroscience [134] 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 [96] [165], grounding [127], situatedness [188], self-organization [192] [164], enaction [198],

and incremental learning [105].

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 [166]. 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.

Internal guiding mechanisms In infant development, one observes a progressive increase of the complexity of activities with an associated progressive increase of capabilities [152], 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” [88] [119]. 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 [92] [114] [116]. Futher, the exploration of critical role of intrinsic motivation as lever of cognitive development for all and for all ages is today expanded to several fields of research, closest to its original study, special education or cognitive aging, and farther away, neuropsychological clinical research. 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 [115] [130] [183]. Based on this, a number of researchers have began in the past few years to build computational implementation of intrinsic motivation [166] [167] [178] [90] [131] [146] [179]. 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 [167, 166, 168, 177]. 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 [108] [128] 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.

Socially Guided and Interactive Learning Social guidance is as important as intrinsic motivation in the cognitive development of human babies [152]. 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 [87]. 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 [193, 95], motivated by the various mechanisms that allow humans to socially guide a robot [175]. 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 [193, 138, 147].

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 [157, 158]. We also work on developing perceptual capabilities by gradually segmenting the perceptual space and identifying objects and their characteristics through interaction with the user [145] and robots experiments [132]. 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 [143].

Exploration mechanisms are combined with research in the following directions:

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.

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).

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.

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.

Emergence of social behavior in multi-agent populations FLOWERS studies how populations of interacting learning agents can collectively acquire cooperative or competitive strategies in challenging simulated environments. This differs from "Social learning and guidance" presented above: instead of studying how a learning agent can benefit from the interaction with a skilled agent, we rather consider here how social behavior can spontaneously emerge from a population of interacting learning agents. We focus on studying and modeling the emergence of cooperation, communication and cultural innovation based on theories in behavioral ecology and language evolution, using recent advances in multi-agent reinforcement learning.

Cognitive variability across Lifelong development and (re)educational Technologies Over the past decade, the progress in the field of curiosity-driven learning generates a lot of hope, especially with regard to a major challenge, namely the inter-individual variability of developmental trajectories of learning, which is particularly critical during childhood and aging or in conditions of cognitive disorders. With the societal purpose of tackling of social inequalities, FLOWERS deals to move forward this new research avenue by exploring the changes of states of curiosity across lifespan and across neurodevelopmental conditions (neurotypical vs. learning disabilities) while designing new educational or rehabilitative technologies for curiosity-driven learning. The information gaps or learning progress, and their awareness are the core mechanisms of this part of research program due to high value as brain fuel by which the individual's internal intrinsic state of motivation is maintained and leads him/her to pursue his/her cognitive efforts for acquisitions /rehabilitations. Accordingly, a main challenge is to understand these mechanisms in order to draw up supports for the curiosity-driven learning, and then to embed them into (re)educational technologies. To this end, two-ways of investigations are carried out in real-life setting (school, home, work place etc): 1) the design of curiosity-driven interactive systems for learning and their effectiveness study ; and 2) the automated personalization of learning programs through new algorithms maximizing learning progress in ITS.

<https://red-radar.inria.fr/document/FLOWERS-RA-2023/domain>

4 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*. Another example is the study of the role of curiosity in learning in the elderly, with a view to assessing its positive value against the cognitive aging as a protective ingredient (i.e. Industrial project with Onepoint and CuriousTECH associate team with M. Fernandes from the Cognitive neuroscience Lab of the University of Waterloo).

Personal and lifelong learning assistive agents Many indicators show that the arrival of personal assistive agents in everyday life, ranging from digital assistants to robots, will be a major fact of the 21st century. These agents 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 agents will have to evolve in unpredictable environments and learn new skills in a lifelong manner while interacting with non-engineer humans, which is out of reach of current technology. In this context, the refoundation of intelligent systems that developmental AI 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 when humans are in-the-loop. 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.

Educational technologies that foster curiosity-driven and personalized learning. 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. The challenges of the school of the 21st century, and in particular to produce conditions for active learning that are personalized to the student's motivations, are challenges shared with other applied fields. Special education for children with special needs, such as learning disabilities, has long recognized the difficulty of personalizing contents and pedagogies due to the great variability between and within medical conditions. More remotely, but not so much, cognitive rehabilitative carers are facing the same challenges where today they propose standardized cognitive training or rehabilitation programs but for which the benefits are modest (some individuals respond to the programs, others respond little or not at all), as they are highly subject to inter- and intra-individual variability. The curiosity-driven technologies for learning and STIs could be a promising avenue to address these issues that are common to (mainstream and specialized) education and cognitive rehabilitation.

Automated discovery in science. Machine learning algorithms integrating intrinsically-motivated goal exploration processes (IMGEPs) with flexible modular representation learning are very promising directions to help human scientists discover novel structures in complex dynamical systems, in fields ranging from biology to physics. The *automated discovery* project lead by the FLOWERS team aims to boost the efficiency of these algorithms for enabling scientist to better understand the space of dynamics of bio-physical systems, that could include systems related to the design of new materials or new drugs with applications ranging from regenerative medicine to unraveling the chemical origins of life. As an example, Grizou et al. [123] recently showed how IMGEPs can be used to automate chemistry experiments addressing fundamental questions related to the origins of life (how oil droplets may self-organize into protocellular structures), leading to new insights about oil droplet chemistry. Such methods can be applied to a large range of complex systems in order to map the possible self-organized structures. The *automated discovery* project is intended to be interdisciplinary and to involve potentially non-expert end-users from a variety of domains. In this regard, we are currently collaborating with Poietis (a bio-printing company) and Bert Chan (an independant researcher in artificial life) to deploy our algorithms. To encourage the adoption of our algorithms by a wider community, we are also working on an interactive software which aims to provide tools to easily use the automated exploration algorithms (e.g. curiosity-driven) in various systems.

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 of 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.

5 Social and environmental responsibility

5.1 Footprint of research activities

AI is a field of research that currently requires a lot of computational resources, which is a challenge as these resources have an environmental cost. In the team we try to address this challenge in two ways:

- by working on developmental machine learning approaches that model how humans manage to learn open-ended and diverse repertoires of skills under severe limits of time, energy and compute: for example, curiosity-driven learning algorithms can be used to guide agent's exploration of their environment so that they learn a world model in a sample efficient manner, i.e. by minimizing the number of runs and computations they need to perform in the environment;
- by monitoring the number of CPU and GPU hours required to carry out our experiments. For instance, our work [11] used a total of 2.5 cpu years. More globally, our work uses large scale computational resources, such as the Jean Zay supercomputer platform, in which we use several hundred thousands hours of GPU and CPU each year.

5.2 Impact of research results

Our research activities are organized along two fundamental research axis (models of human learning and algorithms for developmental machine learning) and one application research axis (involving multiple domains of application, see the Application Domains section). This entails different dimensions of potential societal impact:

- **Towards autonomous agents that can be shaped to human preferences and be explainable** We work on reinforcement learning architectures where autonomous agents interact with a social partner to explore a large set of possible interactions and learn to master them, using language as a key communication medium. As a result, our work contributes to facilitating human intervention in the learning process of agents (e.g. digital assistants, video games characters, robots), which we believe is a key step towards more explainable and safer autonomous agents.
- **Reproducibility of research:** By releasing the codes of our research papers, we believe that we help efforts in reproducible science and allow the wider community to build upon and extend our work in the future. In that spirit, we also provide clear explanations on the statistical testing methods when reporting the results.
- **Digital transformation and Competences' challenges facing schools in the 21st century.** We expect our findings to inform the broader societal challenges inherent to the School of the 21st Century, ranging from helping children (and their teachers) to develop cross-domain skills for

learning such as curiosity and meta-cognition, while improving inclusivity in schools (learners with disabilities, especially cognitive disabilities) as well as promoting lifelong learning in older adults (successful aging), using cognitive-based research findings.

- **AI and personalized educational technologies to reduce inequalities due to neurodiversity** The Flowers team develops AI technologies aiming to personalize sequences of educational activities in digital educational apps: this entails the central challenge of designing systems which can have equitable impact over a diversity of students and reduce inequality in academic achievement. Using models of curiosity-driven learning to design AI algorithms for such personalization, we have been working to enable them to be positively and equitably impactful across several dimensions of diversity: for young learners or for aging populations; for learners with low initial levels as well as for learners with high initial levels; for "normally" developing children and for children with developmental disorders; and for learners of different socio-cultural backgrounds (e.g. we could show in the KidLearn project that the system is equally impactful along these various kinds of diversities).
- **Health: Bio-printing** The Flowers team is studying the use of curiosity-driven exploration algorithm in the domain of automated discovery, enabling scientists in physics/chemistry/biology to efficiently explore and build maps of the possible structures of various complex systems. One particular domain of application we are studying is bio-printing, where a challenge consists in exploring and understanding the space of morphogenetic structures self-organized by bio-printed cell populations. This could facilitate the design and bio-printing of personalized skins or organoids for people that need transplants, and thus could have major impact on the health of people needing such transplants.
- **Tools for human creativity and the arts** Curiosity-driven exploration algorithms could also in principle be used as tools to help human users in creative activities ranging from writing stories to painting or musical creation, which are domains we aim to consider in the future, and thus this constitutes another societal and cultural domain where our research could have impact.
- **Education to AI** As artificial intelligence takes a greater role in human society, it is of foremost importance to empower individuals with understanding of these technologies. For this purpose, the Flowers lab has been actively involved in educational and popularization activities, in particular by designing educational robotics kits that form a motivating and tangible context to understand basic concepts in AI: these include the Inirobot kit (used by >30k primary school students in France ([see](#))) and the Poppy Education kit ([see](#)) now supported by the Poppy Station educational consortium ([see](#)).
- **Health: optimization of intervention strategies during pandemic events** Modelling the dynamics of epidemics helps proposing control strategies based on pharmaceutical and non-pharmaceutical interventions (contact limitation, lock down, vaccination, etc). Hand-designing such strategies is not trivial because of the number of possible interventions and the difficulty to predict long-term effects. This task can be cast as an optimization problem where state-of-the-art machine learning algorithms such as deep reinforcement learning, might bring significant value. However, the specificity of each domain – epidemic modelling or solving optimization problem – requires strong collaborations between researchers from different fields of expertise. Due to its fundamental multi-objective nature, the problem of optimizing intervention strategies can benefit from the goal-conditioned reinforcement learning algorithms we develop at Flowers. In this context, we have developed EpidemioOptim, a Python toolbox that facilitates collaborations between researchers in epidemiology and optimization ([see](#)).

6 Highlights of the year

- **Open-ended learning and autotelic AI with large language models:** The team continued to lay the foundations of autotelic AI [111, 109], i.e. the science studying mechanisms enabling artificial agents to learn to represent and sample their own goals and achieve open-ended learning. In

particular, this year we focused on studying how to build **autotelic AI systems with large language models (LLMs)**. In particular, we studied how LLMs can be used as cognitive tools enabling **creative generation of abstract goals**, either in classical interactive environments such as in the LMA3 system [45] (collab. with M-A. Côté and E. Yuan at Microsoft Research Montreal), or in code/programming environments where **LLM-based autotelic agents learn by self-improving their coding abilities**, e.g. in the CodePlay system [64] (collab. with M. Bowers at MIT) or in the ACES system [77]. This also set the ground for our participation in the new LLM4Code Inria Challenge. While doing this work, we also developed techniques and studied **how language models can be grounded in interactive environments using online RL**, such as in the GLAM system that enables to update and align LLMs to the particular dynamics of the environment [44] (collab. with O. Sigaud at Sorbonne Univ., S. Lamprier at Univ. Angers; T. Wolf at HuggingFace). This work was associated to the development of the **Lamorel** python library, designed for researchers eager to use Large Language Models (LLMs) in interactive environments (e.g. RL setups).

- **Models of cultural evolution in humans and AI systems:** As **generative AI systems become powerful cultural transmission technologies** that influence human cultural evolution in important ways, and can also have their own cultural processes through machine-machine large scale interaction, the study of the dynamics of cultural processes in populations of AI systems/humans becomes crucial. We continued our work in this perspective through computational models of **innovation dynamics in groups of agents** [159] (associated with a collaboration with I. Mommenejad at Microsoft Research), of the **interaction of autotelic learning and self-organization of communication conventions** [52], and of the **formation of language protocols in groups of agents** [70]. We also studied how LLMs can encode superpositions of socio-cultural perspectives, for example value systems, and developed tools to **analyze the robustness and controllability for the expression of these socio-cultural abilities** [74]. We also contributed to a general discussion of a new field studying "machine culture" in [36]. We started a new collaboration with M. Derex from IAST Toulouse, a leading researcher in the field of Cultural Evolution, in the context of the co-direction of the PhD thesis of Jeremy Perez (started in October 2023).
- **Eco-evolutionary AI and meta-reinforcement learning:** We developed an ecological research perspective on AI, highlighting the **interactions between environmental, adaptive, multi-agent and cultural dynamics in sculpting intelligence**. This led to the proposition of a detailed conceptual framework for studying these interactions, summarized in this [blog post](#) and in the HDR thesis of C. Moulin-Frier [154]. This new research perspective is associated with several international and national collaborations: with I. Mommenejad from Microsoft Research (USA), with M. Sanchez-Fibla and R. Solé from the Univ. Pompeu Fabra (Spain), with X. Hinaut from Inria Mnemosyne, with F. d'Errico and L. Doyon from the Univ. of Bordeaux (France) and with M. Derex from IAST Toulouse (France). Several conference papers have been published in this context in 2023. In [48], we studied the **eco-evolutionary dynamics of non-episodic neuroevolution in large multi-agent environments**. In [60], we studied the **emergence of collective open-ended exploration** from decentralized meta-reinforcement learning. In ??, we **extended the framework of autotelic reinforcement learning to multi-agent environments**. In [75], we introduced an eco-devo artificial agent integrating **reservoir computing and meta reinforcement learning**. In [53], we studied the dynamics of **niche construction** in adaptable populations evolving in diverse environments.
- **Generative AI and educational technologies:** We continued key projects studying the use of generative AI in education. First, we published **one of the first international field studies investigating the pedagogical use of LLMs (here GPT-3) in real classrooms**: in [33] (collab. with E. Yuan and T. Wang from MSR Montreal, and with Y-H. Wang), we showed that when used appropriately, GPT-3 can be used to build conversational agents that train efficiently curious question asking in primary school children, enabling to scale up a pedagogical approach for training curiosity skills we developed previously [86]. Also, we studied the **use of LLMs (GPT-3) to partially automate qualitative analysis methods in social sciences** [56] (collab. Z. Xiao, V. Liao, E. Yuan from MSR Montreal), opening new perspectives for studying qualitatively large text corpuses

or verbal data from psychology or educational experiments. Finally, we developed a conceptual framework to think about the **opportunities and challenges associated to using generative AI in the classroom**, and in particular asking how this could be done by enabling children to keep and develop active learning skills [59]. We here identified that one key challenge is to improve the AI literacy of both children and their teachers: with this aim in mind, we started designing a **pedagogical video series explaining in accessible ways various socio-technical dimensions of large language models** (with A. Torres-Leguet). This series, available freely on the [web](#) with a Creative Commons licence, has already been reused in multiple contexts such as the mooc [AI4T](#) (AI for teachers).

- **Meta-cognition in Curiosity- driven educational technologies** We developed several projects leveraging fundamental cognitive science studies of curiosity and meta-cognition to design educational interventions that either directly train these skills, or stimulate them to train other related cognitive skills ranging from maths to languages or other transverse skills like attention, and did this for diverse populations ranging from neurotypical to neurodiverse school children, to healthy young adults and aging populations.

At a fundamental level, we studied the **beneficial role of curiosity on route memory in children**, within a new virtual reality experimental paradigm [79], and in the context of our collaboration with Myra Fernandes at Univ. Waterloo (associated team [CuriousTech](#)). To refine the understanding of metacognitive awareness of one's own learning progress and its role as curiosity-boosters ([31], we designed an **educational software (4MC project) that aims to train curiosity through the practice of meta-cognitive skills in school children**, and pilot studies led to very encouraging results [41]. Also, as a follow-up of our systematic review on the **interactions between curiosity and cognitive load in immersive learning technologies**, we started a field study with 180 university students to test hypotheses about the links between this interaction and learning performances (involving the collaboration with Pr. A. Tricot from University of Montpellier and the [CATIE](#) company).

Leveraging the Learning Progress theory of human curiosity we developed in the past [162] [18], which led us to develop the ZPDES algorithm for personalizing sequences of exercises that foster learning efficiency and motivation [107], we continued studying how **ZPDES can be used to personalize training of attention skills in both young adults and aging populations** (paper under writing, collab. with D. Bavelier at Univ. of Geneva). Related to this project, we wrote a systematic review of the use of AI in cognitive training technologies [2]. We also finalized the analysis of a large-scale experimental study using ZPDES in the context of training maths skills in primary schools, with a focus on the dual impact on learning efficiency and motivation on one hand, and a focus on adding choice possibilities on the other hand, showing positive results of the approach in comparison with hand-made pedagogical sequences. Through a collaboration with the [EvidenceB](#) company and support from the French Ministry of Education, the **ZPDES personalization system was also deployed in the large-scale [AdaptivMaths](#) system now available in all French primary schools (> 68k classrooms)**. EvidenceB further used ZPDES in the MIA seconde system aimed for training high-school students in maths and french.

Finally, in the [Together](#) project, we also experimented a system aimed at **stimulating communication among stakeholders around neurodiverse children in schools** (college level), and in particular trying to foster mutual curiosity among them while taking account possible cross-cultural differences in French and Belgium Schools [55]. This was associated with a systematic review on methods to collaborate and co-educate students with special educational needs [51].

- **Curiosity-driven AI for assisted scientific discovery**: We continued studying how curiosity-driven AI algorithms can enable scientists (physicists, chemists, biologists, etc) explore and map the space of self-organized behaviours in diverse complex systems [69]. In particular, through a collaboration with M. Levin at Tufts University, we studied **how autotelic AI systems (IMGEP algorithms) can enable cost effective discovery of diverse sophisticated and robust behaviors in gene regulatory networks** [73]. This project was associated to the development of the [ADTool](#) software aiming to facilitate the use of such exploration algorithms to scientists with various backgrounds, as well as with the development of [SBMLtoODEjax](#) [61], which aims to automatically parse and

convert SBML models into python models written end-to-end in JAX, enabling fast and easy to use biological models in ML experiments. In another project, we continued our work using exploration algorithms to study **self-organized structures in continuous CAs like Lenia**, enabling to discover self-organization of forms of primitive agency, as described in this [blog post](#). In this context, we designed a new continuous CA called **Flow Lenia**, combining mechanisms for mass conservation and localized embedding of adaptive update rules: this strongly facilitates self-organization of localized patterns and opens possibilities for the self-organization of evolutionary processes. The associated paper, in collaboration with Bert Chan at Google Brain, obtained the **Best paper award at Alife 2023** in Tokyo [54].

- **Workshop/symposium organization:** Laetitia Teodorescu and Cédric Colas have been organizers of the Intrinsically-Motivated and Open-Ended Learning workshop at NeurIPS 2023, <https://imol-workshop.github.io/>; Mayalen Etcheverry was co-organizer of the Workshop on Agent Learning in Open-Endedness (ALOE) at Neurips 2023, <https://sites.google.com/view/aloe2023/home>; H. Sauzéon has been member of the scientific committee of the first "Scientific Day of Gerontopole of New Aquitania", April 2023, Limoges; Pierre-Yves Oudeyer was co-organizer of the Life, Structure and Cognition symposium on Evolution and Learning, at IHES, <https://indico.math.cnrs.fr/event/9963/>, as well as of the Curiosity, Complexity and Creativity conference 2023 at Columbia University, NY, US, <https://zuckermaninstitute.columbia.edu/cc-event>.
- **International Research Visits:** Rania Abdelghani visited the lab of Celeste Kidd at Univ. Berkeley to develop a new project studying how children understand and (mis)use large language models in educational settings. Marion Pech, Matisse Poupart and Maxime Adolphe visited Myra Fernandes's lab and Edith Law's lab at Univ. Waterloo in the context of associated team **CuriousTech**.
- **Collaborations with industry:** The team continued collaborations with various actors in the industry, including HuggingFace, EvidenceB, CATIE, Microsoft Research, Ubisoft, OnePoint.

6.1 Awards

Rémy Portelas obtained the **Best PhD award from University of Bordeaux**, category "Special prize of the jury", for his thesis entitled "Automatic Curriculum Learning for Developmental Machine Learners" [172]

Erwan Plantec, Gautier Hamon, Mayalen Etcheverry, Bert Chan, Pierre-Yves Oudeyer and Clément Moulin-Frier obtained the **Best paper award at Alife 2023** in Tokyo for the paper "Flow-Lenia: Towards open-ended evolution in cellular automata through mass conservation and parameter localization" [54].

6.2 PhD defenses

- Tristan Karch defended his PhD thesis on "Towards Social Autotelic Artificial Agents - Formation and Exploitation of Cultural Conventions in Autonomous Embodied Artificial Agents" [70] on May 11th, 2023.
- Mayalen Etcheverry defended her PhD thesis on "Curiosity-driven AI for Science: Automated Discovery of Self-Organized Structures" [69] on November 16th, 2023.
- Laetitia Teodorescu defended her PhD thesis on "Endless minds most beautiful: building open-ended linguistic autotelic agents with deep reinforcement learning and language models" [71] on November 20th, 2023.

7 New software, platforms, open data

7.1 New software

7.1.1 SocialAI

Name: SocialAI: Benchmarking Socio-Cognitive Abilities in Deep Reinforcement Learning Agents

Keywords: Artificial intelligence, Deep learning, Reinforcement learning, Large Language Models

Functional Description: Source code for the paper <https://arxiv.org/abs/2107.00956>.

A suite of environments for testing socio-cognitive abilities of artificial agents. Environments can be used in the multimodal setting (suitable for RL agents) and in the pure text setting (suitable for Large Language Model-based agents). Also contains RL and LLM baselines.

URL: <https://gitlab.inria.fr/gkovac/act-and-speak>

Contact: Grgur Kovac

7.1.2 AutoDisc

Keyword: Complex Systems

Functional Description: AutoDisc is a software built for automated scientific discoveries in complex systems (e.g. self-organizing systems). It can be used as a tool to experiment automated discovery of various systems using exploration algorithms (e.g. curiosity-driven). Our software is fully Open Source and allows user to add their own systems, exploration algorithms or visualization methods.

URL: <https://gitlab.inria.fr/cromac/AutomatedDiscoveryTool>

Contact: Clément Romic

7.1.3 Kids Ask

Keywords: Human Computer Interaction, Cognitive sciences

Functional Description: Kids Ask is a web-based educational platform that involves an interaction between a child and a conversational agent. The platform is designed to teach children how to generate curiosity-based questions and use them in their learning in order to gain new knowledge in an autonomous way.

URL: <https://github.com/RaniaAbdelghani/KidAsk>

Contact: Rania Abdelghani

7.1.4 ToGather

Keywords: Education, Handicap, Environment perception

Scientific Description: With participatory design methods, we have designed an interactive website application for educational purposes. This application aims to provide interactive services with continuously updated content for the stakeholders of school inclusion of children with specific educational needs.

Functional Description: Website gathering information on middle school students with neurodevelopmental disorders. Authentication is required to access the site's content. Each user can only access the student file(s) of the young person(s) they are accompanying. A student file contains 6 tabs, in which each type of user can add, edit or delete information: 1. Profile: to quickly get to know the student 2. Skills: evaluation at a given moment and evolution over time 3. Compendium of tips: includes psycho-educational tips 4. Meetings: manager and reports 5. News: share information over time 6. Contacts: contact information for stakeholders The student only has the right to view information about him/her.

Publication: <https://hal.inria.fr/hal-03436355/>

Contact: Cécile Mazon

Participants: Isabeau Saint-Supery, Cécile Mazon, Eric Meyer, Hélène Sauzéon

7.1.5 mc_training

Name: Platform for metacognitive training

Keywords: Human Computer Interaction, Education

Functional Description: This is a web platform for children between 9 and 11 years old, designed to help children practice 4 metacognitive skills that are thought to be involved in curiosity-driven learning: - the ability to identify uncertainties - the ability to generate informed hypotheses - the ability to ask questions - the ability to evaluate the value of a preconceived inference.

Children work on a reading-comprehension tasks and, for each of these skills, the platform offers help through a "conversation" with conversational agents that give instructions to perform the task, with respect to every skill, and can give suggestions if the child asks for it.

Contact: Rania Abdelghani

7.1.6 Evolution of adaptation mechanisms in complex environments

Name: Plasticity and evolvability under environmental variability: the joint role of fitness-based selection and niche-limited competition

Keywords: Evolution, Ecology, Dynamic adaptation

Functional Description: This is the code accompanying our paper Plasticity and evolvability under environmental variability: the joint role of fitness-based selection and niche-limited competition" which is to be presented at the Gecco 2022 conference.

In this work we have studied the evolution of a population of agents in a world where the fitness landscape changes with generations based on climate function and a latitudinal model that divides the world in different niches. We have implemented different selection mechanisms (fitness-based selection and niche-limited competition).

The world is divided into niches that correspond to different latitudes and whose state evolves based on a common climate function.

We model the plasticity of an individual using tolerance curves originally developed in ecology. Plasticity curves have the form of a Gaussian the capture the benefits and costs of plasticity when comparing a specialist (left) with a generalist (right) agent.

The repo contains the following main elements :

folder source contains the main functionality for running a simulation
 scripts/run/reproduce_gecco.py can be used to rerun all simulations in the paper
 scripts/evaluate contains scripts for reproducing figures. reproduce_figures.py will produce all figures (provided you have already run scripts/run/reproduce_gecco.py to generate the data)
 folder projects contains data generated from running a simulation
 How to run To install all package dependencies you can create a conda environment as:

```
conda env create -f environment.yml
```

All script executions need to be run from folder source. Once there, you can use simulate.py, the main interface of the codebase to run a simulation, For example:

```
python simulate.py --project test_stable --env_type stable --num_gens 300 --capacity 1000 --num_niches 10 --trials 10 --selection_type NF --climate_mean_init 2
```

will run a simulation with an environment with a climate function whose state is constantly 2 consisting of 100 niches for 300 generations and 10 independent trials. The maximum population size will be 1000*2 and selection will be fitness-based (higher fitness means higher chances of reproduction) and niche limited (individuals reproduce independently in each niche and compete only within a niche),

You can also take a look at scripts/run/reproduce_gecco.py to see which flags were used for the simulations presented in the paper.

Running all simulations requires some days. You can instead download the data produced by running `scripts/run/reproduce_gecco.py` from this google folder and unzip them under the projects directory.

URL: <https://github.com/eleninisioti/ClimateAndLearning>

Contact: Eleni Nisioti

7.1.7 SAPIENS

Name: SAPIENS: Structuring multi-Agent toPology for Innovation through ExperieNce Sharing

Keywords: Reinforcement learning, Multi-agent

Functional Description: SAPIENS is a reinforcement learning algorithm where multiple off-policy agents solve the same task in parallel and exchange experiences on the go. The group is characterized by its topology, a graph that determines who communicates with whom.

All agents are DQNs and exchange experiences have the form of transitions from their replay buffers.

Using SAPIENS we can define groups of agents that are connected with others based on a a) fully-connected topology b) small-world topology c) ring topology or d) dynamic topology.

Install required packages You can install all required python packages by creating a new conda environment containing the packages in `environment.yml`:

```
conda env create -f environment.yml
```

And then activating the environment:

```
conda activate sapiens
```

Example usages Under notebooks there is a Jupyter notebook that will guide you through setting up simulations with a fully-connected and a dynamic social network structure for solving Wordcraft tasks. It also explains how you can access visualizations of the metrics produced during the

Reproducing the paper results Scripts under the scripts directory are useful for reproducing results and figures appearing in the paper.

With `scripts/reproduce_runs.py` you can run all simulations presented in the paper from scratch.

This file is useful for looking at how the experiments were configured but better avoid running it: simulations will run locally and sequentially and will take months to complete.

Instead, you can access the data files output by simulations on this online repo.

Download this zip file and uncompress it under the projects directory. This should create a `projects/paper_done` sub-directory.

You can now reproduce all visualization presented in the paper. Run:

```
python scripts/reproduce_visuals.py
```

This will save some general plots under `visuals`, while project-specific plots are saved under the corresponding project in `projects/paper_done`

URL: <https://github.com/eleninisioti/SAPIENS>

Contact: Eleni Nisioti

7.1.8 architect-builder-abig

Name: Architect-Builder Iterated Guiding

Keyword: Artificial intelligence

Functional Description: Codebase for the paper Learning to guide and to be guided in the Architect-Builder Problem

ABIG stands for Architect-Builder Iterated Guiding and is an algorithmic solution to the Architect-Builder Problem. The algorithm leverages a learned model of the builder to guide it while the builder uses self-imitation learning to reinforce its guided behavior.

URL: <https://github.com/flowersteam/architect-builder-abig>

Contact: Tristan Karch

7.1.9 EAGER

Name: Exploit question-Answering Grounding for effective Exploration in language-conditioned Reinforcement learning

Keywords: Reinforcement learning, Language, Question Generation Question Answering, Reward shaping

Functional Description: A novel QG/QA framework for RL called EAGER In EAGER, an agent reuses the initial language goal sentence to generate a set of questions (QG): each of these self-generated questions defines an auxiliary objective. Here, generating a question consists in masking a word of the initial language goal. Then the agent tries to answer these questions (guess the missing word) only by observing its trajectory so far. When it manages to answer a question correctly (QA) it obtains an intrinsic reward proportional to its confidence in the answer. The QA module is trained using a set of successful example trajectories. If the agent follows a path too different from correct ones at some point in its trajectory, the QA module will not answer the question correctly, resulting in zero intrinsic reward. The sum of all the intrinsic rewards measures the quality of a trajectory in relation to the given goal. In other words, maximizing this intrinsic reward incentivizes the agent to produce behaviour that unambiguously explains various aspects of the given goal.

URL: <https://github.com/flowersteam/EAGER>

Contact: Thomas Carta

7.1.10 IMGC-MARL

Name: Intrinsically Motivated Goal-Conditioned Reinforcement Learning in Multi-Agent Environments

Keywords: Reinforcement learning, Multi-agent, Curiosity

Functional Description: This repo contains the code base of the paper Intrinsically Motivated Goal-Conditioned Reinforcement Learning in Multi-Agent Environments.

In this work, we have studied the importance of the alignment of goals in the training of intrinsically motivated agents in the multi agent goal conditioned RL case. We also proposed a simple algorithm called goal coordination game which allows such agent to learn, in a completely decentralized/selfish way, to communicate in order to align their goal.

The repository contains the code to reproduce the results of the paper. Which includes a custom RL environment (using SimplePlayground "game engine"), model used (architecture + hyperparameters) and custom training (mostly based on RLlib) to train both the model and the communication. We also provide the scripts for the training of every condition we test and notebook to study the results.

URL: <https://github.com/Reytuag/imgc-marl>

Contact: Gautier Hamon



Figure 1: 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.

7.1.11 Flow-Lenia

Name: Flow Lenia: Mass conservation for the study of virtual creatures in continuous cellular automata

Keywords: Cellular automaton, Self-organization

Functional Description: This repo contains the code to run the Flow Lenia system which is a continuous parametrized cellular automaton with mass conservation. This work extends the classic Lenia system with mass conservation and allows to implement new feature like local parameter, environment components etc

Several declination of the system (1 or several channels etc) are available

Please refer to the associated paper for the details of the system

Implemented in JAX

URL: <https://github.com/erwanplantec/FlowLenia>

Contact: Gautier Hamon

7.1.12 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 1). 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.

URL: <https://flowers.inria.fr/research/kidlearn/>

Contact: Benjamin Clement

7.1.13 cognitive-testbattery

Name: Cognitive test battery of human attention and memory

Keywords: Open Access, Cognitive sciences

Scientific Description: Cognitive test batteries are widely used in diverse research fields, such as cognitive training, cognitive disorder assessment, or brain mechanism understanding. Although they need flexibility according to the objectives of their usage, most of the test batteries are not available as open-source software and not be tuned by researchers in detail. The present study introduces an open-source cognitive test battery to assess attention and memory, using a javascript library, p5.js. Because of the ubiquitous nature of dynamic attention in our daily lives, it is crucial to have tools for its assessment or training. For that purpose, our test battery includes seven cognitive tasks (multiple-objects tracking, enumeration, go/no-go, load-induced blindness, task-switching, working memory, and memorability), common in cognitive science literature. By using the test battery, we conducted an online experiment to collect the benchmark data. Results conducted on two separate days showed the high cross-day reliability. Specifically, the task performance did not largely change with the different days. Besides, our test battery captures diverse individual differences and can evaluate them based on the cognitive factors extracted from latent factor analysis. Since we share our source code as open-source software, users can expand and manipulate experimental conditions flexibly. Our test battery is also flexible in terms of the experimental environment, i.e., it is possible to experiment either online or in a laboratory environment.

Functional Description: The evaluation battery consists of 6 cognitive activities (serious games: multi-object tracking, enumeration, go/no-go, Corsi, load-induced blindness, taskswitching, memorability). Easily deployable as a web application, it can be re-used and modified for new experiments. The tool is documented in order to facilitate the deployment and the analysis of results.

URL: <https://github.com/flowersteam/cognitive-testbattery>

Publication: hal-03723887

Contact: Maxime Adolphe

Participants: Pierre-yves Oudeyer, H el ene Sauz eon, Masataka Sawayama, Maxime Adolphe

7.1.14 LLM_stability

Keywords: Artificial intelligence, Deep learning, Large Language Models

Functional Description: Source code for the paper <https://arxiv.org/abs/2307.07870>

Code enabling systematic evaluation of Large Language Models with various psychology questionnaires in different contexts, e.g. following conversations on different topics.

URL: https://gitlab.inria.fr/gkovac/llm_perspectives

Contact: Grgur Kovac

7.1.15 Sensorimotor-lenia

Keywords: Cellular automaton, Gradient descent, Curriculum Learning

Functional Description: Source code for the search of sensorimotor agency in cellular automata associated to this blogpost <https://developmentalsystems.org/sensorimotor-lenia/>. The code allows to find rules in the cellular automata Lenia (through gradient descent, curriculum learning and diversity search) that lead to the self-organization of moving agents robust to perturbation by obstacles.

URL: <https://github.com/flowersteam/sensorimotor-lenia-search>

Contact: Gautier Hamon

7.1.16 MetaIPPO

Keywords: Reinforcement learning, Exploration

Functional Description: Code for the paper "Emergence of collective open-ended exploration from Decentralized Meta-Reinforcement learning" <https://arxiv.org/pdf/2311.00651.pdf>

We train two decentralized agents together on an open ended tasks space to study the emergence of collective exploration behaviors. Our agents are able to generalize to novel objects and tasks, as well as an essentially open ended setting.

URL: <https://github.com/RBorn02/MetaIPPO>

Contact: Gautier Hamon

7.1.17 Lamorel

Keywords: Large Language Models, Reinforcement learning, Distributed computing

Functional Description: Lamorel is a Python library designed for people eager to use Large Language Models (LLMs) in interactive environments (e.g. RL setups).

URL: <https://github.com/flowersteam/lamorel>

Publication: hal-03970122

Contact: Clément Romac

7.1.18 GLAM

Name: Grounding LAnguage Models

Keywords: Large Language Models, Reinforcement learning

Scientific Description: Recent works successfully leveraged Large Language Models' (LLM) abilities to capture abstract knowledge about world's physics to solve decision-making problems. Yet, the alignment between LLMs' knowledge and the environment can be wrong and limit functional competence due to lack of grounding. In this paper, we study an approach (named GLAM) to achieve this alignment through functional grounding: we consider an agent using an LLM as a policy that is progressively updated as the agent interacts with the environment, leveraging online Reinforcement Learning to improve its performance to solve goals. Using an interactive textual environment designed to study higher-level forms of functional grounding, and a set of spatial and navigation tasks, we study several scientific questions: 1) Can LLMs boost sample efficiency for online learning of various RL tasks? 2) How can it boost different forms of generalization? 3) What is the impact of online learning? We study these questions by functionally grounding several variants (size, architecture) of FLAN-T5.

Functional Description: GLAM is a new approach to achieve alignment between a Large Language Model (LLM) and a considered environment/world through functional grounding: we consider an agent using an LLM as a policy that is progressively updated as the agent interacts with the environment, leveraging online Reinforcement Learning to improve its performance to solve goals.

URL: https://github.com/flowersteam/Grounding_LLMs_with_online_RL

Publication: hal-03970122

Contact: Clément Romac

7.1.19 ER-MRL

Keywords: Reservoir Computing, Reinforcement learning

Functional Description: Code for the Evolving-Reservoirs-for-Meta-Reinforcement-Learning (ER-MRL) paper. Our goal is to study the following question : How neural structures, optimized at an evolutionary scale, can enhance the capabilities of agents to learn complex tasks at a developmental scale? To achieve this, we adopt a computational framework based on meta reinforcement learning, modeling the interplay between evolution and development. At the evolutionary scale, we evolve reservoirs, a family of recurrent neural networks generated from hyperparameters. These evolved reservoirs are then utilized to facilitate the learning of a behavioral policy through reinforcement learning. This is done by encoding the environment state through the reservoir before presenting it to the agent.

URL: <https://github.com/corentinleger/ER-MRL>

Contact: Corentin Leger

7.1.20 ecoevojax_analysis

Keywords: Evolutionary Algorithms, Evolution, Grid

Functional Description: Code for the paper "Eco-evolutionary dynamics of non-episodic neuroevolution in large multi-agent environments." <https://dl.acm.org/doi/10.1145/3583133.3590703>. In this work we have proposed a grid-world foraging environment with reset-less neuroevolution. It is implemented in JAX and can support very large population and grid sizes. For example, evolving a population of 1000 agents in a world of 100 K pixels for 1 million time steps takes about 10 minutes on a GPU.

URL: https://github.com/eleninisioti/ecoevojax_analysis

Contact: Gautier Hamon

7.1.21 SBMLtoODEjax

Keywords: SBML, JAX, Python, Numerical simulations, Numerical optimization, Automatic differentiation, Ordinary differential equations, Biomedical data

Scientific Description: Advances in bioengineering and biomedicine demand a deep understanding of the dynamic behavior of biological systems, ranging from protein pathways to complex cellular processes. Biological networks like gene regulatory networks and protein pathways are key drivers of embryogenesis and physiological processes. Comprehending their diverse behaviors is essential for tackling diseases, including cancer, as well as for engineering novel biological constructs. Despite the availability of extensive mathematical models represented in Systems Biology Markup Language (SBML), researchers face significant challenges in exploring the full spectrum of behaviors and optimizing interventions to efficiently shape those behaviors. Existing tools designed for simulation of biological network models are not tailored to facilitate interventions on network dynamics nor to facilitate automated discovery. Leveraging recent developments in machine learning (ML), this paper introduces SBMLtoODEjax, a lightweight library designed to seamlessly integrate SBML models with ML-supported pipelines, powered by JAX. SBMLtoODEjax facilitates the reuse and customization of SBML-based models, harnessing JAX's capabilities for efficient parallel simulations and optimization, with the aim to accelerate research in biological network analysis.

Functional Description: SBMLtoODEjax extends SBMLtoODEpy, a python library developed in 2019 for converting SBML files into python files written in Numpy/Scipy. The chosen conventions for the generated variables and modules are slightly different from the standard SBML conventions (used in the SBMLtoODEpy library) with the aim here to accommodate for more flexible manipulations while preserving JAX-like functional programming style.

URL: <https://developmentalsystems.org/sbmltoodejax/>

Publication: hal-04317246

Contact: Mayalen Etcheverry

Partner: Tufts University

7.2 New platforms

7.2.1 ToGather application

Participants: Cécile Mazon, Hélène Sauzéon, Eric Meyer, Isabeau Saint-Supery.

Name: Application for Specialized education

Keywords: Parent-professional relationships; user-centered design; school inclusion; autism spectrum disorder; ecosystemic approach

Participants: Isabeau Saint-supery, Cécile Mazon, Hélène Sauzéon, Agilonaute

Scientific Description: With participatory design methods, we have designed an interactive website application for educational purposes. This application aims to provide interactive services with continuously updated content for the stakeholders of school inclusion of children with specific educational needs. Especially, the services provide: 1) the student's profile with strengths and weaknesses; 2) an evaluation and monitoring over time of the student's repertoire of acquired, emerging or targeted skills; 3) a shared notebook of effective psycho-educational solutions for the student; 4) a shared messaging system for exchanging "news" about the student and his/her family and, 5) a meeting manager allowing updates of evaluations (student progress). This application is currently assessed with a field study. Then, it will be transferred to the Academy of Nouvelle-Aquitaine-Bordeaux of the National Education Ministry.

URL: The website is not online yet, but all informations such as tutorials are <https://flowers.inria.fr/projet-tous-ensemble/>.

Publication: hal-03436355

8 New results

The team's research program, within the domain of developmental artificial intelligence, aims to study mechanisms of open-ended learning, and in particular the role of curiosity-driven autotelic learning and the role of language as a cognitive tool. We study these topics both in humans and AI systems, both at the level of individuals and at the level of cultural groups, and both at the fundamental and application levels.

Here, we present our recent results along the following research dimensions:

- Open-ended learning and autotelic AI with large language models;
- Models of cultural evolution in humans and AI systems;
- Eco-evolutionary AI and meta-reinforcement learning;
- Generative AI and educational technologies;
- Meta-cognition in curiosity-driven educational technologies;
- Curiosity-driven AI for assisted scientific discovery;

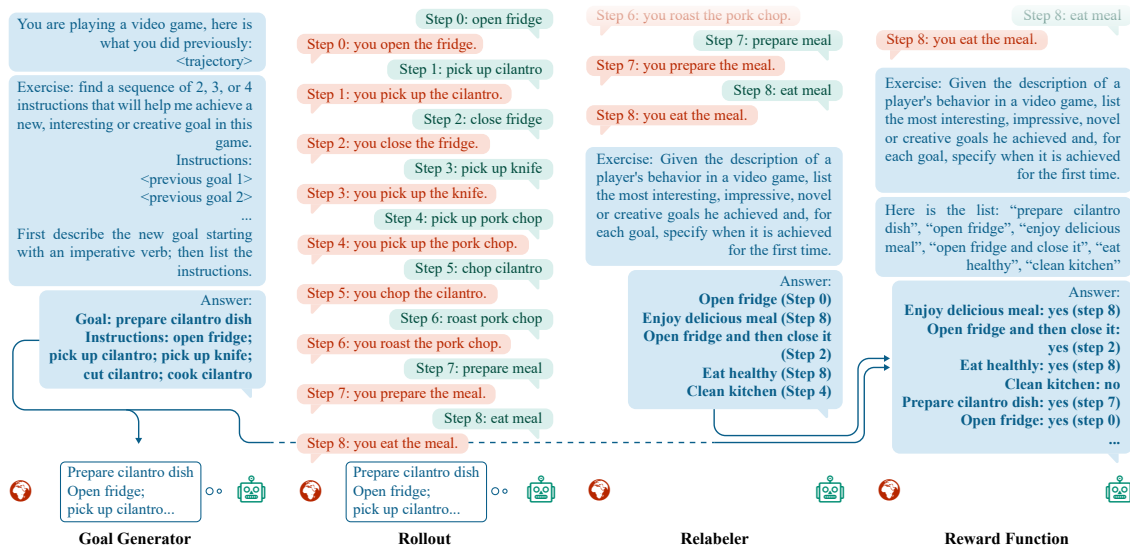


Figure 2: Overview of the LMA3 method. The Goal generation, Policy, Relabeller, and Reward function are represented along with a simplification of the prompts used for the LM.

8.1 Open-ended learning and autotelic AI with large language models

The team continued to lay the foundations of autotelic AI [111, 109], i.e. the science studying mechanisms enabling artificial agents to learn to represent and sample their own goals and achieve open-ended learning.

8.1.1 Language-Model-Augmented Autotelic Agents, Towards Open-Ended Skill Discovery

In this project, we take a step towards truly open-ended language-based autotelic agents by leveraging GPT3 [97], a Large Language Model (LLM) demonstrating impressive language understanding capabilities. For an autotelic agent to be truly open-ended, it needs to be able to:

- Generate its own goals creatively (Goal generator)
- For an arbitrary goal, decide whether a given trajectory achieved the goal or not (Reward function)
- For a given trajectory, give a list of relevant achieved goals (Relabeller or Social Partner)

In this project, we also place ourselves in a textual environment. CookingWorld has been proposed as part of the First TextWorld Problems competition on text agents, and features a house with ingredients that can be cooked to achieve a recipe. We leverage the textual nature of the trajectories of agents in this environment by using GPT (ChatGPT-3.5 in our experiments), with different prompts, as our goal generator, reward function, and relabeller. More precisely (see Figure 2):

- **Goal generator:** we present GPT3 with a prompt composed of a context explanation ("We are in this video game environment"), a list of previously seen rooms, a list of previously seen objects, a list of previously achieved goals, possibly a past trajectory example, and ask it to generate new goals;
- **Reward function:** given an arbitrary language goal and a trajectory including actions and observations, labelled with the steps index at which those actions and observations were issued, we ask GPT to decide whether a given goal is achieved or not in the trajectory, and if it is, at which time step. In this way, we only need to call the GPT-based reward function once per episode (compared with once per step in a traditional RL setup), significantly speeding up our experiments, since the GPT calls are expensive.
- **Relabeller:** Once a trajectory is finished, we show GPT the completed trajectory and ask it to provide a list of goals achieved in this trajectory. We extract these goals and we add the relabelled

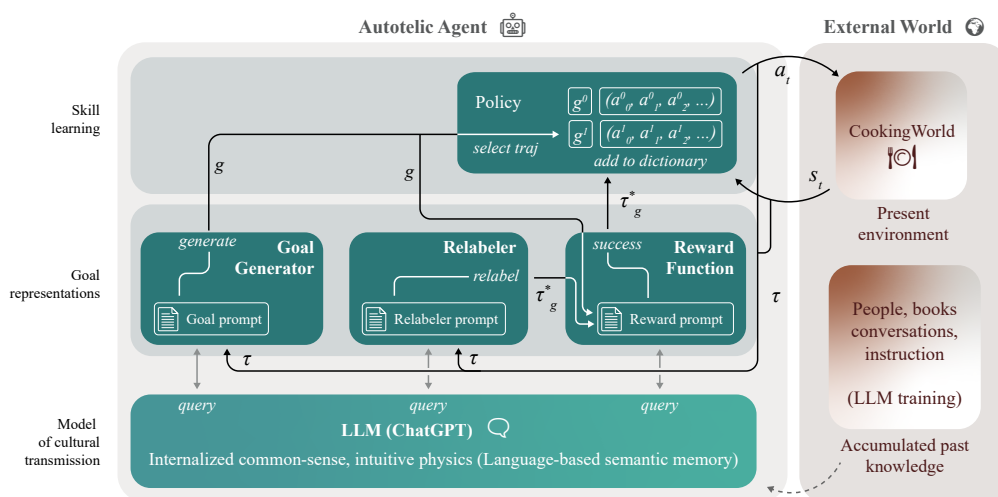


Figure 3: Architecture of LMA3.

trajectories to the replay buffer. The relabelled goals are given again to the reward function for it to double-check correctness.

Goal-conditioned policy implementation Once a goal has been selected by the Goal Generator we need to be able to pursue this goal in the environment. In preliminary experiments we noticed that Deep RL based agents were too sample inefficient to be able to learn an open-ended repertoire of skills in a reasonable time. The solution we implemented is closer to an evolutionary algorithm: the policy maintains a dictionary mapping already experienced goal strings to the shortest experienced sequence of action achieving this goal (according to the relabeller or the reward function). When presented with a new goal, the agent embeds this goal and selects the sequence of actions whose goal key in the dictionary is the closest in embedding space to the one being executed (this match may be exact). The agent then executes this action sequence.

Goal Generator details The LM-based Goal generator starts off with a warmup period of 4000 episodes, where goals are sampled based on the ones that have already been experienced. After this period the goal imagination kicks in, and the language model is shown in its prompt a sample of 60 goals that have been achieved in the past. The goal generator is asked to provide a high-level goal that is composed of the chaining of lower level subgoals among the list of goals in the prompt. The goal-conditioned policy then executes the subgoals one after the other, and there is a probability that the last subgoal be cutoff and random exploration be performed. In the random exploration text actions are given priority inversely proportional to the amount of time they have been seen (to encourage the agent to take rare actions when they become available). The overview of the algorithm, as well as an illustration of the information flow, are given in Figures 2 and 3. Figure 4 shows a confusion matrix quantifying how much the reward function agrees with human judgement.

Ablations We perform several ablation: in one experiment we ablate the human advice and suggestions given to the model (no human tips), in a further ablation we remove the lm goal generator (the warmup phase stays so indefinitely: no lm goals or human tips) and in a final ablation we remove the human tips and the chain of thought prompting in the few-shot examples we give the LM (no human tips or CoT).

Results Our results show that LMA3, across methods, is able to discover a very large number of goals diverse goals, on the order of 1 goal per episode. The goals are distinct and also exhibit high diversity in the predicates and objects they cover. We see that removing human tips reduces the diversity of goals generated, because the LM has less diverse examples on which to base itself. Removing the LM goal generation has minimal influence beyond that, and removing chain-of-thought has the most drastic

	LM success	LM failure
True success	0.91	0.09
True failure	0.09	0.91

Figure 4: Confusion scores between the LM’s judgements of success and failure on a set of 100 goal-trajectory pairs and human judgements.

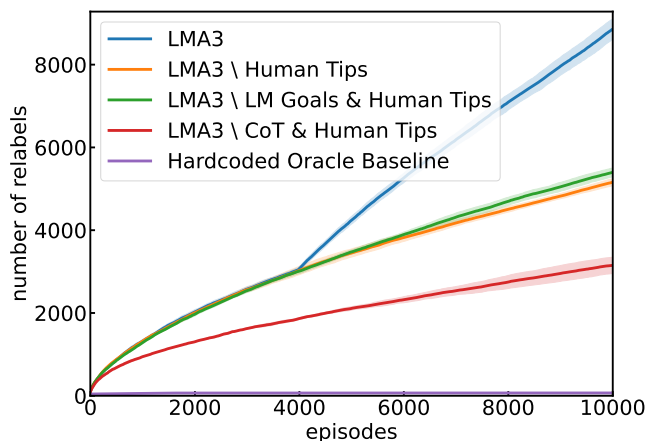


Figure 5: Number of goals discovered by the different versions of LMA3.

effect. The baseline oracle agent is limited to the 69 goals for which we have an exact reward function and relabeler. We additionally evaluate the agent on the oracle goals for which we have an exact reward function. This is represented in Figure 4. We see that most LMA3 variants achieve a high eval score on these held-out goals, despite never having been told what it would be evaluated on.

This work was presented at the Conference on Lifelong Learning Agents (August 2023) in Montreal.

8.1.2 Code-Generating Autotelic Agents

In this section we will present our recent work on code-generating autotelic agents. The idea of exploring this came from the realization that the open-endedness of our agents is upper-bounded by the complexity of the environment. The LMA3 agent was very powerful but could never learn complex skills because CookingWorld is extremely limited. On the other hand, programming is open-ended and an agent writing code has very few limits on the complexity of what it can accomplish.

We will present two works in a programming domain: Codeplay and ACES.

8.1.3 Codeplay: Autotelic Learning through Collaborative Self-Play in Programming Environments

Participants: Laetitia Teodorescu (*correspondant*), Cédric Colas, Matthew Bowers, Thomas Carta, Pierre-Yves Oudeyer.

In this work we propose an approach that is both a way to implement autotelic agents discovering a truly open-ended set of goals, and a way to allow language models to master novel coding skills in

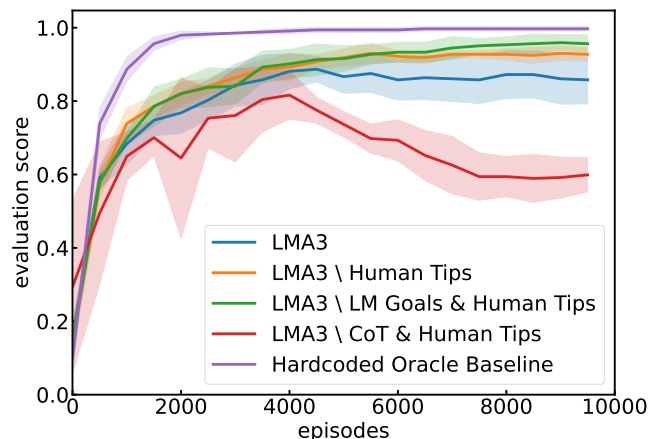


Figure 6: Average success rates for all the 69 oracle goals for all the LMA3 variants. The oracle baseline achieves perfect score insofar as all the oracle goals are discovered by the algorithm.

interaction with an interpreter. We ground LM-agents in a coding environment that provides straightforward binary rewards through code execution. We set ourselves in the Python Programming Puzzles domain [184], and we define an autotelic agent composed of 2 sub-agents: a **Setter** that generates puzzles, and a **Solver** whose objective is to solve the generated puzzles. Both agents are LM policies. They play a collaborative game where the Setter has to create puzzles that push the solver to learn, and the solver sees and tries to solve puzzles in its zone of proximal development (ZPD): hard but still solvable. First experiments presented here (with a fixed Solver) highlight the possibility to tradeoff difficulty of the generated puzzles and their novelty by tuning the reward experienced by the Setter, trained with deep RL (PPO: [182]).

Related work Closely related to this work are approaches to autotelic learning or automatic curriculum learning involving goal setter agents [121, 191, 99, 161, 155], as well as the PowerPlay framework of [180]. Very closely related to this work as well are approaches for generating novel code puzzles for augmenting the capabilities of code-puzzle-solving language models [125], as well as recent attempts to cast program synthesis as a reinforcement learning problem [141, 203].

Python programming puzzles We use as our testbed the Python Programming Puzzle (P3) domain of [184]. Each puzzle is defined by a test program f designed to verify the validity of solution programs g such that valid solutions satisfy $f(g())$ is True when run in an interpreter. P3 puzzles span problems of various difficulties, from trivial to open questions.

Setter We instantiate our Setter as a pretrained language model, finetuned on the P3 domain. We cast the puzzle-generating problem as an MDP where the observation space is the list of all possible sequences of tokens, the action space is the list of all tokens and the reward is the intrinsic motivation measure we compute based on the difficulty of a puzzle. Transitioning from one (fully-observable) state to another is simply appending the emitted token to the observation: the environment is purely deterministic. This training setup is reminiscent of the one in reinforcement learning from human feedback (RLHF: [169]), except in our case we do not use a reward function trained from human preferences but an intrinsic motivation metric based on the Solver’s abilities. Our Setter agent’s stochastic policy is given by the pretrained language model’s logits which are used to sample the tokens (temperature of 0.8 in our experiments). We use PPO [182] as our training deep RL algorithm with an implementation based on the RL4LMs library [174]. As the value head in PPO we use a separate untrained MLP head on top of the language model backbone. In our preliminary experiments we keep the Solver fixed and investigate different rewards for the Setter. We use the small GPT-Neo-125M¹ pretrained on the Codex-generated

¹<https://github.com/EleutherAI/gpt-neo>

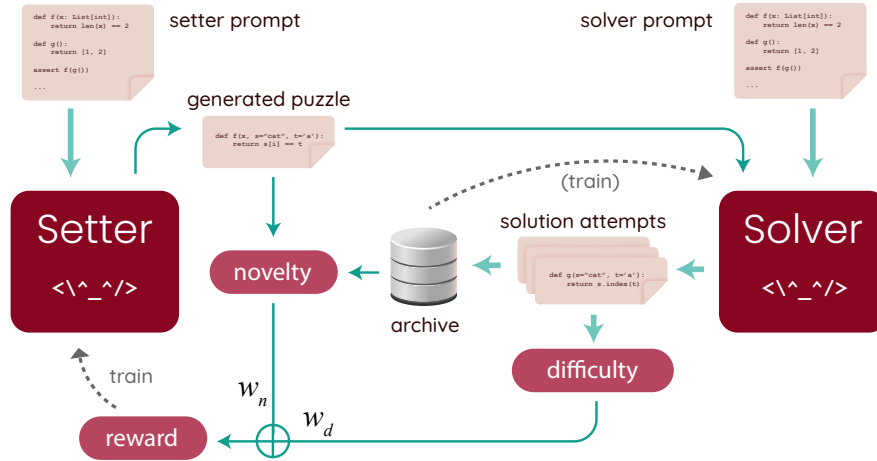


Figure 7: Overview of the proposed Codeplay algorithm. The Setter, a language model, takes in a few-shot prompt and emits a puzzle. This puzzle is given to the problem Solver who appends it to its few-shot prompt, and generates N_a candidate solutions. These problem-solution pairs are given to the Python interpreter, this gives us success or failure information on the solution. The number of successful attempts allows us to compute a difficulty reward. The generated puzzle is also compared with all previously generated puzzles to compute its novelty. Those two rewards are weighted and summed and used as the reward in a deep RL algorithm to train the Setter.

dataset of [125].

Setter difficulty reward We want to reward the Setter for producing puzzles that are hard, while not being empirically unsolvable. To do so, we compute a reward based on the number of solutions generated by our fixed Solver within a maximum number of attempts. Easy problems have a reward of 0, unsolvable or syntactically invalid puzzles have a reward of -1, and hard but solvable puzzles have a reward of 1.

Setter novelty reward Because the Solver is fixed, optimizing only R_d will lead to Setter collapse on a single puzzle (or a small set). To encourage diversity in the generated puzzles we define the *novelty reward* $R_n(p, \mathcal{A})$ of a puzzle p as the average distance between p and its 5 closest neighbors in the archive of previously generated puzzles \mathcal{A} in an embedding space. We normalize this distance by the average pairwise distance of puzzles in the P3 train set in that same space (so R_n is roughly between 0 and 1).

The total reward for the Setter is a weighted sum of the difficulty and novelty rewards:

$$R(p) = w_d R_d(\mathcal{D}(p)) + w_n R_n(p, \mathcal{A}) \quad (1)$$

In the following experimental results we investigate the impact of different values of the weights. We only train the Setter in these experiments.

Prompts We use a fixed few-shot prompt for both the solver and the setter. The few-shot prompt simply includes a series of puzzles and solutions from the tutorial set of the P3 training set (which are pretty short, so we can fit all of them). The puzzles and solutions are separated by `assert(f(g()))` statements. The prompt for the Setter finishes after an assertion, the prompt for the Solver (evaluated N_a times to produce the difficulty estimation) additionally includes the puzzle to be solved at the end.

Results We present here a series of results studying Setter training). We report the learning curves for both reward components R_d and R_n in Figure 8, as well as the total Setter reward, the quantity effectively being optimized. It is the weighed sum of both components (see Equation 1).

The gray curve provides or baseline: no optimization whatsoever is taking place. On the left-hand side, we see that the puzzles generated in this case are close to trivial (R_d close to 0) and that they are relatively repetitive: the gray novelty reward tends to zero. This is because the novelty reward is non-stationary: the

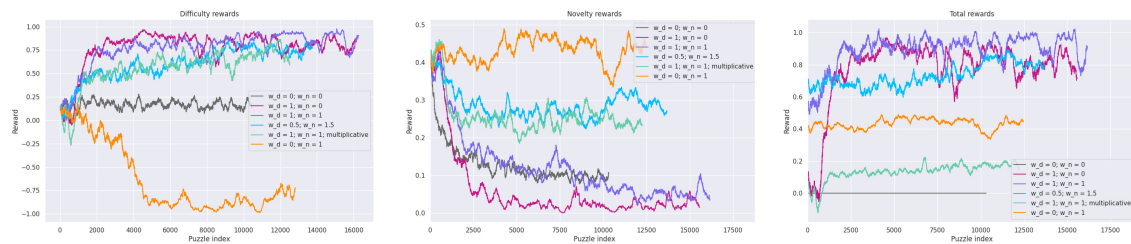


Figure 8: Competence and novelty rewards for different experiments. The values are smoothed over 200 puzzles. The colors correspond to different pairs of weights for Equation 1. See main text for analysis.

more puzzles you generate in the same distribution, the denser the puzzles become in the embedding space used and so the smaller the distance between one puzzle and the next become. The mauve curve shows what happens when only the difficulty reward gets optimized ($w_d = 1, w_n = 0$): the latter goes up pretty quickly but the novelty reward goes to zero fast, indicating highly repetitive generated puzzles (and in effect that is what happens). The orange curve is what happens when only the novelty reward gets optimized ($w_d = 0, w_n = 1$): on the center figure we see that the novelty stays high throughout, which is impressive given that new puzzles are generated continuously, but on the left the difficulty reward goes to -1 quickly. This indicates either generation of unsolvable or of invalid puzzles, and manual inspection reveals the latter to be true. Optimizing only for novelty yields a Solver unable to generate valid programming puzzles. A reward with equal weights ($w_d = 1, w_n = 1$, purple curve) yields puzzle with appropriate difficulty and as repetitive as not optimizing at all (as far as the novelty reward is concerned). Biasing the reward towards novelty ($w_d = 0.5, w_n = 1.5$, blue curve) yields puzzles that maintain a fixed amount of novelty while still eventually converging towards difficult but solvable puzzles. The same effect can be achieved by multiplying the two reward components with a weight of 1 instead of summing them. From these curves we see that there is a tradeoff in the novelty and appropriate difficulty of puzzles for a fixed Solver, and that we are successfully able to navigate this tradeoff by tuning different reward components.

This is still work in progress, and was been presented in the IMOL conference in Paris in September 2023 and in the IMOL workshop at NeurIPS in New Orleans in December 2023.

8.1.4 ACES: Generating diverse programming puzzles with autotelic language models and semantic descriptors

Participants: Julien Pourcel (*correspondant*), Cédric Colas, Pierre-Yves Oudeyer, Laetitia Teodorescu.

Motivation In this project, we examine more specifically how one can generate an interesting diversity of programming puzzles (same domain as Codeplay). We recall that this is an important case study for linguistic autotelic agents because it is a first step towards generalist agents inventing their own problems. Inspired by the Evolution Through Large Models (ELM) method where authors evolve robot morphologies expressed as Sodarace programs using a Large Language Model as a mutation operator, we aim to develop an evolutionary method to create a diverse population of problems using pretrained Language Models. We remark that diversity-producing methods (such as Map-Elites) need a Behavioral Characterization (BC) space in which to measure the diversity of their evolved populations; this is feasible with virtual creatures but seems pretty hard with programming puzzles. We thus introduce the notion of a *Semantic* BC space, composed of abstract categories, and labelling inside this space is done through LLM responses. In our case, we introduce 10 programming descriptors:

- 0 - Sorting and Searching
- 1 - Counting and Combinatorics
- 2 - Trees and Graphs

- 3 - Mathematical Foundations
- 4 - Bit Manipulation
- 5 - String Manipulation
- 6 - Geometry and Grid Problems
- 7 - Recursion and Dynamic Programming
- 8 - Stacks and Queues
- 9 - Optimization Algorithms

We then define an archive of generated programming puzzles and their solutions, and the position of a puzzle in the archive is given by the combination of descriptors that the puzzle-solution pair belongs to (the semantic representation of a puzzle thus being a 10-dimensional vector). The semantic archive is used to store puzzles.

We then perform experiments with the following algorithms:

- **ACES:** our proposed method samples a target cell (combination of descriptors) in the archive at random and populates a few-shot prompt for the language model with puzzles from neighboring cells in the archive. See Figure ?? for an illustration.
- **ELM Semantic:** based on ELM, example puzzles and solutions are given as few-shot in-context examples and a puzzle sampled from the archive is then mutated.
- **ELM:** same as the previous one, except we do not use the semantic archive for sampling: instead we build an archive with centroidal voronoi tessellations, from the embedding of puzzles inside the latent space of a Language Model. This baseline allows us to compare the semantic archive with a more classical one;
- **Static Gen:** In this method, puzzles are sampled from the train set and added as few-shot examples in the prompt;

For all experiments we seed the archive with the P3 train set. (cite the paper)

Results We report results of our runs in Figure 10. Overall, the methods based on semantic archives, ACES and ELM-Semantic, achieve the highest diversity in the semantic space. We report diversity measures inside the embedding spaces of various smaller language models in Figure 11. In these figures we see that overall ACES outperforms other methods in this measure of diversity. We additionally perform tests of the suitability of generated puzzles as finetuning data for smaller LMs. For all methods, we finetune a smaller model (OpenLlama-3b) on the generated set and we test the pass@k metric for different values of k on the P3 test set; we report the scores in Figure 12. From that figure we see that we encounter a tradeoff between how diverse the data is and how useful it is to get a high score on the P3 test set. Further work is needed to get data that is both diverse and useful.

8.1.5 Grounding Large Language Models in Interactive Environments with Online Reinforcement Learning (GLAM)

Participants: Thomas Carta, Clément Romac, Pierre-Yves Oudeyer, Olivier Sigaud (*ISIR Sorbonne Université, Paris, France*), Sylvain Lamprier (*Univ Angers, LERIA*), Thomas Wolf (*Hugging Face*).

The recent rise of Transformer-based Large Language Models (LLMs) trained on massive text datasets has led to models exhibiting impressive capabilities (e.g. natural language generation, question answering, reasoning, translation...). Recently, LLMs were shown to also capture aspects of the physical rules in our world, e.g. about space, colors or even affordances between bodies and objects.

However, LLMs are known to suffer from a lack of grounding (i.e. connecting their inner representation of language to a world) which prevents them from properly dealing with the meaning of concepts and their

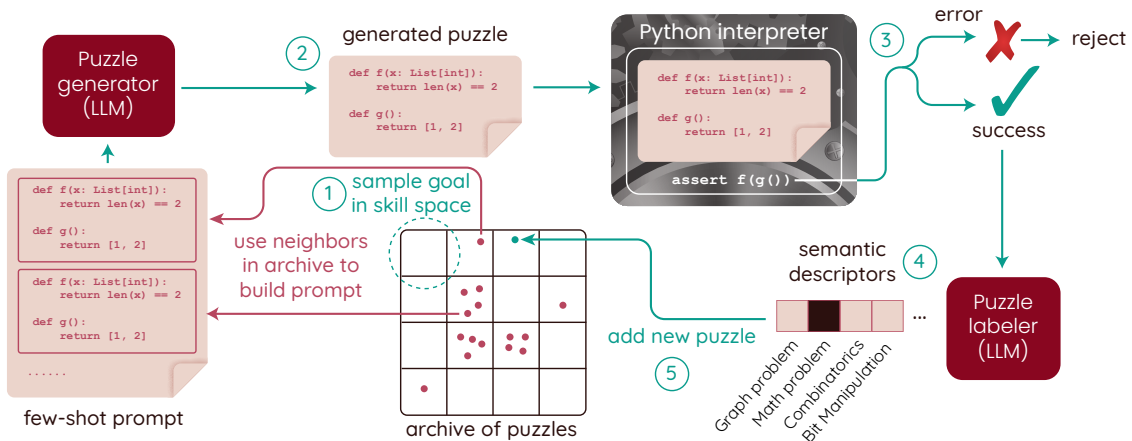


Figure 9: Overview of ACES. ACES maintains an archive of discovered puzzles grouped into cells indexed by their semantic representation (skill combination). ACES runs in several steps: 1) sample a target semantic goal and relevant examples from the archive. 2) given these, generate a puzzle f and its solution g with the puzzle generator. 3) test the validity of that pair by running `assert f(g())` in the interpreter. 4) if the pair is valid, obtain its semantic representation with the puzzle labeler. 5) add the new pair to its corresponding cell in the archive.

direct use to solve tasks in interactive environments. Indeed, alignment between statistical structures in such LLMs and environments can be very limited, or even sometimes entirely wrong. This is partly due to 1) a training process (predicting next words) that is not directly incentivized to solve problems in an environment, 2) lack of abilities to intervene in the environment to identify causal structures; 3) lack in abilities to learn based on data collected as a result of interacting with the environment.

Focusing on such a functional competence, we propose to use an LLM as the policy interacting with a textual environment (i.e. a textual description of the scene is provided by the environment and possible actions are text commands) for decision-making problems. Using Reinforcement Learning (RL) to finetune the LLM to make it solve various tasks in this environment, our method named GLAM “functionally grounds” the LLM on the environment. That is, grounding the dynamics and physical rules of an environment to solve problems and obtain, in the end, an operational LLM able to use natural language to solve tasks in this interactive environment.

Applying GLAM to Flan-T5 780M shows how such an LLM can be functionally grounded and be able to solve the tasks in an interactive textual environment. The resulting LLM also exhibits strong generalization when exposed to variations in the objects it must interact with. Finally, we show that incremental intervention using RL is key by comparing it to passive Imitation Learning. Our results highlight that our GLAM method leads to both better results and generalization even when Imitation Learning is performed using an optimal policy.

8.1.6 Automatic Curriculum Learning for Language Modeling

Participants: Clément Romac (*correspondant*), Rémy Portelas, Pierre-Yves Oudeyer.

We showed in recent works how Automatic Curriculum Learning (ACL) could help Deep Reinforcement Learning methods by tailoring a curriculum adapted to learner’s capabilities [30, 27]. Using ACL can lead to sample efficiency, asymptotic performance boost and help in solving hard tasks.

Parallel to this, recent works in Language Modeling using Transformers (e.g. GPT-2) have starting to get more interested in better understanding convergence and learning dynamics of these models. Trained in a supervised setup, these models are fed with hundred of millions of natural language sequences crawled from the web. The current standard way of training these models (i.e. constructing batches of randomly selected sequences) makes the assumption that all sequences have same interest for the model. However, recent works showed that this does not seem to be the case and that datasets can contain

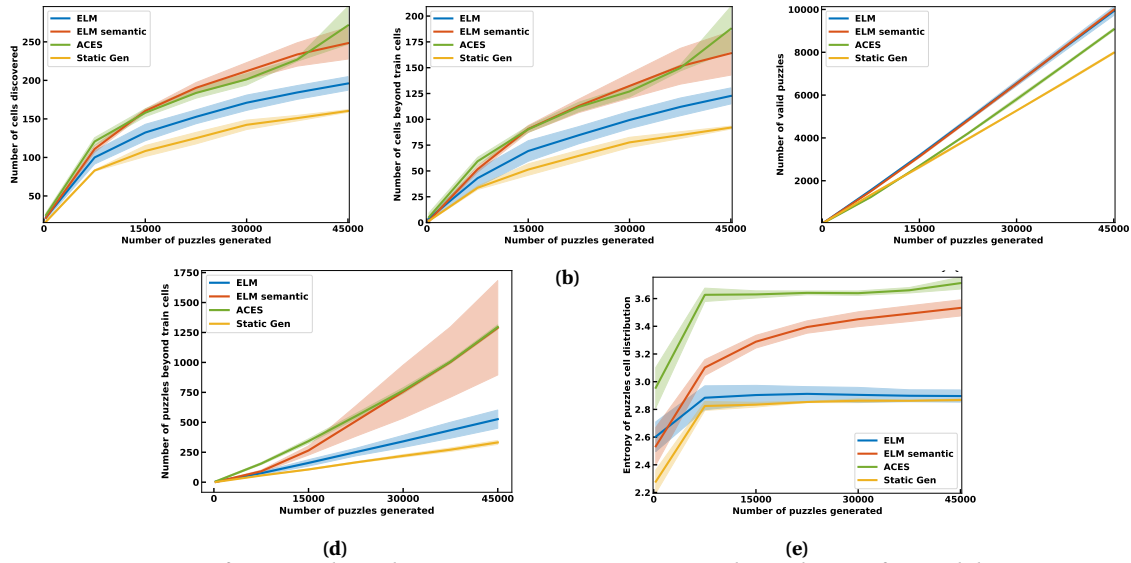


Figure 10: Diversity of generated puzzles in semantic space. We report the evolution of several diversity metrics computed in the semantic space as a function of the number of puzzle-solution pairs generated by the puzzle generator. Semantic algorithms (algnme and ELM semantic) achieve higher diversity in the semantic space.

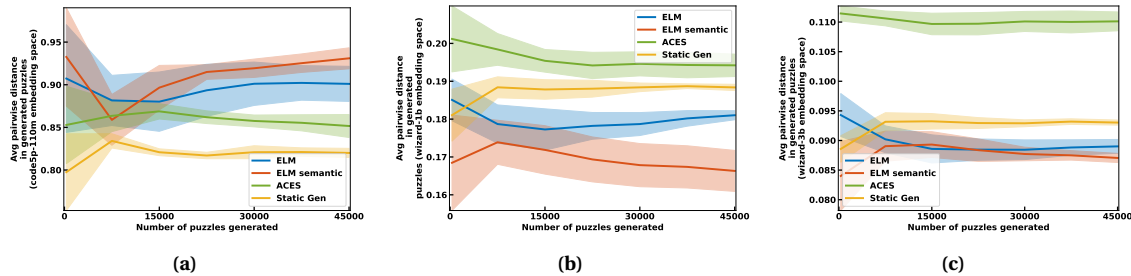


Figure 11: Diversity of generated puzzles in embedding spaces. We report the evolution of the pairwise distance between puzzle-solution pair embeddings as a function of the number of generated puzzle-solution pairs, for three different embedding representation spaces (average across seeds).

outliers harming training. Additionally, some works also showed that hand-designing a curriculum over sequences (e.g. ordered by their length) could speed up and stabilize training.

Building on this, we propose to experiment how ACL could help tailor such a curriculum in an automated way relying on Learning Progress. Our study has several contributions:

- Propose a standardized and more in-depth comparison of current curriculum learning methods used to train language models
- Introduce the first study of ACL in such a training
- Use ACL to propose deeper insights about training dynamics of Transformer models when doing Language Modeling by analysing generated curricula and Learning Progress estimations

We chose to train GPT-2 on the standard OSCAR dataset and use teacher algorithms to select samples that are shown to the model (see fig. 14).

Using ACL, we perform an in-depth analysis of prior methods changing the size of tokens' sequences observed during training following a hand-designed curriculum. Our experiments showed that a Random baseline outperforms these methods. We also provide, thanks to ACL methods based on Learning-Progress Multi-Armed Bandits, hints that while short sequences should not be used as training advances (as Large Language Models quickly learn them), there is no clear evidence that short sequences should be prioritized (and thus long sequences avoided) at the beginning of training.

Additionally, we performed several experiments using more advanced ACL methods on different task spaces and show that these lead to overfitting and underperform in comparison the no-curriculum strategy usually applied in Language Modeling. We hypothesize that, given how large models used in

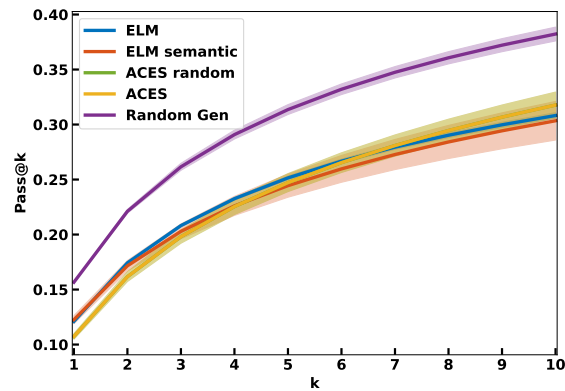


Figure 12: Downstream performance on the P3 test set. Pass@ k is the fraction of puzzles solved after k attempts ($k \in [1:10]$). Green overlaps with yellow.

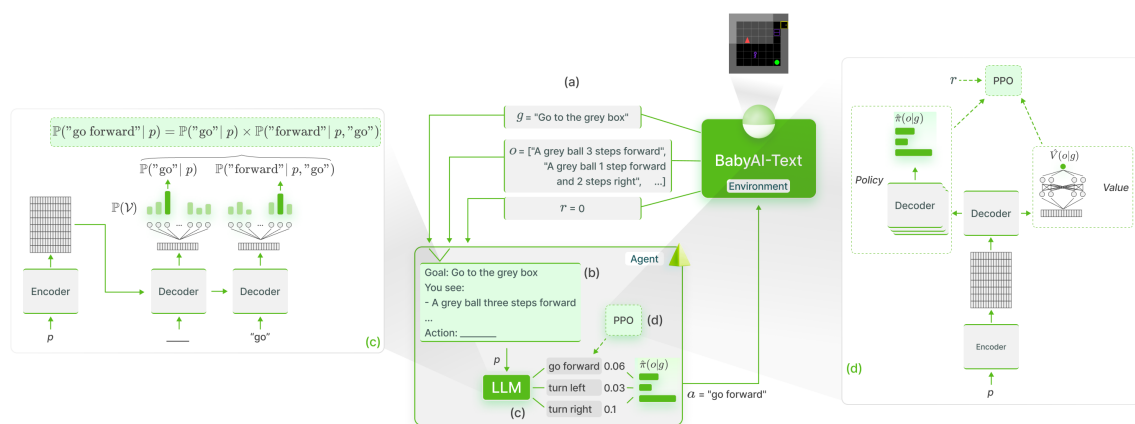


Figure 13: The GLAM method: we use an LLM as agent policy in an interactive textual RL environment (BabyAI-Text) where the LLM is trained to achieve language goals using online RL (PPO), enabling functional grounding. (a) BabyAI-Text provides a goal description for the current episode as well as a description of the agent observation and a scalar reward for the current step. (b) At each step, we gather the goal description and the observation in a prompt sent to our LLM. (c) For each possible action, we use the encoder to generate a representation of the prompt and compute the conditional probability of tokens composing the action given the prompt. Once the probability of each action is estimated, we compute a softmax function over these probabilities and sample an action according to this distribution. That is, the LLM is our agent policy. (d) We use the reward returned by the environment to finetune the LLM using PPO. For this, we estimate the value of the current observation by adding a value head on top of our LLM. Finally, we backpropagate the gradient through the LLM (and its value head).

Language Modeling are, it is better to give a huge amount of very diverse and different samples (even though outliers or harmful samples exist) without any curriculum than using a curriculum that restrains the diversity of samples and introduces duplicates (leading to overfitting).

8.2 Models of cultural evolution in humans and AI systems

As generative AI systems become powerful cultural transmission technologies that influence human cultural evolution in important ways, and can also have their own cultural processes through machine-machine large scale interaction, the study of the dynamics of cultural processes in populations of AI systems/humans becomes crucial.

8.2.1 The effect of social network structure on collective innovation

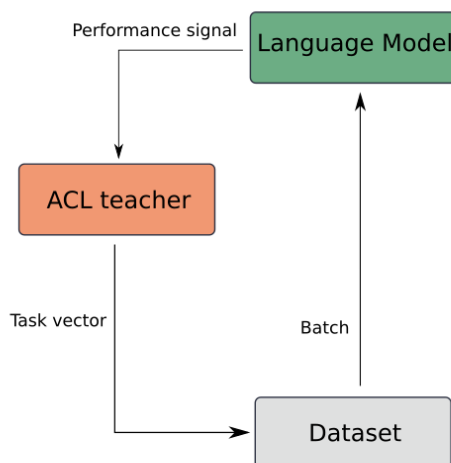


Figure 14: Schema of how ACL was integrated to Language Modeling.

Participants: Eleni Nisioti (*correspondant*), Mateo Mahaut, Pierre-Yves Oudeyer, Ida Momennejad, Clément Moulin-Frier.

Innovations are a central component of open-ended skill acquisition: they denote the emergence of new solutions by the recombination of existing ones and their presence is necessary to ensure a continuous complexification of an agent’s cultural repertoire. While we often tend to attribute discoveries to certain innovative individuals, if we shed a broad perspective at the history of our species we see that human innovation is primarily a collective process. Fields such as psychology and anthropology have been studying the ability of human groups to innovate for some time, with studies indicating that the social network structure has a significant impact: fully-connected structures are better suited for quick convergence in easy problems with clear global optima, while partially-connected structures perform best in difficult tasks where local optima may lure agents away from the globally optimal solution [117]. At the same time a parallel story is unfolding in reinforcement learning (RL): distributed RL is a sub-field where multiple agents solve a task collectively [156]. Compared to the single-agent paradigm, distributed RL algorithms converge quicker and often achieve superior performance. However, these algorithms have only considered full connectivity. In this inter-disciplinary project, we presented a novel learning framework that augments distributed RL with the notion of a social network structure and employed it to study the hypothesis from human studies that partial connectivity performs best in innovation tasks.

We implemented such innovation tasks using Wordcraft, a recently introduced RL playground inspired from the [Little Alchemy 2 game](#) (see left of figure 15 for an illustration of how this task works). We considered a wide diversity of social network structures: static structures that remain constant throughout learning (fully-connected, ring, small-world) and a dynamic structure where the group oscillates between phases of low and high connectivity (we illustrate this dynamic structure on the right of figure 15). Each agent in our implementation employs the DQN learning algorithm and exchanges experiences that have the form of sequences of state-action combinations with its neighbors.

A central conclusion of our empirical analysis was that the dynamic social network structure performs best. In addition to the performance groups achieve we measured behavioral and mnemonic metrics such as behavioral conformity and mnemonic diversity. Such metrics were inspired from human studies and helped us further analyze the behavior of groups. For example, one empirical observation was that sharing experiences did not help the group learn quicker in a very simple innovation task; instead the fully-connected group was the slowest. By looking at the diversity in the memories of the agents we observed that the fully-connected structure had the highest individual diversity (left of figure 16) and

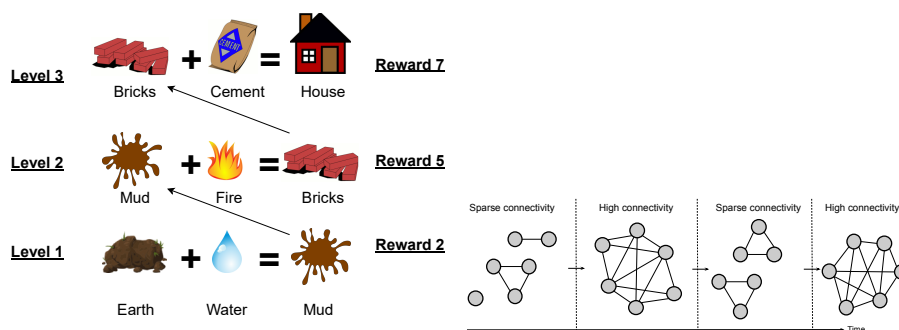


Figure 15: (Left) Illustration of an innovation task, consisting of an initial set of elements (Earth, Water) and a recipe book indicating which combinations create new elements. Upon creating a new element the player moves up an innovation level and receives a reward that increases monotonically with levels. (Right) Dynamic social network structures oscillate between phases of low connectivity, where experience sharing takes place within clusters, and high connectivity, where experiences spread between clusters.

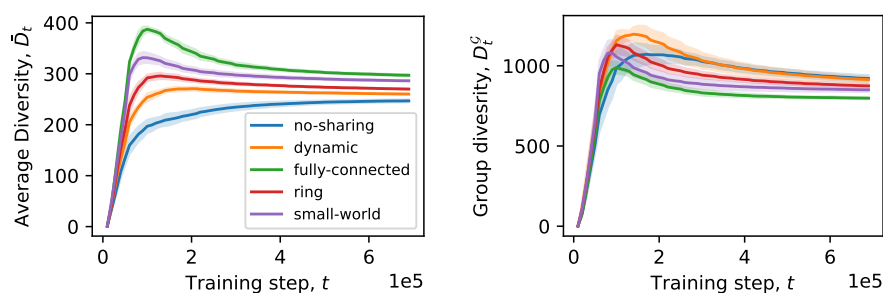


Figure 16: (Left) Illustration of an innovation task, consisting of an initial set of elements (Earth, Water) and a recipe book indicating which combinations create new elements. Upon creating a new element the player moves up an innovation level and receives a reward that increases monotonically with levels. (Right) Dynamic social network structures oscillate between phases of low connectivity, where experience sharing takes place within clusters, and high connectivity, where experiences spread between clusters.

the lowest group diversity (right of figure 16): sharing experiences with others diversifies an individual's experiences but also homogenizes the group, which is bad for its performance.

We see the contribution of this project as two-fold. From the perspective of fields studying human intelligence, we have shown that using RL algorithms as computational tool is a promising direction towards increasing the verisimilitude of simulations and analyzing both behavior and memory. From the perspective of RL, we have shown that distributed RL algorithm should move beyond the fully-connected architecture and explore groups with dynamic topologies. This work is currently a preprint [159] and is about to be submitted in PNAS. We open-source the code [at this link](#).

8.2.2 Cultural evolution in population with heterogeneous and variable preferences

Participants: Jérémy Perez (*correspondant*), Martí Sánchez Fibla, Clément Moulin-Frier.

Theoretical and empirical work in cultural evolution often assume populations in which individuals all agree on the quality of a given cultural trait (i.e. they have homogeneous preferences), and where those preferences are stable in time. Yet, these assumptions are not always met: for example, an uneven distribution of information in a population could lead to heterogeneous preferences; moreover, in some cultural domains (e.g. aesthetic culture), diverse preferences may be the norm rather than the exception. In this project in collaboration with Martí Sánchez Fibla from Universitat Pompeu Fabra,

we designed an agent-based model in which we can control the heterogeneity of preferences, as well as the effect of cultural traits on the evolution of preferences. We find that assuming homogeneous or heterogeneous preferences leads to different predictions on several outcomes. First, populations with greater heterogeneity of preferences converge toward greater cultural diversity. Second, while we replicate the classical result that increasing opportunities to learn socially leads to less diversity in homogeneous populations, we find that this relationship is reversed in heterogeneous populations. We show that this happens because increasing social learning opportunities leads the distribution of cultural traits to converge toward the distribution of preferences. We also look at the consequences of allowing cultural traits to modify the preferences of individuals that possess them. This can for example capture self-reinforcing beliefs, or traits where the acquisition costs make individuals less likely to switch to another trait after possessing them for some time. We find that such “attractive” cultural traits naturally emerge in our model, and that they tend to decrease cultural diversity when preferences are not homogeneous. Overall, by showing that the effect of different parameters on cultural diversity are dependent on the assumed distribution of preferences, we highlight the importance of taking into account the possible heterogeneity of preferences when making predictions about cultural dynamics. An abstract for a poster was submitted to the conference of the European Human Behaviour and Evolution Association (EHBEA 2024).

8.2.3 Interactions between intrinsically motivated goal-exploration processes and cumulative cultural evolution

Participants: Jérémy Perez (*correspondant*), Maxime Derex, Pierre-Yves Oudeyer, Clément Moulin-Frier.

Cumulative culture describes the gradual accumulation and spread of innovations in a population, which results in the formation of a cultural repertoire that no individual could have independently invented on its own. As cumulative culture has been argued to underlie the ecological success of humans, understanding the mechanisms that underpin it has raised a lot of interest. However, computational models of cultural evolution have mostly been restricted agent-based models, which made it difficult to study the consequences of some cognitive capacities on cultural dynamics. In particular, these models often assume that cultural variation is generated in a random manner, which overlooks the sophisticated exploration strategies employed by humans and other animals, such as curiosity-driven exploration. In this project, we aim to fill this gap by modeling agents as reinforcement learners endowed with an intrinsic motivation to generate and pursue their own goals. This will allow to study how a group of curious agents may take cultural trajectories compared to non-curious agents, as well as to look at more sophisticated forms of cultural transmission such as goal transmission. This project is done in collaboration with Maxime Derex (IAST, Toulouse)

8.2.4 Learning and Self-organization of Cultural Conventions Between Artificial Agents

Participants: Tristan Karch (*correspondant*), Clément Moulin-frier, Pierre-Yves Oudeyer.

As introduced earlier, Vygotskian artificial agents internalize cultural conventions in order to transform linguistic production into cognitive tools that help them acquire new skills. A fundamental question is therefore to investigate how such cultural conventions can emerge between agents situated in social contexts.

Self-organizing cultural conventions in a new interactive AI paradigm: the Architect-Builder Problem

In this experiment, we are interested in interactive agents that learn to coordinate, namely, a builder – which performs actions but ignores the goal of the task, i.e. has no access to rewards – and an architect which guides the builder towards the goal of the task. We define and explore a formal setting where artificial agents are equipped with mechanisms that allow them to simultaneously learn a task while at

the same time evolving a shared communication protocol. Ideally, such learning should only rely on high-level communication priors and be able to handle a large variety of tasks and meanings while deriving communication protocols that can be reused across tasks. We present the Architect-Builder Problem (ABP): an asymmetrical setting in which an architect must learn to guide a builder towards constructing a specific structure. The architect knows the target structure but cannot act in the environment and can only send arbitrary messages to the builder. The builder on the other hand can act in the environment, but receives no rewards nor has any knowledge about the task, and must learn to solve it relying only on the messages sent by the architect. Crucially, the meaning of messages is initially not defined nor shared between the agents but must be negotiated throughout learning. The Architect-Builder problem was initially introduced by Vollmer et. al [199] in an experiment named the CoCo game studying the formation of communication protocol between humans in such a context. Diagrams of interactions in the CoCo game and in our numerical adaptation are given in figure 17.

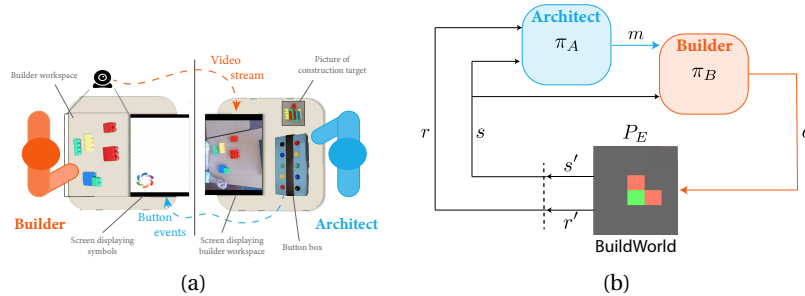


Figure 17: (a) **Schematic view of the CoCo Game (the inspiration for ABP).** The architect and the builder should collaborate in order to build the construction target while located in different rooms. The architecture has a picture of the target while the builder has access to the blocks. The architect monitors the builder workspace via a camera (video stream) and can communicate with the builder only through the use of 10 symbols (button events). (b) **Interaction diagram between the agents and the environment in our proposed ABP.** The architect communicates messages (m) to the builder. Only the builder can act (a) in the environment. The builder conditions its action on the message sent by the builder ($(a|s, m)$). The builder never perceives any reward from the environment.

Under these constraints, we propose Architect-Builder Iterated Guiding (ABIG), a solution to ABP where the architect leverages a learned model of the builder to guide it while the builder uses self-imitation learning to reinforce its guided behavior. We analyze the key learning mechanisms of ABIG and test it in 2D tasks involving grasping cubes, placing them at a given location, or building various shapes. ABIG results in a low-level, high-frequency, guiding communication protocol that not only enables an architect-builder pair to solve the task at hand, but that can also generalize to unseen tasks as illustrated in figure 18. These results were published at the International Conference on Representation Learning (ICLR 2022) [89].

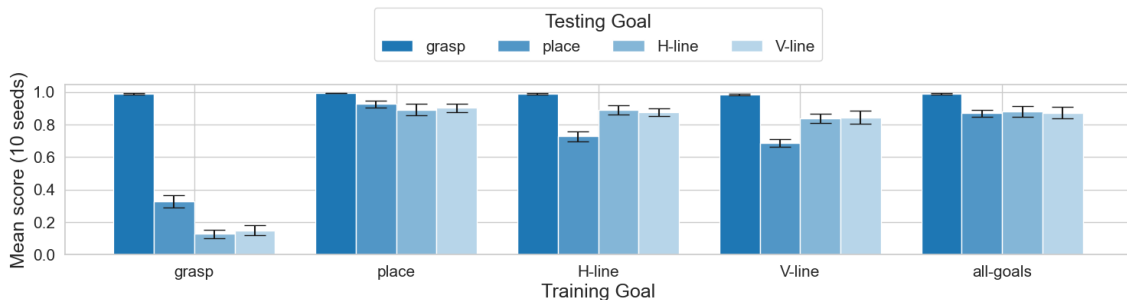


Figure 18: ABIG transfer performances without retraining depending on the training goal. ABIG agents learn a communication protocol that transfers to new tasks. Highest performances reached when training on 'place'.

8.2.5 Autotelic Reinforcement Learning in Multi-Agent Environments

Participants: Eleni Nisioti, Elias Masquil, Gautier Hamon (*correspondant*), Clément Moulin-Frier.

The intrinsically-motivated goal-conditioned learning paradigm is well-established in single-agent settings: by setting its own goals and pursuing them in an environment without external supervision, such an agent is able to acquire a wide diversity of skills [110]. Such agents are called autotelic, from the Greek words auto (self) and telos (end). What happens when you transfer the autotelic paradigm to multi-agent environments, where some skills may require the cooperation of multiple agents (lifting a heavy box for example)? This is the question we aimed at addressing in this project. We believe that multi-agent applications will benefit from agents that can autonomously discover and learn cooperative skills, but entail additional challenges to the ones found in single-agent settings: agents that independently set their own goals will have a very low probability of simultaneously sampling the same cooperative goal, which will make solving these goals difficult.

To explore this question, we implemented what we call cooperative navigation tasks in the Simple Playgrounds environments. This 2-D environment, illustrated on the left of figure 19, consists of a room with 6 landmarks on its walls and two agents that receive continuous-valued observations about the distance and angle to all landmarks and other agents and perform discrete-valued actions that control their angular velocity and longitudinal force. A navigation task is a description of the landmarks that need to be reached: some tasks are individual (for example "at least one agent reaches the red landmark") and some are cooperative (for example "at least one agent reaches the red landmark and at least one agent reaches the blue landmark"). Each autotelic agent is learning using the RL algorithm PPO and, at each training episode, chooses which goal to pursue by random sampling within the goal distribution. In addition to a policy conditioned on its goals, the agent also needs a reward function that indicates whether a goal is achieved (see the schematic on the right of figure 19 for an illustration of the main algorithmic components of autotelic agents). In this project, we assume that the two agents already know this reward function and focus on examining the process of decentralized goal selection.

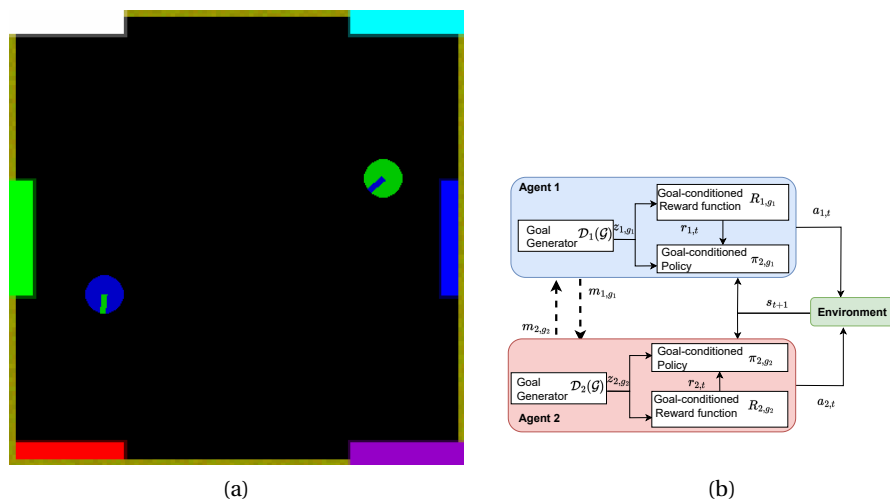


Figure 19: (left) The Cooperative landmarks environment consists of a room with two agents and six (right) landmarks. Landmarks are indicated as colored rectangles and navigation tasks are formulated as a set of landmarks the agent needs to navigate to, which might require coordination between the agents. (Right) Two autotelic agents in a multi-agent environment: the agents can exchange messages and condition their goal selection on them, which enables goal alignment.

Our empirical analysis of this set-up showed that goal alignment is important for solving the cooperative tasks we considered: agents that were independently sampling the goals failed to solve all tasks (see orange curve in figure 20) while agents whose goals were determined by a centralized process (blue curve) that guaranteed that the two agents are always pursuing the same goal performed optimally. We then wondered: can we achieve the same performance without requiring centralization? To achieve

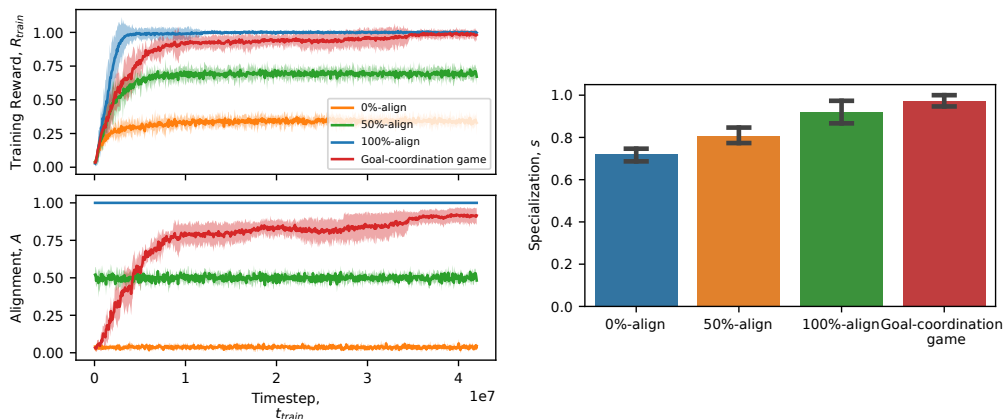


Figure 20: (Left) Performance for the 6-landmarks environment during training (Right) Specialization increases with alignment

this we designed a communication-based algorithm that enables a group to align its goals while remaining decentralized: at the beginning of an episode and before determining their goals the two agents exchange messages and then use these messages to condition their goal-selection (see dashed arrows in the schematic on the right of figure 19). This communication is asymmetric: one randomly chosen agent, the leader, uses its goal generator to choose which goal to pursue and then decides what message to transmit to the follower, which conditions its goal selection on the received message. We observed that the agents learn a communication protocol that leads to the alignment of cooperative goals, even though they were not directly incentivised to do so. They were both independently learning a protocol that maximised their individual rewards but, as we show in our experiments corresponding to figure 20, goal alignment was able to emerge from such decentralized learning. We called this algorithm the Goal-coordination game, as it was inspired from another emergent communication algorithm called the Naming game [189].

To get a better understanding of how alignment helps solve this task we measured specialization, which is the tendency of agents to always go the same landmark when there are two options. For example, if for the goal "at least one agent reaches the red landmark and at least one agent reaches the blue landmark" the first agent always go to red and the second goes to blue, then specialization is maximum. We empirically observed that specialization correlates with alignment and is an optimal strategy in our tasks (see right plot of figure 20).

In 2023, as advised by reviewers, additional experiments on baselines to compare were conducted. We also conducted further experiments trying to understand better what are the causes of this inability to learn in the independent (0%align) case.

This work was accepted at the Conference on Lifelong Learning Agents (CoLLAs) 2023 and presented at the poster session. A preprint version is available on HAL [52]. The source code for reproducing the experiments is available [at this link](#).

8.2.6 The SocialAI School: Insights from Developmental Psychology Towards Artificial Socio-Cultural Agents

Participants: Grgur Kovač (*correspondant*), Remy Portelas, Peter Ford Dominey, Pierre-Yves Oudeyer.

Developmental psychologists have long-established socio-cognitive abilities as a core aspect of human intelligence and development [200] [98]. Those abilities enable us to enter, participate in and benefit from human culture. Humans are then able to push our culture forward by improving on already existing cultural artifacts. This is referred to as the cumulative cultural evolution, and it has been argued that the most impressive of human achievements are the product of it [196]. It seems clear that to

construct artificial agents capable of interacting with and learning from humans one must equip them with socio-cognitive abilities enabling them to enter the human culture. This would enable artificial agents to benefit from our culture, and also push it forwards, i.e. to participate in the cumulative cultural evolution.

Current AI research mostly studies asocial settings or, in the case of Multi-Agent Reinforcement Learning, the emergence of culture (how culture emerges in the first place, rather than how one enters an already existing culture). Furthermore, this research is often done without a strong grounding in developmental psychology.

In this project, we follow the work of Michael Tomasello and Jerome Bruner who outlined a set of core socio-cognitive skills, e.g. social cognition (joint attention, theory of mind), referential and conventionalized communication, imitation, role-reversal, scaffolding, and many more [195, 98]. Importantly, they also showed the importance the (cultural) environment for cognitive development.

To introduce some of those concepts to the AI community, we created a tool for procedural generation of environments - The SocialAI school. With The SocialAI school, experiments studying those socio-cognitive abilities can be easily conducted and cognitive-science experiments reconstructed. Furthermore, The SocialAI School enables us to generate both multimodal environments (suited for RL agents) and pure text version of those environments (suited for LLM-based agents). An example of a SocialAI school environment is shown in figure 21. In it the peer is pointing towards the red box. The agent (the red triangle) has to infer this to mean that the apple is hidden inside the red box.

We conducted many experiments, here we outline a few more important ones. We experimented with multimodal RL agents. We tested generalization of inferring the meaning of referential communication (ex. the pointing gesture) to new scenarios/objects. We found that such a generalization is very hard for standard reinforcement learning agents. We show how a scaffolded environment helps with learning complex interaction sequences (formats). To show how cognitive science experiments can be recreated - we reconstruct a study of role reversal from [120]. Furthermore, we conducted experiments regarding other aspects of social-cognition: joint attention, imitation, perspective taking, etc. We experimented with LLM-based interactive agents. We show that a simple LLM-based agent can achieve some performance but still fails to generalize. This motivates future work for creating more complex LLM-based agents.

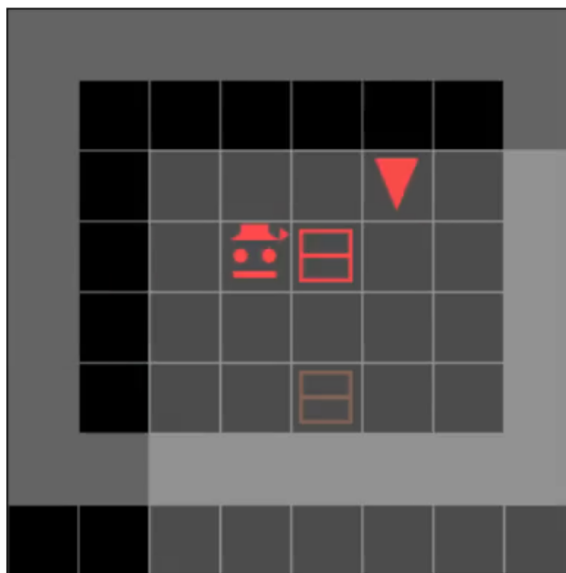


Figure 21: An example of an environment that can be created with the The SocialAI school. The red peer is pointing toward the red box. The task of the agent (the red triangle) is to infer this to mean that the red box contains an apple.

Most of this project was done in 2022. In this year, we extended the project by adding additional experiments with LLM-based agents and updating the implementation.

8.2.7 Value stability in Large Language Models

Participants: Grgur Kovač (*correspondant*), Masataka Sawayama, Remy Portelas, Peter Ford Dominey, Pierre-Yves Oudeyer.

There has been a growing body of research using Large Language Models (LLMs) to simulate individuals or human populations. Those studies usually focus on how a model can express some behavior or values and often overlook the underlying problem that LLMs are highly context-depenant. This problem is even more exacarbated by the use of psychological questionnaires, which were created with the assumption of human-like context-independence. In this project [74], we robustness of LLMs to seemingly unrelated perturbations in the context. Instead of evaluating models on some behavior (and testing robustness along the way), we study how a model's behavior changes over contexts as a primary question. In other words, we study and compare the value stability of LLMs along different contexts.

We leverage the PVQ questionnaire [104] associated with the Schwartz Theory of Basic Values [185], which defines ten values: Self-Direction, Stimulation, Hedonism, Achievement, Power, Security, Conformity, Tradition, Benevolence, Universalism. We study the LLMs abilities to simualte an individual (i.e. without defining a specific persona to simulate) and populations (by instructing the model to simulate various well-known fictional and real-world individuals). Different contexts are induced by simulating conversations on different topics: correcting grammar, writing a poem, playing chess, answering a history question and inventing a joke.

Following psychological methodology we evaluate three types of value stability: mean-level, rank-order, ipsative. We observe that LLMs exhibit low value stability - trivial context changes induce similar or bigger changes in value expression, than those observed in humans as a consequence of much more aggressive circumstances (e.g. 10 years of development or priming in humans compared to a topic change in LLMs). These results push furhter the point that psychological questionnaires administered to LLMs cannot be used and interpreted in the same way as with humans, i.e. much more attention must be given to studing context-dependance. Most importantly, they imply that rather than evaluating many questions from a single context (as is currently common), one should also evaluate same questions from many differt contexts. We propose metrics based on these types of value stability, and systematically compare models in terms of their value stability (see figure 22). To our knowledge we propose the first systematic comparison of many different models on their value stability.

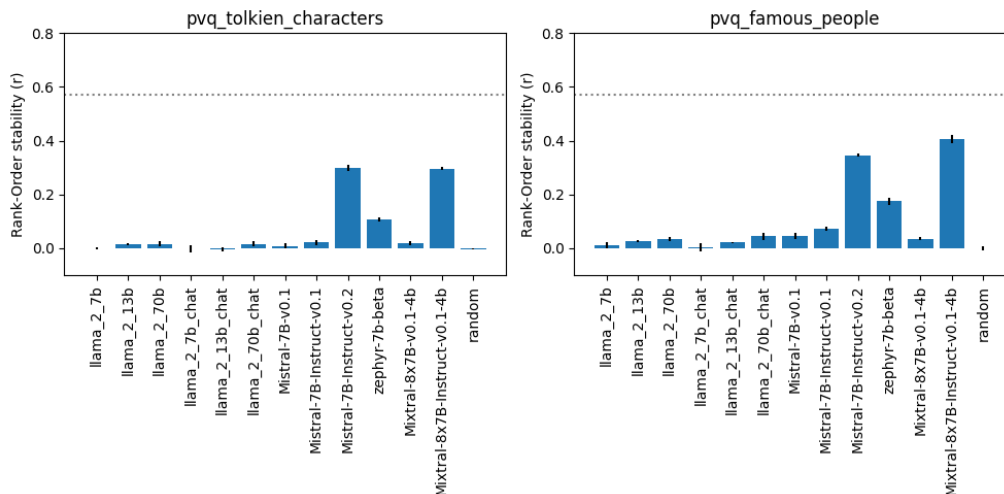


Figure 22: Rank-order stability of different Large Language Models

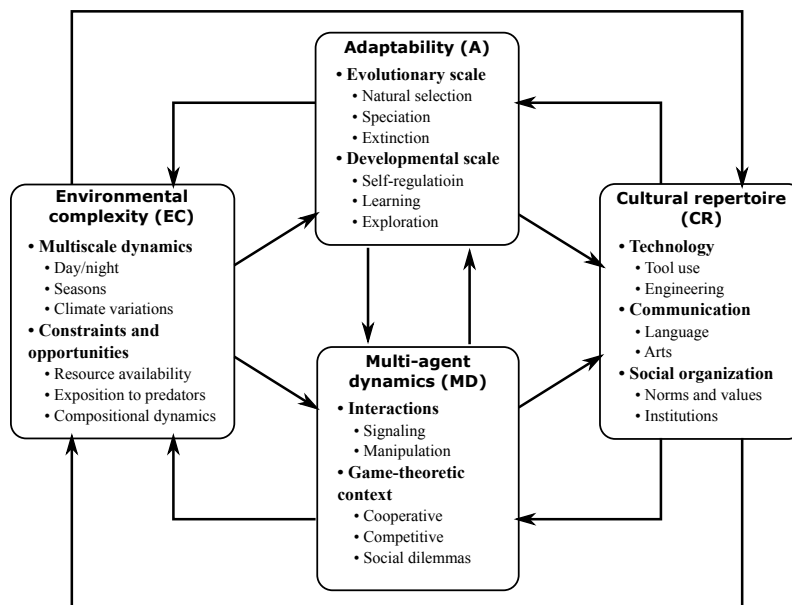


Figure 23: The ORIGINS framework identifies central components (boxes) and their interactions (arrows) driving Open-Ended Skill Acquisition, both in terms of its evolution from environmental complexity (roughly: left to right arrows) as well its open-ended aspect through feedback mechanisms (right to left arrows). The employed terminology reflects a diversity of mechanisms considered in both Artificial Intelligence and Human Behavioral Ecology.

8.3 Ecological Artificial Intelligence

8.3.1 Research perspective: The Ecology of Open-Ended skill Acquisition

Participants: Clément Moulin-Frier (*correspondant*), Eleni Nisioti, Pierre-Yves Oudeyer.

An intriguing feature of the human species is our ability to continuously invent new problems and to proactively acquiring new skills in order to solve them: what is called Open-Ended Skill Acquisition (OESA). Understanding the mechanisms underlying OESA is an important scientific challenge in both cognitive science (e.g. by studying infant cognitive development) and in artificial intelligence (aiming at computational architectures capable of open-ended learning). Both fields, however, mostly focus on cognitive and social mechanisms at the scale of an individual's life. It is rarely acknowledged that OESA, an ability that is fundamentally related to the characteristics of human intelligence, has been necessarily shaped by ecological, evolutionary and cultural mechanisms interacting at multiple spatiotemporal scales.

We have recently initiated a new research direction aiming at understanding, modeling and simulating the dynamics of OESA in artificial systems, grounded in theories studying its eco-evolutionary bases in the human species. For this aim, we have proposed a conceptual framework, called ORIGINS (illustrated Fig. 23 and developed in [154]), expressing the complex interactions between environmental, adaptive, multi-agent and cultural dynamics. This framework raises three main research questions:

- What are the ecological conditions favoring the evolution of autotelic agents?
- How to bootstrap the formation of a cultural repertoire in populations of adaptive agents?
- What is the role of cultural feedback effects in the open-ended dynamics of human skill acquisition?

The contributions described below are addressing some aspects of these research questions. Note that there might be a thematic overlap between the two last research questions outlined above and the previous section on Models of Cultural Evolution 8.2, where we also present related results.

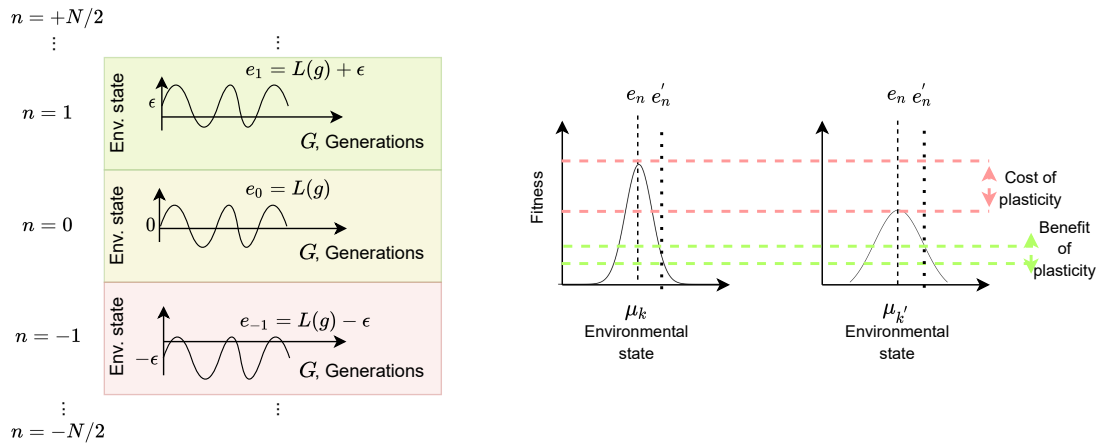


Figure 24: (Left) The latitudinal model we employ to describe how the environmental state varies across niches: a single climate function "sinusoidal curve e_n " L evolves identically for each niche and has a vertical offset equal to $\epsilon \cdot n$ for each niche with index n (Right) Modeling plasticity as a normal distribution $\mathcal{N}(\mu_k, \sigma_k)$. A non-plastic individual (k) has small σ_k and a high peak at their preferred niche, while a plastic individual (k') has large σ_k and a lower peak at their preferred niche. Fitness in a given niche n is computed as the probability density function of this tolerance curve at the environmental state e_n .

8.3.2 Evolution of plasticity and evolvability in variable environments

Participants: Eleni Nisioti (*correspondant*), Clément Moulin-Frier.

The diversity and quality of natural systems have been a puzzle and inspiration for communities studying artificial life. It is now widely admitted that the adaptation mechanisms enabling these properties are largely influenced by the environments they inhabit. Organisms facing environmental variability have two alternative adaptation mechanisms operating at different timescales: *plasticity*, the ability of a phenotype to survive in diverse environments and *evolvability*, the ability to adapt through mutations. Although vital under environmental variability, both mechanisms are associated with fitness costs hypothesized to render them unnecessary in stable environments. In this work, we aimed at studying the interplay between environmental dynamics and adaptation in a minimal model of the evolution of plasticity and evolvability.

To achieve this we designed a simulation environment that attempts to capture the spatial and temporal heterogeneity of real-world environments while keeping the computational complexity low: the user can choose the number of niches, which are arranged in a simple longitudinal model, and a climate function that captures the temporal variation of environmental conditions, which are arranged based on a simple longitudinal model (see left of figure 24 for an illustration of the environment). We defined the evolvability of an agent as its mutation rate and capture plasticity using tolerance curves, a tool developed in ecology [124]. Tolerance curves (which we visualize on the right of Figure 24) have the form of a Gaussian whose mean shows the preferred environmental state of an individual and the variance its plasticity, i.e., its ability to survive under different environmental conditions. This figure also illustrates the cost and benefit of plasticity. If both individuals are at their preferred niche, which coincides with the environmental state, then the plastic individual has lower fitness than the non-plastic (cost of plasticity). If the actual environmental state differs significantly from the preferred one, the plastic individual has higher fitness (benefit of plasticity).

We conducted an extensive empirical study in this environment that aimed at disentangling the effects of different mechanisms: we studied three types of climate functions (stable, sinusoid, noisy) and two types of evolutionary selection pressures (survival of the fittest and niche-limited competition) and environments where the number of niches varies from 1 to 100. Through these simulations we showed that environmental dynamics affect plasticity and evolvability differently and that the selection

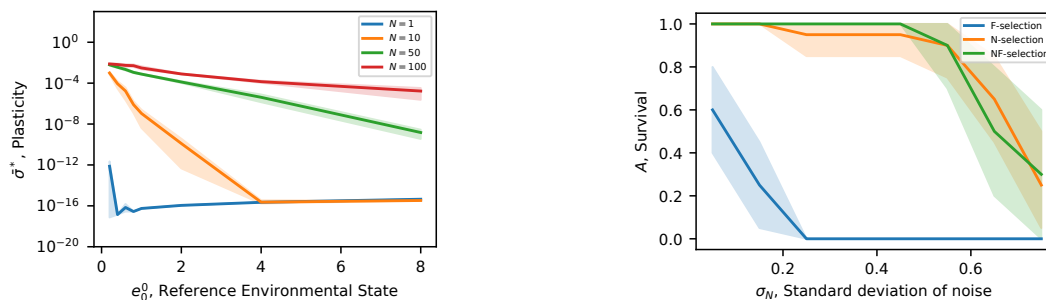


Figure 25: An example of the conclusions derived by our evolutionary study: (left) the plasticity of a population evolving under fitness-based selection and niche-based competition when we vary the number of niches and value of the stable climate function. We observe that plasticity is most favored in environments with low values of climate (sparser fitness) and larger number of niches (right) ability of populations to survive under different selection mechanisms and levels of noise. Populations that do not employ niche-limited competition (F-selection) are not robust.

mechanism matters: a) in stable environments with a large number of niches when both selection-based fitness and niche-limited competition (we call this method NF-selection) are activated, plasticity remains high despite its cost (see left plot in Figure 25) ; b) in a noisy environment introducing niche-limited competition (N-selection and NF-selection) makes populations capable of resisting larger amounts of noise (see right plot in Figure 25). We presented our work at GECCO 2022 [160] and open-sourced the software for reproducing our simulations [in this repository](#). A follow-up of this work has been published at the ALIFE 2023 conference [53], where we introduced mechanisms of niche construction to this model.

8.3.3 Open-ended recipe crafting through meta reinforcement learning

Participants: Gautier Hamon (*correspondant*), Eleni Nisioti, Clément Moulin-Frier.

As a first step towards studying the evolution of open-ended skill acquisition in artificial agents, we studied the environmental conditions favoring the systematic exploration of combinatorial recipes involving various objects. By *combinatorial recipe*, we mean the ability of agents to combine objects in the environment in order to create new ones (in the spirit of the Minecraft video game), some of these crafted objects being associated with a reward. In this work, the training of an agent uses *meta reinforcement learning* where an outer loop, equivalent to an evolutionary mechanism, meta-learns the parameters of an inner loop which can be seen as a developmental mechanism (where the agent acquires skills during its lifetime by interacting with the environment). In the current setup we use RL^2 as our meta-learning algorithm [118, 201] which has already been used for acquiring behaviors efficiently balancing exploration and exploitation in a simple navigation task. Other work studied how different conditions in a bandit favor the evolution of innate vs learned behaviors [140].

Our experiments with recipe crafting are inspired by [the little alchemy game](#). The difference with previous works in similar environments (e.g. 8.2.1) is that at every episode the structure of the recipe is randomly chosen. The agent therefore cannot predict what recipes will be rewarding and have to explore different combinations of objects in order to find the rewarding ones. The agent should also memorize the successful and unsuccessful combinations in order to explore and exploit efficiently.

Our preliminary results use both in a vectorized version of the game (where the agents action are only to choose the 2 objects to combine) and an embodied gridworld version (where the agent has to move, grab objects and put them on top of others in order to craft new ones). In both of these cases, the training efficiently meta learns an exploration/exploitation strategy which is to try new recipes (most of the time it does not try non working recipes more than once) until it finds the rewarding ones and then simply exploits them by making them over and over.

Further work will study how we can change the environment/training in order to evolve open-ended

exploration strategies where an agent will continuously explore new recipes even if it has already found rewarding ones, as a way to be better prepared for future changes in the recipe structure. We hypothesize that such an intrinsic motivation to explore for the sake of acquiring knowledge of the environment, even in the absence of external rewards, could evolve by introducing drastic changes of recipes which the agent has to anticipate in order to survive. During the project, we switched from evolutionary algorithm and Recurrent neural networks to reinforcement learning and transformers. This allowed for more complex environments with more possibilities. We also obtained preliminary results with agents exploring the environment to gain information for the future.

This work uses the [JAX python library](#) for both the model/machine learning part and the environment simulation. JAX allows easy parallelization and fast GPU computation and so learning it through this project will be useful for later projects.

We plan to submit a paper on these experiments in 2024.

8.3.4 Evolving Reservoirs for Meta Reinforcement Learning

Participants: Corentin Leger (*correspondant*), Gautier Hamon, Eleni Nisioti, Xavier Hinaut, Clément Moulin-Frier.

This contribution was realized in the context of the Master internship of Corentin Léger in 2023, as a collaboration between C. Moulin-Frier from the Flowers team and Xavier Hinaut from the Mnemosyne team. It led to a paper which has been accepted at the Evostar conference [75].

Animals demonstrate remarkable adaptability to their environments, a trait honed through the evolution of their morphological and neural structures [194][170]. Animals are born equipped with both hard-wired behavior routines (e.g. breathing, motor babbling) and learning capabilities to adapt from experience. The costs and benefits of evolving hard-wired behaviors vs. learning capabilities depends on different factors, a central one being the level of unpredictability of environmental conditions across generations [190][136]. Phenotypic traits addressing environmental challenges that are shared across many generations are more likely to evolve hard-wired (e.g. breathing), while traits whose utility can hardly be predicted from its utility in previous generations are likely to be learned through individual development (e.g. learning a specific language).

This prompts an intriguing question: How can neural structures, optimized at an evolutionary scale, enhance the capabilities of agents to learn complex tasks at a developmental scale? To address this question, we propose to model the interplay between evolution and development as two nested adaptive loops: evolution optimizes the generation of neural structures through natural selection over generations, shaping developmental learning during an agent's lifetime (Fig. 26).

More precisely, at the evolutionary scale (the outer loop), we use an evolutionary algorithm to optimize a genome specifying Hyper Parameters of Reservoirs [181]. At a developmental scale (the inner loop), a RL agent equipped with a generated reservoir learns an action policy to maximize cumulative reward in a simulated environment. Thus, the objective of the outer evolutionary loop is to optimize macro properties of reservoirs in order to facilitate the learning of an action policy in the inner developmental loop. See Fig.27 for an overview of the method.

Using this computational model, we run experiments in diverse simulated environments, e.g. 2D environments where the agent learns how to balance a pendulum and 3D environments where the agent learns how to control complex morphologies. These experiments provide support to three main hypotheses for how evolved reservoirs can affect intralife learning. First, they can facilitate solving partially-observable tasks, where the agent lacks access to all the information necessary to solve the task. In this case, we test the hypothesis that the recurrent nature of the reservoir will enable learning to infer the unobservable information. Second, it can generate oscillatory dynamics useful for solving locomotion tasks. In this case, the reservoir acts as a meta-learned CPG. Third, it can facilitate the generalization of learned behaviors to new tasks unknown during the evolution phase.

This work was accepted at EvoApplications 27th European Conference on the Applications of Evolutionary and bio-inspired Computation (EvoApps 2024).

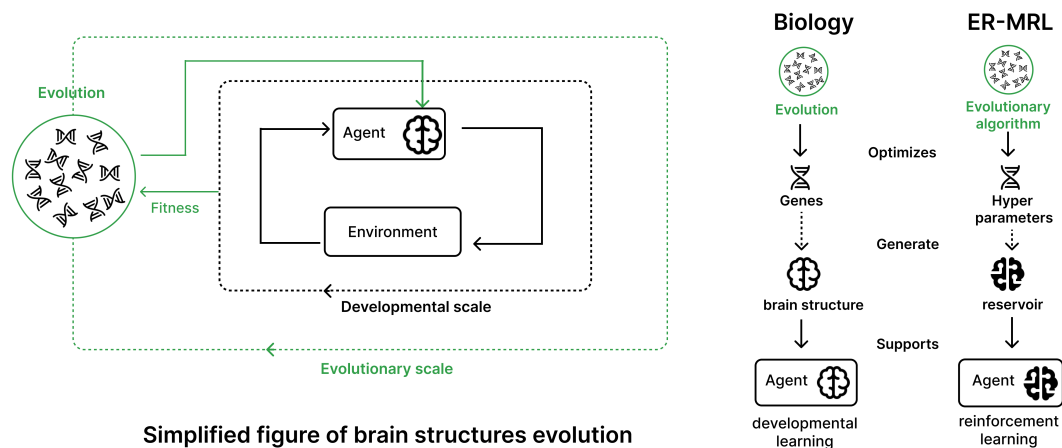


Figure 26: A simplified view of the evolution of brain structures (left) and the parallel with our computational approach (right). We can observe on the left of the figure the interplay between two loops : an evolutionary one that modifies the generating parameters of neural structures, and a developmental one where agents equipped with such neural structures learn to interact with their environment. We propose a computational framework (right) where an evolutionary algorithm optimizes hyperparameters that generate neural structures called reservoirs. These reservoirs are then integrated into RL agents that learn an action policy to maximize their reward in an environment

8.3.5 Emergence of Collective Open-Ended Exploration from Decentralized Meta-Reinforcement Learning

Participants: Richard Bornemann, Gautier Hamon (*correspondant*), Eleni Nisioti, Clément Moulin-Frier.

In this work we want to further investigate the emergence of cooperative exploration strategies of decentralized agents by training them on a open ended distribution of tasks. To this end we introduce a novel environment [28](#) which is conceptually simple yet allows for a complex open ended procedurally generated task space by dynamically combining multiple subtasks sampled from five task types to form a task tree which needs to be solved sequentially, akin to the notion of recipes in [\[91\]](#). We train two agents parameterized by independent recurrent neural networks and optimized using standard proximal policy optimization. As no information is given to the agents about which subtasks have been sampled or how and in which order they should be solved, the agents have to develop general strategies for exploring the environment, effectively learning how to learn from the information obtained by interacting with the environment throughout the episode, in order to solve novel tasks. We show that training independent decentralized agents on only multi agent episodes leads to sub-optimal behavior of the agents, primarily due to the problem of credit assignment when rewards are shared between agents. We propose to include single agent episodes during training to force the agents to learn to solve tasks on their own without relying on any help from other agents. We find that training on a mixture of single and multi agent episodes increases the agents individual performance while simultaneously decreasing the individual performance differences between the agents, leading to a strong improvement in performance in multi agent tasks.

Using this approach we find that decentralized agents trained in our procedurally generated environment learn a powerful collective exploration strategy, allowing them to solve over 70 percent of task trees encountered during training. Moreover, these powerful exploration capabilities lead to strong generalization performance when confronted with objects unseen during training, as well as on novel tasks which require complex coordination to be solved successfully at test time. Additionally we show that the learned collective exploration strategies extend to the open ended task setting, enabling the agents to effectively generalize to task trees with a depth of six, featuring an increased complexity of subtasks, despite being initially trained on task trees comprising only three subtasks.

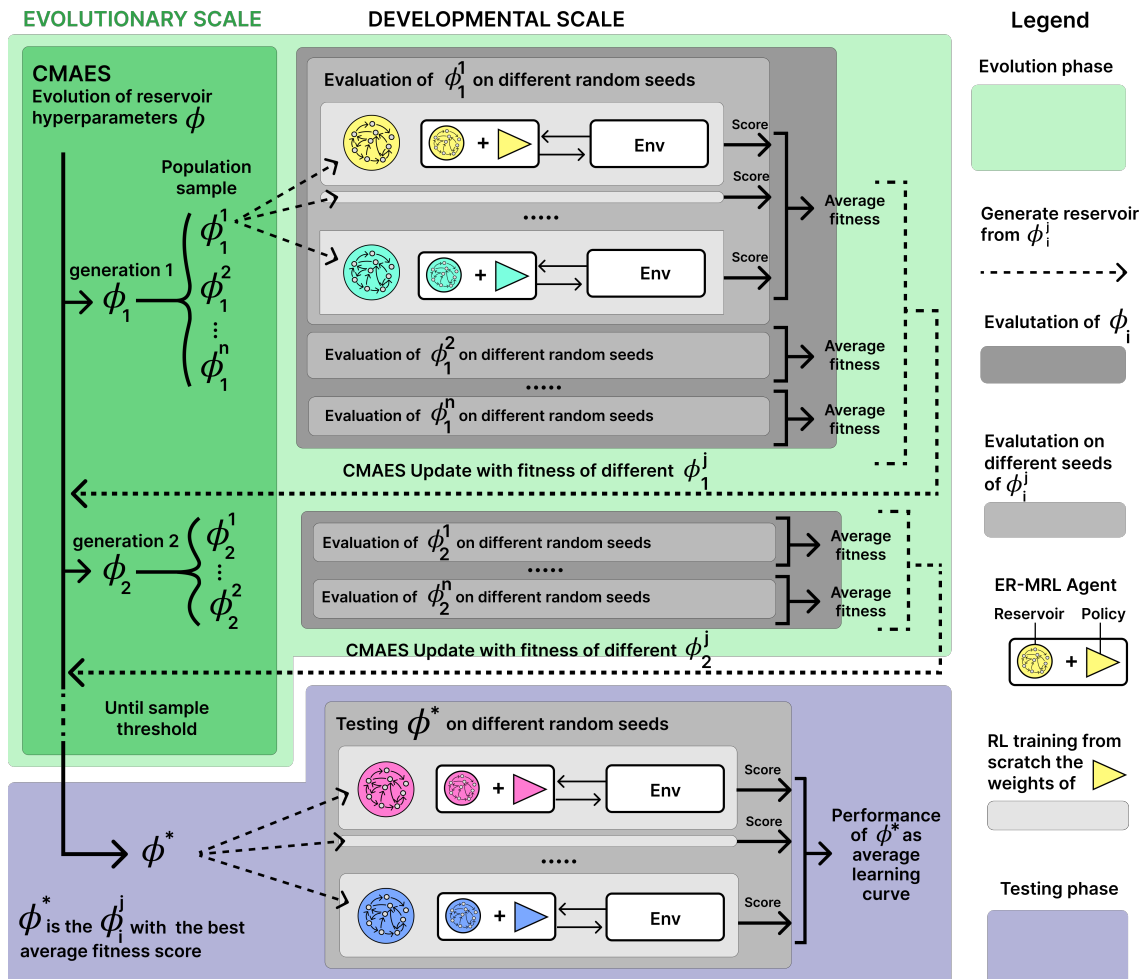


Figure 27: A simplified view of the evolution of brain structures (left) and the parallel with our computational approach (right). We can observe on the left of the figure the interplay between two loops : an evolutionary one that modifies the generating parameters of neural structures, and a developmental one where agents equipped with such neural structures learn to interact with their environment. We propose a computational framework (right) where an evolutionary algorithm optimizes hyperparameters that generate neural structures called reservoirs. These reservoirs are then integrated into RL agents that learn an action policy to maximize their reward in an environment

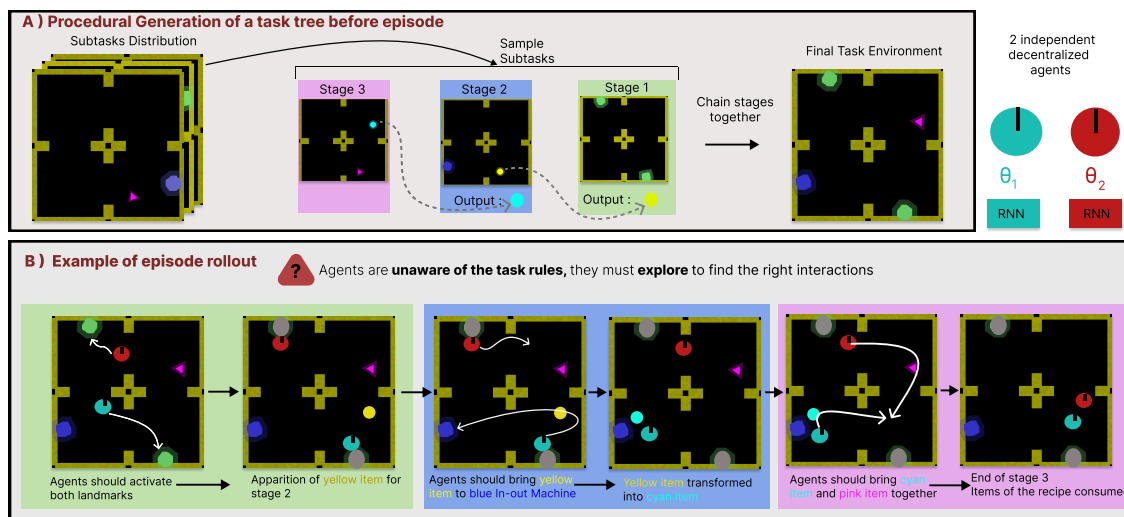


Figure 28: Task Tree Sampling and Episode Rollout. **A)** shows the task tree sampling process. First three subtasks are sampled from a distribution of subtasks, one for each stage of the task tree. **B)** shows an example of a single episode rollout. The agents have to complete the subtasks sequentially in order to create objects which are needed by the subtasks in later stages. Since a new task tree with different subtasks is sampled at the beginning of each episode and no information about the subtasks is given to the agents, the agents have to explore the environment and interact with all present objects so solve the subtask at each stage.

This work was presented as a poster at the Agent Learning in Open-Endedness (ALOE) workshop at NeurIPS 2023. Videos of the agents behaviors can be found on [our companion website](#)

8.3.6 Cooperative control of environmental extremes by artificial intelligent agents

Participants: Martí Sánchez Fibla, Clément Moulin-Frier (*correspondant*), Ricard Solé.

This contribution is the result of a collaboration between Ricard Solé and Martí Sánchez-Fibla from the University Pompeu Fabra (Barcelona, Spain) and Clément Moulin-Frier (Flowers, Inria). A preprint is available [78] and it a paper has been submitted to PNAS.

Humans have been able to tackle biosphere complexities by acting as ecosystem engineers, profoundly changing the flows of matter, energy and information. This includes major innovations that allowed to reduce and control the impact of extreme events. Modelling the evolution of such adaptive dynamics can be challenging given the potentially large number of individual and environmental variables involved. This paper shows how to address this problem by using fire as the source of external, bursting and wide fluctuations. Fire propagates on a spatial landscape where a group of agents harvest and exploit trees while avoiding the damaging effects of fire spreading. The agents need to solve a conflict to reach a group-level optimal state: while tree harvesting reduces the propagation of fires, it also reduces the availability of resources provided by trees. It is shown that the system displays two major evolutionary innovations that end up in an ecological engineering strategy that favours high biomass along with the suppression of large fires. The implications for potential A.I. management of complex ecosystems are discussed.

The computational model is illustrated Fig. 29 and the main results in Fig. 30.

8.3.7 Eco-evolutionary Dynamics of Non-episodic Neuroevolution in Large Multi-agent Environments

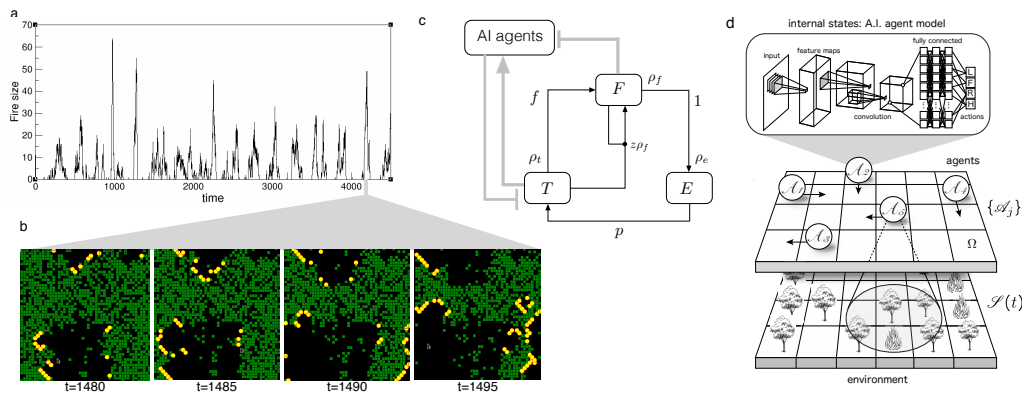


Figure 29: Forest fire dynamics in time and space. In figure (a) a typical time series of the number of burned sites in a forest fire model (FFM) is displayed for a square lattice with $L = 50$ and parameters $p = 0.003$, $f = 0.00003$. The number of sites burning (the fire size) shows marked bursting dynamics. Four spatial snapshots are shown in (b) associated with a fire burst. Here, green, yellow and black correspond to trees, fires and ashes (empty sites) respectively. The basic set of rules is summarized in (c) using black arrows. In our model, we add a set of AI agents whose interactions with the environment are marked with grey arrows where positive and negative interactions are indicated as \rightarrow and \leftarrow , respectively. They benefit from trees but get punished by fire spreading, and can modify tree density by harvesting trees. In (d) we summarize the levels of interaction between forest fire dynamics and its control by neural agents. The bottom layer defines the observed spatial pattern of states of the Forest Fire Model (FFM), which changes stochastically while can be affected by the action of agents (middle layer) that have a limited observation range (indicated as a circle in the bottom layer) and can take decisions about their movement and harvesting trees locally. Each agent (upper layer) makes decisions (implements an action policy, mapping observed states to actions) by means of a convolutional neural network trained with Reinforcement Learning (RL). The RL process eventually defines the behavioural pattern displayed by the agent, which translates into a set of potential actions (*LeftTurn*, *RightTurn*, *ForwardMove*, *Harvest*) in response to the local environment.

Participants: Gautier Hamon (*correspondant*), Eleni Nisioti, Clément Moulin-Frier.

This work focuses on eco-evolutionary dynamics where "organisms are not solely products but, by modifying their niche and therefore its associated fitness landscape, are also causes of evolution" [139]. The main objective of this paper is to propose a method for studying large-scale eco-evolutionary dynamics in agent-based simulations with a reasonable level of biological and ecological plausibility. For this aim, we implement a system with the following properties (see Fig. 31 for illustration):

- *Non-episodic simulation environment with complex intrinsic dynamics.* We model our environment after common-pool resource (CPR) appropriation problems, where a group of agents competes for finite resources. We extend an existing environment of CPR appropriation [171] with the presence of multiple niches, where resources regrow proportionally to the density of nearby resources at different rates in different regions of the environment (Fig 31). We prevent any environment or population reset during a whole simulation run, enabling coupled environmental and population dynamics leading to complex eco-evolutionary feedback effects.
- *Continuous neuroevolution in a large, size-varying agent population* The environment contains thousands of agents, each controlled by a neural network whose weights are optimized using neuroevolution [187]
- *Physiology-driven death and reproduction* There is no notion of rewards, agents are instead equipped with a physiological system modulating their energy level according to the resources they consume, in a non-linear way. At the evolutionary scale, agents reproduce as long as they are able to maintain their energy level within a reasonable range and die if this level goes below a minimum threshold. This is departure from the notion of fitness-based selection and more in line with a minimal criterion selection [94]. Note that the population size can vary with time.

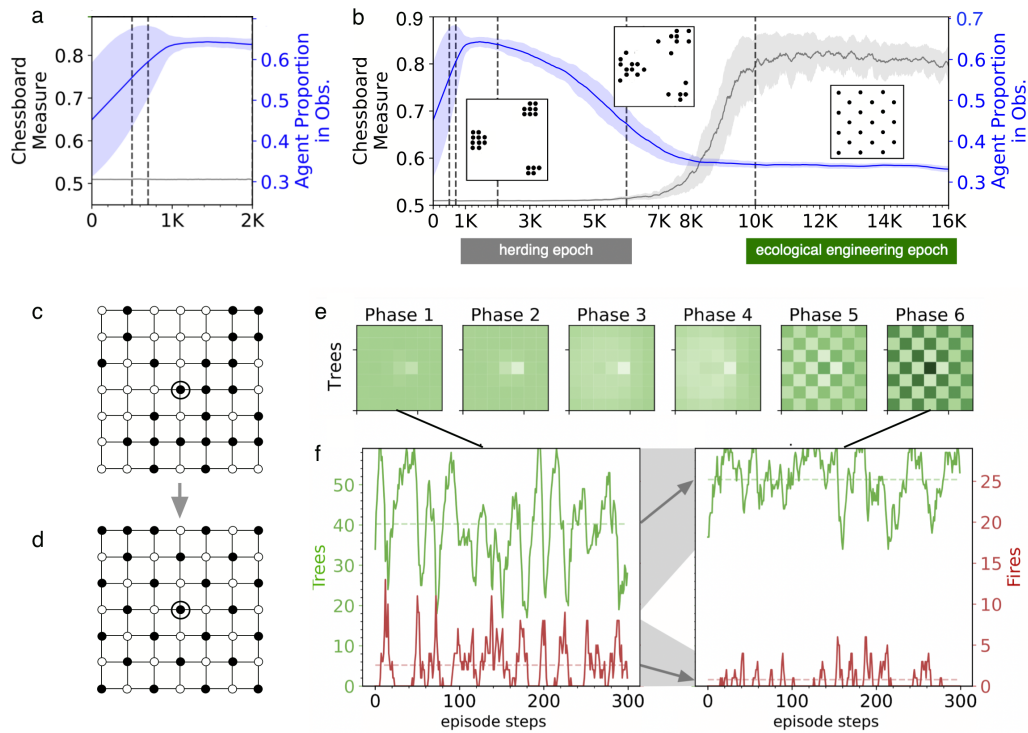


Figure 30: Characterization of the observed cooperative phase transitions. In (a-b), the time series for the herding (blue) and ecological-engineering (grey) measures are displayed. The herding measure characterizes the agents' tendency to form dense herds. It is illustrated in the three insets in (b), sketching typical spatial arrangements of agents: a high herding measure indicates that agents form dense herds (left inset);. In contrast, lower herding measures indicate that they are more uniformly spread in the environment (middle and right inset). The ecological-engineering measure characterizes the agents' ability to create a structured pattern of trees limiting fire propagation. Patterns of trees corresponding to a low and high ecological-engineering measure are illustrated in (c) and (d), respectively. The open circle in the middle indicates an agent, with black and white circles indicating the presence or absence of trees, respectively. Intuitively, the high measure resulting from (d) corresponds to a perfect chessboard pattern preventing fire propagation (which only propagates in the horizontal and vertical dimensions) while maximizing the number of trees. The progressive formation of this structured pattern of trees is displayed in (e), showing the average density of trees in all agent's neighbourhoods during the six emerging phases. In (f), we observe the FFM dynamics in the first and last episodes of the simulation. This demonstrates that the agent population managed to increase (resp. decrease) the average number of trees (resp. fire) as indicated by the horizontal dotted lines while reducing the fluctuation range of trees and fires.

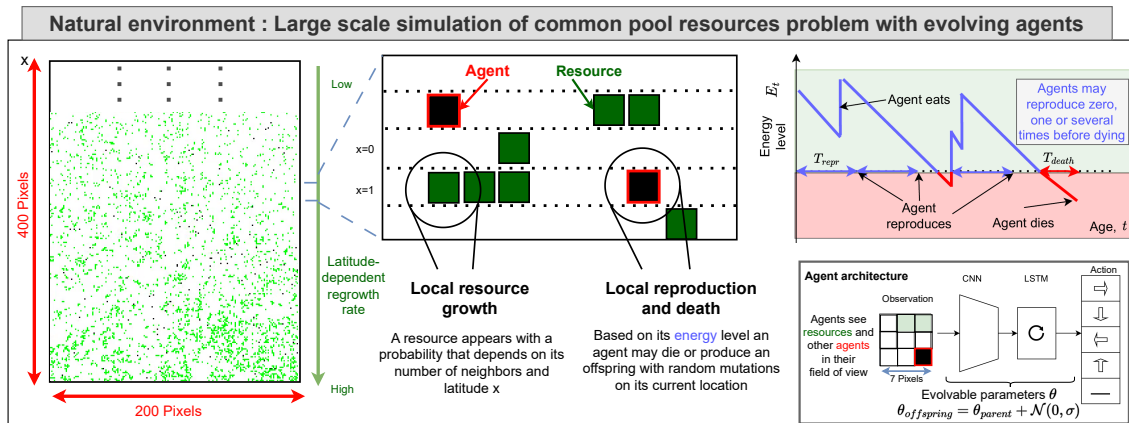


Figure 31: Our simulation environment (Left) is an extension of the Common Pool Resource (CPR) environment [171, 142]: a two-dimensional grid-world where some cells contain resources (in green) that the agents (in black) can collect. Resources grow depending on the presence of other resources around them (local growth, Middle) with an additional very sparse spontaneous growth, which means that over-consumption may lead to their local depletion. We introduce a latitudinal model of resource regrowth similar to 8.3.2. We prevent any environment and population reset during a whole simulation, enabling continual eco-evolutionary dynamics to take place. Each agent may reproduce or die according to a physiological model modulating its energy level as a function of life time and resource consumption (Top-Right). The population size varies during the simulation according to the current amount of available resources and the current ability of agents to collect them. Evolution occurs through the mutation of a parent's network weights when it produces an offspring.

In addition to experiments conducted in the large environment presented, we also conduct experiments in "lab environment" (as opposed to the "natural environment") to isolate the study of certain behavior (which are often intertwined with a lot of dynamics in the natural environment).

One interesting results of these simulation is the emergence of sustainable foragers which as shown in lab environment Fig.32 tends to not overconsume when there is enough resource in their neighbourhood. This allows to keep a certain amount of resource to spread which is therefore beneficial for their future survival as well as the survival of their offspring. (as there is no reset of the environment)

This work was presented as a poster to the Genetic and Evolutionary Computation Conference (GECCO) 2023.

8.4 Generative AI and educational technologies

8.4.1 Fostering curiosity and meta-cognition in children using LLM-based conversational agents

Participants: Pierre-Yves Oudeyer, H el ene Sauz eon (*correspondant*), Mehdi Alami, Rania Abdelghani, Didier Roy, Edith Law, Chlo e Desvaux.

Since 2019 (IDEX cooperation fund between the University of Bordeaux and the University of Waterloo, Canada) and the recent creation of CuriousTECH associate team in 2022 (led by the Flowers team and involving F. Lotte from the Potioc team and M. Fernandes and E. Law from the Waterloo University), we continue our work on the development of new curiosity-driven interaction systems. Substantial progress has been made in this area of application of FLOWERS works (see the website of CuriousTECH team: <https://flowers.inria.fr/curioustech-associate-team/>).

Toward a better understanding of neurocognitive of curiosity-based learning For a better understanding of basic mechanisms of curiosity-based learning, three studies have been completed. The first study regards a new interactive educational application to foster curiosity-driven question-asking in children. This study has been performed during the Master 2 internship of Mehdi Alami co-supervised by H. Sauz eon, E. Law and PY Oudeyer. It addresses a key challenge for 21st-century schools, i.e.,

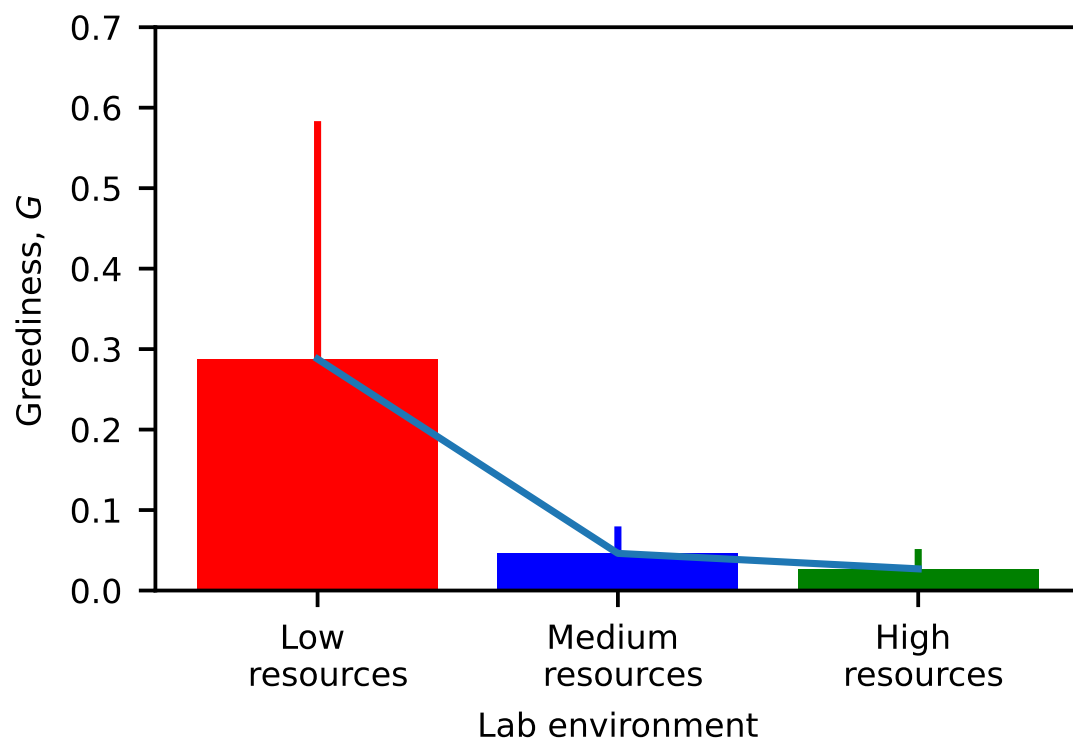


Figure 32: Greediness of a sustainable forager agent across evaluation environments that differ in the amount of resources. Sustainable agents are far less greedy in environments where there is a certain amount of resources available. This strategy allows to keep resources so that they spread and avoid overdepletion of resources.

teaching diverse students with varied abilities and motivations for learning, such as curiosity within educational settings. Among variables eliciting curiosity state, one is known as « knowledge gap », which is a motor for curiosity-driven exploration and learning. It leads to question-asking which is an important factor in the curiosity process and the construction of academic knowledge. However, children questions in classroom are not really frequent and don't really necessitate deep reasoning. Determined to improve children's curiosity, we developed a digital application aiming to foster curiosity-related question-asking from texts and their perception of curiosity. To assess its efficiency, we conducted a study with 95 fifth grade students of Bordeaux elementary schools. Two types of interventions were designed, one trying to focus children on the construction of low-level question (i.e. convergent) and one focusing them on high-level questions (i.e. divergent) with the help of prompts or questions starters models. We observed that both interventions increased the number of divergent questions, the question fluency performance, while they did not significantly improve the curiosity perception despite high intrinsic motivation scores they have elicited in children. The curiosity-trait score positively impacted the divergent question score under divergent condition, but not under convergent condition. The overall results supported the efficiency and usefulness of digital applications for fostering children's curiosity that we need to explore further. The overall results are published in CHI'20 [84]. In parallel to these first experimental works, we wrote this year a review of the existing works on the subject [101].

The second study investigates the neurophysiological underpinnings of curiosity and the opportunities of their use for Brain-computer interactions [85]. Understanding the neurophysiological mechanisms underlying curiosity and therefore being able to identify the curiosity level of a person, would provide useful information for researchers and designers in numerous fields such as neuroscience, psychology, and computer science. A first step to uncovering the neural correlates of curiosity is to collect neurophysiological signals during states of curiosity, in order to develop signal processing and machine learning (ML) tools to recognize the curious states from the non-curious ones. Thus, we ran an experiment in which we used electroencephalography (EEG) to measure the brain activity of participants as they were induced into states of curiosity, using trivia question and answer chains. We used two ML algorithms, i.e. Filter Bank Common Spatial Pattern (FBCSP) coupled with a Linear Discriminant Algorithm (LDA), as well as a Filter Bank Tangent Space Classifier (FBTSC), to classify the curious EEG signals from the non-curious ones. Global results indicate that both algorithms obtained better performances in the 3-to-5s time windows, suggesting an optimal time window length of 4 seconds to go towards curiosity states estimation based on EEG signals. These results have been published [85]

Finally, the third study investigates the role of intrinsic motivation in spatial learning in children [79]. In this study, the state curiosity is manipulated as a preference for a level of uncertainty during the exploration of new environments. To this end, a series of virtual environments have been created and is presented to children. During encoding, participants explore routes in environments according the three levels of uncertainty (low, medium, and high), thanks to a virtual reality headset and controllers and, are later asked to retrace their travelled routes. The exploration area and the wayfinding, i.e. the route overlap between encoding and retrieval phase, (an indicator of spatial memory accuracy) are measured. Neuropsychological tests are also performed. Preliminary results showed that there are better performances under the medium uncertainty condition in terms of exploration area and wayfinding score. These first results supports the idea that curiosity states are a learning booster [79].

Curiosity-driven educational technologies : Combining conversational agents and LLM as lever for field scaling up At the end of 2020, we started an industrial collaboration project with EvidenceB on this topic (CIFRE, contract of Rania Abdelghani validated by the ANRT). The overall objective of the thesis is to propose new educational technologies driven by epistemic curiosity, and allowing children to express themselves more and learn better. To this end, a central question of the work will be to specify the impact of self-questioning aroused by states of curiosity about student performance. Another objective will be to create and study the pedagogical impact of new educational technologies in real situations (schools) promoting an active education of students based on their curiosity. To this end, a web platform called 'Kids Ask' has been designed, developed and tested in three primary schools. The tool offers an interaction with a conversational agent that trains children's abilities to generate curiosity-driven questions and use these questions to explore a learning environment and acquire new knowledge. The results suggest that the configuration helped enhance children's questioning and exploratory behaviors; they also show that learning progress differences in children can be explained by the differences in their

curiosity-driven behaviors [86].

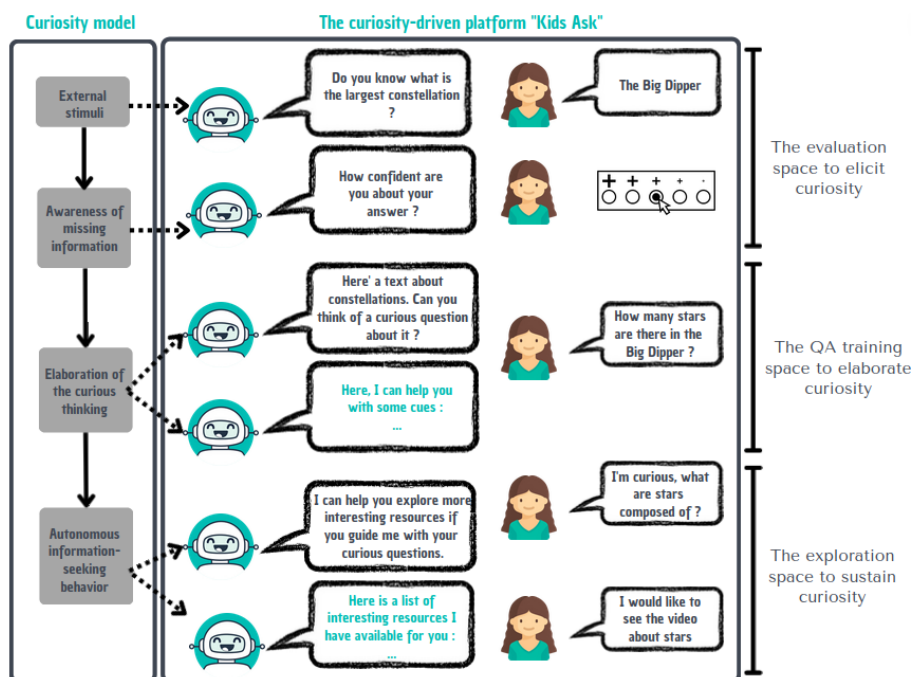


Figure 33: Illustration of a conversational agent's strategies in the different work spaces of the "Kids Ask" platform

Despite showing pedagogical efficiency, the method used in the first study of this PhD is still very limited since it relies on generating curiosity-prompting cues by hand for each educational resource in order to feed the "discussion" with the agent, which can be a very long and costly process. For this reason, a logical follow-up to scale-up and generalize this study was to explore ways to automate the said conversational agents' behaviors in order to facilitate their implementation on a larger scale and for different learning tasks. More particularly, we move towards the natural language processing (NLP) field and the large language models (LLMs) that showed an impressive ability in generating text that resembles the way people write.

In this context, we study using the recent LLM GPT-3 to implement conversational agents that can prompt children's curiosity about a given text-based educational content, by proposing some specific cues. We investigate the validity of this automation method by comparing its impact on children's divergent question-asking skills with respect to the hand-crafted condition we had in our previous work. In a second step, we explore using GPT-3 to propose a new curiosity-prompting behavior for our agent that aims to better support the children's needs of competence, autonomy and relatedness during the question-asking training.

The study was conducted in two primary schools with 75 children aged between 9 and 11. Our first results suggest the validity of using GPT-3 to facilitate the implementation of curiosity-stimulating learning technologies. Indeed, children's performance was similar between the conditions where they had hand-generated or GPT-3-generated cues. In a second step, we also found that GPT-3 can be efficient in proposing the relevant cues that leave children with more autonomy to express their curiosity [33] (publication in process).

Finally, as a follow-up direction to this line of work, we design new digital interventions, that focus on eliciting the metacognitive mechanisms involved in the stimulation and continuity of curiosity, and not just giving the tools to pursue it as done in the previous studies. For this, we use findings from the theories explaining curiosity in order to break this latter up into a set of relevant metacognitive skills. We then take an operational approach to propose adequate digital metacognitive exercises for each one of the said skills (i.e. exercises to identify uncertainty, generate hypotheses etc). We aim to implement this set of metacognitive exercises and investigate its impact on children's abilities to initiate and maintain

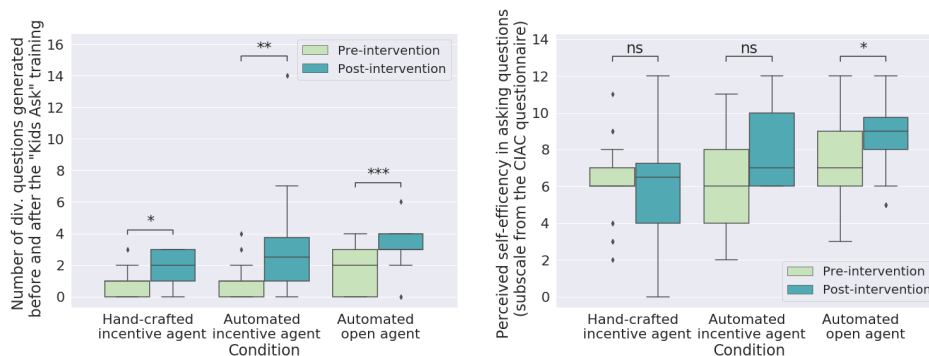


Figure 34: Left: Participants from the three conditions were able to improve their divergent QA abilities after the "Kids Ask" interaction, as shown by the divergent QA fluency test pre- and post-training. **Right:** Children's perception of their QA self-efficacy changed more positively with the intervention for those who interacted with the automated agents.

curious behaviors. We would also be interested in investigating the impact of such training on the learning progress children can achieve. A first study has been conducted with two classrooms to evaluate the accessibility of this new training and the impact on metacognitive efficiency, curiosity-driven question-asking and learning. Our first results being rather positive, we aim to recruit a bigger sample size to validate them [41].

Thanks to the building of primary schools networks (Submission of Léa-Ifé project to ENS-Lyon Call), the next step of this work is to study this digital metacognitive intervention in more ecological settings, specifically when administered by teachers, since this reflects the classical classroom setting. The aim is therefore to support teachers in assimilating the intervention to facilitate its transfer into real classrooms. This has already been initiated in collaboration with the Académie de Bordeaux and elementary school teachers of Bordeaux Métropole, as the first steps of the thesis project that started in October (Chloé Desvaux - Université de Bordeaux). The efficacy of the ecological digital intervention is to be compared with previous results. Another objective of this thesis project is to assess the characteristics of children's curious behaviors more closely. The primary focus being divergent and creative properties and their implication in learning.

On another subject, we also started investigating the importance of curiosity-related metacognitive skills on students' use of the GenAI (Generative AI) tools during learning. Indeed, in [59], we argue about the importance of developing children's sense of critical thinking, epistemic vigilance, etc in order to allow a more active and informed use of these tools during learning. Such skills can help have a more realistic expectations of such tools and evaluate their outputs before integrating them in one's beliefs. A study aiming to understand how children use these tools to solve learning problems is in progress (piloting stage).

8.4.2 Supporting Qualitative Analysis with Large Language Models: Combining Codebook with GPT-3 for Deductive Coding

Qualitative analysis of textual contents unpacks rich and valuable information by assigning labels to the data. However, this process is often labor-intensive, particularly when working with large datasets. While recent AI-based tools demonstrate utility, researchers may not have readily available AI resources and expertise, let alone be challenged by the limited generalizability of those task-specific models. In this study, we explored the use of large language models (LLMs) in supporting deductive coding, a major category of qualitative analysis where researchers use pre-determined codebooks to label the data into a fixed set of codes. Instead of training task-specific models, a pre-trained LLM could be used directly for various tasks without fine-tuning through prompt learning. Using a curiosity-driven questions coding task as a case study, we found, by combining GPT-3 with expert-drafted codebooks, our proposed approach achieved fair to substantial agreements with expert-coded results. We lay out challenges and opportunities in using LLMs to support qualitative coding and beyond. This work was published in [56]

and involved a collaboration with Z. Xiao, V. Liao, E. Yuan from Microsoft Research Montreal.

8.5 Curiosity-driven educational technologies

8.5.1 Machine Learning for Adaptive Personalization in Intelligent Tutoring Systems

Participants: Pierre-Yves Oudeyer (*correspondant*), Benjamin Clément, Didier Roy, H el ene Sauzeon.

The Kidlearn project 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 his 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 [107]. Systematic experiments were also presented with simulated students.

Kidlearn Experiments 2018-2019: Evaluating the impact of ZPDES and choice on learning efficiency and motivation An experiment was held between March 2018 and July 2019 in order to test the Kidlearn framework in classrooms in Bordeaux Metropole. 600 students from Bordeaux Metropole participated in the experiment. This study had several goals. The first goal was to evaluate the impact of the Kidlearn framework on motivation and learning compared to an Expert Sequence without machine learning. The second goal was to observe the impact of using learning progress to select exercise types within the ZPDES algorithm compared to a random policy. The third goal was to observe the impact of combining ZPDES with the ability to let children make different kinds of choices during the use of the ITS. The last goal was to use the psychological and contextual data measures to see if correlation can be observed between the students psychological state evolution, their profile, their motivation and their learning. We first show that LP-based personalization improves learning performance (reproducing and solidifying previous results) while producing a positive and motivating learning experience. We then show that the addition of self-choice as a playful feature triggers intrinsic motivation in the learner and reinforces the learning effectiveness of the LP-based personalizing. In doing so, it strengthens the links between intrinsic motivation and performance progress during the serious game. Conversely, deleterious effects of the playful feature are observed for hand-designed linear paths. Thus, the intrinsic motivation elicited by a playful feature is beneficial only if the curriculum personalization is effective for the learner. Such a result deserves great attention due to the increased use of playful features in non adaptive educational technologies available in the market. Details of these new results, as well as the overall results of this project, are presented in Benjamin Cl ement PhD thesis [106] and are currently being processed to be published.

Kidlearn and Adaptiv’Math The algorithms developed during the Kidlearn project and Benjamin Clement thesis [106] are being used in an innovation partnership for the development of a pedagogical assistant based on artificial intelligence intended for teachers and students of cycle 2. The algorithms are being written in typescript for the need of the project. The expertise of the team in creating the pedagogical graph and defining the graph parameters used for the algorithms is also a crucial part of the role of the team for the project. One of the main goal of the team here is to transfer technologies developed in the team in a project with the perspective of industrial scaling and see the impact and the feasibility of such scaling.

Kidlearn for numeracy skills with individuals with autism spectrum disorders Few digital interventions targeting numeracy skills have been evaluated with individuals with autism spectrum disorder (ASD) [151][150]. Yet, some children and adolescents with ASD have learning difficulties and/or a significant academic delay in mathematics. While ITS are successfully developed for typically developed students to personalize learning curriculum and then to foster the motivation-learning coupling, they are not or fewly proposed today to student with specific needs. The objective of this pilot study is to test the feasibility of a digital intervention using an STI with high school students with ASD and/or intellectual disability. This application (KidLearn) provides calculation training through currency exchange activities, with a dynamic exercise sequence selection algorithm (ZPDES). 24 students with ASD and/or DI enrolled in specialized classrooms were recruited and divided into two groups: 14 students used the KidLearn application, and 10 students received a control application. Pre-post evaluations show that students using KidLearn improved their calculation performance, and had a higher level of motivation at the end of the intervention than the control group. These results encourage the use of an STI with students with specific needs to teach numeracy skills, but need to be replicated on a larger scale. Suggestions for adjusting the interface and teaching method are suggested to improve the impact of the application on students with autism. [148].

8.5.2 Machine learning for adaptive cognitive training

Participants: Pierre-Yves Oudeyer, H el ene Sauz eon (*correspondant*), Masataka Sawayama, Benjamin Cl ement, Maxime Adolphe, Marion Pech.

Because of its cross-cutting nature to all cognitive activities such as learning tasks, attention is a hallmark of good cognitive health throughout life and more particularly in the current context of societal crisis of attention. Recent works have shown the great potential of computerized attention training for an example of attention training, with efficient training transfers to other cognitive activities, and this, over a wide spectrum of individuals (children, elderly, individuals with cognitive pathology such as Attention Deficit and Hyperactivity Disorders). Despite this promising result, a major hurdle is challenging: the high inter-individual variability in responding to such interventions. Some individuals are good responders (significant improvement) to the intervention, others respond variably, and finally some respond poorly, not at all, or occasionally. A central limitation of computerized attention training systems is that the training sequences operate in a linear, non-personalized manner: difficulty increases in the same way and along the same dimensions for all subjects. However, different subjects require in principle a progression at a different, personalized pace according to the different dimensions that characterize attentional training exercises.

To tackle the issue of inter-individual variability, the present project proposes to apply some principles from intelligent tutoring systems (ITS) to the field of attention training. In this context, we have already developed automatic curriculum learning algorithms such as those developed in the KidLearn project, which allow to customize the learner’s path according to his/her progress and thus optimize his/her learning trajectory while stimulating his/her motivation by the progress made. ITS are widely identified in intervention research as a successful way to address the challenge of personalization, but no studies to date have actually been conducted for attention training. Thus, whether ITS, and in particular personalization algorithms, can optimize the number of respondents to an attention training program remains an open question.

Grounded state-of-the-art To investigate this question, we first conducted a systematic review aiming at exploring existing methods in computerized CT and analyzing their outcomes in terms of learning mechanics (intra-training performance) and effectiveness (near, far and everyday life transfer effects of CT) [72]. A search up to June 2023 with multiple databases selecting 19 computerized CT studies revealed that only two studies emphasized the favorable influence of individualization on CT effectiveness, while five underscored its capacity to enhance the training experience by boosting motivation, engagement, and offering diverse learning pathways. In sum, despite promising results in this new research avenue, more research is needed to fully understand and empirically support individualized techniques in cognitive training. Complementing the study of adaptive methods applied to cognitive training, we have attempted through a review of the subjective literature to gain a better understanding of the Multiple Object Tracking (MOT) task, which seems to have the best results in terms of attentional training efficiency in young and older adults. The results of this work highlight that: (1) Multiple cognitive mechanisms are identified as active in the task (divided and sustained attention; foveal and peripheral attention; automatic and controlled inhibition, etc.); (2) a limited number of studies have actually implemented the MOT task in computer-assisted cognitive training; and (3) It's the near (attention tasks) and far (other cognitive tasks) effects that are well documented as positive outcomes of MOT-based training while there is a scarcity of research that has thoroughly analyzed the ecological effects of attentional training, namely the potential transfer effects in everyday life (paper in progress).

ZPDES calibration for MOT training (Young participants) In parallel to this, a web platform has been designed for planning and implementing remote behavioural studies. This tool provides means for registering recruited participants remotely and executing complete experimental protocols: from presenting instructions and obtaining informed consents, to administering behavioural tasks and questionnaires, potentially throughout multiple sessions spanning days or weeks. In addition to this platform, a cognitive test battery composed of seven classical behavioural tasks has been developed. This battery aims to evaluate the evolution of the cognitive performance of participants before and after training. Fully open-source, it mainly targets attention and memory. A preliminary study on a large sample of 50 healthy participants showed that the developed tasks reproduced the results of previous studies, that there were large differences between individuals (no ceiling effect) and that the results were significantly reliable between two measurements taken on two days separated by one night [2].

Randomized and controlled Trial in Young and Olders adults : Predifined vs. ZPDES condition Utilizing these tools, a pilot study campaign was conducted to evaluate the impact of our AI-based personalized cognitive training program. The first pilot experiment involved n=27 participants and aimed to compare the effectiveness of a cognitive training program using a linear difficulty management procedure (staircase procedure) to a program using an ITS for difficulty manipulation. The online training lasted for 10 hours over a period of 2 weeks. The results indicated that the ITS-based intervention produced diverse learning trajectories compared to the linear procedure 35, leading to broader improvements in pre-post cognitive assessment. However, no significant differences were observed in subjective measures of motivation and engagement between the two groups. Subsequent to this initial experiment, two pilot studies (n=11 and n=10, respectively) were conducted with the goal of enhancing motivation and engagement in the game. The first study implemented gamified components such as scores and feedback, while the second study examined hyperparameter updates to the ITS. The analysis of learning trajectories, learning outcomes, and subjective measures yielded promising results in favor of the AI-based personalized procedure.

Building on the preliminary findings, we expanded our research scope with a more comprehensive experimental setup involving two distinct studies. The first study encompassed 64 young adults, sourced through the Prolific platform, while the second study consisted of 49 older adults, recruited from the "Université du temps libre". Our experimental methodology mirrored that of our initial pilot studies, with a notable enhancement: the integration of new gamified elements (including mini-story creation and new visual content) aimed at boosting participant motivation and engagement.

The data analysis encompassed three primary dimensions: initially, an exploratory phase to delineate learning trajectories between control and intervention groups; subsequently, a comparative analysis of pre- and post-test performance on the cognitive battery; and lastly, an examination of participants'

self-reported experiences during training, providing insights into their subjective perceptions of the experiment.

The pilot studies' preliminary outcomes were corroborated in these larger sample groups. Notably, learning trajectories exhibited greater diversity in the group undergoing the intervention procedure. This group also demonstrated a more pronounced improvement across a wider range of cognitive assessment tasks. Although participants engaging in the personalized cognitive training reported a higher cognitive load via questionnaires, the levels of engagement and frustration did not significantly differ between the two groups.

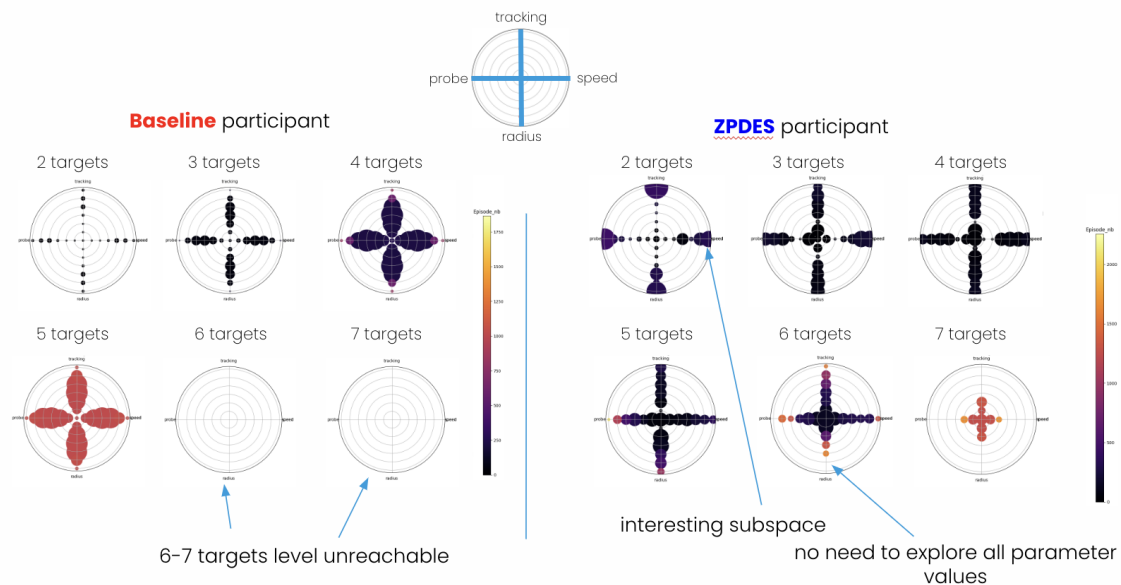


Figure 35: Different learning trajectories for a selected participant in the staircase group (left) and the ITS group (right). The color of a dot indicates the initial presentation of the parameter value, while the size of the dot represents the frequency of the parameter value.

Qualitative Analysis with LLMs: As it is well known that there are more dropouts in older adults compared to young ones, we aimed to better understand the learning experience of trainees with feedback analyses. For this, we designed a new way through several Large Language Models (LLM) enabling to extract hot topics or main dropout's motivations in verbatim that are related to pragmatic, hedonist and/or aesthetic dimensions of cognitive training. The results analyzed through various LLM are encouraging (paper in progress). To support this new approach, we are exploring different prompts on other data corpora in order to ultimately propose a tutorial accessible to anyone wishing to carry out a LLM-based thematic qualitative analysis.

8.5.3 ToGather : Interactive website to foster collaboration among stakeholders of school inclusion for pupils with neurodevelopmental disorders

Participants: H  l  ne Sauz  on (*correspondant*), C  cile Mazon, Eric Meyer, Isabeau Saint-Supery, Christelle Maillart (*Uni. Li  ge, Belgium*), Kam  lia Belassel, Mathieu P  ri  , Valentin Strahm.

Sustain and support the follow-up of the school inclusion of children with neurodevelopmental disorders (e.g., autism, attention disorders, intellectual deficiencies) has become an emergency : the higher is the school level, the lower is the amount of schooled pupils with cognitive disabilities.

Technology-based interventions to improve school inclusion of children with neurodevelopmental disorders have mostly been individual centered, focusing on their socio-adaptive, and cognitive

impairments and implying they have to adapt themselves in order to fit in our society's expectations. Although this approach centered on the normalization of the person has some advantages (reduction of clinical, symptoms), it carries social stereotypes and misconceptions of cognitive disability that are not respectful of the cognitive diversity and intrinsic motivations of the person, and in particular of the student's wishes in terms of school curriculum to achieve his or her future life project [51].

The "ToGather" project aims at enlightening the field of educational technologies for special education by proposing an approach centered on the educational needs of the students and bringing a concerted and informed answer between all the stakeholders including the student and all their support spheres (family, school, medico-social care). To this end, ToGather project that emanates from participatory design methods, primarily consists of having developed a pragmatic tool (interactive website) to help students with cognitive disability and their caregivers to formalize and to visualize the repertoire of academic skills of the student and to make it evolve according to his or her proximal zone of development (in the sense of Vygotsky) on the one hand, and to the intrinsic motivations of the student (his or her own educational and life project) on the other [149].

This project is in partnership with the School Academy of Bordeaux of the French Education Ministry, the ARI association, the Centre of Autism of Aquitaine. It is funded by the FIRAH (foundation) and the Nouvelle-Aquitaine Region (see the dedicated webpages : <https://flowers.inria.fr/projet-tous-ensemble/>).

First, usability studies have been conducted for evaluating ergonomic qualities of the ToGather website, yielding positive results in French and Belgian contexts. Then, we conducted a large field-study to assess the effectiveness of the tool in helping stakeholders to support children with neurodevelopmental disorders (NDD) [176] [83] [55].

The study protocol consisted in a longitudinal non-randomized controlled trial, with baseline, 3-months, and 6-months follow-up assessments. The recruitment was conducted across the entire French territory. Our local partners facilitated the dissemination of the call for participation in Gironde and provided us with contacts to extend it to other regions. Additionally, a recruitment campaign through social media was carried out to communicate about the study and encourage participants to test the ToGather tool.

As the tool was designed to support co-educational process between parents and professionals, a support team had to consist of at least two stakeholders, including at least one of the parents. Initially, 157 participants were recruited in 37 support teams, but 30 individuals did not answer to baseline questionnaire, leading to the exclusion of 11 support teams. After baseline assessment, 13 support teams were allocated to the experimental condition (ToGather app) and 11 to the control condition (usual follow-up).

Primary outcomes measures covered stakeholders' relationships, self-efficacy, and attitudes towards inclusive education, while secondary outcomes measures were related to stakeholders' burden and quality of life, as well as children's school well-being and quality of life.

As the study ended in July 2023, data analysis is still ongoing. Preliminary results after 3 months of use showed encouraging results with an improvement in communication between stakeholders and their respective quality of life (paper in progress)

8.5.4 Curious and therefore not overloaded : Study of the links between curiosity and cognitive load in learning mediated by immersive technologies

Participants: H el ene Sauz eon (*correspondant*), Matisse Poupard,
Andr e Tricot (*Cosupervisor - Univ. Montpellier*),
Florian Larrue (*Industrialist - Le Catie*).

With the ever-increasing interest in digital technologies in education, many questions about their use and effectiveness are emerging. In this context, this project focus on the relationships between three key dimensions of technology-mediated learning: the learner's internal learning processes, the instructional design, and the educational technology used.

In partnership with CATIE (industrial partner) and the EPSYLON laboratory of the University of Montpellier (PR. Andr e Tricot), two main objectives are targeted in this research program started in April 2022:

- To establish connections between the theory of cognitive load and models of curiosity-driven learning.
- To experimentally assess the influence of the choice of educational technology on the associations between pedagogical choices (guided instruction vs. exploration) and learner expertise.

To this end, the program includes 3 main phases of study:

State of the art A systematic review evaluating the contributions and limitations of Virtual Reality (VR) and Augmented Reality (AR) in learning, with a specific focus on examining their impacts on cognitive load and intrinsic motivations, has been completed and is currently in the submission process [173].

The main results are as follows: From a pool of 3250 results, 36 studies with a robust study design investigating the impact of virtual or augmented reality on learning performance and cognitive load, or intrinsic motivation were incorporated. Main results of studies were reported in a grid that we built to determine whether the observed effects were positive, neutral, negative, or inconsistent with established theoretical frameworks. Results of the review indicate that AR effectively optimized cognitive load, leading to enhanced learning outcomes, while VR, on the other hand, tended to overload learners, decreasing learning performance. Regarding intrinsic motivation, results were incoherent with motivational models, likely due to variations in measurement methods. Notably, only a few studies simultaneously investigated cognitive load and intrinsic motivation as integral components of learning efficiency, and they reported conflicting causal relationships between these variables.

Experimentation Based on these results, two experiments involving 140 second-year undergraduate students in the field of medicine were conducted. In the first experiment, we use spatial augmented reality and mixed reality (HoloLens 2) to investigate whether guiding students' drawings during their lectures can reduce their cognitive load and enhance their motivation to learn.

Our hypotheses are as follows:

- We plan to demonstrate that drawing guidance will reduce learners' extrinsic cognitive load, thereby promoting learning.
- We assume that reducing cognitive load will stimulate learners' motivation and engagement, considering cognitive load as a motivational cost.
- However, we assume that the presence of a 3D model (RM + 3D model) will add an additional extrinsic cognitive load, potentially leading to overload and reduced learning.
- We expect to observe a positive correlation between drawing quality and learning.
- We anticipate better acceptability and lower cognitive load for the projector compared with HoloLens (due to the weight of the headset and convergence and interaction issues).
- In addition, this experiment can also be used to understand the effect of learners' prior knowledge on their performance.

In the second experiment, we change the learning paradigm, using virtual reality with different levels of interaction and guidance to examine how exploration and embodied interaction with a 3D model can have a positive impact on learning, cognitive load, and curiosity.

We make the following assumptions:

- Interaction with the system, i.e. the manipulation of information, promotes better learning and stimulates intrinsic motivation. Thus, the passive VR condition should lead to poorer results than the other conditions in terms of learning and motivation.
- Embodied manipulation, in VR, presents an advantage for learning and learner engagement. Manipulating information through the senses and one's own body makes for a more immersive and engaging learning experience.

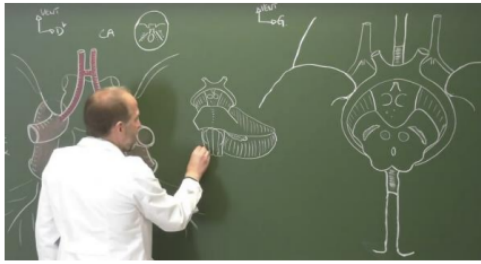

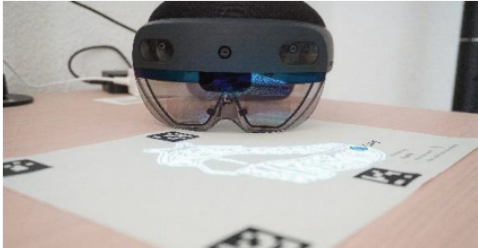

 <p>Control condition: video of the teacher giving the classic lesson</p>	<p>Spatial AR condition: in addition to the video system used in the control condition, a projector placed above the screen will project a 2D image of a drawing, equivalent to the one made by the teacher, which serves as a layer, so that the learner only must draw over the projected layer.</p> 
 <p>Mixed reality condition: like the spatial AR condition, but the projector is replaced by a HoloLens 2 headset.</p>	 <p>RM + 3D model condition: a 3D model of the drawn structure is added to the HoloLens guidance device.</p>

Figure 36: Experimental conditions for experiment 1 : Augmented reality-assisted drawing note-taking

	<p> Passive VR: the user cannot act on the system.</p> <p> Active guided VR: the user can act on the system but is not free to do so.</p> <p> Free active VR: the user chooses how to explore the environment.</p>
---	--

Figure 37: Experimental conditions for experiment 2 : Embodied learning in virtual reality, effect of interactivity

- Guided, motor-only interaction is less interesting in terms of learning and motivation than free interaction, which we can describe as cognitive.
- Free exploration encourages inquisitive behavior, but can overload learners, especially if they are new to the subject.
- Intrinsic motivation minimizes perceived cognitive effort.

Transfer We hope to extend the results obtained to the industrial context in which CATIE's activities are carried out. CATIE's mission is to accelerate technology transfer between the worlds of research and industry. The Human Centered Systems team, in which this research project is part of, supports companies in improving the design of existing or new digital systems, by proposing a human-centered approach. The different questions raised by this project are intended to help CATIE to answer these issues, to improve its know-how in terms of learning and digital systems, and then to transfer this knowledge to EdTech companies.

8.6 Curiosity-driven AI for assisted scientific discovery

8.6.1 Design of an Interactive Software for Automated Discovery in Complex Systems

Participants: Clément Romac (*correspondant*), Jesse Lin, Mathieu Périé, Mayalen Etcheverry, Clément Moulin-Frier, Pierre-Yves Oudeyer.

We further developed our Automated Discovery software and started experimenting with it. First, we released a new version of our standalone Python library. We improved how experiments could be saved and reloaded.

Second, we focused on implementing tools and interfaces for users to give feedback or instructions to the automated discovery algorithm that explores the complex system. As identified by [16], empowering experiments to collaborate with automated discovery methods can be key to obtain interesting discoveries. Integrating such a collaborative process in our tool came with several engineering challenges and we are currently experimenting with our solution and working on making it user-friendly for non-experts end users.

Finally, we released a first open version of [our software](#). We provide a documentation and installation tools using Docker.

8.6.2 Learning Sensorimotor Agency in Cellular Automata

Participants: Gautier Hamon (*correspondant*), Mayalen Etcheverry, Bert Chan, Clément Moulin-Frier, Pierre-Yves Oudeyer.

As a continuation of the previous projects in Automated Discovery in Self-Organizing Systems, we have been working on expanding the set of discoveries of possible structures in continuous CAs such as Lenia [103, 102], and in particular we have been interested to search for emerging *agents* with *sensorimotor* capabilities. Understanding what has led to the emergence of life and sensorimotor agency as we observe in living organisms is a fundamental question. In our work, we initially only assume environments made of low-level elements of matter (called atoms, molecules or cells) locally interacting via physics-like rules. There is no predefined notion of agent embodiment and yet we aim to answer the following scientific question: *is it possible to find environments in which there exists/emerge a subpart that could be called a 'sensorimotor agent'?*

We use Lenia continuous cellular automaton as our artificial "world" [102]. We introduce a novel method based on gradient descent and curriculum learning combined within an intrinsically-motivated goal exploration process (IMGEP) to automatically search parameters of the CA rule that can self-organize

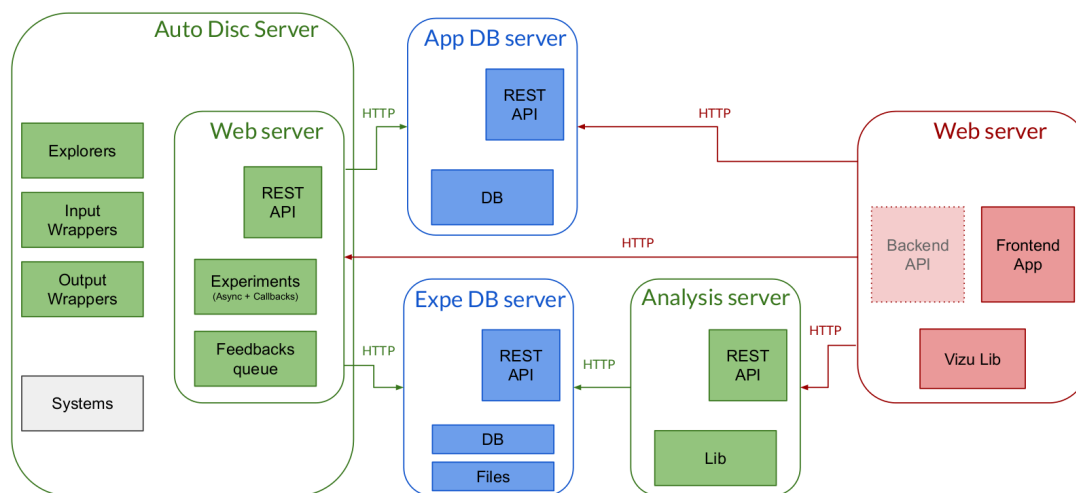


Figure 38: Technical architecture of our software.

spatially localized² and moving patterns³ within Lenia. The IMGEP defines an outer exploratory loop (generation of training goal/loss) and an inner optimization loop (goal-conditioned). We use a population-based version of IMGEP [17],[112] but introduce two novel elements compared to previous papers in the IMGEP literature. First, whereas previous work in [29] and [16] used a very basic nearest-neighbor goal-achievement strategy, our work relies on gradient descent for the local optimization of the (sensitive) parameters of the complex system, which has shown to be very powerful. To do so we made a differentiable version of the Lenia framework, which is also a contribution of this work. Secondly, we propose to control subparts of the environmental dynamics with functional constraints (through predefined channels and kernels in Lenia) to build a curriculum of tasks; and to integrate this stochasticity in the inner optimization loop. This has shown central to train the system to emerge sensorimotor agents that are robust to stochastic perturbations in the environment. In particular, we focus on modeling obstacles in the environment physics and propose to probe the agent sensorimotor capability as its performance to move forward under a variety of obstacle configurations. We also provide in this work tests and metrics to measure the robustness of the obtained agents.

While many complex behaviors have already been observed in Lenia, among which some could qualify as sensorimotor behaviors, they have so far been discovered "by chance" as the result of time-consuming manual search or with simple evolutionary algorithms. Our method provides a more systematic way to automatically learn the CA rules leading to the emergence of basic sensorimotor structures, as shown in Figure 41. Moreover, we investigated and provided ways to measure the (zero-shot) generalization of the discovered sensorimotor agents to several out-of-distribution perturbations that were not encountered during training. Impressively, even though the agents still fail to preserve their integrity in certain configurations, they show very strong robustness to most of the tested variations. The agents are able to navigate in unseen and harder environmental configurations while self-maintaining their individuality (Figure 39). Not only the agents are able to recover their individuality when subjected to external perturbations but also when subjected to internal perturbations: they resist variations of the morphogenetic processes such that less frequent cell updates, quite drastic changes of scales as well as changes of initialization (Figure 40). Furthermore, when tested in a multi-entity initialization and despite having been trained alone, not only the agents are able to preserve their individuality but they show forms of coordinated interactions (attractiveness and reproduction). Our

²**Spatially-localized pattern** a pattern existing within some (fuzzy) boundary i.e. with a limited range in space as opposed to patterns with unbounded growth

³**Moving patterns** a spatially-localized pattern that move and propagate information in space

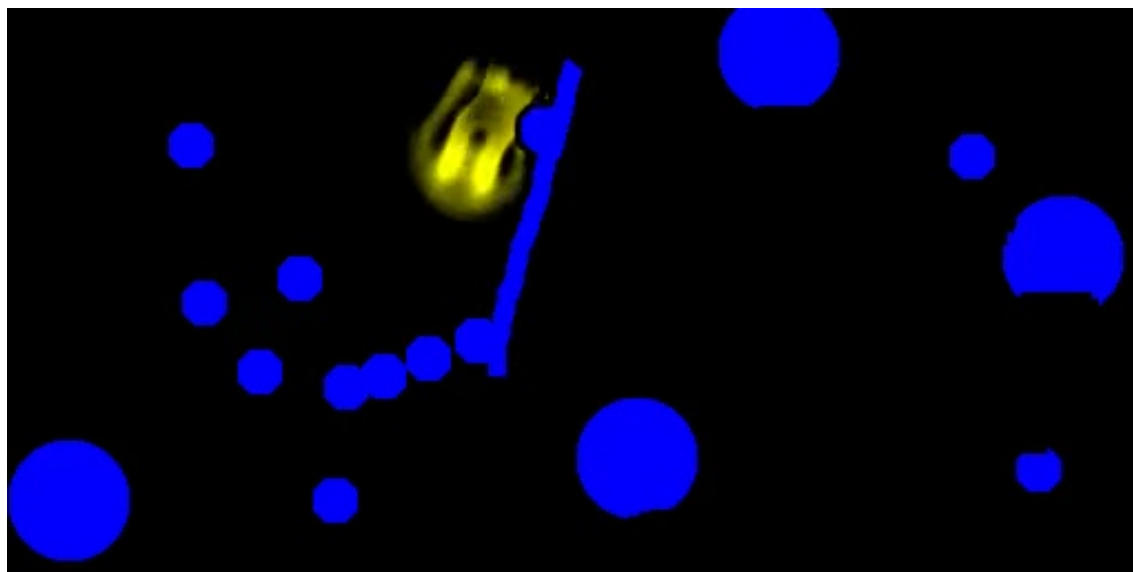


Figure 39: Robustness test to harder/unseen obstacle configurations: straight wall, bigger obstacle, dead ends.

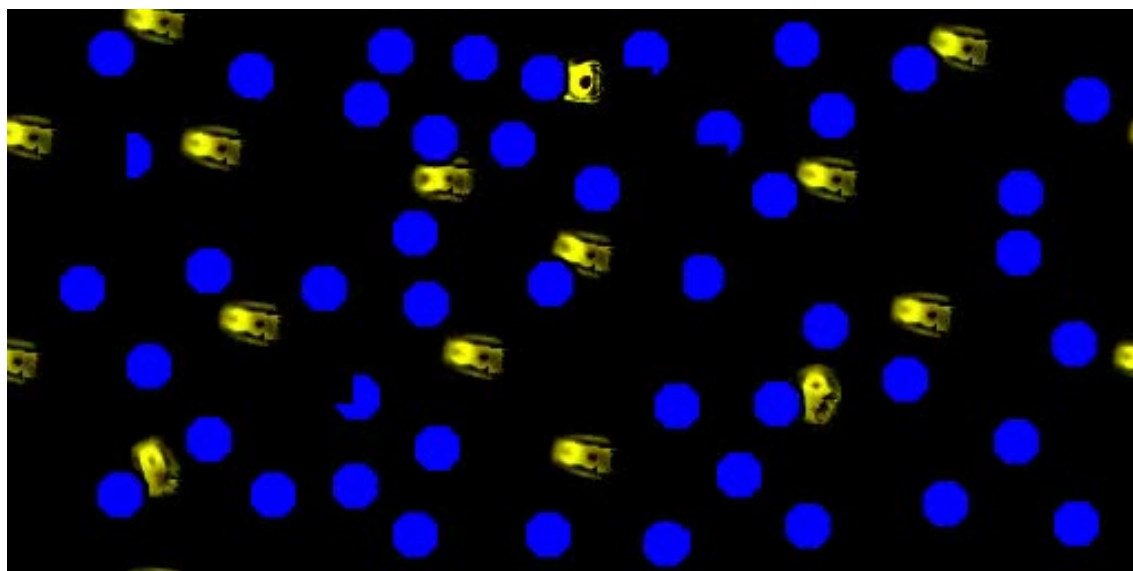


Figure 40: Change of scale changing the kernel size and initialization, the grid is the same size in both

results suggest that, contrary to the (still predominant) mechanistic view on embodiment, biologically-inspired embodiment could pave the way toward agents with strong coherence and generalization to out-of-distribution changes, mimicking the remarkable robustness of living systems to maintain specific functions despite environmental and body perturbations [137]. Searching for rules at the cell-level in order to give rise to higher-level cognitive processes at the level of the organism and at the level of the group of organisms opens many exciting opportunities to the development of embodied approaches in AI in general.

The work has been released in 2022 as a distill-like article which is currently hosted [at this link](#). This article contains an interactive demo in WebGL and javascript, as well as many videos and animations of the results. A colab notebook with the source code of the work is publicly available [at](#).

In 2023, additional quantitative experiments were conducted as well as ablations. And this work was submitted to the Proceedings of the National Academy of Sciences (PNAS) journal.

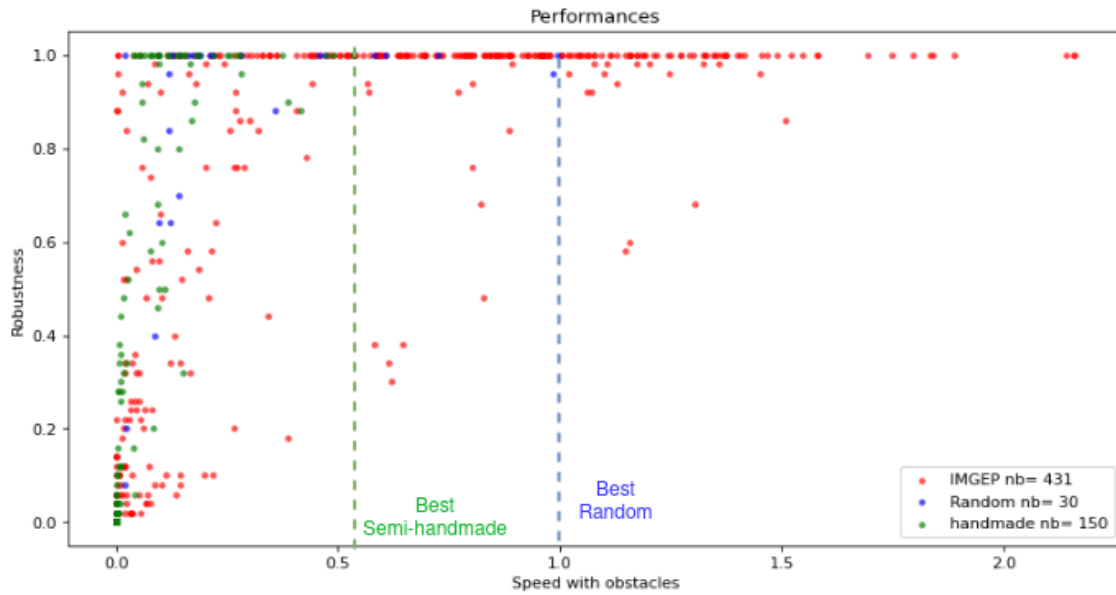


Figure 41: Scatter plot of the agents as their measured performances of robustness to obstacles (y axis) and speed in obstacles (x axis) obtained by IMGEP (red), random search with the same compute resources as IMGEP (blue) and the one from the original lenia paper (green)

8.6.3 Flow lenia: Mass conservation for the study of virtual creatures in continuous cellular automata

Participants: Erwan Plantec, Gautier Hamon (*correspondant*), Mayalen Etcheverry, Pierre-Yves Oudeyer, Clément Moulin-Frier, Bert Chan.

Following our work on trying to find sensorimotor capabilities in cellular automata such as Lenia [103, 102], we kept exploring the search for low level cognition in continuous cellular automata. This led to preliminary search on trying to emerge memory in self-organizing agents as well as work on trying to implement other environmental constraints in the CA in order to emerge interesting behavior. To implement more easily those environmental constraints as well as to ease the emergence of spatially localized patterns (and thus have the optimization/search to focus more on the cognitive ability, removing the need to optimize to prevent uncontrollable growth/explosion of the pattern), we worked on adding mass conservation to the Lenia system.

We propose in this work a mass-conservative (i.e the sum of the CA's activations remains constant over time) extension to Lenia called Flow Lenia [54]. We hypothesize that such conservation laws will help in the search for artificial life-forms by constraining emerging patterns to spatially localized ones. It also allows to implement more easily environmental constraints on the self-organizing agents such as a need for food to grow, etc.

Furthermore, we show that this new model allows for the integration of the update rule parameters within the CA dynamics enabling the emergence of creatures with different parameters and so different properties in the same environment/grid. This leads to multi-species simulation where the grid is filled with agents with different behaviors and properties [42]. Such a feature opens up research perspectives towards the achievement of open-ended intrinsic evolution inside continuous CAs, which means that all the evolutionary part would be a result of the dynamic of the CA (without any external loop/system). We hypothesize that this open-ended intrinsic evolution could, through the competition/cooperation, lead to the emergence of interesting low level cognition in those system.

Simple evolutionary strategy (with an evolutionary loop outside the system) was also used to optimized for pattern with directional and rotational movement.

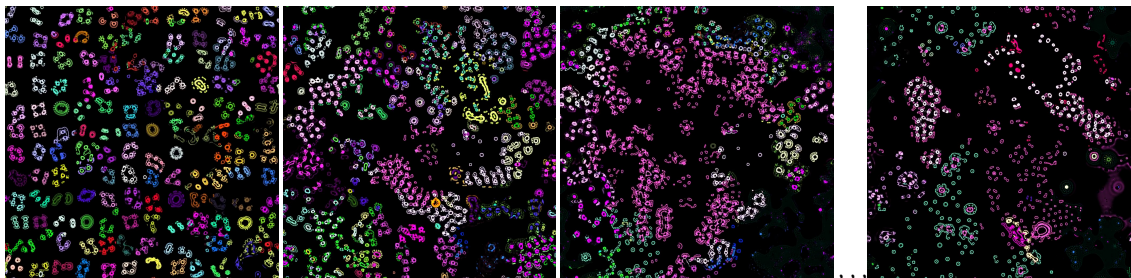


Figure 42: Multi-species simulation in Flow Lenia where each colour represents different values of parameters. Left to right shows the evolution of the system over time with some species stealing the mass of others.

You can find some examples of the system and pattern in this companion website, including the ones trained for movement, random parameters, food in flow Lenia, and multi species simulations: [see. Notebook with the system can be found : here.](#)

This work led to an oral presentation to the WIVACE 2022, 15th International Workshop on Artificial Life and Evolutionary Computation.

In 2023, final quantitative experiments on optimizing the parameters with evolutionary strategies and writing was conducted, as well as some additional exploratory experiments on large simulations for open ended evolution.

This work got the best paper award at the ALIFE 2023 conference where it got an oral presentation.

Collaboration with Bert Chan In the context of the project Automated Discovery in Self-Organizing Systems, we have an ongoing collaboration with Bert Chan, a previously independant researcher on Artificial Life and author of the Lenia system [103, 102] and who is now working as a research engineer at Google Brain. During this collaboration, Bert Chan help us design versions of IMGEP usable by scientists (non ML-experts) end-users, which is the aim of project 8.6.1. Having himself created the Lenia system, he is highly-interested to use our algorithms to automatically explore the space of possible emerging structures and provides us valuable insights into end-user habits and concerns. Bert Chan also co-supervised with Mayalen Etcheverry the master internship of Gautier Hamon which led to the work described in section 8.6.2. He also co-supervised with Gautier Hamon and Mayalen Etcheverry the master internship of Erwan plantec which led to the work described in section 8.6.3.

8.6.4 Exploration of Gene Regulatory Network Behaviors Using Automated Discovery Tools

Participants: Mayalen Etcheverry (*correspondent*), Clément Moulin-Frier, Pierre-Yves Oudeyer, Michael Levin.

In the context of project "Automated Discovery in Self-Organizing Systems", it has been demonstrated that modern tools leveraging computational models of curiosity developed in the Flowers team can be transposed to form efficient AI-driven "discovery assistants." These tools can assist scientists in mapping and navigating the space of possible outcomes in complex systems [29, 16], [126]. In 2022, we initiated a collaboration with Dr. Michael Levin, a renowned biologist at Tufts University, through a 5-month academic exchange with Mayalen Etcheverry in his lab in Boston. This collaboration laid the foundation for continued collaboration throughout 2023, resulting in the submission of one paper [73] (currently under review) and another accepted at the NeurIPS 2023 *AI for Science* workshop [61].

The primary focus of this collaboration was to leverage curiosity-driven exploration algorithms as tools to empower scientific exploration and analysis of basal cognition in biological systems, specifically numerical models of gene regulatory networks (GRNs). Understanding, mapping, predicting, and controlling the complex behavior of these networks is crucial for applications in biomedicine and synthetic bioengineering. However, there are few quantitative tools that facilitate exploration of these networks, especially when their complexity makes unguided exploration infeasible.

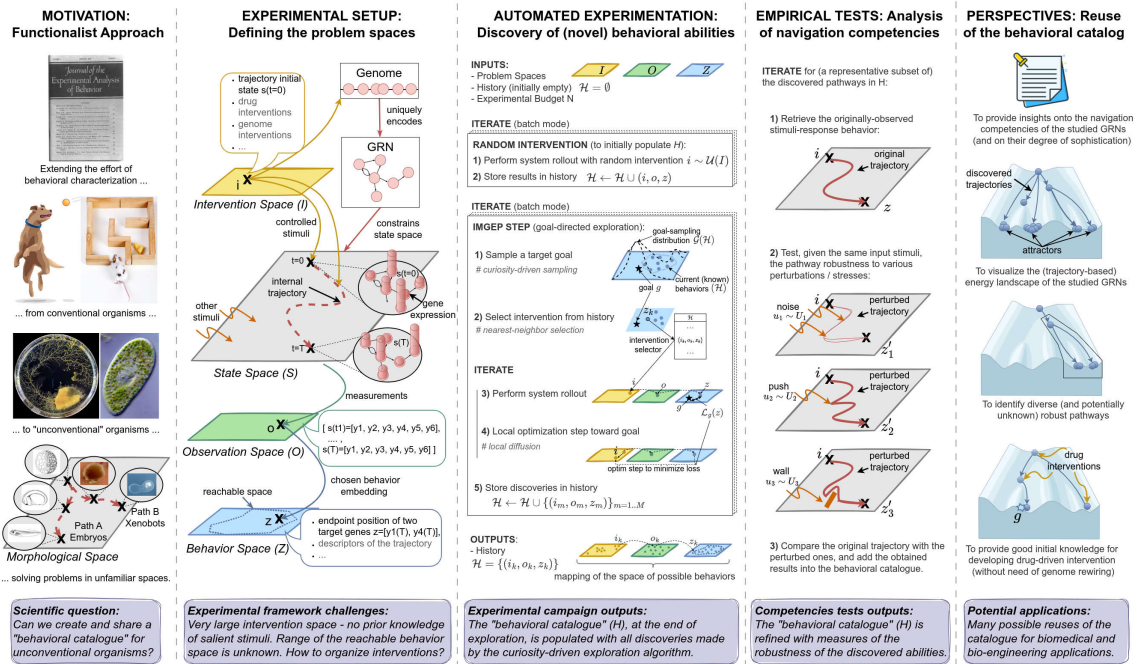


Figure 43: Overview of the proposed framework in [73].

To address these challenges in practice, we proposed an experimental framework summarized in Figure 43. In this framework, we formalized and investigated a view of gene regulatory networks as agents navigating a problem space. We developed automated tools to efficiently map the repertoire of robust goal states that GRNs can reach despite perturbations. These tools rely on two main contributions that we made in this work: (1) The use of curiosity-driven exploration algorithms, originating from the AI community, to explore the range of behavioral abilities of a given system, which we adapted and leveraged to automatically discover the range of reachable goal states of GRNs, and (2) The use of a battery of empirical tests inspired by implementation-agnostic behaviorist approaches that we leveraged to assess the navigation competencies of GRNs.

Our data revealed that models inferred from real biological data can reach a surprisingly wide spectrum of steady states, showcasing various competencies that living agents often exhibit in physiological network dynamics and that do not require structural changes to network properties or connectivity. Furthermore, we investigated the applicability of the discovered "behavioral catalogs" for comparing the evolved competencies across classes of evolved biological networks, as well as for the design of drug interventions in biomedical contexts or for the design of synthetic gene networks in bioengineering. Altogether, these automated tools and the resulting emphasis on behavior-shaping and exploitation of innate competencies can open the path to better interrogation platforms for exploring the complex behavior of biological networks in an efficient and cost-effective manner.

To encourage broader adoption and development of the tools and algorithms, we have released two software packages: SBMLtoODEJax (<https://github.com/flowersteam/sbmltoodejax>) 7.1.21 and AutoDiscJax (<https://github.com/flowersteam/autodiscjax>). SBMLtoODEJax converts Systems Biology Markup Language (SBML) models into Python classes written in JAX, enabling easy simulation and manipulation. AutoDiscJax, built upon JAX and SBMLtoODEJax, facilitates automated discovery and exploration of complex systems, specifically organizing the exploration of computational models of biological GRNs.

9 Bilateral contracts and grants with industry

9.1 Bilateral contracts with industry

Research on lifelong Deep Reinforcement Learning of multiple tasks (Microsoft)

Participants: Pierre-Yves Oudeyer (*correspondant*), Laetitia Teodorescu.

Financing of the PhD grant of Laetitia Teodorescu.

Automated Discovery of Self-Organized Structures (Poietis)

Participants: Pierre-Yves Oudeyer (*correspondant*), Mayalen Etcheverry.

Financing of the CIFRE PhD grant of Mayalen Etcheverry by Poietis.

Machine learning for adaptive cognitive training (OnePoint)

Participants: H el ene Sauz eon (*correspondant*), Pierre-Yves Oudeyer, Maxime Adolphe.

Financing of the CIFRE PhD grant of Maxime Adolphe by Onepoint.

Curiosity-driven interaction system for learning (evidenceB)

Participants: H el ene Sauz eon (*correspondant*), Pierre-Yves Oudeyer, Rania Abdelghani.

Financing of the CIFRE PhD grant of Rania Abdelghani by EvidenceB.

Curious and therefore not overloaded : Study of the links between curiosity and cognitive load in learning mediated by immersive technologies (CATIE)

Participants: H el ene Sauz eon (*correspondant*), Matisse Poupard, Andr e Tricot (*Cosupervisor - Univ. Montpellier*), Florian Larrue (*Industrialist - Le Catie*).

Financing of a PhD grant of Matisse Poupard with CATIE and EPSYLON Lab (Univ. Montpellier).

Augmenting curiosity-driven exploration with very large language models in deep reinforcement learning agents (Hugging Face)

Participants: Pierre-Yves Oudeyer (*correspondant*), Cl ement Romac.

Financing of the PhD grant of Cl ement Romac by Hugging Face.

Autonomous Driving Commuter Car (Renault)

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...).

9.2 Bilateral contracts with industry

We received a 30keuros grant from Google Brain, as well as 30keuros Google cloud credit, for developing projects on automated exploration of continuous cellular automata.

9.3 Bilateral Grants with Foundation

School+ /ToGather project (FIRAH and Region Nouvelle-Aquitaine)

Participants: Hélène Sauzéon (*correspondant*), Cécile Mazon, Isabeau Saint-supery, Eric Meyer.

Financing of one year-postdoctoral position and the app. development by the International Foundation for Applied Research on Disability (FIRAH). The School+ project consists of a set of educational technologies to promote inclusion for children with Autism Spectrum Disorder (ASD). School+ primary aims at encouraging the acquisition of socio-adaptive behaviours at school while promoting self-determination (intrinsic motivation), and has been created according to the methods of the User-Centered Design (UCD). Requested by the stakeholders (child, parent, teachers, and clinicians) of school inclusion, Flowers team works to the adding of an interactive tool for a collaborative and shared monitoring of school inclusion of each child with ASD. This new app will be assessed in terms of user experience (usability and elicited intrinsic motivation), self-efficacy of each stakeholder and educational benefit for child. This project includes the Academie de Bordeaux –Nouvelle Aquitaine, the CRA (Health Center for ASD in Aquitania), and the ARI association.

CLEMENCE Cohort (Fondation de France and Théa Pharma)

Participants: Hélène Sauzéon (*correspondant*), Cécile Mazon, Cécile Delcourt.

The project "Cohorte Longitudinale sur la Myopie et le développement oculaire dans l'ENfanCE(CLEMENCE) is led by C. Delcourt from the lab of Bordeaux Populational Health (2M€). Hélène Sauzéon and Cécile Mazon participate to the research program with the study of developmental changes due to Myopia in visual attention.

10 Partnerships and cooperations

10.1 International initiatives

Clément Moulin-Frier is continuing an active scientific and teaching collaboration with Ricard Solé and Marti Sanchez-Fibla from the University Pompeu-Fabra (UPF) in Barcelona, Spain. The main highlights from this collaboration for 2023 are:

- The obtention in 2023 of a UBGRS-Mob grant from the Collège des Ecoles Doctoral de l'Université Bordeaux, funding a 3-month visit of Gautier Hamon (PhD student supervised by Clément) in the Complex System Lab at UPF, headed by Ricard Solé. Gautier will work on a project on "The emergence of agriculture in artificial agent populations" and the visit will be from January to March 2024.
- The publication of a journal article on "The Morphospace of Consciousness: Three Kinds of Complexity for Minds and Machines" in *NeuroSci* [34].
- The soumission of a paper on "Cooperative control of environmental extremes by artificial intelligent agents" to the PNAS journal (a preprint is online [78])
- The co-responsibility of the course "System Design, Integration and Control" in the CSIM Master at UPF.

- An invited talk of Clément Moulin-Frier at the Symposium **Intelligence: natural, artificial and synthetic** organized by Ricard Solé (see Invited Talks section).
- Two research visits of Clément Moulin-Frier at UPF in January and October 2023.

10.1.1 Inria associate team not involved in an ILL or an international program

We created a new three years- Inria associate team since april 2023, namely CuriousTECH team (see the website of CuriousTECH team : <https://flowers.inria.fr/curioustech-associate-team/>). Its entails two team of Inria (Flowers and Potioc) and two labs of the Waterloo University (HCI Lab, David R. Cheriton School of Computer Science and the Cognitive neuroscience lab. of Pyschology Department). This prosed associate team aims to develop an original, cross-disciplinary approach, joining together two perspectives:

- 1) The fundamental study of curiosity-driven learning across life-span (in children, young adults and older adults) and
- 2) The study of how new (re)educational technologies, using both curiosity-related models and artificial intelligence techniques [3, 8, 9], can personalize learning sequences for each individual, maximizing curiosity and learning efficiency in real world contexts.

Our proposed research will produce new understanding of the role of curiosity in education and healthy aging, through the design and the field assessment of new interactive educational technologies or health-related technologies. Beyond academic contributions, we expect our findings to inform the broader societal challenges inherent to the School of the 21st Century, ranging from helping children (and their teachers) to develop cross-domain skills for learning such as curiosity and meta-cognition, while improving inclusivity in schools (learners with disabilities, especially cognitive disabilities) as well as promoting lifelong learning in older adults (successful aging), using cognitive-based research findings.

Another outcome of our joint program is to use applied research to accelerate the transfer of results to industries and public institutions related to education and healthy aging in both countries. The mixed method approach used in our proposed project (user-centered methods, digital technologies, artificial intelligence, and field assessment) will help demonstrate the effectiveness of our developed technology, and facilitate adoption by industry partners and market stakeholders from various education and health care organizations.

10.2 International research visitors

10.2.1 Visits of international scientists

James McClelland

Status: Professor

Institution of origin: Stanford University

Country: USA

Dates: October 3rd, 2023

Context of the visit: Gave a seminar at Inria on "Capturing Intelligence at the Level of Thought" and visited the Flowers team

Mobility program/type of mobility: (Visit)

Kenji Doya

Status: Professor

Institution of origin: okinawa institute of science and technology (OIST)

Country: Japan

Dates: September 1st, 2023

Context of the visit: Gave a seminar at Inria on "Can robots find their own reward functions?" and visited the Flowers team

Mobility program/type of mobility: (Visit)

10.2.2 Visits to international teams

Research stays abroad

Rania Abdelghani

Visited institution: University of California, Berkeley. The Kidd lab

Country: USA

Dates: 1st October 2023 to 18 December 2023

Context of the visit: Working on designing an experimental protocol to understand children's use of Generative AI tools

Mobility program/ tpe of mobility: research stay

Marion Pech, Matisse Poupard, Maxime Adolphe

Visited institution: University of Waterloo, Fernandes Lab - Augmented Intelligence Lab

Country: Canada

Dates: 10th December 2023 to 19 December 2023

Context of the visit: Presenting results, getting feedbacks, creating new collaborations

Mobility program/ tpe of mobility: research stay

10.2.3 Horizon Europe

INTERACT [INTERACT project on cordis.europa.eu](https://cordis.europa.eu/project/interact)

Title: Help Me Grow: Artificial Cognitive Development via Human-Agent Interactions Supported by New Interactive, Intrinsically Motivated Program Synthesis Methods.

Duration: From October 1, 2022 to September 30, 2025

Partners:

- INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET AUTOMATIQUE (INRIA), France
- MASSACHUSETTS INSTITUTE OF TECHNOLOGY (MIT), United States

Inria contact: Cédric Colas

Coordinator:

Summary: Building machines that interact with their world, discover interesting interactions and learn open-ended repertoires of skills is a long-standing goal in AI. This project aims at tackling the limits of current AI systems by building on three families of methods: Bayesian program induction, intrinsically motivated learning and human-machine linguistic interactions. It targets three objectives: 1) building autonomous agents that learn to generate programs to solve problems with occasional human guidance; 2) studying linguistic interactions between humans and machines via web-based experiments (e.g. properties of human guidance, its impact on learning, human subjective evaluations); and 3) scaling the approach to the generation of constructions in

Minecraft, guided by real players. The researcher will collaborate with scientific pioneers and experts in the key fields and methods supporting the project. This includes supervisors Joshua Tenenbaum (program synthesis, MIT) and Pierre-Yves Oudeyer (autonomous learning, Inria); diverse collaborators, and an advisory board composed of an entrepreneur and leading scientists in developmental psychology and human-robot interactions. The 3rd objective will be pursued via a secondment with Thomas Wolf (CSO) at HuggingFace, a world-leading company in the open source development of natural language processing methods and their transfer to the industry. By enabling users to participate in the training of artificial agents, the project aims to open research avenues for more interpretable, performant and adaptive AI systems. This will result in scientific (e.g. interactive program synthesis approaches), societal (e.g. democratized AI training) and economic impacts (e.g. adaptive AI assistants). The dissemination, communication and exploitation plans support these objectives by targeting scientific (AI, cognitive science), industrial (video games, smart homes) and larger communities (gamers, software engineers, large public).

10.3 National initiatives

ANR Chaire Individuelle Deep Curiosity - PY Oudeyer continued to work on the research program of this Chaire, funding 2 PhDs and 3 postdocs for five years (until 2025).

ANR JCJC ECOCURL - C. Moulin-Frier obtained an ANR JCJC grant. The project is entitled "ECOCURL: Emergent communication through curiosity-driven multi-agent reinforcement learning". The project starts in Feb 2021 for a duration of 48 months. It will fund a PhD student (36 months) and a Research Engineer (18 months) as well as 4 Master internships (one per year).

Projet AIXIA: "Analyse d'Interférences par Intelligence Artificielle". Pierre-Yves Oudeyer and Clément Moulin-Frier obtained a grant from the call for project AIRSTRIP "L'intelligence Artificielle au service de l'Ingénierie des Systèmes Aéronautiques et spatiaux", in collaboration with the IRT Saint Exupéry. The project was accepted in 2023 and will fund 18 months of a research engineer position starting in 2024.

Inria Exploratory Action AIDE - Didier Roy is collaborator of the Inria Exploratory Action AIDE "Artificial Intelligence Devoted to Education", ported by Frédéric Alexandre (Inria Mnemosyne Project-Team), Margarida Romero (LINE Lab) and Thierry Viéville (Inria Mnemosyne Project-Team, LINE Lab). The aim of this Exploratory Action consists to explore to what extent approaches or methods from cognitive neuroscience, linked to machine learning and knowledge representation, could help to better formalize human learning as studied in educational sciences. AIDE is a four year project started middle 2020 until 2024 [see](#).

Inria Exploratory Action I'AM - Hélène Sauzéron is co-PI with P. Dragicevic of the Inria Exploratory Action I'AM "Impact of Augmented Reality on Autobiographical Memory: Examining Involuntary Memories and False Memories" (174,5k€). Starting in last september, the aim of this Exploratory Action consists to explore to what extent augmented reality based devices can produce erroneous autobiographical memories, and more particularly in vulnerable people (Children and older adults or young adults with low memory abilities of source monitoring).

New collaboration with Maxime Derex from IAST Toulouse for the co-direction of the PhD thesis of Jeremy Perez with Clément Moulin-Frier and Pierre-Yves Oudeyer on "Interactions between intrinsically motivated goal-exploration processes and cumulative cultural evolution" (see section [8.2.3](#)).

France 2030 - PPR AUTONOMIE : Vieillesse Et Situations De Handicap - Projet INNOVCare (Lechevalier S., 3,5M€) (2023-26) - Hélène Sauzéron and AS Rigaud will supervise the WP5 dedicated to two care-led innovation experiments with assistive technologies (400k € for Bordeaux).

VBHI project(Vascular Brain Health Institute -IHU, led by S. Debette, 5M€) (2023-26) - H el ene Sauz eon will supervise the WP4.3 dedicated to "Explore Digital Therapeutics To Slow Down Cognitive Decline In Covert Csvd" (150k )

10.3.1 Adaptiv'Math

- Adaptiv'Math
- Program: PIA
- Duration: 2019 - 2020
- Coordinator: EvidenceB
- Partners:
 - EvidenceB
 - Nathan
 - APMEP
 - LIP6
 - INRIA
 - ISOGRAD
 - Daesign
 - Schoolab
 - BlueFrog

The solution Adaptiv'Math comes from an innovation partnership for the development of a pedagogical assistant based on artificial intelligence. This partnership is realized in the context of a call for projects from the Ministry of Education to develop a pedagogical platform to propose and manage mathematical activities intended for teachers and students of cycle 2. The role of Flowers team is to work on the AI of the proposed solution to personalize the pedagogical content to each student. This contribution is based on the work done during the Kidlearn Project and the thesis of Benjamin Clement [106], in which algorithms have been developed to manage and personalize sequence of pedagogical activities. One of the main goal of the team here is to transfer technologies developed in the team in a project with the perspective of industrial scaling.

11 Dissemination

11.1 Promoting scientific activities

11.1.1 Scientific events: organisation

Member of the organizing committees

- Laetitia Teodorescu and C edric Colas have been organizers of the Intrinsically-Motivated and Open-Ended Learning workshop at NeurIPS 2023, <https://imol-workshop.github.io/>
- Mayalen Etcheverry was co-organizer of the Workshop on Agent Learning in Open-Endedness (ALOE) at Neurips 2023, <https://sites.google.com/view/aloe2023/home>
- Sao Mai Nguyen was a co-organizer of the The 6th International Workshop on Intrinsically Motivated Open-ended Learning (IMOL 2023) <https://imolconf2023.github.io/>
- H. Sauz eon has been member of the scientific committee of the first "Scientific Day of Gerontopole of New Aquitania", April 4th 2023, Limoges

- Pierre-Yves Oudeyer was co-organizer of the Life, Structure and Cognition symposium on Evolution and Learning, at IHES, <https://indico.math.cnrs.fr/event/9963/>, as well as of the Curiosity, Complexity and Creativity conference 2023 at Columbia University, NY, US, <https://zuckermaninstitute.columbia.edu/cc-event>.

11.1.2 Scientific events: selection

Reviewer

- Review for the Agent Learning in Open-Endedness (ALOE) workshop at NeurIPS (Gautier Hamon, Grgur Kovač, Clément Romac, Clément Moulin-Frier)
- Review for the Intrinsically Motivated Open-ended Learning (IMOL) workshop at NeurIPS (Grgur Kovač, Clément Romac, Thomas Carta, Laetitia Teodorescu)
- Review for the European Workshop on Reinforcement Learning (EWRL) (Clément Romac)

11.1.3 Journal

Member of the editorial boards PY Oudeyer was member of the editorial board of: IEEE Transactions on Cognitive and Developmental Systems and Frontiers in Neurobotics.

Reviewer - reviewing activities

- H. Sauz on reviewed 5 papers for Digital Health, Journal of Gerontology: Psychological Sciences, British Journal of Educational Technology, Ann e Psy, ACM IDC23.
- C. Mazon reviewed papers for Education and Information Technology.
- PY Oudeyer reviewed for the journals: Trends in Cognitive Science, Journal of Artificial Intelligence Research, Computers in Human Behaviour, Speech Communication.
- Review for Journal Of Artificial Intelligence Research (Cl ment Romac, Thomas Carta)
- Review for Humanities & Social Sciences Communications (Grgur Kovač)
- Review for International Journal for Artificial Intelligence in Education (Rania Abdelghani)

11.1.4 Invited talks

Cl ment Moulin-Frier gave an invited talk at the seminar **Marge, exception, d viation** at Universit  Bordeaux Montaigne on February 24th, 2023. Title of the talk:  cologie de la cognition humaine.

Cl ment Moulin-Frier gave an invited talk at the "2nd r union scientifique de la Soci t  Psych d lique Aquitaine" at the Hopital Saint Andr  in Bordeaux on October 17th, 2023. Title of the talk: The Ecology of Open-Ended Skill Acquisition: Eco-evolutionary, developmental and socio-cultural perspectives.

Cl ment Moulin-Frier gave an invited talk at the symposium **Intelligence: natural,artificial and synthetic** at the Barcelona Collaboratorium on October 5th, 2023. Title of the talk: Promoting behavioral diversity in artificial agents through eco-evolutionary and socio-cultural dynamics.

Cl ment Moulin-Frier gave an invited talk at the seminar of the GIPSA-Lab (Grenoble, France) on December 15th, 2023. Title of the talk: Modelling the eco-evolutionary, developmental and socio-cultural origins of open-ended skill acquisition.

C dric Colas gave an invited talk at the ENS in June 2023 during a Lab meeting of Stefano Palminteri. Title of the talk: Towards Social Autotelic Agents.

C dric Colas gave an invited talk at ICRA in the Life-Long Learning with Human Help Workshop, in July 2023. Title of the talk: Towards Social Autotelic Agents.

C dric Colas gave an invited talk at Brown University in the USA, during Lab meeting of George Konidaris, in September 2023. Title of the talk: Towards Social Autotelic Agents.

Thomas Carta gave an invited talk online during a reading group with Glen Berseth on RL focused on effective generalization and pre-training strategies for control in MILA and ServiceNow Research, in

April 2023. Title of the talk: Grounding Large Language Models in Interactive Environments with Online Reinforcement Learning.

Thomas Carta gave an invited talk online at Naver Labs Europe, in May 2023. Title of the talk: Grounding Large Language Models in Interactive Environments with Online Reinforcement Learning.

Clément Romac gave an invited talk at the COLT team's seminar at Universitat Pompeu Fabra (Barcelona) in May 2023. Title of the talk: Grounding Large Language Models in Interactive Environments with Online Reinforcement Learning.

Hélène Sauzéon gave 4 invited talks :

- Digital technology to support healthy aging : examples in the fields of smart home and computerized cognitive training” Symposium on Research on the Biology and Diseases of Ageing in Bordeaux - Disorders of tissue or cellular homeostasis during ageing and understanding of mechanisms and pathological consequences», 14th of November 2023, University of Bordeaux, Bordeaux
- Que fait le Centre Inria de l'université de Bordeaux pour l'enseignement et l'éducation ? , Journées scientifiques d'Inria, August 30th- September, 1st 2023 Bordeaux
- Participation to TABLE RONDE - Communautés, espaces sociaux, impact de l'IA sur la restructuration des sociétés Forum NAIA -Robocup, July, 6-7 th, Bordeaux
- Les interventions numériques auprès des personnes âgées : le cas des maisons intelligentes, une révolution en marche?, Scientific day on Digital technologie for aging in place, co-organised by The « Conseil de l'Age » (<https://www.hcfea.fr/>) , October 16th, 2023, Paris (Online participation)

Cécile Mazon gave an invited talk for Inria Disability conferences cycle (cycles de conférences Handicap) with Cathy Hémon (Autism Resource center, trainer and specialized teacher) on Autism Spectrum Disorders, coeducation, and participative methods to address field issues

PY Oudeyer gave several invited presentations:

- Keynote talk at the annual conference of the Vision Science Society of Japan, "Developmental AI: machines that learn like children" (Jan 2023);
- Public lecture of the Life, Structure and Cognition conference, "La curiosité: apprentissage, évolution et intelligence artificielle" (June 2023);
- Invited talk at the IMOL Conference, Paris, [Autotelic agents, open-endedness and applications](#) (Sept. 2023)
- Invited talk at the NeuroEconomics seminar of DEC, ENS, "[Open-ended learning in humans and AI: the roles of curiosity and language](#)"
- Invited talk at the NeuroDevRob workshop, Cergy Paris University, [Autotelic agents, open-endedness and applications](#) (Dec. 2023)

11.1.5 Leadership within the scientific community

Helene Sauzéon is the proxis of the Inria center of the University of Bordeaux réseau RT (GDR) CNRS Education since July 2022

Cédric Colas was member of the board of the IMOL community <https://www.imol-community.org/community/>

11.1.6 Scientific expertise

- Clément Moulin-Frier is on active discussion with the start-up [Pontos](#) in view of a future collaboration in 2024.

- Clément Moulin-Frier was a member of the CRCN/ISFP Jury of Inria Bordeaux on May 9th, 2023.

- Helene Sauzeon was member of the selection committee of DR2 09 (Univ. G. Eiffel) in 2023

- Helene Sauzéon was Member of the selection committee of PhD of MSCA COFUND SOUND.AI program (European Prog. at Sorbonne Univ) since 2023
- H  l  ne Sauz  on is member of the ANR committee - CES 38 (Interdisciplinary research section)- since November 2023.
- Helene Sauz  on was member of scientific committee of MAVIE-II -Calyxis [link](#).
- PY Oudeyer was a member of the jury selecting grants for PhDs in AI in the context of SoundAI project at Sorbonne University
- PY Oudeyer reviewed projects for the European Commission (ERC, Marie Curie grants, EU Pathfinder), for the US/Israel Binational Science Foundation, for ANR, for RIF Cyprus, Leverhulme Trust.

11.1.7 Research administration

- H  l  ne Sauz  on and C  cile Mazon are members of directory committee of LILLAB (<https://-www.-lillabneurodev.-fr/>) which is a living and learning lab funded by the “d  l  gation interminist  rielle    la strat  gie nationale    l’autisme et troubles neurod  veloppementaux” and aiming the dissemination of knowledge in connection with the 3 centers of excellence for autism and Neurodevelopmental syndromes; since 2020.
 - H  l  ne Sauz  on is member of directory committee of IFHR / FEDRAH (<https://-ifr-handicap.-inserm.-fr/>) which is a national institute on disability funded by Inserm aiming the researcher networking and dissemination of knowledge on multidisciplinary research on disability; since 2018.
 - H  l  ne Sauz  on is the head of the Innovations and Transfer Committee of the BIND Center of Excellence [link](#) in Bordeaux and member of the BIND directory Committee since 2018
 - C  cile Mazon is co-responsible of the WP digital tools of the PIA AtypieFriendly (ex - AspieFriendly)
 - Pierre-Yves Oudeyer is head of the Flowers project-team, Inria/Univ. Bordeaux/ENSTA ParisTech

11.2 Teaching - Supervision - Juries

11.2.1 Teaching

Teaching Responsibilities:

- Cl  ment Moulin-Frier is responsible professor of the "System Design, Integration and Control" course at the University Pompeu Fabra in Barcelona, Spain.
- C  cile Mazon is responsible of the second year of the curriculum in Technology, Ergonomics, Cognition and Handicap (TECH, Cognitive Sciences - University of Bordeaux) since sept. 2021.
- C  cile Mazon is responsible of the curriculum in Technology, Ergonomics, Cognition and Handicap (TECH, Cognitive Sciences - University of Bordeaux) since sept. 2022.
- David Filliat is in charge since 2012 of the "Robotics and autonomous systems" third year speciality at ENSTA Paris.
- Sao Mai Nguyen is in charge of the "Robot Learning" third year course at ENSTA Paris.

Teaching Involvement in Computer / Engineer science or in cognitive science:

- BS & Master: Cognitive Science, Univ. of Bordeaux- , 37h, Marion Pech
- BS & Master: Neuropsychology, Univ. of Caen-,2H, Marion Pech
- ENSC/ENSEIRB Presentation of developmental artificial intelligence and the Flowers Lab, 2h, Option Robot (Laetitia Teodorescu)
- ENSC Introduction to bayesian analysis, 8h, Option AI (Adolphe Maxime)
- ENSC Transformers and Large Language Models, 8h, Option AI (Cl  ment Romac)
- BS & Master: Cognitive Science, Univ. of Bordeaux- , 192h, C  cile Mazon

- Master: Navigation for Robotics, 21 h, M2, ENSTA Paris, David Filliat
- Master: Navigation for Robotics, 24 h, M2 DataAI, IP Paris - Paris, David Filliat
- Université de Bordeaux - MIASHS Bachelor : Models and measures of high-level cognitive functions - knowledge and representation, 22h, (Poupard Matisse)
- Université de Bordeaux - TECH Master: Virtual reality, interaction and healthcare applications, 7h, (Poupard Matisse)
- Université de Bordeaux - TECH Master: Scientific basis, 2h, (Poupard Matisse)
- 2nd year : Deep Learning, 12h, IMT Atlantique (Sao Mai Nguyen).
- Université de Bordeaux - TECH Master: IT Project Management, 18h (Isabeau Saint-Supery)
- Master UPF-Barcelona: Robotics and AI, 10h (Clément Moulin-Frier)
- PY Oudeyer gave a course on developmental reinforcement learning at ENSEIRB (2h), dec. 2023.
- PY Oudeyer gave a course on developmental reinforcement learning at Cogmaster at Sorbonne Univ./ENS (2h+4h project jury), dec. 2023.
- Master: Cognitive Science, 24h (Eric Meyer)
- Teacher training at the Rectorat of Bordeaux : digital technologies for students with special educational needs , 6h (Eric Meyer)
- 2nd year Master in cognitive science : Assistive technologies (20h), Rania Abdelghani.
- University Degree in Neuropsychological Sciences : Cognitive aging and Digital technologies (3h), Helene Sauzéon
- Clément Moulin-Frier gave a class in the Parcours Robotique at ENSEIRB-MATMECA (Bordeaux INP). 2 hours, November 2023.

11.2.2 Supervision

- PhD defended: Tristan Karch (defended in 2023), "Language acquisition in curiosity-driven Deep RL", beg. in sept. 2019 (supervisors: PY. Oudeyer and C. Moulin-Frier)
 - PhD defended: Mayalen Etcheverry (defended in 2023), "Automated discovery with intrinsically motivated goal exploration processes", beg. in sept. 2020 (supervisors: PY. Oudeyer and C. Moulin-Frier)
 - PhD defended: Laetitia Teodorescu, (defended in 2023), "Graph Neural Networks in Curiosity-driven Exploring Agents", beg. in sept. 2020 (supervisors: PY. Oudeyer and K. Hoffman)
 - PhD in progress: J Grgur Kovac (in progress), "Developmental training of socio-cognitive abilities in AI systems", (supervisors:PF. Dominey and PY. Oudeyer)
 - PhD in progress: J Julien Pourcel (in progress), "Autotelic LLMs that learn how to code", (supervisors: C. Moulin-Frier and PY. Oudeyer)
 - PhD in progress: J Thomas Carta (in progress), "LLM-based Autotelic deep reinforcement learning agents", (supervisors: O. Sigaud, S. Lamprier and PY. Oudeyer)
 - PhD in progress: J Clément Romac (in progress), "Grounding LLMs with online RL", (supervisors: T. Wolf and PY. Oudeyer)
 - PhD in progress: Jeremy Perez, supervised by Clément Moulin-Frier, which started in October 2023
 - PhD in progress : Chloé Desvaux "Design and experiment new metacognitive trainings for fostering curiosity and creativity among children in a school setting: a lever for intrinsically motivated learning?", beg. in October. 2023 (supervised by H. Sauzéon and PY Oudeyer).
 - PhD in progress : Leana Petiot " Study of Augmented reality on the functioning of Implicit autobiographical memory", beg. in October. 2023 (supervised by H. Sauzéon and P. Dragicevic from Potioc team).

- PhD in progress: Maxime Adolphe, "Adaptive personalization in attention training systems", beg. in sept. 2020 (supervisors: H. Sauzéon and PY. Oudeyer)
- PhD in progress: Rania Abdelgani, "Fostering curiosity and meta-cognitive skills in educational technologies", beg. in dec. 2020 (supervisors: H. Sauzéon and PY. Oudeyer).
- PhD in progress: Isabeau Saint-Supery, "Designing and Assessing a new interactive tool fostering stakeholders' cooperation for school inclusion", supervised by H. Sauzéon and C. Mazon.
- PhD in progress : Matisse Poupard "Optimize learning in a digital environment according to learners' level of expertise, epistemic curiosity and mode of instruction" ",beg. in Ap. 2022 (supervised by H. Sauzéon and A. Tricot from Univ. Montpellier).

Gautier Hamon and Clément Moulin-Frier supervised the Master internships of Richard Bornemann and of Corentin Léger (in collaboration with Xavier Hinaut from Mnemosyne) in 2023.

Maxime Adolphe and Hélène Sauzéon supervised the Master internships of Stéphanie Mortemousque in 2023.

Cécile Mazon and Isabeau Saint-Supery supervised the Master internship of Valentin Strahm in 2023.

Rania Abdelghani and Hélène Sauzéon supervised the Master internship of Chloé Desvaux in 2023.

Pierre Dragicevic and Hélène Sauzéon supervised the Master internship of Léana Petiot in 2023.

11.2.3 Juries

H. Sauzéon has been member of scientific board of HDR degree of P. Dragicevic

Clément Moulin-Frier was a member of the PHD jury of Joachim Winther Pedersen (thesis director: Sebastian Risi, Univeristy of Copenhagen).

H. Sauzéon has been president of PhD jury of Axelle Gelineau on " Projet RGS@HOME : Evaluation de l'acceptabilité d'un système de télé-réhabilitation membre supérieur basé sur la réalité virtuelle auprès des patients post-AVC. " at the university of Limoges, December, 8th 2023, Limoges.

H. Sauzéon has been reviewer in the PhD jury of Marine Saba on " Efficacité d'un programme d'entraînement des ressources attentionnelles et de la mémoire de travail chez les personnes âgées avec un trouble cognitif léger : Effets sur les fonctions cognitives et une situation écologique évaluée avec la réalité virtuelle " at the university of Paris, December, 15th 2023, Paris.

Clément Moulin-Frier was a member of the jury of the **Premier hackathon consacré à l'enquête journalistique sur les algorithmes** at IJBA (Bordeaux) on November 30th, 2023.

Clément Moulin-Frier is a member of the "comité de suivi de thèse" of Nathan Trouvain (Mnemosyne, Inria) and Camille Charrier (LPNC, Grenoble).

Hélène sauzéon is a member of the "comité de suivi de thèse" of Hugo Fournier (Lab Psychology, Bordeaux).

Hélène sauzéon is academic tutor for 2 PhD students of the Doctoral School SP2.

Cécile Mazon organized and was president of the jury of the defense of Cognitive Sciences Master students (M1 and M2)

PY Oudeyer was a reviewer in the HdR of Erik Gustaffsson (Univ. Bourgogne Franche Comté), and examiner in the PhD of Enrique Donancio (INSA Rouen Normandie) and Lina Mezghani (Univ. Grenoble).

PY Oudeyer was in the PhD "comité de suivi" of Marc Welter (Univ. Bordeaux), Jean-Baptiste Gaya (Université Paris-Sorbonne), Marie Martin (Univ. Paris Saclay), Elias Najarro (Univ. Copenhagen), Matthis Poupard (Univ. Bordeaux),

11.3 Popularization

11.3.1 Internal or external Inria responsibilities

• Hélène Sauzéon was member of extented office of Project-team committee of the centre of Inria of university of Bordeaux.

• Hélène Sauzéon was student in the Inria MasterClass since sep. 2022.

• PY Oudeyer contributed several internal notes on AI in society, helping Inria direction answer several requests on this topic from governmental organizations.

11.3.2 Articles and contents

- Clément Romac, Thomas Carta and PY Oudeyer wrote down "Quand les IA font l'expérience de la réalité" explaining the recently published paper "Grounding Large Language Models in Interactive Environments with Online Reinforcement Learning" (to be published in "Pour La Science" march 2024)
- Alexandre Torres-Leguet, Clément Romac, Thomas Carta and PY Oudeyer created the (on-going) series of videos "ChatGPT expliqué en 5 minutes", aimed at training AI literacy in a wide diversity of students (e.g. high school), available here: https://developmentalsystems.org/chatgpt_en_5_minutes/. It was already integrated in the MOOC AI4T (<https://www.ai4t.eu>).
- Eleni Nisioti, Clément Moulin-Frier and Pierre-Yves Oudeyer published a Blog post on [An ecological approach to Artificial Intelligence](#).
- Rania Abdelghani, Pierre-Yves Oudeyer and Hélène Sauzéron published a blog post for the ministry of education on [Curiosité, numérique et apprentissage: que dit la recherche?](#).
- Rania Abdelghani, Pierre-Yves Oudeyer, and Maxime Adolphe were interviewed for a [TV program](#) about AI during the ROBOCUP 2023 event.
- Hélène Sauzéron gave an interview for Sud-Ouest Newspaper on "Assistants vocaux : faut-il laisser les enfants utiliser seuls une enceinte connectée ?", December, 14th, 2023 <https://www.sudouest.fr/societe/parentalite/assistants-vocaux-faut-il-laisser-les-enfants-utiliser-seul-une-enceinte-connectee-17789838.php>
- PY Oudeyer gave an interview on the societal dimensions of AI at [France Bleu/France 3 gironde](#) (Dec 2023)
- PY Oudeyer gave an interview on cultural and social stakes of AI in [Les Echos Judiciaires](#)
- PY Oudeyer gave an interview on large language models/ChatGPT at [Sud-Ouest](#).

11.3.3 Education

- 3 "Stagiaire de 3e", Presentation of topics of the team (generative AI, Large language models ,Cellular automata, Reinforcement learning) February 2023 (Clément Romac, Maxime Adolphe)
- "Stagiaire de 3e", Presentation of topics of the team (generative AI, Large language models ,Cellular automata, Reinforcement learning,) July 2023 (Gautier Hamon, Matisse Poupart)
- 2 "Stagiaire de 3e", Presentation of topics of the team (generative AI, Large language models ,Cellular automata, Reinforcement learning, Curiosity and Learning in Human Cognition) November 2023 (Gautier Hamon, Matisse Poupart, Maxime Adolphe)
- 3 "Stagiaire de 3e", Presentation of topics of the team (generative AI, Large language models ,Cellular automata, Reinforcement learning, Curiosity and Learning in Human Cognition) December 2023 (Clément Romac, Matisse Poupart, Chloé Desvaux, Marion Pech, Hélène Sauzéron)
- 2 Classrooms for the CHICHE - One researcher -One classroom- Lycée Václav Havel, Bègles, february 2023
- Podcast - "Etudes scientifiques pluridisciplinaires", Bègles, february 2023
- The 1st edition of the Hackathon [Hack1Robo](#) at the Fablab of Cap Science (June 2-4, 2023) was co-organized by Clément Moulin-Frier.

11.3.4 Interventions

- Clément Moulin-Frier, H el ene Sauzeon and Clement Romac have participated to the Workshop (4h) on AI and its impact for French teachers at Boussoles du Num erique (sep. 2023)
- Workshop (2h) on generative AIs for education at INSPE (nov. 2023)
- Cl ement Moulin-Frier was interviewed by Antoine Beauchamp for the radio program **La Science, CQFD**. The program theme was "Origine de la vie : raconte-moi ta LYFE" and it was broadcasted on France Culture on June 21st, 2023.
- Sauz eon gave a talk on « Outils num eriques et Maladies Neurod eg en eratives : une r evolution en marche ? » for the "Journ ee Grand Public - Chercheurs, patients, aidants : Tous ensemble !, organisez by the patients' association, la Maison du Cerveau, March 9th 2023, Bordeaux.
- Table ronde   NAlA-R, Bordeaux, "**IA G en eratives: ChatGPT and co**" (July 2023)
- Invited presentation at a Cafe P edagogique seminar, « **L'intelligence artificielle peut-elle contribuer   stimuler la curiosit  et la m eta-cognition chez les enfants ?** » (Sept. 2023)
- Invited presentation at FrenchTech Days in Bordeaux, « **Entraîner la curiosit  chez les enfants avec ChatGPT, c'est possible ?** »

12 Scientific production

12.1 Major publications

- [1] R. Abdelghani, P.-Y. Oudeyer, E. Law, C. de Vulpilli eres and H. Sauz eon. 'Conversational agents for fostering curiosity-driven learning in children'. In: *International Journal of Human-Computer Studies* 167 (Nov. 2022), p. 102887. DOI: [10.1016/j.ijhcs.2022.102887](https://doi.org/10.1016/j.ijhcs.2022.102887). URL: <https://hal.archives-ouvertes.fr/hal-03779096>.
- [2] M. Adolphe, M. Sawayama, D. Maurel, A. Delmas, P.-Y. Oudeyer and H. Sauzeon. 'An Open-Source Cognitive Test Battery to Assess Human Attention and Memory'. In: *Frontiers in Psychology* 13 (10th June 2022). DOI: [10.3389/fpsyg.2022.880375](https://doi.org/10.3389/fpsyg.2022.880375). URL: <https://hal.archives-ouvertes.fr/hal-03723887>.
- [3] A. Akakzia, C. Colas, P.-Y. Oudeyer, M. Chetouani and O. Sigaud. 'Grounding Language to Autonomously-Acquired Skills via Goal Generation'. In: ICLR 2021 - Ninth International Conference on Learning Representation. Vienna / Virtual, Austria, 4th May 2021. URL: <https://hal.inria.fr/hal-03121146>.
- [4] M. Alaimi, E. Law, K. D. Pantasdo, P.-Y. Oudeyer and H. Sauz eon. 'Pedagogical Agents for Fostering Question-Asking Skills in Children'. In: *CHI '20 - CHI Conference on Human Factors in Computing Systems*. Honolulu / Virtual, United States, Apr. 2020. DOI: [10.1145/3313831.3376776](https://doi.org/10.1145/3313831.3376776). URL: <https://hal.inria.fr/hal-03109511>.
- [5] A. Baranes and P.-Y. Oudeyer. 'Active Learning of Inverse Models with Intrinsically Motivated Goal Exploration in Robots'. In: *Robotics and Autonomous Systems* 61.1 (Jan. 2013), pp. 69–73. DOI: [10.1016/j.robot.2012.05.008](https://doi.org/10.1016/j.robot.2012.05.008). URL: <https://hal.inria.fr/hal-00788440>.
- [6] H. Caselles-Dupr e, M. Garcia-Ortiz and D. Filliat. 'S-TRIGGER: Continual State Representation Learning via Self-Triggered Generative Replay'. In: IJCNN 2021 - International Joint Conference on Neural Networks. Shenzhen / Virtual, China: IEEE, 18th July 2021, pp. 1–7. DOI: [10.1109/IJCNN52387.2021.9533683](https://doi.org/10.1109/IJCNN52387.2021.9533683). URL: <https://hal.archives-ouvertes.fr/hal-03377783>.
- [7] H. Caselles-Dupr e, M. Garcia-Ortiz and D. Filliat. 'Symmetry-Based Disentangled Representation Learning requires Interaction with Environments'. In: *NeurIPS 2019*. Vancouver, Canada, Dec. 2019. URL: <https://hal.archives-ouvertes.fr/hal-02379399>.

- [8] P.-A. Cinquin, P. Guitton and H. Sauz on. ‘Towards Truly Accessible MOOCs for Persons with Cognitive Impairments: a Field Study’. In: *Human-Computer Interaction* (2021). URL: <https://hal.archives-ouvertes.fr/hal-03436244>.
- [9] C. Colas, P. Fournier, O. Sigaud, M. Chetouani and P.-Y. Oudeyer. ‘CURIOUS: Intrinsically Motivated Modular Multi-Goal Reinforcement Learning’. In: *International Conference on Machine Learning*. Long Beach, France, June 2019. URL: <https://hal.archives-ouvertes.fr/hal-01934921>.
- [10] C. Colas, B. P. Hejblum, S. Rouillon, R. Thi ebaut, P.-Y. Oudeyer, C. Moulin-Frier and M. Prague. ‘EpidemiOptim: a Toolbox for the Optimization of Control Policies in Epidemiological Models’. In: *Journal of Artificial Intelligence Research* (23rd July 2021). DOI: [10.1613/jair.1.12588](https://doi.org/10.1613/jair.1.12588). URL: <https://hal.inria.fr/hal-03099898>.
- [11] C. Colas, T. Karch, N. Lair, J.-M. Dussoux, C. Moulin-Frier, P. F. Dominey and P.-Y. Oudeyer. ‘Language as a Cognitive Tool to Imagine Goals in Curiosity-Driven Exploration’. In: *NeurIPS 2020 - 34th Conference on Neural Information Processing Systems*. Contains main article and supplementaries. Vancouver / Virtual, Canada, Dec. 2020. URL: <https://hal.archives-ouvertes.fr/hal-03083158>.
- [12] C. Colas, O. Sigaud and 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. URL: <https://hal.inria.fr/hal-01890151>.
- [13] C. Craye, T. Lesort, D. Filliat and J.-F. Goudou. ‘Exploring to learn visual saliency: The RL-IAC approach’. In: *Robotics and Autonomous Systems* 112 (Feb. 2019), pp. 244–259. URL: <https://hal.archives-ouvertes.fr/hal-01959882>.
- [14] N. Duminy, S. M. Nguyen, J. Zhu, D. Duhaut and J. Kerdreux. ‘Intrinsically Motivated Open-Ended Multi-Task Learning Using Transfer Learning to Discover Task Hierarchy’. In: *Applied Sciences* 11.3 (Feb. 2021), p. 975. DOI: [10.3390/app11030975](https://doi.org/10.3390/app11030975). URL: <https://hal.inria.fr/hal-03118190>.
- [15] M. Eppe and P.-Y. Oudeyer. ‘Intelligent Behavior Depends on the Ecological Niche’. In: *KI - K unstliche Intelligenz* (3rd Jan. 2021). DOI: [10.1007/s13218-020-00696-1](https://doi.org/10.1007/s13218-020-00696-1). URL: <https://hal.inria.fr/hal-03120583>.
- [16] M. Etcheverry, C. Moulin-Frier and P.-Y. Oudeyer. ‘Hierarchically Organized Latent Modules for Exploratory Search in Morphogenetic Systems’. In: *NeurIPS 2020 - 34th Conference on Neural Information Processing Systems*. Vancouver / Virtual, Canada, Dec. 2020. URL: <https://hal.inria.fr/hal-03099906>.
- [17] S. Forestier, R. Portelas, Y. Mollard and P.-Y. Oudeyer. ‘Intrinsically Motivated Goal Exploration Processes with Automatic Curriculum Learning’. In: *Journal of Machine Learning Research* (1st Apr. 2022). URL: <https://hal.science/hal-01651233>.
- [18] J. Gottlieb and P.-Y. Oudeyer. ‘Towards a neuroscience of active sampling and curiosity’. In: *Nature Reviews Neuroscience* 19.12 (Dec. 2018), pp. 758–770. URL: <https://hal.inria.fr/hal-01965608>.
- [19] T. Karch, L. Teodorescu, K. Hofmann, C. Moulin-Frier and P.-Y. Oudeyer. ‘Grounding Spatio-Temporal Language with Transformers’. In: *NeurIPS 2021 - 35th Conference on Neural Information Processing Systems*. Virtuel, France, 7th Dec. 2021. URL: <https://hal.archives-ouvertes.fr/hal-03439096>.
- [20] A. Laversanne-Finot, A. P er e and P.-Y. Oudeyer. ‘Curiosity Driven Exploration of Learned Disentangled Goal Spaces’. In: *CoRL 2018 - Conference on Robot Learning*. Z urich, Switzerland, Oct. 2018. URL: <https://hal.inria.fr/hal-01891598>.
- [21] T. Lesort, N. D ıaz-Rodr ıguez, J.-F. Goudou and D. Filliat. ‘State Representation Learning for Control: An Overview’. In: *Neural Networks* 108 (Dec. 2018), pp. 379–392. DOI: [10.1016/j.neunet.2018.07.006](https://doi.org/10.1016/j.neunet.2018.07.006). URL: <https://hal.archives-ouvertes.fr/hal-01858558>.
- [22] C. Mazon, B. Cl ement, D. Roy, P.-Y. Oudeyer and H. Sauz on. ‘Pilot study of an intervention based on an intelligent tutoring system (ITS) for instructing mathematical skills of students with ASD and/or ID’. In: *Education and Information Technologies* (2022). DOI: [10.1007/s10639-022-11129-x](https://doi.org/10.1007/s10639-022-11129-x). URL: <https://hal.inria.fr/hal-03676589>.

- [23] M. E. Meade, J. G. Meade, H. Sauz on and M. A. Fernandes. ‘Active Navigation in Virtual Environments Benefits Spatial Memory in Older Adults’. In: *Brain Sciences* 9 (2019). DOI: [10.3390/brainsci9030047](https://doi.org/10.3390/brainsci9030047). URL: <https://hal.inria.fr/hal-02049031>.
- [24] C. Moulin-Frier, J. Brochard, F. Stulp and P.-Y. Oudeyer. ‘Emergent Jaw Predominance in Vocal Development through Stochastic Optimization’. In: *IEEE Transactions on Cognitive and Developmental Systems* 99 (2017), pp. 1–12. DOI: [10.1109/TCDS.2017.2704912](https://doi.org/10.1109/TCDS.2017.2704912). URL: <https://hal.inria.fr/hal-01578075>.
- [25] E. Nisioti, K. Jodogne-del Litto and C. Moulin-Frier. ‘Grounding an Ecological Theory of Artificial Intelligence in Human Evolution’. In: *NeurIPS 2021 - Conference on Neural Information Processing Systems / Workshop: Ecological Theory of Reinforcement Learning*. virtual event, France, 6th Dec. 2021. URL: <https://hal.archives-ouvertes.fr/hal-03446961>.
- [26] A. P er , S. Forestier, O. Sigaud and P.-Y. Oudeyer. ‘Unsupervised Learning of Goal Spaces for Intrinsically Motivated Goal Exploration’. In: *ICLR2018 - 6th International Conference on Learning Representations*. Vancouver, Canada, Apr. 2018. URL: <https://hal.archives-ouvertes.fr/hal-01891758>.
- [27] R. Portelas, C. Colas, K. Hofmann and P.-Y. Oudeyer. ‘Teacher algorithms for curriculum learning of Deep RL in continuously parameterized environments’. In: *CoRL 2019 - Conference on Robot Learning*. <https://arxiv.org/abs/1910.07224>. Osaka, Japan, Oct. 2019. URL: <https://hal.archives-ouvertes.fr/hal-02370165>.
- [28] R. Portelas, C. Colas, L. Weng, K. Hofmann and P.-Y. Oudeyer. ‘Automatic Curriculum Learning For Deep RL: A Short Survey’. In: *IJCAI 2020 - International Joint Conference on Artificial Intelligence*. Kyoto / Virtuelle, Japan, Jan. 2021. URL: <https://hal.archives-ouvertes.fr/hal-02923635>.
- [29] C. Reinke, M. Etcheverry and P.-Y. Oudeyer. ‘Intrinsically Motivated Discovery of Diverse Patterns in Self-Organizing Systems’. In: *International Conference on Learning Representations (ICLR)*. Source code and videos at <https://automated-discovery.github.io/>. Addis Ababa, Ethiopia, Apr. 2020. URL: <https://hal.inria.fr/hal-02370003>.
- [30] C. Romac, R. Portelas, K. Hofmann and P.-Y. Oudeyer. ‘TeachMyAgent: a Benchmark for Automatic Curriculum Learning in Deep RL’. In: *Proceedings of the 38th International Conference on Machine Learning, PMLR 139, 2021*. ICML 2021 - Thirty-eighth International Conference on Machine Learning. Vol. 139. Proceedings of the 38th International Conference on Machine Learning. Vienna / Virtual, Austria, 18th July 2021, pp. 9052–9063. URL: <https://hal.archives-ouvertes.fr/hal-03173198>.
- [31] A. Ten, P. Kaushik, P.-Y. Oudeyer and J. Gottlieb. ‘Humans monitor learning progress in curiosity-driven exploration’. In: *Nature Communications* 12.1 (Dec. 2021). DOI: [10.1038/s41467-021-26196-w](https://doi.org/10.1038/s41467-021-26196-w). URL: <https://hal.inria.fr/hal-03378913>.
- [32] G. Valle Perez, J. Beskow, G. E. Henter, A. Holzapfel, P.-Y. Oudeyer and S. Alexanderson. ‘Transflower: probabilistic autoregressive dance generation with multimodal attention’. In: *SIGGRAPH Asia 2021 - 14th ACM SIGGRAPH Conference and Exhibition on Computer Graphics and Interactive Techniques*. Tokyo, Japan, 14th Dec. 2021. DOI: [10.1145/3478513.3480570](https://doi.org/10.1145/3478513.3480570). URL: <https://hal.archives-ouvertes.fr/hal-03449915>.

12.2 Publications of the year

International journals

- [33] R. Abdelghani, Y.-H. Wang, X. Yuan, T. Wang, P. Lucas, H. Sauz on and P.-Y. Oudeyer. ‘GPT-3-driven pedagogical agents for training children’s curious question-asking skills’. In: *International Journal of Artificial Intelligence in Education*. International Journal of Artificial Intelligence in Education (30th June 2023). DOI: [10.1007/s40593-023-00340-7](https://doi.org/10.1007/s40593-023-00340-7). URL: <https://hal.science/hal-03890447>.

- [34] X. Arsiwalla, R. Solé, C. Moulin-Frier, I. Herreros, M. Sánchez-Fibla and P. Verschure. ‘The Morphospace of Consciousness: Three Kinds of Complexity for Minds and Machines’. In: *NeuroSci* 4.2 (27th Mar. 2023), pp. 79–102. DOI: [10.3390/neurosci4020009](https://doi.org/10.3390/neurosci4020009). URL: <https://inria.hal.science/hal-04356911>.
- [35] D. Brellmann, D. Filliat and G. Frehse. ‘Fourier Features in Reinforcement Learning with Neural Networks’. In: *Transactions on Machine Learning Research Journal* (Oct. 2023). URL: <https://hal-ip-paris.archives-ouvertes.fr/hal-04316346>.
- [36] L. Brinkmann, F. Baumann, J.-F. Bonnefon, M. Derex, T. F. Müller, A.-M. Nussberger, A. Czaplicka, A. Acerbi, T. Griffiths, J. Henrich, J. Leibo, R. Mcelreath, P.-Y. Oudeyer, J. Stray and I. Rahwan. ‘Machine culture’. In: *Nature Human Behaviour* 7.11 (20th Nov. 2023), pp. 1855–1868. DOI: [10.1038/s41562-023-01742-2](https://doi.org/10.1038/s41562-023-01742-2). URL: <https://hal.science/hal-04326925>.
- [37] P.-A. Cinquin, P. Guitton and H. Sauzéon. ‘Towards Truly Accessible MOOCs for Persons with Cognitive Impairments: A Field Study’. In: *Human-Computer Interaction* 38.5-6 (10th May 2023). DOI: [10.1080/07370024.2021.2008250](https://doi.org/10.1080/07370024.2021.2008250). URL: <https://inria.hal.science/hal-03435461>.
- [38] A. Frick, S. Fay, B. Bouazzaoui, H. Sauzéon, L. Angel, S. Vanneste and L. Taconnat. ‘The respective contribution of cognitive control and working memory to semantic and subjective organization in aging’. In: *Psychology and Aging* (2023). DOI: [10.1037/pag0000752](https://doi.org/10.1037/pag0000752). URL: <https://hal.science/hal-04080498>.
- [39] M. Gasse, D. Grasset, G. Gaudron and P.-Y. Oudeyer. ‘Using Confounded Data in Latent Model-Based Reinforcement Learning’. In: *Transactions on Machine Learning Research Journal* (Aug. 2023). URL: <https://inria.hal.science/hal-04404106>.

National journals

- [40] J. Cuadrado, É. Meyer, J. Maire, C. Legigan, C. Sentenac, M. Athane, M. Demont, C. Langa, A. Mence and G. Michel. ‘“Gifted” children, challenges of an integrative psychological assessment in the adolescent clinic. A case study.’ In: *Annales Médico-Psychologiques, Revue Psychiatrique* (May 2023). DOI: [10.1016/j.amp.2023.04.010](https://doi.org/10.1016/j.amp.2023.04.010). URL: <https://hal.science/hal-04108902>.

International peer-reviewed conferences

- [41] R. Abdelghani, E. Law, C. Desvaux, P.-Y. Oudeyer and H. Sauzéon. ‘Interactive environments for training children’s curiosity through the practice of metacognitive skills : a pilot study’. In: IDC 2023 - The 22nd annual ACM Interaction Design and Children Conference. Chicago IL, United States: ACM; ACM, 30th Nov. 2023, pp. 495–501. DOI: [10.1145/3585088.3593880](https://doi.org/10.1145/3585088.3593880). URL: <https://hal.science/hal-04312937>.
- [42] L. Annabi, Z. Ma and S. M. Nguyen. ‘Unsupervised Motion Retargeting for Human-Robot Imitation’. In: Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction. Boulder (CO), United States: ACM, 11th Mar. 2024. DOI: [10.1145/3568294.3580153](https://doi.org/10.1145/3568294.3580153). URL: <https://hal.science/hal-04401885>.
- [43] L. Annabi and S. M. Nguyen. ‘Prerequisite Structure Discovery in Intelligent Tutoring Systems’. In: 2023 IEEE International Conference on Development and Learning (ICDL). Macau, China: IEEE, 9th Nov. 2023, pp. 176–181. DOI: [10.1109/icdl55364.2023.10364416](https://doi.org/10.1109/icdl55364.2023.10364416). URL: <https://inria.hal.science/hal-04401880>.
- [44] T. Carta, C. Romac, T. Wolf, S. Lamprier, O. Sigaud and P.-Y. Oudeyer. ‘Grounding Large Language Models in Interactive Environments with Online Reinforcement Learning’. In: International Conference on Machine Learning 2023. Vol. 202. 3676-3713. Honolulu, Hawaii, United States, 2023. URL: <https://hal.science/hal-03970122>.
- [45] C. Colas, L. Teodorescu, P.-Y. Oudeyer, E. X. Yuan and M.-A. Côté. ‘Augmenting Autotelic Agents with Large Language Models’. In: CoLLAs 2023 - Conference on Lifelong Learning Agents. Montreal, Canada, 22nd Aug. 2023. URL: <https://hal.science/hal-04375117>.

- [46] T. Dupuis, J. Rabarisoa, Q.-C. Pham and D. Filliat. ‘VIBR: Learning View-Invariant Value Functions For Robust Visual Control’. In: CoLLAs 2023 - Conference on Lifelong Learning Agents. Montreal, Canada, 25th Aug. 2023. URL: <https://hal-ip-paris.archives-ouvertes.fr/hal-04202697>.
- [47] G. Franchi, X. Yu, M. Hariat, A. Manzanera and D. Filliat. ‘InfraParis: A multi-modal and multi-task autonomous driving dataset’. In: IEEE/CVF Winter Conference on Applications of Computer Vision (WACV). Waikoloa, United States, 2024. URL: <https://hal.science/hal-04321062>.
- [48] G. Hamon, E. Nisioti and C. Moulin-Frier. ‘Eco-evolutionary Dynamics of Non-episodic Neuroevolution in Large Multi-agent Environments’. In: GECCO ’23 Companion - Companion Conference on Genetic and Evolutionary Computation. Lisbon Portugal, France: ACM; ACM, July 2023, pp. 143–146. DOI: 10.1145/3583133.3590703. URL: <https://inria.hal.science/hal-04379640>.
- [49] G. Kovač, R. Portelas, P. F. Dominey and P.-Y. Oudeyer. ‘The SocialAI School: Insights from Developmental Psychology Towards Artificial Socio-Cultural Agents’. In: TOM 2023 - Theory-of-Mind Workshop at PMLR-202. Honolulu (Hawaii), United States, 23rd Nov. 2023. URL: <https://inria.hal.science/hal-04365787>.
- [50] A. Marusic, S. M. Nguyen and A. Tapus. ‘Evaluating Kinect, OpenPose and BlazePose for Human Body Movement Analysis on a Low Back Pain Physical Rehabilitation Dataset’. In: HRI 2023 - ACM/IEEE International Conference on Human-Robot Interaction. Stockholm, Sweden: ACM, 13th Mar. 2023, pp. 587–591. DOI: 10.1145/3568294.3580153. URL: <https://hal.science/hal-04133691>.
- [51] E. Meyer, H. Sauzéon, I. Saint-Supery and C. Mazon. ‘Systematic review of technologies to collaborate and co-educate students with special educational needs and supporting their schooling’. In: IHMET 2023 - 10th International Conference on Human Interaction and Emerging Technologies. Vol. 111. Nice, France: AHFE International, 31st July 2023, pp. 1–12. DOI: 10.54941/ahfe1004060. URL: <https://inria.hal.science/hal-04222434>.
- [52] E. Nisioti, E. Masquil, G. Hamon, Moulin-Frier and Clément. ‘Autotelic Reinforcement Learning in Multi-Agent Environments’. In: CoLLAs 2023, Conference on Lifelong Learning Agents. Montréal, Canada, Aug. 2023. URL: <https://hal.science/hal-03898121>.
- [53] E. Nisioti and C. Moulin-Frier. ‘Dynamics of niche construction in adaptable populations evolving in diverse environments’. In: ALIFE 2023 - The 2023 Conference on Artificial Life. Sapporo, Japan: MIT Press, 2023. DOI: 10.1162/isal_a_00668. URL: <https://inria.hal.science/hal-04356907>.
- [54] E. Plantec, G. Hamon, M. Etcheverry, P.-Y. Oudeyer, C. Moulin-Frier and B. W.-C. Chan. ‘Flow-Lenia: Towards open-ended evolution in cellular automata through mass conservation and parameter localization’. In: The 2023 Conference on Artificial Life. Tokyo, Japan: MIT Press; MIT Press, July 2023. DOI: 10.1162/isal_a_00651. URL: <https://hal.science/hal-03904389>.
- [55] I. Saint-Supery, H. Sauzéon, C. Maillart, N. Neu, E. Meyer and C. Mazon. ‘Cross-cultural evaluation of a web application to support communication and collaboration among stakeholders of the school inclusion of children with ASD’. In: AAATE 2023 - The 17th International Conference of the Association for the Advancement of Assistive Technology in Europe. Paris, France, 2023. URL: <https://inria.hal.science/hal-04147612>.
- [56] Z. Xiao, X. Yuan, Q. V. Liao, R. Abdelghani and P.-Y. Oudeyer. ‘Supporting Qualitative Analysis with Large Language Models: Combining Codebook with GPT-3 for Deductive Coding’. In: IUI 2023 - 28th International Conference on Intelligent User Interfaces. Sydney, Australia: ACM, 27th Mar. 2023, pp. 75–78. DOI: 10.1145/3581754.3584136. URL: <https://inria.hal.science/hal-04369097>.
- [57] M. Zadem, S. Mover and S. M. Nguyen. ‘Goal Space Abstraction in Hierarchical Reinforcement Learning via Set-Based Reachability Analysis’. In: 2023 IEEE International Conference on Development and Learning (ICDL). Macau, China: IEEE, 9th Nov. 2023, pp. 423–428. DOI: 10.1109/ICDL55364.2023.10364473. URL: <https://inria.hal.science/hal-04401878>.

- [58] K. Zhang, E. Lucet, J. Alexandre Dit Sandretto and D. Filliat. 'Navigation among movable obstacles using machine learning based total time cost optimization'. In: IROS 2023 - IEEE/RSJ International Conference on Intelligent Robots and Systems. Detroit (MI), United States, 1st Oct. 2023. URL: <https://hal-ip-paris.archives-ouvertes.fr/hal-04250608>.

Conferences without proceedings

- [59] R. Abdelghani, H. Sauz on and P.-Y. Oudeyer. 'Generative AI in the Classroom: Can Students Remain Active Learners?' In: NeurIPS 2023 - GAIED Workshop - Conference on Neural Information Processing Systems. New Orleans, USA, United States: arXiv, 2023. DOI: [10.48550/arXiv.2310.03192](https://arxiv.org/abs/10.48550/arXiv.2310.03192). URL: <https://hal.science/hal-04312934>.
- [60] R. Bornemann, G. Hamon, E. Nisioti and C. Moulin-Frier. 'Emergence of Collective Open-Ended Exploration from Decentralized Meta-Reinforcement Learning'. In: ALOE 2023 - Agent Learning in Open-Endedness workshop at NeurIPS 2023. New Orleans, United States, Dec. 2023. URL: <https://inria.hal.science/hal-04354337>.
- [61] M. Etcheverry, M. Levin, C. Moulin-Frier and P.-Y. Oudeyer. 'SBMLTOODEJAX: Efficient Simulation and Optimization of Biological Network Models in JAX'. In: NeurIPS 2023 - AI4Science Workshop. New Orleans (LA), France, 29th Oct. 2023. URL: <https://hal.science/hal-04317246>.
- [62] E. Meyer, I. Saint-Supery, H. Sauz on and C. Mazon. 'A review of technologies to collaborate and co-educate students with special educational needs and support their schooling'. In: Colloque international Education et In galit es. Paris, France, 15th May 2023. URL: <https://inria.hal.science/hal-04154325>.
- [63] I. Saint-Supery, H. Sauz on, C. Maillart, N. Neu, E. Meyer and C. Mazon. 'Cross-cultural evaluation of a web application to support communication and collaboration among stakeholders of the school inclusion of children with ASD'. In: AAATE 2023 - The 17th International Conference of the Association for the Advancement of Assistive Technology in Europe. Paris (France), France, 30th Aug. 2023. URL: <https://inria.hal.science/hal-04233368>.
- [64] L. Teodorescu, C. Colas, M. Bowers, T. Carta and P.-Y. Oudeyer. 'Codeplay: Autotelic Learning through Collaborative Self-Play in Programming Environments'. In: IMOL 2023 - Intrinsically Motivated Open-ended Learning workshop at NeurIPS 2023. New Orleans, United States, 16th Dec. 2023. URL: <https://hal.science/hal-04374993>.
- [65] M. Zadem, S. Mover and S. M. Nguyen. 'Goal Space Abstraction in Hierarchical Reinforcement Learning via Reachability Analysis'. In: IMOL 2023 - The 6th International Workshop on Intrinsically Motivated Open-ended Learning. Paris, France, 13th Sept. 2023. URL: <https://hal.science/hal-04201363>.

Scientific books

- [66] D. Roy. *Manuel d' ducation num rique - Cycle 2 56 - Collection D codage (ouvrage collectif)*. 15th Dec. 2023. URL: <https://inria.hal.science/hal-04410781>.
- [67] D. Roy. *Manuel d' ducation num rique - Cycle 2 78 - Collection D codage (ouvrage collectif)*. 8th Jan. 2024. URL: <https://inria.hal.science/hal-04410788>.

Scientific book chapters

- [68] K. Zhang, E. Lucet, J. A. D. Sandretto, S. Kchir and D. Filliat. 'A review of classical and learning based approaches in task and motion planning'. In: *Informatics in Control, Automation and Robotics*. Vol. LNCS-836. Lecture Notes in Networks and Systems. Springer International Publishing, 30th Nov. 2023, pp. 83–99. DOI: [10.1007/978-3-031-48303-5_5](https://doi.org/10.1007/978-3-031-48303-5_5). URL: <https://hal-ip-paris.archives-ouvertes.fr/hal-04316434>.

Doctoral dissertations and habilitation theses

- [69] M. Etcheverry. ‘Curiosity-driven AI for Science: Automated Discovery of Self-Organized Structures’. Inria & Labri, Université Bordeaux, 16th Nov. 2023. URL: <https://hal.science/tel-04395833>.
- [70] T. Karch. ‘Towards Social Autotelic Artificial Agents : Formation and Exploitation of Cultural Conventions in Autonomous Embodied Artificial Agents’. Université de Bordeaux, 11th May 2023. URL: <https://theses.hal.science/tel-04146927>.
- [71] L. Teodorescu. ‘Endless minds most beautiful: building open-ended linguistic autotelic agents with deep reinforcement learning and language models’. Université de bordeaux, 20th Nov. 2023. URL: <https://hal.science/tel-04396342>.

Reports & preprints

- [72] M. Adolphe, M. Pech, M. Sawayama, D. Maurel, A. Delmas, P.-Y. Oudeyer and H. Sauzéon. *Exploring the Potential of Artificial Intelligence in Individualized Cognitive Training: a Systematic Review*. 26th Dec. 2023. DOI: [10.31234/osf.io/2wg59](https://doi.org/10.31234/osf.io/2wg59). URL: <https://inria.hal.science/hal-04363997>.
- [73] M. Etcheverry, C. Moulin-Frier, P.-Y. Oudeyer and M. Levin. *AI-driven Automated Discovery Tools Reveal Diverse Behavioral Competencies of Biological Networks*. 15th Jan. 2024. DOI: [10.31219/osf.io/s6thq](https://doi.org/10.31219/osf.io/s6thq). URL: <https://hal.science/hal-04395732>.
- [74] G. Kovač, M. Sawayama, R. Portelas, C. Colas, P. F. Dominey and P.-Y. Oudeyer. *Large Language Models as Superpositions of Cultural Perspectives*. 7th Nov. 2023. URL: <https://inria.hal.science/hal-04365783>.
- [75] C. Léger, G. Hamon, E. Nisioti, X. Hinaut and C. Moulin-Frier. *Evolving Reservoirs for Meta Reinforcement Learning*. 2023. URL: <https://inria.hal.science/hal-04354303>.
- [76] J. Perez, M. Sánchez-Fibla and C. Moulin-Frier. *Cultural evolution in populations with heterogeneous and variable preferences*. 14th Dec. 2023. URL: <https://inria.hal.science/hal-04408019>.
- [77] J. Pourcel, C. Colas, P.-Y. Oudeyer and L. Teodorescu. *ACES: Generating diverse programming puzzles with autotelic language models and semantic descriptors*. 25th Oct. 2023. URL: <https://inria.hal.science/hal-04372580>.
- [78] M. Sánchez-Fibla, C. Moulin-Frier and R. Solé. *Cooperative control of environmental extremes by artificial intelligent agents*. 2022. DOI: [10.48550/arXiv.2212.02395](https://doi.org/10.48550/arXiv.2212.02395). URL: <https://inria.hal.science/hal-04356920>.
- [79] Y. Sivashankar, M. Fernandes, P.-Y. Oudeyer and H. Sauzéon. *The Beneficial Role of Curiosity on Route memory in Children*. 16th Oct. 2023. DOI: [10.31234/osf.io/4pdca](https://doi.org/10.31234/osf.io/4pdca). URL: <https://inria.hal.science/hal-04395366>.

Other scientific publications

- [80] L. Annabi and S. M. Nguyen. ‘Prerequisite structure discovery for an intelligent tutoring system based on intrinsic motivation’. In: IMOL 2023 - The 6th International Workshop on Intrinsically Motivated Open-ended Learning. Paris, France, 13th Sept. 2023. URL: <https://hal.science/hal-04201624>.
- [81] J. Pourcel, C. Colas, P.-Y. Oudeyer and L. Teodorescu. ‘ACES: generating diverse programming puzzles with autotelic language models and semantic descriptors’. In: NeurIPS 2023 - The 37th Annual Conference on Neural Information Processing Systems. Nouvelle Orleans, United States, 10th Dec. 2023. URL: <https://inria.hal.science/hal-04366773>.
- [82] M. Zadem, S. Mover and S. M. Nguyen. ‘Emergence of a Symbolic Goal Representation with an Intelligent Tutoring System based on Intrinsic Motivation’. In: NeurIPS 2023 - IMOL Workshop "Intrinsically-Motivated and Open-Ended Learning". New Orleans (Louisiana), United States: IEEE, 16th Dec. 2023, pp. 423–428. URL: <https://inria.hal.science/hal-04403613>.

12.3 Other

Scientific popularization

- [83] I. Saint-Supery, C. Mazon, M. Eric and H. Sauzéon. ‘Design of a web application to support coeducation for school inclusion of pupils with ASD’. In: *Éthiques inclusives en éducation. Recherches, contextes et pratiques* (p. 145-160). Parentalité & Handicap. Champs Social, 2023, p. 260. URL: <https://inria.hal.science/hal-04147610>.

12.4 Cited publications

- [84] M. Alaimi, E. Law, K. D. Pantasdo, P.-Y. Oudeyer and H. Sauzéon. ‘Pedagogical Agents for Fostering Question-Asking Skills in Children’. In: CHI ’20 - CHI Conference on Human Factors in Computing Systems. Honolulu / Virtual, United States, 25th Apr. 2020. DOI: [10.1145/3313831.3376776](https://doi.org/10.1145/3313831.3376776). URL: <https://hal.inria.fr/hal-03109511>.
- [85] A. Appriou, J. Ceha, S. Pramij, D. Dutartre, E. Law, P.-Y. Oudeyer and F. Lotte. ‘Towards measuring states of epistemic curiosity through electroencephalographic signals’. In: IEEE SMC 2020 - IEEE International conference on Systems, Man and Cybernetics. Toronto / Virtual, Canada, 11th Oct. 2020. URL: <https://hal.inria.fr/hal-02971795>.
- [86] R. Abdelghani, P.-Y. Oudeyer, E. Law, C. de Vulpillières and H. Sauzéon. ‘Conversational agents for fostering curiosity-driven learning in children’. In: *International Journal of Human-Computer Studies* 167 (Nov. 2022), p. 102887. DOI: [10.1016/j.ijhcs.2022.102887](https://doi.org/10.1016/j.ijhcs.2022.102887). URL: <https://hal.science/hal-03779096>.
- [87] B. Argall, S. Chernova and M. Veloso. ‘A Survey of Robot Learning from Demonstration’. In: *Robotics and Autonomous Systems* 57.5 (2009), pp. 469–483.
- [88] M. Asada, S. Noda, S. Tawaratsumida and K. Hosoda. ‘Purposive Behavior Acquisition On A Real Robot By Vision-Based Reinforcement Learning’. In: *Machine Learning* 23 (1996), pp. 279–303.
- [89] P. Barde, T. Karch, D. Nowrouzezahrai, C. Moulin-Frier, C. Pal and P.-Y. Oudeyer. ‘Learning to Guide and to Be Guided in the Architect-Builder Problem’. In: *International Conference on Learning Representations*. Virtual, France, Apr. 2022. URL: <https://hal.science/hal-03901793>.
- [90] A. Barto, S. Singh and 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.
- [91] J. Bauer, K. Baumli, F. Behbahani, A. Bhoopchand, N. Bradley-Schmiege, M. Chang, N. Clay, A. Collister, V. Dasagi, L. Gonzalez, K. Gregor, E. Hughes, S. Kashem, M. Loks-Thompson, H. Openshaw, J. Parker-Holder, S. Pathak, N. Perez-Nieves, N. Rakicevic, T. Rocktäschel, Y. Schroecker, S. Singh, J. Sygnowski, K. Tuyls, S. York, A. Zacherl and L. M. Zhang. ‘Human-Timescale Adaptation in an Open-Ended Task Space’. In: *Proceedings of the 40th International Conference on Machine Learning*. Ed. by A. Krause, E. Brunskill, K. Cho, B. Engelhardt, S. Sabato and J. Scarlett. Vol. 202. Proceedings of Machine Learning Research. PMLR, July 2023, pp. 1887–1935. URL: <https://proceedings.mlr.press/v202/bauer23a.html>.
- [92] D. Berlyne. *Conflict, Arousal and Curiosity*. McGraw-Hill, 1960.
- [93] N. Bernstein. *The Coordination and Regulation of Movements*. Pergamon, 1967.
- [94] J. C. Brant and K. O. Stanley. ‘Minimal Criterion Coevolution: A New Approach to Open-Ended Search’. In: *Proceedings of the Genetic and Evolutionary Computation Conference*. GECCO ’17. 2017, pp. 67–74.
- [95] C. Breazeal. *Designing sociable robots*. The MIT Press, 2004.
- [96] R. Brooks, C. Breazeal, R. Irie, C. C. Kemp, B. Scassellati and M. Williamson. ‘Alternative essences of intelligence’. In: *Proceedings of 15th National Conference on Artificial Intelligence (AAAI-98)*. AAAI Press, 1998, pp. 961–968.

- [97] T. Brown, B. Mann, N. Ryder, M. Subbiah, J. D. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell et al. 'Language models are few-shot learners'. In: *Advances in neural information processing systems* 33 (2020), pp. 1877–1901.
- [98] J. Bruner. 'Child's Talk: Learning to Use Language'. In: *Child Language Teaching and Therapy* 1.1 (1985), pp. 111–114. DOI: [10.1177/026565908500100113](https://doi.org/10.1177/026565908500100113). eprint: <https://doi.org/10.1177/026565908500100113>. URL: <https://doi.org/10.1177/026565908500100113>.
- [99] A. Campero, R. Raileanu, H. Küttler, J. B. Tenenbaum, T. Rocktäschel and E. Grefenstette. 'Learning with amigo: Adversarially motivated intrinsic goals'. In: *arXiv preprint arXiv:2006.12122* (2020).
- [100] A. Cangelosi and M. Schlesinger. *Developmental robotics: From babies to robots*. MIT press, 2015.
- [101] J. Ceha, E. Law, D. Kulić, v. Oudeyer Pierre-Y and D. Roy. 'Identifying Functions and Behaviours of Social Robots for In-Class Learning Activities: Teachers' Perspective'. In: *International Journal of Social Robotics* (Sept. 2021). DOI: [10.1007/s12369-021-00820-7](https://hal.inria.fr/hal-03465477). URL: <https://hal.inria.fr/hal-03465477>.
- [102] *Lenia and Expanded Universe*. Vol. ALIFE 2020: The 2020 Conference on Artificial Life. ALIFE 2021: The 2021 Conference on Artificial Life. July 2020, pp. 221–229. DOI: [10.1162/isal_a_00297](https://direct.mit.edu/isal/proceedings-pdf/isal2020/32/221/1908612/isal_a_00297.pdf). eprint: https://direct.mit.edu/isal/proceedings-pdf/isal2020/32/221/1908612/isal_a_00297.pdf. URL: https://doi.org/10.1162/isal%5C_a%5C_00297.
- [103] B. W.-C. Chan. 'Lenia-biology of artificial life'. In: *Complex Systems* 28.3 (2019), pp. 251–286.
- [104] J. Ciecuch and S. H. Schwartz. 'The Number of Distinct Basic Values and Their Structure Assessed by PVQ-40'. In: *J. Pers. Assess.* 94.3 (May 2012), pp. 321–328.
- [105] A. Clark. *Mindware: An Introduction to the Philosophy of Cognitive Science*. Oxford University Press, 2001.
- [106] B. Clément. 'Adaptive Personalization of Pedagogical Sequences using Machine Learning'. Theses. Université de Bordeaux, Dec. 2018. URL: <https://hal.inria.fr/tel-01968241>.
- [107] B. Clément, D. Roy, P.-Y. Oudeyer and M. Lopes. 'Multi-Armed Bandits for Intelligent Tutoring Systems'. In: *Journal of Educational Data Mining (JEDM)* 7.2 (June 2015), pp. 20–48. URL: <https://hal.inria.fr/hal-00913669>.
- [108] D. Cohn, Z. Ghahramani and M. Jordan. 'Active learning with statistical models'. In: *Journal of artificial intelligence research* 4 (1996), pp. 129–145.
- [109] C. Colas, T. Karch, C. Moulin-Frier and P.-Y. Oudeyer. 'Language and culture internalization for human-like autotelic AI'. In: *Nature Machine Intelligence* 4.12 (1st Dec. 2022), pp. 1068–1076. DOI: [10.1038/s42256-022-00591-4](https://doi.org/10.1038/s42256-022-00591-4). URL: <https://doi.org/10.1038/s42256-022-00591-4>.
- [110] C. Colas, T. Karch, O. Sigaud and P.-Y. Oudeyer. 'Autotelic Agents with Intrinsically Motivated Goal-Conditioned Reinforcement Learning: A Short Survey'. en. In: *Journal of Artificial Intelligence Research* 74 (July 2022), pp. 1159–1199. DOI: [10.1613/jair.1.13554](https://www.jair.org/index.php/jair/article/view/13554). URL: <https://www.jair.org/index.php/jair/article/view/13554> (visited on 26/10/2022).
- [111] C. Colas, T. Karch, O. Sigaud and P.-Y. Oudeyer. 'Autotelic agents with intrinsically motivated goal-conditioned reinforcement learning: a short survey'. In: *Journal of Artificial Intelligence Research* 74 (2022), pp. 1159–1199.
- [112] C. Colas, T. Karch, O. Sigaud and P.-Y. Oudeyer. 'Intrinsically Motivated Goal-Conditioned Reinforcement Learning: a Short Survey'. working paper or preprint. Jan. 2021. URL: <https://hal.inria.fr/hal-03099891>.
- [113] W. Croft and D. Cruse. *Cognitive Linguistics*. Cambridge Textbooks in Linguistics. Cambridge University Press, 2004.
- [114] M. Csikszentmihalyi. *Flow-the psychology of optimal experience*. Harper Perennial, 1991.
- [115] P. Dayan and W. Belleine. 'Reward, motivation and reinforcement learning'. In: *Neuron* 36 (2002), pp. 285–298.
- [116] E. Deci and R. Ryan. *Intrinsic Motivation and Self-Determination in Human Behavior*. Plenum Press, 1985.

- [117] M. Derex and R. Boyd. ‘Partial connectivity increases cultural accumulation within groups’. en. In: *Proceedings of the National Academy of Sciences* 113.11 (Mar. 2016), pp. 2982–2987. DOI: [10.1073/pnas.1518798113](https://doi.org/10.1073/pnas.1518798113). URL: <http://www.pnas.org/lookup/doi/10.1073/pnas.1518798113> (visited on 29/01/2021).
- [118] Y. Duan, J. Schulman, X. Chen, P. L. Bartlett, I. Sutskever and P. Abbeel. ‘RL²: Fast Reinforcement Learning via Slow Reinforcement Learning’. In: *arXiv:1611.02779 [cs, stat]* (2016).
- [119] J. Elman. ‘Learning and development in neural networks: The importance of starting small’. In: *Cognition* 48 (1993), pp. 71–99.
- [120] G. E. Fletcher, F. Warneken and M. Tomasello. ‘Differences in cognitive processes underlying the collaborative activities of children and chimpanzees’. In: *Cognitive Development* 27.2 (2012), pp. 136–153. DOI: <https://doi.org/10.1016/j.cogdev.2012.02.003>. URL: <https://www.sciencedirect.com/science/article/pii/S0885201412000093>.
- [121] C. Florensa, D. Held, X. Geng and P. Abbeel. ‘Automatic goal generation for reinforcement learning agents’. In: *International conference on machine learning*. PMLR. 2018, pp. 1515–1528.
- [122] J. Gottlieb, P.-Y. Oudeyer, M. Lopes and A. Baranes. ‘Information-seeking, curiosity, and attention: computational and neural mechanisms’. In: *Trends in Cognitive Sciences* 17.11 (Nov. 2013), pp. 585–93. DOI: [10.1016/j.tics.2013.09.001](https://doi.org/10.1016/j.tics.2013.09.001). URL: <https://hal.inria.fr/hal-00913646>.
- [123] J. Grizou, L. J. Points, A. Sharma and L. Cronin. ‘A curious formulation robot enables the discovery of a novel protocell behavior’. In: *Science Advances* 6.5 (2020), eaay4237. DOI: [10.1126/sciadv.aay4237](https://doi.org/10.1126/sciadv.aay4237). eprint: <https://www.science.org/doi/pdf/10.1126/sciadv.aay4237>. URL: <https://www.science.org/doi/abs/10.1126/sciadv.aay4237>.
- [124] M. Grove. ‘Evolution and dispersal under climatic instability: a simple evolutionary algorithm’. en. In: *Adaptive Behavior* 22.4 (Aug. 2014), pp. 235–254. DOI: [10.1177/1059712314533573](https://doi.org/10.1177/1059712314533573). URL: <http://journals.sagepub.com/doi/10.1177/1059712314533573> (visited on 10/05/2021).
- [125] P. Haluptzok, M. Bowers and A. T. Kalai. ‘Language models can teach themselves to program better’. In: *arXiv preprint arXiv:2207.14502* (2022).
- [126] G. Hamon, M. Etcheverry, B. W.-C. Chan, C. Moulin-Frier and P.-Y. Oudeyer. *Learning Sensorimotor Agency in Cellular Automata*. In this blogpost, we explore the concepts of embodiment, individuality, self-maintenance and sensorimotor agency within a cellular automaton (CA) environment. Whereas those concepts are central in theoretical biology and cognitive science, it remains unclear how such behaviors can emerge in a CA-like environment made only of low-level particles and physical rules. We present a novel set of tools (based on curriculum learning, diversity search and gradient descent over a differentiable CA) to automatically learn the rules leading to the emergence of such behaviors. Our method is able to discover robust self-organizing agents with strong coherence and generalization to out-of-distribution changes, reminiscent of the robustness of living systems to maintain specific functions despite environmental and body perturbations. Jan. 2022. URL: <https://inria.hal.science/hal-03519319>.
- [127] S. Harnad. ‘The symbol grounding problem’. In: *Physica D* 40 (1990), pp. 335–346.
- [128] M. Hasenjager and H. Ritter. *Active learning in neural networks*. Heidelberg, Germany, Germany: Physica-Verlag GmbH, 2002, pp. 137–169.
- [129] J. Haugeland. *Artificial Intelligence: the very idea*. Cambridge, MA, USA: The MIT Press, 1985.
- [130] J.-C. Horvitz. ‘Mesolimbocortical and nigrostriatal dopamine responses to salient non-reward events’. In: *Neuroscience* 96.4 (2000), pp. 651–656.
- [131] X. Huang and 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*. Ed. by C. Prince, Y. Demiris, Y. Marom, H. Kozima and C. Balkenius. Lund University Cognitive Studies 94, 2002, pp. 47–55.

- [132] S. Ivaldi, N. Lyubova, D. Gérardeaux-Viret, A. Droniou, S. Anzalone, M. Chetouani, D. Filliat and 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*. forthcoming. Japan, 2012. URL: <http://hal.inria.fr/hal-00755297>.
- [133] M. Jaderberg, W. M. Czarnecki, I. Dunning, L. Marris, G. Lever, A. G. Castaneda, C. Beattie, N. C. Rabinowitz, A. S. Morcos, A. Ruderman et al. ‘Human-level performance in 3D multiplayer games with population-based reinforcement learning’. In: *Science* 364.6443 (2019). Publisher: American Association for the Advancement of Science, pp. 859–865.
- [134] M. Johnson. *Developmental Cognitive Neuroscience*. 2nd. Blackwell publishing, 2005.
- [135] M. H. Johnson. *Developmental cognitive neuroscience*. Wiley-Blackwell, 2011.
- [136] T. D. Johnston. ‘Selective Costs and Benefits in the Evolution of Learning’. en. In: *Advances in the Study of Behavior*. Ed. by J. S. Rosenblatt, R. A. Hinde, C. Beer and M.-C. Busnel. Vol. 12. Academic Press, Jan. 1982, pp. 65–106. DOI: [10.1016/S0065-3454\(08\)60046-7](https://doi.org/10.1016/S0065-3454(08)60046-7). URL: <http://www.sciencedirect.com/science/article/pii/S0065345408600467> (visited on 22/04/2020).
- [137] H. Kitano. ‘Biological robustness’. In: *Nature Reviews Genetics* 5.11 (2004), pp. 826–837.
- [138] W. B. Knox and 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)*. Toronto, Canada, 2010, pp. 5–12.
- [139] K. N. Laland, T. Uller, M. W. Feldman, K. Sterelny, G. B. Müller, A. Moczek, E. Jablonka and J. Odling-Smee. ‘The extended evolutionary synthesis: its structure, assumptions and predictions’. In: *Proceedings of the royal society B: biological sciences* 282.1813 (2015), p. 20151019.
- [140] R. T. Lange and H. Sprekeler. *Learning not to learn: Nature versus nurture in silico*. 2020. arXiv: [2010.04466](https://arxiv.org/abs/2010.04466) [cs.LG].
- [141] H. Le, Y. Wang, A. D. Gotmare, S. Savarese and S. C. H. Hoi. ‘Coderl: Mastering code generation through pretrained models and deep reinforcement learning’. In: *Advances in Neural Information Processing Systems* 35 (2022), pp. 21314–21328.
- [142] J. Z. Leibo, V. Zambaldi, M. Lanctot, J. Marecki and T. Graepel. ‘Multi-Agent Reinforcement Learning in Sequential Social Dilemmas’. In: *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems. AAMAS ’17*. São Paulo, Brazil, 2017, pp. 464–473.
- [143] M. Lopes, T. Cederborg and 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](https://doi.org/10.1109/DEVLRN.2011.6037359). URL: <http://hal.inria.fr/hal-00636166/en>.
- [144] M. Lungarella, G. Metta, R. Pfeifer and G. Sandini. ‘Developmental Robotics: A Survey’. In: *Connection Science* 15.4 (2003), pp. 151–190.
- [145] N. Lyubova and 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://doi.org/10.1109/IJCNN.2012.6252606). URL: <https://hal.archives-ouvertes.fr/hal-00755298>.
- [146] J. Marshall, D. Blank and 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.
- [147] M. Mason and 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](https://doi.org/10.1145/1957656.1957814). URL: <http://hal.inria.fr/hal-00636164/en>.
- [148] C. Mazon, B. Clément, D. Roy, P.-Y. Oudeyer and H. Sauzéon. ‘Pilot study of an intervention based on an intelligent tutoring system (ITS) for instructing mathematical skills of students with ASD and/or ID’. In: *Education and Information Technologies* (2022). DOI: [10.1007/s10639-022-11129-x](https://doi.org/10.1007/s10639-022-11129-x). URL: <https://inria.hal.science/hal-03676589>.

- [149] C. Mazon, K. Etchegoyhen, I. Saint-Supery, A. Amestoy, M. Bouvard, C. Consel and H. Sauzéon. 'Fostering parents-professional collaboration for facilitating the school inclusion of students with ASD: Design of the "ToGather" web-based prototype'. In: *Educational Technology Research and Development* (Dec. 2021). DOI: [10.1007/s11423-021-10073-w](https://doi.org/10.1007/s11423-021-10073-w). URL: <https://hal.inria.fr/hal-03436355>.
- [150] C. Mazon, C. Fage and 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* 93 (Apr. 2019). DOI: [10.1016/j.chb.2018.12.001](https://doi.org/10.1016/j.chb.2018.12.001). URL: <https://hal.inria.fr/hal-01950078>.
- [151] C. Mazon and H. Sauzéon. 'Utilisation des technologies mobiles auprès des enfants avec TSA.' In: *Autisme et usages du numériques en éducation*. Ed. by P. Bourdon. 2022. URL: <https://inria.hal.science/hal-03120323>.
- [152] P. Miller. *Theories of developmental psychology*. 4th. New York: Worth, 2001.
- [153] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg and D. Hassabis. 'Human-level control through deep reinforcement learning'. en. In: *Nature* 518.7540 (Feb. 2015), pp. 529–533. DOI: [10.1038/nature14236](https://doi.org/10.1038/nature14236). URL: <http://www.nature.com/articles/nature14236> (visited on 22/08/2018).
- [154] C. Moulin-Frier. 'The Ecology of Open-Ended Skill Acquisition'. Habilitation à diriger des recherches. Université de Bordeaux (UB), Dec. 2022. URL: <https://inria.hal.science/tel-03875448>.
- [155] J. Mu, V. Zhong, R. Raileanu, M. Jiang, N. Goodman, T. Rocktäschel and E. Grefenstette. 'Improving intrinsic exploration with language abstractions'. In: *Advances in Neural Information Processing Systems* 35 (2022), pp. 33947–33960.
- [156] A. Nair, P. Srinivasan, S. Blackwell, C. Alcicek, R. Fearon, A. De Maria, V. Panneershelvam, M. Suleyman, C. Beattie, S. Petersen, S. Legg, V. Mnih, K. Kavukcuoglu and D. Silver. *Massively Parallel Methods for Deep Reinforcement Learning*. en. Tech. rep. arXiv:1507.04296. arXiv:1507.04296 [cs]. arXiv, July 2015. URL: <http://arxiv.org/abs/1507.04296> (visited on 15/05/2022).
- [157] S. M. Nguyen, A. Baranes and P.-Y. Oudeyer. 'Bootstrapping Intrinsically Motivated Learning with Human Demonstrations'. In: *IEEE International Conference on Development and Learning*. Frankfurt, Germany, 2011. URL: <http://hal.inria.fr/hal-00645986/en>.
- [158] S. M. Nguyen, A. Baranes and 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. URL: <http://hal.inria.fr/hal-00645995/en>.
- [159] E. Nisioti, M. Mahaut, P.-Y. Oudeyer, I. Momennejad and C. Moulin-Frier. 'Social Network Structure Shapes Innovation: Experience-sharing in RL with SAPIENS'. working paper or preprint. July 2022. URL: <https://hal.science/hal-03715924>.
- [160] E. Nisioti and C. Moulin-Frier. 'Plasticity and evolvability under environmental variability: the joint role of fitness-based selection and niche-limited competition'. In: *GECCO 2022 - The Genetic and Evolutionary Computation Conference*. Boston / Hybrid, United States, July 2022. URL: <https://hal.science/hal-03715928>.
- [161] O. OpenAI, M. Plappert, R. Sampedro, T. Xu, I. Akkaya, V. Kosaraju, P. Welinder, R. D'Sa, A. Petron, H. P. d. O. Pinto et al. 'Asymmetric self-play for automatic goal discovery in robotic manipulation'. In: *arXiv preprint arXiv:2101.04882* (2021).
- [162] P.-Y. Oudeyer, J. v and M. Lopes. 'Intrinsic motivation, curiosity, and learning: Theory and applications in educational technologies'. In: *Progress in brain research* 229 (2016), pp. 257–284.
- [163] P.-Y. Oudeyer. 'Developmental Robotics'. In: *Encyclopedia of the Sciences of Learning*. Ed. by N. Seel. Springer Reference Series. Springer, 2011. URL: <https://hal.inria.fr/hal-00652123/en>.

- [164] 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*. Ed. by S. Dehaene and C. Petit. Odile Jacob, 2009, pp. 83–112. URL: <http://hal.inria.fr/inria-00446908/en/>.
- [165] 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 ?* Ed. by C. Garbay and D. Kaiser. Presses Universitaires de France, 2009. URL: <http://hal.inria.fr/inria-00420309/en/>.
- [166] P.-Y. Oudeyer, F. Kaplan and V. Hafner. 'Intrinsic Motivation Systems for Autonomous Mental Development'. In: *IEEE Transactions on Evolutionary Computation* 11.2 (2007), pp. 265–286. DOI: [10.1109/TEVC.2006.89027](https://doi.org/10.1109/TEVC.2006.89027).
- [167] P.-Y. Oudeyer and F. Kaplan. 'Intelligent adaptive curiosity: a source of self-development'. In: *Proceedings of the 4th International Workshop on Epigenetic Robotics*. Ed. by L. Berthouze, H. Kozima, C. Prince, G. Sandini, G. Stojanov, G. Metta and C. Balkenius. Vol. 117. Lund University Cognitive Studies, 2004, pp. 127–130.
- [168] P.-Y. Oudeyer and F. Kaplan. 'What is intrinsic motivation? A typology of computational approaches'. In: *Frontiers in Neurobotics* 1.1 (2007).
- [169] L. Ouyang, J. Wu, X. Jiang, D. Almeida, C. Wainwright, P. Mishkin, C. Zhang, S. Agarwal, K. Slama, A. Ray et al. 'Training language models to follow instructions with human feedback'. In: *Advances in Neural Information Processing Systems* 35 (2022), pp. 27730–27744.
- [170] K. Pearson. 'Neural adaptation in the generation of rhythmic behavior'. In: *Annual review of physiology* 62.1 (2000), pp. 723–753.
- [171] J. Perolat, J. Z. Leibo, V. Zambaldi, C. Beattie, K. Tuyls and T. Graepel. 'A multi-agent reinforcement learning model of common-pool resource appropriation'. In: *Advances in neural information processing systems* 30 (2017).
- [172] R. Portelas. 'Automatic Curriculum Learning for Developmental Machine Learners'. Theses. Université de Bordeaux, Feb. 2022. URL: <https://theses.hal.science/tel-03633787>.
- [173] M. Poupard, F. Larrue, H. Sauzéon and A. Tricot. 'A systematic review of immersive technologies for education: effects of cognitive load and curiosity state on learning performance'. working paper or preprint. Dec. 2022. URL: <https://hal.science/hal-03906797>.
- [174] R. Ramamurthy, P. Ammanabrolu, K. Brantley, J. Hessel, R. Sifa, C. Bauckhage, H. Hajishirzi and Y. Choi. 'Is Reinforcement Learning (Not) for Natural Language Processing?: Benchmarks, Baselines, and Building Blocks for Natural Language Policy Optimization'. In: 2022. URL: <https://arxiv.org/abs/2210.01241>.
- [175] A. Revel and J. Nadel. 'Imitation and Social Learning in Robots, Humans and Animals: Behavioural, Social and Communicative Dimensions'. In: ed. by K. Dautenhahn and C. Nehaniv. Cambridge University Press, 2004. Chap. How to build an imitator?
- [176] I. Saint-Supery, H. Sauzéon, E. Meyer and C. Mazon. 'ToGather, an interactive website for the stakeholders of school inclusion of children with ASD: an iterative design including user testing'. working paper or preprint. 2022. URL: <https://inria.hal.science/hal-03898867>.
- [177] T. Schatz and 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. URL: <http://hal.inria.fr/inria-00420186/en/>.
- [178] M. Schembri, M. Mirolli and 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. URL: <http://dx.doi.org/10.1109/DEVLRN.2007.4354052>.
- [179] J. Schmidhuber. 'Curious Model-Building Control Systems'. In: *Proceedings of the International Joint Conference on Neural Networks, Singapore*. Vol. 2. IEEE press, 1991, pp. 1458–1463.

- [180] J. Schmidhuber. 'Powerplay: Training an increasingly general problem solver by continually searching for the simplest still unsolvable problem'. In: *Frontiers in psychology* 4 (2013), p. 313.
- [181] B. Schrauwen, D. Verstraeten and J. M. V. Campenhout. 'An overview of reservoir computing: theory, applications and implementations'. In: *The European Symposium on Artificial Neural Networks*. 2007. URL: <https://api.semanticscholar.org/CorpusID:16935574>.
- [182] J. Schulman, F. Wolski, P. Dhariwal, A. Radford and O. Klimov. 'Proximal policy optimization algorithms'. In: *arXiv preprint arXiv:1707.06347* (2017).
- [183] W. Schultz, P. Dayan and P. Montague. 'A neural substrate of prediction and reward'. In: *Science* 275 (1997), pp. 1593–1599.
- [184] T. Schuster, A. Kalyan, O. Polozov and A. T. Kalai. 'Programming puzzles'. In: *arXiv preprint arXiv:2106.05784* (2021).
- [185] S. H. Schwartz. 'An overview of the Schwartz theory of basic values'. In: *Online readings in Psychology and Culture* 2.1 (2012), p. 11.
- [186] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. Van Den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot et al. 'Mastering the game of Go with deep neural networks and tree search'. In: *nature* 529.7587 (2016), pp. 484–489.
- [187] K. O. Stanley, J. Clune, J. Lehman and R. Miikkulainen. 'Designing Neural Networks through Neuroevolution'. In: *Nature Machine Intelligence* 1.1 (Jan. 2019), pp. 24–35. DOI: [10.1038/s42256-018-0006-z](https://doi.org/10.1038/s42256-018-0006-z).
- [188] L. Steels and R. Brooks, eds. *The Artificial Life Route to Artificial Intelligence: Building Embodied, Situated Agents*. Hillsdale, NJ, USA: L. Erlbaum Associates Inc., 1995.
- [189] L. L. Steels. *The Talking Heads experiment. Computational Models of Language Evolution 1*. Berlin: Language Science Press, 2015. DOI: [10.17169/langsci.b49.75](https://doi.org/10.17169/langsci.b49.75).
- [190] D. W. Stephens. 'Change, regularity, and value in the evolution of animal learning'. In: *Behavioral Ecology* 2.1 (1991), pp. 77–89. DOI: [10.1093/beheco/2.1.77](https://doi.org/10.1093/beheco/2.1.77).
- [191] S. Sukhbaatar, Z. Lin, I. Kostrikov, G. Synnaeve, A. Szlam and R. Fergus. 'Intrinsic motivation and automatic curricula via asymmetric self-play'. In: *arXiv preprint arXiv:1703.05407* (2017).
- [192] E. Thelen and L. B. Smith. *A dynamic systems approach to the development of cognition and action*. Cambridge, MA: MIT Press, 1994.
- [193] A. L. Thomaz and C. Breazeal. 'Teachable robots: Understanding human teaching behavior to build more effective robot learners'. In: *Artificial Intelligence Journal* 172 (2008), pp. 716–737.
- [194] A. Tierney. 'Evolutionary implications of neural circuit structure and function'. In: *Behavioural processes* 35.1-3 (1995), pp. 173–182.
- [195] M. Tomasello. 'Becoming human'. In: *Becoming Human*. Harvard University Press, 2019.
- [196] M. Tomasello. *The Cultural Origins of Human Cognition*. Harvard University Press, 1999. URL: <http://www.jstor.org/stable/j.ctvjjsf4jc>.
- [197] A. Turing. 'Computing machinery and intelligence'. In: *Mind* 59 (1950), pp. 433–460.
- [198] F. Varela, E. Thompson and E. Rosch. *The embodied mind: Cognitive science and human experience*. Cambridge, MA: MIT Press, 1991.
- [199] A.-L. Vollmer, J. Grizou, M. Lopes, K. Rohlfing and P.-Y. Oudeyer. 'Studying the co-construction of interaction protocols in collaborative tasks with humans'. In: *4th International Conference on Development and Learning and on Epigenetic Robotics*. IEEE. 2014, pp. 208–215.
- [200] L. S. Vygotsky and M. Cole. *Mind in society: Development of higher psychological processes*. Harvard university press, 1978.
- [201] J. X. Wang, Z. Kurth-Nelson, D. Tirumala, H. Soyer, J. Z. Leibo, R. Munos, C. Blundell, D. Kumaran and M. Botvinick. 'Learning to Reinforcement Learn'. In: *arXiv:1611.05763 [cs, stat]* (2017).
- [202] J. Weng, J. McClelland, A. Pentland, O. Sporns, I. Stockman, M. Sur and E. Thelen. 'Autonomous mental development by robots and animals'. In: *Science* 291 (2001), pp. 599–600.

- [203] J. Yang, A. Prabhakar, K. Narasimhan and S. Yao. 'InterCode: Standardizing and Benchmarking Interactive Coding with Execution Feedback'. In: *arXiv preprint arXiv:2306.14898* (2023).